Power Series

James S. Cook
Liberty University
Department of Mathematics

Spring 2023

Abstract

This article studies power series and their application to calculus in depth. We begin by defining a power series as a function which has a very particular formula:

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_o)^n = a_0 + a_1(x - x_o) + a_2(x - x_o)^2 + \cdots$$

We say x_o is the **center** of the power series and $a_0, a_1, \dots$ are the coefficients of the power series. The center and the coefficients uniquely define the power series. Our first examples of power series are based on the function $f(x) = \frac{1}{1-x}$. We take a whole day to appreciate how the geometric series result and algebra allows us to find power series which represent that function. Then we turned to work on how to understand the domain of a power series. We see how the ratio test determines the domain of the power series up to a couple points. In particular, the domain is a connected subset of real numbers; that is, the domain is an **interval**. Thus we speak of the **Interval Of Convergence** (IOC) which includes all real numbers within $\pm R$ of x_o as well as the possible inclusion of the endpoints. When $R = \infty$ the $IOC = (-\infty, \infty)$ and when $R = 0$ the IOC is just $\{x_o\}$. The Cauchy Hadamard Theorem is also presented, however this is not part of the usual required curriculum of Calculus II. Essentially, the Cauchy Hadamard Theorem is an extension of the root test as it applies to power series and the formulation in terms of subsequences naturally allows for series with infinitely many zero terms.

Next we Term-by-term calculus for power series is described and we find how to use geometric series indirectly to expand our set of examples. Of course not everything is related to the geometric series by some trickery, we need other tools, hence we turn to the discussion of Taylor Polynomials. Our initial section presents Taylor's Theorem for polynomials. We find a generalization of the mean value theorem which bounds the error in the approximation for a function by its n-th Taylor polynomial in terms of the values of the $(n + 1) - th$ derivative. Numerous graphs illustrating the success of this bound are given and applications to physics are discussed. In short, this is how power series are typically used, we replace a complicated nonlinear function with a polynomial in some variable which is small for physical reasons. That trade is an incredible simplification in practice. Many problems in physics and engineering are only solved in the context of such an approximation. Taylor series are defined. A function is **analytic at** x_o if the Taylor series generated at x_o represents the function faithfully on some open interval centered at x_o. We derive the Binomial series as well as the Maclaurin series for the exponential, sine, cosine, hyperbolic sine and cosine and a host of other functions. Special attention is given to the use of substitution to create new series from old. In addition, multiplication and division of series gives another set of tools to create new power series from old. Finally we turn to applications of power series to calculate otherwise intractable integrals as well as to solve differential equations. Applications to limits are also discussed in lecture (but at this time these notes lack examples on the limits)

1 Power Series and their Domains

An *algebraic function* is a function whose formula is given by an algebraic formula. In similar fashion, a power series is a function whose formula is given in terms of a series of power functions.

Definition 1.1.

A function whose formula is given point-wise by a series according to the rule

$$f(x) = \sum_{n=0}^{\infty} c_n(x - x_o)^n = c_0 + c_1(x - x_o) + c_2(x - x_o)^2 + \cdots$$

where $c_n, x_o \in \mathbb{R}$ for $n = 0, 1, \ldots$. We call x_o the **center** of the power series and c_n the n-th **coefficient** of the power series.

The next example shows that we can use the geometric series with a variable in order to identify a given function as a power series.

Example 1.2. *Consider the function* $f_1(x) = \frac{1}{1-x}$ *with domain* $(-1, 1)$*. If* $|x| < 1$ *then by the geometric series we have*

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}.$$

Thus $f_1(x) = 1 + x + x^2 + x^3 + \cdots$ *is a power series. Observe* $x_o = 0$ *is the center of the series and the coefficient* $c_n = 1$ *for each* $n = 0, 1, 2, \ldots$.

For a given power series it may not be possible to find a nice closed-form formula as we saw in the above example. In fact, using power series expands our class of functions well past the narrow lexicon we currently employ. There are many functions which can be represented as a power function and yet fail to have any formula in terms of elementary functions[1]. At the moment, all we have to work with is the geometric series. I'll show you another one of my favorite tricks:

Example 1.3. *Consider the function* $f_2(x) = \frac{1}{1-x}$ *with domain* $(1, 3)$*. Notice* $1 < x < 3$ *implies* $-1 < x - 2 < 1$ *or* $|x - 2| < 1$*. We can apply the geometric series to* $r = \pm(x - 2)$ *since* $|r| = |x - 2| < 1$*. Calculate,*

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

Identify the result above as the sum of a geometric series with $r = -(x - 2)$ *and* $c = -1$ *thus*

$$f_2(x) = \frac{-1}{1+(x-2)} = \sum_{n=0}^{\infty}(-1)(-(x-2))^n = \sum_{n=0}^{\infty}(-1)^{n+1}(x-2)^n.$$

We find the given function $f_2(x)$ *can be viewed as a power series with center* $x_o = 2$ *and coefficients* $c_n = (-1)^{n+1}$ *for* $n = 0, 1, 2, \ldots$.

Great, let's try the same example yet again, once more I'll change the domain to anticipate the calculation I'm about to complete.

[1] elementary functions include polynomials, rational functions, algebraic functions, trigonometric functions, hyperbolic functions, exponentials, logs and products, sums, quotients and composites of all of the cases just listed. It's a lot, but power series give you much much more

Example 1.4. *Consider the function* $f_3(x) = \dfrac{1}{1-x}$ *with domain* $(1,5)$*. Notice* $1 < x < 5$ *implies* $-2 < x-3 < 2$ *or* $|x-3| < 2$ *hence* $|x-3|/2 < 1$*. We can apply the geometric series to* $r = \pm(x-2)/2$ *since* $|r| = |x-2|/2 < 1$*. Calculate,*

$$\frac{1}{1-x} = \frac{1}{1-(x-3+3)} = \frac{-1}{2+(x-3)} = \frac{-1/2}{1+(x-3)/2}$$

Identify the result above as the sum of a geometric series with $r = -(x-3)/2$ *and* $c = -1/2$ *thus*

$$f_3(x) = \frac{-1/2}{1+(x-3)/2} = \sum_{n=0}^{\infty}(-1/2)(-(x-2)/2)^n = \sum_{n=0}^{\infty}\left(\frac{-1}{2}\right)^{n+1}(x-3)^n.$$

We find the given function $f_3(x)$ *can be viewed as a power series with center* $x_o = 3$ *and coefficients* $c_n = \left(\frac{-1}{2}\right)^{n+1}$ *for* $n = 0, 1, 2, \dots$*.*

Notice all three examples we've considered simply look at power series which represent the function $f(x) = \frac{1}{1-x}$ for particular restrictions of the domain. If we begin with $f(x) = \frac{1}{1-x}$ with its natural domain $(-\infty, 1) \cup (1, \infty)$ then we may study how to represent the function **locally** by a power series centered at some point.

Example 1.5. *Consider the function* $f(x) = \dfrac{1}{1-x}$*. Suppose we wish to find the power series centered at* x_o *for the given function. Let's put my trick into play and see where it leads us:*

$$\begin{aligned}\frac{1}{1-x} &= \frac{1}{1-x_o-(x-x_o)} \\ &= \frac{1}{1-x_o}\cdot\frac{1}{1-\frac{x-x_o}{1-x_o}} \\ &= \frac{1}{1-x_o}\sum_{n=0}^{\infty}\left(\frac{x-x_o}{1-x_o}\right)^n \\ &= \sum_{n=0}^{\infty}\frac{1}{(1-x_o)^{n+1}}(x-x_o)^n\end{aligned}$$

where we identified the sum of a geometric series with $c = 1$ *and* $r = \frac{x-x_o}{1-x_o}$*. Notice* $|r| < 1$ *implies* $|x-x_o| < |1-x_o|$ *is the necessary condition for the convergence of the geometric series. Furthermore, there is an obvious reason why this had to happen. Notice that* $|1-x_o|$ *is the distance from* x_o *to* 1*. Thus the condition* $|x-x_o| < |1-x_o|$ *literally means that the point* x *must be closer to the center point* x_o *than the vertical asymptote* $x = 1$ *for* $f(x)$*.*

Remark 1.6.

In retrospect, we can understand why I gave the domains that I did for f_1, f_2 and f_3 earlier in this section. Those were the largest domains I could hope for a power series to represent $f(x) = \frac{1}{1-x}$ for the respective centers of $x_o = 0, 2, 3$.

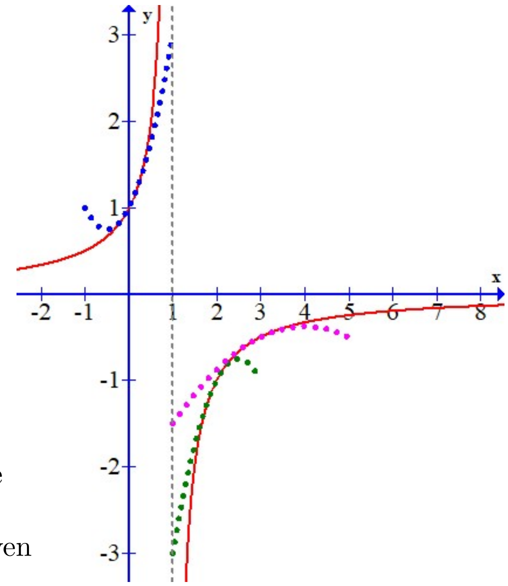
There is much more to learn about manipulating the geometric series to create interesting power series. We will do many more such calculations. That said, I now turn to appreciate visually the meaning of the examples.

Suppose we keep just the first three terms for the first three examples for $f(x) = \dfrac{1}{1-x}$. We can graph the truncated power series to gain some understanding of the situation:

(1.) $y = 1 + x + x^2$
represents $f(x)$ near $x_o = 0$
where $-1 < x < 1$ (blue dots),

(2.) $y = -1 + (x-2) - (x-2)^2$
represents $f(x)$ near $x_o = 2$
where $1 < x < 3$ (green dots),

(3.) $y = -\dfrac{1}{2} + \dfrac{1}{4}(x-3) - \dfrac{1}{8}(x-3)^2$
represents $f(x)$ near $x_o = 3$
where $1 < x < 5$ (magenta dots).



Notice that the quadratic polynomials are closest to the real graph of $y = f(x)$ for values of x which are closest to the center of the power series. Also, notice that a given function $f(x)$ can have many different power series representations. In fact, Example 1.5 shows we can find a power series for $f(x)$ about **any** point $x_o \neq 1$. If we kept more terms then we would see even closer agreement between the function $f(x)$ and the power series which represents the function.

The domain of a power series is a connected subset of real numbers. In particular, we prove this by a simple application of the ratio test.

Theorem 1.7.

Let $f(x) = \displaystyle\sum_{n=0}^{\infty} a_n(x-x_o)^n$ then the domain is one of the following:

(i.) $\{x_o\}$,

(ii.) $[x_o - R, x_o + R]$, $(x_o - R, x_o + R]$, $[x_o - R, x_o + R)$, $(x_o - R, x_o + R)$,

(iii.) $(-\infty, \infty)$.

Proof: I assume $a_n \neq 0$ for all $n \geq 0$ for brevity of argument. Observe the n-th term of the series which defines $f(x)$ is $a_n(x-x_o)^n$ thus consider the ratio

$$\rho = \lim_{n\to\infty} \left| \frac{a_{n+1}(x-x_o)^{n+1}}{a_n(x-x_o)^n} \right| = |x - x_o| \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

If $x = x_o$ then $|x - x_o| = |0| = 0$ and we find $\rho = 0$ thus the series converges and $x_o \in \text{dom}(f)$ for any choice of coefficients a_n. If $\displaystyle\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty$ and $x \neq x_o$ then $\rho = \infty$ and we find the series diverges by the ratio test. If $\displaystyle\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = 0$ then $\rho = 0$ and we find $\text{dom}(f) = (-\infty, \infty)$ by the ratio test. Finally, if $\displaystyle\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = L \neq 0$ then $\rho = |x - x_o|L < 1$ if and only if $|x - x_o| < 1/L$. Identify $R = 1/L$ and note $(x_o - R, x_o + R) \subseteq \text{dom}(f)$ by the ratio test. $\square$

Definition 1.8. *Interval and radius of convergence for power series.*

The domain of a power series $\sum_{n=0}^{\infty} a_n(x-x_o)^n$ is known as the **interval of convergence (IOC)**. We denote the **radius of convergence (ROC)** by R. If $IOC = (-\infty, \infty)$ then we say $R = \infty$. If the IOC has the form $[x_o - R, x_o + R]$ or $(x_o - R, x_o + R]$ or $[x_o - R, x_o + R)$ or $(x_o - R, x_o + R)$ then we say the radius of convergence is R. Finally, if the IOC is a single point then $R = 0$.

The proof of the Corollary below is implicit within the proof of Theorem 1.7.

Corollary 1.9.

Let $f(x) = \sum_{n=0}^{\infty} a_n(x-x_o)^n$ where $a_n \neq 0$ for all n then if $\lim_{n\to\infty} |a_{n+1}/a_n|$ is nonzero and finite then the radius of convergence is given by $R = \lim_{n\to\infty} |a_n/a_{n+1}|$.

Remark 1.10.

The uninterested reader is encouraged to skip ahead to Example 1.14. The material after this remark is arguably what is commonly considered Calculus II.

We commonly face examples where all the even or odd coefficients are zero, so the proof given for Theorem 1.7 would not directly apply. It is fairly clear we can modify the arguments above for series of the form $\sum_{k=0}^{\infty} a_{2k}x^{2k}$ or $\sum_{k=0}^{\infty} a_{2k+1}x^{2k+1}$ and we will derive essentially the same results. However, in the case the limit is finite the relation between the limit of subsequent coefficients and the radius of convergence is a bit different. For instance, $\sum_{k=0}^{\infty} a_{2k}x^{2k}$ suggests we study ratio test:

$$\rho = \lim_{k\to\infty} \left| \frac{a_{2k+2}(x-x_o)^{2k+2}}{a_{2k}(x-x_o)^{2k}} \right| = |x-x_o|^2 \lim_{k\to\infty} \left| \frac{a_{2k+2}}{a_{2k}} \right|$$

If $\lim_{k\to\infty} \left| \frac{a_{2k+2}}{a_{2k}} \right| = L$ then the condition for convergence of the power series is $\rho = |x-x_o|^2 L < 1$ hence $|x-x_o|^2 < 1/L$ which means $|x-x_o| < 1/\sqrt{L}$ hence $R = 1/\sqrt{L}$. Of course, we could also imagine power series such as:

$$\sum_{k=0}^{\infty} a_{3k}(x-x_o)^{3k}, \qquad \text{or} \qquad \sum_{k=0}^{\infty} a_{3k+1}(x-x_o)^{3k+1}, \qquad \text{or} \qquad \sum_{k=0}^{\infty} a_{3k+2}(x-x_o)^{3k+2}$$

where $\lim_{k\to\infty} \left| \frac{a_{3k+3}}{a_{3k}} \right| = L$ would imply $R = 1/\sqrt[3]{L}$. The possibilities are endless. The general formula for the radius must somehow account for all the myriad of possible subsequential variants. The right tool for the job is something called the limit superior

Definition 1.11. *limit superior*

If E is the set of all subsequential limits of $\{a_n\}$ then we define $\limsup_{n\to\infty}(a_n) = \sup(E)$ where $\sup(E)$ is the least upper bound of E which is defined to be ∞ in the case E is unbounded above. Similarly, $\liminf_{n\to\infty}(a_n) = \inf(E)$ where $\inf(E)$ is the greatest lower bound of E which is defined to be $-\infty$ in the case E is unbounded below.

If the limit of a sequence exists then $\liminf_{n\to\infty}(a_n) = \lim_{n\to\infty}(a_n) = \limsup_{n\to\infty}(a_n)$. The limits superior and inferior always exist, for example, $a_n = (-1)^n$ has $\limsup_{n\to\infty}(a_n) = 1$ and $\liminf_{n\to\infty}(a_n) = -1$. A more complete discussion of the limit superior and inferior belong to a different course.

Given the trouble with zero in the ratio test, it is natural to utilize the root test as an alternative method of study. Consider $f(x) = \sum_{n=0}^{\infty} a_n(x - x_o)^n$ and apply the root test to discern the interval of convergence:

$$L = \lim_{n\to\infty} \sqrt[n]{|a_n(x - x_o)^n|} = |x - x_o| \lim_{n\to\infty} \sqrt[n]{|a_n|} < 1$$

implies $|x - x_o| < 1/\lim_{n\to\infty} \sqrt[n]{|a_n|}$ hence we find the radius of convergence is given by $R = 1/\lim_{n\to\infty} \sqrt[n]{|a_n|}$. The calculation above can be refined to allow for series where the $\lim_{n\to\infty} \sqrt[n]{|a_n|}$ does not exist, but the proof is beyond this course. I include this result for the curious:

Theorem 1.12. *Cauchy-Hadamard Theorem*

Let $f(x) = \sum_{n=0}^{\infty} a_n(x - x_o)^n$ then the radius of convergence is given by:

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}$$

where we understand $R = \infty$ in the case $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0$ and $R = 0$ in the case $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \infty$.

The proof is sometimes given in Complex Analysis, this theorem was known to Cauchy in 1821, but Hadamard rediscovered the result in 1888 after which time the theorem was more widely known.

Example 1.13. *Consider* $\sum_{n=1}^{\infty} \frac{1}{(\tan^{-1}(n))^n}(x - x_o)^n$. *I'll apply the root test,*

$$\sqrt[n]{\left|\frac{1}{(\tan^{-1}(n))^n}(x - x_o)^n\right|} = \frac{|x - x_o|}{\tan^{-1}(n)}$$

thus as $n \to \infty$ *we have* $\tan^{-1}(n) \to \pi/2$ *and the root test gives convergence provided* $\frac{2|x-x_o|}{\pi} < 1$ *which is equivalent to* $|x - x_o| < \frac{\pi}{2}$. *We find the radius of convergence is* $R = \pi/2$.

Example 1.14. *Consider* $f(x) = \sum_{n=0}^{\infty} \frac{2^n}{n+1}(x + 3)^n$. *Identify* $c_n = \frac{2^n}{n+1}(x + 3)^n$ *and study,*

$$\lim_{n\to\infty} \left|\frac{c_{n+1}}{c_n}\right| = \lim_{n\to\infty} \frac{2^{n+1}(n + 1)|x + 3|^{n+1}}{2^n(n + 2)|x + 3|^n} = |x + 3| \lim_{n\to\infty} \frac{2(n + 1)}{n + 2} = 2|x + 3| < 1$$

yields $|x + 3| < \frac{1}{2}$. *We find* $R = 1/2$ *and the* **open interval of convergence** *is given by* $(-3 - 1/2, -3 + 1/2) = (-7/2, -5/2)$. *It remains to decide if* $f(-7/2)$ *or* $f(-5/2)$ *are convergent. Consider,*

$$f(-7/2) = \sum_{n=0}^{\infty} \frac{2^n}{n + 1}\left(\frac{1}{-2}\right)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{n + 1}$$

converges by the alternating series test. Likewise, $f(-5/2) = \sum_{n=0}^{\infty} \frac{1}{n+1}$ *is the divergent* $p = 1$ *series. Thus* $IOC = [-7/2, -5/2)$.

Example 1.15. *Consider* $f(x) = \sum_{n=1}^{\infty} \frac{5^n}{n^2} x^{3n}$*. Identify* $c_n = \frac{5^n}{n^2} x^{3n}$ *and study,*

$$\lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right| = \lim_{n\to\infty} \frac{5^{n+1} n^2 |x|^{3(n+1)}}{5^n (n+1)^2 |x^{3n}|} = |x|^3 \lim_{n\to\infty} \frac{5n^2}{(n+1)^2} = 5|x|^3 < 1$$

yields $|x| < \frac{1}{\sqrt[3]{5}}$*. We find* $R = \frac{1}{\sqrt[3]{5}}$ *and the* **open interval of convergence** *is given by* $\left(-\frac{1}{\sqrt[3]{5}}, \frac{1}{\sqrt[3]{5}}\right)$*. It remains to decide if* $f(-1/\sqrt[3]{5})$ *or* $f(1/\sqrt[3]{5})$ *are convergent. Consider,*

$$f\left(-\frac{1}{\sqrt[3]{5}}\right) = \sum_{n=1}^{\infty} \frac{5^n}{n^2} \left(-\frac{1}{\sqrt[3]{5}}\right)^{3n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \qquad \& \qquad f\left(\frac{1}{\sqrt[3]{5}}\right) = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

are convergent series by $p = 2$ *series and the alternating series test. Thus* $IOC = \left[-\frac{1}{\sqrt[3]{5}}, \frac{1}{\sqrt[3]{5}}\right]$.

Example 1.16. *Consider* $f(x) = \sum_{j=0}^{\infty} \frac{3^j}{j!} x^j$*. Identify* $c_j = \frac{3^j}{j!} x^j$ *and study,*

$$\lim_{j\to\infty} \left| \frac{c_{j+1}}{c_j} \right| = \lim_{j\to\infty} \left| \frac{3^{j+1}|x|^{j+1}}{(j+1)!} \cdot \frac{j!}{3^j |x|^j} \right| = |x| \lim_{j\to\infty} \frac{3}{j+1} = 0$$

hence the series converges by the ratio test for any x*. Therefore* $IOC = (-\infty, \infty)$ *and* $R = \infty$.

Example 1.17. $f(x) = \sum_{n=1}^{\infty} n^n (x-2)^n$*. Identify* $c_n = n^n (x-2)^n$ *and study,*

$$\begin{aligned} \lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right| &= \lim_{n\to\infty} \left| \frac{(n+1)^{n+1}|x-2|^{n+1}}{n^n |x-2|^n} \right| \\ &= |x-2| \lim_{n\to\infty} \frac{(n+1)(n+1)^n}{n^n} \\ &= |x-2| \lim_{n\to\infty} (n+1)\left(1 + \frac{1}{n}\right)^n \end{aligned}$$

Thus the series diverges by the ratio test if $x \neq 2$ *since* $\lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right| = \infty$*. In contrast, if* $x = 2$ *then* $\lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right| = 0$ *so the series converges. Thus the* $IOC = \{2\}$ *and* $R = 0$.

Example 1.18. *Consider* $f(x) = \sum_{n=0}^{\infty} 2^{3-2n}(x-7)^n$ *has n-th term* $c_n = 2^{3-2n}(x-7)^n$*. Calculate,*

$$\frac{c_{n+1}}{c_n} = \frac{2^{3-2(n+1)}(x-7)^{n+1}}{2^{3-2n}(x-7)^n} = \frac{2^{-2} \cdot 2^{3-2n}(x-7)^n (x-7)}{2^{3-2n}(x-7)^n} = \frac{x-7}{4}$$

Identify $f(x) = \sum_{n=0}^{\infty} 8 \left(\frac{x-7}{4}\right)^n$ *is geometric series with* $c = 8$ *and* $r = \frac{x-7}{4}$ *thus* $x \in IOC$ *if and only if* $|r| < 1$*. We find* $IOC = (3, 11)$ *where* $R = 4$*. Notice we do not need to check the endpoints for this example because the geometric series result gives us the entire IOC.*

Remark 1.19.

In summary, to find the IOC for a given power series we identify the n-th term c_n then calculate $|c_{n+1}/c_n|$. If the quotient $c_{n+1}/c_n = r$ where r has no n-dependence then the given power series can be identified as a geometric series and the IOC is easily found from the condition $|r| < 1$. Otherwise, we apply the ratio test and we may also need to check the endpoints to decide the precise form of the IOC. If I don't want the students to check endpoints then I'll ask for them to decide the largest open interval of convergence.

Let me conclude our study of power series domain with an example which troubled me when I first found it. At first glance it seems to contradict the domain theorem we stated earlier in this section. However, the resolution of this paradox is simple, the series given in the example below is **not** a power series. I'm not sure what the name is for it, but it serves to illustrate the fact that domains for series formed from other functions need not follow the simple structure of power series domains.

Example 1.20. *Consider* $f(x) = \dfrac{1}{x^2} = \dfrac{1}{1+x^2-1}$. *Apply the geometric series result viewing the sum of a geometric series with* $c = 1$ *and* $r = 1 - x^2$. *If* $|r| < 1$ *we find:*

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

The expression above holds for $|x^2-1| < 1$. *In particular,* $|x^2-1| < 1$ *implies* $-1 < x^2 - 1 < 1$ *hence* $0 < x^2 < 2$ *thus* $0 < |x| < \sqrt{2}$. *Thus the expression holds on* $(-\sqrt{2}, 0) \cup (0, \sqrt{2})$.

2 Calculus on Power Series

The proof of the derivative rule of theorem which follows is technical and I omit the proof.[2]

Theorem 2.1. *Differentiation and integration of power series*

(1.) $\displaystyle \frac{d}{dx}\sum_{n=0}^{\infty} a_n(x-x_o)^n = \sum_{n=0}^{\infty} n a_n(x-x_o)^{n-1}$,

(2.) $\displaystyle \int \left(\sum_{n=0}^{\infty} a_n(x-x_o)^n\right) dx = C + \sum_{n=0}^{\infty}\frac{a_n}{n+1}(x-x_o)^{n+1}$.

Moreover, $\sum_{n=0}^{\infty} a_n(x-x_o)^n$, $\sum_{n=0}^{\infty} n a_n(x-x_o)^{n-1}$ and $\sum_{n=0}^{\infty}\frac{a_n}{n+1}(x-x_o)^{n+1}$ have the same radius of convergence.

Proof: to see that **(2.)** follows from **(1.)** we need only differentiate using **(1.)**,

$$\frac{d}{dx}\left[C + \sum_{n=0}^{\infty}\frac{a_n}{n+1}(x-x_o)^{n+1}\right] = \sum_{n=0}^{\infty}\frac{a_n}{n+1}(n+1)(x-x_o)^n = \sum_{n=0}^{\infty} a_n(x-x_o)^n.$$

[2] see this article I wrote with Daniel Freese, then a student at LU where this result appears as Corollary 5.20 for a power series over an algebra. A less complicated proof is possible here since we merely face real power series.

Suppose $a_n \neq 0$ for all $n \geq 0$ for simplicity of argument and suppose $\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = R$ by Corollary 1.9. Likewise,

$$\lim_{n\to\infty}\left|\frac{na_n}{(n+1)a_{n+1}}\right| = \lim_{n\to\infty}\left|\frac{a_n/(n+1)}{a_{n+1}/(n+2)}\right| = R$$

thus the radius of convergence is shared by all three series. The calculation above holds for the case R is finite or infinite. □

Often it is useful to understand the theorem above with $\sum$-notation and with $x_o = 0$ for brevity:

$$\frac{d}{dx}\left[a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots\right] = a_1 + 2a_2x + 3a_3x^2 + \cdots,$$

and

$$\int\left[a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots\right]\,dx = C + a_0x + \frac{1}{2}a_1x^2 + \frac{1}{3}a_2x^3 + \frac{1}{4}a_3x^4 + \cdots.$$

Example 2.2. *Observe* $\frac{x^2}{1-x^8} = \sum_{n=0}^{\infty} x^2(x^8)^n = \sum_{n=0} x^{8n+2}$ *by the geometric series with* $c = x^2$ *and* $r = x^8$. *Consequently,*

$$\int \frac{x^2\,dx}{1-x^8} = \int \sum_{n=0} x^{8n+2}\,dx = C + \sum_{n=0} \frac{1}{8n+3}x^{8n+3}.$$

Or, if you prefer, without the $\sum$*-notation,*

$$\int \frac{x^2\,dx}{1-x^8} = \int (x^2 + x^{10} + x^{18} + \cdots)\,dx = C + \frac{1}{3}x^3 + \frac{1}{11}x^{11} + \frac{1}{19}x^{19} + \cdots.$$

The example above illustrates the simplicity of integrating a function which is represented as a power series. We will find tools to represent most functions as power series so the example above is not isolated to merely those functions which are so fortunate as to align with the sum of the geometric series. That said, we can do more with geometric series, much more. Let me illustrate:

Example 2.3. *Let* $f(x) = \ln(1+x)$ *then* $\frac{df}{dx} = \frac{1}{1+x} = \sum_{n=0}^{\infty}(-x)^n = \sum_{n=0}^{\infty}(-1)^n x^n$ *where we have used the geometric series with* $r = -x$ *and* $c = 1$ *hence our calculation holds for* $|r| = |x| < 1$. *Integration of the series term-by-term yields*

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

However, we know $f(0) = \ln(1+0) = 0$ *and this implies* $C = 0$. *We find,*

$$\ln(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}x^{n+1} = 1 - \frac{1}{2}x + \frac{1}{3}x^2 + \cdots.$$

Notice $x = 1$ *gives* $\ln(2) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$. *On the other hand, the series for* $\ln(1+x)$ *diverges at* $x = -1$ *since the series for* $\ln(1+x)$ *is the harmonic series. In this case,* $IOC_f = (-1, 1]$ *and we note that integration has added an endpoint to the* $IOC_{\frac{df}{dx}} = (-1, 1)$.

Example 2.4. *Let* $f(x) = \tan^{-1}(x)$ *then* $\frac{df}{dx} = \frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots$. *Then,*

$$\tan^{-1}(x) = C + x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \cdots$$

Again, we find $C = 0$ *since* $\tan^{-1}(0) = 0$. *Thus,*

$$\tan^{-1}(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{2n+1} x^{2n+1}.$$

Example 2.5. *Let* $f(x) = \frac{1}{(1-x)^2}$ *then* $\displaystyle\int \frac{dx}{(1-x)^2} = C + \frac{1}{1-x} = C + 1 + x + x^2 + x^3 + \cdots$. *Then,*

$$\begin{aligned}
\frac{1}{(1-x)^2} &= \frac{d}{dx}\int \frac{dx}{(1-x)^2} \\
&= \frac{d}{dx}\left[C + 1 + x + x^2 + x^3 + \cdots\right] \\
&= 1 + 2x + 3x^2 + \cdots . \\
&= \sum_{n=0}^{\infty} (n+1)x^n.
\end{aligned}$$

In the past two examples I have opted to integrate and differentiate in the $+\cdots$ notation only to discern a pattern in the final step of the calculation. We could also use $\sum$-notation to derive the formulas directly. The next example illustrates the importance of algebra in this game we're playing.

Example 2.6. *Find a power series representation of* $f(x) = \dfrac{1}{x^2 - 6x + 13}$.

$$\begin{aligned}
\frac{1}{x^2 - 6x + 13} &= \frac{1}{(x-3)^2 + 4} \\
&= \frac{1}{4\left(1 + \frac{(x-3)^2}{4}\right)} \\
&= \sum_{n=0}^{\infty} \frac{1}{4}\left(\frac{-(x-3)^2}{4}\right)^n \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{4^{n+1}}(x-3)^{2n}.
\end{aligned}$$

Certainly more examples are possible here, but I think it best for us to continue on to learn about Taylor's Theorem so we can add power series for sine, cosine, hyperbolic sine, hyperbolic cosine and the exponential to our lexicon. Furthermore, we can study the product and quotient of power series to produce even more examples. Geometric series are important, but it is just one of several techniques we need to understand.

3 Taylor's Theorem about polynomial approximation

The idea of a Taylor polynomial is that if we are given a set of initial data $f(a), f'(a), f''(a), \dots, f^{(n)}(a)$ for some function $f(x)$ then we can approximate the function with an n^{th}-order polynomial which fits all the given data. Let's see how it works order by order starting with the most silly case.

3.1 constant functions

Suppose we are given $f(a) = y_o$ then $T_o(x) = y_o$ is the zeroth Taylor polynomial for f centered at $x = a$. Usually you have to be very close to the center of the approximation for this to match the function.

3.2 linearizations again

Suppose we are given values for $f(a)$, $f'(a)$ we seek to find $T_1(x) = c_o + c_1(x-a)$ which fits the given data. Note that

$$\begin{aligned} T_1(a) &= c_o + c_1(a-a) = f(a) & c_o &= f(a). \\ T_1'(a) &= c_1 = f'(a) & c_1 &= f'(a). \end{aligned}$$

Which gives us the first Taylor polynomial for f centered at a: $T_1(x) = f(a) + f'(a)(x-a)$. This function, I hope, is familiar from our earlier study of linearizations. The linearization at a is the best linear approximation to f near a.

3.3 quadratic approximation of function

Suppose we are given values for $f(a)$, $f'(a)$ and $f''(a)$ we seek to find $T_2(x) = c_o + c_1(x-a) + c_2(x-a)^2$ which fits the given data. Note that

$$\begin{aligned} T_2(a) &= c_o + c_1(a-a) + c_2(a-a)^2 = f(a) & c_o &= f(a). \\ T_2'(a) &= c_1 + 2c_2(a-a) = f'(a) & c_1 &= f'(a). \\ T_2''(a) &= 2c_2 = f''(a) & c_2 &= \frac{1}{2}f''(a). \end{aligned}$$

Which gives us the first Taylor polynomial for f centered at a: $T_1(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2$. We would hope this is the best quadratic approximation for f near the point $(a, f(a))$.

3.4 cubic approximation of function

Suppose we are given values for $f(a)$, $f'(a)$, $f''(a)$ and $f'''(a)$ we seek to find $T_2(x) = c_o + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3$ which fits the given data. Note that

$$\begin{aligned} T_3(a) &= c_o + c_1(a-a) + c_2(a-a)^2 + c_3(a-a)^3 = f(a) & c_o &= f(a). \\ T_3'(a) &= c_1 + 2c_2(a-a) + 3c_3(a-a)^2 = f'(a) & c_1 &= f'(a). \\ T_3''(a) &= 2c_2 + 3 \cdot 2c_3(a-a) = f''(a) & c_2 &= \frac{1}{2}f''(a). \\ T_3'''(a) &= 3 \cdot 2c_3) = f'''(a) & c_3 &= \frac{1}{3 \cdot 2}f'''(a). \end{aligned}$$

Which gives us the first Taylor polynomial for f centered at a: $T_1(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{6}f'''(a)(x-a)^3$. We would hope this is the best cubic approximation for f near the point $(a, f(a))$.

3.5 general case

Hopefully by now a pattern is starting to emerge. We see that $T_k(x) = T_{k-1}(x) + \frac{1}{k!}f^{(k)}(a)(x-a)^k$ where $k! = k(k-1)(k-2)\cdots 3\cdot 2\cdot 1$.

Definition 3.1. *Taylor polynomials.*

Suppose f is a function which has k-derivatives defined at a then the k-th Taylor polynomial for f is defined to be $T_k(x)$ where

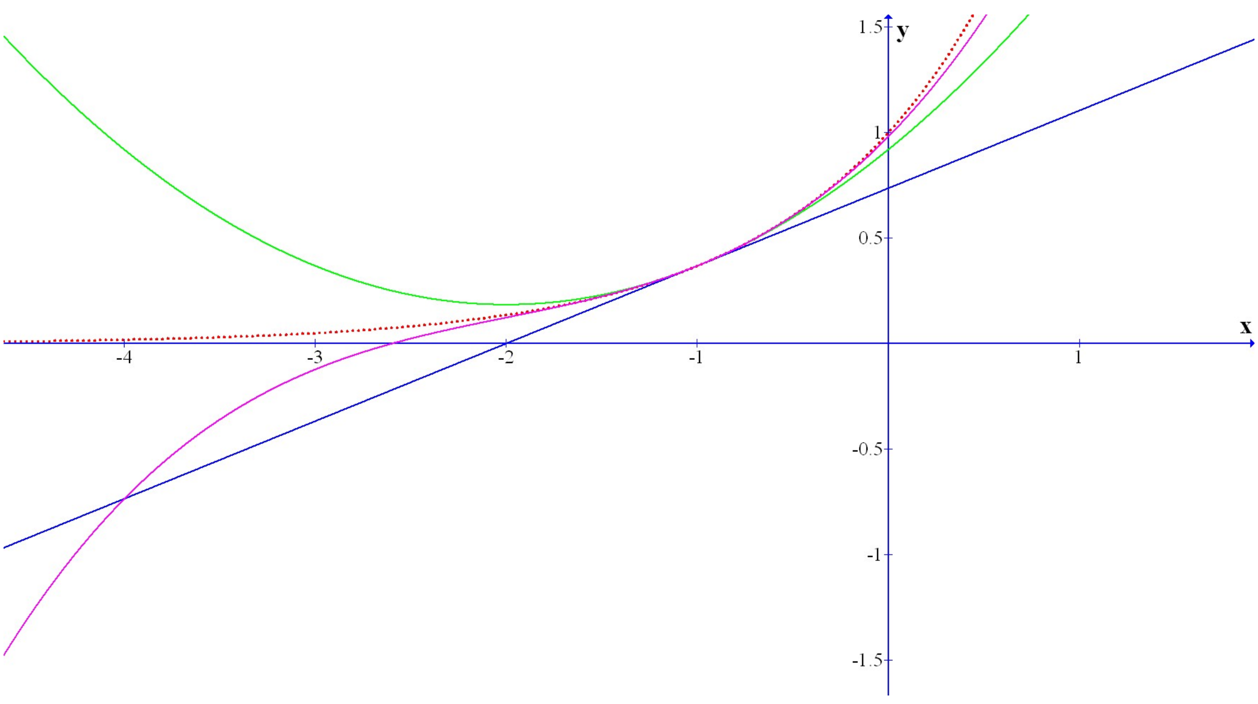
$$T_k(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{k!}f^{(k)}(a)(x-a)^k$$

Let's examine a few examples before continuing with the theory.

Example 3.2. *Suppose $f(x) = e^x$. Calculate the first few Taylor polynomials centered at $a = -1$. Derivatives of the exponential are easy enough to calculate; $f'(x) = f''(x) = f'''(x) = e^x$ therefore we find*

$$\begin{aligned}
T_o(x) &= \frac{1}{e} \\
T_1(x) &= \frac{1}{e} + \frac{1}{e}(x+1) \\
T_2(x) &= \frac{1}{e} + \frac{1}{e}(x+1) + \frac{1}{2e}(x+1)^2 \\
T_3(x) &= \frac{1}{e} + \frac{1}{e}(x+1) + \frac{1}{2e}(x+1)^2 + \frac{1}{6e}(x+1)^3.
\end{aligned}$$

The graph below shows $y = e^x$ as the dotted red graph, $y = T_1(x)$ is the blue line, $y = T_2(x)$ is the green quadratic and $y = T_3(x)$ is the purple graph of a cubic. You can see that the cubic is the best approximation.



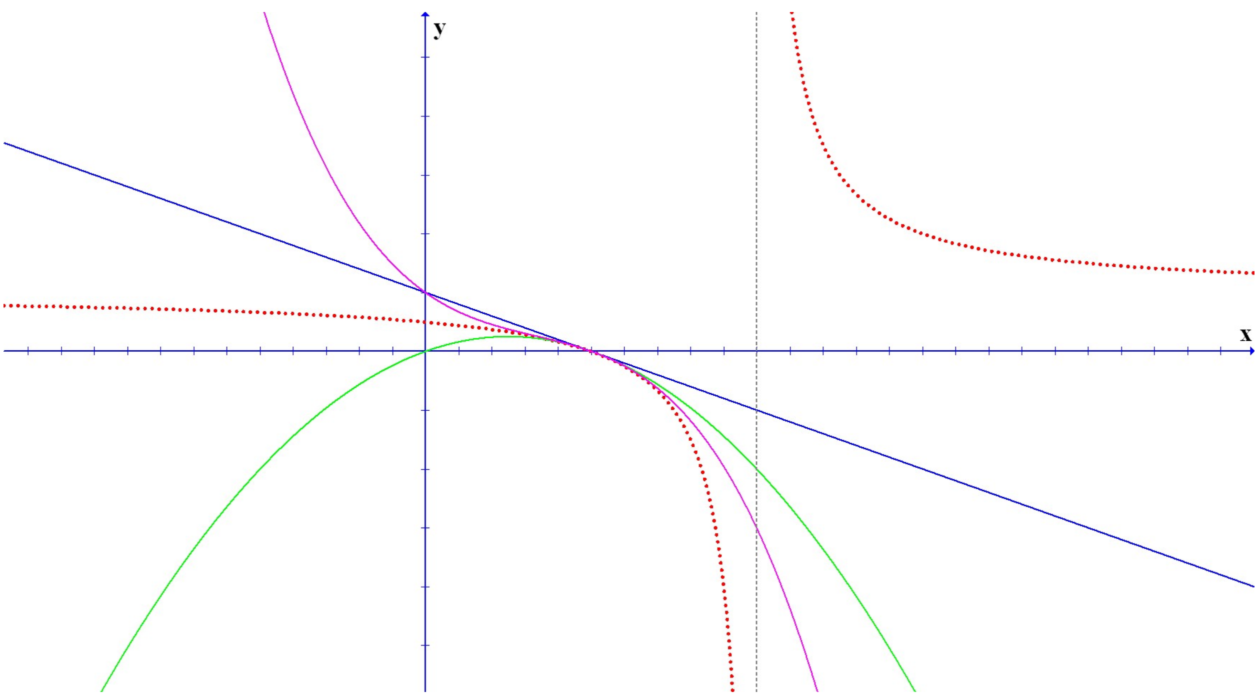
Example 3.3. *Suppose $f(x) = \frac{1}{x-2} + 1$. Calculate the first few Taylor polynomials centered at $a = 1$. Observe*

$$f(x) = \frac{1}{x-2} + 1, \quad f'(x) = \frac{-1}{(x-2)^2}, \quad f''(x) = \frac{2}{(x-2)^3}, \quad f'''(x) = \frac{-6}{(x-2)^4}$$

thus $f(1) = 0, f'(1) = -1, f''(1) = -2$ and $f'''(1) = -6$. Hence,

$$T_1(x) = -(x-1)$$
$$T_2(x) = -(x-1) + (x-1)^2$$
$$T_3(x) = -(x-1) + (x-1)^2 - (x-1)^3$$

The graph below shows $y = \frac{1}{x-2} + 1$ as the dotted red graph, $y = T_1(x)$ is the blue line, $y = T_2(x)$ is the green quadratic and $y = T_3(x)$ is the purple graph of a cubic. You can see that the cubic is the best approximation. Also, you can see that the Taylor polynomials will not give a good approximation to $f(x)$ to the right of the VA at $x = 2$.

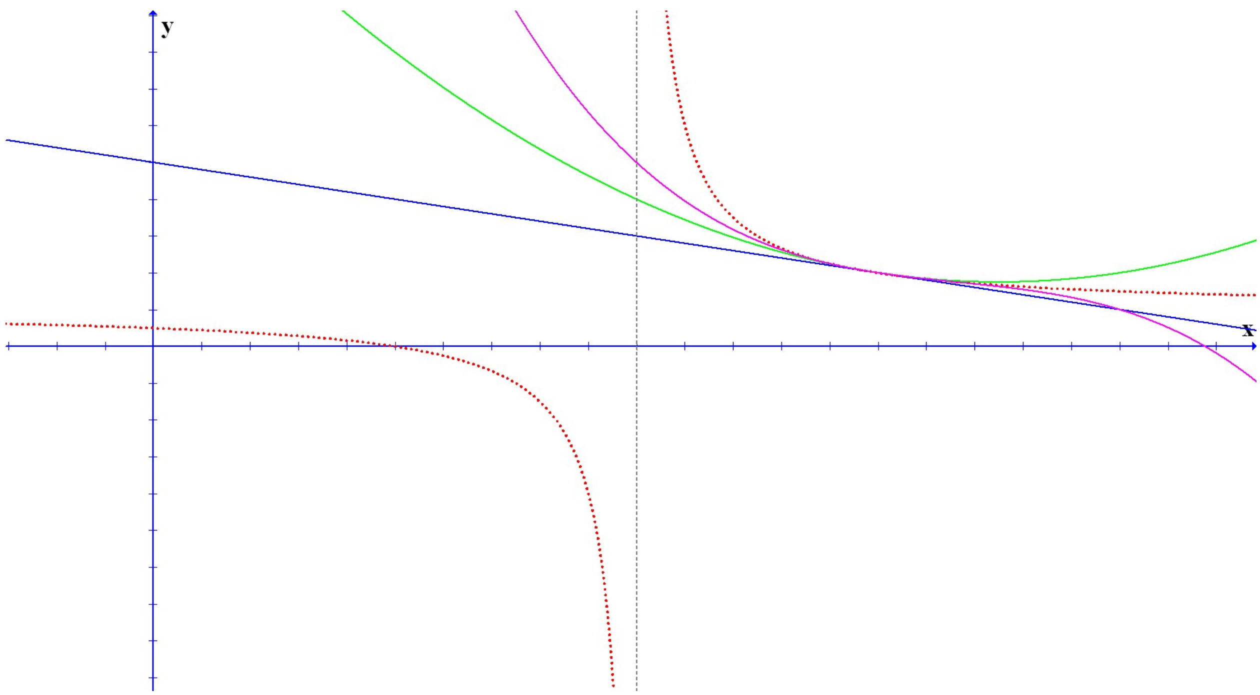


On the next page we examine the same function approximated at a different center point. In other words, for a given $f(x)$ we can consider its Taylor polynomial approximation at different points. Generally, these approximations differ (can you think of a case where they would be the same ?)

Let us continue our study of Taylor polynomials which approximate $f(x) = \frac{1}{x-2} + 1$. *We expand about the center* $a = 3$ *to find*

$$\begin{aligned} T_1(x) &= 2 + (3-x) \\ T_2(x) &= 2 + (3-x) + (3-x)^2 \\ T_3(x) &= 2 + (3-x) + (3-x)^2 + (3-x)^3. \end{aligned}$$

The graph below uses the same color scheme. Notice this time the Taylor polynomials only work well to the right of the vertical asymptote.



Example 3.4. *Let $f(x) = \sin(x)$. It follows that*

$$f'(x) = \cos(x), f''(x) = -\sin(x), f'''(x) = -\cos(x), f^{(4)}(x) = \sin(x), f^{(5)}(x) = \cos(x)$$

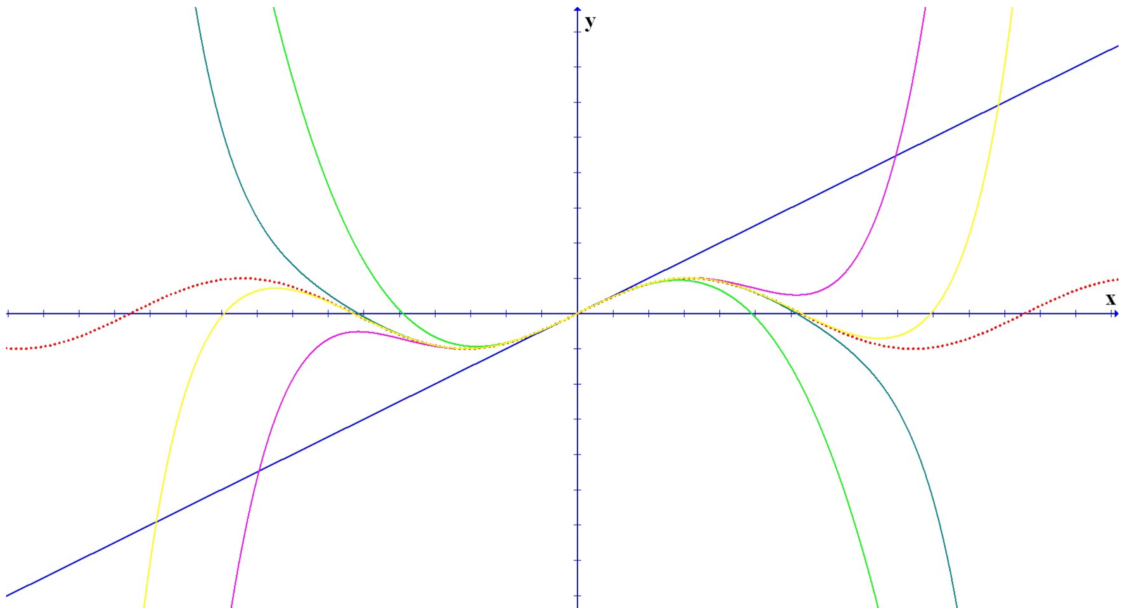
Hence, $f(0) = 0, f'(0) = 1, f''(0) = 0, f'''(0) = -1, f^{(4)}(0) = 0, f^{(5)}(0) = 1$*. Therefore the Taylor polynomials of orders* $1, 3, 5$ *are*

$$T_1(x) = x \qquad \textit{blue graph}$$

$$T_3(x) = x - \frac{1}{6}x^3 \qquad \textit{green graph}$$

$$T_5(x) = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 \qquad \textit{purple graph}$$

The graph below shows the Taylor polynomials calculated above and the next couple orders above. You can see how each higher order covers more and more of the graph of the sine function.



Taylor polynomials can be generated for a given *smooth*[3] function through a certain linear combination of its derivatives. The idea is that we can approximate a function by a polynomial[4], at least locally. We discussed the tangent line approximation to a function. We found that the linearization of a function gives a good approximation for points close to the point of tangency. If we calculate second derivatives we can similarly find a quadratic approximation for the function. Third derivatives go to finding a cubic approximation about some point. I should emphasize from the outset that a Taylor polynomial is just a polynomial, it will not be able to exactly represent a function which is not a polynomial. In order to exactly represent an analytic function we'll need to take infinitely many terms, we'll need a power series. Finally, let me show you an example of how Taylor polynomials can be of fundamental importance in physics.

[3]for $p \in \mathbb{R}$ the notation $f \in C^\infty(p)$ means there exists a nbhd. of $p \in \mathbb{R}$ on which f has infinitely many continuous derivatives.

[4]there do exist pathological examples for which all Taylor polynomials at a point vanish even though the function is nonzero near the point; $f(x) = exp(-1/x^2)$ for $x \neq 0$ and $f(0) = 0$

Example 3.5. *The relativistic energy* E *of a free particle of rest mass* m_o *is a function of its velocity* v*:*

$$E(v) = \frac{m_o c^2}{\sqrt{1 - v^2/c^2}}$$

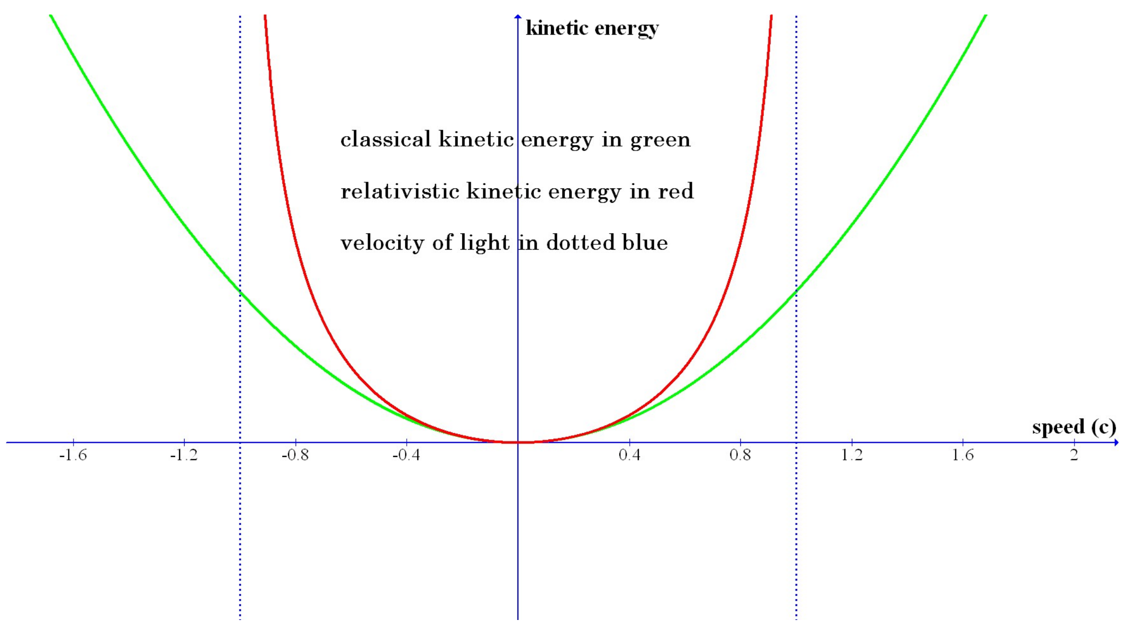
for $-c < v < c$ *where* c *is the speed of light in the space. We calculate,*

$$\frac{dE}{dv} = \frac{m_o v}{(1 - v^2/c^2)^{\frac{3}{2}}}$$

thus $v = 0$ *is a critical number of the energy. Moreover, after a little calculation you can show the* 4*-th order Taylor polynomial in velocity* v *for energy* E *is*

$$E(v) \approx m_o c^2 + \frac{1}{2} m_o v^2 + \frac{3m_o}{8c^2} v^4$$

The constant term is the source of the famous equation $E = m_o c^2$ *and the quadratic term is precisely the classical kinetic energy. The last term is very small if* $v \approx 0$*. As* $|v| \to c$ *the values of the last term become more significant and they signal a departure from classical physics. I have graphed the relativistic kinetic energy* $K = E - m_o c^2$ *(red) as well as the classical kinetic energy* $K_{Newtonian} = \frac{m_o}{2} v^2$ *(green) on a common axis below:*



The blue-dotted lines represent $v = \pm c$ *and if* $|v| > c$ *the relativistic kinetic energy is not even defined. However, for* $v \approx 0$ *you can see they are in very good agreement. We have to get past* 10% *of light speed to even begin to see a difference. In every day physics most speeds are so small that we cannot see that Newtonian physics fails to correctly model dynamics. I may have assigned a homework based on the error analysis of the next section which puts a quantitative edge on the last couple sentences.*

One of the great mysteries of modern science is this fascinating feature of *decoupling*. How is it that we are so fortunate that the part of physics which touches one aspect of our existence is so successfully described. Why isn't it the case that we need to understand relativity before we can pose solutions to the problems presented to Newtonian mechanics? Why is physics so nicely segmented that we can understand just one piece at a time? This is part of the curiosity which leads physicists to state that the existence of physical law itself is bizarre. If the universe is randomly generated as is life then how is it that we humble accidents can so aptly describe what surrounds us. What right have we to understand what we do of nature? Recently some materialists have turned to something called the *anthropomorphic principle* as a tool to describe how this fortunate accident occurred. To the hardcore materialist the allowance of supernatural intervention is abhorrent. They prefer a universe without purpose. Personally, I prefer purpose. Moreover, it is my understanding of my place in this universe and our purpose to glorify God that make me expect to find laws of physics. Laws, or more correctly, approximations of physics reveal the glory of a God we cannot fully comprehend. I guess I digress... back to the math.

3.6 error in Taylor approximations

We've seen a few examples of how Taylor's polynomials will locally mimic a function. Now we turn to the question of extrema. Think about this, if a function is locally modeled by a Taylor polynomial centered at a critical point then what does that say about the nature of a critical point? To be precise we need to know some measure of how far off a given Taylor polynomial is from the function. This is what Taylor's theorem tells us. There are many different formulations of Taylor's theorem[5], the one below is partially due to Lagrange.

Theorem 3.6. *Taylor's theorem with Lagrange's form of the remainder.*

If f has k derivatives on a closed interval I with $\partial I = \{a, b\}$ then

$$f(b) = T_k(b) + R_k(b) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(b-a)^j + R_k(b)$$

where $R_k(b) = f(b) - T_k(b)$ is the k-th remainder. Moreover, there exists $c \in int(I)$ such that

$$R_k(b) = \frac{f^{(k+1)}(c)}{(k+1)!}(b-a)^{k+1}.$$

Proof: We have essentially proved the first portion of this theorem. It's straightforward calculation to show that $T_k(x)$ has the same value, slope, concavity etc... as the function at the point $x = a$. What is deep about this theorem is the existence of c. This is a generalization of the mean value theorem. Suppose that $a < b$, if we apply the theorem to

$$f(x) = T_o(x) + R_1(x)$$

we find Taylor's theorem proclaims there exists $c \in (a, b)$ such that $R_1(b) = f'(c)(b-a)$ and since $T_o(x) = f(a)$ we have $f(b) - f(a) = f'(c)(b-a)$ which is the conclusion of the MVT applied to $[a, b]$.

Proof of Taylor's Theorem: the proof I give here I found in *Real Variables with Basic Metric Space Topology* by Robert B. Ash. Proofs found in other texts are similar but I thought his was particularly lucid.

[5]Chapter 7 of Apostol or Chapter II.6 of Edwards would be good additional readings if you wish to understand this material in added depth.

Since the k-th derivative is given to exist on I it follows that $f^{(j)}$ is continuous for each $j = 1, 2, \ldots, k-1$ (we are not garaunteed the continuity of the k-th derivative, however it is not needed in what follows anyway). Assume $a < b$ and define M implicitly by the equation below:

$$f(b) = f(a) + f'(a)(b-a) + \cdots + \frac{f^{(k-1)}(a)}{(k-1)!}(b-a)^{(k-1)} + \frac{M(b-a)^k}{k!}.$$

Our goal is to produce $c \in (a, b)$ such that $f^{(k)}(c) = M$. Ash suggests replacing a with a variable t in the equation that defined M. Define g by

$$g(t) = -f(b) + f(t) + f'(t)(b-t) + \cdots + \frac{f^{(k-1)}(t)}{(k-1)!}(b-t)^{(k-1)} + \frac{M(b-t)^k}{k!}$$

for $t \in [a, b]$. Note that g is differentiable on (a, b) and continuous on $[a, b]$ since it is the sum and difference of likewise differentiable and continuous functions. Moreover, observe

$$g(b) = -f(b) + f(b) + f'(b)(b-b) + \cdots + \frac{f^{(k-1)}(t)}{(k-1)!}(b-b)^{(k-1)} + \frac{M(b-b)^k}{k!} = 0.$$

On the other hand, the definition of M implies $g(a) = 0$. Therefore, Rolle's theorem applies to g, this means there exists $c \in (a, b)$ such that $g'(c) = 0$. Calculate the derivative of g, the minus signs stem from the chain rule applied to the $b - t$ terms,

$$\begin{aligned} g'(t) = & \frac{d}{dt}\big[-f(b) + f(t)\big] \; + \; \frac{d}{dt}\big[f'(t)(b-t)\big] \; + \cdots + \\ & + \frac{d}{dt}\Big[\frac{f^{(k-1)}(t)}{(k-1)!}(b-t)^{(k-1)}\Big] \; + \; \frac{d}{dt}\Big[\frac{M(b-t)^k}{k!}\Big] \\ & = f'(t) - f'(t) + f''(t)(b-t) - \frac{1}{2}f''(t)2(b-t) \; + \cdots + \\ & + \frac{f^{(k)}(t)}{(k-1)!}(b-t)^{(k-1)} - \frac{f^{(k-1)}(t)}{(k-1)!}k(b-t)^{(k-2)} - \frac{Mk(b-t)^{k-1}}{k!} \\ & = \frac{f^{(k)}(t)}{(k-1)!}(b-t)^{(k-1)} - \frac{Mk(b-t)^{k-1}}{k!} \\ & = \frac{(b-t)^{(k-1)}}{(k-1)!}\Big[f^{(k)}(t) - M\Big] \end{aligned}$$

where we used that $\frac{k}{k!} = \frac{k}{k(k-1)!} = \frac{1}{(k-1)!}$ in the last step. Note that $c \in (a, b)$ therefore $c \neq b$ hence $(b - t) \neq 0$ hence $(b-t)^{(k-1)} \neq 0$ hence $\frac{(b-t)^{(k-1)}}{(k-1)!} \neq 0$. It follows that $g'(c) = 0$ implies $f^{(k)}(c) - M = 0$ which shows $M = f^{(k)}(c)$ for some $c \in (a, b)$. The proof for the case $b > a$ is similar. $\square$

In total, we see that Taylor's theorem is more or less a simple consequence of Rolle's theorem. In fact, the proof above is not much different than the proof we gave previously for the MVT.

Corollary 3.7. *error bound for $T_k(x)$.*

If a function f has $(k+1)$-continuous derivatives on a closed interval $[p,q]$ with length $l = q-p$ and $|f^{(k+1)}(x)| \leq M$ for all $x \in (p,q)$ then for each $a \in (p,q)$

$$|R_k^a(x)| \leq \frac{Ml^{k+1}}{(k+1)!}$$

where $f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j + R_k^a(x)$.

Proof: At each point a we can either look at $[a,x]$ or $[x,a]$ and apply Taylor's theorem to obtain $c_a \in \mathbb{R}$ such that $f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j + R_k^a(x)$ where $R_k^a(x) = \frac{f^{(k+1)}(c_a)}{(k+1)!}(x-a)^{k+1}$. Then we note $|f^{(k+1)}(c_a)| \leq M$ and the corollary follows. □

Consider the criteria for the Second Derivative test. We required that $f'(c) = 0$ and $f''(c) \neq 0$ for a definite conclusion. If f'' is continuous at c with $f''(c) \neq 0$ then it is nonzero on some closed interval $I = [c-\delta, c+\delta]$ where $\delta > 0$. If we also are given that f''' is continuous on I then it follows there exists $M > 0$ such that $|f'''(x)| \leq M$ for all $x \in I$. Observe that

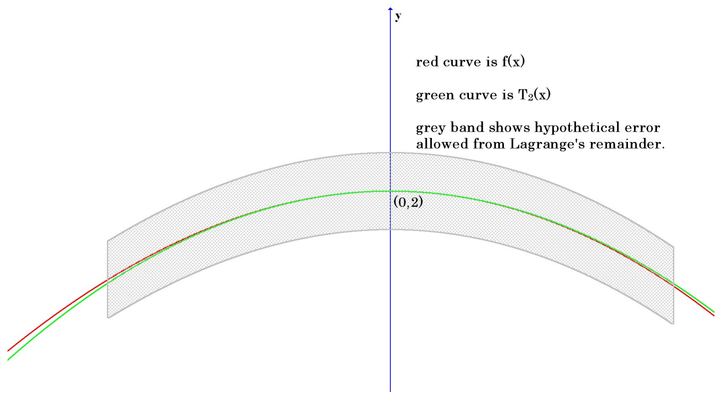
$$|f(x) - f(c) - \frac{1}{2}f''(c)(x-c)^2| = |\frac{1}{6}f'''(\zeta_x)(x-c)^3| \leq \frac{4M\delta^3}{3}$$

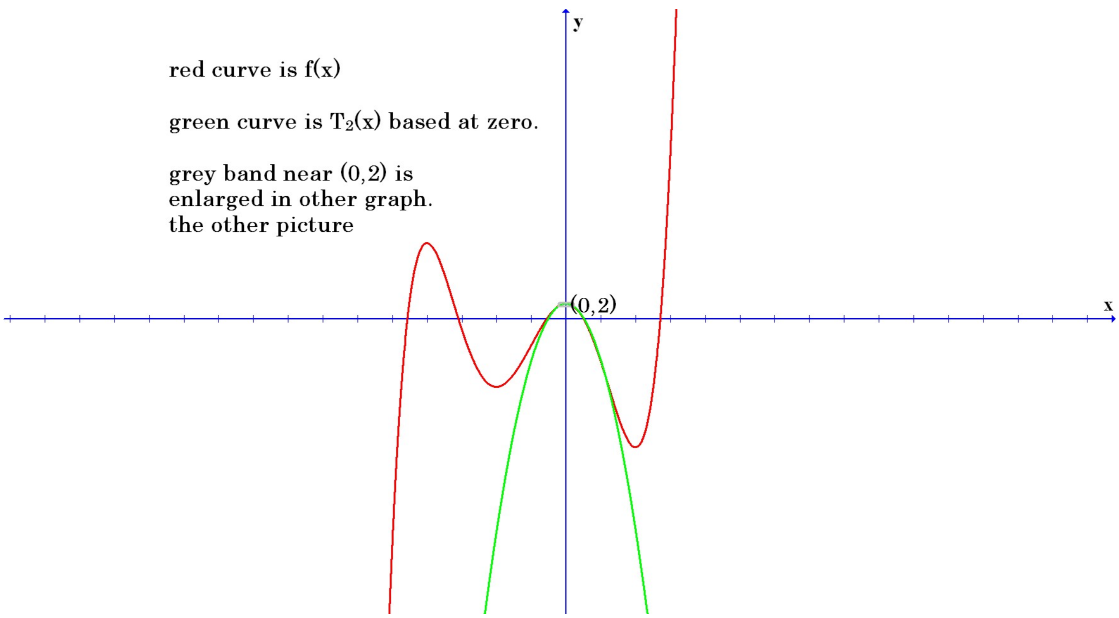
for all $x \in [c-\delta, c+\delta]$. This inequality reveals that we have $f(x) \approx f(c) + \frac{1}{2}f''(c)(x-c)^2$ as $\delta \to 0$. Therefore, locally the graph of the function resembles a parabola which either opens up or down at the critical point. If it opens up ($f''(c) > 0$) then $f(c)$ is the local minimum value of f. If it opens down ($f''(c) < 0$) then $f(c)$ is the local maximum value of f. Of course this is no surprise. However, notice that we may now quantify the error $E_2(x) = |f(x) - T_2(x)| \leq \frac{8M\delta^3}{3}$. If we can choose a bound for $f'''(x)$ independent of x then the error is simply bounded just in terms of the distance from the critical point which we can choose $\delta = |x-c|$ and the resulting error is just $\frac{4M\delta^3}{3}$. Usually, M will depend on the distance from c so the choice of δ to limit error is a bit more subtle. Let me illustrate how this analysis works in an example.

Example 3.8. *Suppose* $f(x) = 6x^5 + 15x^4 - 10x^3 - 30x^2 + 2$. *We can calculate that* $f'(x) = 30x^4 + 60x^3 - 30x^2 - 60x$ *therefore clearly* $(0,2)$ *is a critical point of* f. *Moreover,* $f''(x) = 120x^3 + 180x^2 - 60x - 60$ *shows* $f''(0) = -60$. *I aim to show how the quadratic Taylor polynomial* $T_2(x) = f(2) + f'(2)x + \frac{1}{2}f''(2)x^2 = 2 - 30x^2$ *gives a good approximation for* $f(x)$ *in the sense that the maximum error is essentially bounded by the size of Lagrange's term. Note that*

$$f'''(x) = 360x^2 + 360x - 60 \qquad \textit{and} \qquad f^{(4)}(x) = 720x + 360$$

Suppose we seek to approximate on $-0.1 < x < 0.1$ *then for such* x *we may verify that* $f^{(4)}(x) > 0$ *which means* f''' *is increasing on* $[-0.1, 0.1]$ *thus* $f'''(-0.1) < f'''(x) < f'''(0.1)$ *which gives* $3.6 - 36 - 60 < f'''(x) < 3.6 + 36 - 60$ *thus* $-92.4 < f'''(x) < -20.4$. *Therefore, if* $|x| < 0.1$ *then* $|f'''(x)| < 92.4$. *Using* $\delta = 0.1$ *we should expect a bound on the error of* $\frac{4M\delta^3}{3} = 4(92.4)/3000 = 0.123$. *I have illustrated the global and local qualities of the Taylor Polynomial centered at zero. Notice that the error bound was quite generous in this example.*





Example 3.9. *Here we examine Taylor polynomials for $f(x) = \sin(x)$ on the interval $(-1, 1)$ and second on $(-2, 2)$. In each case we use sufficiently many terms to guarantee an error of less than $\epsilon = 0.1$. Notice that $f^{(2k-1)}(x) = \pm\sin(x)$ whereas $f^{(2k-2)}(x) = \pm\cos(x)$ for all $k \in \mathbb{N}$ therefore $|f^{(n)}(x)| \leq 1$ for all $x \in \mathbb{R}$.*

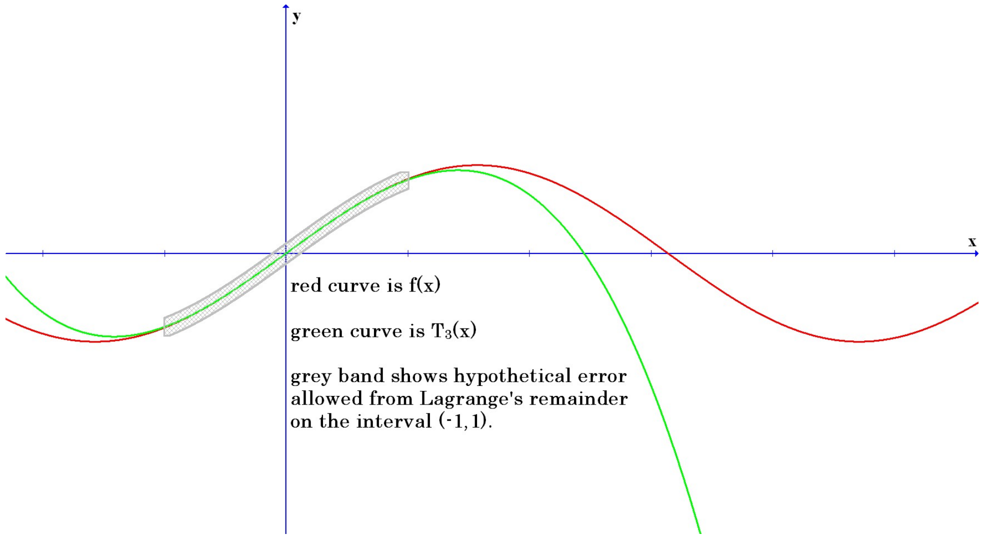
If we wish to bound the error to 0.1 on $-1 < x < 1$ then we to bound the remainder term as follows: (note $-1 < x < 1$ implies $l = 2$ and we just argued $M = 1$ is a good bound for any k)

$$|f(x) - T_k(x)| \leq \frac{Ml^{k+1}}{(k+1)!} = \frac{2^{k+1}}{(k+1)!} = E_k \leq 0.1$$

At this point I just start plugging various values of k until I find a value smaller than the desired bound. For this case,

$$E_1 = \frac{2^2}{2!} = 2, \ E_2 = \frac{2^3}{3!} = \frac{4}{3}, \ E_3 = \frac{2^4}{4!} = \frac{2}{3}, \ E_4 = \frac{2^5}{5!} = \frac{32}{120} \approx 0.25, \ E_5 = \frac{2^6}{6!} = \frac{64}{720} \approx 0.1$$

This shows that $T_4(x)$ will provide the desired accuracy. But, it just so happens that $T_3 = T_4$ in this case so we find $T_3(x) = x - \frac{1}{6}x^3$ will suffice. In fact, it fits the ± 0.1 tolerance band quite nicely:



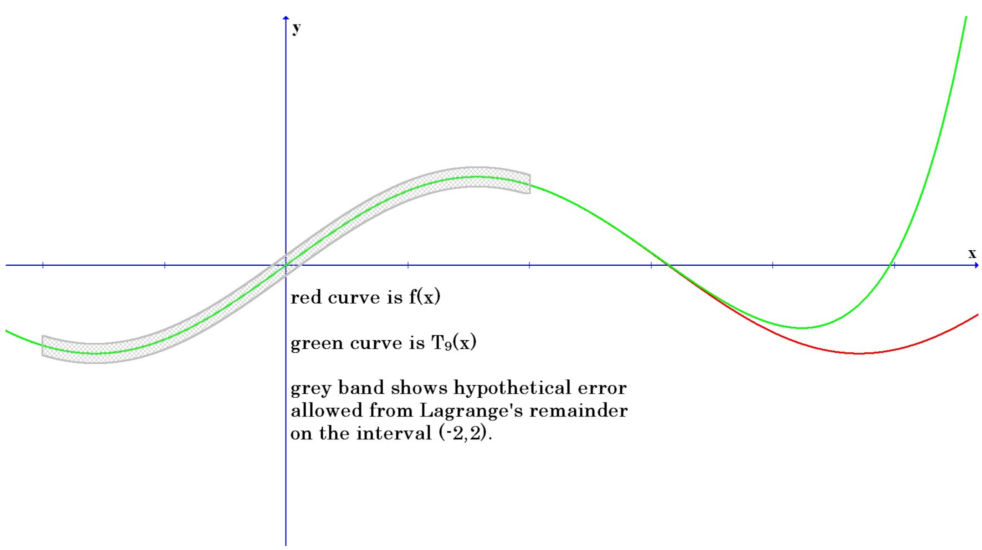
On the next page I determine what is needed to mimic the sine function on the larger interval $-2 < x < 2$*,*
If we wish to bound the error to 0.1 *on* $-2 < x < 2$ *then we to bound the remainder term as follows: (note* $-2 < x < 2$ *implies* $l = 4$ *)*

$$|f(x) - T_k(x)| \leq \frac{Ml^{k+1}}{(k+1)!} = \frac{4^{k+1}}{(k+1)!} = E_k \leq 0.1$$

At this point I just start plugging various values of k *until I find a value smaller than the desired bound. For this case,*

$$E_7 = \frac{4^8}{8!} \approx 1.6, \;\; E_9 = \frac{4^{10}}{10!} \approx 0.3, \;\; E_{11} = \frac{2^{12}}{12!} \approx 0.035$$

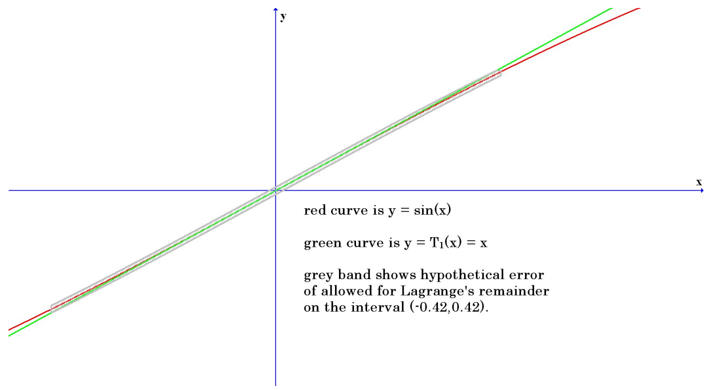
This shows that $T_10(x)$ *will provide the desired accuracy. But, it just so happens that* $T_9 = T_{10}$ *in this case so we find* $T_9(x) = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \frac{1}{362880}x^9$ *will suffice. In fact, as you can see below it fits the* ± 0.1 *tolerance band quite nicely well beyond the target interval of* $-2 < x < 2$*:*



Example 3.10. *Let's think about $f(x) = \sin(x)$ again. This time, answer the following question: for what domain $-\delta < x < \delta$ will $f(x) \approx x$ to within ± 0.01 ? We can use $M = 1$ and $l = 2\delta$. Furthermore, $T_1(x) = T_2(x) = x$ therefore we want*

$$|f(x) - x| \leq \frac{(2\delta)^3}{(3!} = \frac{4\delta^3}{3} \leq 0.1$$

to hold true for our choice of δ. Hence $\delta^3 \leq 0.075$ which suggests $\delta \leq 0.42$. Taylor's theorem thus shows $\sin(x) \approx x$ to within ± 0.01 provided $-0.42 < x < 0.42$. (0.42 radians translates into about 24 degrees). Here's a picture of $f(x) = \sin(x)$ (in red) and $T_1(x) = x$ (in green) as well as the tolerance band (in grey). You should recognize $y = T_1(x)$ as the tangent line.



Example 3.11. *Suppose we are faced with the task of calculating $\sqrt{4.03}$ to an accuracy of 5-decimals. For the purposes of this example assume all calculators are evil. It's after the robot holocaust so they can't be trusted. What to do? We use the Taylor polynomial up to quadratic order: we have $f(x) = \sqrt{x}$ and $f'(x) = \frac{1}{2\sqrt{x}}$ and $f''(x) = \frac{-1}{4(\sqrt{x})^3}$. Apply Taylor's theorem,*

$$\begin{aligned}
\sqrt{4.03} &= f(4) + f'(4)(4.03-4) + \frac{1}{2}f''(4)(4.03-4)^2 + R \\
&= 2 + \frac{1}{4}\frac{3}{100} - \frac{1}{64}\frac{9}{10000} + R \\
&= 2 + 0.0075 - 0.000014062 + R \\
&= 2.\underline{007485}938 + R
\end{aligned}$$

If we bound $f'''(x) = \frac{3}{8(\sqrt{x})^5}$ by M on $[4, 4.03]$ then $|R| \leq \frac{M(0.03)^3}{6}$. Clearly $f''''(x) = \frac{-15}{16(\sqrt{x})^7} < 0$ for $x \in [4, 4.03]$ therefore, f''' is decreasing on $[4, 4.03]$. It follows $f'''(4) \geq f'''(x) \geq f'''(4.03)$. Choose $M = f'''(4) = \frac{3}{8(32)} = \frac{3}{256}$ thus

$$|R| \leq \frac{(0.03)^3}{6}\frac{3}{256} = \frac{27}{256}\frac{1}{10000} \approx \frac{1}{100000} = 0.000001.$$

Therefore, $\boxed{\sqrt{4.03} = 2.007486 \pm 0.000001}$*. As far as I know my TI-89 is still benevolent so we can check our answer; the calculator says $\sqrt{4.03} = 2.\underline{00748598}999$.*

In the last example, we again find that we actually are a whole digit closer to the answer than the error bound suggests. This seems to be typical. Notice, sometimes we could use the alternating series estimation theorem to obtain a bound on the error with greater ease.

Example 3.12. *Newton postulated that the gravitational force between masses m and M separated by a distance of r is*

$$\vec{F} = -\frac{GmM}{r^2}\hat{r}$$

where r is the distance from the center of mass of M to the center of mass m and G is a constant which quantifies the strength of gravity. The minus sign means gravity is always attractive in the direction $\hat{r}$ which points along the line from M to m. Consider a particular case, M is the mass of the earth and m is a small mass a distance r from the center of the earth. It is convenient to write $r = R + h$ where R is the radius of the earth and h is the **altitude** *of m. Here we make the simplifying assumptions that m is a point mass and M is a spherical mass with a homogeneous mass distribution. It turns out that means we can idealize M as a point mass at the center of the earth. All of this said, you may recall that $F = mg$ is the force of gravity in highschool physics where the force points down. But, this is very different then the inverse square law? How are these formulas connected? Focus on a particular ray eminating from the center of the earth so the force depends only on the altitude h. In particular:*

$$F(h) = -\frac{GmM}{(R+h)^2}$$

We calculate,

$$F'(h) = \frac{2GmM}{(R+h)^3}$$

Note that clearly $F''(h) < 0$ hence F' is a decreasing function of h therefore if $0 \leq h \leq h_{max}$ then $F'(0) \geq F'(h) \geq F'(h_{max})$ so $F'(0)$ provides a bound on $F'(h)$. Calculate that

$$F(0) = -\frac{GmM}{R^2} \text{ and } F'(0) = \frac{2GmM}{R^3}$$

Taylor's theorem says that $F(h) = F(0) + E$ and $|E| \leq F'(0)h_{max}$ therefore,

$$F(h) \approx -\frac{GmM}{R^2} \pm \frac{2GmM}{R^3}h$$

Note $G = 6.673 \times 10^{-11}\frac{Nm^2}{kg^2}$ and $R = 6.3675 \times 10^6 m$ and $M = 5.972 \times 10^{24} kg$. You can calculate that $\frac{GmM}{R^2} = 9.83 m/s^2$ which is hopefully familar to some who read this. In contrast, the error term

$$|E| = \frac{2GmM}{R^3}h = (3.1 \times 10^{-6})mh$$

If the altitude doesn't exceed $h = 1,000m$ then the formula $F/m = g$ approximates the true inverse square law to within $0.0031 m/s^2$. At $h = 10,000m$ the error is $0.031 m/s^2$. At $h = 100,000m$ the error is around $0.31 m/s^2$. (100,000 meters is about 60 miles, well above most planes flight ceiling). Taylor's theorem gives us the mathematical tools we need to quantify such nebulous phrases as $F = mg$ "near" the surface of the earth. Mathematically, this is probably the most boring Taylor polynomial you'll ever study, it was just the constant term.

4 Taylor Series and the Binomial Series

Let $C^\infty(a)$ denote the set of functions which have arbitrarily many derivatives which exist at $x = a$. In particular, $f \in C^\infty(a)$ means f is **smooth** near $x = a$. We extend Definition 3.1 as follows:

Definition 4.1. *Taylor series for analytic functions*

If $f \in C^\infty(a)$ then the Taylor series generated by f at $x = a$ is

$$T(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots$$

If $T(x) = f(x)$ for all x close to $x = a$ then we say f is **analytic** at $x = a$. The Taylor series generated by f at $x = 0$ is known as the **Maclaurin series** for f.

The usual example of a function which is **not** analytic is:

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \exp(-1/x) & \text{if } x \leq 0 \end{cases}$$

It can be shown that $f^{(n)}(0) = 0$ for $n = 0, 1, 2, \dots$ hence the Taylor series generated by f at zero is $T(x) = 0$ and clearly $T(x) \neq f(x)$ for all x in any small open set centered at the origin. Thus f is smooth at zero, but f is not analytic at zero.

If the function f is analytic at x_o then f is represented by its Taylor series centered at x_o for x near x_o. To prove f is analytic at x_o we would need to apply Taylor's Theorem with remainder and show the power series converges to the given function for x sufficiently close to x_o. I'll probably illustrate such an argument in lecture, however for the most part I allow students to assume the functions we study are analytic at reasonable points and I do not expect students to give explicit proof of analyticity. We will assume the functions considered are analytic in what follows and I will cease belaboring this point since my main goal is for you to learn how to create and calculate power series.

Example 4.2. *Let $f(x) = (1+x)^\alpha$ where α is a constant. Then calculate,*

$$\begin{aligned}
&f'(x) = \alpha(1+x)^{\alpha-1}, && f'(0) = \alpha \\
&f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}, && f''(0) = \alpha(\alpha-1) \\
&f'''(x) = \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3}, \dots && f'''(0) = \alpha(\alpha-1)(\alpha-2) \\
&f^{(n)}(x) = \alpha(\alpha-1)\cdots(\alpha-(n-1))(1+x)^{\alpha-n}. && f^{(n)}(0) = \alpha(\alpha-1)\cdots(\alpha-(n-1))
\end{aligned}$$

Therefore, the Maclaurin series for $f(x)$ yields:

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-(n-1))}{n!}x^n = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots$$

The result above is known as the **binomial series**.

Digression: *let's investigate the application of our calculation to algebra. If $\alpha = N \in \mathbb{N}$ then notice that $f^{(n)}(0) = 0$ for all $n \geq N$ and we obtain an example of the Binomial Theorem. For example, $\alpha = 4$:*

$$(1+x)^4 = 1 + 4x + \frac{4(3)}{2}x^2 + \frac{4(3)(2)}{3!}x^3 + \frac{4(3)(2)(1)}{4!}x^4 = 1 + 4x + 6x^2 + 4x^3 + x^4.$$

You might recognize the coefficients in the expression above from Pascal's Triangle. The usual notation from combinatorics is read N choose k and we find a nice formula from a little algebra:

$$\binom{N}{k} = \frac{N(N-1)\cdots(N-(k-1))}{k!} = \frac{N!}{(N-k)!\,k!}$$

For example, returning to $\alpha = N = 4$ we have:

$$\binom{4}{0} = \frac{4!}{4!\,0!} = 1, \ \binom{4}{1} = \frac{4!}{3!\,1!} = 4, \ \binom{4}{2} = \frac{4!}{2!\,2!} = 6, \ \binom{4}{3} = \frac{4!}{1!\,3!} = 4, \ \binom{4}{4} = \frac{4!}{0!\,4!} = 1.$$

What do these numbers mean ? For example, if you have 4 toys and you wish to select 2 toys to give your friend then you have $\binom{4}{2} = \frac{4!}{2!\,2!} = 6$ possible selections you can give your friend. Labeling the toys A, B, C, D the six selections would be A, B or A, C or A, D or B, C or B, D or C, D. A more dramatic example, if you have 10 possible toppings to put on a pizza and you are selecting 4 distinct toppings to make the pizza then there are:

$$\binom{10}{4} = \frac{10!}{6!\,4!} = \frac{10(9)(8)(7)6!}{6!\,4!} = \frac{10(9)(8)(7)}{2(3)4} = 5 \cdot 3 \cdot 2 \cdot 7 = 210.$$

Example 4.3. *The formula for relativistic energy is $E = m\gamma c^2$ where $\gamma = \dfrac{1}{\sqrt{1 - v^2/c^2}}$. Here v is the velocity of the mass m and c is the speed of light. The binomial series for $\alpha = -1/2$ yields*

$$\begin{aligned}(1+u)^{-1/2} &= 1 - \frac{1}{2}u - \frac{1}{2!}\cdot\frac{1}{2}\left(\frac{-3}{2}\right)u^2 - \frac{1}{3!}\cdot\frac{1}{2}\left(\frac{-3}{2}\right)\left(\frac{-5}{2}\right)u^3 + \cdots \\ &= 1 - \frac{1}{2}u + \frac{3}{8}u^2 - \frac{5}{16}u^3 + \cdots\end{aligned}$$

thus substituting $u = -v^2/c^2$ gives us the following expansion for the relativistic energy:

$$E = mc^2\left(1 + \frac{v^2}{2c^2} + \frac{3v^4}{8c^4} + \frac{5v^6}{16c^6} + \cdots\right) = mc^2 + \frac{1}{2}mv^2 + \frac{3mv^4}{8c^2} + \frac{5mv^6}{16c^4} + \cdots.$$

The first term is Einstein's famous $E = mc^2$ formula which describes the rest energy of the mass m. Notice $mc^2 >> \frac{1}{2}mv^2$ for $v << c$, the rest energy is much larger than typical kinetic energies from low-velocity physics.

Example 4.4. *The gravitational acceleration at altitude h due to the gravitation from a planet with mass M and radius R is given by*

$$a(h) = -\frac{GM}{(R+h)^2}$$

It is interesting to study this formula in the context $h/R << 1$. Rewrite the formula for the acceleration in terms of the quantity h/R:

$$a(h) = \frac{GM}{R^2(1+h/R)^2} = \frac{GM}{R^2}\left(1 + \frac{h}{R}\right)^{-2} = \frac{GM}{R^2}\left(1 - 2\frac{h}{R} + \cdots\right)$$

where we applied the binomial series for $\alpha = -2$ to see $(1+u)^{-2} = 1 - 2u + \cdots$ with $u = h/R$. In Example 3.12 we noticed the quantity $\frac{GM}{R^2} = 9.83 m/s^2$ for the mass and radius of the earth. The leading error in simplifying the acceleration of gravity to be constant is given by $\frac{2GMh}{R^3}$. In other words, $a = 9.83\, m/s^2 \pm \frac{2GMh}{R^3}$ where $\frac{2GM}{R^3} = 3.1 \times 10^{-6}/s^2$.

5 Power Series Calculation

I do expect students to know Maclaurin series for the common elementary functions from memory. You should memorize and/or learn to derive all the following Maclaurin series. I assume you know these and you are allowed to apply these results to generate new results. This is **considerably** easier than direct generation of the Taylor series from repeated differentiation.

(1.) $e^u = \sum_{n=0}^{\infty} \frac{1}{n!}u^n = 1 + u + \frac{1}{2}u^2 + \frac{1}{6}u^3 + \frac{1}{24}u^4 + \cdots,$

(2.) $\cosh(u) = \sum_{k=0}^{\infty} \frac{1}{(2k)!}u^{2k} = 1 + \frac{1}{2}u^2 + \frac{1}{24}u^4 + \cdots,$

(3.) $\sinh(u) = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}u^{2k+1} = u + \frac{1}{6}u^3 + \frac{1}{120}u^5 + \cdots$

(4.) $\cos(u) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}u^{2k} = 1 - \frac{1}{2}u^2 + \frac{1}{24}u^4 + \cdots,$

(5.) $\sin(u) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}u^{2k+1} = u - \frac{1}{6}u^3 + \frac{1}{120}u^5 + \cdots.$

Now that we know these, let us make use of them to create new power series via simple substitution into the above formulas.

Example 5.1. *Set $u = x^2$ in the Maclaurin series for the exponential:*

$$e^{x^2} = \sum_{n=0}^{\infty} \frac{1}{n!}\left(x^2\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!}x^{2n} = 1 + x^2 + \frac{1}{4}x^4 + \frac{1}{6}x^6 + \frac{1}{24}x^8 + \cdots$$

Example 5.2. *Set $u = x^8$ in the Maclaurin series for sine:*

$$\sin\left(x^8\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\left(x^8\right)^{2k+1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{16k+8} = x^8 - \frac{1}{6}x^{24} + \frac{1}{120}x^{40} + \cdots$$

Example 5.3. *Set $u = x - 2$ in the Maclaurin series for hyperbolic sine:*

$$\sinh(x-2) = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}(x-2)^{2k+1} = x - 2 + \frac{1}{6}(x-2)^3 + \frac{1}{120}(x-2)^5 + \cdots$$

Example 5.4. *Set $u = x^3$ in the Maclaurin series for hyperbolic cosine:*

$$\cosh\left(x^3\right) = \sum_{k=0}^{\infty} \frac{1}{(2k)!}\left(x^3\right)^{2k} = \sum_{k=0}^{\infty} \frac{1}{(2k)!}x^{6k} = 1 + \frac{1}{2}x^6 + \frac{1}{24}x^{12} + \cdots$$

then we can multiply by x^5 to create the power series for $x^5\cosh\left(x^3\right)$

$$x^5\cosh\left(x^3\right) = \sum_{k=0}^{\infty} \frac{1}{(2k)!}x^5\left(x^3\right)^{2k} = \sum_{k=0}^{\infty} \frac{1}{(2k)!}x^{6k+5} = x^5 + \frac{1}{2}x^{11} + \frac{1}{24}x^{17} + \cdots$$

Notice the examples above would be far more difficult if you attempted to directly use Taylor's Theorem (stated below) to generate the power series.

Theorem 5.5. *Taylor's theorem for our convenience*

If f is analytic at $x = a$ then

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots$$

Furthermore, if $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ then $c_n = \frac{f^{(n)}(a)}{n!}$ for $n = 0, 1, 2, \ldots$.

Let's work backwards from an earlier example to use the power series to calculate an absurdly large derivative.

Example 5.6.

$$f(x) = x^5 \cosh\left(x^3\right) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} x^{6k+5} = x^5 + \frac{1}{2}x^{11} + \frac{1}{24}x^{17} + \cdots$$

Since $f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \cdots + \frac{f^{(5)}(0)}{5!}x^5 + \cdots + \frac{1}{11!}x^{11} + \cdots$ *we can compare with the given Maclaurin series to find from equating the coefficients of* x^5 *and* x^{11} *the following:*

$$\frac{f^{(5)}(0)}{5!} = 1, \qquad \& \qquad \frac{f^{(11)}(0)}{11!} = \frac{1}{24}.$$

Thus $f^{(5)}(0) = 5! = 120$ *and* $f^{(11)}(0) = \frac{11!}{24} = 11(10)(9)(8)(7)(6)5$*. On the other hand,* $\frac{f^{(n)}(0)}{n!} = 0$ *for* $n = 0, 1, 2, 3, 4, 6, 7, 8, 9, 10$ *and thus* $f^{(n)}(0) = 0$ *in all those cases. Continuing with this line of investigation, let us calculate* $f^{(n)}(0)$ *for* $n = 65, 66, 67, 68, 69, 70, 71$*. Observe that only powers of the form* x^{6k+5} *are nontrivial in the given Maclaurin series. Solve* $6k + 5 = 65$ *to see* $k = 10$ *then* $k = 11$ *is the next possible matching power with* $6(11) + 5 = 71$*. There are no nontrivial terms of the form* $x^{66}, x^{67}, x^{68}, x^{69}, x^{70}$ *thus we find from Taylor's Theorem,*

$$f^{(66)}(0) = f^{(67)}(0) = f^{(68)}(0) = f^{(69)}(0) = f^{(70)}(0) = 0.$$

On the other hand, from $k = 10$ *we study the coefficient of* x^{65}*:*

$$\frac{f^{(65)}(0)}{65!} = \frac{1}{20!} \quad \Rightarrow \quad f^{(65)}(0) = \frac{65!}{20!}.$$

Similarly, from $k = 11$ *we study the coefficient of* x^{71}*,*

$$\frac{f^{(71)}(0)}{71!} = \frac{1}{22!} \quad \Rightarrow \quad f^{(71)}(0) = \frac{71!}{22!}.$$

Pause to think about finding these derivatives through direct calculation.

6 Multiplication and Division of Series

Multiplication of series is defined in the natural fashion:

$$(a_0 + a_1 + a_2 + \cdots)(b_0 + b_1 + b_2 + \cdots) = a_0b_0 + a_0b_1 + a_1b_0 + a_0b_2 + a_1b_1 + a_2b_0 + \cdots$$

We organize the products such that we group terms such that the sum of their indices is constant. This is known as the **Cauchy Product**. We can prove the Cauchy Product of two absolutely convergent series is a convergent series[6].

Definition 6.1. *product of series*

Given absolutely convergent series $\sum_{n=0}^{\infty} A_n$ and $\sum_{j=0}^{\infty} B_j$ we define:

$$\left(\sum_{n=0}^{\infty} A_n\right)\left(\sum_{j=0}^{\infty} B_j\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} A_k B_{n-k}$$

Our primary interest is in the product of power series. Notice that power series are absolutely convergent on their open interval of convergence so we are certainly free to calculate the product of power series using the Cauchy product. Let us examine how the definition above looks when we apply it to power series: identify $A_n = a_n(x - x_1)^n$ whereas $B_j = b_j(x - x_2)^j$,

$$\left(\sum_{n=0}^{\infty} a_n(x - x_1)^n\right)\left(\sum_{j=0}^{\infty} b_j(x - x_2)^j\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k b_{n-k}(x - x_1)^k(x - x_2)^{n-k}$$

Usually we multiply series with the same center $x_1 = x_2$ hence

$$\left(\sum_{n=0}^{\infty} a_n(x - x_1)^n\right)\left(\sum_{j=0}^{\infty} b_j(x - x_1)^j\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k b_{n-k}(x - x_1)^n$$

Example 6.2. *We use the binomial theorem* $(x + y)^n = \sum_{k=0}^{n} \frac{n!}{(n-k)!\,k!} x^k y^{n-k}$ *for last step:*

$$\begin{aligned}\left(\sum_{n=0}^{\infty} \frac{1}{n!}x^n\right)\left(\sum_{j=0}^{\infty} \frac{1}{j!}y^j\right) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{1}{k!}\frac{1}{(n-k)!}x^k y^{n-k} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}\sum_{k=0}^{n} \frac{n!}{(n-k)!k!}x^k y^{n-k} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(x + y)^n\end{aligned}$$

The calculation above demonstrates that $e^x e^y = e^{x+y}$*. If we defined the exponential by its Maclaurin series then this is a natural argument to prove the law of exponents.*

[6] forgive me for omitting this proof in the current notes

Example 6.3. *A fun example I usually cover is as follows:*

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

Of course this is not surprising since $\frac{1}{1-x} = 1 + x + x^2 + \cdots$ *by the geometric series.*

Division of power series is an important calculational techinque. We can either use long division as I demonstrated in lecture or we can use a recursive technique as I illustrate in the abtract next. Consider $\sum_{n=0}^{\infty} a_n x^n$. We seek to find $b_0, b_1, \dots$ for which $\frac{1}{\sum_{n=0}^{\infty} a_n x^n} = \sum_{n=0}^{\infty} b_n x^n$. In other words, we wish to solve:

$$\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = 1$$

for appropriate values of $b_0, b_1, b_2, \dots$ in terms of the given coefficients $a_0, a_1, a_2, \dots$. Let us expand up to cubic order,

$$a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0)x^3 + \cdots = 1$$

At each order we find an equation,

$$a_0b_0 = 1, \ a_0b_1 + a_1b_0 = 0, \ a_0b_2 + a_1b_1 + a_2b_0 = 0, \ a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0 = 0, \dots$$

Apparently we must require $a_0 \neq 0$ if we are to find a multiplicative inverse for the series $a_0 + a_1x + a_2x^2 + \cdots$. Hence,

$$\boxed{b_0 = \frac{1}{a_0}.}$$

Next,

$$a_0b_1 + a_1b_0 = 0 \ \Rightarrow \ b_1 = -\frac{a_1b_0}{a_0} = -\frac{a_1}{a_0^2} \ \Rightarrow \ \boxed{b_1 = -\frac{a_1}{a_0^2}.}$$

Continuing to the quadratic terms:

$$a_0b_2 + a_1b_1 + a_2b_0 = 0 \ \Rightarrow \ b_2 = -\frac{a_1b_1 + a_2b_0}{a_0} \ \Rightarrow \ \boxed{b_2 = \frac{a_1^2}{a_0^3} - \frac{a_2}{a_0^2}.}$$

Finally the cubic terms:

$$a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0 = 0 \ \Rightarrow \ b_3 = -\frac{a_1b_2 + a_2b_1 + a_3b_0}{a_0} \ \Rightarrow \ b_3 = -\frac{a_1b_2}{a_0} - \frac{a_2b_1}{a_0} - \frac{a_3b_0}{a_0}$$

thus

$$b_3 = -\frac{a_1}{a_0}\left[\frac{a_1^2}{a_0^3} - \frac{a_2}{a_0^2}\right] - \frac{a_2}{a_0}\left[-\frac{a_1}{a_0^2}\right] - \frac{a_3}{a_0^2} \ \Rightarrow \ \boxed{b_3 = -\frac{a_1^3}{a_0^4} + \frac{2a_1a_2}{a_0^3} - \frac{a_3}{a_0^2}.}$$

Example 6.4. *Identify* $\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + \cdots$ *gives* $a_0 = 1, a_1 = 0, a_2 = -1/2, a_3 = 0$ *hence:*

$$\frac{1}{\cos x} = 1 + \frac{1}{2}x^2 + \cdots.$$

We could go on to find further formulas, certainly I don't expect you to memorize these. Rather, I'm doing this calculation as a model for how we can carry out similar calculations for examples.

Example 6.5. *Consider* $\tan x = \dfrac{\sin x}{\cos x} = b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots$. *Thus we wish to solve* $\sin x = (b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots)\cos x$. *Hence study:*

$$x - \frac{1}{6}x^3 - \frac{1}{120}x^5 + \cdots = \big(b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4 + b_5x^5\cdots\big)\left(1 - \frac{1}{2}x^2 + \frac{1}{24}x^4\cdots\right)$$

Equate coefficients of each order to find formulas for $b_0, b_1, \ldots,$

$$\begin{array}{llll} x^0 & : & 0 & = b_0, \\ x^1 & : & 1 & = b_1, \\ x^2 & : & 0 & = -\frac{b_0}{2} + b_2, \\ x^3 & : & \frac{-1}{6} & = -\frac{b_1}{2} + b_3, \\ x^4 & : & 0 & = \frac{b_0}{24} - \frac{b_2}{2} + b_4, \\ x^5 & : & \frac{1}{120} & = \frac{b_1}{24} - \frac{b_3}{2} + b_5, \end{array}$$

We can solve the equations above top-down,

$$b_0 = 0,\ b_1 = 1,\ b_2 = 0,\ b_3 = \frac{1}{2} - \frac{1}{6} = \frac{1}{3},\ b_4 = 0,$$

and

$$b_5 = \frac{1}{120} - \frac{b_1}{24} + \frac{b_3}{2} = \frac{1}{120} - \frac{1}{24} + \frac{1}{6} = \frac{2}{15}$$

Putting it all together,

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots$$

If you ask Wolfram Alpha nicely, it will tell you:

$$\tan x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7 + \frac{62}{2835}x^9 + \frac{1382}{155925}x^{11} + \cdots$$

I intend to derive the previous example via long division in lecture. Let me record a few additional results without proof for our reference.

Example 6.6.

$$\tanh x = \frac{\sinh x}{\cosh x} = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \frac{17}{315}x^7 + \frac{62}{2835}x^9\cdots$$

Example 6.7.

$$\sec x = 1 + \frac{1}{2}x^2 + \frac{5}{24}x^4 + \frac{61}{720}x^6 + \cdots$$

Example 6.8.

$$\frac{5}{x^2 + 4x + 5} = 1 - \frac{4}{5}x + \frac{11}{25}x^2 - \frac{42}{125}x^3 + \frac{41}{625}x^4 + \cdots$$

Example 6.9. *The following result we can obtain from multiplying the maclaurin series for tangent with itself:*

$$\tan^2 x = x^2 + \frac{2}{3}x^4 + \frac{17}{45}x^6 + \frac{62}{315}x^8 + \cdots$$

Example 6.10. *Finally, an example where Taylor's Theorem is the tool to use to derive:*

$$\sinh^{-1}(x) = x - \frac{1}{6}x^3 + \frac{3}{40}x^5 + \cdots$$

or... is it ? What is another approach we could take ? If $f(x) = \sinh^{-1}(x)$ *then* $\dfrac{df}{dx} = \dfrac{1}{\sqrt{1+x^2}}$. *What next ? Think about a binomial series substitution followed by integration. We should work it out in lecture.*

7 Power Series Solution to Differential Equation

In this section we demonstrate a simple, but powerful, technique to solve a wide class of differential equations. In short, assume the solution is a power series, plug-it-in, and work out the consequences.

Example 7.1. *Consider $\frac{dy}{dx} = y$. Of course we already know how to solve this by separation of variables, but this is a particularly simple problem to illustrate the technique so we begin here. Let $y = \sum_{n=0}^{\infty} a_n x^n$ then $\frac{dy}{dx} = \sum_{n=1}^{\infty} n a_n x^{n-1}$. Therefore,*

$$\sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} a_n x^n \;\star$$

It is helpful to relable $n-1=k$ hence $n=1$ gives $k=0$ and thus

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

It follows that: (I made an $n=k$ substitution to the RHS of $\star$ to match the indices)

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

It follows that $a_{k+1}(k+1) = a_k$ for $k = 0, 1, 2, \dots$. Thus,

$$a_{k+1} = \frac{1}{k+1} a_k$$

then

$$a_1 = a_0$$
$$a_2 = \frac{1}{2} a_1 = \frac{1}{2} a_0$$
$$a_3 = \frac{1}{3} a_2 = \frac{1}{3} \cdot \frac{1}{2} a_0$$
$$a_4 = \frac{1}{4} a_3 = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} a_0$$

A pattern is apparent:

$$a_n = \frac{1}{n} \cdot \frac{1}{n-1} \cdots \frac{1}{3} \cdots \frac{1}{2} a_0 = \frac{a_0}{n!}$$

Thus we find solution

$$y = \sum_{n=0}^{\infty} \frac{a_0}{n!} x^n = a_0 \sum_{n=0}^{\infty} \frac{1}{n!} x^n = a_0 e^x.$$

This is not the best way to solve $\frac{dy}{dx} = y$. Notice, we could just separate variables to $\frac{dy}{y} = dx$ then integrate to find $\ln|y| = x + c$ hence $|y| = e^{x+c}$ from which we find $y = \pm e^c e^x$ thus $y = a_0 e^x$ if we set $\pm e^c = a_0$. The next example is not as easy to solve by our previous methods.

Example 7.2. *Solve $y'' - y = 0$. Once more we will posit our solution can be written in the form of a Maclaurin series: $y = \sum_{n=0}^{\infty} a_n x^n$. Differentiate twice:*

$$y'' = \frac{d}{dx}\frac{d}{dx}\sum_{n=0}^{\infty} a_n x^n = \frac{d}{dx}\sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}$$

Let $k = n-2$ then $n = k+2$ and $n-1 = k+1$. Also, $k = 0$ for $n = 2$ and we find

$$y'' = \sum_{k=0}^{\infty}(k+2)(k+1)a_{k+2}x^k = \sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}x^n$$

Consequently, $y'' - y = 0$ gives $y'' = y$ which implies

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

from which we find $(n+2)(n+1)a_{n+2} = a_n$ for $n = 0, 1, 2, \dots$. Thus,

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

Thus, $n = 0$ shows

$$a_2 = \frac{a_0}{2}$$
$$a_4 = \frac{a_2}{4(3)} = \frac{a_0}{4(3)(2)} = \frac{a_0}{4!}$$
$$a_6 = \frac{a_4}{6(5)} = \frac{a_0}{6(5)4!} = \frac{a_0}{6!}$$

Apparently, $a_{2k} = \dfrac{a_0}{(2k)!}$ for $k = 0, 1, \dots$ this covers even n. Odd n follow a similar pattern:

$$a_3 = \frac{a_1}{3(2)} = \frac{a_1}{3!}$$
$$a_5 = \frac{a_3}{5(4)} = \frac{a_1}{5(4)3!} = \frac{a_1}{5!}$$
$$a_7 = \frac{a_5}{7(6)} = \frac{a_1}{7(6)5!} = \frac{a_1}{7!}$$

and we find $a_{2k+1} = \dfrac{a_1}{(2k+1)!}$ for $k = 0, 1, \dots$. Thus we find the general solution of the given second order differential equation breaks into an even and odd function as shown below:

$$y = \sum_{n=0}^{\infty} a_n x^n = \sum_{k=0}^{\infty}\frac{a_0}{(2k)!}x^{2k} + \sum_{k=0}^{\infty}\frac{a_1}{(2k+1)!}x^{2k+1} = a_0\left(\sum_{k=0}^{\infty}\frac{1}{(2k)!}x^{2k}\right) + a_1\left(\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}x^{2k+1}\right).$$

Thus $y = a_o \cosh x + a_1 \sinh x$. I may have given you a similar homework problem where $y'' + y = 0$. The solution is very similar, except the sign from $y'' = -y$ introduces an alternating sign in the formulas for a_n which ultimately changes hyperbolic cosine to cosine and hyperbolic sine to sine. Certainly you can use this example as a template for how to solve $y'' + y = 0$ via the power series technique.