Power Series

Let us talk about analytic functions. That is functions of a complex z that have a power series,

$$f(z) = \sum_{k=0}^{\infty} c_k z^k,$$

or perhaps for some fixed $a \in \mathbb{C}$

$$\sum_{k=0}^{\infty} c_k (z-a)^k.$$

We will sometimes let z be complex and sometimes we will talk about real power series, in which case we will use x. I will always mention which case we are working with.

An analytic function can have a different expansion around different points. Also the convergence does not automatically happen on the entire domain of the function. For example, we know that if $|z| < 1$, then

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

While the left hand side exists on all of $z \neq 1$, the right hand side happens to converge only if $|z| < 1$.

Before we delve into power series, let us prove the root test which we haven't done last semester.

Theorem: (Root test) Let $\sum a_k$ be a series of complex numbers and write

$$\alpha = \limsup_{k \to \infty} \sqrt[k]{|a_k|}$$

Then if $\alpha < 1$, the series converges absolutely. If $\alpha > 1$ then the series diverges.

The test is similar to the ratio test (we have used limit rather than limsup for the ratio test before, but the proof with limsup is also not difficult). If $\alpha = 1$ we get no information.

Proof. Obviously $\alpha \geq 0$. Suppose that $\alpha < 1$, then there must be some β with $\alpha < \beta < 1$. Since $\beta - \alpha > 0$, then by definition of limsup, there must exist some N such that for all $n \geq N$ we have

$$\sqrt[n]{|a_n|} < \beta.$$

In other words $|a_n| < \beta^n$ and $\sum \beta^n$ converges as $0 < \beta < 1$. A series converges if and only if its tail converges and we have by comparison test that

$$\sum_{k=N}^{\infty} |a_n|$$

converges. Therefore $\sum |a_n|$ converges and so $\sum a_n$ converges absolutely.

Next suppose that $\alpha > 1$. We note that this means that for infinitely many n we have $\sqrt[n]{|a_n|} > 1$ and hence $|a_n| > 1$. As for a convergent series we have to have $a_n \to 0$, the series $\sum a_n$ cannot converge. $\square$

Using the root test we can find the so-called *radius of convergence* of a power series.

Theorem 3.39 + (part of 8.1): Let

$$\sum c_k (z-a)^k$$

be a power series. Let

$$\alpha = \limsup_{k \to \infty} \sqrt[k]{|c_k|}$$

If $\alpha = 0$ let $R = \infty$, if $\alpha = \infty$ let $R = 0$, and otherwise let $R = 1/\alpha$. Then the power series converges absolutely if $|z - a| < R$ and diverges when $|z - a| > R$.

Furthermore, if $R > 0$ then the series converges uniformly on $B(a, r)$ for any positive $r < R$.

Proof. Write

$$\limsup_{k\to\infty} \sqrt[k]{\left|c_k(z-a)^k\right|} = |z-a| \limsup_{k\to\infty} \sqrt[k]{|c_k|} = \alpha\,|z-a|$$

Note that this calculation makes sense even if $\alpha = \infty$ when when we decree that $\infty x = \infty$ if $x > 0$. We assume that $|z-a| \neq 0$. Since the series of course always converges if $z = a$.

Now the series converges absolutely if $\alpha\,|z-a| < 1$ and diverges if $\alpha\,|z-a| > 1$. We notice that this is precisely the conclusion of the theorem.

For the "Furthermore" part of the theorem, suppose that $R > 0$ and pick a positive $r < R$. Then if $z \in B(a,r)$ then $|z-a| < r$ then

$$\left|c_k(z-a)^k\right| \leq |c_k|\,r^k$$

Notice that since the series converges absolutely at any point in $B(a,R)$ the series $\sum |c_k|\,r^k$ must converge, it is therefore Cauchy. As for any $m < n$ we have

$$\sum_{k=m+1}^{n} \left|c_k(z-a)^k\right| \leq \sum_{k=m+1}^{n} |c_k|\,r^k$$

the original series is uniformly Cauchy on $B(a,r)$ and hence uniformly convergent on $B(a,r)$. □

The number R is called the *radius of convergence*. It gives us a disk around a where the series converges. We say that the series is *convergent* if $R > 0$, in other words, if the series converges for some point not equal to a.

Note that it is trivial to see that if $\sum c_k(z-a)^k$ converges for some z, then

$$\sum c_k(w-a)^k$$

must converge absolutely whenever $|w-a| < |z-a|$. This follows from the computation of the radius of convergence. Conversely if the series diverges at z, then it must diverge at w whenever $|w-a| > |z-a|$.

This means that to show that the radius of convergence is at least some number, we simply need to show convergence at some point by any method we know.

Examples:

$$\sum_{k=0}^{\infty} z^k$$

has radius of convergence 1.

$$\sum_{k=0}^{\infty} \frac{1}{n^n} z^k$$

has radius of convergence ∞. Similarly

$$\sum_{k=0}^{\infty} \frac{1}{n!} z^k$$

has radius of convergence ∞. Although in this case it is easier to apply the ratio test to the series to note that the series converges absolutely at all $z \in \mathbb{C}$.

On the other hand,

$$\sum_{k=0}^{\infty} n^n z^k$$

has radius of convergence 0, so it converges only if $z = 0$.

Do note the difference between $\frac{1}{1-z}$ and its power series. Let us expand $\frac{1}{1-z}$ as power series around any point $a \neq 1$. Let $c = \frac{1}{1-a}$, then we can write

$$\frac{1}{1-z} = \frac{c}{1-c(z-a)} = c\sum_{k=0}^{\infty} c^k (z-a)^k = \sum_{k=0}^{\infty} \left(\frac{1}{(1-a)^{k+1}} \right) (z-a)^k$$

Then we notice that $\sum c^k (z-a)^k$ converges if and only if the power series on the right hand side converges and

$$\limsup_{k\to\infty} \sqrt[k]{c^k} = c = \frac{1}{|1-a|}$$

So radius of convergence of the power series is $|1-a|$, that is the distance of a from 1. In particular the function has a power series representation around every a not equal to 1. This is what we usually call analytic. Notice that the domain of the function is bigger than the region of convergence of any power series representing the function at any point.

It turns out that if a function has a power series representation converging to the function on some ball, then it has a representation at every point in the ball. We will prove this result later.

Corollary: If

$$f(z) = \sum c_k (z-a)^k$$

is convergent in $B(a,R)$ for some $R > 0$, then $f\colon B(a,R) \to \mathbb{C}$ is continuous.

Proof. For any $z_0 \in B(a,R)$ pick $r < R$ such that $z_0 \in B(a,r)$. If we show that f restricted to $B(a,r)$ is continuous then f is continuous at z_0. This can be easily seen since for example any sequence converging to z_0 will have some tail that is completely in the open ball $B(a,r)$. On $B(a,r)$ the partial sums converge uniformly by the theorem and so the limit is continuous. □

Note that we can always apply an affine transformation $z \mapsto z + a$ which would convert the series to a series at the origin. Therefore it is usually sufficient to just prove results about power series at the origin.

Let us look at derivatives. We will prove the following only for power series of real variable x. We will allow coefficients to be complex valued, but we will only consider the power series on the real axis. Therefore we will consider x a real variable and we will consider convergence an interval $(-R,R)$.

Theorem (rest of 8.1): Let

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

be a power series converging in $(-R,R)$ for some $R > 0$. Then f is differentiable on $(-R,R)$ and

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k.$$

Proof. First notice that

$$\lim_{k\to\infty} \sqrt[k]{k} = 1$$

(exercise LTS). Therefore

$$\limsup_{k\to\infty} \sqrt[k]{|k a_k|} = \limsup_{k\to\infty} \sqrt[k]{k}\,\sqrt[k]{|a_k|} = \limsup_{k\to\infty} \sqrt[k]{|a_k|}.$$

So $\sum_{k=1}^{\infty} k a_k x^{k-1}$ has the same radius of convergence as the series for f.

For any positive $r < R$, we have that both series converge uniformly in $[-r,r]$ by a theorem above. Now for any partial sum

$$\frac{d}{dx}\left(\sum_{k=0}^{n} a_k x^k \right) = \sum_{k=1}^{n} k a_k x^{k-1}$$

So by Theorem 7.17 $f'(x)$ is equal to the limit of the differentiated series as advertised on $[-r, r]$. As this was true for any $r < R$, we have it on $(-R, R)$. $\square$

In fact, if one derivative then by iterating the theorem we obtain that an analytic function is infinitely differentiable.

Corollary: Let

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

be a power series converging in $(-R, R)$ for some $R > 0$. Then f is infinitely differentiable on $(-R, R)$ and

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1)a_k x^{k-n} = \sum_{k=0}^{\infty}(k+n)(k+n-1)\cdots(k+1)a_{k+n}x^k.$$

In particular,

$$f^{(n)}(0) = n!a_n.$$

Note that the coefficients are determined by the derivatives of the function. This in particular means that once we have a function defined in a neighborhood, the coefficients are unique. Also, if we have two power series convergent in $(-R, R)$ such that for all $x \in (-R, R)$ we have

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} b_k x^k$$

then $a_k = b_k$ for all k.

On the other hand, just because we have an infinitely differentiable function doesn't mean that the numbers c_k obtained by $c_n = \frac{f^{(n)}(0)}{n!}$ give a convergent power series. In fact, there is a theorem, which we will not prove, that given an arbitrary sequence $\{c_n\}$, there exists an infinitely differentiable function f such that $c_n = \frac{f^{(n)}(0)}{n!}$. Finally, even if the obtained series converges it may not converge to the function we started with. This counterexample will be in the homework.

(We will skip Theorem 8.2)

We will need a theorem on swapping limits of series. This is sometimes called Fubini's theorem for sums.

Theorem 8.3: Let $\{a_{ij}\}_{i=1,j=1}^{\infty}$ be a double sequence of complex numbers and suppose that for every i the series

$$\sum_{j=1}^{\infty} |a_{ij}|$$

converges and furthermore that

$$\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} |a_{ij}| \right)$$

converges as well. Then

$$\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} a_{ij} \right) = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} a_{ij} \right),$$

where all the series involved converge.

Rudin has a very slick proof of this.

Proof. Let E be the set $\{1/n : n \in \mathbb{N}\} \cup \{0\}$, and treat it as a metric space with the metric inherited from $\mathbb{R}$. Define the sequence of functions $f_i \colon E \to \mathbb{C}$ by

$$f_i(1/n) = \sum_{j=1}^{n} a_{ij} \qquad f_i(0) = \sum_{j=1}^{\infty} a_{ij}$$

As the series converge we get that each f_i is continuous at 0 (since 0 is the only cluster point, they are continuous everywhere, but we don't need that). For all $x \in E$ we have

$$|f_i(x)| \leq \sum_{j=1}^{\infty} |a_{ij}|$$

By knowing that $\sum_i \sum_j |a_{ij}|$ converges (does not depend on x), we know that for any $x \in E$

$$\sum_{i=1}^{n} f_i(x)$$

converges uniformly. So define

$$g(x) = \sum_{i=1}^{\infty} f_i(x),$$

which is therefore a continuous function at 0. So

$$\begin{aligned}\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} a_{ij}\right) &= \sum_{i=1}^{\infty} f_i(0) = g(0) = \lim_{n\to\infty} g(1/n) \\ &= \lim_{n\to\infty} \sum_{i=1}^{\infty} f_i(1/n) = \lim_{n\to\infty} \sum_{i=1}^{\infty}\sum_{j=1}^{n} a_{ij} \\ &= \lim_{n\to\infty} \sum_{j=1}^{n}\sum_{i=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} a_{ij}\right).\end{aligned}$$

□

Now we can prove that once we have a series converging to a function in some interval, we can expand the function around any point.

Theorem 8.4: (Taylor's theorem for real-analytic functions) Let

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

be a power series converging in $(-R, R)$ for some $R > 0$. Given any $a \in (-R, R)$, we obtain for x such that $|x - a| < R - |a|$ that

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

The power series at a could of course converge in a larger interval, but the one above is guaranteed. It is the largest symmetric interval about a that fits in $(-R, R)$.

Proof. Write

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} a_k\big((x-a)+a\big)^k \\ &= \sum_{k=0}^{\infty} a_k \sum_{m=0}^{k} \binom{k}{m} a^{k-m}(x-a)^m \end{aligned}$$

We define $c_{k,m} = a_k\binom{k}{m}a^{k-m}$ if $m \leq k$ and 0 if $m > k$, then we have

$$f(x) = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} c_{k,m}(x-a)^m \tag{1}$$

If we show that the double sum converges absolutely as in Theorem 8.3 we are done. We can swap the order of the summations and obtain the desired coefficients.

$$\begin{aligned}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\left|c_{k,m}(x-a)^m\right| &= \sum_{k=0}^{\infty}\sum_{m=0}^{k}\left|a_k\binom{k}{m}a^{k-m}(x-a)^m\right| \\ &= \sum_{k=0}^{\infty}|a_k|\sum_{m=0}^{k}\binom{k}{m}|a|^{k-m}|x-a|^m \\ &= \sum_{k=0}^{\infty}|a_k|\left(|x-a|+|a|\right)^k\end{aligned}$$

and this converges as long as $(|x-a|+|a|)<R$ or in other words if $|x-a|<R-|a|$.

Now we swap the order of summation in (1), and the following converges when $|x-a|<R-|a|$:

$$f(x)=\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}c_{k,m}(x-a)^m=\sum_{m=0}^{\infty}\left(\sum_{k=0}^{\infty}c_{k,m}\right)(x-a)^m.$$

And we are done. The formula in terms of derivatives at a follows by the corollary of Theorem 8.1. □

Note that if a series converges for $x\in(-R,R)$ it also converges for all complex numbers in $B(0,R)$. We have the following corollary.

Corollary: If $\sum c_k(z-a)^k$ converges to $f(z)$ in $B(a,R)$ and $b\in B(a,R)$, then there exists a power series $\sum d_k(z-b)^k$ that converges to $f(z)$ in $B(b,R-|b-a|)$.

Proof. WLOG assume that $a=0$. We can also rotate to assume that b is real, but since that is harder to picture, let us do it explicitly. Let $\alpha=\frac{\bar{b}}{|b|}$. Notice that

$$\left|{}^1\!/\!_\alpha\right|=|\alpha|=1.$$

Therefore the series converges in $B(0,R)$ if we replace z with ${}^z\!/\!_\alpha$. We apply Theorem 8.4 at $|b|$ and get that a series that converges to $f({}^z\!/\!_\alpha)$ on $B(|b|,R-|b|)$. That is, there are some coefficients a_k such that

$$f({}^z\!/\!_\alpha)=\sum_{k=0}^{\infty}a_k(z-|b|)^k.$$

Notice that $\alpha b=|b|$.

$$f(z)=f({}^{\alpha z}\!/\!_\alpha)=\sum_{k=0}^{\infty}a_k(\alpha z-|b|)^k=\sum_{k=0}^{\infty}a_k\alpha^k(z-{}^{|b|}\!/\!_\alpha)^k=\sum_{k=0}^{\infty}a_k\alpha^k(z-b)^k,$$

and this converges for all z such that $\left|\alpha z-|b|\right|<R-|b|$ or $|z-b|<R-|b|$, which is the conclusion of the theorem. □

Let us define rigorously what an analytic function is. Let $U\subset\mathbb{C}$ be an open set. A function $f\colon U\to\mathbb{C}$ is said to be *analytic* if near every point $a\in U$, f has a convergent power series expansion. That is, for every $a\in U$ if there exists an $R>0$ and numbers c_k such that

$$f(z)=\sum_{k=0}^{\infty}c_k(z-a)^k$$

for all $z\in B(a,R)$.

Similarly if we have an interval (a,b), we will say that $f\colon(a,b)\to\mathbb{C}$ is analytic or perhaps *real-analytic* if for each point $c\in(a,b)$ there is a power series around c that converges in some $(c-R,c+R)$ for some $R>0$.

We have proved above that a convergent power series is an analytic function where it converges. We have also shown before that $\frac{1}{1-z}$ is analytic outside of $z=1$.

Note that just because a real analytic function is analytic on the real line it doesn't necessarily mean that it has a power series representation that converges everywhere. For example, the function

$$f(x) = \frac{1}{1+x^2}$$

happens to be real analytic. It is pretty easy to show this (exercise). A power series around the origin converging to f can have a radius of convergence at most 1, actually it does happen to be exactly 1. Can you see why? (exercise).

Lemma (part of 8.5): Suppose that $f(z) = \sum a_k z^k$ is a convergent power series and $\{z_n\}$ is a sequence of nonzero complex numbers converging to 0, such that $f(z_n) = 0$ for all n. Then $a_k = 0$ for every k.

Proof. By continuity we know $f(0) = 0$ so $a_0 = 0$. Suppose that there exists some nonzero a_k. Let m be the smallest m such that $a_m \neq 0$. Then

$$f(z) = \sum_{k=m}^{\infty} a_k z^k = z^m \sum_{k=m}^{\infty} a_k z^{k-m} = z^m \sum_{k=0}^{\infty} a_{k+m} z^k.$$

Write $g(z) = \sum_{k=0}^{\infty} a_{k+m} z^k$ (this series converges in on the same set as f). g is continuous and $g(0) = a_m \neq 0$. Thus there exists some $\delta > 0$ such that $g(z) \neq 0$ for all $z \in B(0,\delta)$. As $f(z) = z^m g(z)$, then the only point in $B(0,\delta)$ where $f(z) = 0$ is when $z = 0$, but this contradicts the assumption that $f(z_n) = 0$ for all n. □

In a metric space X, a *cluster point* (or sometimes *limit point*) of a set E is a point $p \in X$ such that $B(p,\epsilon) \setminus \{p\}$ contains points of E for all $\epsilon > 0$.

Theorem (better than 8.5): (Identity theorem) Let $U \subset \mathbb{C}$ be an open connected set. If $f\colon U \to \mathbb{C}$ and $g\colon U \to \mathbb{C}$ are analytic functions that are equal on a set $E \subset U$, and E has a cluster point (a limit point) in U, then $f(z) = g(z)$ for all $z \in U$.

Proof. WLOG suppose that E is the set of all points $z \in U$ such that $g(z) = f(z)$. Note that E must be closed as f and g are continuous.

Suppose that E has a cluster point. WLOG assume that 0 is the cluster point. Near 0, we have the expansions

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \qquad \text{and} \qquad g(z) = \sum_{k=0}^{\infty} b_k z^k,$$

which converge in some ball $B(0,R)$. Therefore the series

$$0 = f(z) - g(z) = \sum_{k=0}^{\infty} (a_k - b_k) z^k$$

converges in $B(0,R)$. As 0 is a cluster point of E, then there is a sequence of nonzero points $\{z_n\}$ such that $f(z_n) - g(z_n) = 0$. Therefore by the lemma above we have that $a_k = b_k$ for all k. And therefore $B(0,R) \subset E$.

This means that E is open. As E is also closed, and U is connected, we conclude that $E = U$. □

By restricting our attention to real x we obtain the same theorem for connected open subsets of $\mathbb{R}$, which are just open intervals. Rudin's Theorem 8.5 is essentially this real analogue: if we have two power series around a point that converge on (a,b) and are equal on a set that has a cluster point in (a,b), then the two power series are the same.

The Exponential

Define

$$E(z) = \sum_{k=0}^{\infty} \frac{1}{k!} z^k.$$

This will turn out to be the exponential, but let us call it something else for now. We have shown that this series converges everywhere. We notice that $E(0) = 1$.

By direct calculation we notice that if we restrict to real x we get

$$\frac{d}{dx}\big(E(x)\big) = E(x).$$

This is one of the most important properties of the exponential. In fact, this was the way that we hinted at a possible definition of the exponential last semester using Picard's theorem. Of course by uniqueness part of Picard's theorem, either starting point leads to the same exponential function.

Fix $y \in \mathbb{R}$. Let $f(x) = E(x+y) - E(x)E(y)$ then compute

$$f'(x) = \frac{d}{dx}\big(E(x+y) - E(x)E(y)\big) = E(x+y) - E(x)E(y) = f(x)$$

Furthermore $f(0) = E(y) - E(0)E(y) = 0$. To conclude that $f \equiv 0$ we could at this point apply Picards theorem on existence and uniqueness, but that would be hitting the nail with a sledgehammer. Far easier is to notice that f has a power series at the origin that converges everywhere. And furthermore notice by repeated application of $f'(x) = f(x)$ that $f^{(n)}(x) = f(x)$, in particular $f^{(n)}(0) = f(0) = 0$. That means that f is the series that is identically zero. Therefore we have

$$E(x+y) = E(x)E(y)$$

for all real x and y. Now for any fixed $y \in \mathbb{R}$, we get by the identity theorem (since the real numbers have a cluster point in $\mathbb{C}$) that $E(z+y) = E(z)E(y)$ for all $z \in \mathbb{C}$. Now fixing an arbitrary $z \in \mathbb{C}$ and we get $E(z+y) = E(z)E(y)$ for all $y \in \mathbb{R}$, and hence again by identity theorem we obtain

$$E(z+w) = E(z)E(w)$$

for all $z, w \in \mathbb{C}$.

In particular if $w = -z$ we obtain $1 = E(0) = E(z-z) = E(z)E(-z)$. This implies that $E(z) \neq 0$ for any $z \in \mathbb{C}$.

Let us look at the real numbers. Since E is continuous and never zero, and $E(0) = 1 > 0$, we get that $E(\mathbb{R})$ is connected and $1 \in E(\mathbb{R})$ but $0 \notin E(\mathbb{R})$. Therefore $E(x) > 0$ for all $x \in \mathbb{R}$. Thus $E'(x) > 0$ for all real x and so $E(x)$ is strictly increasing. This means that $E(1) > E(0) = 1$. Then for $n \in \mathbb{N}$

$$E(n) = \big(E(1)\big)^n$$

by the addition formula. As $E(1) > 1$ we conclude that $E(n)$ goes to ∞ as n goes to ∞. Since E is monotone that means that we can even take the continuous limit as x goes to ∞ and get that $E(x)$ goes to ∞. By $E(x)E(-x) = 1$ we obtain that as x goes to $-\infty$ then $E(x)$ must go to zero. Therefore E, when restricted to $\mathbb{R}$, is a one to one and onto mapping of $\mathbb{R}$ to $(0, \infty)$.

Usually we denote

$$e = E(1).$$

As for $n \in \mathbb{N}$ we get that $E(n) = \big(E(1)\big)^n$ we get

$$e^n = E(n).$$

By noting that $E(n)E(-n) = 1$ we obtain $e^n = E(n)$ for all $n \in \mathbb{Z}$. In fact if $r = {}^p/_q$ is rational, then

$$E(r)^q = E(qr) = E(p) = e^p.$$

And hence $E(r) = e^r$ given the usual meaning of e^r (the qth positive root of e^p).

Generally for irrational $x \in \mathbb{R}$ we define

$$e^x = E(x),$$

and we notice that this is the same thing as taking limits of e^r for rational r going to x as $E(x)$ is continuous.

Similarly we simply define that for all $z \in \mathbb{C}$

$$e^z = E(z)$$

and we usually denote $E(z)$ by $\exp(z)$ if we don't use e^z.

We have proved most of the following theorem.

Theorem 8.6: Let us consider e^x a function on the real line only.

(i) e^x is continuous and differentiable everywhere.
(ii) $(e^x)' = e^x$.
(iii) e^x is strictly increasing function of x and $e^x > 0$ for all x.
(iv) $e^{x+y} = e^x e^y$.
(v) $e^x \to \infty$ as $x \to \infty$ and $e^x \to 0$ as $x \to -\infty$. In particular, e^x is a one to one and onto function from $\mathbb{R}$ to $(0, \infty)$.
(vi) $\lim\limits_{x\to\infty} x^n e^{-x} = \lim\limits_{x\to\infty} \dfrac{x^n}{e^x} = 0$ for every $n \in \mathbb{N}$.

Proof. We have proved all these properties except the last one. So let us prove that. From definition it is clear that if $x > 0$ then e^x is bigger than any one of the terms in the series so for example

$$e^x > \frac{x^{n+1}}{(n+1)!}.$$

In other words since $\frac{1}{e^x} = e^{-x}$ we have

$$x^n e^{-x} < \frac{(n+1)!}{x},$$

which proves the claim. □

We can now define the logarithm at least for positive real x. We could define it as the inverse of the exponential, but then we would need to prove some of its properties. It is easier to define $\log\colon (0,\infty) \to \mathbb{R}$ as

$$\log(x) = \int_1^x \frac{1}{t}\, dt.$$

Obviously log is continuous and differentiable with derivative $\frac{1}{x}$ and $\log(1) = 0$. Let us show that it is the inverse. Let $f(x) = \log(e^x)$. Next

$$f'(x) = \frac{d}{dx}\bigl(\log(e^x)\bigr) = \frac{1}{e^x} e^x = 1.$$

Furthermore $f(0) = \log(e^0) = \log(1) = 0$. By the fundamental theorem of calculus we obtain that $f(x) = \int_0^x dt = x$ and hence $\log(e^x) = x$. In other words log and the exponential are inverses of each other. It follows that $\exp\bigl(\log(x)\bigr) = x$.

Let us deduce some properties of the logarithm from the properties of the exponential by using the fact that log is the inverse of the exponential.

From Theorem 8.6 it is clear that $\log(x) \to -\infty$ as $x \to 0$ and $\log(x) \to \infty$ as $x \to \infty$.

Next, suppose that $a = e^x$ and $b = e^y$, then we obtain the addition formula

$$\log(ab) = \log(e^x e^y) = \log(e^{x+y}) = x + y = \log(a) + \log(b).$$

In particular, if $b = 1/a$ then

$$0 = \log(1) = \log(a) + \log(1/a)$$

Putting these two together we obtain that for all $n \in \mathbb{Z}$ and $x > 0$ we get

$$\log(x^n) = n \log(x)$$

Next notice that

$$x^n = \exp\bigl(\log(x^n)\bigr) = \exp\bigl(n \log(x)\bigr).$$

Then

$$\left(\exp\left(\frac{1}{m}\log(x)\right)\right)^m = \exp\left(\frac{m}{m}\log(x)\right) = x$$

or in other words $x^{1/m} = \exp\bigl(\frac{1}{m}\log(x)\bigr)$. So if p is rational and $x > 0$ we obtain that

$$x^p = \exp\bigl(p \log(x)\bigr).$$

We now define the expression

$$x^y = \exp\bigl(y\log(x)\bigr).$$

for $x > 0$ and any $y \in \mathbb{R}$. We then also obtain the useful formula

$$\log(x^y) = y\log(x).$$

The usual rules of exponentials follow from the rules for exp. Let us differentiate x^y

$$\frac{d}{dx}\bigl(x^y\bigr) = \frac{d}{dx}\Bigl(\exp\bigl(y\log(x)\bigr)\Bigr) = \exp\bigl(y\log(x)\bigr)\frac{y}{x} = yx^{y-1}.$$

As exp is a strictly increasing function than as long as $x > 1$ so that $\log(x) > 0$, then x^y is a strictly increasing function y.

Let us also notice that log grows very slowly as $x \to \infty$. In fact it grows slower than any positive power of x. Let $\epsilon > 0$, $y > \epsilon$, and $x > 1$.

$$x^{-y}\log(x) = x^{-y}\int_1^x t^{-1}dt \leq x^{-y}\int_1^x t^{\epsilon-1}dt = x^{-y}\frac{x^{\epsilon}-1}{\epsilon} < \frac{x^{\epsilon-y}}{\epsilon}$$

As $\epsilon - y < 0$, we conclude that $x^{\epsilon-y}$ goes to 0 as $x \to \infty$. The claim follows, that is precisely that

$$\lim_{x\to\infty} x^{-y}\log(x) = \lim_{x\to\infty}\frac{\log(x)}{x^y} = 0$$

this is true for any $y > 0$.

Exercise: Show that log is real-analytic on $(0, \infty)$. Hint: Use the definition.

Trigonometric functions

We can now define the trigonometric functions.

$$\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \qquad \text{and} \qquad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}.$$

Immediately we obtain

$$e^{iz} = \cos(z) + i\sin(z).$$

From the definition of e^z we notice that $\overline{e^z} = e^{\bar{z}}$ and so if x is real then

$$\overline{e^{ix}} = e^{-ix}.$$

Therefore when x is real then we note that from the definition

$$\cos(x) = \operatorname{Re} e^{ix} \qquad \text{and} \qquad \sin(x) = \operatorname{Im} e^{ix}.$$

In other words, sine and cosine are real-valued when we plug in real x.

Also direct from the definition is that $\cos(-z) = \cos(z)$ and $\sin(z) = -\sin(-z)$ for all $z \in \mathbb{C}$.

We will prove that this definition has the geometric properties we usually associate with sin and cos. Let x be real and compute

$$1 = e^{ix}e^{-ix} = \left|e^{ix}\right|^2 = \bigl(\cos(x)\bigr)^2 + \bigl(\sin(x)\bigr)^2.$$

We see that e^{ix} is unimodular, the values lie on the unit circle. By noting that a square is always positive we notice that

$$\bigl(\sin(x)\bigr)^2 = 1 - \bigl(\cos(x)\bigr)^2 \leq 1$$

so $|\sin(x)| \leq 1$ and similarly we show $|\cos(x)| \leq 1$. A fact we have often used in examples.

From the definition we get that $\cos(0) = 1$ and $\sin(0) = 0$. By direct computation (LTS) from the definition we also obtain:

$$\cos(z) = \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}z^{2k}$$

$$\sin(z) = \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}z^{2k+1}$$

Again by direct calculation (LTS) we obtain

$$\frac{d}{dx}\big(\cos(x)\big) = -\sin(x) \qquad \text{and} \qquad \frac{d}{dx}\big(\sin(x)\big) = \cos(x).$$

Another fact that we have often used in examples was that $\sin(x) \leq x$ for $x \geq 0$. Let us look at $f(x) = x - \sin(x)$ and differentiate:

$$f'(x) = \frac{d}{dx}\big(x - \sin(x)\big) = 1 - \cos(x) \geq 0,$$

for all x as $|\cos(x)| \leq 1$. So f is nondecreasing and furthermore $f(0) = 0$. So f must be nonnegative when $x \geq 0$, which is precisely what we wanted to prove.

We claim that there exists a positive x such that $\cos(x) = 0$. As $\cos(0) = 1 > 0$, then cosine is definitely positive at least for some x near 0. Suppose that on $[0, y)$ we have $\cos(x) > 0$, then $\sin(x)$ is strictly increasing on $[0, y)$ and as $\sin(0) = 0$, then $\sin(x) > 0$ for $x \in (0, y)$. Take $a \in (0, y)$. Then for some $c \in (a, y)$

$$2 \geq \cos(a) - \cos(y) = \sin(c)(y - a) \geq \sin(a)(y - a)$$

As $a > 0$, then $\sin(a) > 0$ and so

$$y \leq \frac{2}{\sin(a)} + a$$

Hence there is some largest y such that $\cos(x) > 0$ in $[0, y)$. By continuity, for that y, $\cos(y) = 0$. In fact, y is the smallest positive y such that $\cos(y) = 0$. We can now define

$$\pi = 2y.$$

And we obtain that $\cos(\pi/2) = 0$ and so $\big(\sin(\pi/2)\big)^2 = 1$. As sin was positive on $(0, y)$ we have $\sin(\pi/2) = 1$.

Hence

$$\exp(i\pi/2) = i$$

and by the addition formula we get

$$\exp(i\pi) = -1 \qquad \exp(i2\pi) = 1$$

So $\exp(i2\pi) = 1 = \exp(0)$. The addition formula now says that

$$\exp(z + i2\pi) = \exp(z)$$

for all $z \in \mathbb{C}$. Therefore, we also obtain that

$$\cos(z + 2\pi) = \cos(z) \qquad \text{and} \qquad \sin(z + 2\pi) = \sin(z).$$

So sin and cos are 2π-periodic. We claim that sin and cos are not periodic with a smaller period. It would be enough to show that if $\exp(ix) = 1$ for the smallest positive x, then $x = 2\pi$. Well let x be the smallest positive x such that $\exp(ix) = 1$. Of course, $x \leq 2\pi$. Then by the addition formula

$$\big(\exp(ix/4)\big)^4 = 1$$

If $\exp(ix/4) = a + ib$ then or

$$(a + ib)^4 = a^4 - 6a^2b^2 + b^4 + i(4ab(a^2 - b^2)) = 1$$

As $x/4 \leq \pi/2$, then $a = \cos(x/4) \geq 0$ and $0 < b = \sin(x/4)$. Then either $a = 0$, in which case $x/4 = \pi/2$ or $a^2 = b^2$. But if $a^2 = b^2$, then $a^4 - 6a^2b^2 + b^4 = -4a^4 < 0$ and in particular not equal to 1. Therefore $a = 0$ in which case $x/4 = \pi/2$.

Therefore 2π is the smallest period we could choose for $\exp(ix)$ and hence cos and sin.

Finally we also wish to show that $\exp(ix)$ is 1-to-1 and onto from the set $[0, 2\pi)$ to the set of $z \in \mathbb{C}$ such that $|z| = 1$. First suppose that $\exp(ix) = \exp(iy)$ and assume that $x > y$ then $\exp(i(x - y)) = 1$, which means that $x - y$ is a multiple of 2π and hence only one of them can live in $[0, 2\pi)$. To show onto, pick $(a, b) \in \mathbb{R}^2$ such that $a^2 + b^2 = 1$. If $a, b \geq 0$, then there must exist an $x \in [0, \pi/2]$ such that $\cos(x) = a$, and hence $b^2 = \big(\sin(x)\big)^2$ and since b and $\sin(x)$ are positive we have $b = \sin(x)$. Note that since $-\sin(x)$ is the derivative of $\cos(x)$ and $\cos(-x) = \cos(x)$ then $\sin(x) < 0$ for $x \in [-\pi/2, 0)$. Then using the same reasoning we obtain that if if $a > 0$ and $b \leq 0$, we can find an $x \in [-\pi/2, 0)$, or in other words in $[3\pi/2, 2\pi)$ As multiplying by -1 is the same as multiplying by $\exp(i\pi)$ or $\exp(-i\pi)$ we can always assume that $a \geq 0$ (details left to student).

While we haven't looked at arclength, let us just state without further discussion that the arclength of a curve parametrized by $\gamma\colon [a, b] \to \mathbb{C}$ is given by

$$\int_a^b |\gamma'(t)| \, dt.$$

We have that $\exp(it)$ parametrizes the circle for t in $[0, 2\pi)$. As $\frac{d}{dt}(\exp(it)) = i \exp(it)$ and so the circumference of the circle is

$$\int_0^{2\pi} |i \exp(it)| \, dt = \int_0^{2\pi} 1 \, dt = 2\pi.$$

More generally we notice that $\exp(it)$ parametrizes the circle by arclength. That is, t measures the arclength, and hence a circle of radius 1 by the angle in radians. Hence the definitions of sin and cos we have used above agree with the standard geometric definitions.

Fundamental theorem of algebra

Let us prove the fundamental theorem of algebra, that is the algebraic completeness of the complex field, any polynomial has a root. We will first prove a bunch of lemmas about polynomials.

Lemma: Let $p(z)$ be complex polynomial. If $p(z_0) \neq 0$, then there exist $w \in \mathbb{C}$ such that $|p(w)| < |p(z_0)|$. In fact, we can pick w to be arbitrarily close to z_0.

Proof. Without loss of generality assume that $z_0 = 0$ and $p(0) = 1$. Then write

$$p(z) = 1 + a_k z^k + a_{k+1} z^{k+1} + \cdots + a_d z^d$$

where $a_k \neq 0$. Pick t such that $a_k e^{ikt} = -|a_k|$ which we can do by the discussion on trigonometric functions. Then for any $r > 0$ small enough such that $1 - r^k |a_k| > 0$ we have

$$p(re^{it}) = 1 - r^k |a_k| + r^{k+1} a_{k+1} e^{i(k+1)t} + \cdots + r^d a_d e^{idt}$$

so

$$\begin{aligned} \left|p(re^{it})\right| - \left|r^{k+1} a_{k+1} e^{i(k+1)t} + \cdots + r^d a_d e^{idt}\right| &\leq \left|p(re^{it}) - r^{k+1} a_{k+1} e^{i(k+1)t} - \cdots - r^d a_d e^{idt}\right| \\ &= \left|1 - r^k |a_k|\right| = 1 - r^k |a_k| \end{aligned}$$

In other words

$$\left|p(re^{it})\right| \leq 1 - r^k \left(|a_k| - r \left|a_{k+1} e^{i(k+1)t} + \cdots + r^{d-k-1} a_d e^{idt}\right|\right)$$

For a small enough r the expression the parentheses is positive as $|a_k| > 0$. And then $|p(re^{it})| < 1 = p(0)$. $\square$

Remark: The above lemma holds essentially with an unchanged proof for (complex) analytic functions. A proof of this generalization is left as an exercise to the reader. What the lemma says is that the only minima the modulus of analytic functions (polynomials) has minima precisely at the zeros.

Note also that the lemma does not hold if we restrict to real numbers. For example, x^2+1 has a minimum at $x = 0$, but no zero there. The thing is that there is a w arbitrarily close to 0 such that $|w^2 + 1| < 1$, but this w will necessarily not be real. Letting $w = i\epsilon$ for small $\epsilon > 0$ will work.

Moral of the story is that if $p(0) = 1$, then very close to 0, the polynomial looks like $1 + az^k$ and this has no minimum at the origin. All the higher powers of z are too small to make a difference.

Lemma: Let $p(z)$ be complex polynomial. Then for any M there exists an R such that if $|z| \geq R$ then $|p(z)| \geq M$.

Proof. Write $p(z) = a_0 + a_1 z + \cdots + a_d z^d$ and suppose that $a_d \neq 0$. Suppose that $|z| \geq R$ (so also $|z|^{-1} \leq R^{-1}$). We estimate:

$$\begin{aligned} |p(z)| &\geq \left|a_d z^d\right| - |a_0| - |a_1 z| - \cdots - \left|a_{d-1} z^{d-1}\right| \\ &= |z|^d \left(|a_d| - |a_0|\,|z|^{-d} - |a_1|\,|z|^{-d+1} - \cdots - |a_{d-1}|\,|z|^{-1}\right) \\ &\geq R^d\big(|a_d| - |a_0|\,R^{-d} - |a_1|\,R^{1-d} - \cdots - |a_{d-1}|\,R^{-1}\big) \end{aligned}$$

Then the expression in parentheses is eventually positive for large enough R. In particular, for large enough R we get that it is greater than $|a_d|/2$ and so

$$|p(z)| \geq R^d \frac{|a_d|}{2}.$$

Therefore, we can pick R large enough to be bigger than a given M. $\square$

The above lemma does **not** generalize to analytic functions, even those defined in all of $\mathbb{C}$. The function $\cos(z)$ is an obvious counterexample. Note that we had to look at the term with the largest degree, and we only have such a term for a polynomial. In fact, something that we will not prove is that an analytic function defined on all of $\mathbb{C}$ satisfying the conclusion of the lemma must be a polynomial.

The moral of the story here is that for very large $|z|$ (far away from the origin) a polynomial of degree d really looks like a constant multiple of z^d.

Theorem 8.8: (Fundamental theorem of algebra) Let $p(z)$ be complex polynomial, then there exists a $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$.

Proof. Let $\mu = \inf |p(z)|$. Find an R such that for all z with $|z| \geq R$ we have $|p(z)| \geq \mu + 1$. Therefore, any z with $|p(z)|$ close to μ must be in the closed ball $C(0,R) = \{z : |z| \leq R\}$. As $|p(z)|$ is a continuous real-valued function, it achieves its minimum on the compact set $C(0,R)$ (closed and bounded) and this minimum must be μ. So there is a $z_0 \in C(0,R)$ such that $|p(z_0)| = \mu$. As that is a minimum of $|p(z)|$ on $\mathbb{C}$, then by a lemma above we have that $|p(z_0)| = 0$. $\square$

The theorem doesn't generalize to analytic functions either. For example $\exp(z)$ is an analytic function on $\mathbb{C}$ with no zeros.

Fourier Series

Now that we have trigonometric functions, let us cover Fourier series in more detail. We have seen Fourier series in examples, but let us start at the beginning.

A *trigonometric polynomial* is an expression of the form

$$a_0 + \sum_{n=1}^{N} \big(a_n \cos(nx) + b_n \sin(nx)\big),$$

or equivalently, thanks to Euler's formula:

$$\sum_{n=-N}^{N} c_n e^{inx}.$$

The second form is usually more convenient. Note that if $|z| = 1$ we can write $z = e^{ix}$, and so

$$\sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N} c_n z^n,$$

so a trigonometric polynomial is really a rational function (do note that we are allowing negative powers) evaluated on the unit circle. There is a wonderful connection between power series (actually Laurent series) and Fourier series because of this observation, but we will not investigate this further.

Notice that all the functions are 2π-periodic and hence the trig polynomials are also 2π-periodic. We could rescale x to make the period different, but the theory is the same, so we will stick with the above scale. We compute that the antiderivative of $\exp(inx)$ is $\frac{\exp(inx)}{in}$ and so

$$\int_{-\pi}^{\pi} e^{inx}\, dx = \begin{cases} 2\pi & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let us take

$$f(x) = \sum_{n=-N}^{N} c_n e^{inx},$$

and for $m = -N, \ldots, N$ compute

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-imx}\, dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\sum_{n=-N}^{N} c_n e^{i(n-m)x}\right) dx = \sum_{n=-N}^{N} c_n \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)x}\, dx = c_m.$$

We therefore have a way of computing the coefficients c_m by an integral of f. Of course if $|m| > N$ the integral is just 0. We might as well have included enough zero coefficients to make $|m| \leq N$.

Proposition: Trigonometric polynomial $f(x)$ is real-valued for real x if and only if $c_{-m} = \overline{c_m}$ for all $m = -N, \ldots, N$.

Proof. If $f(x)$ is real-valued, that is $\overline{f(x)} = f(x)$, then

$$\overline{c_m} = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-imx}\, dx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{f(x)e^{-imx}}\, dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{imx}\, dx = c_{-m}.$$

The complex conjugate goes inside the integral because the integral is done on real and imaginary parts separately. On the other hand if $c_{-m} = \overline{c_m}$, then we notice that

$$\overline{c_{-m}e^{-imx} + c_m e^{imx}} = \overline{c_{-m}}e^{imx} + \overline{c_m}e^{-imx} = c_m e^{imx} + c_{-m}e^{-imx}.$$

We also have that $c_0 = \overline{c_0}$. So by pairing up the terms we obtain that f has to be real-valued. □

In fact, the above could also follow from the linear independence of the functions e^{inx}, which we can now prove.

Proposition: If

$$\sum_{n=-N}^{N} c_n e^{inx} = 0$$

for all x, then $c_n = 0$ for all n.

Proof. Proof follows immediately from the integral formula for c_n. □

We now take limits. We call the series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

the *Fourier series*. The numbers c_n we call *Fourier coefficients*. We could also develop everything with sines and cosines, but it is equivalent and slightly more messy.

Several questions arise. What functions are expressible as Fourier series? Obviously, they have to be 2π-periodic, but not every periodic function is expressible with the series. Furthermore, if we do have a Fourier series, where does it converge (if at all)? Does it converge absolutely? Uniformly (and where)?

Also note that the series has two limits. When talking about Fourier series convergence, we often talk about the following limit:

$$\lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{inx}.$$

There are other ways we can sum the series that can get convergence in more situations, but we will refrain from discussion those.

For any function integrable on $[-\pi, \pi]$ we call the numbers

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx$$

its Fourier coefficients. Often these numbers are written as $\hat{f}(n)$ (For those that have seen the Fourier transform, the similarity is not just coincidental, we are really taking a type of Fourier transform here). We can then formally write down a Fourier series. As you might imagine such a series might not even converge. We will write

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

although the $\sim$ doesn't imply anything about the two sides being equal in any way. It is simply that we created a formal series using the formula for the coefficients.

We have seen in examples that if

$$|c_n| \leq \frac{C}{|n|^\alpha}$$

for some C and some $\alpha > 1$, then the Fourier series converges absolutely and uniformly to a continuous function. If $\alpha > 2$, then we had that the series converged to a differentiable function. Let us now investigate convergence of Fourier series from the other side. Given properties of f, what can we say about the series.

Let us first prove some general results about so called *orthonormal* systems. Let us fix an interval $[a, b]$. We will define an *inner product* for the space of functions. We will restrict our attention to Riemann integrable functions since we do not yet have the Lebesgue integral, which would be the natural choice. Let f and g be complex-valued Riemann integrable functions on $[a, b]$ and define

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)}\, dx$$

For those that have seen Hermitian inner products in linear algebra, this is precisely such a product. We have to put in the conjugate as we are working with complex numbers. We then have the "size", that is the L^2 norm $\|f\|_2$ by (defining the square)

$$\|f\|_2^2 = \langle f, f \rangle = \int_a^b |f(x)|^2\, dx.$$

Remark: Notice the similarity to finite dimensions. For $z = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$ we define

$$\langle z, w \rangle = \sum_{k=1}^{n} z_k \overline{w_k}$$

and then the norm is (usually this is just denoted by $\|z\|$ rather than $\|z\|_2$)

$$\|z\|^2 = \langle z, z \rangle = \sum_{k=1}^{n} |z_k|^2$$

This is just the euclidean distance to the origin in $\mathbb{C}^n$.

Let us get back to function spaces. We will assume all functions are Riemann integrable in the following.

Definition: Let $\{\varphi_n\}$ be a sequence of complex-valued functions on $[a,b]$. We will say that this is an *orthonormal system* if

$$\langle \varphi_n, \varphi_m \rangle = \int_a^b \varphi_n(x)\overline{\varphi_m(x)}\,dx = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the above says that $\|\varphi_n\|_2 = 1$ for all n. If we only require that $\langle \varphi_n, \varphi_m \rangle = 0$ for $m \neq n$ then the system would be just an *orthogonal system*.

We have noticed above that for example

$$\left\{ \frac{1}{\sqrt{2\pi}}\, e^{inx} \right\}$$

is an orthonormal system. The factor out in front is to make the norm be 1.

Having an orthonormal system $\{\varphi_n\}$ on $[a,b]$ and an integrable function f on $[a,b]$, we can write a Fourier series relative to $\{\varphi_n\}$. We let

$$c_n = \langle f, \varphi_n \rangle = \int_a^b f(x)\overline{\varphi_n(x)}\,dx,$$

and write

$$f(x) \sim \sum_{n=1}^{\infty} c_n \varphi_n.$$

In other words, the series is

$$\sum_{n=1}^{\infty} \langle f, \varphi_n \rangle \varphi_n(x).$$

Notice the similarity to the expression for the orthogonal projection of a vector onto a subspace from linear algebra. We are in fact doing just that, but in a space of functions.

Theorem 8.11: Let $\{\varphi_n\}$ be an orthonormal system on $[a,b]$ and suppose

$$f(x) \sim \sum_{n=1}^{\infty} c_n \varphi_n(x).$$

If

$$s_n(x) = \sum_{k=1}^{n} c_k \varphi_k(x) \quad \text{and} \quad t_n(x) = \sum_{k=1}^{n} d_k \varphi_k(x).$$

for some other sequence $\{d_k\}$, then

$$\int_a^b \left| f(x) - s_n(x) \right|^2\,dx = \|f - s_n\|_2^2 \leq \|f - t_n\|_2^2 = \int_a^b \left| f(x) - t_n(x) \right|^2\,dx.$$

with equality only if $d_k = c_k$ for all $k = 1, \ldots, n$.

In other words the partial sums of the Fourier series are the best approximation with respect to the L^2 norm.

Proof. Let us write

$$\int_a^b |f - t_n|^2 = \int_a^b |f|^2 - \int_a^b f\overline{t_n} - \int_a^b \overline{f} t_n + \int_a^b |t_n|^2.$$

Now

$$\int_a^b f\overline{t_n} = \int_a^b f \sum_{k=1}^{n} \overline{d_k}\overline{\varphi_k} = \sum_{k=1}^{n} \overline{d_k} \int_a^b f\overline{\varphi_k} = \sum_{k=1}^{n} \overline{d_k} c_k,$$

and

$$\int_a^b |t_n|^2 = \int_a^b \sum_{k=1}^{n} d_k\varphi_k \sum_{j=1}^{n} \overline{d_j}\overline{\varphi_j} = \sum_{k=1}^{n}\sum_{j=1}^{n} d_k\overline{d_j} \int_a^b \varphi_k\overline{\varphi_j} = \sum_{k=1}^{n} |d_k|^2.$$

So

$$\int_a^b |f - t_n|^2 = \int_a^b |f|^2 - \sum_{k=1}^n \overline{d_k} c_k - \sum_{k=1}^n d_k \overline{c_k} + \sum_{k=1}^n |d_k|^2 = \int_a^b |f|^2 - \sum_{k=1}^n |c_k|^2 + \sum_{k=1}^n |d_k - c_k|^2 .$$

This is minimized precisely when $d_k = c_k$. $\square$

When we do plug in $d_k = c_k$, then

$$\int_a^b |f - s_n|^2 = \int_a^b |f|^2 - \sum_{k=1}^n |c_k|^2$$

and so

$$\sum_{k=1}^n |c_k|^2 \leq \int_a^b |f|^2$$

for all n. Note that

$$\sum_{k=1}^n |c_k|^2 = \|s_n\|_2^2$$

by the above calculation. We can take a limit to obtain

Theorem 8.12: (Bessel's inequality)

$$\sum_{k=1}^\infty |c_k|^2 \leq \int_a^b |f|^2 = \|f\|_2^2$$

In particular (as for a Riemann integrable function $\int_a^b |f|^2 < \infty$), we get that $\lim c_k = 0$.

Let us return to the trigonometric Fourier series. Here we note that the system $\{e^{inx}\}$ is orthogonal, but not orthonormal if we simply integrate over $[-\pi, \pi]$. We can also rescale the integral and hence the inner product to make $\{e^{inx}\}$ orthonormal. That is, if we replace

$$\int_a^b \qquad \text{with} \qquad \frac{1}{2\pi} \int_{-\pi}^{\pi}$$

(we are just rescaling the dx really, this is a common trick in analysis) then everything works and we obtain that the system $\{e^{inx}\}$ is orthonormal with respect to the inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \overline{g(x)} \, dx.$$

So suppose we have an integrable function f on $[-\pi, \pi]$. In fact suppose that f is a function defined on all of $\mathbb{R}$ and is 2π periodic. Let

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx.$$

Write

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

We will look at the symmetric partial sums

$$s_N(x) = s_N(f; x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

The inequality leading up to Bessel now reads:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |s_N(x)|^2 \, dx = \sum_{n=-N}^{N} |c_n|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

Let us now define the *Dirichlet kernel*

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}.$$

We claim that

$$D_N(x) = \sum_{n=-N}^{N} e^{inx} = \frac{\sin\big((N + {}^1\!/\!_2)x\big)}{\sin({}^x\!/\!_2)}.$$

at least for x such that $\sin(x/2) \neq 0$. We know that the left hand side is continuous and hence the right hand side extends continuously as well. To show the claim we use a familiar trick:

$$(e^{ix} - 1)D_N(x) = e^{i(N+1)x} - e^{-iNx}.$$

And multiply by $e^{-ix/2}$

$$(e^{ix/2} - e^{-ix/2})D_N(x) = e^{i(N+{}^1\!/\!_2)x} - e^{-i(N+{}^1\!/\!_2)x}.$$

The claim follows.

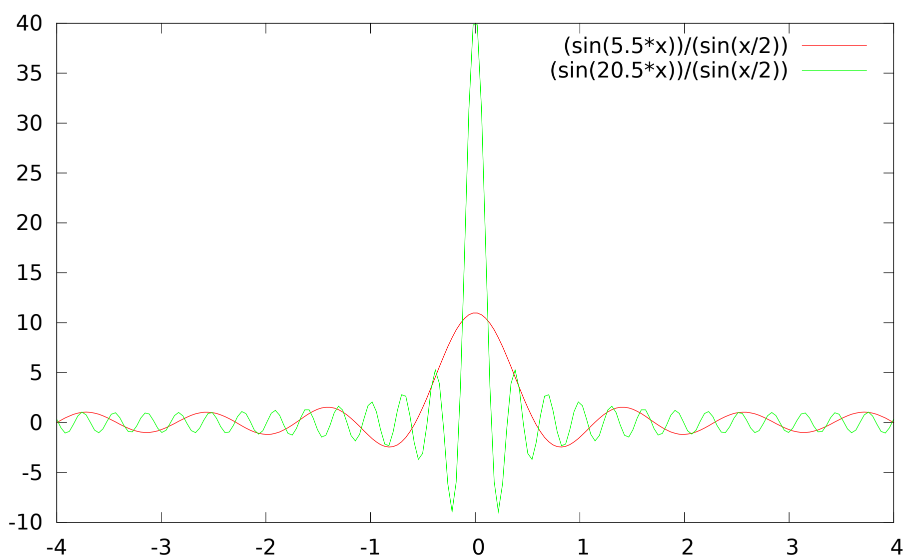
We expand the definition of s_N

$$s_N(x) = \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)e^{-int}\,dt\,e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=-N}^{N} e^{in(x-t)}\,dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)D_N(x-t)\,dt.$$

Convolution strikes again! As D_N and f are 2π-periodic we can also change variables and write

$$s_N(x) = \frac{1}{2\pi} \int_{x-\pi}^{x+\pi} f(x-t)D_N(t)\,dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t)D_N(t)\,dt$$

Look at an example plot of $D_N(x)$ for $N = 5$ and $N = 20$:



Note that the central peak will get taller and taller with N being larger, and the side peaks will stay small (but oscillate wildly). Again, we are looking at in some sense an approximate delta function, although it has all these oscillations away from zero which do not go away. So we expect that s_N goes to f. Things are not so simple, but under some conditions on f, such a conclusion holds.

People write

$$\delta(x) \sim \sum_{n=\infty}^{\infty} e^{inx}$$

although we can't say that as we have not really defined the delta function, no a Fourier series of whatever kind object it is.

Theorem 8.14: Let x be fixed and let f be Riemann integrable on $[-\pi, \pi]$. Suppose that there exist $\delta > 0$ and M such that

$$|f(x+t) - f(x)| \leq M\,|t|$$

for all $t \in (-\delta, \delta)$, then

$$\lim_{N\to\infty} s_N(f;x) = f(x).$$

In other words, if for example f is differentiable at x then we obtain convergence. Generally what the result implies is that if the function is continuous piecewise smooth, then the Fourier series converges (pointwise). By continuous piecewise smooth we mean that f is continuous and periodic so $f(-\pi) = f(\pi)$ and furthermore that there are points $x_0 = -\pi < x_1 < \cdots < x_k = \pi$ such that f restricted to $[x_j, x_{j+1}]$ is continuously differentiable (up to the endpoints) for all j.

Proof. We notice that for all N we get

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N = 1.$$

Write

$$\begin{aligned} s_N(f;x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)\,dt - f(x)\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \big(f(x-t) - f(x)\big)D_N(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(x-t) - f(x)}{\sin(t/2)} \sin\big((N + {}^1\!/_2)t\big)\,dt \end{aligned}$$

Now by the hypotheses we obtain that for small nonzero t we get

$$\left|\frac{f(x-t) - f(x)}{\sin(t/2)}\right| \leq \frac{M\,|t|}{|\sin(t/2)|}$$

As $\sin(t) = t + h(t)$ where $\frac{h(t)}{t} \to 0$ as $t \to 0$, we notice that $\frac{M|t|}{|\sin(t/2)|}$ is continuous at the origin and hence $\frac{f(x-t)-f(x)}{\sin(t/2)}$ must be bounded near the origin. As $t = 0$ is the only place on $[-\pi, \pi]$ where the denominator vanishes, it is the only place where there could be a problem. The function is also Riemann integrable. Now we use a trigonometric identity that follows from the definition (and you've seen it on the homework actually) that

$$\sin\big((N + {}^1\!/_2)t\big) = \cos(t/2)\sin(Nt) + \sin(t/2)\cos(Nt)$$

so

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(x-t) - f(x)}{\sin(t/2)} \sin\big((N + {}^1\!/_2)t\big)\,dt =& \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\frac{f(x-t) - f(x)}{\sin(t/2)}\cos(t/2)\right)\sin(Nt)\,dt \\ &+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \big(f(x-t) - f(x)\big)\cos(Nt)\,dt \end{aligned}$$

Now $\frac{f(x-t)-f(x)}{\sin(t/2)}\cos(t/2)$ and $\big(f(x-t) - f(x)\big)$ are bounded Riemann integrable functions and so their Fourier coefficients go to zero by Theorem 8.12. So the two integrals on the right hand side, which compute the Fourier coefficients for the real version of the Fourier series go to 0 as N goes to infinity. This is because $\sin(Nt)$ and $\cos(Nt)$ are also orthonormal systems. with respect to the same inner product. Hence $s_N(f;x) - f(x)$ goes to 0 and so $s_N(f;x)$ goes to $f(x)$. $\square$

In particular this has the following corollary:

Corollary: If $f(x) = 0$ on an entire open interval J, then $\lim s_N(f;x) = 0$ for all $x \in J$.

In other words, if two functions f and g are equal on an open interval J, then the points on J where $\{s_N(f;x)\}$ and $\{s_N(g;x)\}$ converge are the same. That is, convergence at x is only dependent on the values of the function near x.

We have seen Theorem 8.15 as an example for Stone-Weierstrass theorem. That is, any continuous function on $[-\pi, \pi]$ can be uniformly approximated by trigonometric polynomials. However, these trigonometric polynomials need not be the partial sums s_N. On the other hand, (exercise 15) they can be explicitly constructed from s_N.

We have that the convergence always happens in the L^2 sense and furthermore that formal operations on the (infinite) vectors of Fourier coefficients is the same as the operations using the integral inner product.

We will mostly sketch out the proof and leave some details to the reader as exercises. Some of these are exercises in Rudin.

Theorem 8.16: (Parseval) Let f and g be Riemann integrable 2π-periodic functions with

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad \text{and} \qquad g(x) \sim \sum_{n=-\infty}^{\infty} d_n e^{inx}.$$

Then

$$\lim_{N\to\infty} \|f - s_N(f)\|_2^2 = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 \, dx = 0.$$

Also

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)} \, dx = \sum_{n=-\infty}^{\infty} c_n \overline{d_n},$$

and

$$\|f\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

We will skip the proof in lecture.

Proof. It is not hard too prove (Exercise 12 in chapter 6) that there is a continuous 2π-periodic function h such that

$$\|f - h\|_2 < \epsilon.$$

Now we know that we can approximate h with a trigonometric polynomial uniformly, that is there is a trigonometric polynomial $P(x)$ such that $|h(x) - P(x)| < \epsilon$ for all x. Hence

$$\|h - P\|_2 \le \epsilon.$$

If P is of degree N_0 then for all $N \ge N_0$ we have

$$\|h - s_N(h)\|_2 \le \|h - P\|_2 \le \epsilon$$

as $s_N(h)$ is the best approximation for h in L^2 (Theorem 8.11). Next by the inequality leading up to Bessel we have

$$\|s_N(h) - s_N(f)\|_2 = \|s_N(h - f)\|_2 \le \|h - f\|_2 \le \epsilon$$

It is not difficult (exercise 11 in chapter 6) to show the triangle inequality for the L^2 norm, that is

$$\|f - s_N(f)\|_2 \le \|f - h\|_2 + \|h - s_N(h)\|_2 + \|s_N(h) - s_N(f)\|_2 \le 3\epsilon.$$

For all $N \ge N_0$.

Next

$$\langle s_N(f), g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} s_N(f;x)\overline{g(x)} \, dx = \sum_{k=-N}^{N} c_k \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx}\overline{g(x)} \, dx = \sum_{k=-N}^{N} c_k \overline{d_k}$$

Next we need the Schwarz (or Cauchy-Schwarz) inequality (left as exercise), that is

$$\left| \int_a^b f\bar{g} \right|^2 \le \left(\int_a^b |f|^2 \right) \left(\int_a^b |g|^2 \right)$$

This is left as an exercise. It actually follows by purely formal linear algebra using simple the idea that the integral gives an inner product. So

$$\left|\int_{-\pi}^{\pi} f\bar{g} - \int_{-\pi}^{\pi} s_N(f)g\right| = \left|\int_{-\pi}^{\pi} (f - s_N(f))g\right| \leq \int_{-\pi}^{\pi} |f - s_N(f)|\ |g| \leq \left(\int_{-\pi}^{\pi} |f - s_N(f)|^2\right)^{1/2} \left(\int_{-\pi}^{\pi} |g|^2\right)^{1/2}.$$

Now the right hand side goes to 0 as N goes to infinity. $\square$