1 Material from the Fall Course

The material in the section is from Part I of this course. It appeared as notes titled *The Theorem of Mertens and the Exponential Function*. It is collected here for your/our/my/somebody’s convenience.

1.1 A brief summary on power series

Perhaps the main theorem about power series is:

**Theorem 1** Assume $\sum_{n=0}^{\infty} a_n(x-a)^n$ is a power series. Let $r = \frac{1}{\limsup_{n \to \infty} \sqrt[n]{|a_n|}}$, where if $\limsup_{n \to \infty} \sqrt[n]{|a_n|} = 0$ we let $r = \infty$, if $\limsup_{n \to \infty} \sqrt[n]{|a_n|} = \infty$ we let $r = 0$. The following holds:

1. If $r = 0$ then the series converges only for $x = c$.
2. If $0 < r \leq \infty$, the series converges absolutely for all $x$ with $|x-c| < r$.
3. If $r < \infty$, the series diverges for all $x$ with $|x-c| > r$.

The theorem says nothing about what happens if $|x-c| = r$. That has to be determined in every case. The number $r > 0$ is known as the *radius of convergence* of the series.

Given a power series $\sum_{n=0}^{\infty} a_n(x-a)^n$ with radius $r > 0$ we can define $f : (c-r,c+r)$ by

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

for all $x \in (c-r,c+r)$. If $r = \infty$, then $(c-r,c+r) = \mathbb{R}$.

Later on we will see that a function defined by a power series is infinitely many times differentiable in its interval of definition. For now it will suffice to know that such a function is differentiable at the center.

**Lemma 2** Let

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

for $|x-c| < r$, where $r > 0$. Then $f$ is differentiable at $c$, hence also continuous at $c$. Moreover $f'(c) = a_1$. 
Proof. We have

$$f(x) = a_0 + a_1(x - c) + \sum_{n=2}^{\infty} a_n(x - a)^n = a_0 + a_1(x - c) + (x - c)^2 g(x)$$

where

$$g(x) = \sum_{n=2}^{\infty} a_n(x - a)^{n-2} = \sum_{n=0}^{\infty} a_{n+2}(x - c)^n.$$  

The function $g$ is defined by a series having the same radius of convergence as the one defining $f$ (because $\limsup_{n \to \infty} \sqrt[n]{|a_{n+2}|} = \limsup_{n \to \infty} \sqrt[n]{|a_n|}$), so if $0 < \rho < r$ we have for $|x - c| < \rho$ that

$$|g(x)| \leq \sum_{n=0}^{\infty} |a_{n+2}| |x - c|^n \leq \sum_{n=0}^{\infty} |a_{n+2}| \rho^n =: M,$$

where $M$ is a finite constant. It follows that

$$\left| \frac{f(x) - f(c)}{x - c} - a_1 \right| = \left| \frac{f(x) - a_0}{x - c} - a_1 \right| = |x - c| |g(x)| \leq |x - c|M$$

for $|x - c| < \rho$. It follows that $\lim_{x \to c} |(f(x) - f(c))/(x - c) - a_1| = 0$.  

Incidentally, the coefficients of a power series converging to a function $f(x)$ in an interval of positive length are completely determined by the function. There is a formula which is known from calculus (we’ll also see it later on), but one can give an immediate proof. Suppose that

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

for $|x - c| < r$, but $a_n \neq b_n$ for some $n \in \mathbb{N} \cup \{0\}$. There is then a first such $n$; that is, there is $n \in \mathbb{N} \cup \{0\}$ such that $a_n \neq b_n$ but $a_k = b_k$ for $0 \leq k < n$. We now can write the assumed equality in the form

$$a_0 + \cdots + a_{n-1}(x-c)^{n-1} + \sum_{k=n}^{\infty} a_k(x-c)^k = a_0 + \cdots + a_{n-1}(x-c)^{n-1} + \sum_{k=n}^{\infty} b_k(x-c)^k$$

and one gets

$$(x - c)^n \sum_{k=0}^{\infty} a_{n+k}(x - c)^k = (x - c)^n \sum_{k=0}^{\infty} b_{n+k}(x - c)^k$$

for $|x - c| < r$. If $n = 0$ one has an immediate contradiction since the series on the left converges to $a_n$ the one on the right to $b_n$ for $x = c$. If $n > 0$ it is a little bit more subtle. We can cancel $(x - c)^n$ to get

$$\sum_{k=0}^{\infty} a_{n+k}(x - c)^k = \sum_{k=0}^{\infty} b_{n+k}(x - c)^k$$
for $0 < |x - c| < r$. But here is where Lemma 2 comes in. These two last series define functions in $|x - c| < r$ that are continuous at $c$; since they are equal for $x \neq c$, they must have the same limit for $x \to c$ and this now implies equality also at $c$, hence the contradiction $a_n = b_n$.

1.2 Mertens’ Theorem

We saw this theorem in class. The proof I’m giving here is the same version given in class, minus comments and with more summation symbols.

Theorem 3 Assume $\sum_{n=0}^{\infty} a_n$ is an absolutely convergent series, $\sum_{n=0}^{\infty} b_n$ a convergent series. Let $A = \sum_{n=0}^{\infty} a_n$, $B = \sum_{n=0}^{\infty} b_n$. For $n = 0, 1, 2, \ldots$ define $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Then

$$\sum_{n=0}^{\infty} c_n = AB.$$  

Proof. For $n = 0, 1, 2, \ldots$, set $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$, $\beta_n = B - B_n$. We have

$$\sum_{k=0}^{n} c_k = \sum_{k=0}^{n} \sum_{j=0}^{k} a_j b_{k-j} = \sum_{j=0}^{n} \sum_{k=0}^{n-j} a_j b_{k-j} = \sum_{k=0}^{n} a_j \sum_{k=0}^{n} b_{k-j}$$

$$= \sum_{j=0}^{n} a_j \sum_{k=0}^{n-j} b_k = \sum_{j=0}^{n} a_j B_{n-j} = \sum_{j=0}^{n} a_j (B - \beta_{n-j}) = A_n B - \sum_{j=0}^{n} a_j \beta_{n-j}.$$  

Since $\lim_{n \to \infty} A_n B = AB$, the theorem will be proved if we prove:

$$\lim_{n \to \infty} \sum_{j=0}^{n} a_j \beta_{n-j} = 0 \quad (1)$$

For this purpose, let $\epsilon > 0$ be given. The series $\sum_{n=0}^{\infty} a_n$ converges absolutely so we can let $M = \sum_{n=0}^{\infty} |a_n|$; then $M < \infty$. Since $\lim_{n \to \infty} \beta_n = 0$, there exists $N_1 \in \mathbb{N}$ such that $|\beta_n| < \frac{\epsilon}{2(M+1)}$ if $n \geq N_1$. moreover, the sequence $\{\beta_n\}$ is bounded (because it converges); there is $K$ such that $|\beta_n| \leq K$ for all $n$. In addition, there is $N_2 \in \mathbb{N}$ such that

$$\sum_{k=n}^{\infty} |a_k| < \frac{\epsilon}{2(K+1)}$$

if $n \geq N_2$. Let $n \geq N_1 + N_2$. Then

$$\left|\sum_{j=0}^{n} a_j b_{n-j}\right| \leq \sum_{j=0}^{n} |a_j||\beta_{n-j}| = \sum_{j=0}^{n-N_1} |a_j||\beta_{n-j}| + \sum_{j=n-N_1+1}^{n} |a_j||\beta_{n-j}|$$

In the first of the two summations on the far right side we have $j \leq n - N_1$, thus $n-j \geq N_1$ and $|\beta_{n-j}| < \epsilon/[2(M+1)]$. In the second summation we can
estimate \(|\beta_{n-k}\) by \(K\) leaving \(K\sum_{k=n-N_1+1}^{n} |a_k|\); since \(n-N_1+1 > N_2\), this sum of the \(|a_k|\)'s is bounded by \(\epsilon/[2(K+1)]\). We thus have

\[
\sum_{j=0}^{n} a_j b_{n-j} \leq \frac{\epsilon}{2(M+1)} \sum_{j=0}^{\infty} |a_j| + K \sum_{j=n-N_1+1}^{n} |a_j| \leq \frac{\epsilon}{2(M+1)} M + K \frac{\epsilon}{2(K+1)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\]

This proves (1) and the theorem.

### 1.3 The Exponential Function

We define a function \(\exp : \mathbb{R} \rightarrow \mathbb{R}\) by

\[
\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.
\]

This definition makes sense since the power series converges for all \(x \in \mathbb{R}\).

**Theorem 4** We have

\[
\exp(x) \exp(y) = \exp(x+y)
\]

for all \(x, y \in \mathbb{R}\).

**Proof.** Since the series converge absolutely, we have by Mertens’ Theorem

\[
\exp(x) \exp(y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n = \exp(x+y).
\]

Because of this property it makes sense to define the number \(e\) by \(e = \exp(1) = \sum_{n=0}^{\infty} 1/n!\) and write \(e^x\) for \(\exp(x)\). Theorem 4 states, in fact, that \(\exp(x)\) behaves as one would expect \(e\) to the power \(x\) to behave. For example,

\[
e^n = e \cdot \ldots \cdot e \quad \text{for} \quad n = 1, 2, 3, \ldots,
\]

\[
e^0 = \exp(0) = 1,
\]

\[e^{-x} = \frac{1}{e^x} \quad \text{for all} \quad x \in \mathbb{R},
\]

etc. We can also see quite easily that \(2 < e < 3\). That \(e > 2\) is immediate, since \(e = 1 + \frac{1}{2} + \frac{1}{6} + \sum_{n=4}^{\infty} \frac{1}{n!} < 1 + \frac{1}{2} + \frac{1}{6} + \sum_{n=4}^{\infty} 2^{-n} = \frac{8}{3} + \frac{1}{8} = \frac{67}{24} < 3\).
Theorem 5 The function \( \exp : x \mapsto e^x \) satisfies the following properties: \( \exp \) is a strictly increasing, infinitely many times differentiable function and is one-to-one from \( \mathbb{R} \) onto \((0, \infty)\). It is its own derivative.

Proof. By Lemma 2, \( \exp \) is differentiable at 0 and \( \exp'(0) = 1 \). If \( x \in \mathbb{R} \), then

\[
\frac{e^{x+h} - e^x}{h} = e^x \frac{e^h - 1}{h} \to e^x
\]
as \( h \to 0 \), proving \( \exp \) is differentiable at all \( x \in \mathbb{R} \) and \( \exp' = \exp \). But it is then obvious that \( \exp \) is infinitely many times differentiable with \( \exp^{(n)} = \exp \) for \( n = 0, 1, 2, 3, \ldots \). From

\[
e^x e^{-x} = e^{x-x} = e^0 = 1,
\]
we conclude that \( e^x \neq 0 \) for all \( x \). Then

\[
e^x = \left( e^{x/2} \right)^2 > 0,
\]
since squares of non-zero numbers are positive. It follows that \( e^x > 0 \) for all \( x \in \mathbb{R} \), hence \( \exp(\mathbb{R}) \subset (0, \infty) \). Since \( \exp'(x) = \exp(x) > 0 \) for all \( x \), we also get that \( \exp \) is strictly increasing, hence one-to-one. Since the image of \( \mathbb{R} \) under the continuous function \( \exp \) has to be connected, it has to be a subinterval of \((0, \infty)\), say \( \exp(\mathbb{R}) = (a, b) \), where \( 0 \leq a < b \leq \infty \). Because \( \exp \) is strictly increasing we get \( a = \lim_{x \to -\infty} e^x \), \( b = \lim_{x \to \infty} e^x \). Then for every sequence \( (x_n) \) diverging to \( \infty \) we will have \( \lim_{n \to \infty} e^{x_n} = b \). Since \( e > 2 > 1 \), we have \( \lim_{n \to \infty} e^n = \infty \), thus \( b = \infty \). Since \( e^n \to \infty \) as \( n \to \infty \), we get \( e^{-n} = 1/e^n \to 0 \) as \( n \to \infty \). It follows that \( a = \lim_{x \to -\infty} e^x = 0 \). The Theorem is proved.

The number \( e \) is considered one of the most important numbers in mathematics, so the following result has some interest. Not as much as it used to have before it was proved that \( e \) was transcendental.

Theorem 6 \( e \) is irrational

Proof. By Taylor’s theorem, working with 0 as center, for \( n \in \mathbb{N} \),

\[
\exp(x) = \sum_{k=0}^{n} \frac{1}{k!} \exp^{(k)}(0)x^k + \frac{1}{(n+1)!} \exp^{(n+1)}(\bar{x})x^{n+1} = \sum_{k=0}^{n} \frac{1}{k!} x^k + \frac{1}{(n+1)!} e^{\bar{x}}x^{n+1}
\]

for some point \( \bar{x} \) between 0 and \( k \). Taking \( x = 1 \), using the fact that then all terms in the series for \( e \) are positive, and that if \( 0 < \bar{x} < 1 \) then \( e^{\bar{x}} < e < 3 \), we get

\[
0 < e - \sum_{k=0}^{n} \frac{1}{k!} < \frac{3}{(n+1)!}.
\]
This is true for all \( n \in \mathbb{N} \) assume now, for a contradiction, that \( e = a/b \) where \( a, b \in \mathbb{N} \). Replacing \( e \) by \( a/b \) and multiplying by \( n! \) the last inequalities become

\[
0 < \frac{n!a}{b} - \sum_{k=0}^{n} \frac{n!}{k!} < \frac{3}{n+1}.
\]

This is true for all \( n \) taking \( n > b \), then \( b \) divides \( n! \); \( k! \) divides \( n! \) for \( k = 0, \ldots, n \), thus \( z = \frac{n!a}{b} - \sum_{k=0}^{n} \frac{n!}{k!} \in \mathbb{Z} \), a contradiction since \( 0 < z < 1 \) and there is no integer strictly between 0 and 1.

1.4 The trigonometric functions

We can define \( \cos, \sin \) by their Taylor series. The series

\[
\sum_{k=0}^{\infty} \frac{(-1)^{2k}}{(2k)!}, \quad \sum_{k=0}^{\infty} \frac{(-1)^{2k+1}}{(2k+1)!}
\]

have radius of convergence \( \infty \); thus we can define:

\[
\cos x = \sum_{k=0}^{\infty} \frac{(-1)^{2k}}{(2k)!}, \quad \sin x = \sum_{k=0}^{\infty} \frac{(-1)^{2k+1}}{(2k+1)!}
\]

for \( x \in \mathbb{R} \).

It is clear that \( \cos \) is an even function since only even powers of \( x \) are involved, \( \cos(-x) = \cos x \) for all \( x \in \mathbb{R} \). Similarly, \( \sin \) is odd since only odd powers are involved; \( \sin(-x) = -\sin x \).

It is now a straightforward application of Mertens’ theorem to prove the following classical formulas for the sine and cosine.

**Theorem 7** Let \( x, y \in \mathbb{R} \). Then

\[
\cos(x + y) = \cos x \cos y - \sin x \sin y \quad (2)
\]

\[
\sin(x + y) = \sin x \cos y + \cos x \sin y \quad (3)
\]

**Proof.** Define \( a_n, b_n : \mathbb{R} \to \mathbb{R} \) for \( n = 0, 1, 2, \ldots \), by

\[
a_n(x) = \begin{cases} 
(-1)^k x^{2k} & \text{if } n = 2k \text{ is even}, \\
0 & \text{if } n \text{ is odd}.
\end{cases}
\]

\[
b_n(x) = \begin{cases} 
(-1)^k x^{2k+1} & \text{if } n = 2k + 1 \text{ is odd}, \\
0 & \text{if } n \text{ is even}.
\end{cases}
\]
Then \( \cos x = \sum_{n=0}^{\infty} a_n(x), \sin x = \sum_{n=0}^{\infty} b_n(x). \) Then \( \cos x \cos y = \sum_{n=0}^{\infty} c_n(x, y) \) where

\[
c_n(x, y) = \sum_{j=0}^{n} a_j(x)a_{n-j}(y).
\]

If \( n \) is odd, then one of \( j, n-j \) will be odd, and \( c_n(x, y) = 0. \) If \( n = 2m, \) since \( a_j \neq 0 \) only if \( j = 2k \) for some \( k, \) since \( 2k \leq 2m \) if and only if \( k \leq m, \) we have

\[
c_{2m}(x, y) = \sum_{k=0}^{m} a_{2k}(x)a_{2(m-k)}(y) = \sum_{k=0}^{m} \frac{(-1)^k x^{2k}(-1)^{m-k} x^{2k}}{(2k)! (2m-2k)!}
\]

\[
= (-1)^m \frac{1}{(2m)!} \sum_{k=0}^{m} \binom{m}{k} x^{2k} y^{2m-2k}.
\]

Similarly, \( \sin x \sin y = \sum_{n=0}^{\infty} d_n(x, y) \) where

\[
d_n(x, y) = \sum_{j=0}^{n} b_j(x)b_{n-j}(y).
\]

If \( n = 0 \) then \( d_n = b_0(y) = 0, \) so \( n \geq 1. \) If \( n \) is odd, then one of \( j, n-j \) will be even, and \( d_n(x, y) = 0. \) So once again we can write \( n = 2m, \) but this time \( m \geq 1. \) Now \( b_j \neq 0 \) only if \( j = 2k+1 \) for some \( k \) with \( 2k+1 \leq 2m, \) hence \( k \leq m - \frac{1}{2}; \) \( k \) being an integer we have \( k \leq m - 1. \) In the computations to follow notice that \( 2m - (2k+1) = 2(m-1-k) + 1. \) Thus

\[
d_{2m}(x, y) = \sum_{k=0}^{m-1} b_{2k+1}(x)b_{2(m-1-k)+1}(y) = \sum_{k=0}^{m} \frac{(-1)^k x^{2k+1}(-1)^{m-k} x^{2k}}{(2k+1)! (2m-2k-1)!}
\]

\[
= -(-1)^m \frac{1}{(2m)!} \sum_{k=0}^{m} \binom{m}{k} x^{2k+1} y^{2m-2k-1}.
\]

We now have for \( m \geq 1, \)

\[
c_{2m}(x, y) - d_{2m}(x, y)
\]

\[
= (-1)^m \frac{1}{(2m)!} \left( \sum_{k=0}^{m} \binom{2m}{2k} x^{2k} y^{2m-2k} + \sum_{k=0}^{m} \binom{2m}{2k+1} x^{2k+1} y^{2m-2k-1} \right)
\]

\[
= (-1)^m \frac{1}{(2m)!} \sum_{j=0}^{2m} \binom{2m}{j} x^j y^{2m-j} = (-1)^m \frac{1}{(2m)!} (x+y)^{2m}.
\]

For \( m = 0, \) since \( d_0 = 0, \) we have

\[
c_0(x, y) + d_0(x, y) = 1 = \frac{(-1)^0}{0!} (x+y)^0.
\]

Putting it all together we proved

\[
\cos x \cos y - \sin x \sin y = \sum_{m=0}^{\infty} (c_{2m}(x, y) + d_{2m}(x, y)) = \sum_{m=0}^{\infty} \frac{(-1)^m 1(x+y)^{2m}}{(2m)!} = \cos(x+y)
\]
proving (2). The proof of (3) is quite similar and will be omitted. □

Using these formulas we now can prove:

**Theorem 8** The functions \( \cos \) and \( \sin \) are infinitely many times differentiable and
\[
\cos' = -\sin, \quad \sin' = \cos.
\]

**Proof.** By Lemma 2, \( \cos \) is differentiable at 0 and its derivative at 0 is 0. By the same lemma, \( \sin \) is differentiable at 0 with derivative 1. Let \( x \in \mathbb{R} \).

\[
\frac{\cos(x+h) - \cos x}{h} = \frac{\cos x \cos h - \sin x \sin h - \cos x}{h} = \frac{\cos h - 1}{h} \cos x - \frac{\sin h}{h} \sin x.
\]

Now
\[
\lim_{h \to 0} \frac{\cos h - 1}{h} = 0, \quad \lim_{h \to 0} \frac{\sin h}{h} = 1
\]
since these limits are the derivatives at 0 of \( \cos \) and \( \sin \), respectively. It follows that
\[
\lim_{h \to 0} \frac{\cos(x+h) - \cos x}{h} = -\sin x
\]
proving \( \cos \) is differentiable with derivative \( -\sin \). Similarly
\[
\frac{\sin(x+h) - \sin x}{h} = \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} = \frac{\cos h - 1}{h} \sin x + \frac{\sin h}{h} \cos x \to \cos x
\]
as \( h \to 0 \). This proves that \( \sin \) is differentiable with derivative \( \cos \). It is now clear that both function are infinitely many times differentiable. □

Continuing getting properties of these functions, we now have

**Lemma 9**
\[\cos^2 x + \sin^2 x = 1\]
for all \( x \in \mathbb{R} \).

(We follow the usual convention of writing \( \cos^2 x, \sin^2 x \) for \( (\cos x)^2, (\sin x)^2 \), respectively.)

**Proof.** Let \( f(x) = \cos^2 x + \sin^2 x \). Then \( f'(x) = 2 \cos x (-\sin x) + 2 \sin x \cos x = 0 \) for all \( x \in \mathbb{R} \). Since \( f(0) = 1 \), we are done. □

As a consequence of this lemma we see that \( |\cos x| \leq 1, |\sin x| \leq 1 \) for all \( x \in \mathbb{R} \) and \( \cos x = 0 \) if and only if \( \sin x = \pm 1 \); \( \sin x = 0 \) if and only if \( \cos x = \pm 1 \). Next we bring \( \pi \) into the picture.

**Theorem 10** The function \( \cos \) has a zero in the interval \( (0, 2) \). More precisely, there is \( z_0 \in (0, 2) \) such that \( \cos z_0 = 0 \) and \( \cos x > 0 \) for \( 0 \leq x < z_0 \).
Proof. Since all odd derivatives of \( \cos \) at 0 are 0 (being equal to \( \pm \sin 0 \)) while the even ones alternate between 1 and \(-1\) (being equal to \( \pm \cos 0 \)), Taylor’s formula with remainder for \( \cos \) (centered at 0) gives

\[
\cos x = 1 - \frac{x^2}{2} + R_4(x)
\]

where (since the fourth derivative of \( \cos \) is \( \cos \))

\[
R_4(x) = \frac{\cos(\bar{x})}{24}x^4.
\]

For \( x = 2 \) we have

\[
|R_4(x)| \leq \frac{16|\cos(\bar{x})|}{24} \leq \frac{16}{24} < 1.
\]

Thus

\[
\cos 2 = 1 - \frac{4}{2} + R_4(2) = -1 + R_4(2) < 0.
\]

Because \( \cos 0 = 1 > 0 \), by the intermediate value theorem there is \( z \in (0, 2) \) such that \( \cos z = 0 \). In principle there could be more than one zero in this interval, so we let \( z_0 = \inf\{z \in (0, 2) : \cos z = 0\} \). As an inf, \( z_0 \) is a limit point of the set of zeros of \( \cos \) in \( (0, 2) \); since \( \cos \) is continuous we have \( \cos z_0 = 0 \). This also implies \( z_0 > 0 \) because \( \cos 0 = 1 > 0 \) and then \( \cos x > 0 \) if \( 0 < x < z_0 \).

It is also easy to see that \( z_0 \) is an isolated 0. There are no other close zeroes to the left, but could we have \( z_0 = \lim_{n \to \infty} z_n \) where \( (z_n) \) decreases to \( z_0 \) and \( \cos z_n \to 0 \)? The answer is no, because if this happens then (by the definition of derivative or by L'Hôpital) one would have that

\[
\sin z_0 = \lim_{n \to \infty} \cos z_n - \cos z_0 z_n - z_0 = 0;
\]

but \( \cos \) and \( \sin \) cannot be simultaneously zero by Lemma 9.

We can {\textit{define}} (sort of)

\[
\pi = 2z_0.
\]

From now on we denote this first positive zero of \( \cos \) by \( \pi/2 \). Since \( \cos x > 0 \) in \([0, \pi/2]\) and \( \sin' = \cos \) it follows that \( \sin \) is strictly increasing in the interval \([0, \pi/2]\), hence \( \sin(\pi/2) = 1 \). Incidentally, since then \( \sin x > 0 \) (increasing from 0) in \((0, \pi/2]\), we have that \( \cos \) is strictly decreasing in \([0, \pi/2]\). We also have by (2)

\[
\cos(x - \frac{\pi}{2}) = \cos x \cos(-\frac{\pi}{2}) - \sin x \sin(-\frac{\pi}{2}) = \cos x \cos \frac{\pi}{2} + \sin x \sin \frac{\pi}{2} = \sin x.
\]

Now \( x - \pi/2 \in [0, \pi/2] \) if \( x \in [\pi/2, \pi) \), thus \( \sin x = \cos(x - \pi/2) > 0 \) in \([\pi/2, \pi) \) with \( \sin \pi = \cos \pi/2 = 0 \). Thus \( \pi \) is the first positive 0 of \( \sin \). What does \( \cos \) do past \( \pi/4 \)? Since \( \sin x > 0 \) up to \( \pi \) and \( \cos' = -\sin \), \( \cos \) will decrease strictly
in \([\pi/2, \pi]\): at \(\pi\) the sine is 0, so \(\cos \pi = \pm 1\); but since it has been decreasing, we get \(\cos \pi = -1\). To go past \(\pi\) we use again formulas (2), (3) to see that

\[
\begin{align*}
\cos(x - \pi) &= \cos x \cos \pi + \sin x \sin \pi = -\cos x, \\
\sin(x - \pi) &= \sin x \cos \pi - \cos x \sin \pi = -\sin x.
\end{align*}
\]

We use this in the form \(\cos x = -\cos(x - \pi)\), \(\sin x = -\sin(x - \pi)\). So as \(x\) moves from \(\pi\) to \(2\pi\), the cosine and sine take on minus the values they took on as \(x\) moved from 0 to \(\pi\). From 0 to \(\pi\), cosine decreased from 1 to \(-1\), from \(\pi\) to \(2\pi\) it will increase from \(-1\) to 1. Concerning \(\sin\), from 0 to \(\pi\) it increased first to 1 then decreased to 0; now it will decrease to \(-1\), and then increase to 0. Here are two quite primitive pictures of the behavior so far.

We can now see that both \(\sin\) and \(\cos\) are periodic of period \(2\pi\): Since \(\cos(2\pi) = 1\), \(\sin(2\pi) = 0\),

\[
\begin{align*}
\cos(x + 2\pi) &= \cos x \cos(2\pi) - \sin x \sin(2\pi) = \cos x, \\
\sin(x + 2\pi) &= \sin x \cos(2\pi) + \cos x \sin(2\pi) = \sin x.
\end{align*}
\]

Moreover, \(2\pi\) is the fundamental period; if \(\cos(x + T) = \cos x\) for all \(x\) and some \(T\), then \(T\) has to be a multiple of \(2\pi\). In fact, with \(x = 0\) we get \(\cos T = 1\). Now let \(m \in \mathbb{Z}\) be such that \(2m\pi \leq T < 2(m + 1)\pi\). Then

\[
\cos(T - 2m\pi) = \cos T = 1,
\]

because \(0 \leq T - 2m\pi < 2\pi\) we conclude \(T - 2m\pi = 0\) (since the only point in \([0, 2\pi]\) at which \(\cos 1\) is 0), thus \(T = 2m\pi\). Similarly for \(\sin\): If \(\sin(x + T) = \sin x\) for all \(x\), then \(T = 2m\pi\) for some \(m \in \mathbb{Z}\).

Some of what we proved, and a bit more, can now be collected in the following theorem.

**Theorem 11** Let \(\Phi : \mathbb{R} \to \mathbb{R}^2\) be defined by \(\Phi(t) = (\cos t, \sin t)\) for \(t \in \mathbb{R}\). Then

1. \(\Phi\) is a bijection of the interval \([0, \pi/2]\) onto \(C_1 = \{ (x, y) \in \mathbb{R}^2, x^2 + y^2 = 1, x \geq 0, y \geq 0 \}\).
1. **MATERIAL FROM THE FALL COURSE**

2. \( \Phi \) is a bijection of the interval \([\pi/2, \pi]\) onto \( C_2 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1, x \leq 0, y \geq 0\} \).

3. \( \Phi \) is a bijection of the interval \([\pi, 3\pi/2]\) onto \( C_3 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1, x \leq 0, y \leq 0\} \).

4. \( \Phi \) is a bijection of the interval \([3\pi/2, 2\pi]\) onto \( C_4 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1, x \geq 0, y \geq 0\} \).

5. \( \Phi \) is a bijection of the interval \([0, 2\pi]\) onto the unit circle.

**Proof.** We begin noticing that \( |\Phi(t)|^2 = \cos t + \sin^2 t = 1 \) so that \( \Phi(\mathbb{R}) \subset C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \). Next, suppose \( t, t' \in \mathbb{R} \) and \( \Phi(t) = \Phi(t') \). We then have \( \cos t = \cos t' \) and \( \sin t = \sin t' \). Thus

\[
\cos(t - t') = \cos t \cos t' + \sin t \sin t' = \cos^2 t + \sin^2 t = 1.
\]

But, as we saw above, \( \cos x = 1 \) if and only if \( x = 2m\pi \) for some \( m \in \mathbb{Z} \). One conclusion is that then either \( t = t' \) or \( |t - t'| \geq 2\pi \). It follows that \( \Phi \) restricted to any interval of length \( < 2\pi \) must be one-to-one. This immediately show that \( \Phi \) is one-to-one on each of the intervals \([0, \pi/2], [\pi/2, \pi], [\pi, 3\pi/2], [3\pi/2, 2\pi]\). It is also one-to-one on \([0, 2\pi]\) because we are not including the endpoint \( 2\pi \).

There are several ways of seeing that it is also onto in the intervals in question to the arcs in question. For example, let \((x, y) \in C_1 \). Then \( 0 \leq x \leq 1 \) because the continuous function \( \cos \) decreases from 1 to 0 in \([0, \pi/2]\), by the intermediate value theorem there is \( t \in [0, \pi/2] \) such that \( \cos t = x \). In that interval \( \sin t \geq 0 \); from \( \cos^2 t + \sin^2 t = 1 \) we deduce that

\[
\sin t = +\sqrt{1 - \cos^2 t} = \sqrt{1 - x^2} = y,
\]

since \( y \geq 0 \). Similarly for all the other intervals.

As a final result for these notes we relate, via calculus, our very abstract \( \pi \) with the usual one. For this we compute the area and the length of the unit circle. By calculus, the area of the circle is

\[
A = 4 \int_0^1 \sqrt{1 - x^2} \, dx.
\]

Since \( \sin \) maps \([0, \pi/2]\) bijectively and increasingly onto \([0, 1]\) we get

\[
A = 4 \int_0^{\pi/2} \sqrt{1 - \sin^2 x} \cos x \, dx = 4 \int_0^{\pi/2} \cos^2 x \, dx.
\]

Notice that we used \( \cos x > 0 \) in \((0, \pi/2)\) to compute \( \sqrt{1 - \sin^2 x} \) as \( + \cos x \). By (2),

\[
\cos(2x) = \cos(x + x) = \cos x - \sin^2 x = \cos^2 x - (1 - \cos^2 x) = 2 \cos^2 x - 1
\]

so that

\[
\cos^2 x = \frac{1}{2} + \frac{1}{2} \cos(2x) = \frac{d}{dx} \left( \frac{1}{2} x + \frac{1}{4} \sin(2x) \right).
\]
By the fundamental theorem of calculus,

\[ \int_{0}^{\pi/2} \cos^2 x \, dx = \left. \frac{1}{2} x + \frac{1}{4} \sin(2x) \right|_{0}^{\pi/2} = \frac{\pi}{4} \]

\[ A = 4 \int_{0}^{\pi/2} \cos^2 x \, dx = 4 \frac{\pi}{4} = \pi. \]

For the length, we also use that the equation of the upper half of the circle is given by \( y = \sqrt{1 - x^2} \), thus

\[ ds = \sqrt{1 + \left( \frac{dy}{dx} \right)^2} \, dx = \frac{1}{\sqrt{1 - x^2}} \, dx. \]

Thus, changing variables as before,

\[ L = 4 \int_{0}^{1} \frac{1}{\sqrt{1 - x^2}} \, dx = 4 \int_{0}^{\pi/2} \frac{1}{\cos x} \cos x \, dx = 4 \int_{0}^{\pi/2} \, dx = 2\pi. \]

2 Recalling Taylor’s Theorem

Taylor’s Theorem is a very powerful generalization of the mean value theorem. I give here a very comprehensive statement, including more that what is usually stated.

First of all, a few words on differentiability. Differentiability is defined at points in the interior of the interval of definition of a function. There are several ways of extending this definition to include the endpoints but, since they are not all exactly equivalent one has to be careful by what one means. I will restrict the definition to continuous differentiability and define it as follows. Let \( I \) be an interval in \( \mathbb{R} \). Let \( f : I \to \mathbb{R} \). We will say that \( f \) is continuously differentiable on \( I \), write \( f \in C^1(I) \), if: a) \( f \) is differentiable in the usual sense at all interior points of \( I \); b) \( f' : I^\circ \to \mathbb{R} \) is continuous \((I^\circ = \text{interior of } I)\); c) If \( I \) contains a left endpoint \( a \), then \( \lim_{x \to a^+} f'(x) \) exists, and will be denoted by \( f'(a) \); d) Similarly, if \( I \) contains a right endpoint \( b \), then \( \lim_{x \to b^-} f'(x) \) exists, and will be denoted by \( f'(b) \).

By induction now: If \( m \in \mathbb{N}, m > 1 \), we will say that \( f : I \to \mathbb{R} \) is \( m \)-times continuously differentiable, and write \( f \in C^m(I) \) if \( f \in C^1(I) \) and \( f' \in C^{m-1}(I) \). Values of derivatives at endpoints are thus defined as the limits of the values inside the interval; these limits are supposed to exist.

Finally, \( C^\infty(I) = \bigcap_{m=0}^{\infty} C^m(I) \).

**Theorem 12 Taylor and more** Let \( I \) be an interval in \( \mathbb{R} \), let \( m \) be a non-negative integer, let \( f \in C^m(I) \) and assume that \( f^{(m)} \) is differentiable in \( I^\circ \) (the interior of \( I \)), thus that \( f^{(m+1)} \) exists in \( I^\circ \) without necessarily being continuous. Let \( c \in I \).
1. Let \( P_m(x) = \sum_{k=0}^{m} \frac{f^{(k)}(c)}{k!}(x-c)^k \). Then \( P_m \) is the only polynomial of degree \( \leq m \) such that

\[
P^{(k)}(c) = f^{(k)}(c) \quad \text{for} \quad k = 0, 1, \ldots, m.
\]

Naturally, as a polynomial of degree \( \leq m \), \( P^{(m+1)} \equiv 0 \).

2. We can write (of course) \( f = P_m + R_m \) where we define the remainder

\[ R_m(x) = f(x) - P_m(x) \quad \text{for} \quad x \in I. \]

Clearly \( R(c) = 0 \) since \( f(c) = P(c) \).

If \( x \in I, x \neq c \), there exists \( \bar{x} = \bar{x}(x,m) \) (\( \bar{x} \) depends on both \( x \) and \( m \)) between \( c \) and \( x \) such that

\[ R_m(x) = \frac{f^{(m+1)}(\bar{x})}{(m+1)!} (x-c)^{m+1}. \]

Since \( R_m(c) = 0 \), to avoid writing \( x \neq c \) for the formula for \( R_m \) we simply interpret \( \bar{x} \) between \( x \) and \( a \) as \( \bar{x} = a \) in this case. In the actually unlikely event that \( c \) is an endpoint of \( I \) and \( f^{(m+1)}(c) \) is not defined, no real harm is done; we simply interpret undefined \( \times 0 = 0 \).

This theorem allows one to estimate the error made when cutting of the power series of a function at any given term.

3 Power Series Revisited

Polynomials are the simplest of functions. To evaluate a polynomial we only need to know grade school arithmetic; in fact, even less than that. We only need to know how to add, subtract and multiply; we don’t even need to know how to divide. It is thus a matter of great interest to know which functions can be approximated by polynomials; that will give us a way of evaluating them, study their properties. As it turns out, most functions of interest to mathematics and its applications can be approximated by polynomials. In a way, it is what makes computers usable. At the hardware level computers can only do grade school arithmetic or less in base 2; as it turns out that’s all that is really needed (most of the time). The rightly celebrated theorem due to Karl Weierstrass (and generalized by Marshall Stone) states that if \( I \) is a closed and bounded interval in \( \mathbb{R} \) and \( f : I \to \mathbb{R} \) is continuous, then \( f \) can be uniformly approximated by polynomials. In other words, we can find a sequence of polynomials \( \{p_n\} \) such that

\[
\lim_{n \to \infty} \sup_{x \in I} |f(x) - p_n(x)| = 0.
\]

We saw and proved this theorem in part I of the course. Here we will consider functions that can be approximated by a more restricted sequence of polynomials. In the Theorem of Weierstrass, the \( n+1 \)-st polynomial may have
no term equal to the corresponding term of the $n$-th polynomial. All one can be sure of is that if $f$ is not a polynomial, then the degrees of the approximating polynomials have to increase. But some functions are so nice, so smooth, so calm, so happy to please, that they can be approximated by a very special sequence of polynomials, one in which the $n+1$-th term is obtained by adding one extra term to the $n$-th term. That is,

$$p_{n+1}(x) = p_n(x) + a_{n+1}x^{n+1}.$$ 

These are the power series.

**Definition 1** A **power series centered at** $z_0$ is an expression of the form

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots.$$ 

In this definition the center $z_0$ and the coefficients $a_0, a_1, a_2, \ldots$ are fixed, $z$ is a variable. The theory works out best if one assumes that $z_0, a_0, a_1, \ldots$ are complex numbers, $z$ a complex variable, but if that makes you feel uncomfortable just replace the word “complex” by the word real at all occurrences. The partial sums of a power series are polynomials in the variable $z$:

$$s_0(z) = a_0 \quad \text{(a constant polynomial)},$$

$$s_1(z) = a_0 + a_1(z - z_0) = a_1z + (a_0 - a_1z_0) \quad \text{(a polynomial of degree \leq 1)},$$

$$s_2(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2,$$

$$\cdots \quad \cdots.$$ 

So the theory of power series is, in a way, a chapter in the theory of sequences of functions.

Let $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ be a power series. Let $\{s_n(z)\}$ be the corresponding sequence of partial sums, which are polynomials in $z$. Notice that $s_n(z_0) = a_0$ for all $n \in \mathbb{N}_0$, so the sequence of polynomials always converges for $z = z_0$. In other words,

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n = a_0 \quad \text{for } z = z_0.$$ 

We will say that the power series **converges** if the sequence $\{s_n(z)\}$ has a limit for some $z \neq z_0$. Equivalently, if the series converges for some $z \neq z_0$. If

$$D = \{ z \in \mathbb{C} : \sum_{n=0}^{\infty} a_n(z - z_0)^n \text{ converges} \},$$

then we can define a function $f : D \to \mathbb{C}$ by

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n.$$ 

**Examples.**

*The word *term* is used here with two different meanings.*
1. All polynomials are, of course, examples of power series (centered wherever you want to) with domain of convergence $D = \mathbb{C}$. By “centered wherever you want to” I mean the following. Consider, for example, the polynomial $p(z) = z^3 + 2z^2 - 3z + 4$. We can think of it as being the power series $\sum_{n=0}^{\infty} a_n z^n$ centered at 0 with $a_0 = 4$, $a_1 = -3$, $a_2 = 2$, $a_3 = 1$ and $a_n = 0$ if $n \geq 4$. But we can also do some computations:

\[ p(z) = (z - 5)^3 + 2((z - 5) + 5)^2 - 3((z - 5) + 5) + 4 \]

\[ = \left(z - 5\right)^3 + 15(z - 5)^2 + 75(z - 5) + 125 + 2(z - 5)^2 + 20(z - 5) + 50 \]

\[ -3(z - 5) - 15 + 4 \]

\[ = \left(z - 5\right)^3 + 17(z - 5)^2 + 92(z - 5) + 164 \]

which expresses $p(z)$ as a power series $\sum_{n=0}^{\infty} a_n (z - 5)^n$ centered at 5.

2. The geometric series $\sum_{n=0}^{\infty} z^n$ is the simplest example of a power series that is not a polynomial. In many ways, it is the prototypical power series. From what we saw in the series part, the domain of convergence of this series is $D = \{z \in \mathbb{C} : |z| < 1\}$ and the function $f$ it defines in $D$ is $f(z) = 1/(1 - z)$.

3. One can generate some further examples from the previous one. For example, consider the power series

\[ \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} z^{2n} = 1 - \frac{1}{2} z^2 + \frac{1}{3} z^6 - \frac{1}{4} z^8 \pm \ldots \]

This is a series centered at 0, it is of the form $\sum_{n=0}^{\infty} a_n z^n$ with

\[ a_n = \begin{cases} 
(-1)^k / 2^k & \text{if } n = 2k \text{ is even}, \\
0 & \text{if } n = 2k + 1 \text{ is odd}.
\end{cases} \]

But it also is a geometric series of ratio $-z^2/2$ so its domain of convergence is $D = \{z \in \mathbb{C} : |z| < 2\}$ and there it defines the function

\[ f(z) = \frac{1}{1 + \frac{z^2}{2}} = \frac{2}{2 + z^2}. \]

It turns out that the domain of convergence of a power series is always a disc centered at the point where the series is centered. The next theorem, which states this fact, is perhaps the most important general theorem about power series.

**Theorem 13** Let $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ be a power series and define $r \in [0, \infty]$ by

\[ r = \limsup_{n \to \infty} \sqrt[n]{|a_n|}, \tag{4} \]

where we set (or interpret as usual) $1/0 = \infty$, $1/\infty = 0$. 
1. If \( r = 0 \), the series diverges; i.e., it converges only for \( z = z_0 \).

2. If \( 0 < r \leq \infty \), the series converges absolutely for all \( z \) with \(|z - z_0| < r\) and converges uniformly in every closed disc \( \bar{B}(z_0, \rho) \) with \( \rho < r \).

3. If \( r < \infty \), the series diverges for all \( z \) with \(|z - z_0| > r\).

Before going to the proof, some remarks. Obviously, one either has \( r = 0 \) (if \( \limsup_{n \to \infty} \sqrt[|a_n|] = \infty \), or \( r = \infty \) (if \( \limsup_{n \to \infty} \sqrt[|a_n|] = 0 \)) or \( 0 < r < \infty \) (if \( 0 < \limsup_{n \to \infty} \sqrt[|a_n|] < \infty \)). The theorem takes care of all three cases; it is stated for the convenience of a non-repetitious proof. One could also have stated it as follows:

1. If \( r = 0 \) the series diverges for all \( z \neq z_0 \).

2. If \( 0 < r < \infty \), the series converges absolutely for \(|z - z_0| < r\), diverges for \(|z - z_0| > r\). In addition, the series converges uniformly in every closed disc \( \bar{B}(z_0, \rho) \) with \( \rho < r \).

3. If \( r = \infty \) the series converges absolutely for all \( z \in \mathbb{C} \) and converges uniformly in every compact subset of \( \mathbb{C} \).

We say the series converges uniformly in a set \( S \subset \mathbb{C} \) if the sequence of partial sums \( \{s_n\} \) converges uniformly in \( S \). In particular, the series has to converge at all points of \( z \) and then, if \( f \) is the function defined on the domain of convergence \( D \) by the series, for each \( \epsilon > 0 \) there has to exist \( N \) such that \(|f(z) - s_n(z)| < \epsilon\) if \( n \geq N \), \( z \in S \). This could be the moment to recall that a sequence \( \{s_n\} \) of functions defined on the set \( S \) converges uniformly on \( S \) if and only if for every \( \epsilon > 0 \) there exists \( N \) such that \(|s_n(z) - s_m(z)| < \epsilon\) for all \( z \in S \) if \( n, m \geq N \). It is sort of obvious that if a sequence of functions converges uniformly in some set, it also converges uniformly in all of its subsets. Because of this, because every compact subset of \( \bar{B}(z_0, r) \) is contained in some closed ball \( \bar{B}(z_0, \rho) \), and because these closed discs are compact, the following two statements are equivalent:

1. The series converges uniformly in every closed disc \( \bar{B}(z_0, \rho) \) with \( \rho < r \).

2. The series converges in every compact subset of \( B(z_0, r) \) (with \( B(z_0, \infty) \) interpreted as being all of \( \mathbb{C} \)).

Proof. The absolute convergence part of the theorem is an easy consequence of the root test. Let \( z \in \mathbb{C} \). Then we apply the root test to the series \( \sum_{n=0}^{\infty} a_n(z - z_0)^n \). Suppose first \( z \neq z_0 \) so that \(|z - z_0| \neq 0\). Then

\[
\limsup_{n \to \infty} \sqrt[n]{|a_n| |z - z_0|^n} = |z - z_0| \limsup_{n \to \infty} \sqrt[n]{|a_n| |z - z_0|^n} = |z - z_0| \frac{1}{r}
\]

and we see that this limsup is \(< 1\) if and only if \(|z - z_0| < r\); in particular if \(|z - z_0| < r\) then the limsup is \(< 1\) and the series converges absolutely. If \(|z - z_0| > r\) then the limsup is \(> 1\) and the series diverges. We know, of course, what happens if \( z = z_0 \); we have convergence. All that remains
to be proved is the uniform convergence part of the theorem. Assume thus that $0 < \rho < r$. Then $1/\rho > \limsup_{n \to \infty} \sqrt[n]{|a_n|}$ and there is $s$ such that $1/\rho > s > \limsup_{n \to \infty} \sqrt[n]{|a_n|}$. By the second of these inequalities, there is thus $N \in \mathbb{N}_0$ such that $\sqrt[n]{|a_n|} < s$ for $n \geq N$. Then, if $z \in \bar{B}(z_0, \rho)$, $n \geq N$,

$$|a_n| |z - z_0|^n < s^n \rho^n = (s\rho)^n$$

and $s\rho < 1$. If $n, m \geq N$, say $n \geq N$ and $m = n + p$, $p \geq 0$, then

$$|s_m(z) - s_n(z)| \leq \sum_{k=n+1}^{n+p} |a_k| |z - z_0|^n < \sum_{n=n+1}^{n+p} (s\rho)^n < \sum_{n=n+1}^{\infty} (s\rho)^n$$

$$= \frac{(s\rho)^{n+1}}{1 - s\rho}.$$

Since $\lim_{n \to \infty} (s\rho)^{n+1} = 0$, by taking $N$ sufficiently large once an $\epsilon > 0$ is given, one can get that in addition

$$\frac{(s\rho)^{n+1}}{1 - s\rho} < \epsilon$$

if $n \geq N$, proving uniform convergence since this $N$ is independent of $z$ as long as $|z - z_0| \leq \rho$.

**Exercise 1** Let $S$ be a set, for $n \in \mathbb{N}_0$ let $f_n : S \to \mathbb{C}$ be a function, and assume that $|f_n(p)| \leq b_n$ for $n \in \mathbb{N}_0$, all $p \in S$. Prove what is sometimes known as Weierstrass’ M-test:

If $\sum_{n=0}^{\infty} b_n < \infty$, then the series of functions $\sum_{n=0}^{\infty} f_n$ converges uniformly in $S$.

(By uniform convergence of the series we understand that if we define $s_n : S \to \mathbb{C}$ by $s_n(p) = \sum_{k=0}^{n} f_k(p)$ for $n \in \mathbb{N}_0$, then the sequence of functions $\{s_n\}$ converges uniformly in $S$.)

How can one use this theorem to shorten a bit the proof of Theorem 13?

The number $r$ of Theorem 13 is called the *radius of convergence*. Theorem 13 says nothing about what happens on the boundary of the disc of radius $r$, and there is nothing that can be said at a general level; one has to analyze more or less on a case by case basis, in other words, the domain of convergence could be the open disc $B(z_0, r)$, the closed disc $\bar{B}(z_0, r)$, or anything in between.

**Exercise 2** Prove: Let $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ be a power series. Assume

$$\lim_{n \to \infty} \frac{|a_n|}{|a_{n+1}|}$$

exists. Prove that in this case the radius of convergence of the power series satisfies:

$$r = \lim_{n \to \infty} \frac{|a_n|}{|a_{n+1}|}$$
This frequently provides one with an easier way of computing the radius of convergence.

**Examples.**

1. Consider the four power series:

\[
\sum_{n=1}^{\infty} nz^n, \sum_{n=1}^{\infty} z^n, \sum_{n=1}^{\infty} \frac{1}{n} z^n, \sum_{n=1}^{\infty} \frac{1}{n^2} z^n.
\]

All four have radius 1. Generally speaking, any series of the form

\[
\sum_{n=1}^{\infty} n^\alpha z^n
\]

has radius 1 because

\[
\limsup_n \sqrt[n]{n^\alpha} = \left( \lim_{n \to \infty} n^{1/n} \right)^\alpha = 1^\alpha = 1,
\]

and \(1/1 = 1\). All four converge for \(|z| < 1\), diverge for \(|z| > 1\). However their behavior for \(|z| = 1\) is not the same. The series

\[
\sum_{n=1}^{\infty} nz^n, \sum_{n=1}^{\infty} z^n
\]

diverge for \(|z| = 1\). Of the series \(\sum_{n=1}^{\infty} (1/n) z^n\) we know it diverges for \(z = 1\) because there it is the harmonic series; it converges (conditionally) for \(z = -1\) there it is the alternating harmonic series. We don’t yet what happens for other \(z\) of absolute value 1 (we may or may not see this: one can prove that the series converges for all \(z\) with \(|z| = 1\) except for \(z = 1\)). Finally, the series \(\sum_{n=1}^{\infty} (1/n^2) z^n\) converges absolutely also for all \(z\) with \(|z| = 1\).

2. Consider the power series \(\sum_{n=0}^{\infty} (-7)^n (z + 2)^n\). What is its disc of convergence? If we apply the formula we get

\[
1/r = \lim_{n \to \infty} \sqrt[n]{|(-7)^n|} = 7,
\]

thus \(r = 1/7\) and the domain of convergence is (at least) the disc \(B(-2, 1/7)\). But do we really need this? It is easier to see that we have a geometric series of ratio \(-7(z + 2)\) thus it converges exactly for \(7|z + 2| < 1\); in other words in the open disc of center \(-2\), radius 1/7. So the radius is 1/7, domain of convergence the open disc of radius 1/7, center \(-2\).

3. The series

\[
\sum_{n=1}^{\infty} \frac{1}{n!} z^n = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \cdots
\]

is quite important. Its radius is easier to compute using the last exercise.

\[
r = \lim_{n \to \infty} \frac{1/n!}{1/(n+1)!} = \lim_{n \to \infty} (n + 1) = \infty.
\]

The series converges for all values of \(z\).
Exercise 3 This is an exercise from the last qualifier: Let \( a_n = n\)-th digit in the decimal expansion of \( \pi + 1 \); thus
\[
a_0 = 4, a_1 = 2, a_2 = 5, a_3 = 6, a_4 = 10, \text{etc.}
\]
Prove: The series \( \sum_{n=0}^{\infty} a_n z^n \) has radius of convergence 1.

Exercise 4 If someone were to sell you a power series of the form \( \sum_{n=0}^{\infty} a_n z^n \) and told you it converges for \( z = 1 + i \) but diverges for \( z = -1 \), should you buy it? Why?

4 A brief foray into the exciting realm of analytic functions

Suppose \( U \) is an open subset of \( \mathbb{C} \) and \( f : U \to \mathbb{C} \). Let \( z_0 \in U \). The complex derivative of \( f \) at \( z_0 \) is defined by
\[
f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}
\]
if this limit exists.

A function that is complex differentiable at ALL points of an open subset of \( \mathbb{C} \) is said to be analytic on that set.

I am just scratching the surface of this topic, I don’t want to get into too many details. But here we have some obvious facts:

1. Suppose \( f(z) = z^n \), \( n \) a non-negative integer. Then we have the factorization
\[
z^n - z_0^n = (z - z_0)(z^{n-1} + z^{n-2}z_0 + \cdots + z_0^{n-2} + z_0^{n-1}).
\]
Dividing by \( z - z_0 \) we see that we are left on the right hand side with a sum of \( n \) terms, each one converging to \( z_0^{n-1} \) as \( z \to z_0 \). Thus \( f \) is complex differentiable at \( z_0 \) and \( f'(z_0) = nz_0^{n-1} \). Or, since \( z_0 \) was arbitrary, we can say that \( f \) is complex differentiable at all \( z \in \mathbb{C} \) and \( f'(z) = nz^{n-1} \).

In particular, if \( f(z) = 1 \) for all \( z \), then \( f'(z) = 0 \) for all \( z \); if \( f(z) = z \) for all \( z \), then \( f'(z) = 1 \) for all \( z \).

2. Suppose \( f = u + iv : I \to \mathbb{C} \), where \( I \) is an interval in \( \mathbb{R} \) and \( u, v : I \to \mathbb{R} \). Suppose that \( f \) can be extended to be defined on an open subset \( U \) of \( \mathbb{C} \), \( I \subset U \) in such a way that \( f \) is complex differentiable at \( x_0 \). Most functions we meet in calculus can be so extended, though it isn’t always obvious at that level what is a nice extension. But \( f(x) = x^n \) on \( \mathbb{R} \) extends to \( f(z) = z^n \) on \( \mathbb{C} \).
Then \( u, v : I \to \mathbb{R} \) are differentiable (in the usual sense at \( x_0 \)) and \( f'(x_0) = u'(x_0) + iv'(x_0) \), the LHS being the complex derivative, the RHS the usual derivatives.

In other words, if \( f : I \to \mathbb{C} \) is the restriction of a complex differentiable function to the interval \( I \subset \mathbb{R} \), then \( f \) is differentiable in the usual sense and the derivatives coincide. This should be clear; it is simply part of the following general fact about functions in metric spaces:

If \( X, Y \) are metric spaces, if \( E \subset X \), if \( p \) is a cluster point of \( E \) if \( g : X \setminus \{p\} \to Y \), if \( \lim_{q \to p} g(q) = L \) exists, then

\[
\lim_{q \to p, q \in E} (g|_E)(q) \text{ exists and equals } L.
\]

3. Consider the function \( f : \mathbb{R} \to \mathbb{R} \) defined by \( f(x) = x^2 \). There are many ways to extend it as a nice function to all of \( \mathbb{C} \). Here are two: Define \( f : \mathbb{C} \to \mathbb{C} \) by \( f(z) = z^2 \). As mentioned above, this extension is complex differentiable, thus analytic, on all of \( \mathbb{C} \). The derivative is \( f'(z) = 2z \) which, when restricted to real values of \( z \) gives the usual \( f'(x) = 2x \).

But we can also extend \( f \) to all of \( \mathbb{C} \) by \( f(z) = |z|^2 \). If \( z_0 \in \mathbb{C} \), then (with this definition)

\[
\frac{f(z) - f(z_0)}{z - z_0} = \frac{|z|^2 - |z_0|^2}{z - z_0} = \frac{x^2 + y^2 - x_0^2 - y_0^2}{(x - x_0) + i(y - y_0)}
\]

where \( z = x + iy, z_0 = x_0 + iy_0 \). Is there a limit for \( z \to z_0 \)? Suppose we restrict the difference quotient to a horizontal line through \( z_0 \); this amounts to restricting \( y = y_0 \). Then

\[
\frac{|z|^2 - |z_0|^2}{z - z_0} = \frac{x^2 - x_0^2}{x - x_0} = x + x_0 \to 2x_0
\]

as \( z \to z_0 \). So far, so good. Now we can try to see what happens if \( z \) approaches \( z_0 \) vertically, with \( x = x_0 \). Then we get

\[
\frac{|z|^2 - |z_0|^2}{z - z_0} = \frac{y^2 - y_0^2}{i(y - y_0)} = -i(y + y_0) \to -2iy_0.
\]

And we can ask when is \( 2x_0 = -2iy_0 \)? The answer is hardly ever, if and only if \( x_0 = y_0 = 0 \); i.e., \( z_0 = 0 \). This does not prove complex differentiability at 0, however it is easy to see that if \( z_0 = 0 \) the limit exists and is, in fact, 0. The conclusion is that this extension of \( f(x) = x^2 \) is complex differentiable ONLY at 0.

The reason I introduced all this terminology is so I can state the next theorem in a nice way (but “nice” can be in the eye of the beholder).
Theorem 14 Let \( \sum_{n=0}^{\infty} a_n(z-z_0)^n \) be a power series with radius of convergence \( r > 0 \). Let
\[
f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n
\]
for \( z \in D = B(z_0, r) \). Then \( f \) is analytic in \( D \) and
\[
f'(z) = \sum_{n=0}^{\infty} na_n(z-z_0)^{n-1} = \sum_{n=0}^{\infty} na_n(z-z_0)^{n-1} = \sum_{n=0}^{\infty} (n+1)a_{n+1}(z-z_0)^n
\]
for \( z \in D \). The power series of \( f' \) has radius of convergence \( r \).

Proof. The proof involves some messy computations, but that’s life. First of all, let us prove the last statement. Consider the power series
\[
\sum_{n=0}^{\infty} (n+1)a_{n+1}(z-z_0)^n.
\]
It converges, of course, for \( z = z_0 \). For any other value of \( z \); that is, for \( z \neq z_0 \), we can write
\[
\sum_{n=0}^{\infty} (n+1)a_{n+1}(z-z_0)^n = \frac{1}{z-z_0} \sum_{n=0}^{\infty} (n+1)a_{n+1}(z-z_0)^n+1 = \frac{1}{z-z_0} \sum_{n=0}^{\infty} na_n(z-z_0)^n
\]
and what this shows is that the series converges exactly where the series with coefficients \( na_n \) converges. The radius of the latter is given by
\[
\limsup_{n \to \infty} \frac{1}{n^{1/n} |a_n|^{1/n}} = \limsup_{n \to \infty} \frac{1}{|a_n|^{1/n}} = r,
\]
proving the last assertion of the theorem. We can thus define
\[
g(z) = \sum_{n=1}^{\infty} na_n(z-z_0)^{n-1}
\]
for \( z \in D \). The proof now reduces to proving
\[
\lim_{z \to w} \frac{f(z) - f(w)}{z-w} = g(w)
\]
for all \( w \in D \). So let \( w \in D \). While somewhat messy, there are some obvious things to do. The result is sort of trivial if instead of having a series we had a finite sum, so what we need to do is to see that the higher order terms of the series can be controlled well. Noticing that the term \( a_0 \) cancels when subtracting \( f(w) \) from \( f(z) \), we see that
\[
\frac{f(z) - f(w)}{z-w} - g(w) = \sum_{n=1}^{\infty} a_n \left( \frac{z^n - w^n}{z-w} - nw^{n-1} \right)
\]
for all \( z \in D \). Now
\[
\frac{z^n - w^n}{z - w} - nw^{n-1} = (z^{n-1} + z^{n-2}w + \cdots + w^{n-1}) - nw^{n-1}
\]
and we need to estimate this expression in terms of \( z - w \). Let \( \rho = r - |w - z_0| \); so \( 0 < \rho < r \). We will assume, as we may that \( |z - w| < \rho/2 \). We use again that for \( k \in \mathbb{N}_0 \)
\[
|z^k - w^k| = |z - w||z^{k-1}w + \cdots + w^{k-1}|
\]