# 6 Power series and Taylor's Theorem

Why are we interested in series? In analysis (or calculus, if you will), they provide a very important and rich class of functions, based on *power series*.

## 6.1 Taylor's Theorem and Taylor series

One of the most interesting theorems in analysis is the fact that one can *approximate* sufficiently differentiable functions just by knowing their multiple derivatives at a given point.

We already saw a glimpse of that by observing that if $f$ is differentiable at $x_0$, then $f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)$ with $r(x)$ "small." How can we generalize that to obtain better approximations?

We have also seen that if $f$ is a power series, it may be approximated by polynomials (its partial sums), and the coefficients of those polynomials are related to the derivatives of $f$ at the centre.

This will require some preparation.

### 6.1.1 General Rolle's Theorem

Recall that Rolle's Theorem states that for $f$ differentiable, if $f(a) = f(b)$ for some $a < b$, there must be $c \in (a, b)$ for which $f'(c) = 0$.

Is there an analogue for higher derivatives? Indeed, there is.

**Theoem (General Rolle's Theorem)**
Let $I = [a, b]$ be an interval, and $f$ a continuous function on $I$, $n$ times continuously differentiable on $(a, b]$, where $n \geq 0$ is an integer, and $f(a) = f(b)$. Suppose $f^{(n+1)}$ exists in at least $(a, b)$ and $f^{(k)}(b) = 0$ for $k = 1,2, \dots, n$. Then there is $c \in (a, b)$ such that $f^{(n+1)}(c)$=0.
A similar result holds for intervals $[a, b)$ with the roles of $a$ and $b$ interchanged. EOT.

Proof. We proceed by induction on $n$. If $n = 0$ this is Rolle's Theorem. Now suppose the theorem holds for a particular $n$, and let $a, b$ such that $f(a) = f(b)$ and $f^{(k)}(b) = 0$ for all $1 \leq k \leq n + 1$. We must show that there is $c \in (a, b)$ such that $f^{(n+2)}(c) = 0$.

By induction, there exists $c' \in (a, b)$ such hat $f^{(n+1)}(c') = 0$. Then $c' < b$ and applying Rolle's Theorem to $f^{(n+1)}$ on the interval $[c', b]$, there must be $c \in (c', b) \subseteq (a, b)$ such that ${f^{(n+1)}}'(c) = f^{(n+2)}(c) = 0$. QED.

**Corollary (of Proof)**
Let $I = [a, b]$ be an interval and $f$ a continuous function on $I$, $n + 1$ times differentiable on $[a, b)$, where $n \geq 0$ is an integer, and $f(a) = f(b)$. Suppose $f^{(k)}(a) = 0$ for $k = 1,2, \dots, n$. Then there is $c \in (a, b)$ such that $f^{(n+1)}(c) = 0$. EOC.

Proof. Exercise. QED.

Original Rolle was used to prove the Mean Value Theorem. General Rolle is then used to prove its natural generalization, known as Taylor's Theorem.

### 6.1.2 Taylor polynomials

**Definition**

Let $f$ be a function defined on an interval $I$. Let $c \in I$. If $f$ is $n$-times differentiable at $c$, then the polynomial $P_{f,n,c}(x) = f(c) + f'(c)(x-c) + \frac{1}{2}f''(c)(x-c)^2 + \cdots + \frac{1}{n!}f^{(n)}(c)(x-c)^n$ is called the **degree $n$ Taylor polynomial** (or $n$th Taylor polynomial) of $f$ **at $c$**. It is always a polynomial of degree **at most** $n$.

The polynomials $P_{f,n.c}$ are often also referred to as **Taylor expansions** of $f$ at $c$. We will refer to $c$ as the **centre** of the expansion. EOD.

**Exercise**

Show that if $f$ itself is a polynomial function, then $P_{f,c,n} = f$ for all $c \in \mathbb{R}$ as long as $n \geq \deg f$. EOE.

**Example**

1. Let $f(x) = \sin x$ defined on $\mathbb{R}$. Then $P_{f,0,n} = x - \frac{1}{6}x^3 + \frac{1}{105}x^5 + \cdots = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{(2k-1)!}x^{2k-1}$.
2. Let $f(x) = \log(1+x)$ defined on $(-1,1)$. Then $f'(0) = 1$, $f''(x) = \frac{1}{1+x}' = -\frac{1}{(1+x)^2} = -(1+x)^{-2}$, so $f''(0) = -1$. Continuing, $f^{(n)}(0) = (-1)^{n+1}\,(n-1)!$, and $P_{f,n,0} = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 \pm \cdots + \frac{(-1)^{n+1}}{n}x^n$.

EOE.

**Exercise**

Show that $P_{f,n,c}^{(k)}(c) = f^{(k)}(c)$ for all $1 \leq k \leq n$. EOE.

### 6.1.3 Taylor[1]'s Theorem

One of the remarkable facts of analysis is that one often can say a lot of meaningful things about the difference $\left|f(x) - P_{f,n,c}(x)\right|$. If that difference is small, then the Taylor polynomial can serve as a good approximation for $f$.

**Theorem (Taylor's Theorem; TT)**

Suppose $f$ is $n$ times continuously differentiable on an interval $[a,b]$, and suppose $f^{(n+1)}$ exists on at least $(a,b)$. For every $u \in [a,b)$ there is $d$ strictly between $u$ and $b$ (ie. $d \in (u,b)$) such that

$$f(u) - P_{f,n,b}(u) = \frac{(u-b)^{n+1}}{(n+1)!}f^{(n+1)}(d)$$

A similar theorem holds for intervals of the form $[a,b)$ with the roles of $a$ and $b$ interchanged.
EOT.

For $n = 0$, this is essentially the Mean Value Theorem, where we say a function $f$ is "0-times continuously differentiable" if it is continuous.

Proof. The idea (not mine) is to apply General Rolle. For this, the strategy is to find a function $h(x)$ on $[a,b]$ such that $h(u) = h(b)$ and $h^{(k)}(b) = 0$ for all $1 \leq k \leq n$, and such that $h^{(n+1)}(d) = 0$ for $d$ strictly between $u$ and $b$ if and only if $f(u) - P_{f,n,b}(u) = \frac{(x-b)^{n+1}}{(n+1)!}f^{(n+1)}(d)$.

[1] Brook Taylor (1685 – 1715)

Note that $g(x) = f(x) - P_{f,n,b}(x)$ (defined and $n$ times differentiable on $[a, b]$) satisfies that $g^{(k)}(b) = 0$ for $0 \leq k \leq n$. The only thing that is missing is that $g(u) \neq g(b) = 0$ in general.

To fix this define $h(x) = f(x) - P_{f,n,c}(x) - \frac{(x-b)^{n+1}}{(u-b)^{n+1}}\left(f(u) - P_{f,n,b}(u)\right)$.

Now $h(u) = 0 = h(b)$. Also $h^{(k)}(b) = 0$ for $1 \leq k \leq n$. So, there is $d$ between $u, c$ such that $h^{(n+1)}(d) = 0$. (See also the last exercise in the previous section.)

But $h^{(n+1)}(d) = f^{(n+1)}(d) - 0 - \frac{(n+1)!}{(u-b)^{n+1}}\left(f(u) - P_{f,n,b}(u)\right)$. Solving this equation for $f^{(n+1)}(d)$ gives the result. QED.

**Exercise**
Repeat the proof with the roles of $a$ and $b$ interchanged: show that with the same hypotheses as in TT, that for any $u \in (a, b]$, there is $d \in (a, u)$ such that

$$f(u) - P_{f,n,a}(u) = \frac{(u-a)^{n+1}}{(n+1)!} f^{(n+1)}(d)$$

EOE.

**Definition**
The "error term" $\frac{(u-b)^{n+1}}{(n+1)!} f^{(n+1)}(d)$ is often called the **Lagrange[2] remainder**, denoted $R_n$, or $R_n(u)$. Then $f(x) - P_{f,n,c}(x) = R_n$. (here $c = a, b$, or any fixed element of the domain of $f$. EOD.

**Remark**
Because of the "high" degree in $(u - b)$ the Lagrange remainder may seem large. But note that if $u$ is close to $c$, then $(u - b)^{n+1}$ is small, made even smaller by the division by $(n + 1)!$. Thus, the behaviour of $f^{(n+1)}(x)$ is crucial to analyzing the error term. EOR.

**Warning**
The $d$ in the Lagrange remainder $R_n$ depends on $u$ (or $x$, depending on notation) and the centre $c$. Different $u$ (or $c$) means different $d$ in general. Of course, $d$ may also depend on $n$. But it is usually not useful to include that in the notation. EOW.

**Exercise**
Let $f$ be a smooth function on $[a, b]$. If we know that for all $n \in \mathbb{N}$ $\left|f^{(n)}(x)\right| \leq \alpha C^{n+1}$ for some $\alpha, C > 0$ and all $n$, show that $\lim_{n\to\infty} R_n = 0$. EOE.

**Example**

1. Let $f(x) = \sin x$. Then for $k = 0,1, \ldots$ $f^{(2k)}(x) = (-1)^k \sin x$, and $f^{(2k+1)}(x) = (-1)^k \cos x$. It follows that the Lagrange remainder is always bounded by
$$\frac{|(x-c)^{n+1}|}{(n+1)!}$$
In particular, for fixed $x, c \in \mathbb{R}$, we have $\lim_{n\to\infty} P_{f,n,c}(x) = f(x)$ (why? Quotient Criterion).
2. Let $f(x) = \log(1 + x)$ on $(-1,1)$. Recall $f'(x) = \frac{1}{1+x}$, and for $n \geq 1$

[2] Joseph-Louis Lagrange (aka Guiseppe Luigi Lagrangia) (1736 – 1813)

$$f^{(n)}(x) = (-1)^{n+1}(n-1)!\,(1+x)^{-n}$$

Then for $u \in (-1,1)$, we have that the Lagrange remainder (with $b = 0$) is

$$R_n(x) = \frac{u^{n+1}}{(n+1)!}(-1)^{n+1}n!\,(1+d)^{-n-1}$$

for some $d$ strictly between $u$ and 0. Note that for the Lagrange remainder we get $\frac{|u|^{n+1}}{(n+1)|1+d|^{n+1}} \leq \frac{1}{n+1}$ as long as $x \in [0,1]$, and so $d \geq 0$. Thus, for such $x$, we have $\lim_{n\to\infty} R_n(x) = 0$.

It follows that for $x \in [0,1]$ we have

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

What about $x \in (-1,0)$?

EOE.

**Remark**

Let $f(x) = \log(1+x)$. We know that $f(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$ on $(-1,1)$. We also know that the series diverges for $x = -1$ (harmonic series), and still converges for $x = 1$ (alternating monotone sequence). Could it be that $f(1) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$? The Lagrange Remainder gives the answer: We know that $R_n(1) = f(1) - P_{f,n,0}(1) = \frac{(-1)^{n+1}1^{n+1}}{(n+1)(1+d)^{n+1}}$ for some $d \in (0,1)$. Observe here it is crucial that Taylor's Theorem applies for the *closed* interval $[0,1]$. It follows that $R_n(x) \to 0$ for $n \to \infty$. We conclude that

$$f(1) = \log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

EOR.

The fact that the Lagrange remainder may tend to zero and then $f$ can be evaluated as the limit of a series is enough to motivate a more general treatment of such series.

## 6.2 Review of series

We briefly review some topics regarding series that were discussed in Chapter 3.

### 6.2.1 Definition of series (Reminder)

A series is nothing but a very special form of sequence. Given any sequence $a_n$, we can form the associated series $\sum_{n=1}^{\infty} a_n$. What do we mean by that? Logically the series $\sum_{n=1}^{\infty} a_n$ is just the sequence of numbers $a_1, a_2, \ldots$, but we agree to compute the limit, if it exists, differently namely as

$$\sum_{n=1}^{\infty} a_n := \lim_{n\to\infty} \sum_{i=1}^{n} a_i$$

We could also think of the series as the sequence $a_1, a_1 + a_2, \ldots.$, that is the sequence of its **partial sums**. Note in this sense *any* sequence is a series: the sequence $b_1, b_2, \ldots$ corresponds to the series $\sum_{n=1}^{\infty} a_n$ where $a_1 = b_1, a_2 = b_2 - b_1, a_3 = b_3 - b_2, \ldots$

So $\sum_{n=1}^{\infty} a_n$ is *not* just the number $\lim_{N\to\infty} \sum_{n=1}^{N} a_n$ (if it exists) but also comprises the coefficients $a_1, a_2, \ldots$ In other words, typically two series are considered the same only if all the summands $a_n$ are the same. **It**

**is a slight abuse of notation to denote the series and its limit, if it exists, by the same symbols (namely $\boldsymbol{\sum_{n=1}^{\infty} a_n}$).** But it is very common. If we want to emphasize that we are talking about the value (limit) of a series, we may also write

$\lim \sum_{n=1}^{\infty} a_n$ instead of $\sum_{n=1}^{\infty} a_n$.

As mentioned before, a series should be thought of as the sequence of its partial sums (see below).

**Definition**
The series $\sum_{n=1}^{\infty} a_n$ is **convergent**, if its limit exists and is finite; otherwise, it is **divergent**. Like any sequence, a series can have limit $\pm\infty$.
A **partial sum** of the series $\sum_{n=1}^{\infty} a_n$ is a sum of the form $S_N = \sum_{n=1}^{N} a_n$. The partial sums form a sequence with the same limit (if it exists). EOD.

A series can start at any integer (just like a sequence can), meaning we can form series of the form $\sum_{n=k}^{\infty} a_n$ for any $k \in \mathbb{Z}$, but for simplicity we state most results with base index $1$.

### 6.2.2 Convergence of series

**Lemma**
Let $\sum_{n=1}^{\infty} a_n$ be a convergent series. Then $a_n$ is a zero sequence. EOL.

Recall that a *zero sequence* is a sequence that converges to $0$.

Proof. The sequence of partial sums $S_N$ is a Cauchy sequence. Thus, for $N$ sufficiently large $|a_{N+1}| = |S_{N+1} - S_N|$ is smaller than any predetermined $\varepsilon > 0$. QED.

The converse of the lemma is false:

**Example**
(We have seen this before.) The *harmonic series* $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges to $\infty$. To see this note that

$$\sum_{n=N+1}^{2N} \frac{1}{n} \geq N \frac{1}{2N} = \frac{1}{2}$$

The partial sums therefore cannot form a Cauchy sequence so must diverge. As it is monotone it must diverge to infinity. EOE.

**Remark**
If a series has only nonnegative terms it is monotone increasing. Then it converges if and only if the sequence of partial sums is bounded, by MBC. EOR.

One should also keep in mind that the limit of a series is nothing but the limit of the sequence of its partial sums. Thus, all convergence criteria for sequences apply to series just as well.

In particular, this applies to the Cauchy criterion:

**Lemma (Cauchy Criterion for Series)**
A series $\sum_{n=1}^{\infty} a_n$ is convergent if and only if for every $\varepsilon > 0$ there is $N_0 \in \mathbb{N}$ such that for all $n \geq m > N_0$, $|\sum_{k=m}^{n} a_k| < \varepsilon$. EOL.

Proof. This is the Cauchy criterion applied to the partial sums. QED.

### 6.2.3 Absolute convergence

**Definition.**
A series $\sum_{n=1}^{\infty} a_n$ is **absolutely convergent** if the series $\sum_{n=1}^{\infty}|a_n|$ converges. EOD.

Note if a series converges absolutely, it also converges by the Cauchy-Criterion:

Indeed, if the series converges absolutely, for any $\varepsilon > 0$ there is $n_0$ such that $\left|\sum_{n=N}^{M} a_n\right| \leq \sum_{n=N}^{M}|a_n| < \varepsilon$ as long as $N, M > n_0$, because the partial sums of the series $\sum_{n=1}^{\infty}|a_n|$ form a Cauchy sequence.

The converse is not true: $\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$ converges but does not converge absolutely.

Note that a series $\sum_{n=1}^{\infty} a_n$ converges absolutely if and only if $\sum_{n=1}^{\infty}|a_n| < \infty$, that is, if and only if the sequence of partial sums $S_N = \sum_{n=1}^{N}|a_n|$ is bounded (as mentioned above). Indeed, $S_N$ is a monotone increasing sequence and therefore convergent if and only if it is bounded.

### 6.2.4 The geometric series

One of the most important series is the *geometric series* $\sum_{n=0}^{\infty} a^n$ where $a$ is a real number.

**Proposition**

1. If $a \neq 1,\ \sum_{n=0}^{N} a^n = \frac{a^{N+1}-1}{a-1}$
2. $\sum_{n=0}^{\infty} a^n$ converges if and only if $|a| < 1$.

EOP.

Proof. The first part is an elementary induction on $N$ (done before). For the second part: if $|a| < 1$, then $\lim_{N\to\infty} a^N = 0$, and the first part shows that in this case,

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

If $|a| > 1$, the limit is not finite ($a > 1$) or does not exist ($a < -1$). If $a = 1$, the series clearly diverges to ∞. If $a = -1$, the partial sums alternate between $1$ and $0$. QED.

### 6.2.5 Convergence tests for series

The convergence of the geometric series is at the heart of the following convergence criterion, also known as the *Quotient Criterion* (and we have discussed this before in Chapter 3).

**Theorem (Ratio test for series)**
Let $\sum_{n=1}^{\infty} a_n$ be a series for which almost always $a_n \neq 0$. Suppose $R = \lim_{n\to\infty}\frac{a_{n+1}}{a_n}$ exists. Then

1. If $|R| < 1$, the series converges.
2. If $|R| = 1$ no statement can be made.
3. If $|R| > 1$, or $R = \pm\infty$, the series does not converge.

EOT.

Proof. This is the consequence of the following more general result. QED.

**Theorem (Ratio test, general version)**
Let $\sum_{n=1}^{\infty} a_n$ be a series with $a_n \neq 0$ for $n > N_0$. Let $L = \limsup_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$ and $M = \liminf_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$)

1. If $L < 1$, then the series converges absolutely.
2. If $M > 1$ the series does not converge.
3. If $L = 1$ or $M = 1$ no general statement can be made.

EOT.

**Remark**
To connect the ratio test above to the general ratio test here observe that for a convergent sequence $\limsup$ and $\liminf$ are the same and equal to the limit of the sequence. The only cases not strictly speaking covered are $R = \pm\infty$. But one can adapt the case $M = \infty$ to that ($a_n$ cannot be a zero sequence if $R = \pm\infty$). EOR.

Proof (of the general version). Suppose $L < 1$. Fix $L_0 < 1$ positive such that $L < L_0 < 1$. Then there is $N \in \mathbb{N}$ such that for all $n > N$, we have $\left|\frac{a_{n+1}}{a_n}\right| < L_0$. So $\frac{|a_{N+k}|}{|a_{N+1}|} = \frac{|a_{N+k}|}{|a_{N+k-1}|}\frac{|a_{N+k-1}|}{|a_{N+k-2}|}\cdots\frac{|a_{N+2}|}{|a_{N+1}|} < L_0^{k-1}$. Then $\sum_{n=N+1}^{N+k}|a_n| < \sum_{n=N+1}^{N+k}|a_{N+1}|L_0^{n-(N+1)} = \sum_{i=0}^{k-1}|a_{N+1}|L_0^i$. As $0 < L_0 < 1$, the limit of this for $k \to \infty$ exists and is finite. Thus, $\sum_{n=N+1}^{\infty}|a_n| < \infty$.It follows $\sum_{n=1}^{\infty} a_n$ is absolutely convergent.

If $M > 1$, then for $n$ large enough, we have $\left|\frac{a_{n+1}}{a_n}\right| > M - \varepsilon > 1$ (e.g. for $\varepsilon = \frac{M-1}{2}$). In particular this means $|a_{n+1}| \geq |a_n|$ for $n$ large. Then $|a_n|$ for large $n$ is a monotone increasing positive sequence. It therefore cannot be a zero sequence, and $\sum_{n=1}^{\infty} a_n$ cannot be convergent.

If $L = 1$ there are examples of absolutely convergent sequences, and of sequences that do not converge absolutely: Take $a_n = \frac{(-1)^n}{n}$. Then $L = 1$. The series does not converge absolutely, but it does converge. $a_n = \frac{1}{n^2}$ gives a series that its absolutely convergent, and $L = M = 1$.

$a_n = \frac{1}{n}$ gives an example of $L = M = 1$ that diverges. QED.

Note that if $a_n \geq 0$ for all $n$ absolute convergence is equivalent to convergence.

There is also a root test:

**Theorem (Root test for series)**
Let $\sum_{\mathrm{n}=1}^{\infty} a_n$ be a series and suppose $a_n \geq 0$. Let $L = \limsup \sqrt[n]{a_n}$.

1. If $L < 1$, the series converges.
2. If $L > 1$ the series diverges to $\infty$.
3. If $L = 1$ no general statement can be made.

EOT.

Proof. Suppose $L < 1$. Let $S_n = \sup\limits_{m\geq n} \sqrt[m]{a_m}$. Then for large enough $n$, $n > N_0$, say, we may assume that $S_n \leq L + \varepsilon$ for $\varepsilon = \frac{1-L}{2} > 0$. Then for such $n$ we also have $\sqrt[n]{a_n} \leq L + \varepsilon < 1$, and $a_n \leq (L+\varepsilon)^n$ (some UFOs). But then $\sum_{n=N_0}^{\infty} a_n \leq \sum_{n=N_0}^{\infty}(L+\varepsilon)^n < \infty$, since the geometric series converges.

If $L > 1$, then for $\varepsilon = \frac{L-1}{2}$, and each $n_0 \in \mathbb{N}$, there is (at least one) $n > n_0$ such that $\sqrt[n]{a_n} > L - \varepsilon > 1$. But then $a_n > 1$ and therefore $(a_n)$ is not a zero sequence. The series must therefore diverge.

Finally, for $L = 1$, the series is divergent if $a_n = 1$ for all $n$. But there are examples where it converges. QED.

**Exercise**

Show that $\lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1 = \lim_{n\to\infty} \sqrt[n]{\frac{1}{n^2}}$. EOE.

**Exercise**

Let $S = \sum_{n=1}^{\infty} a_n$ (here $S$ is the series not its limit) be any series. Let $b_n$ be a subsequence of $a_n$ constructed recursively as follows: Let $n_1 = \min\{n \mid a_n \neq 0\}$, $n_2 = \min\{n > n_1 \mid a_n \neq 0\}$, and so on, that is $n_{k+1} = \min\{ n > n_k \mid a_n \neq 0\}$; then $b_i = a_{n_i}$ for $i = 1,2, \ldots$

Show that $S$ and $\sum_{n=1}^{\infty} b_n$ have the same limit if either series has a limit. EOE.

This observation allows us to assume for many purposes that $a_n \neq 0$ for all $n$.

**Exercise**

Modify the arguments given in this section slightly to obtain the following criteria:

Let $\sum_{n=1}^{\infty} a_n$ be any series.

1. If there is $0 \leq q < 1$ such that $\left|\frac{a_{n+1}}{a_n}\right| \leq q < 1$ for almost all (that is, all but finitely many) $n$, then the series converges absolutely.
2. If there is $0 \leq q < 1$ such that $\sqrt[n]{|a_n|} \leq q < 1$ for almost all $n$, then the series is absolutely convergent.

EOE.

**Warning**

It is tempting to conclude that if $\left|\frac{a_{n+1}}{a_n}\right| < 1$ or $\sqrt[n]{|a_n|} < 1$for *almost all* $n$, then the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent. However, this need not be the case. Note that if $a_n = \frac{1}{n}$ and $b_n = \frac{1}{n^2}$, then $\frac{a_{n+1}}{a_n}, \frac{b_{n+1}}{b_n} < 1$ and $\sqrt[n]{a_n}, \sqrt[n]{b_n} < 1$ for all $n$. But $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.

In both cases though, the ratio and root criteria are inconclusive. EOW.

There are several other criteria for convergence of series.

**Compression Test**

Suppose $a_n \geq 0$ is monotone decreasing. Then $\sum_{n=1}^{\infty} a_n$ converges iff $\sum_{n=0}^{\infty} a_{2^n} 2^n$ converges. EOT.

Proof. Suppose $\sum_{n=0}^{\infty} a_{2^n} 2^n$ converges, to $L$, say. Since $a_n$ is monotone decreasing,

$$a_1 + (a_2 + a_3) + (a_4 + a_5 + \cdots + a_7) + \cdots + (a_{2^n} + a_{2^n+1} + \cdots + a_{2^{n+1}-1}) \leq$$

$$a_1 + 2a_2 + 4a_4 + \cdots + 2^n a_{2^n} \leq L$$

Thus $\sum_{n=1}^{2^{n+1}-1} a_n \leq L$ and therefore the series is bounded and convergent.

Conversely suppose $\sum_{n=1}^{\infty} a_n = M$. Then

$$a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + \cdots + a_8) + \cdots + (a_{2^{n-1}+1} + a_{2^{n-1}+2} + \cdots + a_{2^n})$$

$$\geq a_1 + a_2 + 2a_4 + \cdots + 2^{n-1} a_{2^n}$$

Thus $a_1 + \sum_{k=1}^{n} 2^{k-1} a_{2^k} \leq M$. But then $\sum_{k=1}^{n} 2^k a_{2^k} \leq 2M$ is still bounded, and $\sum_{k=0}^{\infty} a_{2^k} 2^k \leq 2M + a_1 < \infty$.

**Example**
The series $\sum_{n=1}^{\infty} \frac{1}{n^a}$ converges (and then absolutely) if and only if $a > 1$. (Here $a$ is any real number.) Indeed, using the compression test, $\sum_{n=0}^{\infty} \frac{1}{(2^n)^a} 2^n = \sum_{n=0}^{\infty} (2^{1-a})^n$, which is a geometric series, and therefore convergent if and only if $2^{1-a} < 1$. But this is equivalent to $1 - a < 0$, or $a > 1$. EOE.

**Definition**
A series $\sum_{n=1}^{\infty} a_n$ is called **bounded**, if its sequence of partial sums is bounded. A sequence $a_n$ is called of **bounded variation** if the series $\sum_{n=1}^{\infty} |a_{n+1} - a_n|$ converges. EOD.

**Fact**
If a sequence $a_n$ is of bounded variation, then it is convergent. EOF.

Proof. The series $\sum_{n=1}^{\infty} (a_{n+1} - a_n)$ is convergent as it is absolutely convergent by definition. Its partial sums $S_N$ satisfy that $S_N = a_{N+1} - a_1$. As $S_N$ is convergent, $a_N$ is. QED.

**Fact**
If $a_n$ is a monotone bounded sequence, then $a_n$ is of bounded variation. EOF.

Proof. Suppose $a_n$ is monotone increasing. Then $\sum_{n=1}^{N} |a_{n+1} - a_n| = \sum_{n=1}^{N} (a_{n+1} - a_n) = a_N - a_1$. This converges to a finite limit for $N \to \infty$ as $a_n$ is convergent. A similar argument works if $a_n$ is monotone decreasing (or apply this argument to the sequence $-a_n$). QED.

**Exercise**
Show that the sequence $\frac{(-1)^n}{n}$ is not of bounded variation. EOE.

**Dirichlet's Rule**
Let $\sum_{n=1}^{\infty} a_n$ be a bounded series and let $b_n$ be a monotone sequence with limit 0. Then $\sum_{n=1}^{\infty} a_n b_n$ converges. EOR.

Proof. Let $S_N = \sum_{n=1}^{N} a_n$. There is $B > 0$ such that $|S_N| \leq B$ for all $N$. Let $N > M > 1$ and $n \geq M$. Then $a_n = S_n - S_{n-1}$ for all $M \leq n \leq N$ and hence

$$\sum_{n=M}^{N} a_n b_n = \sum_{n=M}^{N} b_n (S_n - S_{n-1})$$

We can rewrite the right hand side as

$$S_N b_{N+1} - S_{M-1} b_M + \sum_{n=M}^{N} S_n (b_n - b_{n+1})$$

The absolute value of this can be bounded by $B|b_{N+1}| + B|b_M| + B \sum_{n=M}^{N} |b_n - b_{n+1}|$. For any $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that each of these summands is $< \frac{\varepsilon}{3}$ as long as $N, M > n_0$, as $b_n$ is of bounded variation and a zero sequence. QED.

**Leibniz Rule**

Let $a_n \geq 0$ be a monotone decreasing sequence with limit $0$.
Then $\sum_{n=1}^{\infty}(-1)^n a_n$ converges. EOR.

(This was a homework assignment in MATH 117.)

Proof. $\sum_{n=1}^{\infty}(-1)^n$ is a bounded series, so the result follows from Dirichlet's Rule. QED.

**Example**
$\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$ converges.

### 6.2.6 Rearranging series

Let $\sum_{n=1}^{\infty} a_n$ be a series. Since addition is commutative, it is tempting to assume one can rearrange the summands and have the same limit.

To make precise what this means, we say a **rearrangement** or **reordering** of the series $S = \sum_{n=1}^{\infty} a_n$ is a series of the form $S_\sigma \coloneqq \sum_{n=1}^{\infty} a_{\sigma(n)}$ where $\sigma\colon \mathbb{N} \to \mathbb{N}$ is a **bijection**, in this context often also called a **permutation**.of $\mathbb{N}$. Here $S, S_\sigma$ denote the **series**, and not its limit.

Analogous definitions hold for series where the index starts at some arbitrary integer (rather than 1).

**Notation**
Let $S \subseteq \mathbb{N}$ be a subset. If $a_n$ is a sequence we want to define the symbols $\sum_{\substack{n=1\\ n\in S}}^{N} a_n$ and $\sum_{\substack{n=1\\ n\in S}}^{\infty} a_n$.
The first one is defined as the sum of all $a_n$ for which *both*, $n \in S$ and $n \leq N$. In particular, it is equal to $0$ if $S$ is empty. The second needs more consideration. The *value* is defined as $\lim_{N\to\infty}\sum_{\substack{n=1\\ n\in S}}^{N} a_n$ (if that exists). If we want to think of it as a *series* we must specify a sequence $b_n$ such that $\sum_{\substack{n=1\\ n\in S}}^{\infty} a_n = \sum_{n=1}^{\infty} b_n$ as series. We do this as follows

$$b_n = \vdash \begin{cases} a_n & n \in S \\ 0 & n \notin S \end{cases}$$

We also write $\sum_{\substack{n=1\\ n\in S}}^{N} a_n$ for the partial sum $\sum_{n=1}^{N} b_n$.
EON.

**Example**
Let $S = \sum_{n=1}^{\infty} a_n$ be absolutely convergent. Let $P = \{\, n \mid a_n \geq 0\}$ and $Q = \{n \mid a_n < 0\}$. Then both $\sum_{\substack{n=1\\ n\in P}}^{\infty} a_n$ and $\sum_{\substack{n=1\\ n\in Q}}^{\infty} a_n$ are absolutely convergent.
Indeed, $\sum_{\substack{n=1\\ n\in P}}^{N} |a_n|$, $\sum_{\substack{n=1\\ n\in Q}}^{N} |a_n|$ are both bounded by $\sum_{n=1}^{N} |a_n|$. EOE.

**Theorem**
If a series $S$ is absolutely convergent, then so is any rearrangement $S_\sigma$ and the limit is the same. EOT.

Proof. Let $S = \sum_{n=1}^{\infty} a_n$ (the series, not the limit). We first treat the case that $a_n \geq 0$ for all $n$. Let $L$ be the limit of $S$.

Let $S_N = \sum_{n=1}^{N} a_n$. Then $S_N \leq L$ for all $N$. Let $\sigma$ be any permutation of $\mathbb{N}$. Since $S_\sigma$ has nonnegative summands we must show that $S_\sigma$ is bounded to conclude it is convergent.

For any $n$, let $\Pi_n = \{\,\sigma(1), \sigma(2), \dots, \sigma(n)\} \subseteq \mathbb{N}$. Let $\pi_n = \max \Pi_n$. Then for any $N \in \mathbb{N}$, $\sum_{n=1}^{N} a_{\sigma(n)} \leq \sum_{n=1}^{\pi_N} a_n \leq L$. This uses that $a_n \geq 0$ and that every summand on the left occurs on the right. But this means that $S_\sigma$ is bounded by $L$. And consequently $\sum_{n=1}^{\infty} a_{\sigma(n)} \leq L$.

It follows that $S_\sigma$ has a limit $M \leq L$. We have shown that for every convergent series with nonnegative coefficients, any reordering results in a series with a limit less than or equal to the limit of the original series.

On the other hand, $S$ is a reordering of $S_\sigma$. Indeed, $S = (S_\sigma)_\mu$ where $\mu$ is the inverse permutation defined by $\mu(n) = m$ if $\sigma(m) = n$. By the argument just given, this shows that $L \leq M$. It follows that $L = M$.

Now for the general case. If $S$ is absolutely convergent, we have established that $S_\sigma$ is also absolutely convergent (and therefore, convergent). It remains to show that the limits are equal.

For $n \in \mathbb{N}$, let $P = \{n \mid a_n \geq 0\}$ and $Q = \{n \mid a_n < 0\}$.

Then $S_N = \sum_{n=1}^{N} a_n = \sum_{\substack{n=1 \\ n\in Q}}^{N} a_n + \sum_{\substack{n=1 \\ n\in Q}}^{N} a_n = \sum_{\substack{n=1 \\ n\in P}}^{N} a_n - \sum_{\substack{n=1 \\ n\in Q}}^{N} |a_n|$.

Note that the limits for $N \to \infty$ of both summands on the right exist (as they are both bounded by $\pm \sum_{n=1}^{\infty} |a_n|$).

Thus $\lim\limits_{N\to\infty} S_N = \lim\limits_{N\to\infty} \sum_{\substack{n=1 \\ n\in P}}^{N} a_n - \lim\limits_{N\to\infty} \sum_{\substack{n=1 \\ n\in Q}}^{N} |a_n|$.

The same applies for $S_\sigma$. But note that the sets $P_\sigma = \left\{ n \mid a_{\sigma(n)} \geq 0 \right\}$ and $Q_\sigma = \left\{ n \mid a_{\sigma(n)} < 0 \right\}$ for $S_\sigma$ are $\sigma^{-1}(P)$ and $\sigma^{-1}(Q)$, respectively. Indeed $n \in P_\sigma$ iff $\sigma(n) \in P$, and similarly for $Q_\sigma$.

Therefore the series $\sum_{\substack{n=1 \\ n\in P_\sigma}}^{\infty} a_{\sigma(n)}$ is a reordering of the series $\sum_{\substack{n=1 \\ n\in P}}^{\infty} a_n$ and same for $\sum_{\substack{n=1 \\ n\in Q_\sigma}}^{\infty} |a_{\sigma(n)}|$ and $\sum_{\substack{n=1 \\ n\in Q}}^{\infty} |a_n|$. It follows that these two pairs have the same limits, respectively, by the first part. QED.

**Corollary**
If $a_n \geq 0$ is a sequence such that $\sum_{n=1}^{\infty} a_n = \infty$, then any reordering has the same limit. EOC.

Proof. If there is a rearrangement of the series that has a finite limit, that rearrangement is absolutely convergent, and hence the original series is. QED.

The following proposition shows that this is as good as it gets:

**Rearrangement Theorem**
Let $S = \sum_{n=1}^{\infty} a_n$ be convergent but not absolutely convergent. For any $L \in \mathbb{R}$ there is a permutation $\sigma$ such that $S_\sigma$ has limit $L$. EOP.

Proof. Let $Q = \{\, i \mid a_i < 0\}$ and $P = \{\, i \mid a_i \geq 0\,\}$. Note neither $P$ nor $N$ can be finite, otherwise $S$ is absolutely convergent. In fact $\sum_{\substack{n=1 \\ n\in P}}^{\infty} a_n = \infty$ and $\sum_{\substack{n=1 \\ n\in Q}}^{\infty} a_n = -\infty$. Because if either has a finite limit $\alpha$ say, then the other has also a finite limit $\beta$, say, as well, and the limit of $S$ is $\alpha + \beta$, and $S$ is absolutely convergent.

Let now $L$ be given, and we will assume $L \geq 0$. The case $L < 0$ is similar. Then there is $n_1$ minimal in $P$ such that $A_1 = P_1 := \sum_{\substack{n=1 \\ n \in P}}^{n_1} a_n > L$. Next there is $m_1 \in N$ minimal such that with $N_1 = \sum_{\substack{n=1 \\ n \in Q}}^{m_1} a_n$ we have $U_1 = P_1 + N_1 < L$. Next chose $n_2 > n_1 \in P$ minimal such that with $P_2 = \sum_{\substack{n=n_1+1 \\ n \in P}}^{n_2} a_n$ we have $A_2 = P_1 + N_1 + P_2 > L$. Continuing we $n_1, n_2, \ldots,$ and $m_1, m_2, \ldots$ and $P_i, N_i$ such that

$$A_i = P_1 + N_1 + P_2 + N_2 + \cdots + P_{i-1} + N_{i-1} + P_i > L$$

And $U_i = P_1 + N_1 + \cdots + P_i + N_i < L$.

Also $A_i - L \leq a_{n_i}$ and $L - U_i \leq -a_{m_i}$. This follows from the minimal choice of $n_i$ and $m_i$, respectively.

Indeed, if $A_i - L > a_{n_i}$, then $\left(A_i - a_{n_i}\right) > L$. But then we could choose a smaller number for $n_i$. Same for $L - U_i$.

We now define $\sigma: \mathbb{N} \to \mathbb{N}$ as follows: Think of $P$ and $N$ as two parallel lanes, and $n_i, m_i$ as the respective points where we change lanes.

Both $P, Q$ are infinite, and therefore we can list them by size $P = \{p_1 < p_2 < \cdots\}$ and $Q = \{q_1 < q_2 < \cdots\}$. For $n \in P$ let $p(n)$ be its position in $P$. That is $p(p_i) = i$. Similarly, for $n \in Q$, let $q(n)$ be its position in $Q$. Then $q(q_i) = i$.

For $1 \leq n \leq p(n_1)$ let $\sigma(n) = p_n$. For $p(n_1) < n \leq p(n_1) + q(m_1)$, let $\sigma(n) = q_{n-p(n_1)}$. Then change lanes again, and for $p(n_1) + q(m_1) < n \leq p(n_1) + q(m_1) + p(n_2)$ let $\sigma(n) = p_{n-p(n_1)-q(n_1)}$. Keep doing so. Every natural number appears exactly once as $\sigma(n)$ for some $n$, and therefore $\sigma$ is a permutation.

It follows that if $S$ is the original series, then $\lim S_\sigma = L$. Indeed, $S$ is convergent, so $a_n$ is a zero sequence. Let $\varepsilon > 0$. Then there is $n_0$ such that $|a_n| < \varepsilon$ as long as $n > n_0$.

As long as $i$ is large enough such that $n_i, m_i > n_0$ then the above shows that $A_i - L$ and $L - U_i$ is less than $\varepsilon$. Note that the partial sums of $\sum_{n=1}^{\infty} a_{\sigma(i)}$ are decreasing from $A_i$ to $U_i$ and increasing afterwards from $U_i$ to $A_{i+1}$

$$\sum_{n=1}^{p(n_1)} a_{\sigma(1)} = A_1$$

$$\sum_{n=1}^{p(n_1)+q(m_1)} a_{\sigma(1)} = A_1 + U_1$$

$$\sum_{n=1}^{p(n_1)+q(m_1)+p(n_2)} a_{\sigma(1)} = A_1 + U_1 + A_2$$

and so on, and in between the partial sums never increase their distance from $L$. It follows that the limit is $L$.

If $L < 0$ we apply the same reasoning to $-\sum_{n=1}^{\infty} a_n$ and $-L$. QED.

Note that if we modify the argument as follows, we can also achieve a limit of $\pm\infty$. We will show the case of $L = \infty$.

With $P, Q$ defined as in the proof, let $n_1 \in P$ be minimal such that $A_1 \coloneqq P_1 = \sum_{\substack{n=1 \\ n \in P}}^{n_1} a_n > 1$. Then let $m_1 \in Q$ be minimal such that with $N_1 = \sum_{\substack{n=1 \\ n \in Q}}^{m_1} a_n$ we have $U_1 = P_1 + N_1 < 1$.

If $A_k, U_k, P_k. N_k n_k, m_k$, have been defined we define $n_{k+1}$ as the minimum element $> n_k$ in $P$ such that with $P_{k+1} \coloneqq \sum_{\substack{n=n_k+1 \\ n \in P}}^{n_{k+1}} a_n$,

$$A_{k+1} \coloneqq U_k + P_{k+1} > k$$

Then $m_{k+1}$ is the smallest element $> m_k$ in $Q$ such that with $N_{k+1} = \sum_{\substack{n=m_k+1 \\ n \in Q}}^{m_{k+1}} a_n$ we have

$$U_{k+1} \coloneqq A_{k+1} + N_{k+1} < k$$

Note that both $n_k, m_k \to \infty$ if $k \to \infty$. In particular $|a_{n_k}|$ and $|a_{m_k}|$ are small for large $k$.

The minimality of $n_k, m_k$ then makes sure that $A_k < k + 1$ and $U_k > k - 1$.

We then rearrange the series as in the proof above. Then for $N \geq p(n_1) + q(m_1) + p(n_2) + q(m_2) + \cdots + p(n_k) + q(m_k)$ we have $\sum_{n=1}^{N} a_{\sigma(n)} \geq U_k > k - 1$. Thus, the limit is $\infty$.

### 6.2.7 Products of series

While we haven't yet discussed this in detail, the sum of two series $\sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} c_n$, where $c_n = a_n + b_n$, is again a series, and if $\alpha$ is any real number and $S = \sum_{n=1}^{\infty} a_n$ is a series, then $\alpha S$ is the series $\sum_{n=1}^{\infty} \alpha a_n$.

The situation is a bit more complicated for products of series. In this case it is useful to start the series at $n = 0$. Consider two series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$. Let $S_N, T_M$ be the partial sums, respectively.

Then $S_N T_M$ is a sum of products of the form $a_i b_j$ where $0 \leq i \leq N$ and $0 \leq j \leq M$. But there is no obvious order. Let $p_0, p_1, \ldots$ be any enumeration of all $a_i b_j$. That is each $a_i b_j$ appears once and only once among the $p_k$. More precisely, let $f\colon \mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0$ be a bijection, and for $(i, j) \in \mathbb{N}_0 \times \mathbb{N}_0$ let $p_{i,j} = a_i b_j$. Then $p_k \coloneqq p_{f(k)}$. We call the infinite series $\sum_{n=0}^{\infty} p_n$ a *product series* of the original two series.

For example we could use the antidiagonals in the scheme

*Equation 6-1*

$$\begin{matrix} a_0 b_0 & a_0 b_1 & a_0 b_2 & \cdots \\ a_1 b_0 & a_1 b_1 & a_1 b_2 & \cdots \\ a_2 b_0 & a_2 b_1 & a_2 b_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

$$p_0 = a_0 b_0, p_1 = a_0 b_1, p_2 = a_1 b_0, p_3 = a_0 b_2, \ldots$$

It is then a natural question whether $\sum_{n=0}^{\infty} p_n$ converges, and if so, what is the relationship to the the product of limits $(\sum_{n=0}^{\infty} a_n)(\sum_{n=0}^{\infty} b_n)$? Note that the Rearrangement Theorem destroys any hope that we *always* have equality even if we have convergence.

Note each *row series* $\sum_{n=1}^{\infty} a_n b_m$ converges ($m$ fixed). Similarly, each *column series* $\sum_{m=1}^{\infty} a_n b_m$ ($n$ fixed) converges.

If both series converge absolutely, however, then $|p_0| + |p_1| + \cdots + |p_k| \le \sum_{n=1}^{N} \sum_{m=1}^{M} |a_n||b_m|$ for $N, M$ large enough. And then $|p_0| + |p_1| + \cdots + |p_k| \le (\sum_{n=0}^{\infty} |a_n|)(\sum_{m=0}^{\infty} |b_m|)$, and so *every product series converges* absolutely. They all have the same limit as they are rearrangements of each other. Let $p$ be that common limit. We must show that

$$p = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) =: P$$

This follows from looking at top left squares in Equation 6-1 above. Labeling rows and columns starting at 0, and picking $p_0 = a_0 b_0 .\, p_1 = a_0 b_1, p_2 = a_1 b_1, p_3 = a_1 b_0, \ldots$

Thus we start at the top of column $n$ gow down until we hit the "diagonal" element, and then gow back on row $n$ to the 0th column. Thus, the elements $p_0, p_1, \ldots, p_{n^2-1}$ enumerate the $n^2$ elements in the top left "square" of side length $n$.

But then $p_1 + p_2 + \cdots + p_{n^2-1} = (a_0 + a_1 + \cdots + a_{n-1})(b_0 + b_2 + \cdots + b_{n-1})$.

This converges to $P$. But the left hand side is a subsequence of the sequence of partial sums of the series $\sum_{n=0}^{\infty} p_n$ which converges to $p$. Thus $p = P$.

**Theorem**
Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be absolutely convergent with limit $A$ and $B$, respectively. Then the *Cauchy product* of these two series

$$\sum_{n=0}^{\infty} (a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k}$$

converges absolutely and has limit $AB$. EOT.

Proof. The partial sums of this series form a subsequence of a product series. Any product series converges absolutely to $AB$. QED.

**Remark**
Note that not every product series must converge. Consider the alternating series

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

Since $\sqrt{n}$ is a monotone zero-sequence, we know that this series converges by the Leibniz rule. We also know that it does not converge absolutely since $\sum_{n=0}^{\infty} \frac{1}{\sqrt{n+1}} = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} = \infty$. Let $a_n = \frac{(-1)^n}{\sqrt{n+1}}$.
In this case the series $\sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k a_{n-k}$ does not converge. This series is a subsequence of a product series as in the previous theorem.
To see why it diverges note that

$$|a_0 a_n + a_1 a_{n-1} + \cdots + a_n a_0| = \left|\sum_{k=0}^{n} \frac{(-1)^k (-1)^{n-k}}{\sqrt{k+1}\sqrt{n+1-k}}\right|$$

Now $(-1)^k(-1)^{n-k} = (-1)^n$, so this can be pulled out, and we have

$$\sum_{k=0}^{n} \frac{1}{\sqrt{(k+1)(n+1-k)}} \geq \sum_{k=0}^{n} \frac{1}{\sqrt{(k+1)(n+1)}} \geq \sum_{k=0}^{n} \frac{1}{n+1} = 1$$

The coefficients $c_n = \sum_{k=0}^{n} a_k a_{n-k}$ therefore do not form a zero sequence. EOE.

### 6.2.8 Final interlude on the exponential function

**Example**

Let $E(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$ be the exponential series. We know it is absolutely convergent since $E(|x|)$ converges for all $x \in \mathbb{R}$.

Then $E(x)E(y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!} \frac{1}{(n-k)!} x^k y^{n-k}$. Now observe that $\sum_{k=0}^{n} \frac{1}{k!} \frac{1}{(n-k)!} x^k y^{n-k} = \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k y^{n-k} = \frac{1}{n!}(x+y)^n$.

It follows that for all $x, y \in \mathbb{R}$, we have $E(x)E(y) = E(x+y)$. EOE.

We are more than half-way there to show that $E(x) = \exp(x)$ is actually *the* exponential function, which will close the gap that we haven't yet established that $\exp(x)$ actually exists.

The only thing that is missing is the fact that $E(x)$ defines a differentiable function (it is clearly not zero, since $E(0) = 1$).

We will show this below in greater generality (that every so called *power series* is differentiable in the interior of its domain). But we can deal with $E(x)$ ad hoc.

Note that the example shows that $E$ defines a group homomorphism $\mathbb{R} \to \mathbb{R}_{>0}$. Indeed, $E(x) > 0$ for all $x$. This follows from $E(x) > 0$ for $x \geq 0$ (simply because $E(x) \geq 1$ for these $x$). And for $x < 0$, $E(x) = E(-x)^{-1} > 0$ because $1 = E(0) = E(x-x) = E(x)E(-x)$.

Groups "look" everywhere the same, as we can translate properties around an element $g$ to properties around an element $h$ by multiplying everything by $hg^{-1}$. This is not a precise statement, but it is one of the reasons we are interested in groups.

To illustrate the point: suppose we know that $E(x)$ is differentiable at some point $x_0$. Then to verify that it is differentiable at $y_0$ we can observe that $E(y_0 + h) - E(y_0) = E\big((y_0 - x_0) + x_0 + h\big) - E\big((y_0 - x_0) + x_0\big) = E(y_0 - x_0)(E(x_0 + h) - E(x_0))$.

**Lemma**

$E(x)$ is differentiable at $x_0 = 0$, and $E'(0) = 1$.

Proof. $E(h) - E(0) = E(h) - 1 = \sum_{n=1}^{\infty} \frac{1}{n!} h^n$. Then for $h \neq 0$,

$$\frac{E(h) - 1}{h} = \sum_{n=1}^{\infty} \frac{1}{n!} h^{n-1} = F(h)$$

Note the right hand side is a convergent series for $h \neq 0$, because $hF(h)$ is.

We must show that $\lim_{h \to 0} F(h) = 1$. This may seem obvious but it isn't. It says that $F$ as a function of $h$ is continuous at 0. For $h \neq 0$, $F(h) = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} h^n$ and $F(h) - 1 = \sum_{n=1}^{\infty} \frac{1}{(n+1)!} h^n$.

Then $|F(h)-1| \leq \sum_{n=1}^{\infty} \frac{1}{(n+1)!}|h|^n = |h| \sum_{n=0}^{\infty} \frac{1}{(n+1)!}|h|^n \leq |h|E(|h|)$.

Note that if $|h| \leq 1$, then $E(|h|) \leq E(1)$. Thus, for small $h$, $|F(h)-1| \leq |h|E(1) \to 0$ for $h \to 0$. QED.

**Corollary**
The function $x \mapsto E(x)$ defined on $\mathbb{R}$ is differentiable and $E'(x) = E(x)$ for all $x$.

Proof. $\frac{E(x_0+h)-E(x_0)}{h} = \frac{E(x_0)(E(h)-1)}{h} \to E(x_0)E'(0) = E(x_0)$ for $h \to 0$. QED.

We have now closed a gap in our earlier treatment of exponential functions. We have shown that there is a differentiable function $E \neq 0$ such that $E(x+y) = E(x)E(y)$. We have seen that as soon as we have one non-constant exponential function, we have a unique one for every base $a > 0$.

Note that as $E'(x) = E(x)$ it follows that in out earlier notation $\exp(x) = E(x)$.

We also find a new way of computing $e = E(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$. This converges much **much** faster than $\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{n}\right)^k$.

For example, $\left(1+\frac{1}{20}\right)^{20} \sim 2.6533$. But $\sum_{n=0}^{20} \frac{1}{n!}$ is indistinguishable from $e \sim 2.718282$ for quite a few digits: indeed, note that $(n+k)! \geq k!\,n!$ (binomial coefficients are natural numbers) Therefore $e - \sum_{n=0}^{20} \frac{1}{n!} = \sum_{n=21}^{\infty} \frac{1}{n!} = \sum_{n=0}^{\infty} \frac{1}{(n+21)!} \leq \frac{1}{21!} \sum_{n=0}^{\infty} \frac{1}{n!} \leq \frac{1}{21!} e$. Of course, now we would need to know that $e$ is small compared to 21!. It is: $e < 2.72$. One can show this using the exponential series:

Let $S_N = \sum_{n=0}^{N} \frac{1}{n!}$. Then $\lim_{N\to\infty} S_N = E(1) = e$.

One can show that $S_N \leq e \leq \frac{(N+1)!}{(N+1)!-1} S_N$. Sor for many practical purposes $e = S_N$ if $N > 5$.

### 6.2.9 Double sequences

**Definition**
A **double sequence** is a function $a: \mathbb{N} \times \mathbb{N} \to \mathbb{R}$, so that for each $m, n \in \mathbb{N}$, we have a real number $a_{m,n} = a(m,n)$. Sometimes the indices are allowed to be 0 or bounded negative.
A double sequence $a_{m,n}$ is **bounded** if there is $B > 0$ such that $|a_{m,n}| < B$ for all $m, n$. EOD.

As in the case of regular sequences, one can add double sequences and multiply them by constants, both in the obvious ways (so they form a vector space).

There are obvious extensions of that definitions to subsets of $\mathbb{Z} \times \mathbb{Z}$.

**Definition (Limits of double sequences)**
Let $a_{m,n}$ be a double sequence. A real number $L$ is the **limit** of such a sequence, and the sequence is then called **convergent** to $L$, if for every $\varepsilon > 0$ there is $n_0 \in \mathbb{N}$ such that for all $m, n > n_0$, $|a_{m,n} - L| < \varepsilon$. In this case we write $\lim_{m,n\to\infty} a_{m,n} = L$. There is an analogous definition for improper limits $\pm\infty$:
We say $\lim_{m,n\to\infty} a_{m,n} = \infty$ if for every $M$ there is $n_0$ such that $a_{m,n} > M$ for all $m, n > n_0$.
Similarly, $\lim_{m,n\to\infty} a_{m,n} = -\infty$ if for every $M$ there is $n_0$ such that $a_{m,n} < M$ for all $m, n > n_0$. EOD.

This is a subtle notion. Note that if $a_{m,n}$ converges to, say, $L$, no statement can be made about the sequences $b_n = a_{n,k}$ or $c_n = a_{k,n}$ (for fixed $k$). These sequences may not converge, or converge to different numbers, or to $\pm\infty$.

It is useful to think of $a_{m,n}$ as arranged into an "infinite matrix" whose $i$th row is formed by the sequence $a_{i,n}$ for $n \in \mathbb{N}$. This is just a mental image and doesn't really have a theoretical meaning.

We can then discuss *column* and *row* limits.

**Definition**
Let $a_{n,m}$ be a double sequence. Its $i$th **row limit** is defined as $\lim\limits_{m\to\infty} a_{i,m}$ if it exists.
Similarly, its $t$th **column limit** is defined as $\lim\limits_{n\to\infty} a_{n,i}$, if it exists. EOD.

**Example**

1. A sequence $a_n$ is a Cauchy sequence if and only if the double sequence $b_{m,n} = a_m - a_n$ converges to $0$.
2. Let $a_{m,n} = \frac{(-1)^n}{m}$. Then $\lim\limits_{m\to\infty} a_{m,n} = 0$. For fixed $n$, however, $\lim\limits_{n\to\infty} a_{m,n}$ does not exist.
3. Let $a_{m,n} = \frac{1}{\min\{n,m\}}$. Then $a_{m,n} \to 0$. But for fixed $m, n$ (one at a time) we have $\lim\limits_{n\to\infty} a_{m,n} = \frac{1}{m}$ and $\lim\limits_{m\to\infty} a_{m,n} = \frac{1}{n}$.
4. Let $a_{m,n} = \frac{m}{n}$. Then the row limits are all $0$ and the column limits are all $\infty$. The double sequence has no limit.
5. Let $a_{m,n} = \min\{m, n\}$. Then every row and column converges, but $a_{m,n}$ does not.
6. Let $a_{m,n} = \frac{m}{\max\{m,n\}}$. Then all row limits are $0$ and all column limits are $1$, and $a_{m,n}$ is not convergent as for any $n_0$ there are $m, n > n_0$ such that $a_{m,n} = 1$ and there are $m, n > n_0$ such that $a_{m,n} = 2$.
7. Let
$$a_{m,n} = \vdash\begin{cases} 1 & n \neq m \\ 0 & n = m \end{cases}$$
Then all row and column limits are $1$, but $a_{m,n}$ does not converge.

In the following a statement about natural numbers is **almost always** true, if the set of natural numbers, where the statement is false, is finite.

As an example, $\frac{n}{n-1}$ is defined for almost all integers. Likewise, for a convergent sequence $a_n$ with limit $L$ and any $\varepsilon > 0$, $|a_n - L| < \varepsilon$ is true for almost all $n$.

**Theorem**
Let $a_{m,n}$ be a double sequence with limit $L \in \mathbb{R} \cup \{\pm\infty\}$. Suppose row limits exist and are finite[3] for almost all $i$. Let $L_i$ be the $i$th row limit. Then $\lim\limits_{i\to\infty} L_i = L$.
As similar statement holds for column limits. EOT.

[3] The restriction that row limits are finite is not necessary. But we want the sequence of row limits to be a proper sequence.

Proof. First, let $L$ be finite. Let $n_0 \in \mathbb{N}$ be such that $L_i$ exists for all $i > n_0$. Let $\varepsilon > 0$. There is $N_0$ such that for all $n, m > N_0$ we have $|a_{m,n} - L| < \frac{\varepsilon}{2}$. Therefore, if $i > \max(N_0, n_0)$, we have $|L_i - L| \leq \frac{\varepsilon}{2} < \varepsilon$.

The same reasoning works for column limits.

If $L = \infty$, let $M > 1$. Then there exists $N_0$ such that $a_{m,n} > 2M$ for all $m, n > N_0$. Again, if $i > \max(N_0, n_0)$, then $a_{i,n} > 2M$ for all large enough $n$ and hence $L_i \geq 2M$ >M. So $L_i \to \infty$.

The case $L = -\infty$ is similar. QED.

It is crucial that it is a priori known that the double sequence *does have* a limit.

**Definition**
A double sequence $a_{m,n}$ is called **monotone increasing**, for all $m_0, n_0$ we have $a_{m_0,n_0} \leq a_{m,n}$ whenever $m > m_0$ *and* $n > n_0$. $a_{m,n}$ is **monotone decreasing**, if the double sequence $-a_{m,n}$ is monotone increasing. EOD.

The definitions of **strictly** monotone increasing/decreasing should be clear now.

**Exercise**
Prove the usual statements about limits for limits for double sequences (for sums and products of limits of convergent double sequences). EOE.

**Lemma**
A monotone and bounded double sequence converges. EOL.

Proof. We prove the statement for a bounded monotone increasing double sequence. The case of a monotone decreasing double sequence is similar (and also follows by applying the monotone increasing case to $-a_{m,n}$).

Let $L = \sup\{a_{m,n} \mid m, n \in \mathbb{N}\}$. Then $L \in \mathbb{R}$ because $a_{m,n}$ is bounded. I claim that $\lim_{m,n\to\infty} a_{m,n} = L$.

Indeed, let $\varepsilon > 0$. There is $(m_0, n_0) \in \mathbb{N} \times \mathbb{N}$ such hat $|a_{m_0,n_0} - L| < \varepsilon$. As the double sequence is monotone increasing this means that for all $m, n > \max\{m_0, n_0\}$, we have $|a_{m,n} - L| < \varepsilon$. QED.

**Remark/Exercise**
Note our definition of monotone is not uniformly accepted in the literature. You will also find the (stronger) definition that e.g. $a_{m,n}$ is monotone increasing if $a_{m',n'} \geq a_{m,n}$ whenever $m' \geq m$ and $n' \geq n$.

1. Show that this definition is equivalent to saying that all row and column sequences are monotone increasing. (Here a row sequence is a sequence of the form $n \mapsto a_{m,n}$ for fixed ("row") $m$, and a column sequence is a sequence of the form $m \mapsto a_{m,n}$ for fixed ("column") $n$.
2. Show that this is indeed stronger. That is show that it implies the notion of monotone defined above, but that the converse is not true, that is, there are double sequences, which are monotone by our definition, where the rows and columns are not necessarily monotone.

Because of 1. this stronger definition may be preferrable (it is certainly aesthetically more pleasing). However, since we are mainly interested in monotone and bounded implies convergent, we opt for the weaker definition which consequently applies in more cases. EOR.

### 6.2.10 Double series

As in the case of normal sequences we can define series in the case of double sequences.

**Definition**
Given a sequence $a_{n,m}$, we can form a **double series** as $\sum_{m,n=1}^{\infty} a_{m,n}$, with **partial sums** $S_{M,N} = \sum_{m=1}^{M} \sum_{n=1}^{N} a_{m,n}$. The **limit** or **value** of the double series $\sum_{m,n=1}^{\infty} a_{m,n}$ is $\lim\limits_{M,N\to\infty} S_{M,N}$ if it exists. We call it **convergent** if this limit exists and is finite. A double series is also allowed to start at 0 or any integer.
A double series $\sum_{m,n=1}^{\infty} a_{m,n}$ **convergence absolutely** if the double series $\sum_{m,n=1}^{\infty} |a_{m,n}|$ converges.
The series $\sum_{n=1}^{\infty} a_{m,n}$ is called the $m$th **row series** of the double series. The series $\sum_{m=1}^{\infty} a_{m,n}$ is called the $n$th **column series**. If a row or column series converges its limit is called a **row sum** or **column sum**, respectively.
EOD.

If a double series $\sum_{m,n=1}^{\infty} a_{m,n}$ converges absolutely, then it also converges: to see this, consider

$$b_{m,n} = \vdash \begin{cases} a_{m,n} & a_{m,n} \geq 0 \\ 0 & a_{m,n} \leq 0 \end{cases} \qquad\qquad c_{m,n} = \vdash \begin{cases} -a_{m,n} & a_{m,n} < 0 \\ 0 & a_{m,n} \geq 0 \end{cases}$$

Then $a_{m,n} = b_{m,n} - c_{m,n}$.

The double series $\sum_{m,n=1}^{\infty} b_{m,n}$ and $\sum_{m,n}^{\infty} c_{m,n}$ converge as they have nonnegative terms and are bounded by $\sum_{m,n=1}^{\infty} |a_{m,n}|$. But then $\sum_{m,n=1}^{\infty} a_{m,n} = \sum_{m,n=1}^{\infty} b_{m,n} - \sum_{m,n=1}^{\infty} c_{m,n}$ converges.

**Proposition**
Let $\sum_{m,n=1}^{\infty} a_{m,n}$ be a double series. If it converges, and *all* its row series converge, then the iterated series $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n}$ converges and for the limits we have

$$\sum_{m,n=1}^{\infty} a_{m,n} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n}$$

A similar statement is true if all column series converge and then

$$\sum_{m,n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n}$$

EOP.

Proof. This is an application of the theorem in 6.2.9 to the partial sums of the double series. QED.

We say an *iterated series* $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n}$ converges absolutely, if

$$\lim_{M\to\infty} \lim_{N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} |a_{m,n}| = \lim_{M\to\infty} \sum_{m=1}^{M} \left( \sum_{n=1}^{\infty} |a_{m,n}| \right)$$

is finite (of course the right hand side here makes sense only if $\sum_{n=1}^{\infty} |a_{m,n}|$ converges for all $m$). This is logically imprecise as $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n}$ can also be viewed as a simple series, and then absolute convergence would mean that $\sum_{m=1}^{\infty} \left|\sum_{n=1}^{\infty} a_{m,n}\right|$ converges. Which is not the same, in general.

**Cauchy's Double Series Theorem**

Let $\sum_{m,n=1}^{\infty} a_{m,n}$ be a double series. Suppose $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|a_{m,n}| = \lim_{M\to\infty}\lim_{N\to\infty}\sum_{m=1}^{M}\sum_{n=1}^{N}|a_{m,n}|$ is finite. Then the series $\sum_{m,n}^{\infty} a_{m.n}$ converges absolutely, every column series converges absolutely, and the iterated series of column series converges absolutely, and we have

$$\sum_{m,n=1}^{\infty} a_{m,n} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}$$

A similar statement (with the roles of rows and columns interchanged) holds if the iterated series of column series converge absolutely. EOT.

Proof. Under the assumptions of the theorem, the partial sums $A_{M,N} = \sum_{m=1}^{M}\sum_{n=1}^{N}|a_{m,n}| = \sum_{n=1}^{N}\sum_{m=1}^{M}|a_{m,n}|$ are bounded by $A = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|a_{m,n}|$. But then $\sum_{m,n=1}^{\infty} a_{m,n}$ is absolutely convergent and therefore convergent. Likewise, all row and column sums are absolutely convergent, as for instance $\sum_{m=1}^{\infty}|a_{m,n}| \le A$ since $\sum_{m=1}^{M}|a_{m,n}| \le A_{M,n} \le A$. The theorem in the last section then shows that

$$\sum_{m,n=1}^{\infty} a_{m,n} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}$$

QED.

We will see examples where double series are relevant in short order.

**Example**

Consider the double series $\sum_{m,n=2}^{\infty}\frac{1}{m^n}$.

Note that $\sum_{n=2}^{\infty}\frac{1}{m^n}$ is almost the geometric series $\sum_{n=0}^{\infty}\frac{1}{m^n}$, which, since $\frac{1}{m} < 1$ has limit $\frac{1}{1-\frac{1}{m}} = \frac{m}{m-1}$. We conclude that $\sum_{n=2}^{\infty}\frac{1}{m^n} = \frac{m}{m-1} - 1 - \frac{1}{m} = \frac{1}{m-1} - \frac{1}{m}$.

It follows that the iterated series $\sum_{m=2}^{\infty}\sum_{n=2}^{\infty}\frac{1}{m^n}$ is absolutely convergent with limit $\lim_{M\to\infty}\sum_{m=2}^{M}\left(\frac{1}{m-1} - \frac{1}{m}\right) = \lim_{M\to\infty} 1 - \frac{1}{M} = 1$. It follows that

$$1 = \sum_{m=2}^{\infty}\sum_{n=2}^{\infty}\frac{1}{m^n} = \sum_{n=2}^{\infty}\sum_{m=2}^{\infty}\frac{1}{m^n} = \sum_{m,n=2}^{\infty}\frac{1}{m^n}$$

EOE.

## 6.3 Power series

**Definition**

A **(formal) power series centered at** $c \in \mathbb{R}$ is a sequence $a_n$ $(n \in \mathbb{N}_0)$, written as $\sum_{n=0}^{\infty} a_n(x-c)$. It **converges** at $x_0 \in \mathbb{R}$ if $\sum_{n=0}^{\infty} a_n(x_0-c)^n$ converges and **diverges** otherwise.

We write $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ for this power series to remind us that we think of something that can be "evaluated" at various $x_0 \in \mathbb{R}$. We often write $f(x_0)$ converges/diverges, and write $f(x_0)$ for the **value** (read *limit*) of the power series at $x_0$. EOD.

Note that it is possible that a formal power series doesn't converge anywhere except $c$.

If $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ is a formal power series centered at $c$, and $d \in \mathbb{R}$ we write $f(x+d)$ for the formal power series $\sum_{n=0}^{\infty} a_n\big(x-(c-d)\big)$ centered at $c-d$. In particular, if $d=c$, we obtain a formal power series centred at 0. We will focus our attention mostly on these.

### 6.3.1 Power series as functions

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a (formal) power series (centred at 0). If $I \subseteq \mathbb{R}$ is a set such that for all $x_0 \in I$, $f(x_0)$ converges, we can define a function $F: I \to \mathbb{R}$ defined by $F(z) = f(z)$, where the right hand side means the limit of the series $\sum_{n=0}^{\infty} a_n z^n$.

By abuse of language we often will identify a formal power series with this function it defines. We will then use the same label ($f$) for both the power series and the induced function. While this is imprecise, it rarely causes confusion.

### 6.3.2 The radius of convergence of a power series

**Theorem**

Let $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ be a power series centered at $c$. Let $L := \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Then $f(x_0)$ converges for all $x_0$ with $|x_0 - c| < \frac{1}{L}$, and diverges for all $x_0$ with $|x_0 - c| > \frac{1}{L}$. EOT.

For the purpose of this theorem $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$.

Proof. We may assume $c = 0$ (why?). Let first $L < \infty$. We will show that if $|x_0| < \frac{1}{L}$, then $\sum_{n=0}^{\infty} a_n x_0^n$ converges absolutely. We may thus assume that $x_0 > 0$. Then $|a_n x_0^n| < \frac{|a_n|}{L^n}$, and therefore $\sqrt[n]{|a_n| x_0^n} < \sqrt[n]{\frac{|a_n|}{L^n}} = \frac{\sqrt{|a_n|}}{L}$ and so $\limsup_{n\to\infty} \sqrt[n]{|a_n x_0^n|} < 1$.

By the root test in Convergence of series6.2.2, this shows that the series $f(x_0)$ converges absolutely. Similarly, the inequality is reversed if $|x_0| > \frac{1}{L}$ showing that in this case $f(|x_0|)$ diverges. But note that the argument given in proving divergence in the root test was that if for the series $\sum_{n=1}^{\infty} b_n$ with $b_n \geq 0$, we have $\limsup_{n\to\infty} \sqrt{b_n} > 1$, then $b_n$ is not a zero sequence. Thus $|a_n x_0^n|$ is not a zero sequence, and hence $a_n x_0^n$ is not a zero sequence either. Therefore also $f(x_0)$ diverges.

If $L = \infty$, the same argument again shows divergence for all $x_0 \neq 0$. QED.

**Definition**

For a formal power series $f$, its **radius of convergence** is defined as $\frac{1}{L}$. It is defined to be $\infty$ if $L = 0$, and 0 if $L = \infty$. EOD.

**Remarks**

As an immediate consequence of the theorem, we obtain the following useful remarks:
Let $f(x)$ be a formal power series centered at $c$

- Let $x_0 \in \mathbb{R}$. Suppose $f(x_0)$ converges or that $f(z)$ converges for all $z$ with $|z-c| < |x_0 - c|$. Then the radius of convergence of $f(x)$ is at least $|x_0 - c|$.
- Let $a < b$, and suppose $f(x_0)$ converges for all $x_0 \in (a,b)$. Then the radius of convergence if $f(x)$ is at least $\max\{|a-c|, |b-c|\}$. Note this holds even if $c \notin (a,b)$.

EOR.

**Example**

1. The radius of convergence of $f(x) = \exp(x)$ is $\infty$.
2. The radius of convergence of the geometric series $\sum_{n=0}^{\infty} x^n$ is 1.
3. The radius of convergence of the series $\sum_{n=0}^{\infty} n!\, x^n$ is 0.
4. The radius of convergence of $g(x) = \sum_{n=1}^{\infty} \frac{1}{n} x^n$ is 1 and $g(1)$ diverges, whereas $g(-1)$ converges.
5. The radius of convergence of $h(x) = \sum_{n=0}^{\infty} nx^n$ is 1 and $h(-1)$ and $h(1)$ both diverge ($|(-1)^n n|$ is not a zero sequence).

If a power series $f(x)$ centred at $c$ has radius of convergence $R > 0$ it is in general not clear what happens if $|x_0 - c| = R$. It may converge, diverge, or converge but not absolutely.

**Convention**
We henceforth identify a power series with the function it defines on the interval defined by its radius of convergence. EOC.

The computation of the limit superior is often tedious. In some practical cases, the quotient criterion is more useful.

**Lemma**
Let $f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n$ be a formal power series such that $a_n \neq 0$ for almost all $n$. Suppose $L = \lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists.
Then the radius of convergence of $f$ is equal to $\frac{1}{L}$ (again with $\frac{1}{0} := \infty$ and $\frac{1}{\infty} := 0$). EOL.

Proof. By the quotient criterion, $f(x_0)$ converges (absolutely) if $\limsup \left|\frac{a_{n+1}(x_0-c)^{n+1}}{a_n(x_0-c)^n}\right| = L|x_0 - c| < 1$. Similarly, $f(x_0)$ diverges, if $\liminf \left|\frac{a_{n+1}(x_0-c)^{n+1}}{a_n(x_0-c)^n}\right| = L|x_0 - c| > 1$. (Here we use that the lim inf and lim sup agree with the limit, if the limit exists. QED.

**Fact**
Let $a_n \geq 0$ be a sequence. For any $k \in \mathbb{N}$, we have

$$\limsup_{n\to\infty} \sqrt[n]{a_{n\pm k}} = \limsup_{n\to\infty} \sqrt[n]{a_n}$$

EOF.

Note that $a_{n-k}$ defines a sequence starting at $n = k+1$.

Proof. First note that if $b_n$ is any sequence then $\limsup b_n = \limsup b_{n+k}$ (exercise).

It therefore suffices to show that with $c_n = \sqrt[n+k]{a_{n+k}}$, the sequence $d_n = c_n^{\frac{n+k}{k}}$ has the same limit superior as $c_n$. Let $S = \limsup c_n$. Let $c_{n_\ell}$ be a subsequence with $\lim\limits_{\ell\to\infty} c_{n_\ell} = S$. Then also $\lim c_{n_\ell}^{\frac{n_\ell+k}{k}} = S$.

Indeed, if $S > 0$, $c_{n_\ell}^{1+\frac{n_\ell}{k}} = \exp\left(\frac{1+n_\ell}{k} ln\, c_{n_\ell}\right) \to \exp(\ln S) = S$ (with the definition $\exp(\ln \infty) := \infty$. If $S = 0$, then for almost all $\ell$, $c_{n_\ell} < 1$, and hence almost always $d_{n_\ell} < c_{n_\ell} \to 0$ by the Squeeze Principle.

This shows that the sequence $d_n$ has a subsequence with limit $S$. If $S = \infty$, we are done. Otherwise, if $T > S$ there is no subsequence of $d_n$ with limit $T$. Indeed, if there were, $d_{n_m}$, say, by similar arguments as above, the sequence $c_{n_m} = d_{n_m}^{\frac{k}{k+n_m}}$ would also have limit $T$, a contradiction.

Note that we used the fact that the limit superior is the largest possible limit of a subsequence (exercise). QED.

**Proposition (Radius of convergence for shifted series)**
Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a formal power series (centred at 0) with radius of convergence $R$.
Let $k \in \mathbb{N}$. Then the *shifted series* $g(x) = \sum_{n=0}^{\infty} a_{n+k} x^n$ has the same radius of convergence.
It then follows that also the series $h(x) = \sum_{n=k}^{\infty} a_{n-k} x^n$ has the same radius of convergence. EOP.

Proof. Let $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$ . By the above fact $\limsup_{n\to\infty} \sqrt[n]{|a_{n+k}|} = L$. Thus, the shifted series has the same radius of convergence.

If for $n \geq k$ we put $b_n = a_{n-k}$, this shows that $L = \limsup_{n\to\infty} \sqrt[n]{|b_{n+k}|} = \limsup_{n\to\infty} \sqrt[n]{|b_n|}$.

Then $h(x)$ has the same radius of convergence as $f(x)$.

Note, one could also argue directly and show that $f(x_0)$ converges if and only if the shifted series converges at $x_0$. QED.

**Corollary 1**
Let $f(x)$ be a formal power series centred at $c$ with radius of convergence $R > 0$. Then $f$ is continuous at $x_0 = c$. EOC.

Proof. Let $f(x) = \sum_{n=0}^{\infty} a_n (x - c)$. We must show that $\lim_{x\to c} f(x) = a_0 = f(c)$. To do this note that for every $x_0$, we have

$$|f(x_0) - a_0| = \left|\sum_{n=1}^{\infty} a_n (x_0 - c)^n\right| = |x_0 - c| \left|\sum_{n=0}^{\infty} a_{n+1} (x_0 - c)^n\right|$$

The right hand side involves the shifted series $\sum_{n=0}^{\infty} a_{n+1} (x - c)^n$. This series has the same radius of convergence as $f(x)$.

As $f(x_0)$ is absolutely convergent for $|x_0 - c| < R$, the right hand side is bounded by

$$|x_0 - c| \sum_{n=0}^{\infty} |a_{n+1}||z - c|^n$$

where we choose $z \in (c, c + R)$ such that $|x_0 - c| < |z - c|$ for all $x_0$ close to $c$ (so typically $z$ closer to $c + R$ than any of the $x_0$ close to $c$ in question). Thus $\lim_{x_0\to c} |f(x_0) - a_0| = 0$. QED.

**Corollary 2**
Let $f(x)$ be a formal power series centred at $c$ with radius of convergence $R > 0$. Then $f$ is differentiable at $x_0 = c$, and $f'(x_0) = a_1$. EOC.

Proof. We must show that $\lim_{x_0 \to c} \frac{f(x_0)-f(c)}{x_0-c} = a_1$.But note that

$$f(x_0) - f(c) = (x_0 - c) \sum_{n=0}^{\infty} a_{n+1}(x_0 - c)^n$$

We can therefore identify $\frac{f(x)-f(c)}{x-c}$ with the shifted power series $\sum_{n=0}^{\infty} a_{n+1}(x-c)^n$. Corollary 1 tells us that this is a continuous function at $x_0 = c$, and its value at $c$ is $a_1$. QED.

### 6.3.3 Formal derivatives of power series

**Definition**

Let $f = \sum_{n=0}^{\infty} a_n x^n$ be a formal power series. We define $D(f) := \sum_{n=0}^{\infty}(n+1)a_{n+1}x^n$ and call this the **formal derivative** of $f$. EOD.

**Lemma**

If $f$ has radius of convergence $R$ then so does $D(f)$. EOL.

Proof. Let $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. We must show that $L = \limsup_{n\to\infty} \sqrt[n]{(n+1)|a_{n+1}|}$. As $\lim_{n\to\infty} \sqrt[n]{n+1} = 1$ this follows from our result for the radius of convergence for shifted series. QED.

An alternate argument is outlined after the Transformation Theorem below.

**Exercise**

1. For $k > 0$ show that if we define $D^k(f)$ as applying $D$ $k$ times to $f$, then
$$D^k(f) = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} a_{n+k} x^n$$
Show that the radius of convergence of $D^k(f)$ is the same as that of $f$.
2. For a formal power series $f$ as above we define the **formal antiderivative** as $I(f) = \sum_{n=1}^{\infty} \frac{1}{n} a_{n-1} x^n$. Show that $D\big(I(f)\big) = f$.
3. Show that the radius of convergence of $I(f)$ is that of $f$.

EOE.

**Remark**

It is tempting to conclude from the preceding results that if $f$ is a power series (centred at $0$, say) with radius of convergence $R > 0$, then $f$ is differentiable on $(-R, R)$ and $f' = D(f)$.
This is true, but to prove this is surprisingly involved. Informally speaking it involves to "swap" limits: That means, if $c \in (-R, R)$ to show that $f'(c) = D(f)(c)$, one needs to show that

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{x\to c} \sum_{n=0}^{\infty} a_n \left(\frac{x^n - c^n}{x-c}\right) = \sum_{n=0}^{\infty} a_n \lim_{x\to c}\left(\frac{x^n - c^n}{x-c}\right) = D(f)(c)$$

It is not entirely obvious why that should be true. EOR.

### 6.3.4 The Transformation Theorem

**Theorem (Transformation Theorem)**

Let $f$ be a power series centered at $c$ with convergence radius $R > 0$. For any $d \in (c - R, c + R)$, then (as a function) $f(x) = \sum_{k=0}^{\infty} \frac{D^k(f)(d)}{k!}(x - d)^k$ on $(d - S, d + S)$ where $S = \min\{R - (c - d), R - (d - c)\} = R - |d - c|$. EOT.

Proof. We first treat the case $c = 0$. Let $f = \sum_{n=0}^{\infty} a_n x^n$ and let $d \in (-R, R)$. We want to rewrite $f$ as a power series $g = \sum_{n=0}^{\infty} b_n (x - d)^n$ centered at $d$.

For any $x_0 \in (-R, R)$

$$f(x_0) = f_N\big((x_0 - d) + d\big) = \sum_{n=0}^{\infty} a_n \sum_{k=0}^{n} \binom{n}{k} (x_0 - d)^k d^{n-k}$$

We define $\binom{n}{k} = 0$ if $k > n$, then we can write the above as

$$f(x_0) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \binom{n}{k} (x_0 - d)^k d^{n-k}$$

Note that $f(x_0)$ is absolutely convergent. Now let $|x_0 - d| + |d| < R$.

Then $\sum_{n=0}^{\infty} |a_n|(|x_0 - d| + |d|)^n$ converges, and

$$\sum_{n=0}^{\infty} |a_n|(|x_0 - d| + |d|)^n = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \binom{n}{k} |x_0 - d|^k |d|^{n-k}$$

Thus, the iterated series for $f(x_0)$ above is absolutely convergent. By Cauchy's Double Series Theorem, we can swap the summation (and obtain absolute convergence of the swapped iterated series).

$$f(x_0) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \binom{n}{k} a_n (x_0 - d)^k d^{n-k} = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \binom{n}{k} a_n (x_0 - d)^k d^{n-k}$$
$$= \sum_{k=0}^{\infty} \left( \sum_{n=0}^{\infty} \binom{n}{k} a_n d^{n-k} \right) (x_0 - d)^k$$

Now observe that $\sum_{n=0}^{\infty} \binom{n}{k} a_n d^{n-k} = \sum_{n=k}^{\infty} \binom{n}{k} a_n d^{n-k} = \frac{1}{k!} \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} a_{n+k} d^n = \frac{1}{k!} D^k(f)(d)$.

Note that $|x_0 - d| + |d| < R$ if and only if $|x_0 - d| < R - |d| = \min\{R - d, R + d\}$.

The case of $c \neq 0$ follows by observing that $f(x + c)$ is a power series centered at $0$ and $D^k(f)(x + c) = D^k(f(x + c))$. QED.

**Remark**
Note that the Transformation Theorem implies that $D(f)$ has a radius of convergence at least that of $f$. By the exercise in the last section, $I(D(f))$ has a radius of convergence at least that of $D(f)$. But $I\big(D(f)\big) = f - a_0$ has radius of convergence equal to $f$. Also $D\big(I(f)\big) = f$. Therefore $f, D(f)$, and $I(f)$ all have the same radius of convergence. EOR.

**Corollary 1**
Let $f = \sum_{n=0}^{\infty} a_n (x - c)^n$ be a formal power series centered at $c$ with radius of convergence $R > 0$.

Then for any $x_0 \in (c-R, c+R)$, $f$ is differentiable at $x_0$ and $f'(x_0) = D(f)(x_0) = a_1$. EOC.

Proof. By the Transformation Theorem $f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} D^n(f)(x_0)(x-x_0)^n$ on a small enough interval around $x_0 \in I$. By Corollary 2 in 6.3.2, this means $f$ is differentiable at $x_0$ with $f'(x_0) = D(f)(x_0)$. In particular, this says that the derivative of $f$ is $D(f)$. QED.

**Corollary 2**
Let $f = \sum_{n=0}^{\infty} a_n(x-c)^n$ be a formal power series centered at $c$ with radius of convergence $R > 0$. Then $f$ is smooth on $(c-R, c+R)$, and $f^{(n)}$ is again a power series, namely $D^n(f)$. In particular, $a_n = \frac{1}{n!} D^n(f)(c)$. EOC.

Proof. By Corollary 1, $f$ is differentiable at any point in $(c-R, c+R)$, and its derivative is again a power series, namely $D(f)$, and $a_1 = D(f)(c)$. We now proceed by induction on $n$. If $n = 1$, the assertion is proven. Suppose Corollary 2 holds for a specific natural number $n$. Then $f^{(n)}$ is a power series, defined by $D^n(f)$. Then applying Corollary 1 to $D^n(f)$, we find that $D^n(f)$ is differentiable with derivative $D\big(D^n(f)\big) = D^{n+1}(f)$. This shows that $f^{(n+1)}$ is again a power series, and in fact equal to $D^{n+1}(f)$.

Applying the Transformation Theorem to $x_0 = c$, we get $f = \sum_{n=0}^{\infty} \frac{1}{n!} D^n(f)(c)(x-c)^n$ with the same radius of convergence. We would like to conclude that this means $a_n = \frac{D^n(f)(c)}{n!}$ for all $n$. Bu the Transformation Theorem does not guarantee that. But we know that $a_1 = D(f)(c)$ by Corollary 1. Note that from the definition $D^k(f) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n (x-c)^{n-k}$. Therefore $D^k(f)(c) = k!\, a_k$. QED.

**Convention**
We henceforth identify a power series with the function it defines on the interval defined by its radius of convergence. For example, if $f$ is a power series, we write $f'$ both for $D(f)$ and the derivative of the function defined by $f$. EOC.

**Corollary 3**
Let $f, g$ be two power series centered at $c$, both convergent on the same nonempty interval $I$ containing $c$. Then $f = g$ if and only if $f(x_0) = g(x_0)$ for all $x_0 \in I$. EOC.

**Remark**
This may sound like a tautology, and indeed we are just picking up some loose ends that we have avoided so far, namely the question whether the function defined by a power series determines the power series (that is, its coefficients). To be precise, the issue is, if $f = \sum_{n=0}^{\infty} a_n(x-c)^n$ and $g = \sum_{n=0}^{\infty} b_n(x-c)^n$, then $f = g$ "should" mean $a_n = b_n$ for all $n$, whereas $f(x_0) = g(x_0)$ for all $x_0 \in I$ only means that the functions defined by the two power series are the same. EOR.

Proof. Let $f = \sum_{n=0}^{\infty} a_n(x-c)^n$ and $g = \sum_{n=0}^{\infty} b_n(x-c)^n$. If $f = g$ (as series) then $a_n = b_n$ for all $n$ and therefore $f(x_0) = g(x_0)$ for all $x_0 \in I$.

We must therefore show the converse: if $f(x_0) = g(x_0)$ for all $x_0 \in I$, then $a_n = b_n$ for all $n$. $f, g$ both define the same function on $I$. Both functions are smooth by Corollary 2, and $a_n = \frac{1}{n!} f^{(n)}(c)$ and $b_n = \frac{1}{n!} g^{(n)}(c)$. It follows that $a_n = b_n$. QED.

### 6.3.5 Identity Theorem for power series*

We have seen that if two formal power series define the same function, then they are equal. However, a much stronger statement is true: if they are centred at the same point, and their functions agree at a countable set with that point an accumulation point, then they are already equal.

**Theorem**
Let $f(x), g(x)$ be power series centered at $c$, convergent on the interval $I = (c-R, c+R)$. Let $x_n \neq c \in I$ be any sequence such that $\lim_{n\to\infty} x_n = c$. If $f(x_n) = g(x_n)$ for all $n$, then $f(x) = g(x)$. EOT.

Proof. We may assume that $c = 0$. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $g(x) = \sum_{n=0}^{\infty} b_n x^n$. We will show that $a_n = b_n$ for all $n$.

To do this, for $k = 0,1,2, \ldots$ we define $f_k = \sum_{n=k}^{\infty} a_n x^{n-k} = \sum_{n=0}^{\infty} a_{n+k} x^n$.

We proceed by induction on $n$.

First, $a_0 = b_0$ because both $f$ and $g$ define continuous functions, so $a_0 = f(0) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} g(x_n) = g(0) = b_0$.

Suppose that for a given $n$, $a_m = b_m$ for all $m \leq n$.

Then $f_n - a_n = x\sum_{k=0}^{\infty} a_{k+n+1} x^n = x f_{n+1}$ agrees with $g_n - a_n = x\sum_{k=0}^{\infty} b_{k+n+1} x^n$ on $x_m$ for all $m$.

Thus, the power series $x f_{n+1} = x g_{n+1}$ on all $x_m$. In particular,

$$a_{n+1} = f_{n+1}(0) = \lim_{m\to\infty} f_{n+1}(x_m) = \lim_{m\to\infty} g_{n+1}(x_m) = g_{n+1}(0) = b_{n+1}$$

QED.

**Exercise**
Examine the proof of the previous theorem and see where we used that $x_n \neq c$ for all $n$. EOE.

### 6.3.6 More on power series

We want to discuss some additional features of power series.

A topic we have been touched upon only in passing is how to define/interpret sums and products of power series.

**Remark**
Let $f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n$ , $g(x) = \sum_{n=0}^{\infty} b_n (x-c)^n$ be formal power series centred at $c$.

1. The **sum** $f(x) + g(x)$ is the formal power series $f(x) + g(x) \coloneqq \sum_{n=0}^{\infty} (a_n + b_n)(x-c)^n$.
2. The **product** $f(x)g(x)$ is the formal power series $f(x)g(x) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k}\right)(x-c)^n$.
3. If $a \in \mathbb{R}$, then $af(x)$ is the formal power series $af(x) = \sum_{n=0}^{\infty} a a_n (x-c)^n$.

EOR.

It should be clear that if both $f, g$ are convergent on $(c-R, c+R)$, then so are $f(x) + g(x)$ and $f(x)g(x)$, and the functions given by $f(x) + g(x)$ and $f(x)g(x)$ coincide with the sum (resp. product) of the functions given by $f(x)$ and $g(x)$. Note this uses the theorem about products of absolutely convergent series (6.2.7).

**Proposition**
Let $f(x)$ be a power series centred at $c$, convergent on $(c-R, c+R)$ ($R>0$), and $g(x)$ be a power series centred at $d$ convergent on $(d-S, d+S)$. If $g(d)=c$, there is $\delta>0$ (and $\delta \leq S$) and a power series $h(x)$ centred at $d$, convergent on $I=(d-\delta, d+\delta)$, such that for all $x_0 \in I$, $g(x_0) \in (c-R, c+R)$, and $f\big(g(x_0)\big)=h(x_0)$. EOP.

Note the proposition just says that a composition of power series (when defined) is again a power series close to the centre of $g$.

Proof. This is a bit technical, but not technically difficult.

First, we may assume that $c=d=0$. Indeed, if the proposition is true in this setting, then the general case follows by observing that $f(x+c)$ and $g(x+d)$ are centred at 0 without changing the radius of convergence. Then there is $H$ such that $f\big((g(x_0+d)-c)+c\big)=H(x_0)$, then $f\big(g(x_0)\big)=H(x_0-d)=: h(x_0)$ where $h$ is a power series.

Let $f(x)=\sum_{n=0}^{\infty} a_n x^n$ and $g(x)=\sum_{n=0}^{\infty} b_n x^n$. Since $g(0)=0$, we have $b_0=0$.

Let $f_1=\sum_{n=0}^{\infty} |a_n| x^n$ and $g_1=\sum_{n=0}^{\infty} |b_n| x^n$. $f_1, g_1$ have the same radius of convergence as $f, g$ respectively. $g_1$ is continuous, so there is $\delta>0$ (which we may assume to be $\leq S$) such that $|g(x_0)|<R$ for all $x_0$ with $|x_0|<\delta$ (apply the $\varepsilon\delta$-criterion to $g_1$with $\varepsilon \coloneqq R$).

Then also $|g(x_0)| \leq |g_1(x_0)|<R$. Thus, both $f_1(g_1(x_0))$ and $f\big(g(x_0)\big)$ converge.

Note that $g(x_0)^n=\left(\sum_{k=0}^{\infty} b_k x_0^k\right)^n=\sum_{k=0}^{\infty} c_{kn} x^k$ for some double sequence $c_{kn}$ (only depending on $g(x)$ and not on $x_0$. Indeed, by induction on $n$, each of the powers $g(x)^n$ is a power series (convergent, wherever $g(x)$ converges absolutely); it is after all just an $n$-fold product of power series (see remarks at the beginning of this section).

Likewise, $g_1(x_0)^n=\sum_{k=0}^{\infty} C_{kn} x_0^k$. Note that $|c_{kn}| \leq C_{kn}$ (which is clearly nonnegative by the definition of $g_1(x)$). This would require a proof by induction on $n$. But it is elementary; for example, if $n=2$, $c_{k2}=\sum_{\ell=0}^{k} b_\ell b_{k-\ell}$, whereas $C_{k2}=\sum_{\ell=0}^{k} |b_\ell|\,|b_{k-\ell}|$. Now observe that $c_{k,n+1}=\sum_{\ell=0}^{k} b_\ell c_{k-\ell,n}$ and $C_{k,n+1}=\sum_{\ell=0}^{k} |b_\ell| C_{k-\ell,n}$. So induction really is your friend here.

Now, $f_1\big(g_1(x_0)\big)=\sum_{n=0}^{\infty} |a_n| g_1(x_0)^n=\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |a_n| C_{k,n} x_0^k$, and this iterated series is absolutely convergent (indeed, if $|g_1(x_0)|<R$, then $0 \leq g_1(|x_0|)<R$.

It follows that the iterated series $\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |a_n| C_{k,n} |x_0|^k$ is convergent. But this implies that the iterated series

$$f\big(g(x_0)\big)=\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} a_n c_{n,k} x_0^k$$

is absolutely convergent (each inner sum is absolutely convergent by what we said above, and in fact

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |a_n c_{n,k} x_0^k|$$

converges by the above (each inner sum is bounded, and the sum of inner sums is bounded by $f_1(g_1(|x_0|))$).

By Cauchy's Double Series Theorem, we may swap the order of summation, and

$$f(g(x_0)) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} a_n c_{n,k}\, x_0^k$$

But then $h(x) \coloneqq \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} a_n c_{n,k}\right) x^k$ does the trick. QED.

For us the most important consequence is the following corollary:

**Corollary**
Let $f(x)$ be a power series centered at $c$, convergent on $(c - R, c + R)$ $(R > 0)$, with $f(c) \neq 0$. Then there exists $0 < \delta \leq R$ such that $f(x_0) \neq 0$ for all $x_0 \in (c - \delta, c + \delta)$ and the (well-defined) function $\frac{1}{f}$ is a power series on $(c - \delta, c + \delta)$. EOC.

Proof. WLOG ("without loss of generality") we may assume that $f(c) = 1$. Indeed, if the corollary holds for such power series, and $f(x)$ is a general power series with $f(c) \neq 0$, we may apply the result to the power series $f(c)^{-1} f(x)$, to conclude that $\frac{1}{f(c)^{-1} f(x)}$ is a power series $h(x)$, say, and then conclude that $\frac{1}{f} = \frac{1}{f(c)} h$.

Let $f(x) = \sum_{n=0}^{\infty} a_n (x - c)^n$, with $f(c) = a_0 = 1$. Then $f(x) = 1 - g(x)$, where $g(x) = \sum_{n=1}^{\infty} (-a_n)(x - c)^n$. $g(x)$ has the same radius of convergence as $f(x)$.

Now consider the function $h\colon (-1,1) \to \mathbb{R}$, $h(x) = \frac{1}{1-x}$. We know that $h$ is a power series (namely, the geometric series $\sum_{n=0}^{\infty} x^n$). Since $g(c) = 0$, the proposition says that there is $\delta > 0$ and a power series $k(x)$, centered at $c$ with radius of convergence $\geq \delta$, such that for all $x_0 \in (c - \delta, c + \delta)$, we have $\frac{1}{f(x_0)} = \frac{1}{1 - g(x_0)} = \sum_{n=0}^{\infty} g(x_0)^n = k(x_0)$. QED.

**Corollary 2 (Quotients of power series are power series)**
Let $f(x), g(x)$ be power series centered at $c$, convergent on an open interval containing $c$, and suppose $g(c) \neq 0$. Then there is $\delta > 0$ such that the function $\frac{f}{g}$ is a power series on $(c - \delta, c + \delta)$. EOC.

Proof. There is $\delta > 0$ such that $\frac{1}{g}$ is defined as a function and defines a power series $\frac{1}{g(x)}$ on $(c - \delta, c + \delta)$. Then $\frac{f}{g}$ corresponds to the $f(x) \cdot \frac{1}{g(x)}$, which is a power series by the above. QED.

### 6.3.7 Taylor series

One of the most important examples of so called power series are Taylor series associated to smooth functions. Recall that a function $f$ defined on an interval $I$ is called **smooth** if $f^{(n)}$ exists on all of $I$ for all $n \in \mathbb{N}$.

**Definition**
Let $I$ be an interval, $c \in I^{\circ}$, and $f$ a function defined on $I$, such that $f^{(n)}(c)$ exists for all $n \in \mathbb{N}$. Then the **Taylor series** of $f$ at $c$ is the formal power series

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

EOD.

**Subtlety alert**

The existence of $f^{(n)}$ at $c$ for all $n$ implies that for each $n$ there is a $\delta_n > 0$ such that $f, f', f'', \dots, f^{(n)}$ are defined on $(c-\delta_n, c+\delta_n)$. But nothing prevents $\delta_n$ from being a zero sequence a priori. So $f^{(n)}$ to exist for all $n$ does not directly imply that there is an open interval containing $c$ where $f^{(n)}$ is defined for all $n \in \mathbb{N}$. EOS.

The convergence radius of $T_f$ may be 0. Even if $T_f$ converges at a point $x_0$, it may converge to a value other than $f(x_0)$. We identify the Taylor series with the function it induces on the interval centered at $c$ (of lenth twice the convergence radius).

**Example**

1. Since $\exp'(\mathrm{x}) = \exp(x)$, $T_{\exp,0}(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \exp(x)$.
2. More generally, if $f$ is any power series centered at $c$ with positive convergence radius, then $f$ coincides with $T_{f,c}$.
3. Consider $g(x) = \frac{1}{1+x}$. Then on $(-1,1)$ $g$ is a power series, namely $g(x) = \sum_{n=0}^{\infty}(-1)^n x^n$. Indeed, $g(-x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$.
4. Let $G(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$. Then $D(G) = g$ so $G$ and $g$ have the same radius of convergence. In particular, $G' = g = f'$, so $G = f + C$ for some $C \in \mathbb{R}$. Then $C = G(0) - f(0) = 0$. It follows that $\log(1+x) = G(x)$ on $(-1,1)$. In fact, using Lagrange remainders one can show that this equation also hlds for $x_0 = 1$.
5. For any real number $a \in \mathbb{R}$ and any $n \in \mathbb{N}_0$ we define the **generalized binomial coefficient** $\binom{a}{n}$ as
$$\binom{a}{n} = \frac{a(a-1)(a-2)\cdots(a-n+1)}{n!}$$
if $n > 0$ and 1, if $n = 0$. The **binomial series** for $a$ is defined as the formal power series $B_a(x) = \sum_{n=0}^{\infty}\binom{a}{n} x^n$. Its radius of convergence is 1.
Indeed, for large $n$ and $x \neq 0$, the ratio test gives $\frac{|\binom{a}{n+1} x^{n+1}|}{|\binom{a}{n} x^n|} = \frac{|x|}{n+1}|a-n|$.
For $n \to \infty$ this has limit $|x|$. Thus $B_a(x)$ converges if $|x| < 1$, and diverges if $|x| > 1$ (see the Ratio Test 6.2.5.).
For $a \in \mathbb{R}$, consider $f(x) = (1+x)^a$. Then $T_{f,0}(x) = B_a(x)$. Note that $f^{(n+1)}(x) = a(a-1)(a-2)\cdots(a-n)(1+x)^{a-n-1}$. Then the Lagrange remainder is $\frac{|x|^{n+1}}{(n+1)!} a(a-1)\cdots(a-n)(1+d)^{a-n-1} = |x|^{n+1}\binom{a}{n+1}(1+d)^{a-n-1} \to 0$ for $n \to \infty$ for as long as $d \in (0,1)$.
It follows that $T_{f,0}$ converges to $f(x)$ for all $x \in [0,1)$. What about $x < 0$?
Then $B_a(x)$ still converges, and $B_a'(x) = \sum_{n=0}^{\infty}(n+1)\binom{a}{n+1} x^n \sum_{n=0}^{\infty} a\binom{a-1}{n} x^n$.
Note $(1+x)B_a'(x) = aB_a(x)$ because $\binom{a-1}{n} + \binom{a-1}{n-1} = \binom{a}{n}$. $f$ also satisfies that $(1+x)f'(x) = af(x)$. Also note that $f(x) > 0$ on $(-1,1)$.

Then $\frac{d}{dx}\left(\frac{B_a(x)}{f(x)}\right) = \frac{B_a'(x)f(x) - B_a(x)f'(x)}{f^2(x)} = \frac{\frac{a}{1+x}(B_a(x)f(x) - B_a(x)f(x))}{f^2(x)} = 0.$
Therefore $B_a(x) = Cf(x)$ for some constant $C$. Now $1 = B_a(0) = f(0)$ forces $C = 1$.

6. Consider the function $h$ defined as

$$h(x) = \vdash \begin{cases} 0 & x \leq 0 \\ \exp\left(-\frac{1}{x}\right) & x > 0 \end{cases}$$

Then $h^{(n)}(0) = 0$ for all $n$. Thus $T_{h,0} = \sum_{n=0}^{\infty} 0 \cdot x^n$. But for positive $x$ $h(x) > 0$. So the Taylor series with convergence radius $\infty$ converges nowhere to $h(x)$ for $x > 0$.

EOE.

**Proposition**
If $I$ is an open interval and $f$ is analytic on $I$, then for each $c \in I$, the Taylor series $T_{f,c}$ has a positive convergence radius $R$ and $f(x) = T_{f,c}(x)$ for all $x \in (c - R, c + R) \cap I$. EOP.

Proof. This is just a restatement of the definition of $f$ being analytic on $I$ (see the Transformation Theorem and its corollaries). QED.

**Exercise**

Consider $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = e^{-\frac{1}{x^2}}$ if $x \neq 0$ and $f(0) = 0$. Show that $f$ is smooth.
Show that $T_{f,0} = 0$. EOE.

### 6.3.8 Analytic functions

**Definition**
Let $I$ be an **open** interval, and $f$ a function defined on $I$. We say that $f$ is **analytic at $c \in I$**, if there is a formal power series $g$ centred at $c$ convergent on an interval $(c - \delta, c + \delta) \subseteq I$ for some $\delta > 0$, such that $f(x) = g(x)$ on $(c - \delta, c + \delta)$. We say $f$ is **analytic** if that holds for every $c \in I$. EOD.

For example, the exponential function is analytic at 0. In fact, every power series with positive radius of convergence is analytic at its centre.

Note that by definition a function $f$ on an open interval $I$ is analytic if and only if for every $c \in I$, the Taylor series $T_{f,c}(x)$ converges to $f(x)$ for all $x$ "close" to $c$.

**Remark**
Analytic functions are differentiable. Indeed, for $c \in I$, we can find $\delta > 0$ such $f(x) = g(x)$ for some power series $g$ centered at $c$ and all $x \in (c - \delta, c + \delta)$. But then $f$ is differentiable at $c$ iff $g$ is. Since $g$ is differentiable at $c$, $f$ is. EOR

**Remark**
The Transformation Theorem could be rephrased as follows: given a power series $f(x)$ centred at $c$ and convergent on $I = (c - R, c + R)$ for some $R > 0$, then $f(x)$ is analytic on $I$.

**Example**
Recall the exponential series $\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$. If we take only the "even" or "odd" summands and alternate them, we obtain power series that are also convergent on all of $\mathbb{R}$.

1. Let $S(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)!} x^{2n-1}$. Then the radius of convergence is $\infty$, because $\sum_{n=1}^{\infty} \frac{1}{(2n-1)!} |x|^{2n-1} \leq E(|x|) < \infty$ for all $x \in \mathbb{R}$.
   Then $C(x) := S'(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(2n-1)}{(2n-1)!} x^{2n-2} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$ again has radius of convergence $\infty$.
2. $C'(x) = \sum_{n=1}^{\infty} \frac{(-1)^n 2n}{(2n)!} x^{2n-1} = -\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)!} x^{2n-1} = -S(x)$.

We find $S'(x) = C(x)$, and $C'(x) = -S(x)$. In particular, $S''(x) = -S(x)$ and $C''(x) = -C(x)$. This should remind you of the sin and cos functions. We will see, that in fact $\sin x = S(x)$ and $\cos x = C(x)$. All three functions ($\exp x$, $S(x)$, $C(x)$) are analytic on all of $\mathbb{R}$ by the Transformation Theorem. EOE.

**Exercise**
Show that $\ln x$ is analytic on $(0, \infty)$. EOE.

**Exercise**
Show that sums, products, and quotients (where defined) of analytic functions are analytic. Show that the composition of analytic functions (where defined) is analytic. EOE.

**Exercise**
For those of you that have some familiarity with (linear) algebra:
Let $I$ be an open interval, and let $\mathcal{O}(I)$ be the set of analytic functions on $I$.

1. Show that $\mathcal{O}(I)$ is a subspace of $\mathcal{F}(I)$: show that $\mathcal{O}(I)$ is nonempty; show that for all $f, g \in \mathcal{O}(I)$ and any $a, b \in \mathbb{R}$, also $af + bg \in \mathcal{O}(I)$.
2. Show that $\mathcal{O}(I)$ is a *sub-algebra* of $\mathcal{F}(I)$: show that for all $f, g \in \mathcal{O}(I)$, also $fg \in \mathcal{O}(I)$. (*Hint*: the Cauchy product is helpful.)
3. Show that $D: \mathcal{O}(I) \to \mathcal{O}(I)$ defined by $D(f) = f'$ is a linear transformation.
4. Let $a \in I$. Let $\mathfrak{m}_a \subseteq \mathcal{O}(I)$ be defined as $\mathfrak{m}_a = \{f \in \mathcal{O}(I) \mid f(a) = 0\}$. Show that $\mathcal{O}(I) = \mathbb{R} \oplus \mathfrak{m}_a$, where $\mathbb{R}$ is identified with the constant functions on $I$.

(1. and 2. are essentially part of the previous exercise.) EOE.

**Remark**
Analytic functions are a large class of functions that we can actually describe effectively (at least in principle), since for every point $c$ in their domain, we are given a power series (again, at least in principle) that agrees with the function around $c$. In particular, we have something as close as possible to a "formula" for determining the value of the function.
Maybe even more important is the following fact: one can extend most definitions we have encountered so far to complex valued functions defined on subsets of the complex numbers. Surprisingly, if a function $f: \mathbb{C} \to \mathbb{C}$ is differentiable everywhere (the definition of "differentiable" is the same as the one we have seen, namely that $\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$ exists and is a complex number), then it is automatically analytic! Moreover, if a real valued function $f$ is a power series on the interval $I = (c - R, c + R)$, then there is a *unique* complex differentiable function $\tilde{f}$ defined on $D = \{z \in \mathbb{C} \mid |z - c| < R\}$ that agrees with $f$ on $I$.
Consider the following function $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^x}$. We know that $f(x_0)$ converges for every $x_0 > 1$. One can show that this function is analytic, for $x > 1$ and that the series also converges for $x \in \mathbb{C}$ where $\Re(x) > 1$.

**Warning**
Tying up another loose end, if $f$ is a power series with centre $c$ and *finite* radius of convergence $R > 0$, then we have seen that $f$ is analytic on $I = (c - R, c + R)$. For $d \in I$, we can rewrite $f$ as a power series $g$ with centre $d$, and we have seen that the radius of convergence of $g$ is *at least* $S = R - |d - c|$. However, it *may be strictly larger*. Thus, $g$ may actually converge for points strictly outside $I$ (not just on the boundary of $I$.
Consider $f(x) = \sum_{n=0}^{\infty} x^n$. For $|x| < 1$ this is equal to $\frac{1}{1-x}$. Let $h(x) = \frac{1}{1-x}$. Note that $h$ is defined at $-1$, and $h(-1) = \frac{1}{2}$, whereas $f(x)$ does not converge for $x = -1$ (but $f(x)$ is bounded close to $x - 1$ and that is no coincidence).
Consider $d = -\frac{1}{2}$. Put $g(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(d)}{n!}(x - d)^n$.
Next, observe that $f'(x) = \frac{1}{(1-x)^2}, f''(x) = \frac{2}{(1-x)^3}, \dots, f^{(n)}(x) = \frac{n!}{(1-x)^{n+1}}$. Hence

$$g(x) = \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^{n+1} \left(x + \frac{1}{2}\right)^n$$

The inverse of the radius of convergence of this series is $\limsup_{n\to\infty} \sqrt[n]{\left(\frac{2}{3}\right)^{n+1}} = \frac{2}{3}$, and the radius of convergence is $\frac{3}{2}$. One can show that $h$ and $g$ coincide on all of $(-2,1)$.
We know this for the interval $(-1,0)$, as by the Idendity Theorem $f = g = h$ on $(-1,1)$.
One can show that $h$ is analytic everywhere in $\mathbb{R} \setminus \{1\}$. But there is no single power series convergent in all its domain.
On the other hand, if a function is analytic on all of $\mathbb{R}$, then there exists a power series (centred at 0) with convergence radius $\infty$ agreeing with the function. EOW.

### 6.3.9 *Excursion: Abel's Limit Theorem

When discussing power series, we observed that if a series $f(x)$ has radius of convergence $R > 0$ nothing can be said in general about the series evaluated at the boundary of its interval of convergence.

**Theorem (Abel's Limit Theorem)**
Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $R > 0$. Suppose $f(R)$ converges. Then $\lim_{x\to R^-} f(x) = f(R)$. A similar statement holds if $f(-R)$ converges, or if $f$ is centred at a number different from $c$. EOT.

This might sound obvious, but while we showed that $f$ is continuous on $(-R, R)$, we did not claim anything about the boundary points $\pm R$, even if $f$ happens to converge there.

Proof. We must show that $f(x) \to f(R)$ as $x \to R^-$. Let us first assume that $R = 1$.

Then $f(1) = \sum_{n=0}^{\infty} a_n$.

Let $0 < x < 1$.

$$f(x) - f(1) = \frac{1 - x}{1 - x}\big(f(x) - f(1)\big)$$

Note that $\frac{1}{1-x} f(x) = \sum x^n \sum a_n x^n = \sum_{n=0}^{\infty} S_n x^n$ where $S_n = \sum_{k=0}^{n} a_n$.

On the other hand, $\frac{f(1)}{1-x} = \sum_{n=0}^{\infty} f(1)x^n$.

$$f(x) - f(1) = (1-x)\left(\sum_{n=0}^{\infty} \big(S_n - f(1)\big)x^n\right)$$

Taking absolute values

$$|f(x) - f(1)| \leq |1-x| \sum_{n=0}^{\infty} |S_n - f(1)||x|^n$$

Note at this point we don't know yet that the right hand side is finite. Let $\varepsilon > 0$ be arbitrary. Then there is certainly some $n_0 \in \mathbb{N}$ such that $|S_n - f(1)| < \frac{\varepsilon}{2}$ for all $n > n_0$. Rewriting the above we find

$$|f(x) - f(1)| \leq |1-x| \sum_{n=0}^{n_0} |S_n - f(1)||x|^n + \frac{\varepsilon}{2}|1-x| \sum_{n=n_0+1}^{\infty} |x|^n$$

Now $0 < x < 1$ means $|1-x| = 1-x$ and $\sum_{n=n_0+1}^{\infty} |x|^n \leq \frac{1}{1-x}$. Together,

$$|f(x) - f(1)| < (1-x) \sum_{n=0}^{n_0} |S_n - f(1)|x^n + \frac{\varepsilon}{2} \leq (1-x) \sum_{n=0}^{n_0} |S_n - f(1)| + \frac{\varepsilon}{2}$$

There is $\delta > 0$ such that for all $0 < x < 1$ with $1 - x < \delta$ we have

$$(1-x) \sum_{n=0}^{n_0} |S_n - f(1)| < \frac{\varepsilon}{2}$$

This is trivial if $\sum_{n=0}^{\infty} |S_n - f(1)| = 0$, and otherwise $\delta < \frac{\varepsilon}{2\sum_{n=0}^{n_0} |S_n - f(1)|}$ works.

For all such $x$ we then have

$$|f(x) - f(1)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

By definition this means $\lim_{x \to 1^-} f(x) = f(1)$.

If $R \neq 1$, let $g(x) = f(Rx)$. Then the radius of convergence of $g$ is 1, and $g(1)$ converges if and only if $f(R)$ converges, and then $f(1) = g(1)$. In this case we conclude that $\lim_{x \to R^-} f(x) = \lim_{x \to 1^-} g(x) = g(1) = f(1)$.

The case of $-R$ is similar. We leave the case of a centre other than 0 as an exercise. QED.

For an example see the discussion of $\arctan x$ in 6.4.5 below.

It might be tempting to try to prove a theorem of the form if $\lim_{x \to R} f(x)$ exists, then $f(R)$ converges and equals that limit. However, as the geometric series shows, this is not true in general:

$$\lim_{x \to 1^-} \sum_{n=0}^{\infty} x^n = \lim_{x \to 1^-} \frac{1}{1-x} = \frac{1}{2}$$

But $\sum_{n=0}^{\infty}(-1)^n$ diverges.

## 6.4 Power series and differential equations

Power series are tremendously helpful when trying to solve certain differential equations.

### 6.4.1 Differential equations

Let us briefly discuss in a quasi-formal way what we mean by a differential equation. We have discussed examples of the form $f' = f$ or $f'' = -f$. But what does this precisely mean?

**Definition**

An **ordinary differential equation** (ODE, for short) **of order** $n$ is a function $F: \Omega \to \mathbb{R}$ where $\Omega \subseteq \mathbb{R}^{n+2}$ is some subset. A **solution** of such an ODE is a function $u$ defined on some interval $I$, $n$-times differentiable, such that for all $x \in I$, $\left(x, u(x), u'(x), \dots, u^{(n)}(x)\right) \in \Omega$, and

$$F\left(x, u(x), u'(x), \dots, u^{(n)}(x)\right) = 0$$

EOD.

This is a bit technical, and we will not dwell on this. We are for the time being mostly interested in linear ODEs with constant coefficients.

**Examples**

1. $\Omega = \mathbb{R}^3$, $F(x, x_0, x_1) = x_0 - x_1$. A solution of this ODE is a function $u$ such that $F\left(x, u(x), u'(x)\right) = 0$, or, in other words, $u'(x) = u(x)$ for all $x$ in some interval $I$. Usually we write this ODE simply as $f' = f$.
2. $\Omega = \mathbb{R}^4$, $F(x, x_0, x_1, x_2) = x_0 + x_2$. A solution of this ODE is a function $u$ such that $F\left(x, u(x), u'(x), u''(x)\right) = u(x) + u''(x) = 0$, ie. $u''(x) = -u(x)$ for all $x$ in some interval $I$. We would write this ODE as $f'' = -f$.
3. In general, a *homogeneous linear ODE with constant coefficients* is a linear function $F: \mathbb{R}^{n+2} \to \mathbb{R}$, so that $F(x, x_0, x_1, \dots, x_n) = c_0 x_0 + c_1 x_1 + \cdots + c_n x_n$. A solution is then a function $u$ defined on an interval $I$ such that $c_0 u(x) + c_1 u'(x) + \cdots + c_n u^{(n)}(x) = 0$ for all $x \in I$. Usually, we require $c_n \neq 0$ (otherwise it could be rewritten as an ODE of smaller order). Again, we would write this ODE as $c_n f^{(n)} = -c_0 f - c_1 f' - \cdots - c_{n-1} f^{(n-1)}$.
   Note that "constant coefficients" refers to the fact that the $c_i$ are constants. In general, they are allowed to be functions of the first variable $x$. "Homogeneous" refers to the fact that the "linear" function $F$ has no "constant" term, that is no term depending only on $x$.

EOE.

To complete and finish the discussion we also define the notion of an initial value problem:

**Definition**

An **initial value problem** (IVP) is an ODE $F$ of order $n$ as above, together with an element $(t_0, x_0, x_1, \dots, x_{n-1}) \in \mathbb{R}^{n+1}$. A **solution** of the IVP is a solution $u$ of the ODE $F$ on an interval $I$ containing $t_0$ such that $u(t_0) = x_0$, $u'(t_0) = x_1, \dots, u^{(n-1)}(t_0) = x_{n-1}$. EOD.

**Example**
For the ODE $f' = f$ above, an IVP would be determined by the additional condition $f(t_0) = x_0$. For the ODE $f'' = -f$, it would be $f(t_0) = x_0$ and $f'(t_0) = x_1$. EOE.

### 6.4.2 ODEs and power series

Suppose we want to solve an ODE given by some $F: \Omega \to \mathbb{R}$.

A natural approach would be to assume that our solution is analytic or a power series. If $u(x) = \sum_{k=0}^{\infty} a_k (x-c)^k$, say, the condition that $F\big(x, u(x), D(u)(x), \dots, D^n(u)(x)\big) = 0$ should result in conditions on the coefficients $a_k$ that we hopefully can solve.

We illustrate this at two examples.

### 6.4.3 The equation $f' = f$

We have discussed this equation previously in MATH 117. And of course, we all know that the exponential function is a solution. Let us pretend we haven't done so, though.

Suppose $u(x) = \sum_{n=0}^{\infty} a_n x^n$ is a solution of $f' = f$. Then

$$u'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n = \sum_{n=0}^{\infty} a_n x^n$$

Assuming the solution $u(x)$ converges around 0 we know that this equation then means $a_n = (n+1)a_{n+1}$ for all $n$ (e.g. by the Identity Theorem).

We obtain the following recursive formula: $a_{n+1} = \frac{1}{n+1} a_n$. If we turn the question into an IVP by requiring $u(0) = 1$, say, the Recursive Definition Theorem guarantees that there is a unique sequence $a_0, a_1, \dots$ that satisfies this recursive formula with $a_0 = 1$.

One easily verifies that $a_n = \frac{1}{n!}$ is this unique solution, and we get that $u(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ (hopefully, this does not come as a surprise).

We know that this series has radius of convergence $\infty$, and we obtain a solution of the IVP $f' = f$ with $f(0) = 1$.

We have discussed such solutions already in MATH 117, but let us recap the main points. (The "Excursion" there contained a small gap: it did not show that $f(x) \neq 0$ for all $x$. We will close this gap now).

Suppose $f' = f$ and there is $x_0$ with $f'(x_0) > 0$. Then $f(x) > 0$ for all $x \geq x_0$: Let $S = \{ x > x_0 \mid f(x) < 0 \}$. If $S \neq \emptyset$, then as $S$ is bounded below we have $x_0 \leq s \coloneqq \inf S$. By continuity, $f(s) = 0$: indeed, there is $\delta > 0$ such that $f(x) > 0$ for all $x \in [x_0, x+\delta)$. But then $s \geq x + \delta$; thus, $f(x) > 0$ for all $x \in [x_0, s)$. Continuity forces $f(s) \geq 0$. Thus $s \notin S$, and there is a sequence $x_n \in S$ with

$x_n \to s$. But then $f(s) = \lim f(x_n) \leq 0$. It follows that $f(s) = 0$. But $f$ is strictly monotone increasing on $[x_0, s]$, and so $f(s) > f(x_0) > 0$, a contradiction. Therefore $f(x) > 0$ for all $x \geq x_0$.

From this we conclude that if $g$ is any other solution, then $g(x) = cf(x)$ for all $x \geq x_0$ and some constant $c$. Indeed, consider (on $[x_0, \infty)$) the function $h = \frac{g}{f}$. Then $h'(x) = \frac{g'(x)f(x)-g(x)f'(x)}{f^2(x)} = 0$ for all $x$, and thus $h(x) = c$ is constant.

It now follows that $f(x) > 0$ for all $x \in \mathbb{R}$. Indeed, let $x_1 < x_0$. Let $z \coloneqq x_0 - x_1$, and put $g(x) \coloneqq f(x - z)$. Then $g' = g$, and so for $x \geq x_0$, we have $g(x) = cf(x)$ for some constant $c$. Now $g(x_0) = f(x_0 - z) = f(x_1) = cf(x_0)$. But $c > 0$, since for $x > x_0 + z$, we know that $g(x) = cf(x - z) > 0$.

We have shown the following Lemma.

**Lemma**
Let $f$ be defined on $\mathbb{R}$ and suppose $f' = f$.

1. If $f(x_0) > 0$ for some $x_0$, then $f$ is strictly monotone increasing, and $f > 0$.
2. If $f(x_0) < 0$ for some $x_0$, then $f$ is strictly monotone decreasing, and $f < 0$.
3. If $f(x_0) = 0$ for some $x_0$, then $f = 0$.

EOL.

Proof. 1. we just did. 2. is 1. applied to $-f$. And 3. follows since $f(x)$ cannot be positive or negative anywhere otherwise 1. or 2. Applied. QED.

**Corollary**
Let $f$ be a solution of $f' = f$ with $f(0) = 1$. Then

1. Every solution of $f' = f$ is of the form $g = cf$ where $c = g(0)$.
2. $f$ is an exponential function, that is, $f(x + y) = f(x)f(y)$ for all $x, y$.

EOC.

Proof. The first part rests on the fact that $f > 0$, so as above one shows that $\frac{d}{dx}\frac{g}{f} = 0$ and so $g = cf$. And $g(0) = cf(0) = c$.

For the second part, fix $y \in \mathbb{R}$, and apply 1. to the function $g(x) = f(x + y)$. Then $g = g(0)f = f(y)f$, as needed. QED.

**Remark**
When we discussed exponential functions, we defined them as follows: $f$ defined on $\mathbb{R}$ is an exponential function if

1. $f$ is continuous at $0$.
2. $f$ is not constant $0$.
3. $f(x + y) = f(x)f(y)$ for all $x, y$.

We then proceeded to show: such a function is continuous everywhere. We showed that $\exp(x)$ is an exponential function, by verifying 1., 3. and we also showed it is differentiable everywhere.
We then showed that every exponential function is of the form $\exp(cx)$ for some constant $c$.
We could have proceeded as follows: first show that $\exp x$ is a solution of $f' = f$ with $f(0) = 1$, bypassing proving 3. first. EOR.

### 6.4.4 The equation $f'' = -f$

In the following a "solution" is always a solution of the equation $f'' = -f$ defined on $\mathbb{R}$.

**Observation 1**
If $u$ is a solution of $f'' = -f$, then so is $u'$. EOO.

Proof. $(u')'' = (u'')' = (-u)' = -u'$. QED.

**Observation 2**
If $a, b \in \mathbb{R}$, and $u, v$ are solutions, then $au + bv$ is also a solution. EOO.

Proof. Differentiate twice. QED.

**Observation 3**
If $u$ is a solution, then the function $g = u^2 + (u')^2$ is constant. EOO.

Proof. $g$ is differentiable and $g'(x) = 2u(x)u'(x) + 2u'(x)u''(x) = 2u'(x)\big(u(x) + u''(x)\big) = 0$ for all $x$. QED.

**Corollary**
If $u$ is a solution such that $u(x_0) = u'(x_0) = 0$ for some $x_0$, then $u = 0$. EOC.

Proof. As above the function $g = u^2 + (u')^2$ is constant. But $g(x_0) = 0$. Therefore $u(x)^2 + u'(x)^2 = 0$ for all $x$. This means $u(x) = 0$ for all $x$. QED.

To construct a solution, let us assume $u(x) = \sum_{n=0}^{\infty} a_n x^n$ is a power series and a solution. Let us assume $u(0) = a_0 = 0$, and $u'(0) = a_1 = 1$.

Then $u''(x) = \sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}x^n$. Comparing coefficients for each $n$, we get

$$a_{n+2} = -\frac{a_n}{(n+2)(n+1)}$$

for all $n$. This defines $a_n$ recursively for even and odd numbers:

$b_k \coloneqq a_{2k}$ and $c_k := a_{2k+1}$ results in the two recursions

$$b_{k+1} = -\frac{b_k}{(2k+2)(2k+1)}$$

and

$$c_{k+1} = -\frac{c_k}{(2k+3)(2k+2)}$$

subject to the starting conditions $c_0 = 1$ and $b_0 = 0$. The Recursive Definition Theorem guarantees unique sequences $b_k$, $c_k$: $b_k = 0$ for all $k$ and $c_k = \frac{(-1)^k}{(2k+1)!}$.

We obtain the formal power series

$$u(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

It has radius of convergence $\infty$: indeed, $\sum_{n=0}^{\infty} |\frac{(-1)^n}{(2n+1)!} x^{2n+1}| \leq \sum_{n=0}^{\infty} \left|\frac{x^n}{n!}\right| < \infty$.

Let us write $S(x)$ for this distinguished solution. Then by Observation 1, $C(x) \coloneqq S(x)$ is also a solution.

Note that $C(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$.

Finally, observe $S(0) = -C'(0) = 0$, $S'(0) = C(0) = 1$.

**Lemma**
If $u$ is a solution, then $u = u'(0)S + u(0)C$. Conversely, $u = aS + bC$ is the unique solution with $u(0) = b$ and $u'(0) = a$. EOL.

Proof. We know that $h = aS + bC$ is a solution by Observation 2. Moreover, $h(0) = b$ and $h'(0) = a$.

On the other hand, if $u$ is any solution, let $g = u'(0)S + u(0)C$. Then $u - g$ is a solution. But $(u - g)(0) = (u - g)'(0) = 0$, and therefore $u - g = 0$. Thus, $u = g$. This shows that $u$ is uniquely determined by $u(0)$ and $u'(0)$. QED.

### 6.4.5 Trigonometric functions revisited
We have discussed $\sin x$ and $\cos x$ before. But we have never properly shown that these functions actually exist. We will remedy this now.

When discussing these functions, we postulated that they satisfy the following conditions based on geometric considerations:

T1 $\sin(x)$ is odd, and $\cos(x)$ is even.
T2 $\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y)$
T3 $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$
T4 For every sequence $x_n \to 0$ with $x_n \neq 0$ for all $n$, $\frac{\sin x_n}{x_n} \to 1$.
T5 For every sequence $x_n \to 0$, $\cos(x_n) \to 1$.

We showed that this forced $\sin x$ and $\cos x$ to be differentiable and $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$. In particular, $\sin''(x) = -\sin(x)$. It also follows that $\sin(0) = 0$ (by T1) and $\cos(0) = 1$ (by T5).

This shows, that if $\sin x$ and $\cos x$ really exist as well defined functions satisfying these properties, then we must have $\sin x = S(x)$ and $\cos x = C(x)$ by the previous section:

$$\sin x = \cos(0)\, S(x) + \sin(0)\, C(x)$$

What therefore remains to be seen is that $S(x)$ and $C(x)$ satisfy T1 – T5 with $S$ playing the role of $\sin x$ and $C$ that of $\cos x$.

T1 This follows from a homework exercise since $S(x)$ has only odd powers of $x$ and $C(x)$ has only even powers of $x$.
T2 One could compute $S(x + y)$ as a series and compare to the right hand side. However, a quicker method is the following: Fix $y \in \mathbb{R}$ and put $u(x) = S(x + y)$. Then $u'' = -u$, and therefore $u = u'(0)S + u(0)C$. But $u(0) = S(y)$ and $u'(0) = C(y)$. Therefore $u(x) = S(x + y) = C(y)S(x) + S(y)C(x)$ for all $x, y$.

T3 This is the same argument as in T2 applied to $u(x) = C(x + y)$.

T4 $S(x)$ is a power series convergent everywhere, so it is differentiable at 0 and $S'(0) = C(0) = 1$. But that is what T4 says.

T5 $C$ is continuous so $C(x_n) \to C(0) = 1$.

What remains is to verify some of the properties we expect $\sin x$ and $\cos x$ to have.

We already know that $\sin^2(x) + \cos^2(x) = \sin^2(0) + \cos^2(0) = 1$ by Observation 3 above.

**Lemma**
There is a minimal positive number $x_0$ such that $\cos x_0 = 0$. EOL.

Proof. We first show that $A \coloneqq \{ x > 0 \mid \cos(x) = 0 \}$ is nonempty. We proceed by a proof by contradiction. Suppose $A$ is empty. Then $\cos(x) > 0$ for all $x \geq 0$: indeed, if there is $x_1 > 0$ such that $\cos x_1 < 0$, then by the IVT (applied to $[0, x_1]$) there is $x \in (0, x_1)$ with $\cos x = 0$.

This shows that $\sin(x)$ is strictly monotone increasing on $[0, \infty)$. It is also bounded above by 1 (since $\sin^2(x) + \cos^2(x) = 1$).

This means that $s_0 \coloneqq \lim\limits_{x\to\infty} \sin(x)$ exists and is finite. (Indeed, $s_0 = \sup\limits_{x \geq 0} \sin x$). Also note that as $\sin x$ is strictly increasing, $\sin(x) > 0$ for all $x > 0$, which forces $s_0 > 0$ and also $\cos x$ to be strictly decreasing on $[0, \infty)$. By assumption $\cos x > 0$, so $t_0 \coloneqq \lim\limits_{x\to\infty} \cos x$ exists, is finite, and $t_0 \geq 0$.

T2 applied to $x = y$ says

$$\sin(2x) = 2\sin(x)\cos(x)$$

Taking the limit for $x \to \infty$ on both sides, we get (1) $s_0 = 2 s_0 t_0$. Since $s_0 > 0$, (1) implies $t_0 = \frac{1}{2}$. As $\cos x$ approaches $t_0$ from above, we find $\cos(x) \geq \frac{1}{2}$ on $[0, \infty)$. By the Horse Race Theorem, this means $\sin(x) \geq \frac{1}{2}x$ for all $x \in [0, \infty)$, which contradicts that $\sin x$ is bounded. We are forced to conclude that $A \neq \emptyset$. Let $x_0 \coloneqq \inf A$. Since $\cos(x)$ is continuous and $\cos(0) = 1$, there is $\delta > 0$ such that $\cos(x) > 0$ for $x \in [0, \delta)$. This shows $x_0 \geq \delta > 0$. If $x_0 \notin A$, there is a sequence $x_n \in A$ with $x_n \to x_0$. As $\cos(x)$ is continuous $0 = \cos(x_n) \to \cos(x_0)$, so $x_0 \in A$. This is a contradiction, and we must have $\cos(x_0) = 0$. QED.

**Definition**
We define the real number $\pi$ as $\pi \coloneqq 2x_0$. EOD.

**Remark**
This definition of $\pi$ is consistent with the geometric definition. However, since we don't know what a circle, let alone its circumference, is, we use this as an independent definition. If one does introduce geometry rigorously, then it does follow that our $\pi$ here is equal to the "geometric" $\pi$. EOR.

**Exercise**
Show that $\sin\left(\frac{\pi}{2}\right) = 1$, and $\sin\left(-\frac{\pi}{2}\right) = -1$. EOE.

$\cos(x) > 0$ for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. This follows from the fact that we know this by definition of $\pi$ for $x \in \left[0, \frac{\pi}{2}\right)$, and it follows for $x \in \left(-\frac{\pi}{2}, 0\right]$ because $\cos x$ is even.

Therefore, the function

$$\tan x = \frac{\sin x}{\cos x}$$

(tangent function) is well defined on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Its derivative is $\frac{1}{\cos^2(x)} > 0$. Since $\sin\left(\frac{\pi}{2}\right) = 1$, and $\cos(x) \to 0$ for $x \to \frac{\pi}{2}$, it follows that $\tan x \to \infty$ for $x \to \frac{\pi}{2}$. $\tan(x)$ is obviously an odd function, and therefore $\tan x \to -\infty$ for $x \to -\frac{\pi}{2}$.

We conclude that $\tan x$ has range $\mathbb{R}$, and is strictly monotone increasing. It follows that $\tan x$ is invertible, with inverse function denoted "arcus tangent" $\arctan x$ defined on $\mathbb{R}$ with range $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Recall from MATH 117, that

$\frac{d}{dx} \arctan x = \cos^2(arctan\, x)$, and $\tan^2(x) = \frac{\sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} - 1$, and therefore $\cos^2(x) = \frac{1}{1+\tan^2(x)}$. Thus,

$$\frac{d}{dx} \arctan(x) = \frac{1}{1 + x^2}$$

This holds for all $x \in \mathbb{R}$. For $|x| < 1$, we can use the geometric series (evaluated in $-x^2$) and find

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

Computing the formal anti-derivative, we get $f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$. Since $f(0) = 0 = \arctan 0$, $f$ is the Taylor series of $\arctan x$ around $0$, and $\arctan x = f(x)$ for $x \in (-1,1)$.

Note that $f(\pm 1)$ still converges by the Leibniz Rule. By Abel's Limit Theorem, it follows that $f(1) = \lim_{x \to 1^-} f(x) = \lim_{x \to 1^-} \arctan x = \arctan 1$. But $\arctan 1 = \frac{\pi}{4}$: to see this note that $1 = \sin\left(\frac{\pi}{2}\right) = \sin\left(2 \cdot \frac{\pi}{4}\right) = 2 \sin\left(\frac{\pi}{4}\right) \cos\left(\frac{\pi}{4}\right)$. On the other hand, $0 = \cos\left(2 \cdot \frac{\pi}{4}\right) = \cos^2\left(\frac{\pi}{4}\right) - \sin^2\left(\frac{\pi}{4}\right)$, so $\sin\left(\frac{\pi}{4}\right) = \pm \cos\left(\frac{\pi}{4}\right)$. Taken together $\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}$.

This shows that $f(1) = \frac{\pi}{4}$, and we obtain the formula

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

### 6.4.6 *Excursion: The general homogenous linear ODE with constant coefficients

The observations above were made *ad hoc*. One should study ODEs from a more conceptual viewpoint. As such it is just as much an object of Linear Algebra as Calculus. In particular eigenvalue theory can be effectively used to study such equations. We don't have time (and the necessary linear algebra background) to go into depth.

Suppose we want to solve an $n$-th order ODE of the form

$$f^{(n)} + c_{n-1}f^{(n-1)} + \cdots + c_1 f' + c_0 f = 0$$

(The general case with a nonzero coefficient $c_n$ in front of $f^{(n)}$ can be easily reduced to this case.)

Associated to this equation is the polynomial $P(x) = x^n + c_{n-1}x^n + \cdots + c_1 x + c_0$. (It is what is called the *characteristic polynomial* of a special matrix associated to this equation.)

Suppose $\lambda$ is a (real) root of $P$. Then $u(x) = e^{\lambda x}$ is a solution. Indeed, one easily checks that

$$u^{(n)}(x) + c_{n-1}u^{(n-1)}(x) + \cdots + c_0 u(x) = P(\lambda)u(x) = 0$$

If $\lambda = i\omega \in \mathbb{C}$ is a purely imaginary root (then also $-\lambda$ is a root), then $u(x) = \cos(\omega x)$ and $v(x) = \sin(\omega x)$ are solutions. (This is seen by looking at it from a complex point of view and observe that $e^{\lambda x} = \cos(\omega x) + i\sin(\omega x)$ is still a solution, and then both the real and imaginary parts are solutions.)

In general, if $\lambda = a + i\omega$ is a complex root, then $\cos(\omega x)e^{ax}$ and $\sin(\omega x)\, e^{ax}$ are solutions.

If $P$ has $n$ distinct roots, this gives $n$ linearly independent solutions, and one can show that these are all in the sense that every other solution is a linear combination of those. If $P$ has multiplicities, then there are more complicated solutions.

For example, if $P = (x-1)^2 = x^2 - 2x + 1$, then one solution is $u(x) = e^x$:

$$u''(x) - 2u'(x) + u(x) = e^x - 2e^x + e^x = 0$$

But another solution is $v(x) = xe^x$. $v'(x) = e^x + xe^x$, and $v''(x) = e^x + e^x + xe^x = 2e^x + xe^x$. And so

$$v''(x) - 2v'(x) + v(x) = 2e^x + xe^x - 2e^x - 2xe^x + xe^x = 0$$

It is a valid question what we would need linear algebra for. The point is that the equation above of order $n$ is a special case of a system of $n$ first order ODEs. A general such system has the form

$$\begin{aligned} f_1' &= c_{11}f_1 + c_{12}f_2 + \cdots + c_{1n}f_n \\ f_2' &= c_{21}f_1 + c_{22}f_2 + \cdots + c_{2n}f_n \\ &\vdots \\ f_n' &= c_{n1}f_1 + c_{n2}f_2 + \cdots + c_{nn}f_n \end{aligned}$$

A solution consists of $n$ functions $f_1, f_2, \dots, f_n$ satisfying these equations. If we form the matrix $A = [c_{ij}]$, we can formally write

$$\begin{bmatrix} f_1' \\ f_2' \\ \vdots \\ f_n' \end{bmatrix} = A \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

where the right hand side means matrix multiplication. This is to be interpreted pointwise (for each $x \in \mathbb{R}$). If one can *diagonalize* $A$, then one can solve this system immediately. This is referred to as “decoupling” the system. We won’t go into details, but if $A$ is already diagonal, this should be clear: in this case $c_{ij} = 0$ unless $i = j$. And the system is equivalent to $f_1' = c_{11}f_1, f_2' = c_{22}f_2, \dots, f_n' = c_{nn}f_n$, and the solutions are $f_1 = e^{c_{11}x}$, $f_2 = e^{c_{22}x}, \dots, f_n = e^{c_{nn}x}$ and linear combinations thereof.

Our $n$th order ODE fits into this scheme, by putting $f_1' = f_2, f_2' = f_3, \dots, f_{n-1}' = f_n, f_n' = -c_0f_1 - c_1f_2 - \dots - c_{n-1}f_n$. Note that $f_i = f^{(i-1)}$. The characteristic polynomial of the matrix $A$ is in this case the polynomial $P$ (up to a sign).