Chapter 10

Series and Approximations

An important theme in this book is to give **constructive** definitions of mathematical objects. Thus, for instance, if you needed to evaluate

$$\int_0^1 e^{-x^2}\,dx,$$

you could set up a Riemann sum to evaluate this expression to any desired degree of accuracy. Similarly, if you wanted to evaluate a quantity like $e^{.3}$ from first principles, you could apply Euler's method to approximate the solution to the differential equation

$$y'(t) = y(t), \text{ with initial condition } y(0) = 1,$$

using small enough intervals to get a value for $y(.3)$ to the number of decimal places you needed. You might pause for a moment to think how you would get $\sin(5)$ to 7 decimal places—you wouldn't do it by drawing a unit circle and measuring the y-coordinate of the point where this circle is intersected by the line making an angle of 5 radians with the x-axis! Defining the sine function to be the solution to the second-order differential equation $y'' = -y$ with initial conditions $y = 0$ and $y' = 1$ when $t = 0$ is much better if we actually want to construct values of the function with more than two decimal accuracy.

Ordinary arithmetic lies at the heart of all calculations

What these examples illustrate is the fact that the only functions our brains or digital computers can evaluate directly are those involving the arithmetic operations of addition, subtraction, multiplication, and division. Anything else we or computers evaluate must ultimately be reducible to these four operations. But the only functions directly expressible in such terms are polynomials and rational functions (i.e., quotients of one polynomial by another). When you use your calculator to evaluate $\ln 2$, and the calculator shows .69314718056, it is really doing some additions, subtractions, multiplications, and divisions to compute this 11-digit *approximation* to $\ln 2$. There are no obvious connections to logarithms at all in what it does. One of the triumphs of calculus is the development of techniques for calculating highly accurate approximations of this sort quickly. In this chapter we will explore these techniques and their applications.

10.1 Approximation Near a Point or Over an Interval

Suppose we were interested in approximating the sine function—we might need to make a quick estimate and not have a calculator handy, or we might even be designing a calculator. In the next section we will examine a number of other contexts in which such approximations are helpful. Here is a third degree polynomial that is a good approximation in a sense which will be made clear shortly:

$$P(x) = x - \frac{x^3}{6}.$$

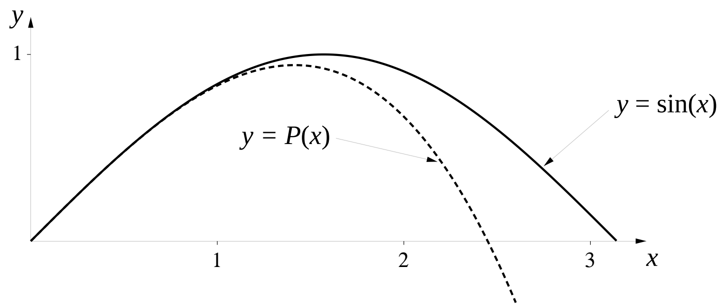
(You will see in section 2 where $P(x)$ comes from.)

If we compare the values of $\sin(x)$ and $P(x)$ over the interval $[0, 1]$ we get the following:

x	$\sin x$	$P(x)$	$\sin x - P(x)$
0.0	0.0	0.0	0.0
.2	.198669	.198667	.000002
.4	.389418	.389333	.000085
.6	.564642	.564000	.000642
.8	.717356	.714667	.002689
1.0	.841471	.833333	.008138

The fit is good, with the largest difference occurring at $x = 1.0$, where the difference is only slightly greater than .008.

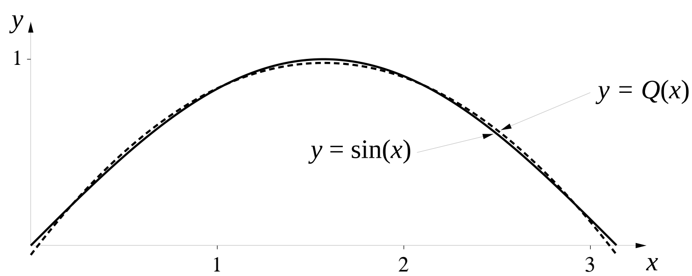
If we plot $\sin(x)$ and $P(x)$ together over the interval $[0, \pi]$ we see the ways in which $P(x)$ is both very good and not so good. Over the initial portion of the graph—out to around $x = 1$—the graphs of the two functions seem to coincide. As we move further from the origin, though, the graphs separate more and more. Thus if we were primarily interested in approximating $\sin(x)$ near the origin, $P(x)$ would be a reasonable choice. If we need to approximate $\sin(x)$ over the entire interval, $P(x)$ is less useful.



On the other hand, consider the second degree polynomial

$$Q(x) = -.4176977x^2 + 1.312236205x - .050465497$$

(You will see how to compute these coefficients in section 6.) When we graph $Q(x)$ and $\sin(x)$ together we get the following:



While $Q(x)$ does not fits the graph of $\sin(x)$ as well as $P(x)$ does near the origin, it is a good fit overall. In fact, $Q(x)$ exactly equals $\sin(x)$ at 4 values of x, and the greatest separation between the graphs of $Q(x)$ and $\sin(x)$ over the interval $[0, \pi]$ occurs at the endpoints, where the distance between the graphs is .0505 units.

What we have here, then, are two kinds of approximation of the sine function by polynomials: we have a polynomial $P(x)$ that behaves very much like the sine function near the origin, and we have another polynomial $Q(x)$ that keeps close to the sine function over the entire interval $[0, \pi]$. Which one is the "better" approximation depends on our needs. Each solves an important problem. Since finding approximations near a point has a neater

There's more than one way to make the "best fit" to a given curve solution—Taylor polynomials—we will start with this problem. We will turn to the problem of finding approximations over an interval in section 6.

10.2 Taylor Polynomials

The general setting. In chapter 3 we discovered that functions were locally linear at most points—when we zoomed in on them they looked more and more like straight lines. This fact was central to the development of much of the subsequent material. It turns out that this is only the initial manifestation of an even deeper phenomenon: Not only are functions locally linear, but, if we don't zoom in quite so far, they look locally like parabolas. From a little further back still they look locally like cubic polynomials, etc. Later in this section we will see how to use the computer to visualize these "local parabolizations", "local cubicizations", etc. Let's summarize the idea and then explore its significance:

> The functions of interest to calculus look locally like polynomials at most points of their domain. The higher the degree of the polynomial, the better typically will be the fit.

Comments The "at most points" qualification is because of exceptions like those we ran into when we explored locally linearity. The function $|x|$, for instance, was not locally linear at $x = 0$—it's not locally like any polynomial of higher degree at that point either. The issue of what "goodness of fit" means and how it is measured is a subtle one which we will develop over the course of this section. For the time being, your intuition is a reasonable guide—one fit to a curve is better than another near some point if it "shares more phosphor" with the curve when they are graphed on a computer screen centered at the given point.

The behavior of a function can often be inferred from the behavior of a local polynomialization

The fact that functions look locally like polynomials has profound implications conceptually and computationally. It means we can often determine the behavior of a function locally by examining the corresponding behavior of what we might call a "local polynomialization" instead. In particular, to find the values of a function near some point, or graph a function near some point, we can deal with the values or graph of a local polynomialization instead. Since we can actually evaluate polynomials directly, this can be a major simplification.

There is an extra feature to all this which makes the concept particularly attractive: not only are functions locally polynomial, it is easy to find the coefficients of the polynomials. Let's see how this works. Suppose we had some function $f(x)$ and we wanted to find the fifth degree polynomial that best fit this function at $x = 0$. Let's call this polynomial

We want the best fit at $x = 0$

$$P(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5.$$

To determine P, we need to find values for the six coefficients a_0, a_1, a_2, a_3, a_4, a_5.

Before we can do this, we need to define what we mean by the "best" fit to f at $x = 0$. Since we have six unknowns, we need six conditions. One obvious condition is that the graph of P should pass through the point $(0, f(0))$. But this is equivalent to requiring that $P(0) = f(0)$. Since $P(0) = a_0$, we thus must have $a_0 = f(0)$, and we have found one of the coefficients of $P(x)$. Let's summarize the argument so far:

The best fit should pass through the point $(0, f(0))$

> The graph of a polynomial passes through the point $(0, f(0))$ if and only if the polynomial is of the form
>
> $$f(0) + a_1x + a_2x^2 + \cdots .$$

But we're not interested in just any polynomial passing through the right point; it should be headed in the right direction as well. That is, we want the slope of P at $x = 0$ to be the same as the slope of f at this point—we want $P'(0) = f'(0)$. But

The best fit should have the right slope at $(0, f(0))$

$$P'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4,$$

so $P'(0) = a_1$. Our second condition therefore must be that $a_1 = f'(0)$. Again, we can summarize this as

> The graph of a polynomial passes through the point $(0, f(0))$ and has slope $f'(0)$ there if and only if it is of the form
>
> $$f(0) + f'(0)x + a_2x^2 + \cdots .$$

Note that at this point we have recovered the general form for the local linear approximation to f at $x = 0$: $L(x) = f(0) + f'(0)x$.

But there is no reason to stop with the first derivative. Similarly, we would want the way in which the slope of $P(x)$ is changing—we are now talking about $P''(0)$—to behave the way the slope of f is changing at $x = 0$, etc. Each higher derivative controls a more subtle feature of the shape of the graph. We now see how we could formulate reasonable additional conditions which would determine the remaining coefficients of $P(x)$:

> Say that $P(x)$ is the **best fit** to $f(x)$ at the point $x = 0$ if
> $P(0) = f(0),\ P'(0) = f'(0),\ P''(0) = f''(0), \ldots, P^{(5)}(0) = f^{(5)}(0).$

The final criterion for best fit at $x = 0$

Since $P(x)$ is a fifth degree polynomial, all the derivatives of P beyond the fifth will be identically 0, so we can't control their values by altering the values of the a_k. What we are saying, then, is that we are using as our criterion for the best fit that all the derivatives of P as high as we can control them have the same values at $x = 0$ as the corresponding derivatives of f.

While this is a reasonable definition for something we might call the "best fit" at the point $x = 0$, it gives us no direct way to tell how good the fit really is. This is a serious shortcoming—if we want to approximate function values by polynomial values, for instance, we would like to know how many decimal places in the polynomial values are going to be correct. We will take up this question of goodness of fit later in this section; we'll be able to make measurements that allow us to to see how well the polynomial fits the function. First, though, we need to see how to determine the coefficients of the approximating polynomials and get some practice manipulating them.

Notation for higher derivatives

Note on Notation: We have used the notation $f^{(5)}(x)$ to denote the fifth derivative of $f(x)$ as a convenient shorthand for $f'''''(x)$, which is harder to read. We will use this throughout.

Finding the coefficients We first observe that the derivatives of P at $x = 0$ are easy to express in terms of $a_1, a_2, \ldots$. We have

$$\begin{aligned}
P'(x) &= a_1 + 2\,a_2x + 3\,a_3x^2 + 4\,a_4x^3 + 5\,a_5x^4,\\
P''(x) &= 2\,a_2 + 3\cdot 2\,a_3x + 4\cdot 3\,a_4x^2 + 5\cdot 4\,a_5x^3,\\
P^{(3)}(x) &= 3\cdot 2\,a_3 + 4\cdot 3\cdot 2\,a_4x + 5\cdot 4\cdot 3\,a_5x^2,\\
P^{(4)}(x) &= 4\cdot 3\cdot 2\,a_4 + 5\cdot 4\cdot 3\cdot 2\,a_5x,\\
P^{(5)}(x) &= 5\cdot 4\cdot 3\cdot 2\,a_5.
\end{aligned}$$

Thus $P''(0) = 2\,a_2$, $P^{(3)}(0) = 3 \cdot 2\,a_3$, $P^{(4)}(0) = 4 \cdot 3 \cdot 2\,a_4$, and $P^{(5)}(0) = 5 \cdot 4 \cdot 3 \cdot 2\,a_5$.

We can simplify this a bit by introducing the **factorial** notation, in which we write $n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$. This is called "$n$ factorial". Thus, for example, $7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$. It turns out to be convenient to extend the factorial notation to 0 by defining $0! = 1$. (Notice, for instance, that this makes the formulas below work out right.) In the exercises you will see why this extension of the notation is not only convenient, but reasonable as well!

Factorial notation

With this notation we can express compactly the equations above as $P^{(k)}(0) = k!\,a_k$ for $k = 0, 1, 2, \ldots 5$. Finally, since we want $P^{(k)}(0) = f^{(k)}(0)$, we can solve for the coefficients of $P(x)$:

The desired rule for finding the coefficients

$$a_k = \frac{f^{(k)}(0)}{k!} \quad \text{for } k = 0, 1, 2, 3, 4, 5.$$

We can now write down an explicit formula for the fifth degree polynomial which best fits $f(x)$ at $x = 0$ in the sense we've put forth:

$$P(x) = f(0) + f'(0)x + \frac{f^{(2)}(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \frac{f^{(5)}(0)}{5!}x^5.$$

We can express this more compactly using the Σ–notation we introduced in the discussion of Riemann sums in chapter 6:

$$P(x) = \sum_{k=0}^{5} \frac{f^{(k)}(0)}{k!}\,x^k.$$

We call this the **fifth degree Taylor polynomial for** $f(x)$. It is sometimes also called the fifth **order** Taylor polynomial.

It should be obvious to you that we can generalize what we've done above to get a best fitting polynomial of any degree. Thus

> The **Taylor polynomial of degree n** approximating the function $f(x)$ at $x = 0$ is given by the formula
>
> $$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}\,x^k.$$

General rule for the Taylor polynomial at $x = 0$

We also speak of the Taylor polynomial *centered at* $x = 0$.

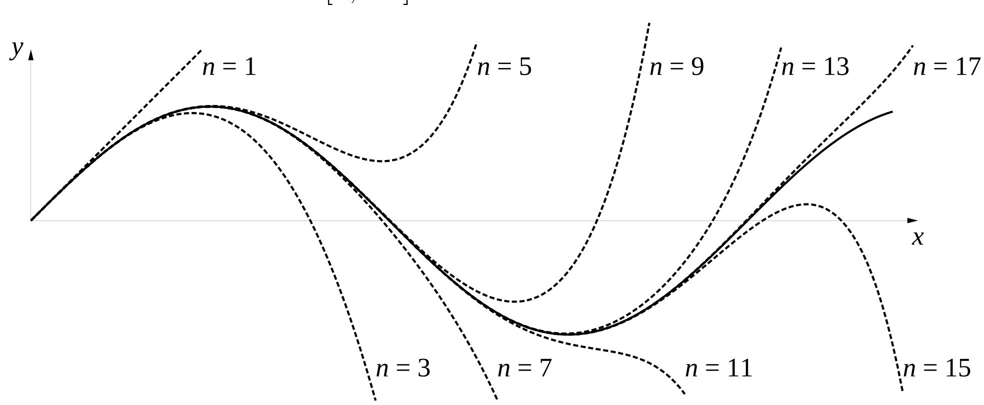
Example. Consider $f(x) = \sin(x)$. Then for $n = 7$ we have

$$\begin{aligned}
f(x) &= \sin(x), & f(0) &= 0,\\
f'(x) &= \cos(x), & f'(0) &= +1,\\
f^{(2)}(x) &= -\sin(x), & f^{(2)}(0) &= 0,\\
f^{(3)}(x) &= -\cos(x), & f^{(3)}(0) &= -1,\\
f^{(4)}(x) &= \sin(x), & f^{(4)}(0) &= 0,\\
f^{(5)}(x) &= \cos(x), & f^{(5)}(0) &= +1,\\
f^{(6)}(x) &= -\sin(x), & f^{(6)}(0) &= 0,\\
f^{(7)}(x) &= -\cos(x), & f^{(7)}(0) &= -1.
\end{aligned}$$

From this we can see that the pattern 0, +1, 0 , −1, . . . will repeat forever. Substituting these values into the formula we get that for any odd integer n the n-th degree Taylor polynomial for $\sin(x)$ is

$$P_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \pm x^n n!.$$

Note that $P_3(x) = x - x^3/6$, which is the polynomial we met in section 1. We saw there that this polynomial seemed to fit the graph of the sine function only out to around $x = 1$. Now, though, we have a way to generate polynomial approximations of higher degrees, and we would expect to get better fits as the degree of the approximating polynomial is increased. To see how closely these polynomial approximations follow $\sin(x)$, here's the graph of $\sin(x)$ together with the Taylor polynomials of degrees $n = 1, 3, 5, \ldots, 17$ plotted over the interval $[0, 7.5]$:



While each polynomial eventually wanders off to infinity, successive polynomials stay close to the sine function for longer and longer intervals—the Taylor polynomial of degree 17 is just beginning to diverge visibly by the time x reaches 2π. We might expect that if we kept going, we could find Taylor polynomials that were good fits out to $x = 100$, or $x = 1000$. This is indeed the case, although they would be long and cumbersome polynomials to work with. Fortunately, as you will see in the exercises, with a little cleverness we can use a Taylor polynomial of degree 9 to calculate $\sin(100)$ to 5 decimal place accuracy.

The higher the degree of the polynomial, the better the fit

Other Taylor Polynomials: In a similar fashion, we can get Taylor polynomials for other functions. You should use the general formula to verify the Taylor polynomials for the following basic functions. (The Taylor polynomial for $\sin(x)$ is included for convenient reference.)

Approximating polynomials for other basic functions

$f(x)$	$P_n(x)$
$\sin(x)$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \pm \frac{x^n}{n!}$ (n odd)
$\cos(x)$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \pm \frac{x^n}{n!}$ (n even)
e^x	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots + \frac{x^n}{n!}$
$\ln(1-x)$	$-\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots + \frac{x^n}{n}\right)$
$\frac{1}{1-x}$	$1 + x + x^2 + x^3 + \cdots + x^n$

Taylor polynomials at points other than $x = 0$. Using exactly the same arguments we used to develop the best-fitting polynomial at $x = 0$, we can derive the more general formula for the best-fitting polynomial at any value of x. Thus, if we know the behavior of f and its derivatives at some point $x = a$, we would like to find a polynomial $P_n(x)$ which is a good approximation to $f(x)$ for values of x close to a.

General rule for the Taylor polynomial at $x = a$

Since the expression $x - a$ tells us how close x is to a, we use it (instead of the variable x itself) to construct the polynomials approximating f at $x = a$:

$$P_n(x) = b_0 + b_1(x-a) + b_2(x-a)^2 + b_3(x-a)^3 + \cdots + b_n(x-a)^n.$$

You should be able to apply the reasoning we used above to derive the following:

> The **Taylor polynomial of degree n centered at $x = a$** approximating the function $f(x)$ is given by the formula
>
> $$P_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^n(a)}{n!}(x-a)^n$$
> $$= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Program: TAYLOR

Set up GRAPHICS
```
DEF fnfact(m)
     P = 1
     FOR r = 2 TO m
         P = P * r
     NEXT r
     fnfact = P
END DEF
DEF fnpoly(x)
     Sum = x
     Sign = -1
     FOR k = 3 TO 17 STEP 2
         Sum = Sum + Sign * x^k/fnfact(k)
         Sign = (-1) * Sign
     NEXT k
     fnpoly = Sum
END DEF
FOR x = 0 TO 3.14 STEP .01
```
Plot the line from `(x, fnpoly(x))` *to* `(x + .01, fnpoly(x + .01))`
```
NEXT x
```

A computer program for graphing Taylor polynomials Shown above is a program that evaluates the 17-th degree Taylor polynomial for $\sin(x)$ and graphs it over the interval $[0, 3.14]$. The first seven lines of the program constitute a subroutine for evaluating factorials. The syntax of such subroutines varies from one computer language to another, so be sure to use the format that's appropriate for you. You may even be using a language that already knows how to compute factorials, in which case you can omit the subroutine. The second set of 9 lines defines the function `poly` which evaluates the 17-th degree Taylor polynomial. Note the role of the variable `Sign`—it simply changes the sign back and forth from positive to negative as each new term is added to the sum. As usual, you will have to put in commands to set up the graphics and draw lines in the format your computer language uses. You can modify this program to graph other Taylor polynomials.

New Taylor Polynomials from Old

Given a function we want to approximate by Taylor polynomials, we could always go straight to the general formula for deriving such polynomials. On the other hand, it is often possible to avoid a lot of tedious calculation of derivatives by using a polynomial we've already calculated. It turns out that any manipulation on Taylor polynomials you might be tempted to try will probably work. Here are some examples to illustrate the kinds of manipulations that can be performed on Taylor polynomials.

Substitution in Taylor Polynomials. Suppose we wanted the Taylor polynomial for e^{x^2}. We know from what we've already done that for any value of u close to 0,

$$e^u \approx 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \frac{u^4}{4!} + \cdots + \frac{u^n}{n!}.$$

In this expression u can be anything, including another variable expression. For instance, if we set $u = x^2$, we get the Taylor polynomial

$$\begin{aligned}
e^{x^2} &= e^u \\
&\approx 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \frac{u^4}{4!} + \cdots + \frac{u^n}{n!} \\
&= 1 + (x^2) + \frac{(x^2)^2}{2!} + \frac{(x^2)^3}{3!} + \frac{(x^2)^4}{4!} + \cdots + \frac{(x^2)^n}{n!} \\
&= 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \frac{x^8}{4!} + \cdots + \frac{x^{2n}}{n!}.
\end{aligned}$$

You should check to see that this is what you get if you apply the general formula for computing Taylor polynomials to the function e^{x^2}.

Similarly, suppose we wanted a Taylor polynomial for $1/(1+x^2)$. We could start with the approximation given earlier:

$$\frac{1}{1-u} \approx 1 + u + u^2 + u^3 + \cdots + u^n.$$

If we now replace u everywhere by $-x^2$, we get the desired expansion:

$$\begin{aligned}
\frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \frac{1}{1-u} & \\
&\approx 1 + u + u^2 + u^3 + \cdots + u^n \\
&= 1 + (-x^2) + (-x^2)^2 + (-x^2)^3 + \cdots + (-x^2)^n \\
&= 1 - x^2 + x^4 - x^6 + \cdots \pm x^{2n}.
\end{aligned}$$

Again, you should verify that if you start with $f(x) = 1/(1+x^2)$ and apply to f the general formula for deriving Taylor polynomials, you will get the preceding result. Which method is quicker?

Multiplying Taylor Polynomials. Suppose we wanted the 5-th degree Taylor polynomial for $e^{3x} \cdot \sin(2x)$. We can use substitution to write down polynomial approximations for e^{3x} and $\sin(2x)$, so we can get an approximation for their product by multiplying the two polynomials:

$$\begin{aligned}
&e^{3x} \cdot \sin(2x) \\
&\quad \approx \left(1 + (3x) + \frac{(3x)^2}{2!} + \frac{(3x)^3}{3!} + \frac{(3x)^4}{4!} + \frac{(3x)^5}{5!}\right)\left((2x) - \frac{(2x)^3}{3!} + \frac{(2x)^5}{5!}\right) \\
&\quad \approx 2x + 6x^2 + \frac{23}{3}x^3 + 5x^4 - \frac{61}{60}x^5.
\end{aligned}$$

Again, you should try calculating this polynomial directly from the general rule, both to see that you get the same result, and to appreciate how much more tedious the general formula is to use in this case.

In the same way, we can also divide Taylor polynomials, raise them to powers, and chain them by composition. The exercises provide examples of some of these operations.

Differentiating Taylor Polynomials. Suppose we know a Taylor polynomial for some function f. If g is the derivative of f, we can immediately get a Taylor polynomial for g (of degree one less) by differentiating the polynomial we know for f. You should review the definition of Taylor polynomial to see why this is so. For instance, suppose $f(x) = 1/(1-x)$ and $g(x) = 1/(1-x)^2$. Verify that $f'(x) = g(x)$. It then follows that

$$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right) \approx \frac{d}{dx}(1 + x + x^2 + \cdots + x^n)$$
$$= 1 + 2x + 3x^2 + \cdots + nx^{n-1}.$$

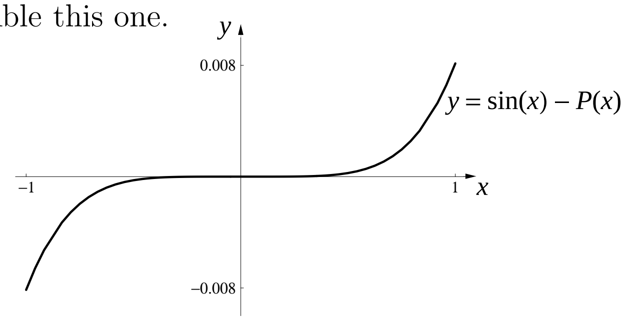
Integrating Taylor Polynomials. Again suppose we have functions $f(x)$ and $g(x)$ with $f'(x) = g(x)$, and suppose this time that we know a Taylor polynomial for g. We can then get a Taylor polynomial for f by antidifferentiating term by term. For instance, we find in chapter 11 that the derivative of $\arctan(x)$ is $1/(1+x^2)$, and we have seen above how to get a Taylor polynomial for $1/(1+x^2)$. Therefore we have

$$\arctan x = \int_0^x \frac{1}{1+t^2}dt \approx \int_0^x \left(1 - t^2 + t^4 - t^6 + \cdots \pm t^{2n}\right) dt$$
$$= t - \frac{1}{3}t^3 + \frac{1}{5}t^5 - \cdots \pm \frac{1}{2n+1}t^{2n+1}\Bigg|_0^x$$
$$= x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots \pm \frac{1}{2n+1}x^{2n+1}.$$

Goodness of fit

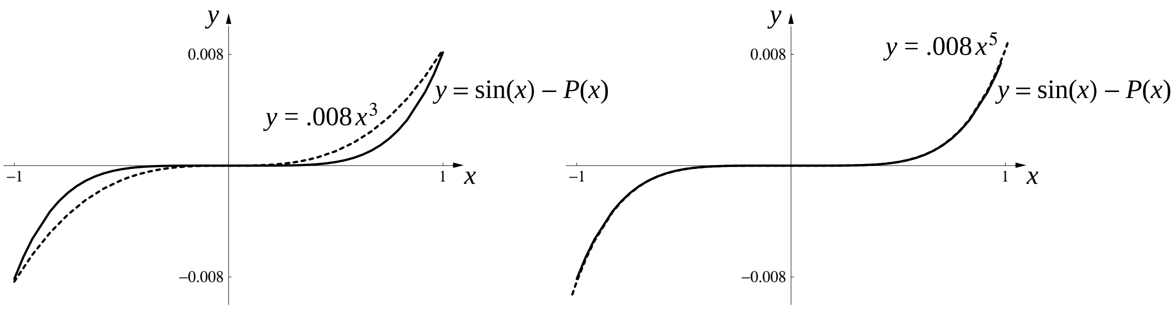
Graph the difference between a function and its Taylor polynomial

Let's turn to the question of *measuring* the fit between a function and one of its Taylor polynomials. The ideas here have a strong geometric flavor, so you should use a computer graphing utility to follow this discussion. Once again, consider the function $\sin(x)$ and its Taylor polynomial $P(x) = x - x^3/6$. According to the table in section 1, the difference $\sin(x) - P(x)$ got smaller as x got smaller. Stop now and graph the function $y = \sin(x) - P(x)$ near $x = 0$. This will show you exactly how $\sin(x) - P(x)$ depends on x. If you choose the interval $-1 \le x \le 1$ (and your graphing utility allows its vertical and horizontal scales to be set independently of each other), your graph should resemble this one.



The difference looks like a power of x

This graph looks very much like a cubic polynomial. If it really is a cubic, we can figure out its formula, because we know the value of $\sin(x) - P(x)$ is about .008 when $x = 1$. Therefore the cubic should be $y = .008\,x^3$ (because then $y = .008$ when $x = 1$). However, if you graph $y = .008\,x^3$ together with $y = \sin(x) - P(x)$, you should find a poor match (the left-hand figure, below.) Another possibility is that $\sin(x) - P(x)$ is more like a *fifth* degree polynomial. Plot $y = .008\,x^5$; it's so close that it "shares phosphor" with $\sin(x) - P(x)$ near $x = 0$.



Finding the multiplier

If $\sin(x) - P(X)$ were *exactly* a multiple of x^5, then $(\sin x - P(x))/x^5$ would be constant and would equal the value of the multiplier. What we actually find is this:

x	$\dfrac{\sin x - P(x)}{x^5}$
1.0	.0081377
0.5	.0082839
0.1	.0083313
0.05	.0083328
0.01	.0083333

suggesting $\displaystyle\lim_{x\to 0} \frac{\sin x - P(x)}{x^5} = .008333\ldots.$

How $P(x)$ fits $\sin(x)$

Thus, although the ratio is not constant, it appears to converge to a definite value—which we can take to be the value of the multipier:

$$\sin x - P(x) \approx .008333\,x^5 \quad \text{when} \quad x \approx 0.$$

We say that $\sin(x) - P(x)$ *has the same order of magnitude as* x^5 as $x \to 0$. So $\sin(x) - P(x)$ is about as small as x^5. Thus, if we know the size of x^5 we will be able to tell how close $\sin(x)$ and $P(x)$ are to each other.

Comparing two numbers

A rough way to measure how close two numbers are is to count the number of decimal places to which they agree. But there are pitfalls here; for instance, none of the decimals of 1.00001 and 0.99999 agree, even though the difference between the two numbers is only 0.00002. This suggests that a good way to compare two numbers is to look at their difference. Therefore, we say

$$A = B \text{ to } k \text{ decimal places} \quad \textit{means} \quad A - B = 0 \text{ to } k \text{ decimal places}$$

Now, a number equals 0 to k decimal places precisely when it *rounds off to* 0 (when we round it to k decimal places). Since X rounds to 0 to k decimal places if and only $|X| < .5 \times 10^{-k}$, we finally have a precise way to compare the size of two numbers:

$$\boxed{A = B \text{ to } k \text{ decimal places} \quad \textit{means} \quad |A - B| < .5 \times 10^{-k}.}$$

What the fit means computationally

Now we can say how close $P(x)$ is to $\sin(x)$. Since x is small, we can take this to mean $x = 0$ *to k decimal places*, or $|x| < .5 \times 10^{-k}$. But then,

$$|x^5 - 0| = |x - 0|^5 < (.5 \times 10^{-k})^5 < .5 \times 10^{-5k-1}$$

(since $.5^5 = .03125 < .5 \times 10^{-1}$). In other words, if $x = 0$ to k decimal places, then $x^5 = 0$ to $5k + 1$ places. Since $\sin(x) - P(x)$ has the same order of magnitude as x^5 as $x \to 0$, $\sin(x) = P(x)$ to $5k + 1$ places as well. In fact, because the multiplier in the relation

$$\sin x - P(x) \approx .008333\, x^5 \qquad (x \approx 0)$$

is .0083. . . , we gain two more decimal places of accuracy. (Do you see why?) Thus, finally, we see how reliable the polynomial $P(x) = x - x^3/6$ is for calculating values of $\sin(x)$:

> When $x = 0$ to k decimal places of accuracy, we can use $P(x)$ to calculate the first $5k+3$ decimal places of the value of $\sin(x)$.

Here are a few examples comparing $P(x)$ to the *exact* value of $\sin(x)$:

x	$P(x)$	$\sin(x)$
.0372	.0371914201920	.037194207856 . . .
.0086	.0085998939907	.008599893991 . . .
.0048	.0047999815680000	.0047999815680212 . . .

The underlined digits are guaranteed to be correct, based on the number of decimal places for which x agrees with 0. (Note that, according to our rule, $.0086 = 0$ to *one* decimal place, not two.)

Taylor's theorem

Order of magnitude

Taylor's theorem is the generalization of what we have just seen; it describes the goodness of fit between an arbitrary function and one of its Taylor polynomials. We'll state three versions of the theorem, gradually uncovering more information. To get started, we need a way to compare the order of magnitude of *any* two functions.

> We say that $\varphi(x)$ has **the same order of magnitude** as $q(x)$ as $x \to a$, and we write $\varphi(x) = O(q(x))$ as $x \to a$, if there is a constant C for which
>
> $$\lim_{x \to a} \frac{\varphi(x)}{q(x)} = C.$$

Now, when $\lim_{x\to a} \varphi(x)/q(x)$ is C, we have

$$\varphi(x) \approx Cq(x) \quad \text{when } x \approx a.$$

We'll frequently use this relation to express the idea that $\varphi(x)$ has the same order of magnitude as $q(x)$ as $x \to a$.

'Big oh' notation

The symbol O is an upper case "oh". When $\varphi(x) = O(q(x))$ as $x \to a$, we say $\varphi(x)$ *is 'big oh' of* $q(x)$ *as* x *approaches* a. Notice that the equal sign in $\varphi(x) = O(q(x))$ does *not* mean that $\varphi(x)$ and $O(q(x))$ are equal; $O(q(x))$ isn't even a function. Instead, the equal sign and the O together tell us that $\varphi(x)$ stands in a certain relation to $q(x)$.

> **Taylor's theorem, version 1**. If $f(x)$ has derivatives up to order n at $x = a$, then
>
> $$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R(x),$$
>
> where $R(x) = O((x-a)^{n+1})$ as $x \to a$. The term $R(x)$ is called the **remainder**.

Informal language

This version of Taylor's theorem focusses on the general shape of the remainder function. Sometimes we just say the remainder has "order $n+1$", using this short phrase as an abbreviation for "the order of magnitude of the function $(x-a)^{n+1}$". In the same way, we say that *a function and its* n*-th degree Taylor polynomial at* $x = a$ *agree to order* $n+1$ *as* $x \to a$.

Notice that, if $\varphi(x) = O(x^3)$ as $x \to 0$, then it is also true that $\varphi(x) = O(x^2)$ (as $x \to 0$). This implies that we should take $\varphi(x) = O(x^n)$ to mean "φ has *at least* order n" (instead of simply "φ has order n"). In the same way, it would be more accurate (but somewhat more cumbersome) to say that $\varphi = O(q)$ means "φ has *at least* the order of magnitude of q".

Decimal places of accuracy

As we saw in our example, we can translate the order of agreement between the function and the polynomial into information about the number of decimal places of accuracy in the polynomial approximation. In particular, if $x - a = 0$ to k decimal places, then $(x - a)^n = 0$ to nk places, at least. Thus, as the order of magnitude n of the remainder increases, the fit increases, too. (You have already seen this illustrated with the sine function and its various Taylor polynomials, in the figure on page 600.)

A formula for the remainder

While the first version of Taylor's theorem tells us that $R(x)$ looks like $(x - a)^{n+1}$ in some general way, the next gives us a concrete formula. At least, it *looks* concrete. Notice, however, that $R(x)$ is expressed in terms of a number c_x (which depends upon x), but the formula doesn't tell us *how* c_x depends upon x. Therefore, if you want to use the formula to *compute* the value of $R(x)$, you can't. The theorem says only that c_x exists; it doesn't say how to find its value. Nevertheless, this version provides useful information, as you will see.

Taylor's theorem, version 2. Suppose f has continuous derivatives up to order $n + 1$ for all x in some interval containing a. Then, for each x in that interval, there is a number c_x between a and x for which

$$R(x) = \frac{f^{(n+1)}(c_x)}{(n+1)!}\,(x - a)^{n+1}.$$

This is called **Lagrange's form of the remainder**.

Another formula for the remainder

We can use the Lagrange form as an aid to computation. To see how, return to the formula

$$R(x) \approx C(x - a)^{n+1} \quad (x \approx a)$$

that expresses $R(x) = O((x - a)^{n+1})$ as $x \to a$ (see page 608). The constant here is the limit

$$C = \lim_{x \to a} \frac{R(x)}{(x - a)^{n+1}}.$$

If we have a good estimate for the value of C, then $R(x) \approx C(x-a)^{n+1}$ gives us a good way to estimate $R(x)$. Of course, we could just evaluate the limit to determine C. In fact, that's what we did in the example; knowing $C \approx .008$ there gave us two more decimal places of accuracy in our polynomial approxmation to the sine function.

Determining C from f at $x = a$

But the Lagrange form of the remainder gives us another way to determine C:

$$\begin{aligned} C = \lim_{x \to a} \frac{R(x)}{(x-a)^{n+1}} &= \lim_{x \to a} \frac{f^{(n+1)}(c_x)}{(n+1)!} \\ &= \frac{f^{(n+1)}(\lim_{x \to a} c_x)}{(n+1)!} \\ &= \frac{f^{(n+1)}(a)}{(n+1)!}. \end{aligned}$$

In this argument, we are permitted to take the limit "inside" $f^{(n+1)}$ because $f^{(n+1)}$ is a continuous function. (That is one of the hypotheses of version 2.) Finally, since c_x lies between x and a, it follows that $c_x \to a$ as $x \to a$; in other words, $\lim_{x \to a} c_x = a$. Consequently, we get C *directly* from the function f itself, and we can therefore write

$$R(x) \approx \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} \qquad (x \approx a).$$

An error bound

The third version of Taylor's theorem uses the Lagrange form of the remainder in a similar way to get an *error bound* for the polynomial approximation based on the size of $f^{(n+1)}(x)$.

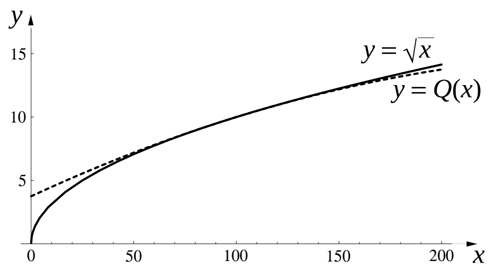
Taylor's theorem, version 3. Suppose that $|f^{(n+1)}(x)| \leq M$ for all x in some interval containing a. Then, for each x in that interval,

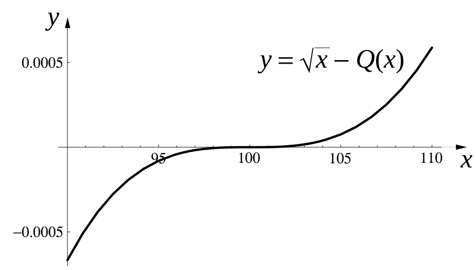
$$|R(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1}.$$

With this error bound, which is derived from knowledge of $f(x)$ near $x = a$, we can determine quite precisely how many decimal places of accuracy a Taylor polynomial approximation achieves. The following example illustrates the different versions of Taylor's theorem.

Example. Consider $\sqrt{x}$ near $x = 100$. The second degree Taylor polynomial for $\sqrt{x}$, centered at $x = 100$, is

$$Q(x) = 10 + \frac{(x-100)}{20} - \frac{(x-100)^2}{8000}.$$





Plot $y = Q(x)$ and $y = \sqrt{x}$ together; the result should look like the figure on the left, above. Then plot the remainder $y = \sqrt{x} - Q(x)$ near $x = 100$. This graph should suggest that $\sqrt{x} - Q(x) = O((x-100)^3)$ as $x \to 100$. In fact, this is what version 1 of Taylor's theorem asserts. Furthermore,

Version 1: the remainder is $O((x-100)^3)$

$$\lim_{x \to 100} \frac{\sqrt{x} - Q(x)}{(x-100)^3} \approx 6.25 \times 10^{-7};$$

check this yourself by constructing a table of values. Thus

$$\sqrt{x} - Q(x) \approx C(x-100)^3 \quad \text{where } C \approx 6.25 \times 10^{-7}.$$

We can use the Lagrange form of the remainder (in version 2 of Taylor's theorem) to get the value of C another way—directly from the third derivative of $\sqrt{x}$ at $x = 100$:

Version 2: determining C in terms of $\sqrt{x}$ at $x = 100$

$$C = \left.\frac{(x^{1/2})'''}{3!}\right|_{x=100} = \frac{\frac{1}{2} \cdot \frac{-1}{2} \cdot \frac{-3}{2} \cdot (100)^{-5/2}}{6} = \frac{1}{2^4 \cdot 10^5} = 6.25 \times 10^{-7}.$$

This is the *exact* value, confirming the estimate obtained above.

Let's see what the equation $\sqrt{x} - Q(x) \approx 6.25 \times 10^{-7}(x-100)^3$ tells us about the accuracy of the polynomial approximation. If we assume $|x-100| < .5 \times 10^{-k}$, then

Accuracy of the polynomial approximation

$$\begin{aligned} |\sqrt{x} - Q(x)| &< 6.25 \times 10^{-7} \times (.5 \times 10^{-k})^3 \\ &= .78125 \times 10^{-(3k+7)} \;<\; .5 \times 10^{-(3k+6)}. \end{aligned}$$

Thus

$$x = 100 \text{ to } k \text{ decimal places} \Longrightarrow \sqrt{x} = Q(x) \text{ to } 3k + 6 \text{ places.}$$

For example, if $x = 100.47$, then $k = 0$, so $Q(100.47) = \sqrt{100.47}$ to 6 decimal places. We find

$$Q(100.47) = \underline{10.02347}23875,$$

and the underlined digits should be correct. In fact,

$$\sqrt{100.47} = \underline{10.023472}4521\ldots.$$

Here is a second example. If $x = 102.98$, then we can take $k = -1$, so $Q(102.98) = \sqrt{102.98}$ to $3(-1) + 6 = 3$ decimal places. We find

$$Q(102.98) = \underline{10.147}88995, \qquad \sqrt{102.98} = \underline{10.147}906187\ldots.$$

Version 3: an explicit error bound

Let's see what additional light version 3 sheds on our investigation. Suppose we assume $x = 100$ to $k = 0$ decimal places. This means that x lies in the open interval $(99.5, 100.5)$. Version 3 requires that we have a bound on the size of the third derivative of $f(x) = \sqrt{x}$ over this interval. Now $f'''(x) = \frac{3}{8}x^{-5/2}$, and this is a decreasing function. (Check its graph; alternatively, note that its derivative is negative.) Its maximum value therefore occurs at the left endpoint of the (closed) interval $[99.5, 100.5]$:

$$|f'''(x)| \le f'''(99.5) = \tfrac{3}{8}\,(99.5)^{-5/2} < 3.8 \times 10^{-6}.$$

Therefore, from version 3 of Taylor's theorem,

$$|\sqrt{x} - Q(x)| < \frac{3.8 \times 10^{-6}}{3!}\,|x - 100|^3$$

Since $|x - 100| < .5$, $|x - 100|^3 < .125$, so

$$|\sqrt{x} - Q(x)| < \frac{3.8 \times 10^{-6} \times .125}{6} = .791667 \times 10^{-7} < .5 \times 10^{-6}.$$

This proves $\sqrt{x} = Q(x)$ to 6 decimal places—confirming what we found earlier.

Applications

Evaluating Functions. An obvious use of Taylor polynomials is to evaluate functions. In fact, whenever you ask a calculator or computer to evaluate a function—trigonometric, exponential, logarithmic—it is typically giving you the value of an appropriate polynomial (though not necessarily a Taylor polynomial).

Now you can do anything your calculator can!

Evaluating Integrals. The fundamental theorem of calculus gives us a quick way of evaluating a definite integral provided we can find an antiderivative for the function under the integral (cf. chapter 6.4). Unfortunately, many common functions, like e^{-x^2} or $(\sin x)/x$, don't have antiderivatives that can be expressed as finite algebraic combinations of the basic functions. Up until now, whenever we encountered such a function we had to rely on a Riemann sum to estimate the integral. But now we have Taylor polynomials, and it's easy to find an antiderivative for a polynomial! Thus, if we have an awkward definite integral to evaluate, it is reasonable to expect that we can estimate it by first getting a good polynomial approximation to the integrand, and then integrating this polynomial. As an example, consider the **error function**, erf(t), defined by

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2}\,dx\,.$$

The error function

This is perhaps the most important integral in statistics. It is the basis of the so-called "normal distribution" and is widely used to decide how good certain statistical estimates are. It is important to have a way of obtaining fast, accurate approximations for erf(t). We have already seen that

$$e^{-x^2} \approx 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots \pm \frac{x^{2n}}{n!}.$$

Now, if we antidifferentiate term by term:

$$\begin{aligned}
\int e^{-x^2}\,dx &\approx \int \left(1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots \pm \frac{x^{2n}}{n!}\right) dx \\
&= \int 1\,dx - \int x^2\,dx + \int \frac{x^4}{2!}\,dx - \int \frac{x^6}{3!}\,dx + \cdots \pm \int \frac{x^{2n}}{n!}\,dx \\
&= x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \cdots \pm \frac{x^{2n+1}}{(2n+1)\cdot n!}.
\end{aligned}$$

Thus,

$$\int_0^t e^{-x^2}\,dx \approx x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \cdots \pm \frac{x^{2n+1}}{(2n+1)\cdot n!}\Bigg|_0^t,$$

giving us, finally, an approximate formula for erf(t):

A formula for approximating the error function

$$\operatorname{erf}(t) \approx \frac{2}{\sqrt{\pi}}\left(t - \frac{t^3}{3} + \frac{t^5}{5\cdot 2!} - \frac{t^7}{\cdot 3!} + \cdots \pm \frac{t^{2n+1}}{(2n+1)\cdot n!}\right).$$

Thus if we needed to know, say, erf(1), we could quickly approximate it. For instance, letting $n = 6$, we have

$$\begin{aligned}\operatorname{erf}(1) &\approx \frac{2}{\sqrt{\pi}}\left(1 - \frac{1}{3} + \frac{1}{5\cdot 2!} - \frac{1}{7\cdot 3!} + \frac{1}{9\cdot 4!} - \frac{1}{11\cdot 5!} + \frac{1}{13\cdot 6!}\right)\\ &\approx \frac{2}{\sqrt{\pi}}\left(1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + \frac{1}{216} - \frac{1}{1320} + \frac{1}{9360}\right)\\ &\approx .746836\,\frac{2}{\sqrt{\pi}} \approx .842714,\end{aligned}$$

a value accurate to 4 decimals. If we had needed greater accuracy, we could simply have taken a larger value for n. For instance, if we take $n = 12$, we get the estimate .8427007929. . . , where all 10 decimals are accurate (i.e., they don't change as we take larger values n).

Evaluating Limits. Our final application of Taylor polynomials makes explicit use of the order of magnitude of the remainder. Consider the problem of evaluating a limit like

$$\lim_{x\to 0}\frac{1-\cos(x)}{x^2}.$$

Since both numerator and denominator approach 0 as $x \to 0$, it isn't clear what the quotient is doing. If we replace $\cos(x)$ by its third degree Taylor polynomial with remainder, though, we get

$$\cos(x) = 1 - \frac{1}{2!}x^2 + R(x),$$

and $R(x) = O(x^4)$ as $x \to 0$. Consequently, if $x \neq 0$ but $x \to 0$, then

$$\begin{aligned}\frac{1-\cos(x)}{x^2} &= \frac{1-\left(1-\frac{1}{2}x^2+R(x)\right)}{x^2}\\ &= \frac{\frac{1}{2}x^2 - R(x)}{x^2} = \frac{1}{2} - \frac{R(x)}{x^2}.\end{aligned}$$

Since $R(x) = O(x^4)$, we know that there is some constant C for which $R(x)/x^4 \to C$ as $x \to 0$. Therefore,

$$\begin{aligned}\lim_{x\to 0}\frac{1-\cos(x)}{x^2} &= \frac{1}{2} - \lim_{x\to 0}\frac{R(x)}{x^2} = \frac{1}{2} - \lim_{x\to 0}\frac{x^2\cdot R(x)}{x^4} \\ &= \frac{1}{2} - \lim_{x\to 0} x^2 \cdot \lim_{x\to 0}\frac{R(x)}{x^4} = \frac{1}{2} - 0\cdot C = \frac{1}{2}.\end{aligned}$$

Extending the 'big oh' notation

There is a way to shorten these calculations—and to make them more transparent—by extending the way we read the 'big oh' notation. Specifically, we will read $O(q(x))$ as "some (unspecified) function that is the same order of magnitude as $q(x)$".

Then, instead of writing $\cos(x) = 1 - \frac{1}{2}x^2 + R(x)$, and then noting $R(x) = O(x^4)$ as $x \to 0$, we'll just write

$$\cos(x) = 1 - \tfrac{1}{2}x^2 + O(x^4) \qquad (x \to 0).$$

In this spirit,

$$\begin{aligned}\frac{1-\cos(x)}{x^2} &= \frac{1 - \left(1 - \frac{1}{2}x^2 + O(x^4)\right)}{x^2} \\ &= \frac{\frac{1}{2}x^2 - O(x^4)}{x^2} = \tfrac{1}{2} + O(x^2) \qquad (x \to 0).\end{aligned}$$

We have used the fact that $\pm O(x^4)/x^2 = O(x^2)$. Finally, since $O(x^2) \to 0$ as $x \to 0$ (do you see why?), the limit of the last expression is just $1/2$ as $x \to 0$. Thus, once again we arrive at the result

$$\lim_{x\to 0}\frac{1-\cos(x)}{x^2} = \frac{1}{2}.$$

Exercises

1. Find a seventh degree Taylor polynomial centered at $x = 0$ for the indicated antiderivatives.

a) $\displaystyle\int \frac{\sin(x)}{x}\,dx.$

[Answer: $\displaystyle\int \frac{\sin(x)}{x}\,dx \approx x - \frac{x^3}{3\cdot 3!} + \frac{x^5}{5\cdot 5!} - \frac{x^7}{7\cdot 7!}$.]

b) $\displaystyle\int e^{x^2}\,dx.$

c) $\displaystyle\int \sin(x^2)\,dx.$

2. Plot the 7-th degree polynomial you found in part (a) above over the interval $[0, 5]$. Now plot the 9-th degree approximation on the same graph. When do the two polynomials begin to differ visibly?

3. Using the seventh degree Taylor approximation

$$E(t) \approx \int_0^t e^{-x^2}\,dx = t - \frac{t^3}{3} + \frac{t^5}{5 \cdot 2!} - \frac{t^7}{7 \cdot 3!},$$

calculate the values of $E(.3)$ and $E(-1)$. Give only the significant digits—that is, report only those decimals of your estimates that you think are fixed. (This means you will also need to calculate the ninth degree Taylor polynomial as well—do you see why?)

4. Calculate the values of $\sin(.4)$ and $\sin(\pi/12)$ using the seventh degree Taylor polynomial centered at $x = 0$

$$\sin(x) \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}.$$

Compare your answers with what a calculator gives you.

5. Find the third degree Taylor polynomial for $g(x) = x^3 - 3x$ at $x = 1$. Show that the Taylor polynomial is actually equal to $g(x)$—that is, the remainder is 0. What does this imply about the *fourth* degree Taylor polynomial for g at $x = 1$?

6. Find the seventh degree Taylor polynomial centered at $x = \pi$ for
(a) $\sin(x)$; (b) $\cos(x)$; (c) $\sin(3x)$.

7. In this problem you will compare computations using Taylor polynomials centered at $x = \pi$ with computations using Taylor polynomials centered at $x = 0$.

a) Calculate the value of $\sin(3)$ using a seventh degree Taylor polynomial centered at $x = 0$. How many decimal places of your estimate appear to be fixed?

b) Now calculate the value of $\sin(3)$ using a seventh degree Taylor polynomial centered at $x = \pi$. Now how many decimal places of your estimate appear to be fixed?

8. Write a program which evaluates a Taylor polynomial to print out $\sin(5°)$, $\sin(10°)$, $\sin(15°)$, ..., $\sin(40°)$, $\sin(45°)$ accurate to 7 decimals. (Remember to convert to radians before evaluating the polynomial!)

9. **Why 0! = 1**. When you were first introduced to exponential notation in expressions like 2^n, n was restricted to being a positive integer, and 2^n was defined to be the product of 2 multiplied by itself n times. Before long, though, you were working with expressions like 2^{-3} and $2^{1/4}$. These new expressions weren't defined in terms of the original definition. For instance, to calculate 2^{-3} you wouldn't try to multiply 2 by itself -3 times—that would be nonsense! Instead, 2^{-m} is defined by looking at the key *properties* of exponentiation for positive exponents, and extending the definition to other exponents in a way that preserves these properties. In this case, there are two such properties, one for adding exponents and one for multiplying them:

$$\begin{array}{lcl} \text{Property A:} & 2^m \cdot 2^n = 2^{m+n} & \text{for all positive } m \text{ and } n, \\ \text{Property M:} & (2^m)^n = 2^{mn} & \text{for all positive } m \text{ and } n. \end{array}$$

a) Show that to preserve property A we have to define $2^0 = 1$.

b) Show that we then have to define $2^{-3} = 1/2^3$ if we are to continue to preserve property A.

c) Show why $2^{1/4}$ must be $\sqrt[4]{2}$.

d) In the same way, you should convince yourself that a basic property of the factorial notation is that $(n+1)! = (n+1) \cdot n!$ for any positive integer n. Then show that to preserve this property, we have to define $0! = 1$.

e) Show that there is no way to define $(-1)!$ which preserves this property.

10. Use the general rule to derive the 5-th degree Taylor polynomial centered at $x = 0$ for the function

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

Use this approximation to estimate $\sqrt{1.1}$. How accurate is this?

11. Use the general rule to derive the formula for the n-th degree Taylor polynomial centered at $x = 0$ for the function

$$f(x) = (1 + x)^c \text{ where } c \text{ is a constant.}$$

12. Use the result of the preceding problem to get the 6-th degree Taylor polynomial centered at $x = 0$ for $1/\sqrt[3]{1 + x^2}$.
[Answer: $1 - \frac{1}{3}x^2 + \frac{2}{9}x^4 - \frac{14}{81}x^6$.]

13. Use the result of the preceding problem to approximate

$$\int_0^1 \frac{1}{\sqrt[3]{1 + x^2}}\, dx.$$

14. Calculate the first 7 decimals of erf(.3). Be sure to show why you think all 7 decimals are correct. What degree Taylor polynomial did you need to produce these 7 decimals?
[Answer: $\mathrm{erf}(.3) = .3286267\ldots$.]

15. a) Apply the general formula for calculating Taylor polynomials centered at $x = 0$ to the tangent function to get the 5-th degree approximation.
[Answer: $\tan(x) \approx x + x^3/3 + 2x^5/15$.]
b) Recall that $\tan(x) = \sin(x)/\cos(x)$. Multiply the 5-th degree Taylor polynomial for $\tan(x)$ from part a) by the 4-th degree Taylor polynomial for $\cos(x)$ and show that you get the fifth degree polynomial for $\sin(x)$ (discarding higher degree terms).

16. Show that the n-th degree Taylor polynomial centered at $x = 0$ for $1/(1 - x)$ is $1 + x + x^2 + \cdots + x^n$.

17. Note that

$$\int \frac{1}{1 - x}\, dx = -\ln(1 - x).$$

Use this observation, together with the result of the previous problem, to get the n-th degree Taylor polynomial centered at $x = 0$ for $\ln(1 - x)$.

18. a) Find a formula for the n-th degree Taylor polynomial centered at $x = 1$ for $\ln(x)$.

b) Compare your answer to part (a) with the Taylor polynomial centered at $x = 0$ for $\ln(1 - x)$ you found in the previous problem. Are your results consistent?

19. a) The first degree Taylor polynomial for e^x at $x = 0$ is $1 + x$. Plot the remainder $R_1(x) = e^x - (1 + x)$ over the interval $-.1 \leq x \leq .1$. How does this graph demonstrate that $R_1(x) = O(x^2)$ as $x \to 0$?

b) There is a constant C_2 for which $R_1(x) \approx C_2 x^2$ when $x \approx 0$. Why? Estimate the value of C_2.

20. This concerns the second degree Taylor polynomial for e^x at $x = 0$. Plot the remainder $R_2(x) = e^x - (1 + x + x^2/2)$ over the interval $-.1 \leq x \leq .1$. How does this graph demonstrate that $R_2(x) = O(x^3)$ as $x \to 0$?

a) There is a constant C_3 for which $R_2(x) \approx C_3 x^3$ when $x \approx 0$. Why? Estimate the value of C_3.

21. Let $R_3(x) = e^x - P_3(x)$, where $P_3(x)$ is the third degree Taylor polynomial for e^x at $x = 0$. Show $R_3(x) = O(x^4)$ as $x \to 0$.

22. At first glance, Taylor's theorem says that

$$\sin(x) = x - \frac{1}{6}x^3 + O(x^4) \quad \text{as } x \to 0.$$

However, graphs and calculations done in the text (pages 605–607) make it clear that

$$\sin(x) = x - \frac{1}{6}x^3 + O(x^5) \quad \text{as } x \to 0.$$

Explain this. Is Taylor's theorem wrong here?

23. Using a suitable formula (that is, a Taylor polynomial with remainder) for each of the functions involved, find the indicated limit.

a) $\displaystyle\lim_{x\to 0} \frac{\sin(x)}{x}$ [Answer: 1]

b) $\displaystyle\lim_{x\to 0} \frac{e^x - (1 + x)}{x^2}$ [Answer: 1/2]

c) $\displaystyle\lim_{x\to 1} \frac{\ln x}{x - 1}$

d) $\displaystyle\lim_{x\to 0} \frac{x - \sin(x)}{x^3}$ [Answer: 1/6]

e) $\lim\limits_{x\to 0} \dfrac{\sin(x^2)}{1-\cos(x)}$

24. Suppose $f(x) = 1 + x^2 + O(x^4)$ as $x \to 0$. Show that

$$(f(x))^2 = 1 + 2x^2 + O(x^4) \quad \text{as } x \to 0.$$

25. a) Using $\sin x = x - \frac{1}{6}x^3 + O(x^5)$ as $x \to 0$, show

$$(\sin x)^2 = x^2 - \tfrac{1}{3}x^4 + O(x^6) \quad \text{as } x \to 0.$$

b) Using $\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + O(x^5)$ as $x \to 0$, show

$$(\cos x)^2 = 1 - x^2 + \tfrac{1}{3}x^4 + O(x^5) \quad \text{as } x \to 0.$$

c) Using the previous parts, show $(\sin x)^2 + (\cos x)^2 = 1 + O(x^5)$ as $x \to 0$. (Of course, you already know $(\sin x)^2 + (\cos x)^2 = 1$ *exactly*.)

26. a) Apply the general formula for calculating Taylor polynomials to the tangent function to get the 5-th degree approximation.
b) Recall that $\tan(x) = \sin(x)/\cos(x)$, so $\tan(x)\cdot\cos(x) = \sin(x)$. Multiply the fifth degree Taylor polynomial for $\tan(x)$ from part a) by the fifth degree Taylor polynomial for $\cos(x)$ and show that you get the fifth degree Taylor polynomial for $\sin(x)$ plus $O(x^6)$—that is, plus terms of order 6 and higher.

27. a) Using the formulas

$$e^u = 1 + u + \tfrac{1}{2}u^2 + \tfrac{1}{6}u^3 + O(u^4) \qquad (u \to 0),$$
$$\sin x = x - \tfrac{1}{6}x^3 + O(x^5) \qquad (x \to 0),$$

show that $e^{\sin x} = 1 + x + \frac{1}{2}x^2 + O(x^4)$ as $x \to 0$.
b) Apply the general formula to obtain the third degree Taylor polynomial for $e^{\sin x}$ at $x = 0$, and compare your result with the formula in part (a).

28. Using $\dfrac{e^x - 1}{x} = 1 + \frac{1}{2}x + \frac{1}{6}x^2 + \frac{1}{24}x^3 + O(x^4)$ as $x \to 0$, show that

$$\frac{x}{e^x - 1} = 1 - \tfrac{1}{2}x + \tfrac{1}{12}x^2 + O(x^4) \qquad (x \to 0).$$

29. Show that the following are true as $x \to \infty$.

a) $x + 1/x = O(x)$.

b) $5x^7 - 12x^4 + 9 = O(x^7)$.

c) $\sqrt{1 + x^2} = O(x)$.

d) $\sqrt{1 + x^p} = O(x^{p/2})$.

30. a) Let $f(x) = \ln(x)$. Find the smallest bound M for which

$$|f^{(4)}(x)| \le M \quad \text{when } |x - 1| \le .5.$$

b) Let $P_3(x)$ be the degree 3 Taylor polynomial for $\ln(x)$ at $x = 1$, and let $R_3(x)$ be the remainder $R_3(x) = \ln(x) - P_3(x)$. Find a number K for which

$$|R(x)| \le K\,|x - 1|^4$$

for all x satisfying $|x - 1| \le .5$.

c) If you use $P_3(x)$ to approximate the value of $\ln(x)$ in the interval $.5 \le x \le 1.5$, how many digits of the approximation are correct?

d) Suppose we restrict the interval to $|x-1| \le .1$. Repeat parts (a) and (b), getting *smaller* values for M and K. Now how many digits of the polynomial approximation $P_3(x)$ to $\ln(x)$ are correct, if $.9 \le x \le 1.1$?

"Little oh" notation. Similar to the "big oh" notation is another, called the "little oh": if

$$\lim_{x \to a} \frac{\phi(x)}{q(x)} = 0,$$

then we write $\phi(x) = o(q(x))$ and say ϕ is '*little oh*' of q as $x \to a$.

31. Suppose $\phi(x) = O(x^6)$ as $x \to 0$. Show the following.

a) $\phi(x) = O(x^5)$ as $x \to 0$.

b) $\phi(x) = o(x^5)$ as $x \to 0$.

c) It is false that $\phi(x) = O(x^7)$ as $x \to 0$. (One way you can do this is to give an explicit example of a function $\phi(x)$ for which $\phi(x) = O(x^6)$ but for which you can show $\phi(x) = O(x^7)$ is false.)

d) It is false that $\phi(x) = o(x^6)$ as $x \to 0$.

32. Sketch the graph $y = x\ln(x)$ over the interval $0 < x \le 1$. Explain why your graph shows $\ln(x) = o(1/x)$ as $x \to 0$.

10.3 Taylor Series

In the previous section we have been talking about approximations to functions by their Taylor polynomials. Thus, for instance, we were able to write statements like

$$\sin(x) \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!},$$

where the approximation was a good one for values of x not too far from 0. On the other hand, when we looked at Taylor polynomials of higher and higher degree, the approximations were good for larger and larger values of x. We are thus tempted to write

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots,$$

indicating that the sine function is equal to this "infinite degree" polynomial. This infinite sum is called the **Taylor series** centered at $x = 0$ for $\sin(x)$. But what would we even mean by such an infinite sum? We will explore this question in detail in section 5, but you should already have some intuition about what it means, for it can be interpreted in exactly the same way we interpret a more familiar statement like

You have seen infinite sums before

$$\begin{aligned}\frac{1}{3} &= .33333\ldots \\ &= \frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \frac{3}{10000} + \frac{3}{100000} + \cdots.\end{aligned}$$

Every decimal number is a sum of fractions whose denominators are powers of 10; 1/3 is a number whose decimal expansion happens to need an infinite number of terms to be completely precise. Of course, when a practical matter arises (for example, typing a number like 1/3 or π into a computer) just the beginning of the sum is used—the "tail" is dropped. We might write 1/3 as 0.33, or as 0.33333, or however many terms we need to get the accuracy we want. Put another way, we are saying that 1/3 is the *limit* of the finite sums of the right hand side of the equation.

Infinite degree polynomials are to be viewed like infinite decimals

Our new formulas for Taylor series are meant to be used exactly the same way: when a computation is involved, take only the beginning of the sum, and drop the tail. Just where you cut off the tail depends on the input value x and on the level of accuracy needed. Look at what happens when we we approximate the value of $\cos(\pi/3)$ by evaluating Taylor polynomials of increasingly higher degree:

$$\begin{aligned}
1 &= 1.0000000\\
1-\frac{1}{2!}\left(\frac{\pi}{3}\right)^2 &\approx 0.4516887\\
1-\frac{1}{2!}\left(\frac{\pi}{3}\right)^2+\frac{1}{4!}\left(\frac{\pi}{3}\right)^4 &\approx 0.5017962\\
1-\frac{1}{2!}\left(\frac{\pi}{3}\right)^2+\frac{1}{4!}\left(\frac{\pi}{3}\right)^4-\frac{1}{6!}\left(\frac{\pi}{3}\right)^6 &\approx 0.4999646\\
1-\frac{1}{2!}\left(\frac{\pi}{3}\right)^2+\frac{1}{4!}\left(\frac{\pi}{3}\right)^4-\frac{1}{6!}\left(\frac{\pi}{3}\right)^6+\frac{1}{8!}\left(\frac{\pi}{3}\right)^8 &\approx 0.5000004\\
1-\frac{1}{2!}\left(\frac{\pi}{3}\right)^2+\frac{1}{4!}\left(\frac{\pi}{3}\right)^4-\frac{1}{6!}\left(\frac{\pi}{3}\right)^6+\frac{1}{8!}\left(\frac{\pi}{3}\right)^8-\frac{1}{10!}\left(\frac{\pi}{3}\right)^{10} &\approx 0.5000000
\end{aligned}$$

These sums were evaluated by setting $\pi = 3.141593$. As you can see, at the level of precision we are using, a sum that is six terms long gives the correct value. However, five, four, or even three terms may have been adequate for the needs at hand. The crucial fact is that these are all honest calculations using *only* the four operations of elementary arithmetic.

Note that if we had wanted to get the same 6 place accuracy for $\cos(x)$ for a larger value of x, we might need to go further out in the series. For instance $\cos(7\pi/3)$ is also equal to .5, but the tenth degree Taylor polynomial centered at $x = 0$ gives

$$1-\frac{1}{2!}\left(\frac{7\pi}{3}\right)^2+\frac{1}{4!}\left(\frac{7\pi}{3}\right)^4-\frac{1}{6!}\left(\frac{7\pi}{3}\right)^6+\frac{1}{8!}\left(\frac{7\pi}{3}\right)^8-\frac{1}{10!}\left(\frac{7\pi}{3}\right)^{10} = -37.7302,$$

which is not even close to .5 . In fact, to get $\cos(7\pi/3)$ to 6 decimals, we need to use the Taylor polynomial centered at $x = 0$ of degree 30, while to get $\cos(19\pi/3)$ (also equal to .5) to 6 decimals we need the Taylor polynomial centered at $x = 0$ of degree 66!

The key fact, though, is that, for any value of x, if we go out in the series far enough (where what constitutes "far enough" will depend on x), we can approximate $\cos(x)$ to any number of decimal places desired. For any x, the value of $\cos(x)$ is the limit of the finite sums of the Taylor series, just as $1/3$ is the limit of the finite sums of its infinite series representation.

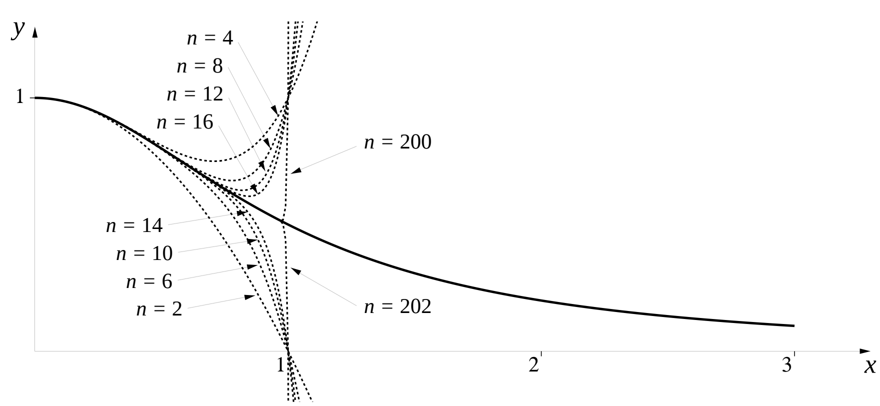
In general, given a function $f(x)$, its Taylor series centered at $x = 0$ will be

$$f(0)+f'(0)x+\frac{f^{(2)}(0)}{2!}x^2+\frac{f^{(3)}(0)}{3!}x^3+\frac{f^{(4)}(0)}{4!}x^4+\cdots=\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}x^k.$$

We have the following Taylor series centered at $x = 0$ for some common functions:

$f(x)$	Taylor series for $f(x)$
$\sin(x)$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$
$\cos(x)$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$
e^x	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$
$\ln(1-x)$	$-\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots\right)$
$\frac{1}{1-x}$	$1 + x + x^2 + x^3 + \cdots$
$\frac{1}{1+x^2}$	$1 - x^2 + x^4 - x^6 + \cdots$
$(1+x)^c$	$1 + cx + \frac{c(c-1)}{2!}x^2 + \frac{c(c-1)(c-2)}{3!}x^3 + \cdots$

While it is true that $\cos(x)$ and e^x equal their Taylor series, just as $\sin(x)$ did, we have to be more careful with the last four functions. To see why this is, let's graph $1/(1+x^2)$ and its Taylor polynomials $P_n(x) = 1 - x^2 + x^4 - x^6 + \cdots \pm x^n$ for $n = 2$, 4, 6, 8, 10, 12, 14, 16, 200, and 202. Since all the graphs are symmetric about the y-axis (why is this?), we draw only the graphs for positive x:



It appears that the graphs of the Taylor polynomials $P_n(x)$ approach the graph of $1/(1+x^2)$ very nicely *so long as* $x < 1$. If $x \geq 1$, though, it looks like there is no convergence, no matter how far out in the Taylor series we go. We can thus write

A Taylor series may not converge for all values of x

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots \qquad \text{for } |x| < 1,$$

where the restriction on x is essential if we want to use the = sign. We say that the interval $-1 < x < 1$ is the **interval of convergence** for the Taylor series centered at $x = 0$ for $1/(1+x^2)$. Some Taylor series, like those for $\sin(x)$ and e^x, converge for all values of x—their interval of convergence is $(-\infty, \infty)$. Other Taylor series, like those for $1/(1+x^2)$ and $\ln(1-x)$, have finite intervals of convergence.

Brook Taylor (1685–1731) was an English mathematician who developed the series that bears his name in his book *Methodus incrementorum* (1715). He did not worry about questions of convergence, but used the series freely to attack many kinds of problems, including differential equations.

Remark On the one hand it is perhaps not too surprising that a function should equal its Taylor series—after all, with more and more coefficients to fiddle with, we can control more and more of the behavior of the associated polynomials. On the other hand, we are saying that a function like $\sin(x)$ or e^x has its behavior for all values of x completely determined by the value of the function and all its derivatives at a single point, so perhaps it is surprising after all!

Exercises

1. a) Suppose you wanted to use the Taylor series centered at $x = 0$ to calculate $\sin(100)$. How large does n have to be before the term $(100)^n/n!$ is less than 1?

b) If we wanted to calculate $\sin(100)$ directly using this Taylor series, we would have to go very far out before we began to approach a limit at all closely. Can you use your knowledge of the way the circular functions behave to calculate $\sin(100)$ much more rapidly (but still using the Taylor series centered at $x = 0$)? Do it.

c) Show that we can calculate the sine of any number by using a Taylor series centered at $x = 0$ *either* for $\sin(x)$ *or* for $\cos(x)$ to a suitable value 0f x between 0 and $\pi/4$.

2. a) Suppose we wanted to calculate $\ln 5$ to 7 decimal places. An obvious place to start is with the Taylor series centered at $x = 0$ for $\ln(1 - x)$:

$$-\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots\right)$$

with $x = -4$. What happens when you do this, and why? Try a few more values for x and see if you can make a conjecture about the interval of convergence for this Taylor series.

[Answer: The Taylor series converges for $-1 \leq x < 1$.]

b) Explain how you could use the fact that $\ln(1/A) = -\ln A$ for any real number $A > 0$ to evaluate $\ln x$ for $x > 2$. Use this to compute $\ln 5$ to 7 decimals. How far out in the Taylor series did you have to go?

c) If you wanted to calculate $\ln 1.5$, you could use the Taylor series for $\ln(1-x)$ with either $x = -1/2$, which would lead directly to $\ln 1.5$, or you could use the series with $x = 1/3$, which would produce $\ln(2/3) = -\ln 1.5$. Which method is faster, and why?

3. We can improve the speed of our calculations of the logarithm function slightly by the following series of observations:

a) Find the Taylor series centered at $u = 0$ for $\ln(1 + u)$.

[Answer: $u - u^2/2 + u^3/3 - u^4/4 + u^5/5 + \cdots$]

b) Find the Taylor series centered at $u = 0$ for

$$\ln\left(\frac{1-u}{1+u}\right).$$

(Remember that $\ln(A/B) = \ln A - \ln B$.)

c) Show that *any* $x > 0$ can be written in the form $(1 - u)/(1 + u)$ for some suitable $-1 < u < 1$.

d) Use the preceding to evaluate $\ln 5$ to 7 decimal places. How far out in the Taylor series did you have to go?

4. a) Evaluate $\arctan(.5)$ to 7 decimal places.

b) Try to use the Taylor series centered at $x = 0$ to evaluate arctan(2) directly—what happens? Remembering what the arctangent function means geometrically, can you figure out a way around this difficulty?

5. a) **Calculating** π The Taylor series for the arctangent function,

$$\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots \pm \frac{1}{2n+1}x^{2n+1} + \cdots,$$

lies behind many of the methods for getting lots of decimals of π rapidly. For instance, since $\tan\left(\frac{\pi}{4}\right) = 1$, we have $\frac{\pi}{4} = \arctan 1$. Use this to get a series expansion for π. How far out in the series do you have to go to evaluate π to 3 decimal places?

b) The reason the preceding approximations converged so slowly was that we were substituting $x = 1$ into the series, so we didn't get any help from the x^n terms in making the successive corrections get small rapidly. We would like to be able to do something with values of x between 0 and 1. We can do this by using the addition formula for the tangent function:

$$\tan(\alpha + \beta) = \frac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta}.$$

Use this to show that

$$\frac{\pi}{4} = \arctan\left(\frac{1}{2}\right) + \arctan\left(\frac{1}{5}\right) + \arctan\left(\frac{1}{8}\right).$$

Now use the Taylor series for each of these three expressions to calculate π to 12 decimal places. How far out in the series do you have to go? Which series did you have to go the farthest out in before the 12th decimal stabilized? Why?

6. **Raising** e **to imaginary powers** One of the major mathematical developments of the last century was the extension of the ideas of calculus to **complex numbers**—i.e., numbers of the form $r + s\,i$, where r and s are real numbers, and i is a new symbol, defined by the property that $i \cdot i = -1$. Thus $i^3 = i^2\,i = -i$, $i^4 = i^2\,i^2 = (-1)(-1) = 1$, and so on. If we want to extend our standard functions to these new numbers, we proceed as we did in the previous section and look for the crucial *properties* of these functions

to see what they suggest. One of the key properties of e^x as we've now seen is that it possesses a Taylor series:

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

But this property only involves operations of ordinary arithmetic, and so makes perfectly good sense even if x is a complex number

a) Show that if s is any real number, we must define $e^{i\,s}$ to be $\cos(s)+i\,\sin(s)$ if we want to preserve this property.

b) Show that $e^{\pi\,i} = -1$.

c) Show that if $r + s\,i$ is any complex number, we must have

$$e^{r+s\,i} = e^r(\cos s + i\,\sin s)$$

if we want complex exponentials to preserve all the right properties.

d) Find a complex number $r + si$ such that $e^{r+s\,i} = -5$.

7. **Hyperbolic trigonometric functions** The hyperbolic trigonometric functions are defined by the formulas

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \qquad \sinh(x) = \frac{e^x - e^{-x}}{2}.$$

(The names of these functions are usually pronounced "cosh" and "cinch.") In this problem you will explore some of the reasons for the adjectives *hyperbolic* and *trigonometric*.

a) Modify the Taylor series centered at $x = 0$ for e^x to find a Taylor series for $\cosh(x)$. Compare your results to the Taylor series centered at $x = 0$ for $\cos(x)$.

b) Now find the Taylor series centered at $x = 0$ for $\sinh(x)$. Compare your results to the Taylor series centered at $x = 0$ for $\sin(x)$.

c) Parts (a) and (b) of this problem should begin to explain the *trigonometric* part of the story. What about the *hyperbolic* part? Recall that the familiar trigonometric functions are called *circular functions* because, for any t, the point $(\cos t, \sin t)$ is on the unit circle with equation $x^2 + y^2 = 1$ (cf. chapter 7.2). Show that the point $(\cosh t, \sinh t)$ lies on the hyperbola with equation $x^2 - y^2 = 1$.

8. Consider the Taylor series centered at $x = 0$ for $(1 + x)^c$.

a) What does the series give if you let $c = 2$? Is this reasonable?

b) What do you get if you set $c = 3$?

c) Show that if you set $c = n$, where n is a positive integer, the Taylor series will terminate. This yields a general formula—the **binomial theorem**—that was discovered by the 12th century Persian poet and mathematician, Omar Khayyam, and generalized by Newton to the form you have just obtained. Write out the first three and the last three terms of this formula.

d) Use an appropriate substitution for x and a suitable value for c to derive the Taylor series for $1/(1 - u)$. Does this agree with what we previously obtained?

e) Suppose we want to calculate $\sqrt{17}$. We might try letting $x = 16$ and $c = 1/2$ and using the Taylor series for $(1 + x)^c$. What happens when you try this?

f) We can still use the series to help us, though, if we are a little clever and write

$$\sqrt{17} = \sqrt{16+1} = \sqrt{16\left(1+\frac{1}{16}\right)} = \sqrt{16}\cdot\sqrt{1+\frac{1}{16}} = 4\cdot\sqrt{1+\frac{1}{16}}.$$

Now apply the series using $x = 1/16$ to evaluate $\sqrt{17}$ to 7 decimal place accuracy. How many terms does it take?

g) Use the same kind of trick to evaluate $\sqrt[3]{30}$.

Evaluating Taylor series rapidly Suppose we wanted to plot the Taylor polynomial of degree 11 associated with $\sin(x)$. For each value of x, then, we would have to evaluate

$$P_{11}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!}.$$

Since the length of time it takes the computer to evaluate an expression like this is roughly proportional to the number of multiplications and divisions involved (additions and subtractions, by comparison, take a negligible amount of time), let's see how many of these operations are needed to evaluate $P_{11}(x)$. To calculate x^{11} requires 10 multiplications, while 11! requires 9 (if we are clever and don't bother to multiply by 1 at the end!), so the evaluation of the term $x^{11}/11!$ will require a total of 20 operations (counting the final division). Similarly, evaluating $x^9/9!$ requires 16 operations, $x^7/7!$ requires 12, on down to $x^3/3!$, which requires 4. Thus the total number of multiplications and divisions needed is

$$4 + 8 + 12 + 16 + 20 = 60.$$

This is not too bad, although if we were doing this for many different values of x, which would be the case if we wanted to graph $P_{11}(x)$, this would begin to add up. Suppose, though, that we wanted to graph something like $P_{51}(x)$ or $P_{101}(x)$. By the same analysis, evaluating $P_{51}(x)$ for a single value of x would require

$$4 + 8 + 12 + 16 + 20 + 24 + \cdots + 96 + 100 = 1300$$

multiplications and divisions, while evaluation of $P_{101}(x)$ would require 5100 operations. Thus it would take roughly 20 times as long to evaluate $P_{51}(x)$ as it takes to evaluate $P_{11}(x)$, while $P_{101}(x)$ would take about 85 times as long.

9. Show that, in general, the number of multiplications and divisions needed to evaluate $P_n(x)$ is roughly $n^2/2$.

We can be clever, though. Note that $P_{11}(x)$ can be written as

$$x\left(1 - \frac{x^2}{2\cdot 3}\left(1 - \frac{x^2}{4\cdot 5}\left(1 - \frac{x^2}{6\cdot 7}\left(1 - \frac{x^2}{8\cdot 9}\left(1 - \frac{x^2}{10\cdot 11}\right)\right)\right)\right)\right).$$

10. How many multiplications and divisions are required to evaluate this expression?
[Answer: $3 + 4 + 4 + 4 + 4 + 1 = 20$.]

11. Thus this way of evaluating $P_{11}(x)$ is roughly three times as fast, a modest saving. How much faster is it if we use this method to evaluate $P_{51}(x)$?
[Answer: The old way takes roughly 13 times as long.]

12. Find a general formula for the number of multiplications and divisions needed to evaluate $P_n(x)$ using this way of grouping.

Finally, we can extend these ideas to reduce the number of operations even further, so that evaluating a polynomial of degree n requires only n multiplications, as follows. Suppose we start with a polynomial

$$p(x) = a_0 + a_1\, x + a_2\, x^2 + a_3\, x^3 + \cdots + a_{n-1}\, x^{n-1} + a_n\, x^n.$$

We can rewrite this as

$$a_0 + x\,(a_1 + x\,(a_2 + \ldots + x\,(a_{n-2} + x\,(a_{n-1} + a_n\, x))\ldots))\,.$$

You should check that with this representation it requires only n multiplications to evaluate $p(x)$ for a given x.

13. a) Write two computer programs to evaluate the 300th degree Taylor polynomial centered at $x = 0$ for e^x, with one of the programs being the obvious, standard way, and the second program being this method given above. Evaluate $e^1 = e$ using each program, and compare the length of time required.

b) Use these two programs to graph the 300th degree Taylor polynomial for e^x over the interval $[0, 2]$, and compare times.

10.4 Power Series and Differential Equations

So far, we have begun with functions we already know, in the sense of being able to calculate the value of the function and all its derivatives at at least one point. This in turn allowed us to write down the corresponding Taylor series. Often, though, we don't even have this much information about a function. In such cases it is frequently useful to assume that there is some infinite polynomial—called a **power series**—which represents the function, and then see if we can determine the coefficients of the polynomial.

This technique is especially useful in dealing with differential equations. To see why this is the case, think of the alternatives. If we can approximate the solution $y = y(x)$ to a certain differential equation to an acceptable degree of accuracy by, say, a 20-th degree polynomial, then the only storage space required is the insignificant space taken to keep track of the 21 coefficients. Whenever we want the value of the solution for a given value of x, we can then get a quick approximation by evaluating the polynomial at x. Other alternatives are much more costly in time or in space. We could use Euler's method to grind out the solution at x, but, as you've already discovered, this can be a slow and tedious process. Another option is to calculate lots of values and store them in a table in the computer's memory. This not only takes up a lot of memory space, but it also only gives values for a finite set of values of x, and is not much faster than evaluating a polynomial. Until 30 years ago, the table approach was the standard one—all scientists and mathematicians had a handbook of mathematical functions containing hundreds of pages of numbers giving the values of every function they might need.

To see how this can happen, let's first look at a familiar differential equation whose solutions we already know:

$$y' = y.$$

Of course, we know by now that the solutions are $y = ae^x$ for an arbitrary constant a (where $a = y(0)$). Suppose, though, that we didn't already know how to solve this differential equation. We might see if we can find a power series of the form

$$y = a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3 + \cdots + a_n\,x^n + \cdots$$

that solves the differential equation. Can we, in fact, determine values for the coefficients $a_0, a_1, a_2, \cdots, a_n, \cdots$ that will make $y' = y$?

Using the rules for differentiation, we have

$$y' = a_1 + 2a_2\,x + 3a_3\,x^2 + 4a_4\,x^3 + \cdots + na_n\,x^{n-1} + \cdots .$$

Two polynomials are equal if and only if the coefficients of corresponding powers of x are equal. Therefore, if $y' = y$, it would have to be true that

$$\begin{aligned} a_1 &= a_0 \\ 2a_2 &= a_1 \\ 3a_3 &= a_2 \\ &\vdots \\ na_n &= a_{n-1} \\ &\vdots \end{aligned}$$

Therefore the values of a_1, a_2, a_3, ... are not arbitrary; indeed, each is determined by the preceding one. Equations like these—which deal with a sequence of quantities and relate each term to those earlier in the sequence—are called **recursion relations**. These recursion relations permit us to express every a_n in terms of a_0:

$$\begin{aligned} a_1 &= a_0 \\ a_2 &= \frac{1}{2}\,a_1 & & &&= \frac{1}{2}\,a_0 \\ a_3 &= \frac{1}{3}\,a_2 &&= \frac{1}{3}\cdot\frac{1}{2}\,a_0 &&= \frac{1}{3!}\,a_0 \\ a_4 &= \frac{1}{4}\,a_3 &&= \frac{1}{4}\cdot\frac{1}{3!}\,a_0 &&= \frac{1}{4!}\,a_0 \\ &\vdots &&\vdots &&\vdots \\ a_n &= \frac{1}{n}\,a_{n-1} &&= \frac{1}{n}\cdot\frac{1}{(n-1)!}\,a_0 &&= \frac{1}{n!}\,a_0 \\ &\vdots &&\vdots &&\vdots \end{aligned}$$

Notice that a_0 remains "free": there is no equation that determines its value. Thus, without additional information, a_0 is *arbitrary*. The series for y now becomes

$$y = a_0 + a_0\,x + \frac{1}{2!}\,a_0\,x^2 + \frac{1}{3!}\,a_0\,x^3 + \cdots + \frac{1}{n!}\,a_0\,x^n + \cdots$$

or

$$y = a_0\left[1 + x + \frac{1}{2!}\,x^2 + \frac{1}{3!}\,x^3 + \cdots + \frac{1}{n!}\,x^2 + \cdots\right].$$

But the series in square brackets is just the Taylor series for e^x—we have derived the Taylor series from the differential equation alone, without using any of the other properties of the exponential function. Thus, we again find that the solutions of the differential equation $y' = y$ are

$$y = a_0 e^x,$$

where a_0 is an arbitrary constant. Notice that $y(0) = a_0$, so the value of a_0 will be determined if the initial value of y is specified.

Note In general, once we have derived a power series expression for a function, that power series will also be the Taylor series for that function. Although the two series are the same, the term Taylor series is typically reserved for those settings where we were able to evaluate the derivatives through some other means, as in the preceding section.

Bessel's Equation

For a new example, let's look at a differential equation that arises in an enormous variety of physical problems (wave motion, optics, the conduction of electricity and of heat and fluids, and the stability of columns, to name a few):

$$x^2 \cdot y'' + x \cdot y' + (x^2 - p^2) \cdot y = 0\,.$$

This is called the **Bessel equation of order *p***. Here p is a parameter specified in advance, so we will really have a different set of solutions for each value of p. To determine a solution completely, we will also need to specify the initial values of $y(0)$ and $y'(0)$. The solutions of the Bessel equation of order p are called **Bessel functions of order *p***, and the solution for a given value of p (together with particular initial conditions which needn't concern us here) is written $J_p(x)$. In general, there is no formula for a Bessel function in terms of simpler functions (although it turns out that a few special cases like $J_{1/2}(x), J_{3/2}(x), \ldots$ can be expressed relatively simply). To evaluate such a function we could use Euler's method, or we could try to find a power series solution.

Friedrich Wilhelm Bessel (1784–1846) was a German astronomer who studied the functions that now bear his name in his efforts to analyze the perturbations of planetary motions, particularly those of Saturn.

Consider the Bessel equation with $p = 0$. We are thus trying to solve the differential equation

$$x^2 \cdot y'' + x \cdot y' + x^2 \cdot y = 0.$$

By dividing by x, we can simplify this a bit to

$$x \cdot y'' + y' + x \cdot y = 0.$$

Let's look for a power series expansion

$$y = b_0 + b_1\,x + b_2\,x^2 + b_3\,x^3 + \cdots .$$

We have

$$y' = b_1 + 2b_2x + 3b_3x^2 + 4b_4x^3 + \cdots + (n+1)b_{n+1}x^n + \cdots$$

and

$$y'' = 2b_2 + 6b_3x + 12b_4x^2 + 20b_5x^3 + \cdots + (n+2)(n+1)b_{n+2}x^n + \cdots ,$$

We can now use these expressions to calculate the series for the combination that occurs in the differential equation:

$$\begin{array}{rcccccc}
xy'' = & & 2b_2\,x & + & 6b_3\,x^2 & + \cdots \\
y' = b_1 & + & 2b_2\,x & + & 3b_3\,x^2 & + \cdots \\
xy = & & b_0\,x & + & b_1\,x^2 & + \cdots \\
\hline
xy'' + y' + xy = b_1 & + & (4b_2 + b_0)x & + & (9b_3 + b_1)x^2 & + \cdots
\end{array}$$

In general, the coefficient of x^n in the combination will be

Finding the coefficient of x^n

$$(n+1)n\,b_{n+1} + (n+1)\,b_{n+1} + b_{n-1} = (n+1)^2\,b_{n+1} + b_{n-1}.$$

If the power series y is to be a solution to the original differential equation, the infinite series for $xy'' + y' + xy$ must equal 0. This in turn means that every coefficient of that series must be 0. We thus get

$$\begin{aligned}
b_1 &= 0, \\
4b_2 + b_0 &= 0, \\
9b_3 + b_1 &= 0, \\
&\vdots \\
n^2b_n + b_{n-2} &= 0. \\
&\vdots
\end{aligned}$$

If we now solve these recursively as before, we see first off that since $b_1 = 0$, it must also be true that

$$b_k = 0 \qquad \text{for every odd } k.$$

For the even coefficients we have

$$b_2 = -\frac{1}{2^2} b_0,$$
$$b_4 = -\frac{1}{4^2} b_2 = \frac{1}{2^2 4^2} b_0,$$
$$b_6 = -\frac{1}{6^2} b_4 = -\frac{1}{2^2 4^2 6^2} b_0,$$

and, in general,

$$b_{2n} = \pm \frac{1}{2^2 4^2 6^2 \cdots (2n)^2} b_0 = \pm \frac{1}{2^{2n}(n!)^2} b_0.$$

Thus any function y satisfying the Bessel equation of order 0 must be of the form

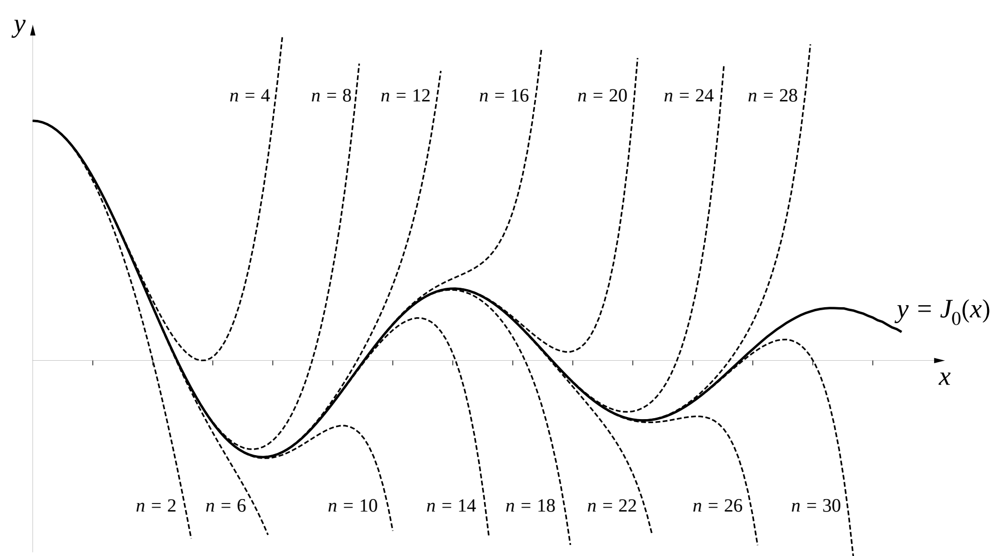
$$y = b_0 \left(1 - \frac{x^2}{2^2} + \frac{x^4}{2^4(2!)^2} - \frac{x^6}{2^6(3!)^2} + \cdots \right).$$

In particular, if we impose the initial condition $y(0) = 1$ (which requires that $b_0 = 1$), we get the 0-th order Bessel function $J_0(x)$:

$$J_0(x) = 1 - \frac{x^2}{4} + \frac{x^4}{64} - \frac{x^6}{2304} + \frac{x^8}{147456} + \cdots .$$

The graph of the Bessel function J_0

Here is the graph of $J_0(x)$ together with the polynomial approximations of degree 2, 4, 6, …, 30 over the interval $[0, 14]$:



The graph of J_0 is suggestive: it appears to be oscillatory, with decreasing amplitude. Both observations are correct: it can in fact be shown that J_0 has infinitely many zeroes, spaced roughly π units apart, and that $\lim_{x\to\infty} J_0(x) = 0$.

The S-I-R Model One More Time

In exactly the same way, we can find power series solutions when there are several interacting variables involved. Let's look at the example we've considered at a number of points in this text to see how this works. In the S-I-R model we basically wanted to solve the system of equations

The S-I-R model

$$\begin{aligned} S' &= -aSI, \\ I' &= aSI - bI, \\ R' &= bI, \end{aligned}$$

where a and b were parameters depending on the specific situation. Let's look for solutions of the form

$$\begin{aligned} S &= s_0 + s_1\,t + s_2\,t^2 + s_3\,t^3 + \cdots\,, \\ I &= i_0 + i_1\,t + i_2\,t^2 + i_3\,t^3 + \cdots\,, \\ R &= r_0 + r_1\,t + r_2\,t^2 + r_3\,t^3 + \cdots\,. \end{aligned}$$

If we put these series in the equation $S' = -a\,S\,I$, we get

$$\begin{aligned} s_1 + 2s_2t + 3s_3t^2 + \cdots &= -a(s_0 + s_1t + s_2t^2 + \cdots)(i_0 + i_1t + i_2t^2 + \cdots) \\ &= -a(s_0i_0 + (s_0i_1 + s_1i_0)t + (s_0i_2 + s_1i_1 + s_2i_0)t^2 + \cdots). \end{aligned}$$

As before, if the two sides of the differential equation are to be equal, the coefficients of corresponding powers of t must be equal:

Finding the coefficients of the power series for $S(t)$

$$\begin{aligned} s_1 &= -as_0i_0, \\ 2s_2 &= -a(s_0i_1 + s_1i_0), \\ 3s_3 &= -a(s_0i_2 + s_1i_1 + s_2i_0), \\ &\ \vdots \\ ns_n &= -a(s_0i_{n-1} + s_1i_{n-2} + \ldots + s_{n-2}i_1 + s_{n-1}i_0) \\ &\ \vdots \end{aligned}$$

Recursion again

While this looks messy, it has the crucial recursive feature—each s_k is expressed in terms of previous terms. That is, if we knew all the s and the i coefficients out through the coefficients of, say, t^6 in the series for S and I, we could immediately calculate s_7. We again have a *recursion relation.*

Finding the power series for $I(t)$

We could expand the equation $I' = aSI - bI$ in the same way, and get recursion relations for the coefficients i_k. In this model, though, there is a shortcut if we observe that since $S' = -aSI$, and since $I' = aSI - bI$, we have $I' = -S' - bI$. If we substitute the power series in this expression and equate coefficients, we get

$$ni_n = -ns_n - bi_{n-1},$$

which leads to

$$i_n = -s_n - \frac{b}{n} i_{n-1}$$

—so if we know s_n and i_{n-1}, we can calculate i_n.

We are now in a position to calculate the coefficients as far out as we like. For we will be given values for a and b when we are given the model. Moreover, since $s_0 = S(0) =$ the initial S-population, and $i_0 = I(0) =$ the initial I-population, we will also typically be given these values as well. But knowing s_0 and i_0, we can determine s_1 and then i_1. But then, knowing these values, we can determine s_2 and then i_2, and so on. Since the arithmetic is tedious, this is obviously a place for a computer. Here is a program that calculates the first 50 coefficients in the power series for $S(t)$ and $I(t)$:

Program: SIRSERIES

```
DIM S(0 to 50), I(0 to 50)
a = .00001
b = 1/14
S(0) = 45400
I(0) = 2100
FOR k = 1 TO 50
     Sum = 0
     FOR j = 0 TO k - 1
          Sum = Sum + S(j) * I(k - j - 1)
     NEXT j
     S(k) = -a * SUM/k
     I(k) = -S(k) - b * I(k - 1)/k
NEXT k
```

Comment: The opening command in this program introduces a new feature. It notifies the computer that the variables `S` and `I` are going to be **arrays**—strings of numbers—and that each array will consist of 51 elements. The element `S(k)` corresponds to what we have been calling s_k. The integer `k` is called the **index** of the term in the array. The indices in this program run from 0 to 50.

The effect of running this program is thus to create two 51-element arrays, `S` and `I`, containing the coefficients of the power series for S and I out to degree 50. If we just wanted to see these coefficients, we could have the computer list them. Here are the first 35 coefficients for S (read across the rows):

45400	−953.4	−172.3611	−17.982061	−.86969127
5.4479852e-2	1.5212707e-2	1.4463108e-3	4.3532884e-5	−7.9100481e-6
−1.4207959e-6	−1.0846994e-7	−6.512610e-10	9.304633e-10	1.256507e-10
7.443310e-12	−2.191966e-13	−9.787285e-14	−1.053428e-14	−4.382620e-16
4.230290e-17	9.548369e-18	8.321674e-19	1.760392e-20	−5.533369e-21
−8.770972e-22	−6.101928e-23	3.678170e-25	6.193375e-25	7.627253e-26
4.011923e-27	−1.986216e-28	−6.318305e-29	−6.271724e-30	−2.150100e-31

Thus the power series for S begins

$$45400 - 953.4t - 172.3611t^2 - 17.982061t^3 - .86969127t^4 + \cdots - 2.15010\times 10^{-31}t^{34} + \cdots$$

In the same fashion, we find that the power series for I begins

$$2100 + 803.4t + 143.66824t^2 + 14.561389t^3 + .60966648t^4 + \cdots + 2.021195\times 10^{-31}t^{34} + \cdots$$

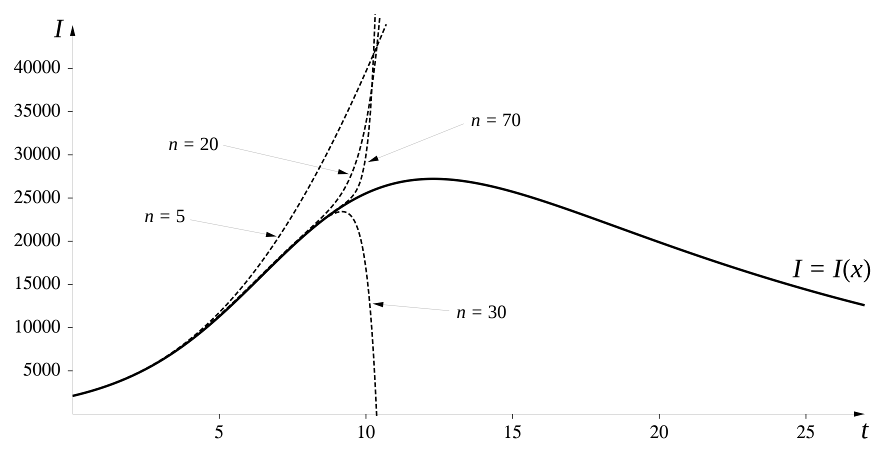
Extending SIRSERIES to graph S and I

If we now wanted to graph these polynomials over, say, $0 \le t \le 10$, we can do it by adding the following lines to SIRSERIES. We first define a couple of short subroutines `SUS` and `INF` to calculate the polynomial approximations for $S(t)$ and $I(t)$ using the coefficients we've derived in the first part of the program. (Note that these subroutines calculate polynomials in the straightforward, inefficient way. If you did the exercises in section 3 which developed techniques for evaluating polynomials rapidly, you might want to modify these subroutines to take advantage of the increased speed available.) Remember, too, that you will need to set up the graphics at the beginning of the program to be able to plot.

Extension to **SIRSERIES**

```
DEF SUS(x)
     Sum = S(0)
     FOR j = 1 TO 50
         Sum = Sum + S(j) * x^j
     NEXT j
     SUS = Sum
END DEF
DEF INF(x)
     Sum = I(0)
     FOR j = 1 TO 50
         Sum = Sum + I(j) * x^j
     NEXT j
     INF = Sum
END DEF
FOR x = 0 TO 10 STEP .01
     Plot the line from (x, SUS(x)) to (x + .01, SUS(x + .01))
     Plot the line from (x, INF(x)) to (x + .01, INF(x + .01))
NEXT x
```

Here is the graph of $I(t)$ over a 25-day period, together with the polynomial approximations of degree 5, 20, 30, and 70.

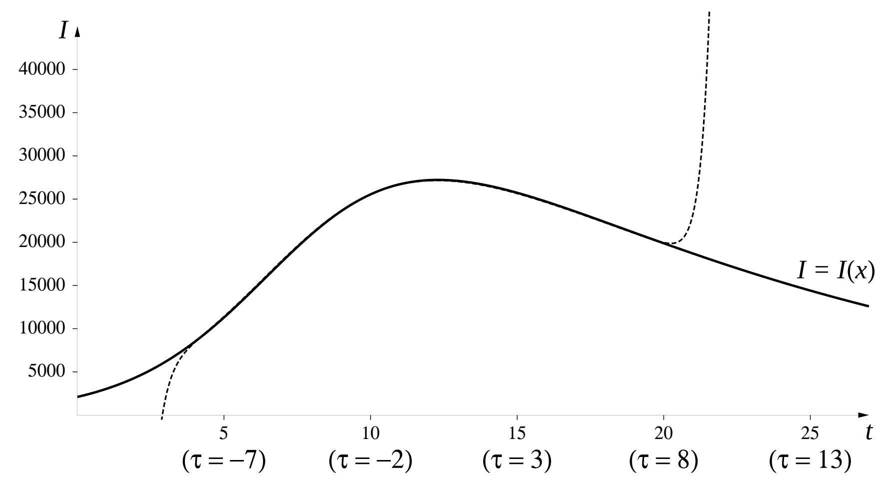


Note that these polynomials appear to converge to $I(t)$ only out to values of t around 10. If we needed polynomial approximations beyond that point, we could shift to a different point on the curve, find the values of I and S there by Euler's method, then repeat the above process. For instance, when

$t = 12$, we get by Euler's method that $S(12) = 7670$ and $I(12) = 27,136$. If we now shift our clock to measure time in terms of $\tau = t - 12$, we get the following polynomial of degree 30:

$$27136 + 143.0455\tau - 282.0180\tau^2 + 23.5594\tau^3 + .4548\tau^4 + \cdots + 1.2795 \times 10^{-25}\tau^{30}$$

Here is what the graph of this polynomial looks like when plotted with the graph of I. On the horizontal axis we list the t-coordinates with the corresponding τ-coordinates underneath.



The interval of convergence seems to be approximately $4 < t < 20$. Thus if we combine this polynomial with the 30-th degree polynomial from the previous graph, we would have very accurate approximations for I over the entire interval $[0, 20]$.

Exercises

1. Find power series solutions of the form

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots$$

for each of the following differential equations.

a) $y' = 2xy$.

b) $y' = 3x^2y$.

c) $y'' + xy = 0$.

d) $y'' + xy' + y = 0$.

2. a) Find power series solutions to the differential equation $y'' = -y$. Start with

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n + \cdots .$$

Notice that, in the recursion relations you obtain, the coefficients of the even terms are completely independent of the coefficients of the odd terms. This means you can get two separate power series, one with only even powers, with a_0 as arbitrary constant, and one with only odd powers, with a_1 as arbitrary constant.

b) The two power series you obtained in part a) are the Taylor series centered at $x = 0$ of two familiar functions; which ones? Verify that these functions do indeed satisfy the differential equation $y'' = -y$.

3. a) Find power series solutions to the differential equation $y'' = y$. As in the previous problem, the coefficients of the even terms depend only on a_0, and the coefficients of the odd terms depend only on a_1. Write down the two series, one with only even powers and a_0 as an arbitrary constant, and one with only odd powers, with a_1 as an arbitrary constant.

b) The two power series you obtained in part a) are the Taylor series centered at $x = 0$ of two *hyperbolic trigonometric functions* (see the exercises in section 3). Verify that these functions do indeed satisfy the differential equation $y'' = y$.

4. a) Find power series solutions to the differential equation $y' = xy$, starting with

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n + \cdots .$$

What recursion relations do you get? Is $a_1 = a_3 = a_5 = \cdots = 0$?

b) Verify that

$$y = e^{x^2/2}$$

satisfies the differential equation $y' = xy$. Find the Taylor series for this function and compare it with the series you obtained in a) using the recursion relations.

5. **The Bessel Equation.**

a) Take $p = 1$. The solution satisfying the initial condition $y' = 1/2$ when $x = 0$ is defined to be the first order Bessel function $J_1(x)$. (It will turn out that y has to be 0 when $x = 0$, so we don't have to specify the initial value of y; we have no choice in the matter.) Find the first five terms of the power series expansion for $J_1(x)$. What is the coefficient of x^{2n+1}?

b) Show by direct calculation from the series for J_0 and J_1 that

$$J_0' = -J_1.$$

c) To see, from another point of view, that $J_0' = -J_1$, take the equation

$$x \cdot J_0'' + J_0' + x \cdot J_0 = 0$$

and differentiate it. By doing some judicious cancelling and rearranging of terms, show that

$$x^2 \cdot (J_0')'' + x \cdot (J_0')' + (x^2 - 1)(J_0') = 0.$$

This demonstrates that J_0' is a solution of the Bessel equation with $p = 1$.

6. a) When we found the power series expansion for solutions to the 0-th order Bessel equation, we found that all the odd coefficients had to be 0. In particular, since b_1 is the value of y' when $x = 0$, we are saying that all solutions have to be flat at $x = 0$. This should bother you a bit. Why can't you have a solution, say, that satisfies $y = 1$ and $y' = 1$ when $x = 0$?

b) You might get more insight on what's happening by using Euler's method, starting just a little to the right of the origin and moving left. Use Euler's method to sketch solutions with the initial values

i.	$y = 2$	$y' = 1$	when	$x = 1$,
ii.	$y = 1.1$	$y' = 1$	when	$x = .1$,
iii.	$y = 1.01$	$y' = 1$	when	$x = .01$.

What seems to happen as you approach the y-axis?

7. **Legendre's differential equation**

$$(1 - x^2)y'' - 2xy' + \ell(\ell + 1)y = 0$$

arises in many physical problems—for example, in quantum mechanics, where its solutions are used to describe certain orbits of the electron in a hydrogen atom. In that context, the parameter ℓ is called the *angular momentum* of the electron; it must be either an integer or a "half-integer" (i.e., a number like 3/2). Quantum theory gets its name from the fact that numbers like the angular momentum of the electron in the hydrogen atom are "quantized", that is, they cannot have just any value, but must be a multiple of some "quantum"—in this case, the number 1/2.

a) Find power series solutions of Legendre's equation.

b) Quantization of angular momentum has an important consequence. Specifically, when ℓ is an integer it is possible for a series solution to stop—that is, to be a polynomial. For example, when $\ell = 1$ and $a_0 = 0$ the series solution is just $y = a_1 x$—all higher order coefficients turn out to be zero. Find polynomial solutions to Legendre's equation for $\ell = 0, 2, 3, 4$, and 5 (consider $a_0 = 0$ or $a_1 = 0$). These solutions are called, naturally enough, **Legendre polynomials**.

8. It turns out that the power series solutions to the S-I-R model have a finite interval of convergence. By plotting the power series solutions of different degrees against the solutions obtained by Euler's method, estimate the interval of convergence.

9. a) **Logistic Growth** Find the first five terms of the power series solution to the differential equation

$$y' = y(1 - y).$$

Note that this is just the logistic equation, where we have chosen our units of time and of quantities of the species being studied so that the carrying capacity is 1 and the intrinsic growth rate is 1.

b) Using the initial condition $y = .1$ when $x = 0$, plot this power series solution on the same graph as the solution obtained by Euler's method. How do they compare?

c) Do the same thing with initial conditions $y = 2$ when $x = 0$.

10.5 Convergence

We have written expressions such as

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots,$$

meaning that for any value of x the series on the right will converge to $\sin(x)$. There are a couple of issues here. The first is, what do we even mean when we say the series "converges", and how do we prove it converges to $\sin(x)$? If x is small, we can convince ourselves that the statement is true just by trying it. If x is large, though, say $x = 100^{100}$, it would be convenient to have a more general method for proving the stated convergence. Further, we have the example of the function $1/(1 + x^2)$ as a caution—it seemed to converge for small values of x ($|x| < 1$), but not for large values.

Let's first clarify what we mean by convergence. It is, essentially, the intuitive notion of "decimals stabilizing" that we have been using all along. To make explicit what we've been doing, let's write a "generic" series

Convergence means essentially that "decimals stabilize"

$$b_0 + b_1 + b_2 + \cdots = \sum_{m=0}^{\infty} b_m.$$

When we evaluated such a series, we looked at the **partial sums**

$$\begin{aligned}
S_1 &= b_0 + b_1 &&= \sum_{m=0}^{1} b_m, \\
S_2 &= b_0 + b_1 + b_2 &&= \sum_{m=0}^{2} b_m, \\
S_3 &= b_0 + b_1 + b_2 + b_3 &&= \sum_{m=0}^{3} b_m, \\
&\vdots && \vdots \\
S_n &= b_0 + b_1 + b_2 + \ldots + b_n &&= \sum_{m=0}^{n} b_m, \\
&\vdots && \vdots
\end{aligned}$$

Typically, when we calculated a number of these partial sums, we noticed that beyond a certain point they all seemed to agree on, say, the first 8 decimal places. If we kept on going, the partial sums would agree on the first 9 decimals, and, further on, on the first 10 decimals, etc. This is precisely what we mean by convergence:

The infinite series

$$b_0 + b_1 + b_2 + \cdots = \sum_{m=0}^{\infty} b_m$$

converges if, no matter how many decimal places are specified, it is always the case that the partial sums eventually agree to at least this many decimal places.

Put more formally, we say the series converges if, given any number D of decimal places, it is always possible to find an integer N_D such that if k and n are both greater than N_D, then S_k and S_n agree to at least D decimal places.

The number defined by these stabilizing decimals is called the **sum** of the series.

If a series does not converge, we say it **diverges**.

What it means for an infinite sum to converge

In other words, for me to prove to you that the Taylor series for $\sin(x)$ converges at $x = 100^{100}$, you would specify a certain number of decimal places, say 5000, and I would have to be able to prove to you that if you took partial sums with enough terms, they would all agree to at least 5000 decimals. Moreover, I would have to be able to show the same thing happens if you specify *any* number of decimal places you want agreement on.

How can this be done? It seems like an enormously daunting task to be able to do for any series. We'll tackle this challenge in stages. First we'll see what goes wrong with some series that don't converge—divergent series. Then we'll look at a particular convergent series—the **geometric series**—that's relatively easy to analyze. Finally, we will look at some more general rules that will guarantee convergence of series like those for the sine, cosine, and exponential functions.

Divergent Series

Suppose we have an infinite series

$$b_0 + b_1 + b_2 + \cdots = \sum_{m=0}^{\infty} b_m,$$

and consider two successive partial sums, say

$$S_{122} = b_0 + b_1 + b_2 + \ldots + b_{122} = \sum_{m=0}^{122} b_m$$

and

$$S_{123} = b_0 + b_1 + b_2 + \ldots + b_{122} + b_{123} = \sum_{m=0}^{123} b_m.$$

Note that these two sums are the same, except that the sum for S_{123} has one more term, b_{123}, added on. Now suppose that S_{122} and S_{123} agree to 19 decimal places. In section 2 we defined this to mean $|S_{123} - S_{122}| < .5 \times 10^{-19}$. But since $S_{123} - S_{122} = b_{123}$, this means that $|b_{123}| < .5 \times 10^{-19}$. To phrase this more generally,

> Two successive partial sums, S_n and S_{n+1}, agree out to k decimal places if and only if $|b_{n+1}| < .5 \times 10^{-k}$.

But since our definition of convergence required that we be able to fix any specified number of decimals provided we took partial sums lengthy enough, it must be true that *if the series converges*, the individual terms b_k must become arbitrarily small if we go out far enough. Intuitively, you can think of the partial sums S_k as being a series of approximations to some quantity. The term b_{k+1} can be thought of as the "correction" which is added to S_k to produce the next approximation S_{k+1}. Clearly, if the approximations are eventually becoming good ones, the corrections made should become smaller and smaller. We thus have the following necessary condition for convergence:

A necessary condition for convergence

> If the infinite series $b_0 + b_1 + b_2 + \cdots = \sum_{m=0}^{\infty} b_m$ converges, then $\lim_{k\to\infty} b_k = 0$.

Remark: It is important to recognize what this criterion does and does not say—it is a **necessary** condition for convergence (i.e., every convergent sequence has to satisfy the condition $\lim_{k\to\infty} b_k = 0$)—but it is not a **sufficient** condition for convergence (i.e., there are some divergent sequences

Necessary and *sufficient* mean different things that also have the property that $\lim_{k\to\infty} b_k = 0$). The criterion is usually used to detect some divergent series, and is more useful in the following form (which you should convince yourself is equivalent to the preceding):

> If $\lim\limits_{k\to\infty} b_k \neq 0$, (either because the limit doesn't exist at all, or it equals something besides 0), then the infinite series
>
> $$b_0 + b_1 + b_2 + \cdots = \sum_{m=0}^{\infty} b_m \qquad \text{diverges.}$$

Detecting divergent series

This criterion allows us to detect a number of divergent series right away. For instance, we saw earlier that the statement

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

appeared to be true only for $|x| < 1$. Using the remarks above, we can see why this series has to diverge for $|x| \geq 1$. If we write $1 - x^2 + x^4 - x^6 + \cdots$ as $b_0 + b_1 + b_2 + \cdots$, we see that $b_k = (-1)^k x^{2k}$. Clearly b_k does not go to 0 for $|x| \geq 1$—the successive "corrections" we make to each partial sum just become larger and larger, and the partial sums will alternate more and more wildly from a huge positive number to a huge negative number. Hence the series converges *at most* for $-1 < x < 1$. We will see in the next subsection how to prove that it really does converge for all x in this interval.

Using exactly the same kind of argument, we can show that the following series also diverge for $|x| > 1$:

$f(x)$	Taylor series for $f(x)$
$\ln(1-x)$	$-\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots\right)$
$\frac{1}{1-x}$	$1 + x + x^2 + x^3 + \cdots$
$(1+x)^c$	$1 + cx + \frac{c(c-1)}{2!}x^2 + \frac{c(c-1)(c-2)}{3!}x^3 + \cdots$
$\arctan x$	$x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \pm \frac{x^{2n+1}}{2n+1}$

The details are left to the exercises. While these common series all happen to diverge for $|x| > 1$, it is easy to find other series that diverge for $|x| > 2$ or $|x| > 17$ or whatever—see the exercises for some examples.

The Harmonic Series

We stated earlier in this section that simply knowing that the individual terms b_k go to 0 for large values of k does not guarantee that the series

$$b_0 + b_1 + b_2 + \cdots$$

will converge. Essentially what can happen is that the b_k go to 0 slowly enough that they can still accumulate large values. The classic example of such a series is the **harmonic series**:

An important counterexample

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \sum_{i=1}^{\infty} \frac{1}{i}.$$

It turns out that this series just keeps getting larger as you add more terms. It is eventually larger than 1000, or 1 million, or 100^{100} or .... This fact is established in the exercises. A suggestive argument, though, can be quickly given by observing that the harmonic series is just what you would get if you substituted $x = 1$ into the power series

$$x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots .$$

But this is just the Taylor series for $-\ln(1-x)$, and if we substitute $x = 1$ into this we get $-\ln 0$, which isn't defined. Also, $\lim_{x\to 0} -\ln x = +\infty$.

The Geometric Series

A series occurring frequently in a wide range of contexts is the **geometric series**

$$G(x) = 1 + x + x^2 + x^3 + x^4 + \cdots ,$$

This is also a sequence we can analyze completely and rigorously in terms of its convergence. It will turn out that we can then reduce the analysis of the convergence of several other sequences to the behavior of this one.

By the analysis we performed above, if $|x| \geq 1$ the individual terms of the series clearly don't go to 0, and the series therefore diverges. What about the case where $|x| < 1$?

To avoid divergence, $|x|$ must be less than 1

The starting point is the partial sums. A typical partial sum looks like:

$$S_n = 1 + x + x^2 + x^3 + \cdots + x^n\,.$$

A simple expression for the partial sum S_n

This is a finite number; we must find out what happens to it as n grows without bound. Since S_n is finite, we can calculate with it. In particular,

$$xS_n = x + x^2 + x^3 + \cdots + x^n + x^{n+1}.$$

Subtracting the second expression from the first, we get

$$S_n - xS_n = 1 - x^{n+1},$$

and thus (if $x \neq 1$)

$$S_n = \frac{1 - x^{n+1}}{1 - x}.$$

(What is the value of S_n if $x = 1$?)

This is a handy, compact form for the partial sum. Let us see what value it has for various values of x. For example, if $x = 1/2$, then

$$\begin{array}{rcccccccc} n: & 1 & 2 & 3 & 4 & 5 & 6 & \cdots \to & \infty \\ S_n: & 1 & \frac{3}{2} & \frac{7}{4} & \frac{15}{8} & \frac{31}{16} & \frac{63}{32} & \cdots \to & 2 \end{array}$$

Finding the limit of S_n as $n \to \infty$

It appears that as $n \to \infty$, $S_n \to 2$. Can we see this algebraically?

$$\begin{aligned} S_n &= \frac{1 - (1/2)^{n+1}}{1 - \frac{1}{2}} \\ &= \frac{1 - (1/2)^{n+1}}{1/2} \\ &= 2 \cdot (1 - (1/2)^{n+1}) = 2 - (1/2)^n. \end{aligned}$$

As $n \to \infty$, $(1/2)^n \to 0$, so the values of S_n become closer and closer to 2. The series converges, and its sum is 2.

Summing another geometric series

Similarly, when $x = -1/2$, the partial sums are

$$S_n = \frac{1 - (-1/2)^{n+1}}{3/2} = \frac{2}{3}\left(1 \pm \frac{1}{2^{n+1}}\right).$$

The presence of the $\pm$ sign does not alter the outcome: since $(1/2^{n+1}) \to 0$, the partial sums converge to $2/3$. Therefore, we can say the series converges and its sum is $2/3$.

In exactly the same way, though, for any x satisfying $|x| < 1$ we have

$$S_n = \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x}(1 - x^{n+1}),$$

and as $n \to \infty$, $x^{n+1} \to 0$. Therefore, $S_n \to 1/(1 - x)$. Thus the series converges, and its sum is $1/(1 - x)$.

To summarize, we have thus proved that

Convergence and divergence of the geometric series

> The geometric series
>
> $$G(x) = 1 + x + x^2 + x^3 + x^4 + \cdots$$
>
> converges for all x such that $|x| < 1$. In such cases the sum is
>
> $$\frac{1}{1 - x}.$$
>
> The series diverges for all other values of x.

As a final comment, note that the formula

$$S_n = \frac{1 - x^{n+1}}{1 - x}$$

is valid for all x except $x = 1$. Even though the partial sums aren't converging to any limit if $x > 1$, the formula can still be useful as a quick way for summing powers. Thus, for instance

$$1 + 3 + 9 + 27 + 81 + 243 = \frac{1 - 3^6}{1 - 3} = \frac{1 - 729}{-2} = \frac{-728}{-2} = 364,$$

and

$$1 - 5 + 25 - 125 + 625 - 3125 = \frac{1 - (-5)^6}{1 - (-5)} = \frac{1 - 15625}{6} = -264.$$

Alternating Series

A large class of common power series consists of the **alternating series**—series in which the terms are alternately positive and negative. The behavior of such series is particularly easy to analyze, as we shall see in this section. Here are some examples of alternating series we've already encountered :

Many common series are alternating, at least for some values of x

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots ,$$
$$[.15in] \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots ,$$
$$[.15in] \frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots , \qquad (\text{for } |x| < 1).$$

Other series may be alternating for some, but not all, values of x. For instance, here are two series that are alternating for negative values of x, but not for positive values:

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

$$\ln(1-x) = -(x + x^2 + x^3 + x^4 + \cdots) \qquad (\text{for } |x| < 1).$$

Convergence criterion for alternating series. Let us write a generic alternating series as

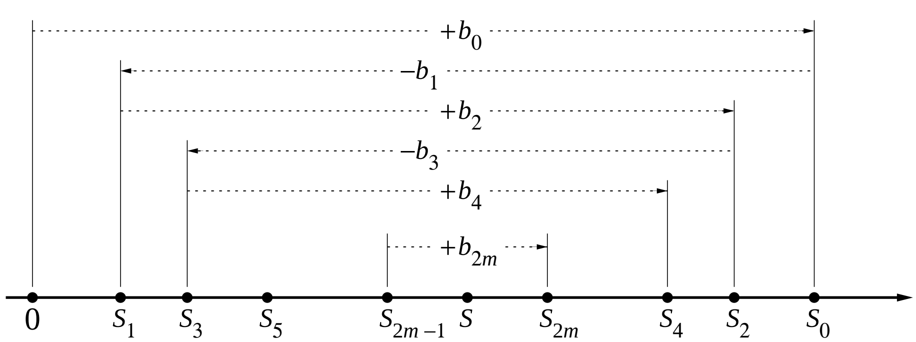
$$b_0 - b_1 + b_2 - b_3 + \cdots + (-1)^m b_m + \cdots ,$$

where the b_m are positive. It turns out that an alternating series converges if the terms b_m both consistently shrink in size and approach zero:

$$\boxed{\begin{array}{c} b_0 - b_1 + b_2 - b_3 + \cdots + (-1)^m b_m + \cdots \qquad \text{converges if} \\ 0 < b_{m+1} \le b_m \quad \text{for all } m \quad \text{and} \quad \lim_{m\to\infty} b_m = 0. \end{array}}$$

Alternating series are easy to test for convergence

It is this property that makes alternating series particularly easy to deal with. Recall that this is *not* a property of series in general, as we saw by the example of the harmonic series. The reason it is true for alternating series becomes clear if we view the behavior of the partial sums geometrically:



We mark the partial sums S_n on a number line. The first sum $S_0 = b_0$ lies to the right of the origin. To find S_1 we go to the left a distance b_1. Because $b_1 \leq b_0$, S_1 will lie between the origin and S_0. Next we go to the right a distance b_2, which brings us to S_2. Since $b_2 \leq b_1$, we will have $S_2 \leq S_0$. The next move is to the left a distance b_3, and we find $S_3 \geq S_1$. We continue going back and forth in this fashion, each step being less than or equal to the preceding one, since $b_{m+1} \leq b_m$. We thus get

$$0 \leq S_1 \leq S_3 \leq S_5 \leq \ldots \leq S_{2m-1} \leq \cdots \leq S_{2m} \leq \ldots \leq S_4 \leq S_2 \leq S_0.$$

The partial sums oscillate, with the exact sum trapped between consecutive partial sums

The partial sums oscillate back and forth, with all the odd sums on the left increasing and all the even sums on the right decreasing. Moreover, since $|S_n - S_{n-1}| = b_n$, and since $\lim_{n\to\infty} b_n = 0$, the difference between consecutive partial sums eventually becomes arbitrarily small—the oscillations take place within a smaller and smaller interval. Thus given any number of decimal places, we can always go far enough out in the series so that S_k and S_{k+1} agree to that many decimal places. But if n is any integer greater than k, then, since S_n lies between S_k and S_{k+1}, S_n will also agree to that many decimal places—those decimals will be fixed from k on out. The series therefore converges, as claimed—the sum is the unique number S that is greater than all the odd partial sums and less than all the even partial sums.

A simple estimate for the accuracy of the partial sums

For a convergent alternating series, we also have a particularly simple bound for the error when we approximate the sum S of the series by partial sums.

> If $$S_n = b_0 - b_1 + b_2 - \cdots \pm b_n,$$
> and if $\quad 0 < b_{m+1} \leq b_m \quad$ for all $m \quad$ and $\quad \lim\limits_{m\to\infty} b_m = 0,$
> (so the series converges), then
> $$|S - S_n| < b_{n+1}.$$
> In words, the error in approximating S by S_n is less than the next term in the series.

Proof: Suppose n is odd. Then we have, as above, that $S_n < S < S_{n+1}$. Therefore $0 < S - S_n < S_{n+1} - S_n = b_{n+1}$. If n is even, a similar argument shows $0 < S_n - S < S_n - S_{n+1} = b_{n+1}$. In either case, we have $|S - S_n| < b_{n+1}$, as claimed.

Note further that we also know whether S_n is too large or too small, depending on whether n is even or odd.

Estimating the error in approximating $\cos(.7)$

Example. Let's apply the error estimate for an alternating series to analyze the error if we approximate $\cos(.7)$ with a Taylor series with three terms:

$$\cos(.7) \approx 1 - \frac{1}{2!}(.7)^2 + \frac{1}{4!}(.7)^4 = 0.765004166\ldots.$$

Since the last term in this partial sum was an addition, this approximation is too big. To get an estimate of how far off it might be, we look at the next term in the series:

$$\frac{1}{6!}(.7)^6 = .0001634\ldots.$$

We thus know that the correct value for $\cos(.7)$ is somewhere in the interval

$$.76484 = .76500 - .00016 \leq \cos(.7) \leq .76501,$$

so we know that $\cos(.7)$ begins $.76\ldots$ and the third decimal is either a 4 or a 5. Moreover, we know that $\cos(.7) = .765$ rounded to 3 decimal places.

If we use the partial sum with four terms, we get

$$\cos(.7) \approx 1 - \frac{1}{2!}(.7)^2 + \frac{1}{4!}(.7)^4 - \frac{1}{6!}(.7)^6 = .764840765\ldots,$$

and the error would be less than

$$\frac{1}{8!}(.7)^8 = .0000014\ldots < .5 \times 10^{-5},$$

so we could now say that $\cos(.7) = .76484\ldots.$

How many terms are needed to obtain accuracy to 12 decimal places?

If we wanted to know in advance how far out in the series we would have to go to determine $\cos(.7)$ to, say, 12 decimals, we could do it by finding a value for n such that

$$b_n = \frac{1}{n!}(.7)^n \leq .5 \times 10^{-12}.$$

With a little trial and error, we see that $b_{12} \approx .3 \times 10^{-10}$, while $b_{14} < 10^{-13}$. Thus if we take the value of the 12th degree approximation for $\cos(.7)$, we can be assured that our value will be accurate to 12 places.

We have met this capability of getting an error estimate in a single step before, in version 3 of Taylor's theorem. It is in contrast to the approximations made in dealing with general series, where we typically had to look at the pattern of stabilizing digits in the succession of improving estimates to get a sense of how good our approximation was, and even then we had no guarantee.

It may be possible to convert a given problem to one involving alternating series

Computing e. Because of the fact that we can find sharp bounds for the accuracy of an approximation with alternating series, it is often desirable to convert a given problem to this form where we can. For instance, suppose we wanted a good value for e. The obvious thing to do would be to take the Taylor series for e^x and substitute $x = 1$. If we take the first 11 terms of this series we get the approximation

$$e = e^1 \approx 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{10!} = 2.718281801146\ldots,$$

but we have no way of knowing how many of these digits are correct.

Suppose instead, that we evaluate e^{-1}:

$$e^{-1} \approx 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{1}{10!} = .367879464286\ldots.$$

Since $1/(11!) = .000000025\ldots$, we know this approximation is accurate to at least 7 decimals. If we take its reciprocal we get

$$1/.3678794624286\ldots = 2.718281657666\ldots,$$

which will then be accurate to 6 decimals (in the exercises you will show why the accuracy drops by 1 decimal place), so we can say $e = 2.718281\ldots$.

The Radius of Convergence

We have seen examples of power series that converge for all x (like the Taylor series for $\sin x$) and others that converge only for certain x (like the series for $\arctan x$). How can we determine the convergence of an arbitrary power series of the form

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n + \cdots ?$$

We must suspect that this series *may not* converge for all values of x. For example, does the Taylor series

$$1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots$$

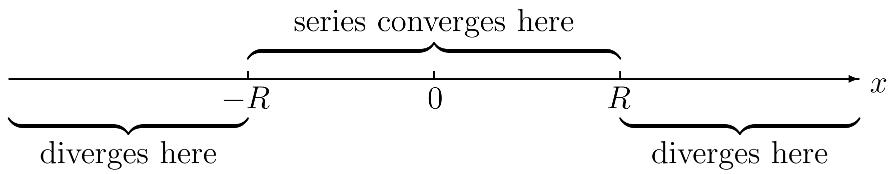
converge for all values of x, or only some? When it converges, does it converge to e^x or to something else? After all, this Taylor series is designed to look like e^x only near $x = 0$; it remains to be seen how well the function and its series match up far from $x = 0$.

The question of convergence has a definitive answer. It goes like this: if the power series

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots .$$

The answer to the convergence question

converges for a particular value of x, say $x = s$, then it automatically converges for any *smaller* value of x (meaning any x that is closer to the origin than s is; i.e, any x for which $|x| < |s|$). Likewise, if the series *diverges* for a particular value of x, then it also diverges for any value farther from the origin. In other words, the values of x where the series converges are not interspersed with the values where it diverges. On the contrary, within a certain distance R from the origin there is only convergence, while beyond that distance there is only divergence. The number R is called the **radius of convergence** of the series, and the range where it converges is called its **interval of convergence**.



The radius of convergence of a power series

An obvious example of the radius of convergence is given by the geometric series

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

Radius of convergence of the geometric series

We know that this converges for $|x| < 1$ and diverges for $|x| > 1$. Thus the radius of convergence is $R = 1$ in this case.

It is possible for a power series to converge for all x; if that happens, we take R to be ∞. At the other extreme, the series may converge only for $x = 0$. (When $x = 0$ the series collapses to its constant term a_0, so it certainly converges *at least* when $x = 0$.) If the series converges only for $x = 0$, then we take R to be 0.

At $x = R$ the series may diverge or converge; different things happen for different series. The same is true when $x = -R$. The radius of convergence tells us *where* the switch from convergence to divergence happens. It does not tell us *what* happens at the place where the switch occurs. If we know that the series converges for $x = \pm R$, then we say that $[-R, R]$ is the interval of convergence. If the series converges when $x = R$ but *not* when $x = -R$, then the interval of convergence is $(-R, R]$, and so on.

The Ratio Test

There are several ways to determine the radius of convergence of a power series. One of the simplest and most useful is by means of the **ratio test**. Because the power series to which we apply this test need not include *consecutive* powers of x (think of the Taylor series for $\cos x$ or $\sin x$) we'll write a "generic" series as

$$b_0 + b_1 + b_2 + \cdots = \sum_{m=0}^{\infty} b_m$$

Here are three examples of the use of this notation.

1. The Taylor series for e^x is $\sum_{m=0}^{\infty} b_m$, where

$$b_0 = 1, \quad b_1 = x, \quad b_2 = \frac{x^2}{2!}, \cdots, b_m = \frac{x^m}{m!}.$$

2. The Taylor series for $\cos x$ is $\sum_{m=0}^{\infty} b_m$, where

$$b_0 = 1, \quad b_1 = \frac{-x^2}{2!}, \quad b_2 = \frac{x^4}{4!}, \cdots, b_m = (-1)^m \frac{x^{2m}}{(2m)!}.$$

3. We can even describe the series

$$17 + x + x^2 + x^4 + x^6 + x^8 + \cdots = 17 + x + \sum_{m=2}^{\infty} x^{2m-2}$$

in our generic notation, in spite of the presence of the first two terms "$17 + x$" which don't fit the pattern of later ones. We have $b_0 = 17$, $b_1 = x$, and then $b_m = x^{2m-2}$ for $m = 2, 3, 4, \ldots$.

Convergence is determined by the "tail" of the series

The question of convergence for a power series is unaffected by the "beginning" of the series; only the pattern in the "tail" matters. (Of course the *value* of the power series is affected by all of its terms.) So we can modify our generic notation to fit the circumstances at hand. No harm is done if we don't begin with b_0.

Using this notation we can state the ratio test (but we give no proof).

Ratio Test: the series $b_0 + b_1 + b_2 + b_3 + \cdots + b_n + \cdots$ converges if $\displaystyle\lim_{m\to\infty} \frac{|b_{m+1}|}{|b_m|} < 1$.

The ratio test for the geometric series . . .

Let's see what the ratio test says about the geometric series:

$$1 + x + x^2 + x^3 + \cdots .$$

We have $b_m = x^m$, so the ratio we must consider is

$$\frac{|b_{m+1}|}{|b_m|} = \frac{|x^{m+1}|}{|x^m|} = \frac{|x|^{m+1}}{|x|^m} = |x|.$$

(Be sure you see why $|x^m| = |x|^m$.) Obviously, this ratio has the same value for all m, so the limit

$$\lim_{m\to\infty} |x| = |x|$$

exists and is less than 1 precisely when $|x| < 1$. Thus the geometric series converges for $|x| < 1$—which we already know is true. This means that the radius of convergence of the geometric series is $R = 1$.

. . . and for e^x

Look next at the Taylor series for e^x:

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{m=0}^{\infty} \frac{x^m}{m!}.$$

For negative x this is an alternating series, so by the criterion for convergence of alternating series we know it converges for all $x < 0$. The radius of convergence should then be ∞. We will use the ratio test to show that in fact this series converges for *all* x.

In this case

$$b_m = \frac{x^m}{m!},$$

so the relevant ratio is

$$\frac{|b_{m+1}|}{|b_m|} = \left|\frac{x^{m+1}}{(m+1)!}\right| \cdot \left|\frac{m!}{x^m}\right| = \frac{|x^{m+1}|}{|x^m|} \cdot \frac{m!}{(m+1)!} = |x| \cdot \frac{1}{m+1} = \frac{|x|}{m+1}.$$

Unlike the example with the geometric series, the value of this ratio depends on m. For any particular x, as m gets larger and larger the numerator stays the same and the denominator grows, so this ratio gets smaller and smaller. In other words,

$$\lim_{m\to\infty} \frac{|b_{m+1}|}{|b_m|} = \lim_{m\to\infty} \frac{|x|}{m+1} = 0.$$

Since this limit is less than 1 for any value of x, the series converges for *all* x, and thus the radius of convergence of the Taylor series for e^x is $R = \infty$, as we expected.

One of the uses of the theory developed so far is that it gives us a new way of specifying functions. For example, consider the power series

$$\sum_{m=0}^{\infty} (-1)^m \frac{2^m}{m^2+1}\, x^m = 1 - x + \frac{4}{5}x^2 - \frac{8}{10}x^3 + \frac{16}{17}x^4 + \cdots .$$

In this case $b_m = (-1)^m \dfrac{2^m}{m^2+1} x^m$, so to find the radius of convergence, we compute the ratio

$$\frac{|b_{m+1}|}{|b_m|} = \frac{2^{m+1}|x|^{m+1}}{(m+1)^2+1} \cdot \frac{m^2+1}{2^m|x|^m} = 2|x| \frac{m^2+1}{m^2+2m+2}.$$

To figure out what happens to this ratio as m grows large, it is helpful to rewrite the factor involving the m's as

Finding the limit of $|b_{m+1}|/|b_m|$ may require some algebra

$$\frac{m^2 \cdot (1 + 1/m^2)}{m^2 \cdot (1 + 2/m + 2/m^2)} = \frac{1 + 1/m^2}{1 + 2/m + 2/m^2}.$$

Now we can see that

$$\lim_{m\to\infty} \frac{|b_{m+1}|}{|b_m|} = 2|x| \cdot \frac{1}{1} = 2|x|.$$

The limit value is less than 1 precisely when $2|x| < 1$, or, equivalently, $|x| < 1/2$, so the radius of convergence of this series is $R = 1/2$. It follows that for $|x| < 1/2$, we have a new function $f(x)$ defined by the power series:

$$f(x) = \sum_{m=0}^{\infty} (-1)^m \frac{2^m}{m^2+1}\, x^m.$$

We can also discuss the radius of convergence of a power series

$$a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_m(x-a)^m + \cdots$$

centered at a location $x = a$ other than the origin. The radius of convergence of a power series of this form can be found by the **ratio test** in exactly the same way it was when $a = 0$.

Example. Let's apply the ratio test to the Taylor series centered at $a = 1$ for $\ln(x)$:

$$\ln(x) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m}(x-1)^m.$$

We can start our series with b_1, so we can take $b_m = \dfrac{(-1)^{m-1}}{m}(x-1)^m$. Then the ratio we must consider is

$$\frac{|b_{m+1}|}{|b_m|} = \frac{|x-1|^{m+1}}{m+1} \cdot \frac{m}{|x-1|^m} = |x-1| \cdot \frac{m}{m+1} = |x-1| \cdot \frac{1}{1+1/m}.$$

Then

$$\lim_{m\to\infty} \frac{|b_{m+1}|}{|b_m|} = |x-1| \cdot 1 = |x-1|.$$

From this we conclude that this series converges for $|x-1| < 1$. This inequality is equivalent to

$$-1 < x - 1 < 1,$$

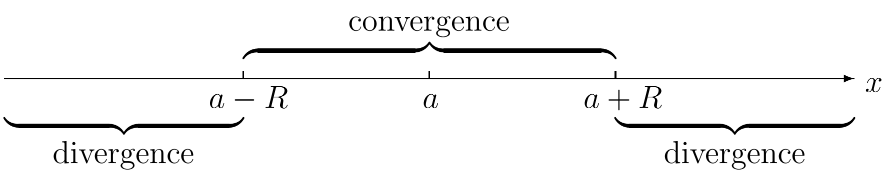
which is an interval of "radius" 1 about $a = 1$, so the radius of convergence is $R = 1$ in this case. We may also write the interval of convergence for this power series as

$$0 < x < 2.$$

More generally, using the ratio test we find that a power series centered at a converges in an interval of "radius" R (and width $2R$) around the point $x = a$ on the x-axis. Ignoring what happens at the endpoints, we say the **interval of convergence** is

$$a - R < x < a + R.$$

Here is a picture of what this looks like:



The convergence of a power series centered at a

Exercises

1. Find a formula for the sum of each of the following power series by performing suitable operations on the geometric series and the formula for *its* sum.

a) $1 - x^3 + x^6 - x^9 + \cdots$.

b) $x^2 + x^6 + x^{10} + x^{14} + \cdots$.

c) $1 - 2x + 3x^2 - 4x^3 + \cdots$.

d) $x + 2x^2 + 3x^3 + 4x^4 + \cdots$.

e) $x + \dfrac{x^2}{2} + \dfrac{x^3}{3} + \dfrac{x^4}{4} + \cdots$.

2. Determine the value of each of the following infinite sums. (Each os these sums is a geometric or related series evaluated at a particular value of x.)

a) $\dfrac{1}{4} + \dfrac{1}{16} + \dfrac{1}{64} + \dfrac{1}{256} + \cdots$.

b) .02020202 ....

c) $-\dfrac{5}{2} + \dfrac{5}{4} - \dfrac{5}{8} + \cdots$.

d) $\dfrac{1}{1} - \dfrac{2}{2} + \dfrac{3}{4} - \dfrac{4}{8} + \dfrac{5}{16} - \dfrac{6}{32} + \cdots$.

e) $\dfrac{1}{1 \cdot 10} + \dfrac{1}{2 \cdot 10^2} + \dfrac{1}{3 \cdot 10^3} + \dfrac{1}{4 \cdot 10^4} + \cdots$.

3. **The Multiplier Effect**. Economists know that the effect on a local economy of tourist spending is greater than the amount actually spent by the tourists. The *multiplier effect* quantifies this enlarged effect. In this problem you will see that calculating the multiplier effect involves summing a geometric series.

Suppose that, over the course of a year, tourists spend a total of A dollars in a small resort town. By the end of the year, the townspeople are therefore A dollars richer. Some of this money leaves the town—for example, to pay state and federal taxes or to pay off debts owed to "big city" banks. Some of it stays in town but gets put away as savings. Finally, a certain fraction of the original amount is spent in town, by the townspeople themselves. Suppose

3/5-ths is spent this way. The tourists and the townspeople *together* are therefore responsible for spending

$$S = A + \frac{3}{5}A \quad \text{dollars}$$

in the town that year. The second amount—$\frac{3}{5}A$ dollars—is *recirculated* money.

Since one dollar looks much like another, the recirculated money should be handled the same way as the original tourist dollars: some will leave the town, some will be saved, and the remaining 3/5-ths will get recirculated a *second* time. The twice-recirculated amount is

$$\frac{3}{5} \times \frac{3}{5}A \quad \text{dollars,}$$

and we must revise the calculation of the total amount spent in the town to

$$S = A + \frac{3}{5}A + \left(\frac{3}{5}\right)^2 A \quad \text{dollars.}$$

But the twice-recirculated dollars look like all the others , so 3/5-ths of them will get recirculated a *third* time. Revising the total dollars spent yet again, we get

$$S = A + \frac{3}{5}A + \left(\frac{3}{5}\right)^2 A + \left(\frac{3}{5}\right)^3 A \quad \text{dollars.}$$

This process never ends: no matter how many times a sum of money has been recirculated, 3/5-ths of it is recirculated once more. The total amount spend in the town is thus given by a *series.*

a) Write the series giving the total amount of money spent in the town and calculate its sum.

b) Your answer in a) is a certain multiple of A—what is the multiplier?

c) Suppose the recirculation rate is r instead of 3/5. Write the series giving the total amount spent and calculate its sum. What is the multiplier now?

d) Suppose the recirculation rate is 1/5; what is the multiplier in this case?

e) Does a lower recirculation rate produce a smaller multiplier effect?

4. Which of the following alternating series converge, which diverge? Why?

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

b) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{1}{\sqrt{3n+2}}$

c) $\displaystyle\sum_{n=2}^{\infty}(-1)^n\frac{n}{\ln n}$

d) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{n}{5n-4}$

e) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{\arctan n}{n}$

f) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{(1.0001)^n}{n^{10}+1}$

g) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{1}{n^{1/n}}$

h) $\displaystyle\sum_{n=2}^{\infty}(-1)^n\frac{1}{\ln n}$

i) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{n!}{n^n}$

j) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{n!}{1\cdot 3\cdot 5\cdots(2n-1)}$

5. For each of the sums in the preceding problem that converges, use the alternating series criterion to determine how far out you have to go before the sum is determined to 6 decimal places. Give the sum for each of these series to this many places.

6. Find a value for n so that the nth degree Taylor series for e^x gives at least 10 place accuracy for all x in the interval $[-3, 0]$.

7. We defined the harmonic series as the infinite sum

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots=\sum_{i=1}^{\infty}\frac{1}{i}.$$

a) Use a calculator to find the partial sums

$$S_n = 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots\frac{1}{n}$$

for $n = 1, 2, 3, \ldots, 12$.

b) Use the following program to find the value of S_n for $n = 100$. Modify the program to find the values of S_n for $n = 500$, 1000, and 5000.

Program: HARMONIC

```
n = 100
sum = 0
FOR i = 1 TO n
     sum = sum + 1/i
NEXT i
PRINT  n, sum
```

c) Group the terms in the harmonic series as indicated by the parentheses:

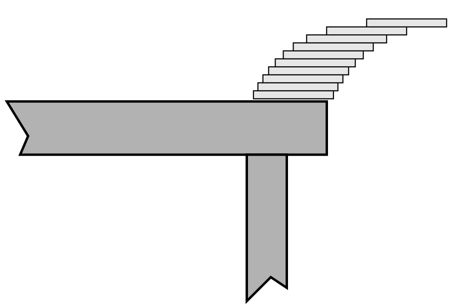
$$1+\left(\frac{1}{2}\right)+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)+ \\ +\left(\frac{1}{9}+\cdots+\frac{1}{16}\right)+\left(\frac{1}{17}+\cdots+\frac{1}{32}\right)+\cdots.$$

Explain why each parenthetical grouping totals at least 1/2.

d) Following the pattern in part (c), if you add up the terms of the harmonic series forming S_n for $n = 2^k$, you can arrange the terms as $1 + k$ such groupings. Use this fact and the result of c) to explain why S_n exceeds $1 + k \cdot \frac{1}{2}$.

e) Use part (d) to explain why the harmonic series *diverges*.

f) You might try this problem if you've studied physics—enough to know how to locate the center of mass of a system. Suppose you had n cards and wanted to stack them on the edge of a table with the top of the pile leaning out over the edge. How far out could you get the pile to reach if you were careful? Let's choose our units so the length of each card is 1. Clearly if $n = 1$, the farthest reach you could would be $\frac{1}{2}$. If $n = 2$, you could clearly place the top card to extend half a unit beyond the bottom card. For the system to be stable, the center of mass of the two cards must be to the left of the edge of the table. Show that for this to happen, the bottom card can't extend more than 1/4 unit beyond the edge. Thus with $n = 2$, the maximum extension of the pile is $\frac{1}{2} + \frac{1}{4} = \frac{3}{4}$. The picture at the right shows 10 cards stacked carefully.



Prove that if you have n cards, the stack can be built to extend a distance of

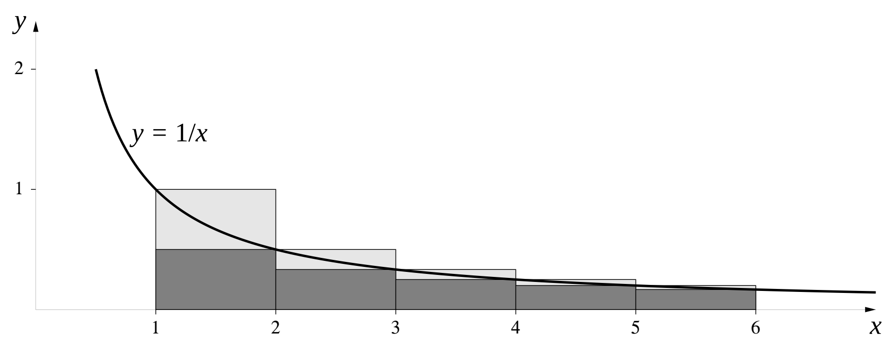
$$\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\frac{1}{8}+\cdots+\frac{1}{2n}=\frac{1}{2}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{n}\right).$$

In the light of what we have just proved about the harmonic series, this shows that if you had enough cards, you could have the top card extending 100 units beyond the edge of the table!

8. **An estimate for the partial sums of the harmonic series**. You may notice in part (d) of the preceding problem that the sum $S_n = 1+1/2+1/3+\cdots+1/n$ grows in proportion to the exponent k of $n = 2^k$; i.e., the sum grows like the logarithm of n. We can make this more precise by comparing the value of S_n to the value of the integral

$$\int_1^n \frac{1}{x}\,dx = \ln(n).$$

a) Let's look at the case $n = 6$ to see what's going on. Consider the following picture:



Show that the lightly shaded region plus the dark region has area equal to S_5, which can be rewritten as $S_6 - \frac{1}{6}$. Show that the dark region alone has area $S_6 - 1$. Hence prove that

$$S_6 - 1 < \int_1^6 \frac{1}{x}\,dx < S_6 - \frac{1}{6},$$

and conclude that

$$\frac{1}{6} < S_6 - \ln(6) < 1.$$

b) Show more generally that $\displaystyle \frac{1}{n} < S_n - \ln(n) < 1.$

c) Use part (b) to get upper and lower bounds for the value of S_{10000}.
[Answer: $9.21044 < S_{10000} < 10.21035$.]

d) Use the result of part (b) to get an estimate for how many cards you would need in part (f) of the preceding problem to make the top of the pile extend 100 units beyond the edge of the table.
[Answer: It would take approximately 10^{87} cards—a number which is on the same order of magnitude as the number of atoms in the universe!]

Remarkably, partial sums of the harmonic series exceed $\ln(n)$ in a very regular way. It turns out that

$$\lim_{n\to\infty} \{S_n - \ln(n)\} = \gamma,$$

where $\gamma = .5772\ldots$ is called **Euler's constant.** (You have seen another constant named for Euler, the base $e = 2.7183\ldots$.) Although one can find the decimal expansion of γ to any desired degree of accuracy, no one knows whether γ is a rational number or not.

9. Show that the power series for $\arctan x$ and $(1+x)^c$ diverge for $|x| > 1$. Do the series converge or diverge when $|x| = 1$?

10. Find the radius of convergence of each of the following power series.

a) $1 + 2x + 3x^2 + 4x^3 + \cdots$.

b) $x + 2x^2 + 3x^3 + 4x^4 + \cdots$.

c) $1 + \frac{1}{1^2}x + \frac{1}{2^2}x^2 + \frac{1}{3^2}x^3 + \frac{1}{4^2}x^4 + \cdots$. [Answer: $R = 1$]

d) $x^3 + x^6 + x^9 + x^{12} + \cdots$.

e) $1 + (x+1) + (x+1)^2 + (x+1)^3 + \cdots$.

f) $17 + \frac{1}{3}x + \frac{1}{3^2}x^2 + \frac{1}{3^3}x^3 + \frac{1}{3^4}x^4 + \cdots$.

11. Write out the first five terms of each of the following power series, and determine the radius of convergence of each.

a) $\displaystyle\sum_{n=0}^{\infty} nx^n$. [Answer: $R = 1$]

b) $\displaystyle\sum_{n=0}^{\infty} \frac{n^2}{2^n}x^n$. [Answer: $R = 2$]

c) $\displaystyle\sum_{n=0}^{\infty} (n+5)^2\, x^n$. [Answer: $R = 1$]

d) $\displaystyle\sum_{n=0}^{\infty} \frac{99}{n^n} x^n.$ [Answer: $R = \infty$]

e) $\displaystyle\sum_{n=0}^{\infty} n!\, x^n.$ [Answer: $R = 0$]

12. Find the radius of convergence of the Taylor series for $\sin x$ and for $\cos x$. For which values of x can these series therefore represent the sine and cosine functions?

13. Find the radius of convergence of the Taylor series for $f(x) = 1/(1+x^2)$ at $x = 0$. (See the table of Taylor series in section 3.) What is the radius of convergence of this series? For which values of x can this series therefore represent the function f? Do these x values constitute the *entire* domain of definition of f?

14. In the text we used the alternating series for e^x, $x < 0$, to approximate e^{-1} accurate to 7 decimal places. The claim was made that in taking the reciprocal to obtain an estimate for e, the accuracy drops by one decimal place. In this problem you will see why this is true.

a) Consider first the more general situation where two functions are reciprocals, $g(x) = 1/f(x)$. Express $g'(x)$ in terms of $f(x)$ and $f'(x)$.

b) Use your answer in part a) to find an expression for the *relative error* in g, $\Delta g/g(x) \approx g'(x)\Delta x/g(x)$, in terms of $f(x)$ and $f'(x)$. How does this compare to the relative error in f?

c) Apply your results in part b) to the functions e^x and e^{-x} at $x = 1$. Since e is about 7 times as large as $1/e$, explain why the error in the estimate for e should be about 7 times as large as the error in the estimate for $1/e$.

10.6 Approximation Over Intervals

A powerful result in mathematical analysis is the **Stone-Weierstrass Theorem**, which states that given any continuous function $f(x)$ and any interval $[a, b]$, there exist polynomials that fit f over this interval to any level of accuracy we care to specify. In many cases, we can find such a polynomial simply by taking a Taylor polynomial of high enough degree. There are several ways in which this is not a completely satisfactory response, however. First, some functions (like the absolute value function) have corners or other places where they aren't differentiable, so we can't even build a Taylor series at such points. Second, we have seen several functions (like $1/(1 + x^2)$) that have a finite interval of convergence, so Taylor polynomials may not be good fits no matter how high a degree we try. Third, even for well-behaved functions like $\sin(x)$ or e^x, we may have to take a very high degree Taylor polynomial to get the same overall fit that a much lower degree polynomial could achieve.

In this section we will develop the general machinery for finding polynomial approximations to functions over given intervals. In chapter 12.4 we will see how this same approach can be adapted to approximating periodic functions by **trigonometric polynomials**.

Approximation by polynomials

Two possible criteria for best fit

Example. Let's return to the problem introduced at the beginning of this chapter: find the second degree polynomial which best fits the function $\sin(x)$ over the interval $[0, \pi]$. Just as we did with the Taylor polynomials, though, before we can start we need to agree on our criterion for the best fit. Here are two obvious candidates for such a criterion:

1. The second degree polynomial $Q(x)$ is the best fit to $\sin(x)$ over the interval $[0, \pi]$ if the *maximum* separation between $Q(x)$ and $\sin(x)$ is smaller than the maximum separation between $\sin(x)$ and any other second degree polynomial:

$$\max_{0 \le x \le \pi} |\sin(x) - Q(x)| \qquad \text{is the smallest possible.}$$

2. The second degree polynomial $Q(x)$ is the best fit to $\sin(x)$ over the interval $[0, \pi]$ if the *average* separation between $Q(x)$ and $\sin(x)$ is smaller

than the average separation between $\sin(x)$ and any other second degree polynomial:

$$\frac{1}{\pi}\int_0^{\pi} |\sin(x) - Q(x)|\,dx \qquad \text{is the smallest possible.}$$

Unfortunately, even though their clear meanings make these two criteria very attractive, they turn out to be largely unusable—if we try to apply either criterion to a specific problem, including our current example, we are led into a maze of tedious and unwieldy calculations.

Why we don't use either criterion

Instead, we use a criterion that, while slightly less obvious than either of the two we've already articulated, still clearly measures some sort of "best fit" and has the added virtue of behaving well mathematically. We accomplish this by modifying criterion 2 slightly. It turns out that the major difficulty with this criterion is the presence of absolute values. If, instead of considering the average separation between $Q(x)$ and $\sin(x)$, we consider the average of the *square* of the separation between $Q(x)$ and $\sin(x)$, we get a criterion we can work with. (Compare this with the discussion of the best-fitting line in the exercises for chapter 9.3.) Since this is a definition we will be using for the rest of this section, we frame it in terms of arbitrary functions g and h, and an arbitrary interval $[a, b]$:

> Given two functions g and h defined over an interval $[a, b]$, we define the **mean square separation** between g and h over this interval to be
>
> $$\frac{1}{(b-a)}\int_a^b (g(x) - h(x))^2\,dx.$$

Note: In this setting the word **mean** is synonymous with what we have called "average". It turns out that there is often more than one way to define the term "average"—the concepts of median and mode are two other natural ways of capturing "averageness", for instance—so we use the more technical term to avoid ambiguity.

We can now rephrase our original problem as: find the second degree polynomial $Q(x)$ whose mean squared separation from $\sin(x)$ over the interval $[0, \pi]$ is as small as possible. In mathematical terms, we want to find

The criterion we shall use

coefficients a_0, a_1, and a_2 which such that the integral

$$\int_0^\pi \left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)^2\, dx$$

is minimized. The solution $Q(x)$ is called the quadratic **least squares approximation** to $\sin(x)$ over $[0, \pi]$.

The key to solving this problem is to observe that a_0, a_1, and a_2 can take on any values we like and that this integral can thus be considered a function of these three variables. For instance, if we couldn't think of anything cleverer to do, we might simply try various combinations of a_0, a_1, and a_2 to see how small we could make the given integral. Therefore another way to phrase our problem is

A mathematical formulation of the problem

Find values for a_0, a_1, and a_2 that minimize the function

$$F(a_0, a_1, a_2) = \int_0^\pi \left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)^2\, dx\,.$$

We know how to find points where functions take on their extreme values—we look for the places where the partial derivatives are 0. But how do we differentiate an expression involving an integral like this? It turns out that for all continuous functions, or even functions with only a finite number of breaks in them, we can simply interchange integration and differentiation. Thus, in our example,

$$\begin{aligned}\frac{\partial}{\partial a_0} F(a_0, a_1, a_2) &= \frac{\partial}{\partial a_0} \int_0^\pi \left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)^2\, dx \\ &= \int_0^\pi \frac{\partial}{\partial a_0} \left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)^2\, dx \\ &= \int_0^\pi 2\left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)(-1)\, dx.\end{aligned}$$

Similarly we have

$$\begin{aligned}\frac{\partial}{\partial a_1} F(a_0, a_1, a_2) &= \int_0^\pi 2\left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)(-x)\, dx, \\ \frac{\partial}{\partial a_2} F(a_0, a_1, a_2) &= \int_0^\pi 2\left(\sin(x) - (a_0 + a_1\, x + a_2\, x^2)\right)(-x^2)\, dx.\end{aligned}$$

We now want to find values for a_0, a_1, and a_2 that make these partial derivatives simultaneously equal to 0. That is, we want

Setting the partials equal to zero gives equations for a_0, a_1, a_2

$$\int_0^\pi 2\left(\sin(x) - (a_0 + a_1\,x + a_2\,x^2)\right)(-1)\,dx = 0,$$

$$\int_0^\pi 2\left(\sin(x) - (a_0 + a_1\,x + a_2\,x^2)\right)(-x)\,dx = 0,$$

$$\int_0^\pi 2\left(\sin(x) - (a_0 + a_1\,x + a_2\,x^2)\right)(-x^2)\,dx = 0,$$

which can be rewritten as

$$\int_0^\pi \sin(x)\,dx = \int_0^\pi \left(a_0 + a_1\,x + a_2\,x^2\right)\,dx,$$

$$\int_0^\pi x\,\sin(x)\,dx = \int_0^\pi \left(a_0\,x + a_1\,x^2 + a_2\,x^3\right)\,dx,$$

$$\int_0^\pi x^2\,\sin(x)\,dx = \int_0^\pi \left(a_0\,x^2 + a_1\,x^3 + a_2\,x^4\right)\,dx.$$

All of these integrals can be evaluated relatively easily (see the exercises for a hint on evaluating the integrals on the left–hand side). When we do so, we are left with

Evaluating the integrals gives three linear equations

$$\begin{aligned} 2 &= \pi a_0 + \frac{\pi^2}{2}a_1 + \frac{\pi^3}{3}a_2, \\ \pi &= \frac{\pi^2}{2}a_0 + \frac{\pi^3}{3}a_1 + \frac{\pi^4}{4}a_2, \\ \pi^2 - 4 &= \frac{\pi^3}{3}a_0 + \frac{\pi^4}{4}a_1 + \frac{\pi^5}{5}a_2. \end{aligned}$$

But this is simply a set of three linear equations in the unknowns a_0, a_1, and a_2, and they can be solved in the usual ways. We could either replace each expression in π by a corresponding decimal approximation, or we could keep everything in terms of π. Let's do the latter; after a bit of tedious arithmetic we find

$$
\begin{aligned}
a_0 &= \frac{12}{\pi} - \frac{120}{\pi^3} &&= -.050465\ldots, \\
a_1 &= \frac{-60}{\pi^2} + \frac{720}{\pi^4} &&= 1.312236\ldots, \\
a_2 &= \frac{60}{\pi^3} - \frac{720}{\pi^5} &&= -.417697\ldots,
\end{aligned}
$$

and we have

$$Q(x) = -.050465 + 1.312236\,x - .417698\,x^2,$$

which is the equation given in section 1 at the beginning of the chapter.

The analysis we gave for this particular case can clearly be generalized to apply to any function over any interval. When we do this we get

How to find least squares polynomial approximations in general

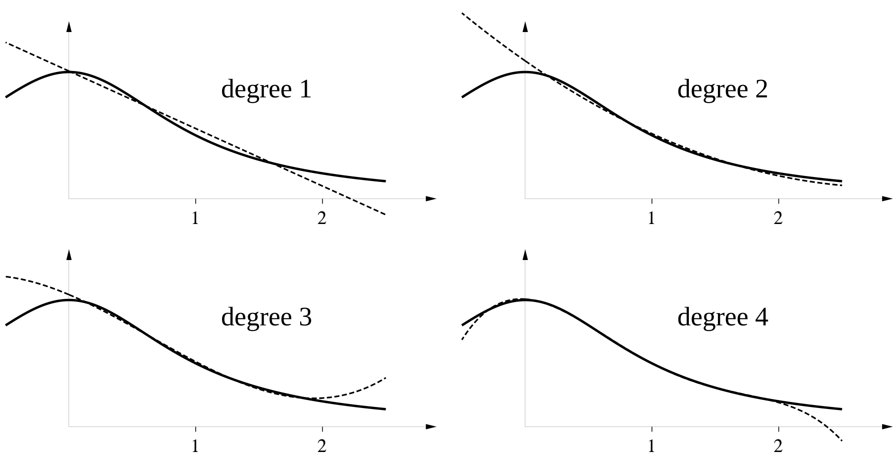
> Given a function g over an interval $[a, b]$, then the n-th degree polynomial
>
> $$P(x) = c_0 + c_1\,x + c_2\,x^2 + \cdots + c_n\,x^n$$
>
> whose mean square distance from g is a minimum has coefficients that are determined by the following $n+1$ equations in the $n+1$ unknowns c_0, c_1, c_2, …, c_n:
>
> $$
\begin{aligned}
\int_a^b g(x)\,dx &= c_0 \int_a^b dx + c_1 \int_a^b x\,dx + \cdots + c_n \int_a^b x^n\,dx, \\
\int_a^b x\,g(x)\,dx &= c_0 \int_a^b x\,dx + c_1 \int_a^b x^2\,dx + \cdots + c_n \int_a^b x^{n+1}\,dx, \\
&\vdots \\
\int_a^b x^n\,g(x)\,dx &= c_0 \int_a^b x^n\,dx + c_1 \int_a^b x^{n+1}\,dx + \cdots + c_n \int_a^b x^{2n}\,dx.
\end{aligned}
$$

All the integrals on the right-hand side can be evaluated immediately. The integrals on the left-hand side will typically need to be evaluated numerically, although simple cases can be evaluated in closed form. Integration by parts is often useful in these cases. The exercises contain several problems using this technique to find approximating polynomials.

The real catch, though, is not in obtaining the equations—it is that solving systems of equations by hand is excruciatingly boring and subject to frequent arithmetic mistakes if there are more than two or three unknowns involved. Fortunately, there are now a number of computer packages available which do all of this for us. Here are a couple of examples, where the details are left to the exercises.

Solving the equations is a job for the computer

Example. Let's find polynomial approximation for $1/(1+x^2)$ over the interval $[0, 2]$. We saw earlier that the Taylor series for this function converges only for $|x| < 1$, so it will be no help. Yet with the above technique we can derive the following approximations of various degrees (see the exercises for details):

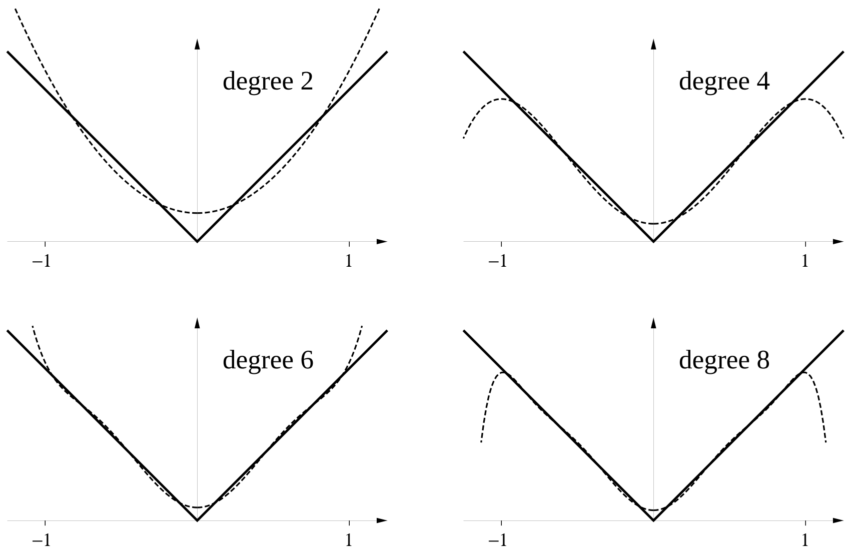


Here are the corresponding equations of the approximating polynomials:

degree	polynomial
1	$1.00722 - .453645\,x$
2	$1.08789 - .695660\,x + .121008\,x^2$
3	$1.04245 - .423017\,x - .219797\,x^2 + .113602\,x^3$
4	$1.00704 - .068906\,x - 1.01653\,x^2 + .733272\,x^3 - .154916\,x^4$

Example. We can even use this new technique to find polynomial approximations for functions that aren't differentiable at some points. For instance, let's approximate the function $h(x) = |x|$ over the interval $[-1, 1]$. Since this function is symmetric about the y–axis, and we are approximating it over an interval that is symmetric about the y–axis, only even powers of x will

The technique works even when differentiability fails appear. (See the exercises for details.) We get the following approximations of degrees 2, 4, 6, and 8:



Here are the corresponding polynomials:

degree	polynomial
2	$.1875 + .9375\,x^2$
4	$.117188 + 1.64062\,x^2 - .820312\,x^4$
6	$.085449 + 2.30713\,x^2 - 2.81982\,x^4 + 1.46631\,x^6$
8	$.067291 + 2.960821\,x^2 - 6.415132\,x^4 + 7.698173\,x^6 - 3.338498\,x^8$

The technique is useful for data functions

A Numerical Example. If we have some function which exists only as a set of data points—a numerical solution to a differential equation, perhaps, or the output of some laboratory instrument—it can often be quite useful to replace the function by an approximating polynomial. The polynomial takes up much less storage space and is easier to manipulate. To see how this works, let's return to the S-I-R model we've studied before

$$S' = -.00001\,SI,$$
$$I' = .00001\,SI - I/14,$$
$$R' = I/14,$$

with initial values $S(0) = 45400$ and $I(0) = 2100$.

Let's find an 8th degree polynomial $Q(t) = i_0 + i_1 t + i_2 t^2 + \cdots + i_8 t^8$ approximating I over the time interval $0 \le t \le 40$. We can do this by a minor modification of the Euler's method programs we've been using all along. Now, in addition to keeping track of the current values for S and I as we go along, we will also need to be calculating Riemann sums for the integrals

$$\int_0^{40} t^k I(t)\,dt \qquad \text{for } k = 0, 1, 2, \ldots, 8,$$

as we go through each iteration of Euler's method.

Since the numbers involved become enormous very quickly, we open ourselves to various sorts of computer roundoff error. We can avoid some of these difficulties by **rescaling** our equations—using units that keep the numbers involved more manageable. Thus, for instance, suppose we measure S, I, and R in units of 10,000 people, and suppose we measure time in "decadays", where 1 decaday = 10 days. When we do this, our original differential equations become

The importance of using the right-sized units

$$\begin{aligned} S' &= -SI, \\ I' &= SI - I/1.4, \\ R' &= I/1.4, \end{aligned}$$

with initial values $S(0) = 4.54$ and $I(0) = 0.21$. The integrals we want are now of the form

$$\int_0^{4} t^k I(t)\,dt \qquad \text{for } k = 0, 1, 2, \ldots, 8.$$

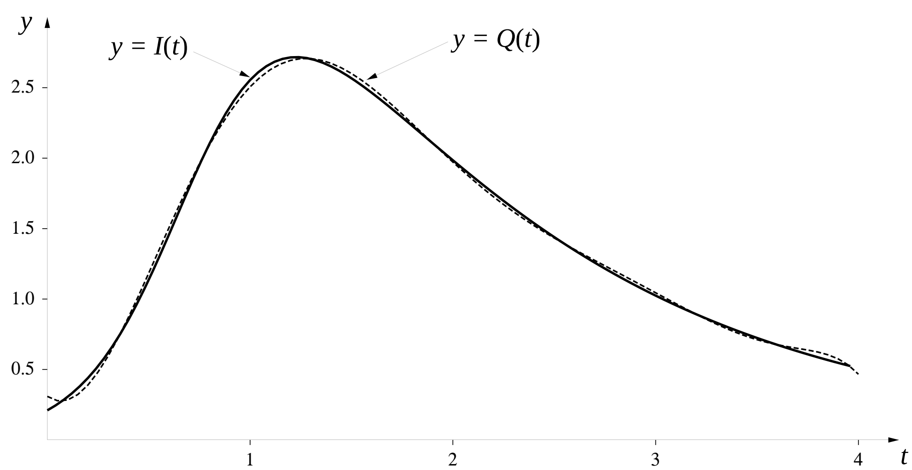
The use of Simpson's rule (see chapter 11.3) will also reduce errors. It may be easiest to calculate the values of I first, using perhaps 2000 values, and store them in an array. Once you have this array of I values, it is relatively quick and easy to use Simpson's rule to calculate the 9 integrals needed. If you later decide you want to get a higher-degree polynomial approximation, you don't have to re-run the program.

Using Simpson's rule helps reduce errors

Once we've evaluated these integrals, we set up and solve the corresponding system of 9 equations in the 9 unknown coefficients i_k. We get the following 8-th degree approximation

$$\begin{aligned} Q(t) = {} & .3090 - .9989\,t + 7.8518\,t^2 - 3.6233\,t^3 - 3.9248\,t^4 + 4.2162\,t^5 \\ & - 1.5750\,t^6 + .2692\,t^7 - .01772\,t^8 . \end{aligned}$$

When we graph Q and I together over the interval $[0, 4]$ (decadays), we get



—a reasonably good fit.

A Caution. Numerical least-squares fitting of the sort performed in this last example fairly quickly pushes into regions where the cumulative effects of the inaccuracies of the original data, the inaccuracies of the estimates for the integrals, and the immense range in the magnitude of the numbers involved all combine to produce answers that are obviously wrong. Rescaling the equations and using good approximations for the integrals can help put off the point at which this begins to happen.

Exercises

1. To find polynomial approximations for $\sin(x)$ over the interval $[0, \pi]$, we needed to be able to evaluate integrals of the form

$$\int_0^\pi x^n \sin(x)\, dx.$$

The value of this integral clearly depends on the value of n, so denote it by I_n.

a) Evaluate I_0 and I_1. Suggestion: use integration by parts (Chapter 11.3) to evaluate I_1.

[Answer: $I_0 = 2$, and $I_1 = \pi$.]

b) Use integration by parts twice to prove the general **reduction formula**:

$$I_{n+2} = \pi^{n+2} - (n+2)(n+1)I_n \qquad \text{for all } n \geq 0.$$

c) Evaluate I_2, I_3, and I_4.

[Answer: $I_2 = \pi^2 - 4$, $I_3 = \pi^3 - 6\pi$, and $I_4 = \pi^4 - 12\pi^2 + 48$.]

d) If you have access to a computer package that will solve a system of equations, find the 4-th degree polynomial that best fits the sine function over the interval $[0, \pi]$. What is the maximum difference between this polynomial and the sine function over this interval?

[Answer: $.00131 + .98260\, x + .05447\, x^2 - .23379\, x^3 + .03721\, x^4$, with maximum difference occuring at the endpoints.]

2. To find polynomial approximations for $|x|$ over the interval $[-1, 1]$, we needed to be able to evaluate integrals of the form

$$\int_{-1}^{1} x^n |x| \, dx.$$

As before, let's denote this integral by I_n.

a) Show that

$$I_n = \begin{cases} \dfrac{2}{n+2} & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

b) Derive the quadratic least squares approximation to $|x|$ over $[-1, 1]$.

c) If you have access to a computer package that will solve a system of equations, find the 10-th degree polynomial that best fits $|x|$ over the interval $[-1, 1]$. What is the maximum difference between this polynomial and $|x|$ over this interval?

3. To find polynomial approximations for $1/(1+x^2)$ over the interval $[0, 2]$, we needed to be able to evaluate integrals of the form

$$\int_0^2 \frac{x^n}{1+x^2} \, dx.$$

Call this integral I_n.

a) Evaluate I_0 and I_1.

[Answer: $I_0 = \arctan(1) = \pi/4$, and $I_1 = (\ln 2)/2 = .3465736$.]

b) Prove the general reduction formula:

$$I_{n+2} = \frac{2^{n+1}}{n+1} - I_n \qquad \text{for } n = 0,\, 1,\, 2,\, \ldots.$$

c) Evaluate I_2, I_3, and I_4.

[Answer: $I_2 = 2 - \frac{\pi}{4}, I_3 = 2 - \frac{\ln 2}{2}, I_4 = \frac{2}{3} + \frac{\pi}{4}$.]

d) If you have access to a computer package that will solve a system of equations, find the 4-th degree polynomial that best fits $1/(1+x^2)$ over the interval $[0,2]$. What is the maximum difference between this polynomial and the function over this interval?

4. Set up the equations (including evaluating all the integrals) for finding the best fitting 6-th degree polynomial approximation to $\sin(x)$ over the interval $[-\pi, \pi]$.

5. In the S-I-R model, find the best fitting 8-th degree polynomial approximation to $S(t)$ over the interval $0 \le t \le 40$.

10.7 Chapter Summary

The Main Ideas

- **Taylor polynomials** approximate functions at a point. The Taylor polynomial $P(x)$ of degree n is the **best fit** to $f(x)$ at $x = a$; that is, P satisfies the following conditions: $P(a) = f(a)$, $P'(a) = f'(a)$, $P''(a) = f''(a)$, …, $P^{(n)}(a) = f^{(n)}(a)$.

- **Taylor's theorem** says that a function and its Taylor polynomial of degree n agree to order $n+1$ near the point where the polynomial is centered. Different versions expand on this idea.

- If $P(x)$ is the Taylor polynomial approximating $f(x)$ at $x = a$, then $P(x)$ approximates $f(x)$ for values of x near a; $P'(x)$ approximates $f'(x)$; and $\int P(x)\,dx$ approximates $\int f(x)\,dx$.

- A **Taylor series** is an infinite sum whose partial sums are Taylor polynomials. Some functions *equal* their Taylor series; among these are the sine, cosine and exponential functions.

- A **power series** is an "infinite" polynomial

$$a_0 + a_1 x + a_x x^2 + \cdots + a_n x^n + \cdots .$$

If the solution of a differential equation can be represented by a power series, the coefficients a_n can be determined by **recursion relations** obtained by substituting the power series into the differential equation.

- An infinite series **converges** if, no matter how many decimal places are specified, all the partial sums eventually agree to at least this many decimal places. The number defined by these stabilizing decimals is called the **sum** of the series. If a series does not converge, we say it **diverges**.

- If the series $\sum_{m=0}^{\infty} b_m$ converges, then $\lim_{m\to\infty} b_m = 0$. The important counter-example of the **harmonic series** $\sum_{m=1}^{\infty} 1/m$ shows that $\lim_{m\to\infty} b_m = 0$ is a necessary but not sufficient condition to guarantee convergence.

- The **geometric series** $\sum_{m=0}^{\infty} x^m$ converges for all x with $|x| < 1$ and diverges for all other x.

- An **alternating series** $\sum_{m=0}^{\infty}(-1)^m b_m$ converges if $0 < b_{m+1} \leq b_m$ for all m and $\lim_{m\to\infty} b_m = 0$. For a convergent alternating series, the error in approximating the sum by a partial sum is less than the next term in the series.

- A convergent power series converges on an **interval of convergence** of width $2R$; R is called the **radius of convergence**. The **ratio test** can be used to find the radius of convergence of a power series: $\sum_{m=0}^{\infty} b_m$ converges if $\lim_{m\to\infty} |b_{m+1}|/|b_m| < 1$.

- A polynomial $P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ is the **best fitting** approximation to a function $f(x)$ on an interval $[a, b]$ if $a_0, a_1, \cdots, a_n$ are chosen so that the **mean squared separation** between P and f

$$\frac{1}{b-a}\int_a^b (P(x) - f(x))^2\,dx$$

is as small as possible. The polynomial P is also called the **least squares approximation** to f on $[a, b]$.

Expectations

- Given a differentiable function $f(x)$ at a point $x = a$, you should be able to write down any of the **Taylor polynomials** or the **Taylor series** for f at a.
- You should be able to use the program TAYLOR to graph Taylor polynomials.
- You should be able to obtain new Taylor polynomials by substitution, differentiation, anti-differentiation and multiplication.
- You should be able to use Taylor polynomials to find the value of a function to a specified degree of accuracy, to approximate integrals and to find limits.
- You should be able to determine the order of magnitude of the agreement between a function and one of its Taylor polynomials.
- You should be able to find the power series solution to a differential equation.
- You should be able to test a series for divergence; you should be able to check a series for convergence using either the **alternating series test** or the **ratio test**.
- You should be able to find the sum of a **geometric series** and its interval of convergence.
- You should be able to estimate the error in an approximation using partial sums of an alternating series.
- You should be able to find the **radius of convergence** of a series using the ratio test.
- You should be able to set up the equations to find the **least squares** polynomial approximation of a particular degree for a given function on a specified interval. Working by hand or, if necessary, using a computer package to solve a system of equations, you should be able to find the coefficients of the least squares approximation.