Chapter 4
Approximating with Polynomials and Taylor Series

4.1 The Idea:

As we said in Chapter 1, the basic idea on which all of differential Calculus is based is this

Lines are simple, approximate complicated functions with lines.

The lines used in differential Calculus are tangent lines. They are good approximations near the point of tangency, as we see from the most common picture from differential Calculus.

Figure 4.1: Graph of $f(x) = x^2$ and line tangent at $x = 1/3$.

For example, the function $f(x) = x^2$ is close to the line $y = 2x/3 - 1/9$ near the point of tangency which occurs at $x = 1/3$ (see Figure above). Far from $x = 1/3$ then line and the function $f(x) = x^2$ are not close at all. Similarly, $g(x) = \sin(x)$ is approximately equal to $y = x$ near $x = 0$, but only near $x = 0$.

Figure 4.2: Graph of $g(x) = \sin x$ and line tangent at $x = 0$.

Once we know the slope of the tangent line (i.e., the derivative) of a function at a point, then computing the tangent line is straight forward. From the slope of the tangent line, we can tell if the function is increasing or decreasing near the point of tangency or if it is a candidate for being a local maximum of minimum.

But why only use lines? Certainly lines are simple, both algebraically and geometrically, but we pay for this simplicity because the distance between the tangent line of a function and the function generally grows quickly as we move far from the point of tangency. Could we approximate with another type of function which is almost as simple as a line, but for which the degree of approximation was much better than with a line? Quadratic equations (those involving x^2, squared terms) are pretty simple, why not approximate with them instead of lines? Cubics (polynomials involving x^3) are not that much more complicated than quadratics. Why not use cubics? Or why not use fourth degree polynomials? Certainly, the more complicated the approximating functions we use, the more accurate we can make the approximation over a larger interval from the point of tangency.

We should begin at the beginning.

4.2 Approximating Lines:

First we consider carefully just how good an approximation the tangent line really is. Consider a function $f(x)$ and fix a point x_0. How close is the tangent line of $f(x)$ at x_0 to $f(x)$?

Well, recall that the derivative $f'(x_0)$ is defined to be

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Hence, for x close to x_0 we have that

$$f'(x_0)(x - x_0) \approx f(x) - f(x_0)$$

or

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

The right hand side of this equation is just the formula for the tangent line for $f(x)$ at $x = x_0$. This is another way to say that $f(x)$ is close to its tangent line at $x = x_0$. This begs the question "How close is it?"

Let the error betweenthe function and its tangent line at x_0 be $E(x)$, that is

$$E(x) = f(x) - (f(x_0) + f'(x_0)(x - x_0).$$

Then we can compute that

$$E(x_0) = 0 \text{ and } E'(x_0) = 0.$$

Suppose we somehow know that there is a number M such that

$$-M \leq f''(x) \leq M$$

for all x. Note that

$$E'(x) = f'(x) - f'(x_0)$$

and

$$E''(x) = f''(x),$$

so we ahve

$$-M \leq E''(x) \leq M$$

for all x. We integrate all sides of this inequality

$$\int_{x_0}^{x} -M\,dx \leq \int_{x_0}^{x} E''(x)\,dx \leq \int_{x_0}^{x} M\,dx$$

to obtain (for $x > x_0$),

$$-M(x-x_0) \leq E'(x) - E'(x_0) \leq M(x-x_0).$$

But $E'(x) = f'(x) - f'(x_0)$ so $E'(x_0) = 0$ and we have

$$-M(x-x_0) \leq E'(x) \leq M(x-x_0).$$

Integrating again, we get

$$\int_{x_0}^{x} M(x-x_0)\,dx \leq \int_{x_0}^{x} E'(x)\,dx \leq \int_{x_0}^{x} M(x-x_0)\,dx,$$

so, evaluating and using $E(x_0) = 0$, we have

$$\frac{-M}{2}(x-x_0)^2 \leq E(x) \leq \frac{M}{2}(x-x_0)^2,$$

that is

$$|E(x)| \leq \frac{M}{2}(x-x_0)^2.$$

To show this inequality for $x < x_0$, we repeat the above computation with integrals

$$\int_{x}^{x_0} \quad .$$

We have established a bound on the size of the "error" $E(x)$ term between the function and its tangent line at $x = x_0$. The bound depends on knowing the number M such that

$$|f''(x)| \leq M$$

for all x. So what we need to know is that the second derivative of $f(x)$ satisfies

$$|f''(x)| < M$$

for all x. This shows the following:

Theorem: Suppose a function $f(x)$ is twice continuously differentiable and suppose there is a number M such that

$$|f''| \leq M$$

for all x. Then

$$|E(x)| = |f(x) - (f(x_0) + (x - x_0)f'(x_0))| \leq \frac{M}{2}|x - x_0|^2.$$

We have accomplished two good things. First, we have shown that the tangent line to a function $f(x)$ at $x = x_0$ is given by

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Second, we have given an explicit estimate of the difference between $f(x)$ and its tangent line at x_0,

$$|f(x) - (f(x_0) + f'(x_0)(x - x_0))| \leq \frac{M}{2}|x - x_0|^2,$$

where $|f''(x)| \leq M$ for all x. The important thing to remember about this estimate is that the distance between $f(x)$ and its tangent line at x_0 grows like a constant times $|x - x_0|^2$ as x moves away from x_0.

In order to get the estimate we need some extra information about $f(x)$ (a bound on its second derivative). This is only to be expected since the distance between a tangent line and the function varies widely depending on how fast the function changes.

Getting a bound on the derivative $f''(x)$ for all x can be very difficult. Looking at the discussion above, we only used the bound

$$|f''(s)| \leq M$$

for s between x_0 and x.

The tangent line to $f(x)$ at $x = x_0$ is also called "the first order approximation to $f(x)$ at x_0" or "the first degree Taylor polynomial for $f(x)$ at x_0". The words "first order" and "first degree" indicate that the highest power of $(x - x_0)$ that appears in the formula for the tangent line is the first power. Likewise, the distance between the tangent line and the function is "second order" in $x - x_0$ because it is less than a constant times $|x - x_0|^2$.

When we say that the tangent line approximates a function near the point of tangency, we now guarentee on how good an approximation this is. For example, take $f(x) = \sin(x)$ and $x_0 = \pi/6$. We compute that

$$\sin(x) \approx \sin\left(\frac{\pi}{6}\right) + \cos\left(\frac{\pi}{6}\right)\left(x - \frac{\pi}{6}\right).$$

Because $f''(x) = -\sin(x)$ is satisfies $|-\sin(x)| \leq 1$ for all x, we have that

$$\left|\sin(x) - \left(\sin\left(\frac{\pi}{6}\right) + \cos\left(\frac{\pi}{6}\right)\left(x - \frac{\pi}{6}\right)\right)\right| < \frac{1}{2}\left|x - \frac{\pi}{6}\right|^2.$$

Note that the line tangent line at $x = \pi/6$ and the function $\sin x$ can be closer than the estimate on the error. In fact, the graphs of the line and $\sin x$ actually cross again for $x < 0$. The estimate is a guarentee on the largest the error could be.

Figure 4.3: Tangent line to $\sin x$ at $x = \pi/6$.

Similarly, for the function e^x, the approximation at $x_0 = 1$ is

$$e^x \approx e + e \cdot (x - 1),$$

where we used that $d^2(e^x)/dx^2 = e^x$ evaluated at $x_0 = 1$ is e. So

$$\left|\frac{d^2(e^x)}{dx^2}\right| = |e^x| < e^2 \qquad \text{for } 0 \leq x \leq 2,$$

and so

$$|e^x - (e + e(x - 1))| \leq \frac{e^2}{2}|x - 1|^2.$$

We get a different approximation if we take $x_0 = 0$. Using that $e^0 = 1$, we get that

$$e^x \approx 1 + 1 \cdot x.$$

Using the same estimate on the second derivative for e^x on $0 \leq x \leq 2$ we have that

$$|e^x - (1 + x)| \leq \frac{e^2}{2}x^2,$$

for $0 \leq x \leq 2$. This approximation has the advantage that it doesn't use the number e, so no error is introduced in approximating e.

Figure 4.4: Approximations of e^x

4.3 Quadratic Approximations:

Like everything else in life, using linear approximations involves a trade off. Lines are simple and easy to compute with, but away from the point of tangency the tangent line is not a very accurate approximation. To obtain greater accuracy, we can use quadratic polynomials tangent to the graph of the given function. We get better accuracy, but must pay the price in greater complication.

There are lots of ways to choose a quadratic that approximates a given function just like there are lots of ways to choose a line that approximates the function. In Calculus, we "work locally", which means we choose the quadratic that best approximates the given function near a particular point.

Given a function $f(x)$ and a point x_0, we wish to find the quadratic polynomial, $P_2(x)$, of the form

$$P_2(x) = a + b(x - x_0) + c(x - x_0)^2$$

that best approximates $f(x)$ near x_0 (we use the subscript 2 on $P_2(x)$ to remind us that it is a quadratic polynomial). To do this we must choose the coefficients a, b and c.

Since $P_2(x_0) = a$, in order to have $f(x)$ and $P_2(x)$ be tangent at x_0 the two functions must have the same value at x_0, so we must pick

$$a = f(x_0).$$

Also, in order for $f(x)$ and $P_2(x)$ be tangent at x_0, they must have the same derivative at x_0). We compute

$$P_2'(x) = b + 2c(x - x_0)$$

so $P_2'(x_0) = b$. We choose

$$P_2'(x_0) = f'(x_0)$$

or

$$b = f'(x_0).$$

To continue this pattern, we should choose c so that

$$P_2''(x_0) = f''(x_0).$$

Since

$$P_2''(x_0) = 2c,$$

this implies that we should take

$$c = \frac{f''(x_0)}{2}.$$

So our candidate for "tangent quadratic" to $f(x)$ at $x = x_0$ is

$$P_2(x) = f(x_0) + f'(x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

Note that the tangent line to $f(x)$ at $x = x_0$ gives the constant and linear terms of this polynomial, as we would expect. Forming the tangent quadratic involves choosing only a second degree term to add to the tangent line.

To determine if this is the choice that we want, let

$$E_2(x) = f(x) - P_2(x),$$

hence,

$$f'(x_0) = P_2'(x_0)$$

and

$$\left(f'\right)'(x_0) = \left(P_2'\right)'(x_0).$$

That is, we have chosen $P_2(x)$ so that

$$E_2(x_0) = E_2'(x_0) = E_2''(x_0) = 0.$$

Now $P_2'(x)$ is linear function, so it gives the equation for the line tangent to $f'(x)$ at $x = x_0$. Suppose M_3 is a constant such that

$$|f'''| \leq M_3$$

for all x. We can use the estimate of the last section to show that

$$E_2'(x) = f'(x) - P'(x)$$

satisfies

$$|E_2'(x)| \leq \frac{M_3}{2}\left|x - x_0\right|^2.$$

To turn this into an estimate on $E_2(x)$, we use that

$$-\frac{M_3}{2}(x-x_0)^2 \le E_2'(x) \le \frac{M_3}{2}(x-x_0)^2,$$

and integrate to obtain

$$\int_{x_0}^{x} -\frac{M_3}{2}(x-x_0)^2\,dx \le \int_{x_0}^{x} E_2'(x)\,dx \le \int_{x_0}^{x} \frac{M_3}{2}(x-x_0)^2\,dx$$

so

$$-\frac{M_3}{3\cdot 2}(x-x_0)^3 \le E_2(x)-E_2(x_0) \le \frac{M_3}{3\cdot 2}(x-x_0)^3,$$

or

$$|E_2(x)| \le \frac{M_3}{3\cdot 2}|x-x_0|^3.$$

This is the best we could hope for. We have taken the first order approximation of $f(x)$, added a quadratic term to obtain a better approximation near x_0. The difference between the quadratic approximation and the original function involves a constant that depends on the third derivative of $f(x)$ and on the distance from x_0 cubed.

As an example, consider $f(x)=\cos(x)$ with $x_0=0$. To find the tangent quadratic equation at $x=0$, we compute $f'(x)=-\sin(x)$ and $f''(x)=-\cos(x)$. Hence, near $x_0=0$ we have

$$\cos(x) \approx 1+0\cdot(x-0)-\frac{1}{2}\cdot(x-0)^2 = 1-\frac{x^2}{2}.$$

The third derivative of $\cos(x)$ is $-\sin(x)$, and $|-\sin(x)| \le 1$ for all x. Hence

$$\left|\cos(x)-\left(1-\frac{x^2}{2}\right)\right| \le \frac{1}{6}|x|^3.$$

For x near zero, this is quite accurate. If $x<0.1$ then

$$\left|\cos(x)-\left(1-\frac{x^2}{2}\right)\right| < 0.00017.$$

We can collect the discussion above into a definition and a theorem.

Definition: Given a function $f(x)$, the **second order** or **second degree** or **quadratic Taylor polynomial** for $f(x)$ at $x=x_0$ is

$$P_2(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2}(x-x_0)^2.$$

Theorem: Suppose $f(x)$ is three times continuously differentiable. Fix two numbers x_0 and x and suppose there is a number M_3 such that for all z between x_0 and x we have

$$|f'''(z)| \le M_3.$$

Then

$$|f(x) - P_2(x)| \leq \left| f(x) - \left(f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \right) \right| \leq \frac{M_3}{2 \cdot 3} |x - x_0|^3 .$$

Note that the first degree Taylor polynomial gives the first order term of the second degree Taylor polynomial. So to compute the second degree Taylor polynomial of $f(x) = e^x$ at $x = 1$ we can use the calculation at the end of the previous section that

$$e^x \approx e + e(x - 1).$$

Then, since

$$f''(x) = f'''(x) = e^x$$

the second degree Taylor polynomial of $f(x) = e^x$ at $x_0 = 1$ is

$$e^x \approx e + e(x - 1) + \frac{e}{2}(x - 1)^2 .$$

Also,

$$|f'''(x)| \leq e^2$$

for $0 \leq x \leq 2$, so we have

$$\left| e^x - \left(e + e(x - 1) + \frac{e}{2}(x - 1)^2 \right) \right| < \frac{e^2}{6} |x - 1|^3 .$$

Similarly, for $x_0 = 0$, we get the second order Taylor polynomial of e^x is

$$e^x \approx 1 + x + \frac{1}{2}x^2$$

and for $0 \leq x \leq 2$

$$\left| e^x - \left(1 + x + \frac{1}{2}x^2 \right) \right| \leq \frac{e^2}{6} |x|^3 .$$

Figure 4.5 Quadratic approximations for e^x centered at $x = 0$ and $x = 1$.

4.4 The n^{th} Degree Approximation

We step up now to the cubic or third degree or cubic $P_3(x)$ tangent to a function $f(x)$ at $x = x_0$ by taking the quadratic approximation

$$P_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$$

and adding a cubic term, that is

$$P_3(x) = P_2(x) + k(x - x_0)^3.$$

We choose the constant k so that

$$f'''(x_0) = P_3'''(x_0),$$

this will guarentee that the first three derivatives of the error term

$$E_3(x) = f(x) - P_3(x)$$

are zero at $x = x_0$. Since

$$P_3'''(x_0) = 3 \cdot 2k$$

we get that $k = f'''(x_0)/3!$ and

$$P_3(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \frac{f'''}{3!}(x_0)(x - x_0)^3.$$

We could continue, constructing the fourth degree (quartic) polynomial tangent to $f(x)$ at $x = x_0$, but the pattern is becoming clear, so we jump to the general n^{th} case.

Definition: The n^{th} **degree** or n^{th} **order Taylor Polynomial** or **power series** of a function $f(x)$ centered at $x = x_0$ is

$$P_{x_0,n}(x) = f(x_0) + \frac{1}{1!}\frac{df}{dx}\bigg|_{x_0}(x - x_0) + \frac{1}{2!}\frac{d^2 f}{dx^2}\bigg|_{x_0}(x - x_0)^2 + \ldots + \frac{1}{n!}\frac{d^n f}{dx^n}\bigg|_{x_0}(x - x_0)^n$$

or, in summation notation

$$P_{x_0,n}(x) = \sum_{j=0}^{n} \frac{1}{j!}\frac{d^j f}{dx^j}\bigg|(x - x_0)^j$$

(where we use the conventions that $0! = 1$ and

$$\frac{d^0 f}{dx^0}\bigg|_x = f(x).)$$

The notation $P_{x_0,n}$ is not standard. That means you must say what it is before you use it.

Theorem (Taylor's Theorem) : Suppose $f(x)$ is an $n+1$-times continuously differentiable function. Suppose there is a number M_{n+1} such that for all z between x and x_0 we have

$$\left| \left. \frac{d^{n+1}f}{dx^{n+1}} \right|_z \right| \leq M_{n+1}.$$

If $P_{x_0,n}$ is the n^{th}-degree Taylor polynomial for $f(x)$ at x_0 Then

$$|f(x) - P_{x_0,n}(x)| \leq \frac{M_{n+1}}{(n+1)!} |x - x_0|^{n+1}.$$

This expression is called **Taylor's inequality**.

The $n = 1$ case of this theorem is discussed in Section 4.2 and the $n = 2$ case in Section 4.3. The definition and theorem above just continue the pattern established in those sections.

Checking this theorem is not as hard as you might imagine. We can use a bootstrapping technique called "mathematical induction". We already have shown the theorem for $n = 1$ and $n = 2$. Now, let $n = 3$. Consider a function $f(x)$, points x_0, and the constant M_{n+1} such that

$$\left| \frac{d^{n+1}f}{dx^{n+1}} \right| \leq M_{n+1},$$

for z between x and x_0. If we let

$$h(x) = f'(x)$$

then we have that

$$\left| \frac{d^n h}{dx^n} \right| \leq M_{n+1},$$

for z between x and x_0. We construct the Taylor polynomial for the function $h(x)$ of degree $n-1$ at x_0,

$$Q(x) = h(x_0) + \ldots + \frac{1}{(n-1)!} \left. \frac{d^n h}{dx^n} \right|_{x_0} (x - x_0)^{n-1}$$

and we know that

$$|h(x) - Q(x)| \leq \frac{M_{n+1}}{n!} |x - x_0|^n.$$

We make a couple of simplifying assumptions. Fix $x > x_0$ and assume $Q(z) < h(z)$ for $x_0 < z < x$. This is just one possible configuration of these functions. To get a complete proof, we would have to deal with all the other possibilities as well. This assumption allows us to remove the absolute values in the estimate above.

Now,

$$\int_{x_0}^{x} h(s)\,dx = f(x) - f(x_0)$$

since h is the derivative of f. Also, we have

$$\int_{x_0}^{x} Q(s)\,ds = \left.\frac{df}{dx}\right|_{x_0} (x - x_0) + \ldots \frac{1}{n!} \left.\frac{d^n f}{dx^n}\right|_{x_0} (x - x_0)^n.$$

Using these formulas, if we integrate both sides of

$$h(x) - Q(x) \leq \frac{M_{n+1}}{n!}(x - x_0)^n$$

we have

$$f(x) - f(x_0) - \left(\left.\frac{df}{dx}\right|_{x_0} (x - x_0) + \ldots \frac{1}{n!} \left.\frac{d^n f}{dx^n}\right|_{x_0} (x - x_0)^n \right) \leq \frac{M_{n+1}}{(n+1)!}(x - x_0)^{n+1}.$$

But this is precisely what we wanted to show because

$$f(x_0) + \left(\left.\frac{df}{dx}\right|_{x_0} (x - x_0) + \ldots \frac{1}{n!} \left.\frac{d^n f}{dx^n}\right|_{x_0} (x - x_0)^n \right)$$

is the n^{th} degree Taylor polynomial of f at x_0

At the beginning of this argument we said that we knew the theorem was true for $n = 2$ so let $n = 3$. During the argument, we never specifically used the fact that $n = 3$. What we showed is that "If the theorem is true for $n = N$ then it is also true for $n = N + 1$." Since we already know it is true for $n = 2$ it must be true for $n = 3$ and since it is true for $n = 3$ it must be true for $n = 4$, and so on. For any value of n we are given, we could repeat the argument above until we know the theorem is true for that n. Hence, it must be true for all n.

4.5 Remarks on Taylor Polynomials

There are a number of Taylor polynomials you should have at your fingertips (that is, memorize). Verifying that these really are the formulas make a bunch of nice exercises. In the formulas below we use the convention that $0! = 1$. While this does not make a lot of sense from the definition of factorial ($n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$), it does make it a lot easier to write compact summation formulas.

- $$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \frac{x^n}{n!} = \sum_{j=0}^{n} \frac{x^j}{j!}.$$

- $$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots + x^n = \sum_{j=0}^{n} x^j.$$

- $$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \pm \frac{x^{2n+1}}{(2n+1)!} = \sum_{j=0}^{n} \frac{(-1)^j x^{2j+1}}{(2j+1)!}.$$

  Note that this is the $(2n+2)^{\text{nd}}$-degree Taylor polynomial of $sin(x)$ at $x = 0$. The $\pm$ in the formula means that this term is either plus or minus depending on n because the coefficient of x^{2n+2} is zero.

- $$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots \pm \frac{x^{2n}}{(2n)!} = \sum_{j=0}^{n} \frac{(-1)^j x^{2j}}{(2j)!}.$$

  Note that this is the $(2n+1)^{\text{st}}$-degree Taylor polynomial for $\cos(x)$ at $x = 0$ because the x^{2n+1} coefficient is zero.

More examples of Taylor polynomials for well-known functions are given in the exercises. The more of these you have memorized the easier life will be.

Mathematics is made up of two kinds of facts. There are very general theorems applying to large sets of functions and detailed calculations applying to single functions. We should not overlook the detailed theorems as occasionally they come in just as handy as the general theorems. There are a number of relationships between the Taylor polynomials above that can be exploited to great benefit.

For example, we can obtain the Taylor polynomial for $1/(1-x)$ by the method of "synthetic division" you learned long ago. Synthetic division in this case is just long division of $(1-x)$ into 1 as follows:

Synthetic division

So,

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

But we already knew this. The partial sums of the geometric series are given by the formula

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

or

$$1 + x + x^2 + \ldots + x^n = \frac{1}{1-x} - \frac{x^{n+1}}{1-x}.$$

We have come up with the Taylor series for $1/(1-x)$ centered at $x = 0$ in two different ways.

Substitution allows us to generate new power series from known power series. For example, we know that

$$\frac{1}{1-x} \approx 1 + x + x^2 + x^3 + \ldots + x^n.$$

Replacing x with $-x$ in this formula, we obtain

$$\frac{1}{1-(-x)} \approx 1 + (-x) + (-x)^2 + (-x)^3 + \ldots + (-x)^n,$$

or

$$\frac{1}{1+x} \approx 1 - x + x^2 - x^3 + \ldots + (-1)^n x^n.$$

A more complicated application of the same idea can sometimes be used to compute the Taylor polynomial centered at zero for the function $e^{\sin(x)}$. Say we want to compute the 3$^{\text{rd}}$ degree Taylor polynomial centered at $x = 0$ for this function. We first recall that

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}.$$

Replacing x with $\sin(x)$ we have

$$e^{\sin(x)} \approx 1 + \sin(x) + \frac{\sin^2(x)}{2!} + \frac{\sin^3(x)}{3!}.$$

Replacing $\sin(x)$ on the right hand side of this equation with its Taylor polynomial centered at zero, we have

$$e^{\sin(x)} \approx 1 + (x - (x^3/3!)) + \frac{(x-(x^3/3!))^2}{2!} + \frac{(x-(x^3/3!))^3 3!}{.}$$

Now,

$$\left(x - \frac{x^3}{3!}\right)^2 = x^2 - 2\frac{x^4}{3!} + \frac{x^6}{3! \cdot 3!}$$

and only the first term on the right is of degree three or less. Similarly,

$$\left(x - \frac{x^3}{3!}\right)^3 = x^3 + \ldots$$

where ... represents the terms of degree greater than three. Substituting again, we have

$$e^{\sin(x)} \approx 1 + (x - (x^3/3!)) + \frac{x^2 - 2\frac{x^4}{3!} + \frac{x^6}{3!\cdot 3!}}{2!} + \frac{x^3 + \dots}{3!}.$$

Dropping all terms of degree greater than three, we get

$$e^{\sin(x)} \approx 1 + x - \frac{x^3}{3!} + \frac{x^2}{2!} + \frac{x^3}{3!}$$

or

$$e^{\sin(x)} \approx 1 + x + \frac{x^2}{2!} + 0 \cdot x^3$$

is the third degree Taylor polynomial centered at $x = 0$. We computed this polynomial without computing any derivatives. We are able to compute the Taylor polynomial of $e^{\sin(x)}$ using substitution in this way because there is no constant term in the Taylor polynomial of $\sin(x)$ centered at $x = 0$. Using this idea on $e^{\cos(x)}$ is more messy (see Exercises to ponder).

4.6 Examples:

We should think of Taylor polynomials as a natural extension of the idea of tangent lines. Taylor polynomials give a way of approximating complicated functions with polynomials. Hopefully, the polynomials are easier to compute and work with than the original function, but we must always remember that the Taylor polynomial is only an approximation.

As an example, consider the following curious situation. You have been using the function e^x for several years already. You can manipulate it using algebra and tools from Calculus. Yet, answering the question, "What is the value of e?" is not so easy. We remember that $e \approx 2.71828\dots$, but where did this come from and what is the next digit? Taylor polynomials provide a way to compute e as accurately as anyone might like as follows: We know that the Taylor polynomial for e^x centered at $x = 0$ is

$$e^x \approx 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}.$$

So, using $x = 1$ we have

$$e = e^1 = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}.$$

We can use Taylor's theorem to determine the accuracy of this approximation of e, if we can come up with a bound on the $(n+1)^{\text{st}}$ derivative of e^x. We already know that

$$\frac{d^{n+1}e^x}{dx^n} = e^x,$$

So for $0 \le x \le 1$ we have that

$$\left|\frac{d^{n+1}e^x}{dx^n}\right| \le e \quad \text{for } 0 \le x \le 1.$$

Using Taylor's inequality we now have

$$\left|e - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!}\right)\right| \leq \frac{e}{(n+1)!}.$$

This is starting to look circular because e appears on both sides of the equation. We can save ourselves by noting that this is only an inequality. We do not need the exact value of e for the right hand side. Any number greater than e gives us a bound on the size of the expression on the left.

We can show (without any prior knowledge of the value of e) that $e < 4$ (how we know this depends on the definition of e^x and its relationship to $\ln x$, see the Exercises). Using this we have

$$\left|e - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!}\right)\right| \leq \frac{4}{(n+1)!}.$$

So, for example, the difference between e and the expression

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!}$$

is less than

$$\frac{4}{6!} = \frac{4}{720} = \frac{1}{180}.$$

Since this is less than $1/100$, the Taylor polynomial at 0 of degree five evaluated at $x = 1$ gives an approximation of e that is accurate to at least two decimal places. To get a more accurate approximation we simply take more terms. We know that $1/n!$ tends to zero very quickly, so the accuracy of the approximation grows very quickly with n.

As a second example, consider the motion of an object thrown straight up. Let x denote the height of the object above the ground, so x will be a function of time t and the usual equation we use is

$$\frac{d^2x}{dt^2} = -g$$

where g is the acceleration of gravity near the earth's surface (approximately 9.8 meters per second^2).

This works well for everyday things like baseballs, but rockets and spacecraft can reach much higher altitudes. We know that the force of gravity decreases as the distance from the object grows. The formula above is really just the 0^{th} order approximation of the more accurate formula given by

$$\frac{d^2x}{dt^2} = -\frac{GM}{(R_0 + x)^2}$$

where M is the mass of the earth, G is the universal gravitation constant and R_0 is the radius of the earth. (This is still just an approximation because it ignores the effect of the atmosphere, the fact that the earth is not perfectly spherical, etc.) Let

$$f(x) = \frac{GM}{(R_0 + x)^2}.$$

We compute that

$$f'(x) = -\frac{2GM}{(R_0 + x)^3},$$

$$f''(x) = \frac{6GM}{(R_0 + x)^4},$$

$$\vdots$$

$$\frac{d^n f}{dx^n} = \pm\frac{(n+1)!GM}{(R_0 + x)^{n+2}}.$$

Hence, the n^{th}-degree Taylor polynomial for f centered at $x = 0$ is

$$f(x) \approx P_n(x) = \frac{GM}{R_0^2} - \frac{2GM}{R_0^3}x + \frac{6GM}{R_0^4}x^2 - \ldots \pm \frac{(n+1)!GM}{(R_0)^{n+2}}x^n.$$

The 0^{th} order approximation of f gives

$$f(x) \approx -\frac{GM}{R_0^2}$$

and letting $g = GM/R_0^2$ we get the simple formula above. The error between f and its zero-th order approximation is "order x", that is, it is bounded by a constant times $|x|$.

The 1^{st} order Taylor polynomial gives the formula

$$f(x) \approx -g + \frac{2GM}{R_0^3}x.$$

Replacing $2GM/R_0^3$ with $2g/R_0$ we have

$$f(x) \approx -g + \frac{2gx}{R_0} = -g\left(1 - \frac{2x}{R_0}\right).$$

Once x is a significant fraction of R_0, this approximation will be much more accurate. The second order approximation is

$$f(x) \approx -g + \frac{2GMx}{R_0^3} - \frac{6GM}{R_0^4} = -g\left(1 - \frac{2x}{R_0} + \frac{6x^2}{R_0^2}\right)$$

and the second degree term becomes significant when $(x/R_0)^2$ is significant. The error term is "order x^3"

The word "significant" is a relative term. How large x/R_0 must be before the first or second order terms are significant depends on how much error we are willing to put up with in predictions about our body. We can estimate the error in each approximation using Taylor's inequality, but the decision of how many terms to include in the Taylor expansion depends on the situation. Each extra degree term we add to the Taylor Polynomial improves the approximation with the actual function. The error is like a constant times the next higher power of x. But be warned–the constant in the error estimate also changes each time we add a term.

4.7 Taylor Series and Power Series:

Given a function $f(x)$ which can be differentiated as many times as we like, we can approximate it near a point $x = x_0$ as closely as we like by computing the n^{th} degree Taylor polynomial with large n. What if we want an "infinitely good" approximation. Since improving the degree or approximation involves increasing the degree of the Taylor polynomial, an infinitely good approximation should require an "infinite degree" polynomial. This motivates the following

Definition: Given a function $f(x)$ which is infinitely differentiable (that is, derivatives of all orders exist) and a point x_0, the **Taylor series** for f centered at x_0 is

$$f(x_0) + f'x_0(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!} \left.\frac{d^n f}{dx^n}\right|_{x_0} (x - x_0)^n,$$

where we use the convention that $d^0 f/dx^0 = f(x)$.

A Taylor series is just a Taylor polynomial where we forget to stop adding terms. It is an example of a **power series**, a series of the form

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots = \sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

The numbers $a_0, a_1, a_2, \ldots$ are called the **coefficients**.

What we are doing is using infinite series to define a function of the variable x. Given a value of x, we plug it into the summation and get an infinite series. If the series converges, the limit of the partial sums is the value of the function. If the series diverges then x is not in the domain of the function.

Definition: We say a function $f(x)$ is **analytic** at x_0 if there is an $\epsilon > 0$ such that for all x with

$$|x - x_0| < \epsilon$$

the Taylor series

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left.\frac{d^n f}{dx^n}\right|_{x_0} (x - x_0)^n$$

converges and

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{d^n f}{dx^n} \right|_{x_0} (x - x_0)^n.$$

That is, a function is analytic at a point x_0 if its Taylor series converges to $f(x)$ for x near x_0. Luckily, most of the functions we are used to dealing with in Calculus, like e^x, sine, cosine, tangent, etc. are analytic where ever they are defined.

4.8 A Curious and Useful Observation

Recall (as you should be able to do for the rest of your life), the Taylor series of e^x is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots.$$

This series converges absolutely for all real numbers x. It turns out that it also converges absolutely when we plug in other types of things for x.

You have probably encountered the number $i = \sqrt{-1}$. Adding this "number" to our set of real numbers is a very useful. It allows us to solve many more equations, for example

$$\begin{aligned} x^2 + 9 &= 0 \\ x^2 &= -9 \\ x &= \sqrt{-9} \\ x &= 3\sqrt{-1} = 3i. \end{aligned}$$

Arithmetic and Calculus with complex numbers, perhaps surprisingly, turns out to be remarkably useful in fields from physics to electrical engineering.

If we replace x in the Taylor series for e^x with ix, we obtain

$$e^{ix} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} = 1 + ix + \frac{(ix^2)}{2!} + \frac{(ix)^3}{3!} + \ldots.$$

Using the facts that $i^2 = -1$, $i^3 = i \cdot i^2 = -i$, $i^4 = i^2 \cdot i^2 = -1 \cdot -1 = 1$ and so on, we get that

$$e^{ix} = 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \ldots.$$

It turns out that the commutative law holds for absolutely convergent series. That is, we can rearrange the terms of an absolutely convergent series however we like without altering the sum. (The situation for a conditionally convergent series is much more subtle, see Exercises.) If we collect together all the terms containing i and all those without i we have

$$e^{ix} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots\right).$$

You should now have that "tingly" feeling down your spine that happens when you see something really weird. The two series on the right are familiar. They are just the Taylor series for sine and cosine centered at $x = 0$. That is, we have that

$$e^{ix} = \cos(x) + i\sin(x).$$

This is called **Euler's formula**. It is one of the great surprises of mathematics. It tells us that the function e^x, discovered as the inverse of the natural log function, is related to the trigonometric functions sine and cosine. If we plug in $x = \pi$ to this formula we get a particularly striking formula

$$e^{i\pi} = -1.$$

All of the weird numbers you have learned from elementary school to high school, $-1, \pi, e$ and $i = \sqrt{-1}$, are related to each other.

This formula is much more than just a curiosity. As a first example, we use the rules of manipulating exponents to derive some standard (but hard to remember) trigonometric identities. Since the rules of arithmetic of complex numbers are the same as for real numbers, we know that

$$e^{i(\theta+\phi)} = e^{i\theta+i\phi} = e^{i\theta}e^{i\phi}.$$

Using Euler's formula on both ends of this equation we have

$$\begin{aligned}
\cos(\theta+\phi) &+ i\sin(\theta+\phi) = (\cos(\theta) + i\sin(\theta))(\cos(\phi) + i\sin(\phi)) \\
&= \cos(\theta)(\cos(\phi) + i\sin(\phi)) + i\sin(\theta)(\cos(\phi) + i\sin(\phi)) \\
&= \cos(\theta)\cos(\phi) + i\cos(\theta)\sin(\phi) + i\sin(\theta)\cos(\phi) + i^2\sin(\theta)\sin(\phi) \\
&= \cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi) + i(\cos(\theta)\sin(\phi) + \sin(\theta)\cos(\phi)).
\end{aligned}$$

Two complex numbers $a + ib$ and $c + id$ are equal if and only if $a = c$ and $b = d$, hence the equation above implies the two equations

$$\begin{aligned}
\cos(\theta+\phi) &= \cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi) \\
\sin(\theta+\phi) &= \cos(\theta)\sin(\phi) + \sin(\theta)\cos(\phi).
\end{aligned}$$

That is, the usual angle sum formulas for sine and cosine are really consequences of the rules for arithmetic with exponents.

Exercises:

1. Compute the Taylor polynomial of each of the following as indicated

   (a)
   $$f(x) = 3 + 3x^2 - 5x^3 \text{ centered at } x = 0 \text{ to degree } 4$$

(b)

$$f(x) = 3 + 3x^2 - 5x^3 \text{ centered at } x = 1 \text{ to degree } 4$$

(c)

$$f(x) = \sin(x) \text{ centered at } x = \pi/4 \text{ to degree } 4$$

(d)

$$f(x) = \cos(x) \text{ centered at } x = \pi/4 \text{ to degree } 4$$

(e)

$$f(x) = \ln(x) \text{ centered at } x = 1 \text{ to degree } 4$$

(f)

$$f(x) = \tan(x) \text{ centered at } x = 0 \text{ to degree } 4$$

(g)

$$f(x) = \sec(x) \text{ centered at } x = 0 \text{ to degree } 4$$

(h)

$$f(x) = \arcsin(x) \text{ centered at } x = 0 \text{ to degree } 4$$

(i)

$$f(x) = \arccos(x) \text{ centered at } x = 0 \text{ to degree } 4$$

2. For each of the following, find the indicated Taylor polynomial without computing any derivatives. (Use substitution into a Taylor polynomial you already know, synthetic division, composition, ...)

(a)

$$f(x) = e^{(1+x)} \text{ centered at } x = 0 \text{ to degree } 4$$

(b)

$$f(x) = \frac{1}{1+x^2} \text{ centered at } x = 0 \text{ to degree } 4$$

(c)

$$f(x) = \ln(1-x) \text{ centered at } x = 0 \text{ to degree } 4$$

(d)

$$f(x) = \cos^3(x^2) \text{ centered at } x = 0 \text{ to degree } 4$$

(e)

$$f(x) = \sec(x) \text{ centered at } x = \pi/4 \text{ to degree } 4$$

(f)

$$f(x) = \frac{1}{1+2x-x^2} \text{ centered at } x = 0 \text{ to degree } 4$$

(g)

$$f(x) = \frac{1}{x^2} \text{ centered at } x = 1 \text{ to degree } 4$$

(Hint: First substitute $z = x - 1$.)

3. How large a degree Taylor polynomial must you use in order to accurately compute $\sin(.5)$ to 8 decimal places?

4. How large a degree Taylor polynomial must you use in order to accurately compute e^2 to 8 decimal places?

5. How large a degree Taylor polynomial must you use in order to accurately compute e^{-2} to 8 decimal places? (Does this change your answer to the previous problem?)

6. Suppose we know that $0 \leq g(x) \leq f(x)$ for all x.

   (a) Verify that

   $$\int_a^b g(x)\,dx \leq \int_a^b f(x)\,dx$$

   for all $a < b$.

   (b) Can you say

   $$g'(x) \leq f'(x)?$$

   Why or why not?

7. Find an expression for the n^{th} term of the Taylor series for $\tan(x)$ centered at $x = 0$ and memorize it.

8. Find an expression for the n^{th} term of the Taylor series for $\sec(x)$ centered at $x = 0$ and memorize it.

9. Find an expression for the n^{th} term of the Taylor series for $\ln(x)$ centered at $x = 1$ and memorize it.

10. Draw the graphs of the 3^{rd}, 5^{th}, 7^{th} and 9^{th} degree Taylor polynomials for $\sin(x)$ centered at $x = 0$. Looking at the polynomials and the graphs, what would graph of the 33^{rd} degree Taylor polynomial of $\sin(x)$ centered at $x = 0$?

11. Use Euler's formula to derive trigonometric identities from the following formulas

$$e^{(i\theta)^2} = e^{2i\theta}$$

Exercises to Ponder

1. (a) Show that

$$\sum_{n=1}^{\infty} \frac{1}{2n}$$

and

$$\sum_{n=0}^{\infty} \frac{1}{2n+1}$$

both diverge.

(b) What does this say about applying the commutative law to infinite series? For example, can we rearrange the terms of

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

arbitrarily and obtain the same sum?

2. We know that the radius of convergence of the geometric series

$$\sum_{n=0}^{\infty} x^n$$

must be less than or equal to 1 because

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

which tends to infinity as x tends to 1. Can you use the same idea to put upper bounds on the radius of convergence of the Taylor series centered at zero for $1/(1+x^2)$.

3. Compare the Taylor series for $1/(1-x)$ and $\ln(1-x)$ centered at $x = 0$. How could you guess one if you knew the other?

4. What is the Taylor series centered at $x = 0$ for the function

$$f(x) = \left(\begin{array}{ll} e^{-1/x^2} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{array} \right.$$

(Hint: If you assume the derivatives exist, then computing the derivatives of $f(x)$ at $x = 0$ is easy, just take limits from the left. Verifying that the derivatives are correct is hard.) What is strange about this Taylor series and this function?

5. Compute the Taylor polynomial of $e^{\cos x}$ centered at $x = 0$ using what you know about e^x and $\cos x$.

6. Recall the "definition" of e via natural logs

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

Check that

$$\int_1^4 \frac{1}{x}\,dx > 1$$

by estimating the area under the graph. Show how this implies $e < 4$.