17 Power series

17.1 Introduction to Taylor series

Recall that as an application of the ratio test, we said that

$$\sum_{n\geq 0} \frac{x^{2n}}{(2n)!}$$

converges to a limit for all real x. Let $f(x)$ be limit. We claim that $f(x) = \cosh x$. Indeed, by Taylor's theorem with Lagrange form of the remainder term, we have that for each fixed x there is some t between 0 and x for which

$$\cosh x = 1 + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!} + (\cosh)^{(2n+1)}(t)\frac{x^{2n+1}}{(2n+1)!}.$$

so

$$\left|\cosh x - \left(1 + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!}\right)\right| = \left|\sinh t \frac{x^{2n+1}}{(2n+1)!}\right|.$$

Now between 0 and x, $|\sinh t|$ is never more than $\sinh |x|$, and so

$$\left|\cosh x - \left(1 + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!}\right)\right| \leq |\sinh x| \frac{|x|^{2n+1}}{(2n+1)!}.$$

For each fixed real x, we have that $|\sinh x||x^{2n+1}|/(2n+1)! \to 0$ as $n \to \infty$ and so, by definition of convergence of a sequence to a limit, we get that

$$\left(1 + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!}\right)_{n=0}^{\infty} \to \cosh x$$

for each fixed real x, that is,

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

Note that we used the ratio test to conclude that $(x^{2n}/(2n)!)$ converges for all real x, but in fact that was unnecessary, given Taylor's theorem. Indeed, suppose we have a function $f(x)$, and we know that for all fixed x in some set A, the remainder term $R_{n,a,f}(x)$ goes to 0 as n goes to infinity. Then we have

$$|f(x) - P_{n,a,f}(x)| \to 0$$

as n goes to infinity, so that, directly from the definition of convergence, for each fixed $x \in A$, $(f^{(n)}(a)(x-a)^n/n!)$ is summable, and the sum is $f(x)$. That is, for $x \in A$

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

So, for example, we have

- $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ for all real x,
- $\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ for all real x,
- $\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ for all real x,
- $\tan^{-1} x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$ for $-1 \leq x \leq 1$ (and note that when $|x| > 1$, the sum *does not converge*, by the vanishing criterion, although $\tan^{-1}$ is defined for $|x| > 1$).

The expression

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

is called the *Taylor series*[250] *of f at* (or *about*) a (and evaluated at x); think of it as an "infinite degree Taylor polynomial". Sometimes, the Taylor series evaluates to the function it is the Taylor series of, at all points in the domain of the function (e.g. when $f(x) = e^x$ and $a = 0$); sometimes, it evaluates to the function at some points of the domain, but not all of the function (e.g. when $f(x) = \tan^{-1}(x)$ and $a = 0$); and sometimes, it evaluates to the function at the single points of the domain a (e.g. when

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0, \end{cases}$$

at $a = 0$). We will discuss this in more detail later.

We now consider the degree to which the Taylor series of a function is a useful tool for understanding the function. We start with an example. For clarity, write

- $\sin_n(x)$ for $P_{2n+1,0,\sin}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$

and

- $\cos_n(x)$ for $P_{2n,0,\cos}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!}$.

As we have earlier observed, we have that for all real x

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = \lim_{n \to \infty} \sin_n(x)$$

and

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \lim_{n \to \infty} \cos_n(x).$$

Notice that

$$\int_0^x \cos_n(t)\ dt = \sin_n(x)$$

[250] When $a = 0$, the Taylor series is often referred to as the *Maclaurin* series.

so that

$$\sin x = \lim_{n\to\infty} \int_0^x \cos_n(t)\ dt. \quad (\star)$$

Can we use $(\star)$ to conclude that $\sin_x = \int_0^x \cos(t)\ dt$ for all real x? In other words, is it true that

$$\lim_{n\to\infty} \int_0^x \cos_n(t)\ dt = \int_0^x \lim_{n\to\infty} \cos_n(t)\ dt?$$

In general, we cannot draw such a conclusion. That is, if $f_n : \mathbb{R} \to \mathbb{R}$ is a sequence of functions, one for each $n \in \mathbb{N}$, with the properties that for all real x, each of

- $\lim_{n\to\infty} f_n(x)$,
- $\int_0^x \lim_{n\to\infty} f_n(t)\ dt$
- $\int_0^x f_n(t)\ dt$ and
- $\lim_{n\to\infty} \int_0^x f_n(t)\ dt$

exist, we *cannot* conclude that

$$\lim_{n\to\infty} \int_0^x f_n(t)\ dt = \int_0^x \lim_{n\to\infty} f_n(t)\ dt \quad (\star\star)$$

for all x.

Here is a counter-example. Define f_n via:

$$f_n(x) = \begin{cases} 0 & \text{if } x \leq 0 \text{ or } x \geq 2/n \\ n^2 x & \text{if } 0 \leq x \leq 1/n \\ 2n - n^2 x & \text{if } 1/n \leq x \leq 2/n. \end{cases}$$

Observe[251] that for each real x,

- $\lim_{n\to\infty} f_n(x) = 0$ (if $n \leq 0$, this is automatic; if $x > 0$ then as long as $n > 2/x$ we have $f_n(x) = 0$),
- $\int_0^x \lim_{n\to\infty} f_n(t)\ dt = 0$ (this is automatic from the last observation),
- $\int_0^x f_n(t)\ dt$ exists for all n (f_n is continuous and bounded) and
- $$\lim_{n\to\infty} \int_0^x f_n(t)\ dt = \begin{cases} 0 & \text{if } x \leq 0 & \text{(immediate)} \\ 1 & \text{if } x > 0 & \text{(for all } n > 2/x,\ \int_0^x f_n(t)\ dt = 1). \end{cases}$$

So

$$\lim_{n\to\infty} \int_0^x f_n(t)\ dt \neq \int_0^x \lim_{n\to\infty} f_n(t)\ dt$$

for all x; only for $x \leq 0$.

In the next section, we isolate the condition that is needed to make $(\star\star)$ hold, and explore its consequences.

[251] Draw a graph to see what's going on!

17.2 Pathologies of pointwise convergence

At the end of the last section we saw that if a sequence of integrable functions (f_n) converges *pointwise* to an integrable limit f on some set A, meaning that for each $x \in A$ we have $f_n(x) \to f(x)$ as $n \to \infty$, then it is not necessarily that case that the limit of the integrals of the f_n's is the integral of the limit.

It is fairly clear what is going on in the counter-example: although the functions f_n are converging, at each real x, to the value 0, there is some sense in which the functions f_n as a whole are not converging to 0: we cannot put an ε-width window around the x-axis, and then find an N large enough so that for all $n > N$ the graph of f_n is lying completely inside that window. What we can do is, for each specific x, find an N so that for all $n > N$ we have $f_n(x)$ within ε of $f(x)$; but as the x's get closer to 0, the required N gets larger and larger.

Here's another, perhaps more worrying example. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x^n & \text{if } 0 \leq 1 \leq 1 \\ 1 & \text{if } x \geq 1. \end{cases}$$

For each x we have that $f_n(x)$ converges to a limit as n grows. Specifically,

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1 & \text{if } x \geq 1. \end{cases}$$

Notice that

> Each f_n is continuous, but the limit f is not.

So, the limit of continuous functions need not be continuous. This example can be modified slightly, by rounding the corners of f_n, to also yield a sequence of differentiable functions that tends to a limit, but not to a differentiable limit.

It's not as clear what the problem might be in this second (pair of) example(s); but a little thought shows that it suffers from the same issue as the first example. Given $\varepsilon > 0$, for each x there is an N large enough so that $f_n(x)$ is within ε of $\lim_{n\to\infty} f_n(x)$ for all $n > N$. But if we put an ε-width window around the limit function (a strip of width ε around the x-axis, for $x < 1$, and a strip of width ε around the line $y = 1$, for $x \geq 1$), then there is no single N with the property that for all $n > N$ the graph of f_n lies inside this window: for each $n \in \mathbb{N}$, the point $(\sqrt[n]{(1/2)}, 1/2)$ is on the graph of f_n, but it is not in the window as long as $\varepsilon < 1/2$.

So: we can talk about a sequence (f_n) of functions (all with the same domain A) converging to a limit f (also with domain A), if for each $x \in A$ we have $f_n(x) \to f(x)$ as $n \to \infty$. This kind of convergence is called *pointwise* convergence. But pointwise convergence appears to be quite weak; it preserves neither continuity nor integrability.

Or: we can talk about a sequence (f_n) of functions (all with the same domain A) converging to a limit f (also with domain A) if the graphs of the f_n's converge to the graph of f, in the sense that for all $\varepsilon > 0$ there is N such that for all $n > N$ the graph of f_n lies completely inside a window of width ε around the graph of f. This notion of converges is easily seen to imply pointwise convergence; and as we are about to see, it behaves much better with respect to continuity, integrability and differentiability. Before exploring this, we made the definition of this kind of convergence, which we call *uniform* convergence, precise.

17.3 Definition and basic properties of uniform convergence

We begin with the definition.

Definition of uniform convergence Let (f_n) be a sequence of functions, all defined on some domain A, and let f be a function, also defined on the domain A. Say that (f_n) *converges uniformly to f on A* if for all $\varepsilon > 0$ there is N such that $n > N$ implies that for all $x \in A$,

$$|f_n(x) - f(x)| < \varepsilon.$$

It's worth symbolically comparing the definition of "$f_n \to f$ pointwise on A":

$$(\forall \varepsilon > 0)(\forall x \in A)(\exists N)((n > N) \Rightarrow (|f_n(x) - f(x)| < \varepsilon))$$

with that of "$f_n \to f$ uniformly on A":

$$(\forall \varepsilon > 0)(\exists N)(\forall x \in A)((n > N) \Rightarrow (|f_n(x) - f(x)| < \varepsilon)).$$

In pointwise convergence, N depends on both ε *and* x; but in uniform convergence, N only depends on ε — for each ε, the *same* N works for every x. This should seem familiar — it is essentially the same distinction as between continuity and uniform continuity.

The value of uniform convergence is conveyed in the following three theorems, which (roughly) say that

- the uniform limit of continuous functions is continuous,
- the integral of the uniform limit is the limit of the integrals, and
- (modulo some extra conditions) the derivative of the uniform limit is the limit of the derivatives.

Theorem 17.1. *Let (f_n) be a sequence of functions that are all defined and continuous on $[a, b]$, and that converge uniformly on $[a, b]$ to some a function f. Then f is continuous on $[a, b]$.*

Proof: Unsurprisingly, we will show the slightly stronger statement that f is *uniformly* continuous on $[a, b]$. Given $\varepsilon > 0$ we want a $\delta > 0$ such that $|x - y| < \delta$ (with $x, y \in [a, b]$) implies $|f(x) - f(y)| < \varepsilon$. We write

$$\begin{aligned} |f(x) - f(y)| &= |(f(x) - f_n(x)) + (f_n(x) - f_n(y)) + (f_n(y) - f(y))| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f(y) - f_n(y)|. \quad (\star) \end{aligned}$$

By *uniform* convergence of f_n to f, we know that there is N such that $n > N$ implies both

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3}, \qquad |f(y) - f_n(y)| < \frac{\varepsilon}{3}. \quad (\star\star)$$

Fix any $n > N$. Since f_n is uniformly continuous (it is a continuous function on a closed interval), there is $\delta > 0$ such that $|x - y| < \delta$ (with $x, y \in [a, b]$) implies $|f_n(x) - f_n(y)| < \varepsilon/3$. But this implies, via $(\star)$ and $(\star\star)$ that whenever $|x - y| < \delta$ (with $x, y \in [a, b]$) we have

$$|f(x) - f(y)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

□

Theorem 17.2. *Let (f_n) be a sequence of functions that are all defined and integrable on $[a, b]$, and that converge uniformly on $[a, b]$ to some a function f. Then f is integrable*[252] *on $[a, b]$ and*

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

Proof: We first check that f is bounded. By uniform convergence of (f_n) to f there is N such that $n > N$ implies $|f_n(x) - f(x)| < 1$ for all $x \in [a, b]$, which says that $|f(x)| < |f_n(x)| + 1$. Now fix such an $n > N$. Since f_n is integrable, it is bounded on $[a, b]$, say by M; but then that says that f is bounded on $[a, b]$ by $M + 1$.

Next we show that f is integrable on $[a, b]$. Fix $\varepsilon > 0$. By uniform convergence of (f_n) to f there is N such that $n > N$ implies

$$|f_n(x) - f(x)| < \frac{\varepsilon}{4(b - a)}$$

for all $x \in [a, b]$. Fix such an $n > N$. Since f_n is integrable on $[a, b]$, there is a partition $P = (t_0, \ldots, t_m)$ of $[a, b]$ with

$$U(f_n, P) - L(f_n, P) = \sum_{i=1}^{m} (M_i^n - m_i^n)(t_i - t_{i-1}) < \varepsilon/2,$$

where M_i^n is the supremum of f_n on $[t_{i-1}, t_i]$, and m_i^n is the infimum.

Because $|f_n(x) - f(x)| < \frac{\varepsilon}{4(b-a)}$ for all $x \in [a, b]$, we have that

$$M_i \le M_i^n + \frac{\varepsilon}{4(b - a)}, \qquad m_i \ge m_i^n - frac\varepsilon 4(b - a)$$

where M_i is the supremum of f on $[t_{i-1}, t_i]$, and m_i is the infimum. It follows that

$$M_i - m_i \le M_i^n - m_i^n + \frac{\varepsilon}{2(b - a)}$$

[252]Spivak takes integrability of f as a hypothesis, but it is in fact implied by uniform convergence, and is a good exercise in definition of the integral.

for each i, and so

$$\begin{aligned}
U(f_n,P)-L(f_n,P) &= \sum_{i=1}^{m}(M_i-m_i)(t_i-t_{i-1}) \\
&\leq \sum_{i=1}^{m}\left(M_i^n-m_i^n+\frac{\varepsilon}{2(b-a)}\right)(t_i-t_{i-1}) \\
&= \sum_{i=1}^{m}(M_i^n-m_i^n)(t_i-t_{i-1})+\frac{\varepsilon}{2(b-a)}\sum_{i=1}^{m}(t_i-t_{i-1}) \\
&< \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,
\end{aligned}$$

and so we conclude that f is integrable on $[a,b]$.

Finally, we show that $\int_a^b f = \lim_{n\to\infty}\int_a^b f_n$. Fix $\varepsilon>0$. By uniform convergence of (f_n) to f there is N such that $n>N$ implies

$$|f_n(x)-f(x)|<\frac{\varepsilon}{b-a}$$

for all $x\in[a,b]$. So (using integrability of f_n and f to justify the existence of the integrals below)

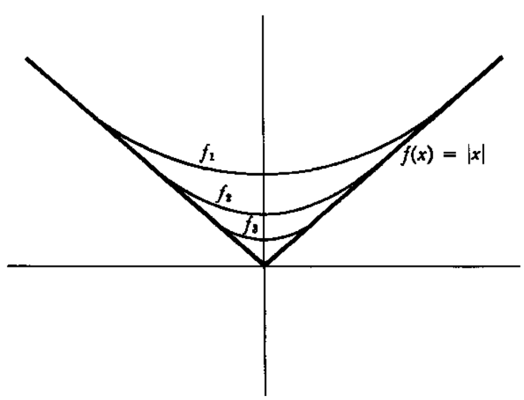
$$\left|\int_a^b f_n-\int_a^b f\right|\leq\int_a^b|f_n-f|<(b-a)\frac{\varepsilon}{b-a}=\varepsilon,$$

which says $\int_a^b f_n\to\int_a^b f$ as $n\to\infty$. □

It would be nice to now have a theorem that says

> Let (f_n) be a sequence of functions that are all defined and differentiable on $[a,b]$, and that converge uniformly on $[a,b]$ to some a function f. Then f is differentiable on $[a,b]$ and $f'(x)=\lim_{n\to\infty}f_n'(x)$ for each $x\in[a,b]$.

Unfortunately, such a theorem is false; the uniform limit of differentiable functions on a closed interval need not be differentiable, as shown in this picture taken from Spivak (Chapter 24, Figure 8):



And Spivak has a further example showing that even if f is differentiable, it need not be the case that $f'(x) = \lim_{n\to\infty} f_n'(x)$ for each $x \in [a, b]$. Some quite careful hypothesis are needed to get an analog of Theorems 17.1 and 17.2 for differentiability. Unlike Theorem 17.2, where the hypothesis were natural but the proof involved, here the proof is very easy, given the correct hypotheses.

Theorem 17.3. *Let (f_n) be a sequence of functions that are all defined and differentiable on $[a, b]$, and that converge pointwise on $[a, b]$ to some a function f. Suppose also that each f_n' is integrable on $[a, b]$, and that the f_n' converge uniformly to* some *continuous function g. Then f is differentiable on $[a, b]$ and $f'(x) = \lim_{n\to\infty} f_n'(x)$ for each $x \in [a, b]$.*

Proof: From Theorem 17.2, then the fundamental theorem of calculus we have that for each $x \in [a, b]$,

$$\begin{aligned} \int_a^x g &= \lim_{n\to\infty} \int_a^x f_n' \\ &= \lim_{n\to\infty} (f_n(x) - f_n(a)) \\ &= f(x) - f(a). \end{aligned}$$

So

$$f(x) = f(a) + \int_a^x g,$$

and (using continuity of g) another application of the fundamental theorem of calculus gives that f is differentiable at x, and that

$$f'(x) = g(x) = \lim_{n\to\infty} f_n'(x).$$

□

17.4 Application to power series

Definition of a power series A *power series* (at, or about, 0) is a series of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$$

where x and $a_0, a_1, \ldots$ are real numbers. More generally, for a real number a a *power series* at, or about, a is a series of the form

$$a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n + \cdots .$$

At least initially, we'll only think about power series at 0. An example of a power series is the Taylor series of function f that is infinitely differentiable at 0:

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots .$$

Consistent with our previous notation, we say that a power series *converges* if the sequence $(a_n x^n)$ is summable, and *converges absolutely* if $(|a_n||x|^n)$ is summable. By definition of summability, saying that a power series converges is equivalent to saying that the sequence $(s_n(x))$ of partial sums converges, where

$$s_n(x) = \sum_{k=0}^{n} a_k x^k,$$

and saying that a power series converges absolutely is equivalent to saying that the sequence $(\bar{s}_n(x))$ of "absolute" partial sums converges, where $\bar{s}_n(x) = \sum_{k=0}^{n} |a_k||x|^k$.

Our aim now is to use the results of the last section to explore situations in which a power series converges, and the degree to which a power series can be manipulated in "natural" ways (specifically, via term-by-term integration and differentiation). Rather than stating a general theorem (which would be rather long) we will develop the theory piece by piece.

So, let $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$ be a power series, with partial sums $s_n(x) = \sum_{k=0}^{n} a_k x^k$ and "absolute" partial sums $\bar{s}_n(x) = \sum_{k=0}^{n} |a_k||x|^k$. Note that f here is not (yet) naming a function, since the power series may not be summable; it is simply convenient shorthand for the expression $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$.

Basic working hypothesis Suppose that $x_0 \geq 0$ is a real number satisfying that

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} x_0 \text{ exists and equal some } \ell < 1.$$

We begin by establishing that under this hypothesis,

for all $x \in [-x_0, x_0]$ the power series $f(x)$ converges to a finite limit.

Indeed, for $x = 0$ the claim is trivial, and for $x \neq 0$ apply the ratio test to the series $\sum_{n=0}^{\infty} |a_n x^n|$. We get that

$$\frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \left(\frac{|a_{n+1}|}{|a_n|} x_0\right) \frac{|x|}{x_0} \to \ell \frac{|x|}{x_0} < 1,$$

the existence of the limit following from the basic working hypothesis. From the ratio test we conclude that $(a_n x^n)$ is absolutely summable, and so in particular is summable, i.e., the power series is both convergent and absolutely convergent. So now f is a function with domain $[-x_0, x_0]$, and $f(x)$ is more than just shorthand for the power series; it is the value of the function.

Next we establish that

the function f is continuous on $[-x_0, x_0]$.

What we will actually show is that

the sequence of partial sums $s_n(x)$ converges *uniformly* to $f(x)$ on $[-x_0, x_0]$

(we already know that is converges pointwise). From Theorem 17.1, the continuity of f on $[-x_0, x_0]$ follows immediately. To see uniform convergence, note that for $x \in [-x_0, x_0]$ we have

$$\begin{aligned} |f(x) - s_n(x)| &= \left|\sum_{k=n+1}^{\infty} a_k x^k\right| \\ &\leq \sum_{k=n+1}^{\infty} |a_k||x|^k \\ &\leq \sum_{k=n+1}^{\infty} |a_k||x_0|^k. \end{aligned}$$

Because $(a_k x_0^k)$ is absolutely summable (we just proved that above), for every $\varepsilon > 0$, n can be chosen large enough that $\sum_{k=n+1}^{\infty} |a_k||x_0|^k < \varepsilon$,[253] and since the choice of n depends only on x_0 (and ε) it works for all $x \in [-x_0, x_0]$.

Notice that in the first inequality above we have slipped in a "triangle inequality" for series: if (c_n) is absolutely summable (so both $\sum_{n=1}^{\infty} c_n$ and $\sum_{n=1}^{\infty} |c_n|$ exist), then

$$\left|\sum_{n=1}^{\infty} c_n\right| \leq \sum_{n=1}^{\infty} |c_n|.$$

This is an easy exercise.

In summary, so far we have show that if $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$ is a power series, and if x_0 is a non-negative number for which $\lim_{n\to\infty} |a_{n+1}|x_0/|a_n|$ exists and is less than 1, then

- for each $x \in [-x_0, x_0]$, $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$ converges to a finite limit $f(x)$, and in fact converges absolutely,
- the convergence is uniform on $[-x_0, x_0]$, and
- f is continuous on $[-x_0, x_0]$.

The situation described above is often referred to as *absolute uniform convergence* of the power series; and the power series is said to converge *absolutely uniformly*.

Next, we establish that

f is differentiable on $[-x_0, x_0]$, and its derivative is given by

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots + (n+1)a_{n+1}x^n + \cdots .$$

Moreover, the sequence described above is absolutely uniformly convergent on $[-x_0, x_0]$.

[253] Formally: by the Cauchy condition, n can be chosen large enough so that for all $m > n$ we have $\sum_{k=n+1}^{m} |a_k||x_0|^k < \varepsilon$, so by the boundedness criterion, for suitably large n we have that $\sum_{k=n+1}^{\infty} |a_k||x_0|^k$ exists and is at most ε.

To verify all this, the first thing we do is check that $a_1+2a_2x+3a_3x^2+\cdots+(n+1)a_{n+1}x^n+\cdots$ is absolutely uniformly convergent on $[-x_0, x_0]$. But this follows immediately from what we have just done:

$$\lim_{n\to\infty}\frac{(n+2)|a_{n+2}|}{(n+1)|a_{n+1}|}|x_0| = \lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}|x_0| < 1$$

(by the basic working hypothesis for $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$), so the basic working hypothesis is satisfied for $a_1 + 2a_2x + 3a_3x^2 + \cdots + (n+1)a_{n+1}x^n + \cdots$, and we can apply all the results we have proven about $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$ to $a_1 + 2a_2x + 3a_3x^2 + \cdots + (n+1)a_{n+1}x^n + \cdots$.

So now we know that the sequence $(t_n(x))$ of partial sums of $a_1 + 2a_2x + 3a_3x^2 + \cdots + (n+1)a_{n+1}x^n + \cdots$ (where $t_n(x) = \sum_{k=0}^n (k+1)a_{k+1}x^k$) converges uniformly to a limit that is continuous on $[-x_0, x_0]$. But $t_n(x) = s'_{n+1}(x)$, where recall $s_n(x) = \sum_{k=0}^n a_kx^k$ is the partial sum of $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$. It follows that $(s_n(x))$ converges uniformly to a limit that is continuous on $[-x_0, x_0]$. Now also (s_n) is a sequence of functions that are all defined and differentiable on $[-x_0, x_0]$, and that converge pointwise (in fact, uniformly) on $[-x_0, x_0]$ to f. And finally, clearly each s'_n is integrable on $[-x_0, x_0]$. So all the hypotheses of Theorem 17.3 are satisfied, and we conclude that f is differentiable on $[-x_0, x_0]$ and

$$f'(x) = \lim_{n\to\infty} s'_n(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots + (n+1)a_{n+1}x^n + \cdots$$

for each $x \in [-x_0, x_0]$.

This whole argument can be repeated of f'; we conclude that f is *infinitely* differentiable on $[-x_0, x_0]$, and that the successive derivatives can be found by differentiating the power series term-by-term.

Finally, we turn to integrability.

> f is integrable on $[-x_0, x_0]$, and the function $g : [-x_0, x_0] \to \mathbb{R}$ defined by $g(x) = \int_0^x f(t)\ dt$ is given by
>
> $$g(x) = a_0x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \cdots + \frac{a_{n-1}}{n}x^n + \cdots.$$
>
> Moreover, the sequence described above is absolutely uniformly convergent on $[-x_0, x_0]$.

The absolute uniform convergence of the sequence is verified exactly as in the case of differentiation. For the rest, note that since $(s_n(x))$ converges uniformly to f on $[-x_0, x_0]$, we can apply Theorem 17.3 to conclude that f is integrable on $[-x_0, x_0]$[254] and that

$$\begin{aligned}\int_0^x f(t)\ dt &= \lim_{n\to\infty}\int_0^x s_n(t)\ dt \\ &= \lim_{n\to\infty}\int_0^x \left(a_0x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \cdots + \frac{a_n}{n+1}x^{n+1}\right) \\ &= a_0x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \cdots + \frac{a_{n-1}}{n}x^n + \cdots.\end{aligned}$$

[254]We already knew that f was integrable, since it is continuous. But there is no harm in discovering this fact by a second route.

This whole argument can be repeated of g; we conclude that f is *infinitely* integrable on $[-x_0, x_0]$[255], and that the successive integrals can be found by integrating the power series term-by-term.

Time for a second (and final) summary: if $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$ is a power series, and if x_0 is a non-negative number for which $\lim_{n\to\infty} |a_{n+1}|x_0/|a_n|$ exists and is less than 1, then

- for each $x \in [-x_0, x_0]$, $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$ converges absolutely uniformly to a finite limit $f(x)$,
- f is continuous on $[-x_0, x_0]$,
- f is differentiable arbitrarily many times on $[-x_0, x_0]$; each derivative is found by differentiating the power series term-by-term, and the resulting power series is still absolutely uniformly convergent on $[-x_0, x_0]$,
- $g(x) = \int_0^x f(t)\ dt$ is defined on $[-x_0, x_0]$; it is found by integrating the power series term-by-term, and the resulting power series is still absolutely uniformly convergent on $[-x_0, x_0]$ (so the process can be repeated arbitrarily many times).

We refer to all this as the power series being "nice".

Of course, we would like the power series to be nice for as large an x_0 as possible. If $\lim_{n\to\infty} |a_{n+1}|/|a_n|$ doesn't exist, then there is no x_0 (other than $x_0 = 0$) for which $\lim_{n\to\infty} |a_{n+1}|x_0/|a_n|$ exists. So in this case, we will say nothing.

If $\lim_{n\to\infty} |a_{n+1}|/|a_n| = 0$ then for *all* x_0 we have $\lim_{n\to\infty} |a_{n+1}|x_0/|a_n| = 0 < 1$. So in this case, the power series is nice on the whole real line.

If $\lim_{n\to\infty} |a_{n+1}|/|a_n| = 1/R > 0$, then as long as $x_0 < R$, the power series is nice on $[-x_0, x_0]$. On the other hand, it is easy to see (by the ratio test, or just by the fact that the nth term does not go to zero) that if $|x_0| > R$ then the power series does not converge.

If $\lim_{n\to\infty} |a_{n+1}|/|a_n| = +\infty$ then only for $x_0 = 0$ do we have $\lim_{n\to\infty} |a_{n+1}|x_0/|a_n| < 1$, and again it is easy to see (by the fact that the nth term does not go to zero) that if $x_0 \neq 0$ then the power series does not converge. So in this case, the power series is nice only at 0.

So, in summary:

> if $\lim_{n\to\infty} |a_{n+1}|/|a_n| = 1/R > 0$, then the power series is nice on every closed interval contained in $(-R, R)$, but not on any closed interval that includes anything outside $[-R, R]$.[256] R is referred to as the *radius of convergence* of the power series. If $\lim_{n\to\infty} = 0$, then the power series is nice on every closed interval contained in the reals; in this case we say that the radius of convergence is infinite. If $\lim_{n\to\infty} |a_{n+1}|/|a_n| = +\infty$ then the power series is nice only at 0; in this case we say that the radius of convergence is 0. In all three cases, the power series has

[255] No surprise — unlike with differentiation, once a function is integrable once, it is infinitely integrable

[256] What happens at R and $-R$ has to be dealt with on a case-by-case basis.

a radius of convergence, inside of which it is nice, and outside of which it does not converge.

If $\lim_{n\to\infty} |a_{n+1}|/|a_n|$ does not exist, then the power series might still have a radius of convergence; that is a story for another day (year?) involving the notion of $\limsup$.

For power series at values a other than zero, that is, power series of the form

$$a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n + \cdots,$$

the story is essentially the same: with R defined exactly as before, the power series is nice on every closed interval contained in $(a-R, a+R)$; the proof is almost exactly the same, just a little bit more annoying as it involves the extra parameter a.

If a power series $a_0 + a_1x + a_2x^2 + \cdots$ has some radius of convergence R, what can we say about the function $f(x)$ that the power series converges to, inside $(-R, R)$? Well,

- $f(0) = a_0$.
- $f'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \cdots$ so $f'(0) = a_1$.
- $f''(x) = 2a_2 + 6a_3x + 12a_4x^2 + \cdots$ so $f''(0) = 2a_2$ or $a_2 = f''(0)/2$.
- $f'''(x) = 6a_3 + 24a_4x + \cdots$ so $f'''(0) = 6a_3$ or $a_3 = f'''(0)/6 = f'''(0)/3!$.
- In general, $a_n = f^{(n)}(0)/n!$

In other words:

> If a power series $a_0 + a_1x + a_2x^2 + \cdots$ has some radius of convergence R, and so converges to a function $f(x)$ inside $(-R, R)$, then $a_0 + a_1x + a_2x^2 + \cdots$ is in fact the Taylor series of f. In particular, that means that is any *other* power series converges to f on $(-R, R)$, that other power series must be identical to $a_0 + a_1x + a_2x^2 + \cdots$.

We finish up this section with some examples.

- $\exp, \sin, \cos$: we already know, from Taylor's theorem, that
  - $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ for all real x,
  - $\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ for all real x,
  - $\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ for all real x,

  and it is easy to apply the theory we have developed to show that each of these power series are absolutely uniformly convergent on every closed interval in the reals (so, radius of convergence is $+\infty$ in each case). This allows us to confirm, by term-by-term differentiation, that all the well-known relations between these functions hold, such as

– the derivative of e^x is e^x,

– an antiderivative of cos is sin, and

– $\sin'' + \sin = 0$, $\cos'' + \cos = 0$.

Recall that we proved that the only continuous functions $f : \mathbb{R} \to \mathbb{R}$ satisfying $f' = f$ are functions of the form $f(x) = ae^x$, so in particular the only $f : \mathbb{R} \to \mathbb{R}$ satisfying $f' = f$ and $f(0) = 1$ is $f(x) = e^x$; and that the only continuous functions $f : \mathbb{R} \to \mathbb{R}$ satisfying $f'' + f = 0$ are functions of the form $f(x) = a \sin x + b \cos x$, so in particular the only $f : \mathbb{R} \to \mathbb{R}$ satisfying $f'' + f =$ and $f(0) = 0$, $f'(0) = 1$ is $f(x) = \sin x$, while the only $f : \mathbb{R} \to \mathbb{R}$ satisfying $f'' + f =$ and $f(0) = 1$, $f'(0) = 0$ is $f(x) = \cos x$. But from the theory we have just developed, we can conclude (without ever knowing anything about the exponential function) that if $f(x)$ is defined by

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

then $f' = f$ and $f(0) = 1$; so we could use this as an alternate *definition* of the exponential function. Similarly, we could formally define sin and cos via power series. Many authors take this approach.

- log: We can easily check that the power series

$$f(x) = 1 + x + x^2 + \cdots + x^n + \cdots$$

has radius of convergence 1. We also already know that $f(x) = 1/(1 - x)$ for all $x \in (-1, 1)$. So, substituting $-x$ for x, we have

$$\frac{1}{1+x} = 1 - x + x^2 - \cdots + (-1)^n x^n + \cdots .$$

By the theory we have just developed, we have

$$\int_0^x \frac{dt}{1+t} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots ,$$

valid for all x between -1 and 1. But $\int_0^x \frac{dt}{1+t} = \log(1 + x)$, and the power series on the right above is the Taylor series at 0 of $\log(1+x)$; so we have just shown that the Taylor series of $\log(1 + x)$ converges to $\log(1 + x)$, that is

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

for $x \in (-1, 1)$. Recall from a homework exercise that it was only easy to analyze the remainder term of the Taylor polynomial, and obtain the above result, for $-1/2 < x < 1$. The theory of uniform convergence allows us to easily fill in the gap.

- Estimating integrals. Power series can be used to estimate integrals. For example,

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots$$

(just evaluate the series for e^x at $-x^2/2$; since the series is nice for all real x, it is nice for $-x^2/2$). The above series for e^{-x^2} is easily seen to have infinite radius of convergence, so we can write

$$\int_0^x e^{-t^2}\, dt = x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7.3!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)n!} + \cdots .$$

For example, what is $\int_0^1 e^{-t^2}\, dt$? Setting $x = 1$ above, and using Leibniz test for alternating series (in the strong form that gives an error estimate of the magnitude of the next term after truncation) we get that for each n

$$\int_0^1 e^{-t^2}\, dt \left(\sum_{k=0}^{n} (-1)^k \frac{1}{(2k+1)k!} \right) \pm \frac{1}{(2n+3)(n+1)!}$$

Taking $n = 10$ we get

$$\int_0^1 e^{-t^2}\, dt = 0.746824133 \pm 0.000000001.$$

- More general series. What if $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$ doesn't exist? If $(|a_{n+1}|/|a_n|)$ is eventually bounded above by $M > 0$, then whole theory described above goes through, as long as $Mx_0 < 1$[257].

  For an example, consider the Fibonacci numbers, defined by $F_0 = 0, F_1 = 1$ and $F_n = F_{n_1} + F_{n-2}$ for $n \geq 2$. We can form the "generating function" of the Fibonacci numbers:

$$F(x) = F_0 + F_1 x + F_2 x^2 + \cdots .$$

  Since (F_n) is increasing we have

$$\frac{F_{n+1}}{F_n} \leq \frac{2F_n}{F_n} \leq 2,$$

  and so $F(x)$ converges absolutely uniformly on some interval around 0 (for example, on $[-0.49, 0.49]$). We have

$$\begin{aligned} F(x) &= F_0 + F_1 x + F_2 x^2 + \cdots \\ &= x + (F_1 + F_0)x^2 + (F_2 + F_1)x^3 + \cdots \\ &= x + x(F_1 x + F_2 x^2 + \cdots) + x^2(F_0 + F_1 x + F_2 x^2 + \cdots) \\ &= x + xF(x) + x^2 F(x) \end{aligned}$$

[257] Exercise!

(with all manipulations easily seen to be valid inside the interval $[-0.49, 0.49]$) so

$$\begin{aligned} F(x) &= \frac{x}{1-x-x^2} \\ &= \frac{A}{1-\varphi_1 x} + \frac{B}{\varphi_2 x}, \quad \text{(partial fractions)} \end{aligned}$$

where $(1-\varphi_1 x)(1-\varphi_2 x) = 1-x-x^2$, so $\varphi_1 + \varphi_2 = 1$, $\varphi_1\varphi_2 = -1$, so

$$\varphi_1 = \frac{1+\sqrt{5}}{2}, \quad \varphi_2 = \frac{1-\sqrt{5}}{2}$$

and A, B are constants. So

$$F_n = A\varphi_1^n + B\varphi_2^n.$$

(This is obtained by using the obvious power series for $1/(1-\varphi_1 x)$ and $1/(1-\varphi_2 x)$). At $n = 0$ we get $A + B = 0$, and at $n = 1$ we get $A\varphi_1 + B\varphi_2 = 1$, so

$$A = \frac{1}{\sqrt{5}}, \quad B = \frac{-1}{\sqrt{5}},$$

and so finally we get the remarkable *Binet's formula*:

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$$

All manipulations were valid because all power series involved converge in some non-zero interval around 0, and everything can be done inside an interval of convergence that works for all power series involved.

There is a vast literature on power series; we have only scratched the surface. Hopefully we have done enough to see that power series can be a very useful way of viewing functions.