MATH 1220

# Mathematical Discovery 2

current editor

Dr Paul Vrbik

last edit

September 26, 2017

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

# Contents

# 1 Sequences

*"Confusion is the sweat of learning."*

– Folklore

A **sequence** is a list of numbers, in this course they will be real numbers $\mathbb{R}$. These sequences can be viewed as a function from $\mathbb{N}$ to $\mathbb{R}$ and will be denoted $\{x_n\}$ for $n \in \mathbb{N}$ and this means

$$\{x_n\} := x_1, x_2, x_3, \dots$$

where $x_1$ is the first **element** of the sequence, $x_2$ the second, and so on.... Sometimes it will be convenient to begin the sequence with $x_0$ instead.

We study sequences of numbers because they arise in many situations. Consider the sequence of

1. scores of a batsman in cricket.

2. digits in a decimal expansion: $\pi = 3.14\dots = x_0 \,.\, x_1\, x_2 \dots$

3. decimal approximations to a number $x_1 = 3.1$, $x_2 = 3.14$, $\dots$, $x_n =$ expansion of $\pi$ to $n$ decimal places.

4. approximations given by Newton's Method:

$$x_0 = 1, \qquad\qquad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

5. of **binomial coefficients** $\left\{\binom{\frac{1}{2}}{k}\right\} = 1, \frac{1}{2}, -\frac{1}{8}, \dots.$

In many situations, we are interested in whether the sequence converges. The situations are typically when we have some algorithm that produces a sequence of approximations to some desired number. Late on, we come to study sequences of functions that approximate a desrired function, and sequences of other things.

**Convergence.** We say the sequence $\{x_n\}$ **converges** to $x$ when

$$\forall \varepsilon > 0; \exists N \in \mathbb{N} : n \geqslant N \implies |x_n - x| < \varepsilon. \tag{1.1}$$

When this is the case we can write/say

1. $x_n \to x$ as $n \to \infty$,
2. $\lim_{n\to\infty} x_n = x$, and
3. $x$ is the *limit* of the sequence $\{x_n\}$.

All mean the same thing.

**Proposition 1.1.** $\lim_{n\to\infty} \frac{1}{n^r} = 0$ whenever $r > 0$. That is to say

$$n \in \mathbb{N}, r > 0 \implies \frac{1}{n^r} \to 0 \text{ as } n \to \infty$$

or equivalently: 0 is the limit limit of $\{\frac{1}{n^r}\}$.

PROOF. Fix $\varepsilon > 0$. We need to find $N$ such that

$$n \geqslant N \implies \left|\frac{1}{n^r} - 0\right| < \varepsilon.$$

Since $n > 0$, $\left|\frac{1}{n^r} - 0\right| = \frac{1}{n^r}$ and so

$$\begin{aligned} \left|\frac{1}{n^r} - 0\right| < \varepsilon &\iff \frac{1}{n^r} < \varepsilon \\ &\iff \frac{1}{\varepsilon} < n^r \\ &\iff n > \frac{1}{\varepsilon^{\frac{1}{n}}}. \end{aligned}$$

Thus any $N \geqslant \frac{1}{\varepsilon^{\frac{1}{n}}}$ will do. ■

Note the smaller $\varepsilon$ is, the larger $N$ must be. To guarantee that $x_n$ is closer to $x$, need to go further along sequence. That is, iterate Newton's Method further.

Fortunately, we do not need to check the definition every time. The following theorem allows us to reduce many limits to a few standard ones.

**Algebra of Limits.** Suppose that $x_n \to x$ and $y_n \to y$ and let $a, b, \in \mathbb{R}$. Then

1. $ax_n + by_n \to ax + by$,
2. $x_n y_n \to xy$,
3. if $y_n \neq 0$ for all $n$ and $y \neq 0$ then $\dfrac{x_n}{y_n} \to \dfrac{x}{y}$.

as $n \to \infty$.

These statements need to be justified in terms of the definition of convernce. We will use what we are given about $\{x_n\}$ and $\{y_n\}$ to show what we want about $\{ax_n + by_n\}$.

PROOF OF 1. Fix $\varepsilon > 0$. We must find $N \in \mathbb{N}$ such that $n \geqslant N \implies |(ax_n + by_n) - (ax + by)| < \varepsilon$. We estimate:

$$\begin{aligned} |(ax_n + by_n) - (ax + by)| &= |a(x_n - x) + b(y_n - y)| \\ &\leqslant |a|\,|x_n - x| + |b|\,|y_n - y|. \end{aligned}$$

When $a = 0$ we take $N_1 = 1$ otherwise, since $x_n \to x$, we can choose $N_1 \in \mathbb{N}$

$$n \geqslant N_1 \implies |x_n - x| < \frac{\varepsilon}{2|a|}.$$

When $b = 0$ take $N_2 = 1$ otherwise, since $y_n \to y$, we can choose $N_2 \in \mathbb{N}$ such that

$$n \geqslant N_2 \implies |y_n - y| < \frac{\varepsilon}{2|b|}.$$

Choose $N = \max(N_1, N_2)$. Then

$$\begin{aligned} n \geqslant N &\implies n \geqslant N_1 \text{ and } n \geqslant N_2 \\ &\implies |a|\,|x_n - x| < \frac{|a|\varepsilon}{2|a|} \text{ and } |b|\,|y_n - y| < \frac{|b|\varepsilon}{2|b|} \\ &\implies |a|\,|x_n - x| + |b|\,|y_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Hence this $N$ suffices. ■

Other basic examples (in addition to $\frac{1}{n^r}$ as $n \to \infty$) are provided by limits of functions.

**Proposition 1.2.** $f(x) \to L$ as $x \to \infty \implies f(n) \to L$ as $n \to \infty$.

PROOF. The statements are trivially equivalent when $x$ is replaced by $n$ in the definitions of the two statements. ■

**Example 1.3.** [1]

$$\lim_{n\to\infty} n\left(1-e^{-\frac{1}{n}}\right) = \lim_{x\to\infty}\frac{1-e^{-\frac{1}{x}}}{\frac{1}{x}}$$
$$= \lim_{x\to\infty}\frac{-e^{-\frac{1}{x}}\cdot\frac{1}{x^2}}{-\frac{1}{x^2}}$$
$$\overset{H}{=} \lim_{x\to\infty} +e^{-\frac{1}{x}}$$
$$= 1.$$

**Example 1.4.** $\sin\frac{1}{n} \to 0$ as $n\to\infty$ because $\sin\frac{1}{x}\to \mathbf{0}$ as $x\to\infty$.

**Example 1.5.**

$$\frac{n^2-2n+3}{2n^3+n+7} = \frac{\frac{n^2-2n+3}{n^3}}{\frac{2n^3+n+7}{n^3}} = \frac{\frac{1}{n}-\frac{2}{n^2}+\frac{3}{n^3}}{2+\frac{1}{n^2}+\frac{7}{n^3}}$$

Since we have

$$\frac{1}{n}\to 0 \qquad \frac{1}{n^2}\to 0 \qquad \frac{1}{n^3}\to 0$$

as $n\to\infty$. We have, by the Algebra of Limits, that

$$\frac{1}{n}-\frac{2}{n^2}+\frac{3}{n^2}\to 0-2\cdot 0+3\cdot 0 = 0$$

and

$$2+\frac{1}{n^2}+\frac{7}{n^3}\to 2+0+7\cdot 0 = 2.$$

Hence, by the algebra of limits,

$$\frac{\frac{1}{n}-\frac{2}{n^3}+\frac{3}{n^3}}{2+\frac{1}{n^2}+\frac{7}{n^3}} \to \frac{0}{2} = 0$$

Not all limits can be evaluated in this way.

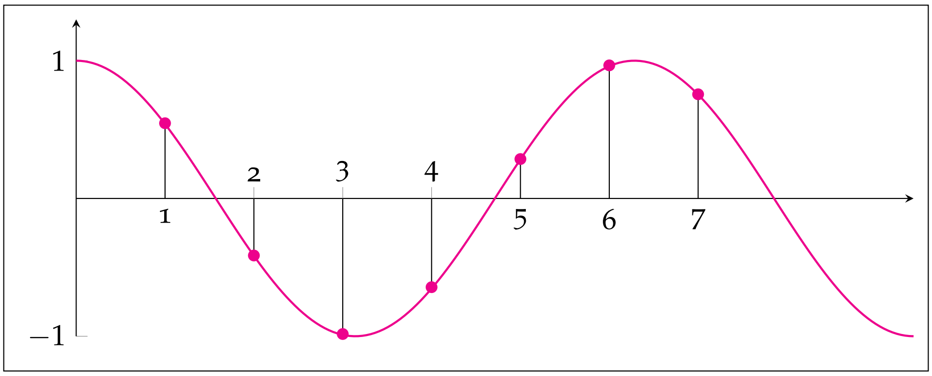
**Example 1.6.**

$$\frac{n^2-n\cos n}{6n^2+n+1} = \frac{\frac{n^2-n\cos n}{n^2}}{\frac{6n^2+n+1}{n^2}} = \frac{1-\frac{1}{n}\cos n}{6+\frac{1}{n}+\frac{1}{n^2}}.$$

Can apply the algebra of limits to conclude that $6+\frac{1}{n}+\frac{1}{n^2}\to 6$ as $n\to\infty$. We need to know $\lim_{n\to\infty}\frac{1}{n}\cos n$. $\frac{1}{n}\to 0$ as $n\to\infty$ but $\{\cos n\}$ does not

[1] $\overset{H}{=}$ denotes the step is done via L'Hoptials Rule.

converge:




And thus we cannot apply the Algebra of Limits here.

**The Squeeze Principle.** Suppose that

1. $x_n \to x$ as $n \to \infty$,
2. $y_n \to x$ as $n \to \infty$, and
3. $\forall n\,(x_n \leqslant z_n \leqslant y_n)$.

Then $z_n \to x$ as $n \to \infty$.

PROOF. Proof is given in Math 2330. ■

**Example 1.7.** Continuing Example 1.6. Since, for all $n$,

$$-1 \leqslant \cos n \leqslant 1 \implies -\frac{1}{n} \leqslant \frac{1}{n}\cos n \leqslant \frac{1}{n}.$$

We have $-\frac{1}{n} \to 0$ and $\frac{1}{n} \to 0$ as $n \to \infty$. Hence $\frac{1}{n}\cos n \to 0$ as $n \to \infty$ by the squeeze principle.

Hence $1 - \frac{1}{n}\cos n \to 1$ as $n \to \infty$ and thereby

$$\frac{n^2 - n\cos n}{6n^2 + n + 1} = \frac{1 - \frac{1}{n}\cos n}{6 + \frac{1}{n} + \frac{1}{n^2}} \to \frac{1}{6} \text{ as } n \to \infty$$

by the algebra of limits.

**Proposition 1.8.** Suppose that $\{x_n\}$ is an increasing sequence such that there is $M \in |RR$ with $x_n \leqslant M$ for all $n$. Then $\{x_n\}$ converges.

This will be proved in Math 2330 Analysis. It is an important property of $\mathbb{R}$.

# 2 Series

A **series** is the formal sum of a sequence. Thus, if $\{x_n\}$ is a sequence, the corresponding series is $\sum_{k=1}^{\infty} x_k$.

**Definition 2.1.** The $n$**th partial sum** of the series $\sum_{k=1}^{\infty} x_k$ is

$$s_n := \sum_{k=1}^{n} x_k = x_1 + x_2 + \cdots + x_n.$$

The series $\sum_{k=1}^{n}$ **converges with sum** $s$ if the sequence $\{s_n\}$ of the partial sums converges to $s$. In this case we write

$$\sum_{k=1}^{\infty} x_k = s.$$

If $\sum_{k=1}^{\infty} x_k$ does not converge, we say that it **diverges**.

Intuitively, $\sum_{k=1}^{\infty} x_k = s$ if $x_1 + x_2 + \cdots + x_n$ is close to $s$ for big $n$.
We will often start the sum at a different point like

$$\sum_{k=a}^{\infty} x_k.$$

This converges if the sequence of partial sums $\{\sum_{k=a}^{n} x_k\}$ converges. It follows from the algebra of limits that $\sum_{k=1}^{\infty} x_k$ converges if and only if $\sum_{k=a}^{\infty} x_k$ converges and

$$\sum_{k=1}^{\infty} = \sum_{k=1}^{a-1} x_k + \sum_{k=a}^{\infty} x_k = (x_1 + x_2 + \cdots + x_{a-1}) + \sum_{k=a}^{\infty} x_k.$$

**Example 2.2.** Let $|r| < 1$. Then

$$\sum_{k=0}^{\infty} r^k = \lim_{n\to\infty} \sum_{k=0}^{n} r^k$$

converges. To see this note

$$\begin{aligned} s_n = \sum_{k=0}^{n} r^k &= \frac{1-r^{n+1}}{1-r} && \text{sum of geometric series} \\ &\to \frac{1-0}{1-r} = \frac{1}{1-r} && \text{as } n \to \infty \text{ by algebra of limits.} \end{aligned}$$

Hence $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$.

**Example 2.3.** Let $0 < x < 1$ have decimal expansion

$$x = 0.d_1\, d_2\, d_3 \cdots$$

where $d_k \in \{0, 1, \ldots, 9\}$. Then the series $\sum_{k=1}^{\infty} d_k 10^{-k}$ has partial sums given by

$$s_n = \sum_{k=1}^{n} d_k 10^{-k} = \text{ expansion of } x \text{ to } n \text{ decimal places.}$$

Hence $\sum_{k=1}^{\infty} d_k 10^{-k} = \lim_{n\to\infty} s_n = x$.

The sum

$$\sum_{k=2}^{\infty} \left( \frac{1}{k-1} - \frac{1}{k} \right) \tag{2.1}$$

has partial sums given by

$$s_n = \sum_{k=2}^{n} \left( \frac{1}{k-1} - \frac{1}{k} \right) = \left( \frac{1}{1} - \frac{1}{2} \right) + \left( \frac{1}{2} - \frac{1}{3} \right) + \cdots = 1 - \frac{1}{n}.$$

Hence

$$\sum_{k=2}^{\infty} \left( \frac{1}{k-1} - \frac{1}{k} \right) = \lim_{n\to\infty} s_n = \lim_{n\to\infty} 1 - \frac{1}{n} = 1$$

and thereby

$$\sum_{k=2}^{\infty} \frac{1}{k^2-k} = \sum_{k=2}^{\infty} \left( \frac{1}{k-1} - \frac{1}{k} \right) = 1. \tag{2.2}$$

**Algebra of Series.** Suppose that $\sum_{k=a}^{\infty} x_n$ and $\sum_{k=n}^{\infty} y_n$ are *convergent* series. Then

$$\sum_{k=a}^{\infty} bx_n + cy_n = b \sum_{k=a}^{\infty} x_n + c \sum_{k=a}^{\infty} y_n.$$

PROOF. Let $s_n = \sum_{k=a}^{n} x_k$ and $t_n = \sum_{k=a}^{n} y_k$, so that

$$s_n \to s = \sum_{k=a}^{\infty} x_k \qquad \text{and} \qquad t_n \to t = \sum_{k=a}^{\infty} y_k$$

■

PROOF. Let $s_n = \sum_{k=a} x_k$ and $t_n = \sum_{k=a} y_k$, so that

$$s_n \to s = \sum_{k=a}^{\infty} x_k \qquad \text{and} \qquad t_n \to t = \sum_{k=a}^{\infty} y_k.$$

Then

$$\begin{aligned} \sum_{k=a}^{n} (bx_k + cy_k) &= b \sum_{k=a}^{n} x_k + c \sum_{k=a}^{n} y)k \\ &= bs_n + ct_n \\ &\to bs + ct \qquad \text{by Algebra of Limits} \\ &= b \sum_{k=a}^{\infty} x_k + c \sum_{k=a}^{\infty} y_k. \end{aligned}$$

■

Multiplication of series is much more difficult. Consider

$$s_2 t_2 = (x_1 + x_2)(y_1 + y_2) \quad \text{and} \quad s_3 t_3 \quad = (x_1 + x_2 + x_3)(y_1 + y_2 y + 3).$$

Rearranging into order for a series is a problem.

**Example 2.4.**

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{2}{3^n} + \frac{5}{(-4)^{n+1}} &= \sum_{n=0}^{\infty} 2\frac{1}{3^n} - \frac{5}{4} \cdot \frac{1}{(-4)^n} \\ &= 2 \sum_{n=0}^{\infty} \frac{1}{3^n} - \frac{5}{4} \sum_{n=0}^{\infty} \frac{1}{(-4)^n} \\ &= 2 \cdot \frac{1}{1 - \frac{1}{3}} + \frac{5}{4} \cdot \frac{1}{1 + \frac{1}{4}} \\ &= 3 + 1 = 4. \end{aligned}$$

**Lemma 2.5.** If $\sum x_n$ converges, then $x_n \to 0$ as $n \to \infty$.

PROOF. Since $\sum x_n$ converges, we have $s_n \to s$, say. Then $s_{n-1} \to s$ as well and so

$$x_n = s_n - s_{n-1} \to s - s = 0$$

by the algebra of limits. ■

**Example 2.6.** $\sum_{n=2}^{\infty} \frac{1}{n^2 - n}$ converges and so $\frac{1}{n^2-n} \to 0$.

The contrapositive is more useful.

**Example 2.7.** The contrapositive is more useful. $\psi \implies \varphi$ is equivalent to $\neg B \implies \neg A$. If $x_n \not\to 0$, then $\sum x_n$ does not converge. For example $(-1)^n \not\to 0$ and so $\sum_{n=1}^{\infty}(-1)^n$ does not converge.

There are series such that $x_n \to 0$ and $\sum_n x_n$ diverges.

**Proposition 2.8.** If $\{x_n\}$ is a sequence of nonnegative numbers, then $\sum_k x_k$ either converges or diverges to $\infty$. i.e.

$$s_n = \sum_{k=1}^{n} x_k n \to \infty \text{ as } n \to \infty.$$

PROOF. The sequence $\{s_n\}$ of partial sums is nondecreasing: $s_{n+1} = s_n + x_n \geqslant s_n$. Suppose that there is some $M \in \mathbb{R}$ with $s_n \leqslant M$ for every $n$. Then, by the Monotone Convergence Theorem $\{s_n\}$ converges, i.e. $\sum_k x_n$ converges.

Otherwise, for every $M$ there is an $N \in \mathbb{N}$ such that $s_N > M$. But then,

$$n \geqslant N \implies s_n = s_N + \sum_{k=N+1}^{n} \geqslant s_N \geqslant M$$

and so $s_n \to \infty$. ∎

$$\sum_{n=1}^{\infty} \frac{1}{n} \text{ diverges.}$$

PROOF.

$$\begin{aligned} s_n = \sum_{k=1}^{n} \frac{1}{k} &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots \\ &> 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{16} \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2} \\ &= 1 + \frac{\ell}{2} \quad \text{if } n = 2^{\ell} \end{aligned}$$

Given $M \in \mathbb{R}^+$, choose $R \geqslant 2M$ and put $n = 2^\ell$. Then

$$s_{2^\ell} = \sum_{k=1}^{2^\ell} \frac{1}{k} > 1 + \frac{\ell}{2} > M.$$

Hence $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$. ■

Even though in the example we have $\frac{1}{k} \to 0$ as $k \to \infty$, $\sum_{k=1}^{\infty} \frac{1}{l} = \infty$. We shall see that $\sum_k x_k$ converges if $x_k \to 0$ "fast enough".

$$\frac{1}{k} \to 0 \qquad \text{but not fast enough.}$$

$$\frac{1}{k^2 - k} \to 0 \qquad \text{fast enough}$$

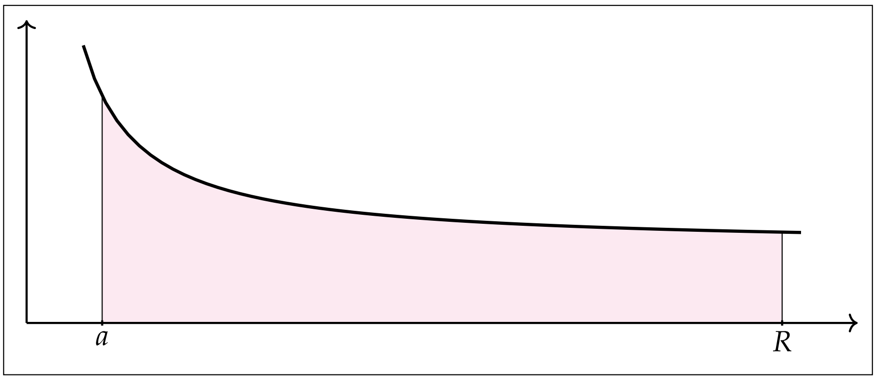
We will compare rates of convergence of sequences. Here is a good source of convergent series.

**The Integral Test.** Suppose $f : [a, \infty) \to (0, \infty)$ is a continuous and decreasing function. Then

$$\sum_{k=a}^{\infty} f(k) \text{ converges} \iff \lim_{R \to \infty} \int_a^R f(x)\,\mathrm{d}x \text{ exists.}$$

PROOF OF $\Longleftarrow$. Suppose that

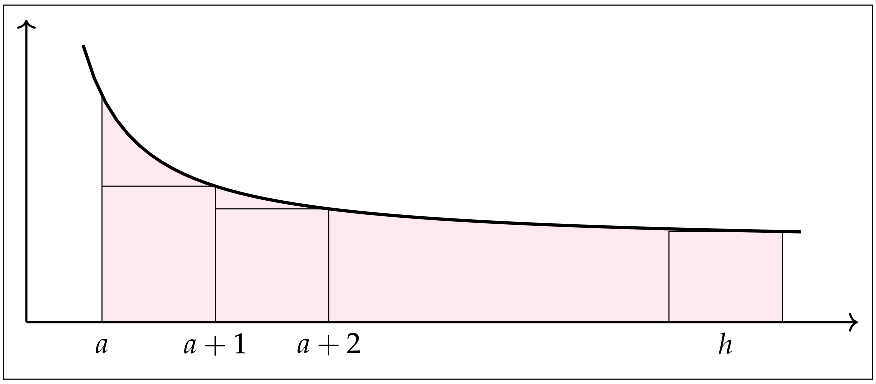
$$\lim_{R \to \infty} \int_a^R f(x)\,\mathrm{d}x = I < \infty.$$




$I$ is the area under the curve between $a$ and $R$.

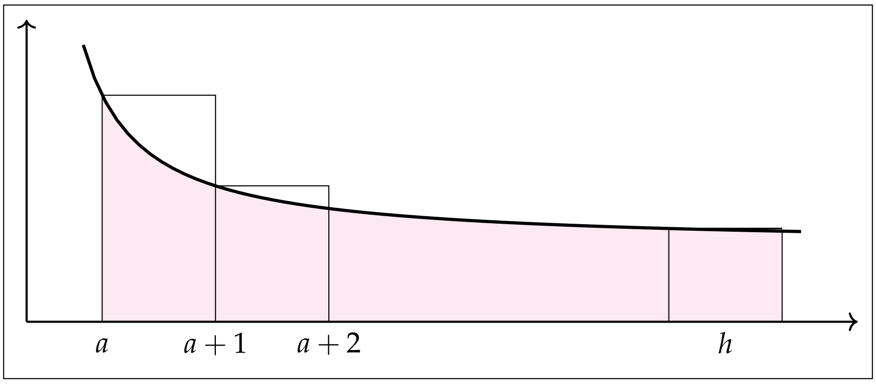
Then, $\forall n$,

$$\begin{aligned} s_n = \sum_{k=a}^{n} f(k) &= f(a) + \sum_{k=n+1}^{n} f(k) \\ &\leqslant f(a) + \int_a^n f(x)\,\mathrm{d}x \\ &\leqslant f(a) + I \end{aligned}$$




Hence $\{s_n\}$ is an increasing sequence which is bounded above. Therefore $\{s_n\}$ converges i.e. $\sum_{k=a}^{\infty} f(k)$ converges. ■

PROOF OF $\Longrightarrow$. Suppose that $\int_a^R f(x)\,\mathrm{d}x \to \infty$ as $R \to \infty$. On the other hand,




$$s_n \geqslant \int_a^{n+1} f(x)\,\mathrm{d}x \to \infty \text{ as } n \to \infty$$

and it follows that $s_n \to \infty$ as $n \to \infty$. Hence $\sum_k f(k)$ diverges to $\infty$. ■

**Example 2.9.** $\sum_{k=1}^{\infty} \frac{1}{k^s}$ converges $\iff$ $s > 1$, $\frac{1}{k^5} = f(k)$, where $f(x) = \frac{1}{x^5}$. We have

$$\int_1^R f(x)\,\mathrm{d}x = \int_1^R \frac{1}{x^5}\,\mathrm{d}x$$

$$= \begin{cases} \left[\dfrac{x^{-s+1}}{-s+1}\right]_1^R & s \neq 1 \\ [\ln x]_1^R & s = 1 \end{cases}$$

$$= \begin{cases} \dfrac{R^{-s+1}}{-s+1} - \dfrac{1}{-s+1} & s \neq 1 \\ \ln R & s = 1 \end{cases}$$

<u>**If** $s > 1$</u> then $-s+1 < 0$ and so

$$-\frac{1}{-s+1} \to \frac{1}{1-s} \quad \text{as} \quad R \to \infty.$$

Hence $\sum_{k=1}^{\infty} \frac{1}{k^s}$ converges when $s < 1$.

<u>**If** $s = 1$</u> then

$$\ln R \to \infty \quad \text{as} \quad R \to \infty.$$

Hence $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges to $\infty$. (Already seen directly.)

<u>**If** $s < 1$</u> then $-s+1 > 0$ and so

$$-\frac{1}{1-s} \to \infty \quad \text{as} \quad R \to \infty.$$

Hence $\sum_{i=1}^{\infty} \frac{1}{k^s}$ diverges to $\infty$ when $s < 1$.

Thus, in particular, $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges, ($\frac{1}{n^2} \to 0$ fast enough) and $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ diverges to $\infty$, ($\frac{1}{\sqrt{k}}) \to 0$, but no fast enough.)

These examples, and the geometric series, from the basis of the theory of convergent series. Many other series can be shown to be convergent (or divergent) by comparing with them.

**The Comparison Test.** Let $\{x_n\}$ and $(y_n)$ be sequences and suppose that there is an $N$ such that

$$\forall n \geqslant N; 0 \leqslant x_n \leqslant y_n$$

1. $\sum_n y_n$ converges $\implies$ $\sum_n x_n$ converges.
2. $\sum_n x_n$ diverges $\implies$ $\sum_n y_n$ diverges.

PROOF. For part 1. let

$$s_n = \sum_{k=N}^{n} x_n \qquad \text{and} \qquad t_n = \sum_{k=N}^{n} y_n.$$

Thus $\{s_n\}$ is a non-decreasing sequences and $s_n \leqslant t_n$ for all $n \geqslant N$. Since $\sum_n y_n$ converges, we have $t_n \to t$ for some and $t_n \leqslant t$ for all $n$ because $\{t_n\}$ is nondecreasing. Hence $s_n \leqslant t$ for all $n \geqslant N$ and so $\{s_n\}$ is a non-decreasing and bounded above sequence. Hence, by (the MCT) Proposition 1.8, so converges Therefore $\sum_n x_n$ converges.

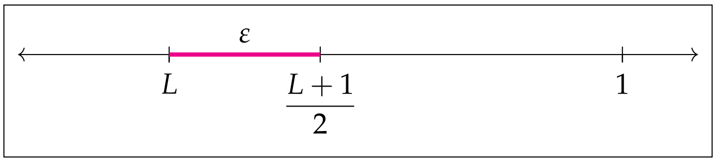
Part 2. is the contrapositive of 1. ■

**Example 2.10.** $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ diverges. We have $0 \leqslant \frac{1}{n} \leqslant \frac{1}{\sqrt{n}}$ for all $n$ and $\sum \frac{1}{n}$ diverges.

**The Ratio Test.** Let $x_n > 0$ for every $n$ and suppose that $\frac{x_{n+1}}{x_n} \to L$ as $n \to \infty$. Then

1. $L < 1 \implies \sum x_n$ converges,
2. $L > 1 \implies \sum x_n$ diverges,
3. $L = 1 \implies \sum x_n$ many converge or diverge.

PROOF OF 1. Since $L < 1$, we have $L < L + \frac{1-L}{2} = \frac{L+1}{2} < 1$.




Hence there is $N \in \mathbb{N}$ such that

$$\begin{aligned} n \geqslant N &\implies \frac{x_{n+1}}{x_n} < \frac{L+1}{2} \\ &\iff x_{n+1} < \frac{L+1}{2} x_n. \end{aligned}$$

It may be shown by induction that

$$x_{N+j} < \left(\frac{i+1}{2}\right)^j x_N, \quad j \geqslant 0.$$

The series

$$\sum_{i=N}^{\infty} \left(\frac{i+1}{2}\right)^{n-N} x_N$$

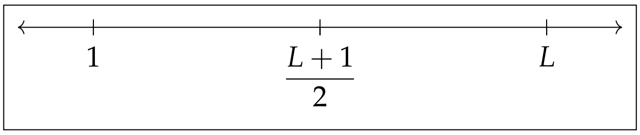
i.e. a geometric series and converges to $\frac{2x_N}{1-2}$ because $\frac{i+1}{2} < 1$. Hence $\sum_n x_n$

converges by the Comparison Test. ■

Thus the Ratio Test relies on comparing the given series with the basic example of the geometric series.

PROOF OF 2. Since $L > 1$, we have

$$1 < L - \frac{i-1}{2} = \frac{L+1}{2} < L.$$




Hence there is $N \in \mathbb{N}$ such that

$$\begin{aligned} n \geqslant N \implies \frac{x_{n+1}}{x_n} &> \frac{i+1}{2} \\ \iff x_{n+1} &> \left(\frac{i+1}{2}\right) x_n. \end{aligned}$$

Hence $x_{N+j} > \left(\frac{i+1}{2}\right)^j x_N$, $j \geqslant 0$. Since $\left(\frac{i+1}{2}\right) > 1$, it follows that $x_n \to \infty$ as $n \to \infty$. Hence $\sum_n x_n$ diverges. ■

$$\sum_{n=1}^{\infty} \frac{n^2}{2^n} \text{ converges}$$

PROOF. We have $x_n = \frac{n^2}{2^n}$ and so

$$\frac{x_{n+1}}{x_n} = \frac{\frac{(n+1)^2}{2^{n+1}}}{\frac{n^2}{2^n}} = \frac{1}{2}\left(1 + \frac{1}{n}\right)^2 \to \frac{1}{2} \quad \text{as} \quad n \to \infty.$$

Since $L = \frac{1}{2} < 1$, the series converges. ■

$$\sum_{k=0}^{\infty} \frac{2^n}{n!} \text{ converges}$$

PROOF. We have $x_n = \frac{2^n}{n!}$ and so

$$\frac{x_{n+1}}{x^n} = \frac{\frac{2^{n+1}}{(n+1)!}}{\frac{2^n}{n!}} = \frac{2}{n+1} \to 0 \quad \text{as} \quad n \to \infty.$$

Since $L = 0 < 1$, the series converges. ■

$$\sum_{n=1}^{\infty} \frac{n^n}{n!} \text{ diverges}$$

PROOF. We have $x_n = \frac{n^n}{n!}$ and so

$$\frac{x_{n+1}}{x_n} = \frac{\frac{(n+1)^{n+1}}{(n+1)!}}{\frac{n^n}{n!}} = \left\{1 + \frac{1}{n}\right\}^n \to e > 1.$$

■

$$\sum_{n=1}^{\infty} \frac{1}{n} \text{ diverges}$$

PROOF.

$$\frac{x_{n+1}}{x_n} == \frac{\frac{1}{n+1}}{\frac{1}{n}} = \frac{n}{n+1} \to 1 \text{ as } n \to \infty.$$

■

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converges}$$

PROOF.

$$\frac{x_{n+1}}{x_n} = \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} = \left(\frac{n}{n+1}\right)^2 \to 1$$

■

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converges}$$

PROOF.

$$\frac{x_{n+1}}{x_n} = \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} = \left(\frac{n}{n+1}\right)^2 \to 1 \text{ as } n \to \infty$$

■

**Remark.** Thus the series may converge or may diverge when $L = 1$.

Other tests for convergence of series include the $n$th root test and the limit form of the comparison test.

The comparison test applies to non-negative series and such series either convergence or they diverge to infinity. Convergence of series where not all terms have the same sign is more copmlicated.

**Proposition 2.11.** Let $\{x_n\}$ be a sequence of real numbers. If $\sum_n |x_n|$ converges, the $\sum_n x_n$ converges.

PROOF. We have that

$$\forall n; 0 \leqslant x_n + |x_n| \leqslant 2|x_n|$$

Since $\sum |x_n|$ is convergent $\sum 2|x_n|$ is a converged by Proposition **??**. Hence, by the comparison test, $\sum(x_n + |x_n|)$ is convergent.

We have $x_n = (x_n + |x_n|) - |x_n|$ and both $\sum(x_n + |x_n|)$ and $\sum |x_n|$ are convergent. Therefore,

$$\sum_n x_n = \sum_n (x_n + |x_n|) - |x_n|$$

is convergent by Proposition **??**. ■

**Corollary 2.12.** Let $\{x_n\}$ be a sequence of non-zero real numbers and suppose that $\frac{|x_{n+1}|}{|x_n|} \to L$.

1. $L < 1 \implies \sum x_n$ converges,
2. $L > 1 \implies \sum x_n$ diverges.

The converse of Proposition **??** is false. It is possible for $\sum x_n$ to converge while $\sum |x_n|$ does not.

**Example 2.13.** We have already seen that $\sum \frac{1}{n}$ diverges, but $\sum \frac{(-1)^{n+1}}{n}$ converges.

Consider the even partial sums

$$\begin{aligned} s_{ik} &= \sum_{n=1}^{ik} \frac{(-1)^{n+1}}{n} \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{2k-1} - \frac{1}{2k}\right) \end{aligned}$$

is an increasing sequence because

$$s_{2(k+1)} = s_{2k} + \left(\frac{1}{2k+1} - \frac{1}{2k+2}\right).$$

It is also bounded above because

$$s_{2k} = 1 - \left(\frac{1}{2} - \frac{1}{3}\right) - \left(\frac{1}{8} - \frac{1}{5}\right) - \cdots - \frac{1}{2k} < 1.$$

Hence $\{s_k\}$ converges, $s_{2k} \to s$ say.

For the odd partial sums we have

$$s_{2k+1} = s_{2k} + \frac{1}{2k+1} \to s + 0 = s \text{ as } k \to \infty$$

by the Algebra of Limits. Hence $s_k \to s$ as $k \to \infty$.

> The fact that, if the even terms and the odd terms of a sequence converge to the same limit, then the sequence as a whole converges to that limit too, must be proved.

Thus $\frac{(-1)^{n+1}}{n}$ converges.

**Definition 2.14.** Suppose that $\sum x_n$ is a series of real numbers.

1. $\sum |x_n|$ converges $\implies$ $\sum x_n$ converges absolutely, **converges absolutely**.
2. $\sum x_n$ converges but $\sum |x_n|$ diverges, we say that $\sum x_n$ **converges conditionally**.

> **Alternating Series Test.** Suppose that $\{x_n\}$ is a decreasing sequence (i.e. $x_n > x_{n+1}$, $\forall n$) and that $x_n \to 0$. Then the alternating series $\sum(-1)^{n+1} x_n$ converges.

PROOF. (The same argument as the example.)

Consider $s_{2k} = \sum_{n=1}^{2k} (-1)^{n+1} x_n$. $\{s_{2k}\}$ is an increasing sequence boudned by $x_1$. Hence $s_{2k} \to s$, say, and

$$s_{2k+1} = s_{2k} + x_{sk+1} \to s + 0 = s \text{ as } k \to \infty.$$

Therefore $\sum_{n=1}^{\infty} (-1)^{n+1} x_n$ sums to $s$. ∎

**Exercise 2.1.** $\sum \frac{(-1)^n}{n^2}$ converges absolutely

**Exercise 2.2.** $\sum \frac{1}{\sqrt{n}}$ diverges

**Exercise 2.3.** $\sum \frac{(-1)^n}{\sqrt{n}}$ converges

**Exercise 2.4.** Find an non-alternating series which converge conditionally.

Conditional convergence is much more delicate than absolute convergence. Care must be taken.

It can be shown that, if $\sum x_n$ s absolutely convergent, then we can rearrange the terms in any way and still get the same sum. e.g. Let $E = \sum x_{2n}$ be the sum of the even terms and $0 = \sum_{2n-1}$ be the sum of the odd terms. Then $\sum x_n = E + 0$ when $\sum x_n$ is absolutely convergent. (Also, sum of the positive and sum of negative.)

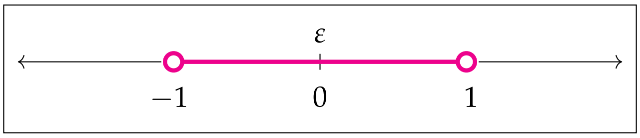
Consider $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ though. The sum of the even terms is

$$\sum_{k=1}^{\infty} \frac{(-1)^{2k+1}}{2k} = \sum_{k=1}^{\infty} \frac{-1}{2k} = -p$$

and the sum of the odd terms is

$$\sum_{k=1}^{\infty} \frac{(-1)^{2k}}{2k-1} = \sum_{k=1}^{\infty} \frac{1}{2k-1} = \infty.$$

In fact, the order in which the terms in the series are added can be rearranged so that it sums to any number at all.




For example, the series can be rearranged to sum to 2002: Add positive (odd) terms until the sum is just greater than 2002.

$$1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2\ell - 1} > 2002.$$

Then add negative (even) terms until the sum is just less than 2002.

$$1 + \frac{1}{3} + \cdots + \frac{1}{2k+1} - \frac{1}{2} < 2002.$$

Then add more positive terms:

$$() - \frac{1}{2} + () > 2002$$

$$(\cdots) - \frac{1}{4} < 2002.$$

and so on.

All terms are eventually included in a series which sums to 2002. The same is true of any conditionally convergent series.

**Proposition 2.15.** Let $\sum x_n$ be conditionally convergent. Then the terms can be rearranged so that the new series diverges or sums to any given value $S$.

SKETCH OF PROOF. The proof is essentially the same argument as in the example. If $\sum x_n$ is conditionally convergent, then

$$\sum(\text{positive terms}) = \infty \qquad \text{and} \qquad \sum(\text{negative terms}) = \infty$$

■

## 2.1 Power Series

**Power Series.** A **power series** is a series of the form

$$\sum_{n=0}^{\infty} a_n (x-c)^n \tag{2.3}$$

where $x$ is a variable and $a_n$ and $c$ are constants. A series in this form is said to be *centred at* $c$.

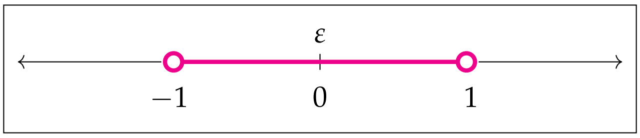
The series may converge for some values of $x$ and diverge for others. When it converges, the sum of the power series is a number, son, which depends on $x$. Thus $s$ is a function whose domain is the set of all $x$ such that 2.3 converges. Such power series functions are very useful.

The natural questions to ask are:

1. For which values of $x$ does a power series converge?
2. What are the properties of functions given as the sum of a power series?
3. Which functions can be realised as the sum of a power series?

We have already seen that the geometric series $\sum_{n=0}^{\infty} x^n$ converges for $|x| < 1$ and diverges for $|x| > 1$. For $-1 < x < 1$ we have

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$




Thus the series converges on an interval and the centre of the interval is the same as the centre of the series. The series diverges for $x$ outside the interval. We shall see that general power series behave in the same way.

The proofs of the theorems will not be given but they use the comparison test to compare a given power series with the geometric series.

**Proposition 2.16.** Suppose that $|a_n| \leqslant 2$ for every $n$. Then $\sum_{n=0}^{\infty} a_n(x-3)^n$ converges absolutely for $x \in (2, 4)$.

PROOF. We have $|a_n(x-3)^n| \leqslant 2|x-3|^n$ for all $n$. Since

$$\sum_n 2|x-3|^n = 2\sum_n |x-3|^n \text{ converges}$$

when $|x-3| < 1$, it follows by the comparison test that $\sum_n |a_n(x-3)^n|$ converges when $|x-3| < 1$. Hence $\sum_n a_n(x-3)^3$ converges when $|x-3| < 1$. ■

**Theorem 2.17.** Let $\sum_{n=0}^{\infty} a_n(x-c)^n$ be a power series centred at $c$. Then either

1. the series converges for all $x$; or
2. the series converges for $x = c$ only, or
3. there is a $R > 0$ such that the series converges for $|x-c| < R$ and diverges for $|x-c| > R$.

(Note. In case 3. we have not said what happens when $|x-c| = R$. Anything can happen.

**Radius of Convergence.** The power series

$$\sum_n^{\infty} a_n(x-c)^n$$

has **radius of convergence** equal to

$\infty$ when the series converges for all $x$,

$0$ when the series converges for $x = c$ only,

$R$ when $\exists R > 0$ such that the series converges for $|x - c| < R$ and diverges for $|x - c| > R$.

The set on which the series converges is called the **interval of convergence**. In case (1) the interval is $\mathbb{R}$. In case (2) the interval of conergence is $\{c\}$. In case (3) the interval of convergence will be $(c - R, c + R)$, $[c - R, c + R)$, $(c_R, c + R]$, or $[c_R, c + R]$. All possibilities may occur.

The ratio test may be used to find the radius of convergence. (Not useful for proofs of general results.)

**Question 2.18.** For which values of $x$ does the following series converge?

$$\sum_{n=5}^{\infty} \frac{n^2 - x^n}{3^{n-2}}.$$

ANSWER. The terms of the series are $a_N = \frac{n^2 x^n}{3^{n-2}}$. We have

$$\frac{|a_{n+1}|}{a_n} = \frac{\frac{(n+1)^2 x^{n+1}}{3^{n-1}}}{\frac{n^2 x^n}{3^{n-2}}} \to \frac{|x|}{3} \quad \text{as} \quad n \to \infty$$

The limit $L = \frac{|x|}{3} < 1$ when $|x| < 3$. Hence the series converges absolutely when $|x| < 3$. The limit $L = \frac{|x|}{2} > 1$ when $|x| > 3$. Hence the series diverges when $|x| > 3$.

When $x = 3$, the series is

$$\sum_{n=5}^{\infty} 9n^2$$

which diverges because $9n^2 \not\to 0$.

When $x = -3$, the series is

$$\sum_{n=5}^{\infty} (-1)^n 9n^2$$

which diverges because

$$(-1)^n 9n^2 \not\to 0.$$

Hence the interval of convergence is $(-3, 3)$. (The series is centred at 0). ◆

**Exercise 2.5.** 1. $\sum_{n=1}^{\infty} \frac{(x-2)^n}{n}$,

2. $\sum_{n=1}^{\infty} \frac{(2x)^n}{n^2}$,

3. $\sum_{n=0}^{\infty} \frac{x^n}{n!}$,

4. $\sum_{n=1}^{\infty} \frac{n^n}{(x+5)^n}$.

The answer to the first question is therefore that power series converge on an interval. Each power series therefore defines a function

$$F(x) = \sum_{n=0}^{\infty} a_n (x-c)^n$$

on its interval of convergence. What are the properties of functions defined in this way?

**Theorem 2.19.** Suppose that

$$f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n$$

converges on $(c-R, c+R)$. Then $f$ is differentiable on $(c-R, c+R)$ and

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x-c)^{n-1},$$

and this power series for $f'$ also converges absolutely on $(c-R, c+R)$.

The function $f$ has an antiderivative

$$\int_c^x f(t)\,\mathrm{d}t = \sum_{n=0}^{\infty} \frac{a_n (x-c)^{n+1}}{n+1}$$

and this power series is absolutely convergent on $(c-R < c+R)$.

PROOF. We will not prove this theorem in this course. It is harder than you might think! ■

**Example 2.20.** We have

$$x \in (-1,1) \implies \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right) = \sum_{n=1}^{\infty} nx^{n-1}.$$

Hence

$$x \in (-1,1) \implies \frac{1}{1-x}^{2} = \sum_{n=0}^{\infty}(n+1)x^n.$$

**Example 2.21.** Since

$$\frac{1}{1+x} = \sum_{n=0}^{\infty}(-1)^n x^n$$

we have

$$\int_0^x \frac{1}{1+t}\,\mathrm{d}t = \sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}x^{n+1}$$

or $\ln(1+x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n$, $-1 < x < 1$.

Note: Unlike $\sum_{n=0}^{\infty} x^n$, the series on the right hand side converges when $x = 1$ by the alternating series test. It can be shown that the equation holds for $x = 1$ as well:

$$\begin{aligned}\ln 2 &= \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n} \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \cdots .\end{aligned}$$

This requires a separate proof, the theorem does not imply it.

The theorem may be stated in words as “power series may be integrated and differentiated term by term”. Other operations on power series may be carried out term by term.

**Addition.** If $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ and $g(x) = \sum_{n=0}^{\infty} b_n(x-c)^n$ are two power series which both converge on $(c-T, c+T)$ (i.e. $T \leqslant$ the radius of convergence of both series)< then

$$f(x) + g(x) = \sum_{n=0}^{\infty}(a_n + b_n)(x-c)^n.$$

**Multiplication.** This is more complicated but works just as multiplication of polynomials does. e.g. $f(x) = 1 + x$ and $g(x) = 1 + x - 2^4$ are power series (with infinite radius of convergence).

$$f(x)g(x) = 1 + 2x + x^2 - 2x^4 - 2x^5$$

$$
\begin{aligned}
\frac{1}{1-x} &= 1 + x + x^2 + \cdots \\
\left(\frac{1}{1-x^2}\right)^2 &= 1 + x + x^2 + \cdots \\
&\quad + x + x^2 + x^3 + \cdots \\
&\quad\quad + x^2 + x^3 + \cdots \\
&= 1 + 2x + 3x^2 + \cdots .
\end{aligned}
$$

**Composition.** Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $(x) = \sum_{n=0}^{\infty} b_n x^N$ and suppose that $g(0) = b_0$ has modulus less than the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$. Then a *composite power series* is obtained:

$$
\sum_{n=0}^{\infty} a_n \left(\sum_{k=0}^{\infty} b_k x^k\right)^n .
$$

The series converges to $f \circ g(x)$.

**Example 2.22** (Very simple example)**.**

$$
\begin{aligned}
f(x) &= \frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots \\
g(x) &= x^2.
\end{aligned}
$$

Then

$$
\begin{aligned}
(f \circ g)(x) = \frac{1}{1+x^2} &= 1 - x^2 + (x^2)^2 - (x^2)^3 + \cdots \\
&= 1 - x^2 + x^4 - x^6 + \cdots
\end{aligned}
$$

This series converges when $x^2 \leqslant 1$, i.e. $-1 \leqslant x \leqslant 1$.

Integrating, we have

$$
\arctan x = \int_0^x \frac{1}{1+t^2}\,\mathrm{d}t = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots
$$

and when $x = 1$, this yields

$$
1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \arctan 1 = \frac{\pi}{4}.
$$

In all these operations power series behave like infinite degree polynomials and it often helps to PAGE CUT OFF 51.

**Example 2.23** (A less simple example). $\frac{1}{1+x+2x^2} = f \circ g(x)$ where $f(x) = \frac{1}{1+x}$, $g(x) = x + 2x^2$. Hence

$$\begin{aligned}
&\frac{1}{1+(x+2x^2)} \\
&\quad = 1 - (x+2x^2) + (x+2x^2)^2 + (x+2x^2)^3 + \cdots \\
&\quad = (1 - x - 2x^2 + \cdots) + (x^2 + 4x^3 + 4x^4 + \cdots) - (x^3 + 6x^4 + 12x^5 + 8x^6 + \cdots) + (x^4 + 6 \\
&\quad = 1 - x - x^2 + 3x^3 - x^4 - 4x^5 - \cdots
\end{aligned}$$

Converges for $|x + 2x^2| < 1$ i.e. $-1 < x < \frac{1}{2}$. Since the centre of convergence is 0, radius of convergence is in fact at least 1 for $-1 < x < 1$.

### 2.1.1 *Which functions may be represented as power series?*

If $f$ is represented by a power series on $(c-R, c+R)$, then by Theorem **??** $f$ is differentiable and $f'$ is the sum of a power series on $(c-R, c+R)$. Hence $f$ cannot be $|x-c|$ for example.

Applying the theorem to $f'$, we see that $f$ is twice differentiable. An induction argument shows that $f$ is infinitely differentiable on $(c-R, c+R)$. The coefficients of the power series can be recovered from the derivatives.

**Proposition 2.24.** Suppose that $f(x) = \sum_{n=0}^{\infty} a_n x^n$ on $(c-R, c+R)$. Then $f$ is infinitely differentiable on $(c-R, c+R)$ and

$$a_n = \frac{f^{(n)}(c)}{n!}.$$

PROOF. First of all,

$$f(c) = \sum_{n=0}^{\infty} a_n (0)^n = a_0 + a_1 0 + a_2 0 + \cdots = a_0.$$

Next, by Theorem **??**

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x-c)^{n-1} \implies f'(c) = a.$$

By induction,

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) a_n (x-c)^{n-k}.$$

Hence $f^{(k)}(c) = k! a_k$. ■

**Taylor Series.** Let $f$ be infinitely differentiable at $c$. Then the **Taylor series** for $f$ centred at $c$ is

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

When $c = 0$, the series $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}$ is also known as the **MacClaurin series** for $f$.

The $n$th sum of the Taylor series

$$T_n(f,c)(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

is called the degree $n$ **Taylor polynomial** for $f$ (centred at $c$).

Note:

1. The degree $n$ Taylor polynomial is the unique degree $n$ (or less) polynomial $p$ such that $p^{(k)}(c) = f^{(k)}(c)$ for $k = 0, 1, \ldots, n$.
2. $y = T_1(f,c)(x)$ is the equation of the straight tangent to $y = f(x)$ at $x = c$.

PICURE PAGE 55.

**Question 2.25.** Does the Taylor series for $f$ at $c$ converge to $f$ near $c$?

ANSWER. In general, no.

However, for most familiar functions it does and is used to compute the functions. The next result is used to prove convergence of Taylor series to the functions. ◆

**Taylor's Theorem.** Suppose that $f$ is $n+1$ times continuously differentiable on $(c-R, c+R)$. Then for each $x \in (c-R, c+R)$, there is $d$ between $x$ and $c$ such that

$$\begin{aligned} f(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + \frac{f^{(n+1)}(d)}{(n+1)!}(x-c)^{n+1} \\ &= \text{degree } n \text{ Taylor polynomial at } c + \text{Remainder term.} \\ &= T_n(f,c) + R_n(x). \end{aligned}$$

It is always possible to write

$$f(x) = \text{Taylor polynomial} + \text{something}.$$

The point of the theorem is that it tells what the something is. This allows us to estimate the size of the error if we approximate $f$ by its Taylor polynomial.

The proof of Taylor's Theorem is beyond the scope of this course.

**Example 2.26.** Let $f(x) = \sin x$.

$$f^{(n)}(x) = \begin{cases} \cos x & n = 4j+1 \\ -\sin x & n = 4j+2 \\ -\cos x & n = 4j+3 \\ \sin x & n = 4j. \end{cases}$$

For the Taylor series centred at 0,

$$\begin{aligned} |R_n(x)| &= \left| \frac{\pm \sin d \text{ or } \pm \cos d}{(n+1)!} (x-c)^n \right| \\ &\leqslant \frac{|x-c|^n}{(n+1)!} \to 0 \quad \text{as} \quad n \to \infty \end{aligned}$$

and

$$f^{(n)} = \begin{cases} 1 & n = 4j+1 \\ 0 & n = 4j \text{ or } 4j+2 \\ -1 & n = 4j+3. \end{cases}$$

Hence, for odd $n$, the degree $n$ Taylor polynomial is

$$\begin{aligned} T_n(\sin, 0)(x) &= x - \frac{1}{3!}x^3 + \frac{1}{5!} + \cdots \pm \frac{1}{n!}x^n \\ &= \sum_{j=0}^{k} \frac{(-1)^j}{(2j+1)!} x^{2j+1}, \end{aligned}$$

where $n = 2k+1$.

Since $R_n(x) \to 0$ for all $x$, we have $\sin x = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{ij+1}$.

Taylor's Theorem allows us to estimate the error when we approximate by a Taylor polynomial. e.g. Use the degree 5 Taylor polynomial to approximate $\sin 1$ and estimate the error. We have

$$T_5(\sin, 0)(1) = 1 - \frac{1}{3!} + \frac{1}{5!}$$

$$= 1 - \frac{1}{6} + \frac{1}{1210}$$
$$= 0.841\overline{6}.$$

The error is

$$|R_5(1)| = \left|\frac{\pm \sin d}{6!}\right| \leqslant \frac{1}{6!} = 0.0013\overline{8}.$$

In fact, $T_5(\sin, 0) = T_6(\sin, 0)$ and so the error is

$$|R_6(1)| = \left|\frac{\pm \cos d}{7!}\right| \leqslant \frac{1}{7!} = 0.00019\ldots$$

when $d \in (0, 1)$.

## 2.2 The Generalised Binomial Theorem

**Theorem 2.27** (Generalised Bionmail Theorem)**.** Let $r$ be nay real number. Then

$$(1+x)^r = \sum_{n=0}^{\infty} \binom{r}{n} x^n \quad \text{for} \quad |x| < 1$$

where $\binom{r}{n} = \dfrac{r(r-1)\cdots(r-n+1)}{n!}$.

(Note: For some values of $r$ the series converges for some $x$ with $|x| \geqslant 1$ as well.)

PROOF. The $n$th derivative of $(1+x)^r$ is

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(1+x)^r = r(r-1)\cdots(r-n+1)(1+x)^{n-r}.$$

When $x = 0$, this equals $r(r-1)\cdots(r-n+1)$.

It follows the coefficient of $x^n$ is

$$\frac{r(r-1)\cdots(r-n+1)}{n!} = \binom{r}{n}$$

To show that the series converges for $|x| < 1$ use the ration test. We have $a_n = \binom{r}{n}x^n$ and do

$$\frac{a_{n+1}}{a_n} = \left|\frac{\binom{r}{n+1}}{x^{n+1}}\binom{r}{n}x^n\right| = \left|\frac{r-n}{n+1}\cdot x\right| \to |x| \quad \text{as} \quad n \to \infty.$$

Therefore **??** converges when $|x| < 1$. ■

Notice we have not shown that the series converges to $(1+x)^r$. For this we must show that $\binom{r}{n}|x|^n \to 0$ as $n \to \infty$.

**Exercise 2.6.** Prove $\displaystyle\binom{r}{n}|x|^n \to 0$ as $n \to \infty$.

We check a couple of cases:

<u>$r = 1$</u>

$$\binom{-1}{n} = \frac{-1(-2)\cdots(-n)}{n!} = (-1)^n$$

and hence the binomial theorem in this case is;

$$(1+x)^{-1} = 1 - x + x^2 - \cdots,$$

That is to say, the sum of a geometric series.

<u>$r = \frac{1}{2}$</u>

$$\binom{\frac{1}{2}}{n} = \frac{\frac{1}{2}\left(-\frac{2}{1}\right)\left(-\frac{3}{2}\right)n!\cdots\left(\frac{3}{2}-n\right)}{n!}$$

Note that $\left|\binom{\frac{1}{2}}{n}| < \frac{1}{2n}\right|$ and so $|R_n(x) < \frac{|x|^n}{2n} \to 0$ whenever $|x| \leqslant 1$.

Hence $(1+x)^{\frac{1}{2}} = \displaystyle\sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}x^n$ for $|x| \leqslant 1$.

# 3 Differential Equations

**Example 3.1.** Population growth and compound interest are modelled by exponential growth.

**Example 3.2** (Concrete Situation)**.** A lump of lead at 100°C is out in a room where the temperature is 20°C. After one hour the lead is 60°C. What will its temperature be after three hours?

**Example 3.3** (Scientific Observation)**. Newton's Law of Cooling**, the rate of change of temperature of a cooling body is proportional to the difference between its temperature and the surrounding temperature.

Let $T(t)$ be the temperature of the lead at time $t$. Then

$$\begin{aligned} T(0) &= 100°\text{C} \\ T'(t) &= k(T(t) - T_a) \text{ for some k} \end{aligned}$$

where $T_a = 20°\text{C}$ is the *ambient temperature*.

## 3.1 Differential Equations

A **differential equation** (DE) is an equation involving an unknown function and its derivatives. An **ordinary differential equation** (ODE) is a DE in which the unknowns is a function of one variable.

**Example 3.4.** If $y$ is a function of $x$: $y' = 2y$.

**Definition 3.5** (Order)**.** The **order** of an ODE is the highest order derivative appearing.

**Example 3.6.** $y' = 2y$ is a first-order ODE

$$y'' + y' + y = x^2 + x + 1$$

is a second-order ODE.

**Definition 3.7.** A **solution** to an ODE is a function which satisfies the equation.

**Example 3.8.** $y = e^{2x}$ is a solution of $y' = y$ because

$$\begin{aligned} y' &= 2e^{2x} \quad \text{by chain rule} \\ 2y &= 2e^{2x} \end{aligned}$$

**Example 3.9.** $y = x^2 - x$ is a solution of $y'' + y' + y = x^2 + x + 1$ because

$$\begin{aligned} y &= x^2 - x \\ y' &= 2x - 1 \\ y'' &= x^2 + x + 1. \end{aligned}$$

**Example 3.10.** $y' = f(x)$ has solution $y = \int f(x)\,dx$. The solution is not unique, there is a constant of integration. We shall see that there are typically infinitely many solutions to an ODE.

Differential Equations generally arise from mathematical models.

picture page 64

### 3.1.1 *Application: Radioactive Decay*

**Concrete Situation:** A radioactive material (e.g. uranium) emits radiation as it decays to another element. (Uranium decays to thorium and then in a chain of other reactions to lead eventually.)
**Scientific Observation:** The rate of decay is proportional to the mass of radioactive material (provided there is not enough mass for a chain reaction). Let $M(t)$ be the mass of material at time $t$. The

$$M'(t) = kM(t)$$

**Mathematics:** Every solution to this equation has the form $M(t) = Ce^{kt}$ for some $C$. In fact, $C = M(0) =$ mass at time 0. Note that $k$ is a constant determined by measurement (experiment).

We can then *predict* how much radioactive material will remain in 100 years (say). This situation, where the rate of change of a function is proportional to the value of the function is very common.

**Proposition 3.11.** If $y(x)$ is a differentiable function such that $y'(x) =$

$ky(x)$, for some constant $k$, then

$$y(x) = y(0)e^{kx}.$$

PROOF. Consider $e^{-kx}y$. We have

$$\begin{aligned}\frac{d}{dx}\left(e^{-kx}y\right) &= -ke^{-kx}y + e^{-kx}y' && \text{product rule}\\ &= -ke^{-kx}y + e^{-kx}(ky) && y \text{ substituting the DE}\\ &= 0.\end{aligned}$$

Hence $e^{-kx}y = c$, constant. Hence $y(x) = ce^{-kx}$. Substituting $x = 0$ yield $c = y(0)$. ■

It follows that situations of this sort, where the rate of change is proportional to the value of the function lead to

**exponential decay** when $k < 0$ for which the derivative is negative, or

**exponential growth** when $k > 0$ for which the derivative is positive.

**Example 3.12 (Exponential Decay).** Radioactive decay. Newton's law of coding. Transmission of light.

**Mathematics:** Hence, if $D(t) = T(t) - T_a$, then $D'(t) = kD(t)$, $D(0) = 100 - 20 = 80$. Therefore,

$$D(t) = 80e^{kt}.$$

We do not know what $k$ is yet but have one more piece of information

$$T(1) = 60, \text{ and so } D(1) = 60 - 20 = 40.$$

Therefore, $40 = 80e^k$ which implies

$$e^k = \frac{1}{2}, \; k = \ln\frac{1}{2}$$

and so $D(t) = 80\left(\frac{1}{2}\right)^t$.

**Prediction:** At $t = 3$, $T(3) = T_n + D(3) = 20 + 80\left(\frac{1}{2}\right)^3 = 30°C$.

**Example 3.13 (Radioactive Decay).** $^{238}_{92}U$ has a half-life of $4 \cdot 5 \times 10^9$ year (about the age of the solar system) before it decays to the $^{234}_{90}Th$ (which is also radioactive and decays in steps to lead).

**Exercise 3.1.** How long does it take 10 grams of $^{235}_{92}U$ to decay to decay in 1 gram of $^{235}_{92}U$ (and other elements)?

**Scientific Observation:** The rate of decay of an element is proportional to the amount:

$$M'(t) = kM(t),$$

if $M(t)$ is the amount of material at time $t$. Hence $M(t) = M(0)e^{kt}$. At $t_1 = 4.5 \times 10^9$ years we have $M(t_1) = \frac{1}{2}M(0)$ (the time for half to decay is independent of $M(0)$ and so

$$e^{kt_1} = \frac{1}{2} \implies kt_1 = \ln\frac{1}{2} \implies k = \frac{\ln\frac{1}{2}}{4}.$$

**Prediction:** If $M(0) = 10$, the time $t$ when $M(t) = 1$ satisfies

$$\begin{aligned} 1 = 10e^{kt} &\implies \frac{1}{10} = e^{kt} \\ &\implies t = \frac{\ln\frac{1}{10}}{k} = \frac{\ln\frac{1}{10}}{\ln\frac{1}{2}} \times 4.5 \times 10^9. \end{aligned}$$

We have seen:

1. $y' = f(x) \implies y = \int f(x)\,dx + c$.
2. $y' = ky \implies y = ce^{kt}$.

In both cases the general solution has arbitrary constant. Further information is used to evaluate the constant. This is usually the value of the function at some value of $x$ (or $t$).

**Question 3.14.** How do we find the general solution?

ANSWER. There is no general method for finding closed solutions (like integration). Numerical techniques may be used. ◆

There are techniques which apply to certain commonly occuring types of ODE's.

## 3.1.2 Separable ODE's

**Separable ODE.** A 1st-order ODE of the form

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

is called **separable**.

To solve a separable ODE, we "separate the variables":

$$\begin{aligned} y' = f(x)g(y) &\iff \frac{1}{g(y)}y' = f(x) \\ &\implies \int \frac{1}{g(y)}y'\,dx = \int f(x)\,dx \\ &\implies \int \frac{1}{g(y)}\,dy = \int f(x)\,dx. \end{aligned}$$

Where we have used a change of variable on the left by substitution.

Integrating both sides gives

1. function of $y(x) =$ function of $x$,
2. solution for $y$ if possible.

**Question 3.15.** Solve $y' = y^2 = 1$.

ANSWER. Write as

$$\frac{dy}{dx} = y^2 + 1$$

separate variables

$$\frac{1}{y^2+1}\,dy = dx$$

Integrates

$$\begin{aligned} \int \frac{1}{y^2+1}\,dy = \int 1\,dx &\implies \arctan y = x + c \\ &\implies y = \tan(x+c). \end{aligned}$$

If we are given the initial value problem $y' = y^2 + 1$, $y(0) = 1$, then

$$1 = \tan(0+c) \implies c = \frac{\pi}{4}$$

and so $y = \tan\left(x + \frac{\pi}{4}\right)$. ◆

Notes:

1. Only put constant of integration on one side.
2. Put it in when integrating $y = \tan x + c$ is not a solution.
3. Check your answer. $y = \tan\left(x + \frac{\pi}{4}\right) \implies y' = \sec^2\left(x + \frac{\pi}{4}\right)$.

$$\begin{aligned} y^2 + 1 &= \tan^2\left(x + \frac{\pi}{4}\right) + 1 \\ &= \frac{\sin^2\left(x+\frac{\pi}{4}\right) + \cos^2\left(x + \frac{\pi}{4}\right)}{\cos^2\left(x+\frac{\pi}{4}\right)} \end{aligned}$$

$$= \sec^2\left(x + \frac{\pi}{4}\right).$$

Hence $y' = y^2 + 1$.

**Question 3.16.** Solve $y' = x^2 y(y-1)$.

ANSWER. ◆

**Question 3.17.** Solve $y' = 2xy$.

ANSWER. ◆

## 3.2 *First Order Linear ODEs*

**Linear ODE.** An ODE in the function $y$ is **linear** if all terms in $y$, $y'$, $y''$, ... are linear.

**Example 3.18.**

1. $y' + P(x)y = Q(x)$ is a 1st order linear ODE.
2. $y'' + P(x)y' + Q(x)y = R(x)$ is a 2nd order linear ODE.
3. $y' = y^2$ is not linear (but you know how to solve it because it is separable).

Equations such as the ones in Example 3.18 are called *linear* by analogy with the linear transformations you have been studying. In fact, they do correspond to a linear transformation on the (infinite-dimensional) space of functions. (See 3rd year courses on Hilbert Space and Linear Operators).

Every 1st order linear ODE may be put in the form $y' + P(x)y = Q(x)$. (If the coefficient of $y'$ is not 1, divide through by it.) To solve it, we use the method **integrating factor** which involves putting the equation into a different form. The **integrating factor** is the function

$$I(x) = \mathrm{e}^{\int P(x)\,\mathrm{d}x}.$$

We multiply through by this factor to obtain the equation.

$$\mathrm{e}^{\int P(x)\,\mathrm{d}x} y' + P(x)\mathrm{e}^{\int P(x)\,\mathrm{d}x} y = \mathrm{e}^{\int P(x)\,\mathrm{d}x} Q(x).$$

Notice now that $I'(x) = P(x)\mathrm{e}^{\int P(x)\,\mathrm{d}x}$ and so the equation is

$$I(x)y' + I'(x)y = \mathrm{e}^{\int P(x)\,\mathrm{d}x} Q(x).$$

The left-hand-side if $(I(x)y)'$, by the product rule and so

$$(I(x)y)' = \mathrm{e}^{\int P(x)\,\mathrm{d}x}Q(x).$$

Hence

$$I(x)y = \int \mathrm{e}^{\int P(x)\,\mathrm{d}x}Q(x)\,\mathrm{d}x + c$$

and so

$$\begin{aligned} y &= \frac{1}{I(x)}\left(\int \mathrm{e}^{\int P(x)\,\mathrm{d}x}Q(x)\,\mathrm{d}x + c\right) \\ &= \mathrm{e}^{-\int P(x)\,\mathrm{d}x}\left(\int \mathrm{e}^{\int P(x)\,\mathrm{d}x}Q(x)\,\mathrm{d}x + c\right). \end{aligned}$$

**Question 3.19.** Solve the initial value problem

$$y' + \frac{2}{x}y - x^3 = 0,\ y(1) = 1.$$

ANSWER. This is equivalent to

$$y' + \frac{2}{x}y = x^3,$$

so $P(x) = \frac{2}{x}$, $Q(x) = x^3$ and the integrating factor is

$$I(x) = \mathrm{e}^{\int \frac{2}{x}\frac{\mathrm{d}}{\mathrm{d}x}} = \mathrm{e}^{2\ln x} = x^2.$$

Multiplying by $I(x)$ gives the equation

$$x^2y' + 2xy = x^5$$

i.e. $(x^2y)' = x^5$. Integrating yields

$$x^2y = \frac{1}{6}x^6 + c$$

and so $y(x) = \frac{1}{6}x^4 + \frac{c}{x^2}$.

When $x = 1$ then

$$1 = \frac{1}{6} + \frac{c}{1}$$

and so $c = \frac{5}{6}$.

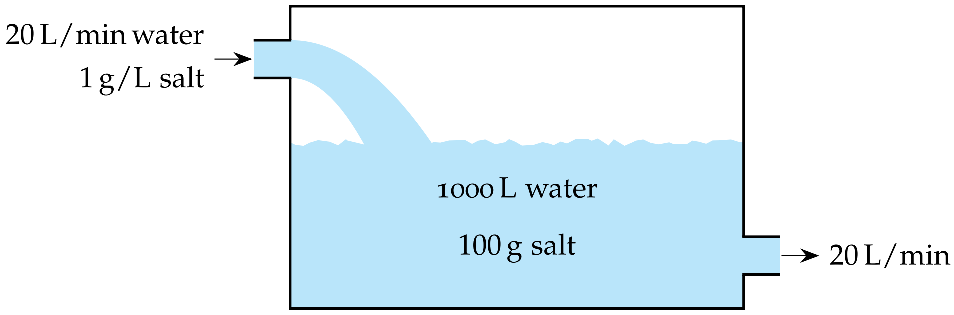
The solution to the initial value problem is

$$y(x) = \frac{1}{6} + \frac{5}{6x^2}.$$

◆

### 3.2.1 Mixing Problems

**Question 3.20.** A tank contains 1000L of water in which 100 gm of salt are dissolved. Solutions flows out of the tank at a rate of 20 L/min but

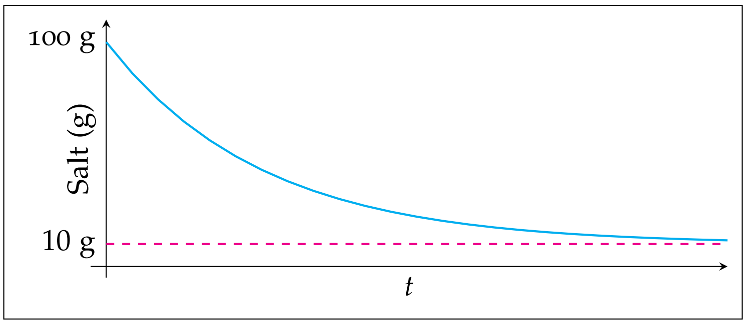



but a 1gm/100L solution flows in at 20 L/min.

Find the amount of salt in the tank as a function of time.

ANSWER. The initial concentration of salt is 10gm/100L. Over time, we expect the concentration to approach that of the incoming solution, i.e. 1gm/100L.

Total amount of salt will be $\frac{1}{100} \times 1000 = 10$ gm.




Let $A(t)$ = number of grams of salt in the tank after $t$ minutes and thereby $A(0) = 100$.

$$\begin{aligned} A'(t) &= (\text{in flow of salt}) - (\text{out flow of salt}) \\ &= \frac{1}{100} \times 20 - \frac{20}{1000} A(t). \end{aligned}$$

Hence

$$A'(t) + 0.02A(t) = 0.2.$$

The integrating factor is

$$I(t) = e^{\int 0.02\,dt} = e^{0.02t}.$$

Multiplying through by $I(t)$ yields

$$e^{0.02t}A'(t) + 0.02e^{0.02t}A(t) = 0.2e^{0.02t}$$

i.e. $\left(e^{0.02t}A(t)\right)' = 0.2e^{0.02t}$. Hence

$$e^{0.02t}A(t) = 10e^{0.02t} + c \implies A(t) = 10 + ce^{-0.02t}.$$

The initial condition yields

$$100 = 10 + ce^{0} \implies c = 90.$$

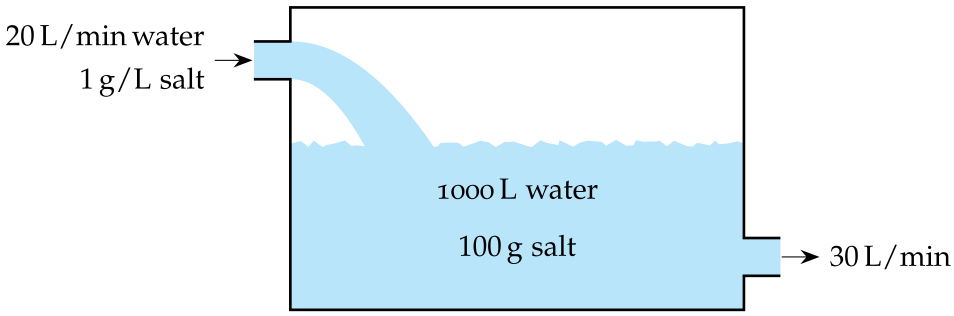
Hence

$$A(t) = 10 + 90e^{-0.02t}.$$

Note that $e^{-0.02t} \to 0$ as $t \to \infty$ and so $A(t) \to 10$ as $t \to \infty$, as expected.

◆

**Question 3.21.** Now suppose that 30L/min flows out.




Express the amount of salt in the tank as a function of time. In this question, the volume of water changes.

ANSWER. Let $V(t)$ = volume of solution after $t$ minutes then

$$\begin{aligned} V'(t) &= \text{in flow} - \text{out flow} \\ &= 20\text{L/min} - 30\text{L/min} \\ &= -10\text{L/min}. \end{aligned}$$

$$V(0) = 1000\text{L},$$

$$V(t) = 1000 - 10t$$

and after 100 minutes the tank is empty.

Let $A(t) =$ number of grams of salt in the tank after $t$ minutes then $A(0) = 1000$ and

$$A'(t) = \text{in flow} - \text{out flow} = \frac{1}{100} \times 20 - \frac{30}{1000 - 10t} A(t).$$

Hence

$$A'(t) + \frac{3}{100 - t} A(t) = 0.2.$$

$$I(t) = e^{\int \frac{3}{100-t}\, dt} = e^{-3\ln(100-t)} = (100 - t)^{-3}.$$

$$(100 - t)^{-3} A'(t) + 3(100 - t)^{-4} A(t) = 0.2(100 - t)^{-3}$$

that is

$$\left( \frac{A(t)}{(100 - t)^3} \right)' = 0.2(100 - t)^{-3}$$

◆

## 3.3 *The Logistic Equation*

**Logistic Equation.** The ODE

$$\frac{dp}{dt} = k(M - P)P$$

where $P$ is a function of $t$ and $k$ and $M$ are constants is called the **logistic equation**.

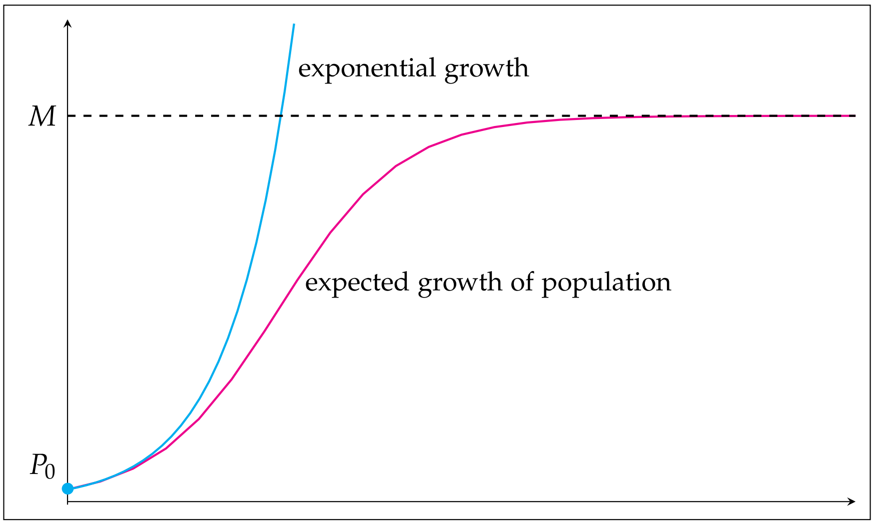
The logistic equation usually is used to model populations,

$$P(t) = \text{population at time } t$$
$$M = \text{max population which can be supported (e.g. by food supply).}$$

When $P$ is small we have

$$\frac{dP}{dt} \approx kMP, \; P(t) \approx P_0 e^{kMt}$$

and so population grown is exponential




As $P$ increases, the rate of growth decreases. As $P \to M$ the rate of growth levels out.

$$P' = k(M-P)P$$

is a separable ODE.

**Question 3.22.** Solve

$$\frac{\frac{dp}{dt}}{(M-P)P} = k$$

by separation of variables.

ANSWER.

$$\begin{aligned}
\int \frac{1}{(M-P)P}\frac{dP}{dt}\,dt &= \int k\,dt \\
\iff \int \frac{1}{(M-P)P}\,dP &= kt + C \\
\iff \frac{1}{M}\int \frac{1}{M-P} + \frac{1}{P}\,dP &= kt + c \\
\iff \frac{1}{M}[-\ln(M-P) + \ln P] &= kt + c \\
\iff \ln PM - P &= kMt + cM \\
\iff \frac{P}{M-P} &= e^{cM}e^{kMt} \\
\iff \frac{M}{P} - 1 &= e^{-cM}e^{-kMt} \\
\implies P(t) &= \frac{M}{1 + e^{-cM}e^{-kMt}}
\end{aligned}$$

We still have to find the integration constant $c$: Let $P(0) = P_0$. Then

$$P_0 = \frac{M}{1 + e^{-cm}} \implies e^{-cM} = \frac{M}{P_0} - 1.$$

Hence substitute in:

$$\begin{aligned} P(t) &= \frac{M}{1 + (\frac{M}{P_0} - 1)e^{-kMt}} \\ &= \frac{MP_0}{P_0 + (M - P_0)e^{-kMt}}. \end{aligned}$$

Therefore $P(t) < M$ for all $t$.

When $P_0 << M$,

$$P(t) \approx \frac{MP_0}{Me^{-kMt}} = P_0 e^{kMt}$$

(exponential growth). As $t \to \infty$, $e^{-kMt} \to 0$. Hence

$$P(t) \to \frac{MP_0}{P_0} = M \text{ as } t \to \infty.$$

◆

## 3.4 2nd Order Linear ODE's with constant coefficients

These are ODE's of the form

$$ay'' + by' + cy = f(x),$$

where $a, b, c$ are constants.

Such equations occur in models of electric circuits and also in models of weights on springs and pendulums. The same mathematics describes these physically quite different situations.

Note: going to 2nd order makes the linear ODE more difficult but we have simplified by assuming constant coefficients.

There is a general solution for such equations. We saw that the general solution of 1st order equations involves one constant (a constant of integration). The solution of 2nd order equations generally involves the constants (need to 'integrate' twice to find the solution).

In the case of 1st order equations, an initial value condition, $y(a) = b$, is used to determine the constant and find a unique solution.

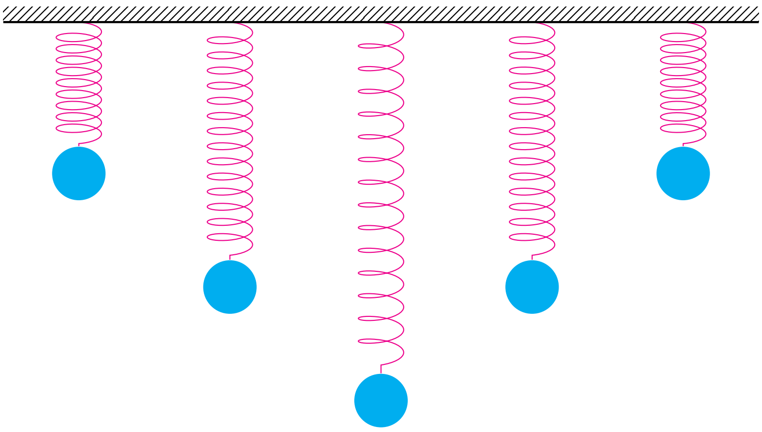
In the case of 2nd order equations, there are two types of conclusions use:

### *Boundary Value Problems*

Where $y(a) = y_1, y(b) = y_2$ Typically, we are interested in the values of the function between $a$ and $b$, given that we know its values at the endpoints (boundary).

### *Initial Value Problems*

Specifically $y(0)$ and $y'(0)$. Mass bouncing on a spring.



Given position and speed of the mass at time $t = 0$, describes its subsequent motion. Because the equation is linear the set of solutions to the ODE has a special form.

**Proposition 3.23.** Let $y_p$ be a particular solution to the ODE

$$ay'' + by' + cy = f(x).$$

Then ever solution to the equation has the form

$$y = y_p + y_n,$$

where $y_n$ is a solution to the *homogeneous equation*

$$ay'' + by' + cy = 0.$$

PROOF. Since the differential map $y \mapsto y'$ is a linear tfn on functions, i.e.

$$(y_1 + y_2)' = y_1' + y_2' \text{ and}$$
$$(\,dy)' = \,dy'$$

$d$ a constant. The proof is exactly the same as the proof that the general solution to the system of linear equations.

$$Ax = b$$

has the form $x = x_0 + x_n$, where $x_0$ is a particular solution and $Ax_n = 0$. ■

The similarly extends to the solution set of the homogenous equation as well.

**Proposition 3.24.**

1. The solutions of the homogeneous ODE

$$ay''' + by' + cy = 0$$

from a linear space, i.e. if $y_1$ and $y_2$ are two solutions, then so is $\alpha y_1 + \beta y_2$.

2. The solution space is two-dimensional. Thus if $y - 1$ and $y_2$ are independent solutions, then every solution has the form $\alpha y_1 + \beta y_2$. ($(y_1, y_2)$ is a basis for the space of solutions.)

PROOF. Exercise. ■

Solving the second order linear ODE with constant coefficients thus reduces to two steps

1. Find two independent solutions $y_1$ and $y_2$ to $ay'' + by' + cy = 0$.
2. Find a particular solution $y_p$ to $ay'' + by' + cy = f(x)$.

The general is then

$$y = y_p + \alpha y_1 + \beta y_2.$$

The constants $\alpha$ and $\beta$ are these emerging when we 'integrate' the second order equation:

| | |
|---|---|
| a B.V.P. | $y(a) = A,\ y(b) = B$ |
| or an I.V.P. | $y(0) = A,\ y'(0) = B,$ |

when substituted into the general solution

$$y = y_p + \alpha y_1 + p y_2,$$

give two linear equations in the two unknowns $\alpha$ and $\beta$.

In fact, all of the above discussion applies to linear second order ODEs with constant coefficients. However, with constant coefficients, the first problem is easily solved.

**Problem 1** Find two independent solutions of the homogeneous equation. As for the first order equation ($ay' + by = 0$), solutions have the form $y = \mathrm{e}^{mx}$. Then

$$y' = m\mathrm{e}^{mx} \qquad \text{and} \qquad y'' = m^2\mathrm{e}^{mx}.$$

Hence

$$ay'' + by' + cy = 0 \iff am^2\mathrm{e}^{mx} + bm\mathrm{e}^{mx} + c\mathrm{e}^{mx} = 0 \iff am^2 + bm + c = 0$$

(because $\mathrm{e}^{mx} \neq 0$ always).

This is a quadratic equation in the unknown $m$. It is called the **characteristic equation** of the ODE.

If the characteristic equation has two distinct real solutions, $m_1$ and $m_2$, then

$$y_1 = \mathrm{e}^{m_1 x} \qquad \text{and} \qquad y_2 = \mathrm{e}^{m_2 x}$$

are independent solutions of[1]

$$ay'' + by' + cy = 0$$

and the general solution is

$$y = \alpha\mathrm{e}^{m_1 x} + \beta\mathrm{e}^{m_2 x}.$$

If the characteristic equation has one real (repeated) root, $k$ say, then $\mathrm{e}^{kx}$ and $x\mathrm{e}^{kx}$ are independent solutions and the general solution is

$$y = \alpha\mathrm{e}^{kx} + \beta x\mathrm{e}^{kx} = (\alpha + \beta x)\mathrm{e}^{kx}.$$

**Lemma 3.25.** $x\mathrm{e}^{kx}$ is a solution.

[1] Prove $\mathrm{e}^{m_1 x}$ and $\mathrm{e}^{m_2 x}$ are independent.

PROOF. If the characteristic equation has a repeated root $m_1$, then it is

$$a(m-k)^2 = 0 \qquad \text{or} \qquad a(m^2 - 2km + k^2) = 0$$

and the ODE was

$$ay'' - 2aky' + ak^2y = 0.$$

**Exercise 3.2.** Check that these solutions are independent.

Try $y = xe^{kx}$. Then

$$y' = e^{kx} + kxe^{k+x} \qquad \text{and} \qquad y'' = 2ke^{kx} + k^2xe^{kx}.$$

Hence

$$\begin{aligned} &ay'' - 2aky' + ak^2y \\ &\quad = 2ake^{kx} + ak^2xe^{kx} - 2ake^{kx} - 2ak^2xe^{kx} + ak^2xe^{kx} \\ &\quad = 0. \end{aligned}$$

■

**Question 3.26.** Write down the general solution to

$$y'' + 3y' - 4y = 0.$$

ANSWER. The characteristic equation is given by

$$m^2 + 3m - 4 = 0 \iff (m+4)(m-1) = 0$$

and hence $m = -4$ or $m = 1$. The general solution is given by $y = \alpha e^{-4x} + \beta e^x$. ◆

**Question 3.27.** Write down the general solution to

$$y'' - 4y'_4y = 0.$$

ANSWER. The characteristic equation is given by

$$m^2 - 4m + 4 = 0 \iff (m-2)^2 = 0$$

and hence $m = 2$ is a repeated root. The general solution is given by $y = (\alpha + \beta x)e^{2x}$. ◆

What if the characteristic equation does not have real roots? Then the roots are complex and have the form $m = r \pm is$.

(Note: the ODE has real coefficients and so the complex roots are conjugates.)

We still try for a solution of the form

$$e^{mx} = e^{(r+is)x} \qquad \text{or} \qquad e^{(r-is)x}$$

but we have to make sense of this so that

$$\frac{d}{dx}\left(e^{(r+is)x}\right) = (r+is)e^{(r+is)x}$$

That is what is needed to make the argument using the characteristic equation work.

Define

$$\begin{aligned} e^{(r+is)x} &= e^{rx} \cdot e^{isx} \\ &= e^{rx}(\cos sx + i \sin sx). \end{aligned}$$

Then

$$\frac{d}{dx} e^{(r+is)x} = \cdots$$

Thus $e^{rx}(\cos sx + i \sin sx)$ and $e^{rx}(\cos sx - i \sin sx)$ are solutions, but are complex.

However

$$\frac{1}{2}\left(e^{(r+is)x} + e^{(r-is)x}\right) = e^{rx}\cos x$$

and

$$\frac{1}{2i}\left(e^{(r+is)x} - e^{(r-is)x}\right) = e^{rx}\sin x$$

and these are independent real solutions. Hence the general solution is

$$\begin{aligned} y &= \alpha e^{rx}\cos sx + \beta e^{rx}\sin sx \\ &= e^{rx}(\alpha\cos sx + \beta \sin sx) \end{aligned}$$

where $r \pm is$ are the solutions to the characteristic equation.

**Question 3.28.** Solve

$$y'' + y' + y = 0.$$

ANSWER. The characteristic equation

$$m^2 + m + 1 = 0 \iff m = \frac{-1 \pm \sqrt{-3}}{2} = -\frac{1}{2} \pm \frac{\sqrt{3}i}{2}$$

and so

$$y = e^{-\frac{1}{2}x}\left(\alpha\cos\frac{\sqrt{3}}{2}x + \beta\sin\frac{\sqrt{3}}{2}x\right)$$

is the general solution ◆

**Problem 2** Find a particular solution to

$$ay'' + by' + cy = f(x).$$

# 4 Calculus of Several Variables

## 4.1 Functions of Two Real Variables

These are functions defined on some $D \subseteq \mathbb{R}^2$. The function $f : D \to \mathbb{R}$ associates to each $(x, y) \in D$ a real number $f(x, y)$. $D$ is called the **domain** of $f$.

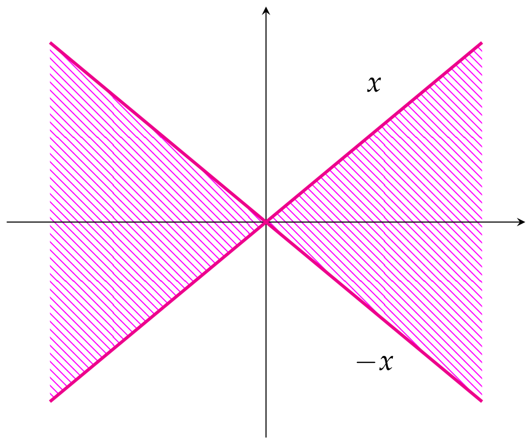
**Example 4.1.** The following are functions on two real variables:

$$f(x, y) = \text{distance of } (x, y) \text{ from } (1, 2)$$
$$f(x, y) = x^2 + y^2 - 1$$
$$f(x, y) = \sin(xy)$$
$$f(x, y) = \sqrt{x^2 - y^2}$$

all have domain $\mathbb{R}^2$.

**Question 4.2.** What is the domain of $f(x, y) = \sqrt{x^2 - y^2}$?

ANSWER.



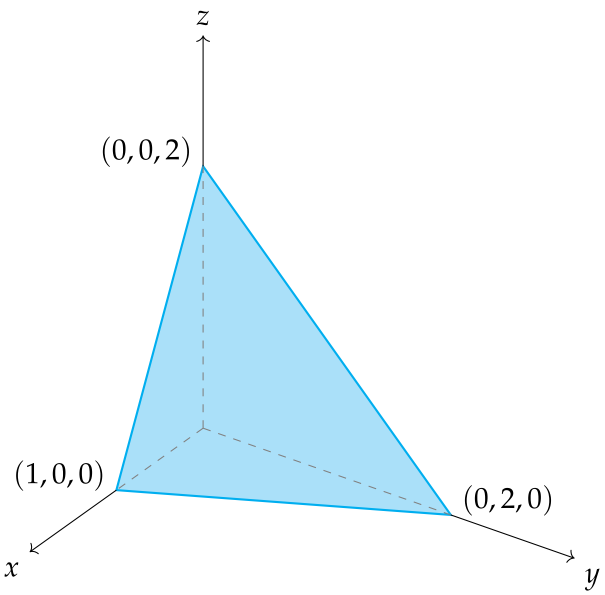

◆

**Graph.** The **graph** of $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ is the set of points given by

$$\{(x, y, z) : z = f(x, y),\ (x, y) \in D\} \subseteq \mathbb{R}^3.$$

The graph of a function of 1 variable is a curve in the plane. The graph of a function of two variables is a surface in three dimensional space — it is harder to draw and to imagine.

**Example 4.3.** Let $z = 2 - 2x - y$. This is the equation of a plane. First plot in the intercepts with the axes

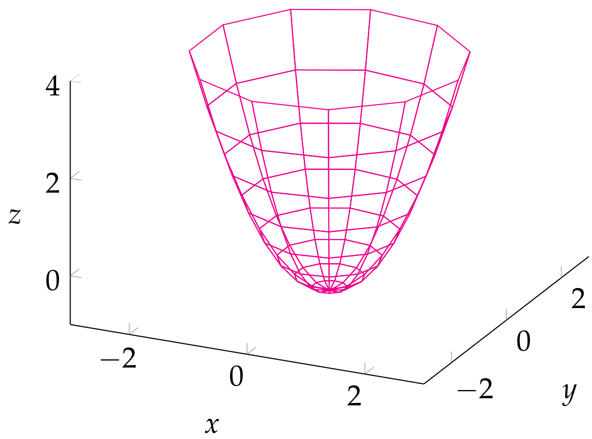



Note: the vertical line through each point in the domain intercepts the *graph* exactly once.

**Example 4.4.** $z =$ distance of $(x, y)$ from $(1, 2)$.

Picture page 102

**Example 4.5.** $z = x^2 + y^2 - 1$.




In the $(x,z)$-plane, (i.e. $y = 0$) and $z = x^2 - 1$ is a parabola.
In the $(y,z)$-plane, (i.e. $x = 0$) and $z = y^2 - 1$ is a parabola.

Some ideas for graphing functions of two variables was seen in the examples, namely, find the intercepts with the axes and graph the functions found by restricting to the axes.

Another idea is to look at the **contour liner** or **level sets** of the function. These are the curves in the $(x,y) - plane$ that satisfy $f(x,y) = c$.

**Example 4.6.** $f(x,y) = 2 - 2x - y$.

picture.

**Example 4.7.** $f(x,y) = \text{distance from } (x,y) \text{ to } (1,2) = ||(x,y) - (1,2)||$

picture page 105.

**Example 4.8.** $f(x,y) = x^2 + y^2 - 1$.

picture page 105.

Note: contours get closer together as the curve gets steeper.

You see contours of a two-variable function in a weather report.

page 106

Introduce coordinates to the map. Then the air pressure at each point (measured by a barometer) gives a function of two variables. The **isobars** seen on a weather map are curves of constant air pressure, i.e. are the contours of the air pressure function.

Making weather forecasts involves making a mathematical model which includes many functions of two (or more) variables and differential equations relating them. To predict the weather, solve the DEs.

## 4.2 Limits and Continuity

**Limit.** Suppose that $f(x,y)$ is defined for all $(x,y)$ near $(a,b)$.

$$f(x,y) \to L \text{ as } (x,y) \to (a,b) \iff$$
$$\forall \varepsilon > 0; \exists \delta > 0 : 0 \leqslant ||(x,y) - (a,b)|| < \delta \implies |f(x,y) - L| < \varepsilon.$$

Where $||(x,y) - (a,b)|| = \sqrt{(x-a)^2 + (y-b)^2}$.

Hence the definition says that $f(x,y)$ is close to $L$ wherever $(x,y)$ is sufficiently close to $(a,b)$. This may also be written:

$$\lim_{(x,y)\to(a,b)} f(x,y) = L.$$

Note: One important difference between limits of one-variable functions and limits of two-variable functions is that, in the one-variable case we may think of approaching $a$ from the right or left:

picture page 108

but in the two-variable case we may approach $(a,b)$ from infinitely many directions

picture page 108.

**Question 4.9.** Show $\lim_{(x,y)\to(1,1)} \frac{xy}{x^2+y^2} = \frac{1}{2}$.

ANSWER. We have

$$\frac{xy}{x^2+y^2} - \frac{1}{2} = \frac{-\frac{1}{2}(x-y)^2}{x^2+y^2}.$$

When $(x,y)$ is close to $(1,1)$, $x - y$ is close to $1 - 1 = 0$. The top line can be made as close to 0 as we wish by choosing $(x,y)$ sufficiently close to $(1,1)$. When $(x,y)$ is close to $(1,1)$, $x^2 + y^2$ is close to $1^2 + 1^2 = 2$. Hence

$$\frac{xy}{x^2+y^2} - \frac{1}{2}$$

can be made as close to 0 as we wish by choosing $(x, y)$ sufficiently close to $(1, 1)$. ◆

**Question 4.10.** Show $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ does not exist.

ANSWER. Consider what happens as we approach $(0, 0)$ from different directions.

picture page 109

Along the $x$-axis we have $y = 0$ and so $\frac{xy}{x^2+y^2} = 0$. Along the $y$-axis we have $x = 0$ and so $\frac{xy}{x^2+y^2} = 0$. However, observe the limit is *not* 0.

Approaching along the $y = x$ axis. Then

$$\frac{xy}{x^2+y^2} = \frac{x^2}{x^2+x^2} = \frac{1}{2} \not\to 0 \quad \text{as} \quad (x, x) \to (0, 0).$$

Note: the function need not be defined at $(a, b)$. ◆

**Continuous.** Suppose that $f(x, y)$ is defined at all $(x, y)$ near $(a, b)$ and also at $(a, b)$. Then $f$ is **continuous** at $(a, b)$ if

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b).$$

Moreover, $f$ is continuous on $D$ if it is continuous at every every point in $D$.

**Proposition 4.11.** The usual algebra of limits applies to functions of two (or more) variables. It follows that sums, products and quotients of continuous except where the denominator is zero.

**Corollary 4.12.** Polynomials in two variables are continuous everywhere in $\mathbb{R}^2$. Quotients of polynomials are continuous except where the denominator is zero.

**Example 4.13.**

1. $f(x, y) = \frac{xy}{x^2+y^2}$ is continuous everywhere except at $(0, 0)$.
2. $f(x, y) = \frac{x^2y^2}{x^2+y^2}$ is continuous everywhere except at $(0, 0)$.

In the second case however

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2} = 0.$$

Could use a similar argument to earlier example, or could use polar coordinates: Set

$$x = r\cos\theta \qquad \text{and} \qquad y = r\sin\theta$$

Then $(x, y) \to (0, 0)$ means $r \to 0$.

picture page 112

We have

$$\frac{x^2y^2}{x^2+y^2} = \frac{r^2\cos^2\theta r^2\sin^2\theta}{r^2} \to 0 \text{ as } r \to 0$$

(independent of the value of $\theta$).

## 4.3 Partial Derivatives

The derivative of a function of one variable gives a measure of the rate of change of the function as the variable increase. With functions of two or more variables the situation is more complicated because the variables may change in many directions. We will see that the general situation can be understood by varying the variables one at a time. This leads to the notion of a partial derivative.

**Partial derivative.** Let $f$ be a function defined at and near $(a, b) \in \mathbb{R}^2$. Then the **partial derivative** of $f$ with respect to $x$ at $(a, b)$ is

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h\to 0} \frac{f(a+h,\, b) - f(a,\, b)}{h}$$

and the partial derivative with respect to $y$ is

$$\frac{\partial f}{\partial y}(a, b) = \lim_{h\to 0} \frac{f(a,\, b+h) - f(a,\, b)}{h}.$$

Thus the partial derivatives are found by keeping one variable fixed and taking the derivative as usual with respect to the other.

**Question 4.14.** Let $f(x, y) = x^2 + y^3$. Find $\frac{\partial f}{\partial x}(1, 1)$.

ANSWER.

$$\frac{\partial f}{\partial x}(1, 1) = \lim_{h\to 0} \frac{(1+h)^2 + 1^3 - (1^2 + 1^3)}{h}$$

$$= \lim_{h \to 0} \frac{2h + h^2}{h} = 2.$$

◆

$\frac{\partial f}{\partial x}(a, b)$ is a number. Finding the partial derivative provided it exists at every point in the domain of $f$ gives as another 2-variable function $\frac{\partial f}{\partial x}$. We find $\frac{\partial f}{\partial x}$ by regarding $y$ as a constant and differentiating wrt $x$ in the usual way.

**Question 4.15.** Find $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for the functions

1. $f(x, y) = x^2y + y^3$,
2. $f(x, y) = \sin(x^2y + y^3)$.

ANSWER.

1. $$\frac{\partial}{\partial x}(x^2y + y^3) = 2xy,$$
   $$\frac{\partial}{\partial y}(x^2y + y^3) = x^2 + 3y^2.$$
2. $$\frac{\partial}{\partial x}\left(\sin(x^2y + y^3)\right) = \cos(x^2y + y^3) \cdot (2xy),$$
   and $$\frac{\partial}{\partial y}\left(\sin(x^2y + y^3)\right) = \cos(x^2y + y^3) \cdot (x^2 + 3y^2).$$

   Find the derivative, at $(1, 1)$ say, simply by substituting.

$$\frac{\partial}{\partial x}(x^2y + y^3)\Big|_{(1,1)} = 2xy\Big|_{(1,1)} = 2.$$

◆

The familiar differentiation rules for sums, product etc apply. We have used some in the above examples.

### 4.3.1 Higher Order Partial Derivatives

Since $\frac{\partial f}{\partial x}$ is a function of two variables, we may find its partial derivative with respect to $x$.

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

and also with respect to $y$.

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

There are four 2nd order partial derivatives

$$\frac{\partial^2 f}{\partial x^2}, \qquad \frac{\partial^2 f}{\partial y \partial x}, \qquad \frac{\partial^2 f}{\partial x \partial y}, \qquad \frac{\partial^2 f}{\partial y^2}.$$

**Example 4.16.**

1. $\dfrac{\partial}{\partial x}(x^2y+y^3) = 2xy,$
2. $\dfrac{\partial^2}{\partial x^2}(x^2y+y^3) = \dfrac{\partial}{\partial x}(2xy) = 2y,$
3. $\dfrac{\partial^2}{\partial y \partial x}(x^2y+y^3) = \dfrac{\partial}{\partial y}(2xy) = 2x,$
4. $\dfrac{\partial}{\partial y}(x^2y+y^3) = x^2+3y^2,$
5. $\dfrac{\partial^2}{\partial y^2}(x^2y+y^3) = \dfrac{\partial}{\partial y}(x^2+3y^2) = 6y,$
6. $\dfrac{\partial^2}{\partial x \partial y}(x^2y+y^3) = \dfrac{\partial}{\partial x}(x^2+3y^2) = 2x.$

In the example, $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$. Is this coincidence?

**Proposition 4.17.** Suppose that $f$ is a function of two variables and that both and order derivatives $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ exist and are continuous near $(a, b)$. Then

$$\frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b).$$

PROOF. Consider the function, $d$, of $h$ defined by

$$d(h) = f(a+h,\, b+h) - f(a+h,\, b) - f(a,\, b+h) + f(a,\, b).$$

If we unite $g_1(x) = f(x,\, b+h) - f(x,\, b)$, then

$$d(h) = g_1(a+b) - g_1(a).$$

By the *mean value theorem*, there is $c_1$ between $a$ and $a+h$ such that

$$\begin{aligned} d(h) = hg_1^1(c_1) &= h\left(\frac{\partial f}{\partial x}(c_1,\, b+h) - \frac{\partial f}{\partial x}(a,\, b)\right) \\ &= h\left(h\frac{\partial^2 f}{\partial y \partial x}(c_1,\, d_1)\right) \end{aligned}$$

for sure $d_1$ between $b$ and $b+h$.

On the other hand, if we write $g_2(y) = f(a+h, y) - f(a, y)$, then

$$\begin{aligned} d(h) &= g_2(b+h) - g_2(b) \\ &= h g_2'(d_2) && \text{some } d_2 \text{ between } b \text{ and } b+h \\ &= h\left(\frac{\partial f}{\partial y}(a+h, d_2) - \frac{\partial f}{\partial y}(a, d_2)\right) \\ &= h^2 \frac{\partial^2 f}{\partial x \partial y}(c_2, d_2) && \text{some } c_2 \text{ between } a \text{ and } a+h \end{aligned}$$

As $h \to 0$, $c_1, c_2 \to a$, and $d_1, d_2 \to b$. Hence

$$\begin{aligned} \frac{d(h)}{h^2} &= \frac{\partial^2 f}{\partial y \partial x}(c_1, d_1) \to \frac{\partial^2}{\partial y \partial x}(a, b), \quad \text{and} \\ \frac{d(h)}{h^2} &= \frac{\partial^2 f}{\partial x \partial y}(c_2, d_2) \to \frac{\partial^2}{\partial x \partial y}(a, b). \end{aligned}$$

These limit must be equal. Hence

$$\frac{\partial^2}{\partial y \partial x}(a, b) = \frac{\partial^2}{\partial x \partial y}(a, b).$$

■

There are many notations for the partial derivative:

$$\frac{\partial f}{\partial x} \qquad f_x \qquad f_1 \qquad D_1 f \qquad D_2 f.$$

## 4.4 Linear Approximation and tangent Planes

Partial derivatives have a geometrical interpretation, just as the derivative of a function of one variable does. The derivative of a 1-variable function is the slope of the tangent line to the graph. The **partial derivative** give the equation of the **tangent plane** to a surface.

**Tangent Plane.** Suppose that $f$ has continuous partial derivatives near the point $(x_0, y_0)$. Then the surface $z = f(x, y)$ has a **tangent plane** with equation

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

PROOF. The proof is omitted but note that, if the surface has a tangent plane, then its equation must be this one. The equation of a plane passing through $(x_0, y_0, z_0)$ is

$$z - z_0 = a(x - x_0) + b(y - y_0).$$

Differentiating this equation yields

$$\frac{\partial z}{\partial x} = a \text{ and } \frac{\partial z}{\partial y} = b$$

and so, if the plane is tangent to the surface, then

$$a = f_x(x_0, y_0) \text{ and } b = f_y(x_0, y_0).$$

■

**Question 4.18.** Find the tangent plane to the surface $z = -(2x^2 + y^2)$ at $(1, 1)$.

ANSWER. The level surfaces $2x^2 + y^2 = -c$ are ellipses. Cross-sections are parabolas. At $(1, 1)$, $z = -(2 + 1) = -3$.

$$\left.\frac{\partial - (2x^2 + y^2)}{\partial x}\right|_{(1,1)} = -4$$
$$\left.\frac{\partial - (2x^2 + y^2)}{\partial y}\right|_{(1,1)} = -2.$$

Equation of tangent plane is

$$z - (-3) = -4(x - 1) - 2(y - 1) \iff z = 3 - 4x - 2y.$$

◆

The part of the proof which is omitted is that of the statement "Then the surface $z = f(x, y)$ has a tangent plane." The statement of the proposition includes the hypothesis,

"Suppose that the partial derivatives are continuous."

This hypothesis is necessary and must be used in the proof.

**Example 4.19.** The function

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

has partial derivatives at every point in $\mathbb{R}^2$:

$$f_x(x,y) = \begin{cases} \frac{y(y^2-x^2)}{(x^2+y^2)^2} & (x,y) \neq (0,0) \\ 0 & x = (0,0). \end{cases}$$

$$\boxed{f_x(0,0) = \lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x-0} = \lim_{x\to 0} \frac{0-0}{x-0} = 0}$$

and

$$f_y(x,y) = \begin{cases} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

However $f_x$ and $f_y$ are not continuous at $(0,0)$ — and neither is $f$ as we have already seen.

$$z = \frac{xy}{x^2+y^2}$$

does not have a tangent plane at $(0,0)$.

When a surface has a tangent plane at a point, the plane may be used to approximate the surface near the point just as we approximate a curve by its tangent line in the 1-variable case. Thus the partial derivatives may be used to make linear approximations to functions of more than one variable.

Hence, if $f$ has a continuous partial derivatives near $(x_0, y_0)$. Then, for $\Delta x$ and $\Delta y$ small,

$$f(x+\Delta x, y_0+\Delta y) \approx f(x_0,\, y_0) + f_x(x_0,\, y_0)\Delta x + f_y(x_0,\, y_0)\Delta y$$

(c.f. equation of tangent plane).

**Example 4.20.** Let $f(x,y) = \sqrt{2x^2+y^2}$. Estimate $f(1.9,\, 1.1)$. Take $x_0 = 2$, $y_0 = 1$, $\Delta x = -0.1$, and $\delta y = 0.1$.

We have

$$f_x(x,y) = \frac{1}{2}(2x^2+y^2)^{-\frac{1}{2}} \cdot (4x)$$

$$f_x(2,1) = \frac{1}{2}(2\cdot 4+1)^{-\frac{1}{2}} \cdot 8 = \frac{4}{3}$$

$$f_y(x,y) = \frac{1}{2}(2x^2+y^2)^{-\frac{1}{2}} \cdot (2y)$$

so $f_y(2,1) = \frac{1}{2}\cdot\frac{1}{3}\cdot 2 = \frac{1}{3}$ and thus

$$\sqrt{2(1.9)^2+(1.1)^2} \approx \sqrt{2\cdot 4+1} + \frac{4}{3}(-0.1) + \frac{1}{3}(0.1) = 3 - 0.1 = 2.9.$$

## 4.5 The Chain Rule

**The Chain Rule.** Suppose that $f(x, y)$ has continuous partial derivatives and that $x = x(t)$, $y = y(t)$ are differentiable functions of the parameter $t$. Then $f(x(t), y(t))$ is a differentiable function of $t$ and

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

The corresponding formula holds for functions of more than two variables.

PROOF. Omitted. Recall that even for functions of one variable the proof of the chain rule is tricky. Given in MATH 2330 analysis. ■

There is a similar formula if $x$ and $y$ are functions of 2 variables, $s$ and $t$ say, with $x = s(x_0, t)$ and $y = y(s, t)$.

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

**Example 4.21.** Polar coordinate

$$x = r\cos\theta \qquad y = r\sin\theta$$

$$\frac{\partial x}{\partial r} = \cos\theta \qquad \frac{\partial x}{\partial \theta} = -r\sin\theta$$

$$\frac{\partial y}{\partial y} = \sin\theta \qquad \frac{\partial y}{\partial \theta} = r\cos\theta$$

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta \qquad \frac{\partial f}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}$$

(Away form $(x, y) = (0, 0)$.)

## 4.6 The Gradient

The equation for the approximation to $f(x_0 + \Delta x,\, y_0 + \Delta y)$ may be written in a vector form which makes it look more like the 1-variable version.

Let $\boldsymbol{x} = (x, y)$, $\boldsymbol{x}_0 = (x_0, y_0)$, $\Delta\boldsymbol{x} = (\Delta x, \Delta y)$. Then

$$f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) \approx f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot \Delta\boldsymbol{x},$$

where $\nabla f = (f_x, f_y)$ is called the **gradient** of the 2-variable function $f$.

**Gradient.** If $f = f(x_1, \ldots, x_n)$ is a function of $n$ variables the gradient of $f$ is the $n$-dimensional vector valued function

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

The same approximation formula holds in $n$-dimensions, provided that all partial derivatives are continupus next $\boldsymbol{x}_0$.

The gradient also appears in directional derivatives.

## 4.7 Directional Derivatives

The partial derivatives show the rate of change of a 2-variable function in the $x$- and $y$- directions. There is no reason why these directions should be preferred, $x$ and $y$, are often just a matter of how a problem has been parametrised. We may find a derivative in any direction.

**Directional Derivative.** Directions in $\mathbb{R}^2$ are specified by unit vectors,

$$\boldsymbol{u} = (r, s) = r\,\mathbf{i} + s\,\mathbf{j}, \quad r^2 + s^2 = 1.$$

The **directional derivative** of $f$ at $(a, b)$ in the direction of $\boldsymbol{u}$ is

$$\begin{aligned} D_{\boldsymbol{u}} f(a,b) &= \frac{\mathrm{d}}{\mathrm{d}h} f((a,b) + h\boldsymbol{u})\Big|_{h=0} \\ &= \lim_{h \to 0} \frac{f(a + hr,\, b + hs) - f(a,b)}{h}. \end{aligned}$$

Note: $D_{\boldsymbol{u}} f(a,b)$ is a number. It is the rate of change of $f$ in the direction of $\boldsymbol{u}$.

**Proposition 4.22.** Suppose that the partial derivatives of $f$ are continuous at $(a, b)$. Then

$$D_{\boldsymbol{u}} f(a,b) = \frac{\partial f}{\partial x} \cdot r + \frac{\partial f}{\partial y} \cdot s = f(a,b) \cdot \boldsymbol{u}.$$

PROOF. (Use the chain rule.)

$$\begin{aligned}
&\frac{\partial}{\partial h} f(a+hr,\, b+hs) \\
&\qquad = \frac{\partial f}{\partial x}(a+hr,\, b+hs) \cdot \frac{\mathrm{d}(a+hr)}{\mathrm{d}r} + \frac{\mathrm{d}f}{\mathrm{d}y}(a+hr,\, b+hs) \cdot \frac{\mathrm{d}(b+hs)}{\mathrm{d}h} \\
&\qquad = \frac{\partial f}{\partial x}(a+hr,\, b+hs) \cdot r + \frac{\mathrm{d}f}{\mathrm{d}y}(a+hr,\, b+hs) \cdot s.
\end{aligned}$$

■

**Exercise 4.1.** Find the direction derivative of

$$f(x,y) = x^3y - x^2y^2 + 2x^2 - 3y$$

at $(2, -3)$ in the direction of the vector $(1, 2)$.