Taylor's Theorem and Applications

By James S. Cook, November 11, 2018, For Math 132 Online.

I wrote this for Math 131 notes when I taught Calculus I at Liberty University around 2009-2013. In short, I see this topic as a natural extension of the discussion of tangent lines and tangent line approximation of functions in first semester calculus. However, most curricula include this in Calculus II. Of course, we don't stop at polynomials in Calculus II. We consider the Taylor series for a function which is formed from the limit of the sequence of Taylor polynomials. That said, we will focus on Taylor Polynomials and the error in replacing $f(x)$ with its k-th Taylor polynomial centered at some point. Naturally, in the context of Calculus II this amounts to studying the error in truncating the Taylor series with a mere polynomial. Let's get to it:

0.1 Taylor's Theorem about polynomial approximation

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

0.1.1 constant functions

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

0.1.2 linearizations again

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

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

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

0.1.3 quadratic approximation of function

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

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

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

0.1.4 cubic approximation of function

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

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

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

0.1.5 general case

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

Definition 0.1.1. *Taylor polynomials.*

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

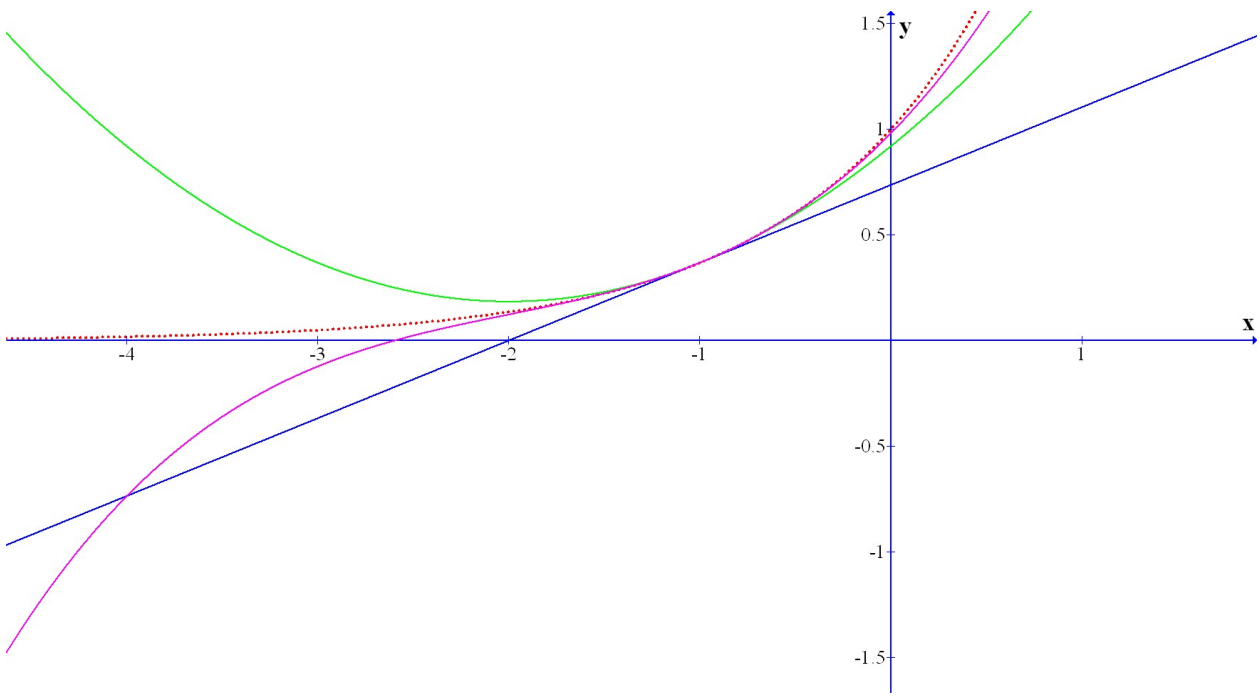
$$T_k(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{k!}f^{(k)}(a)(x-a)^k$$

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

Example 0.1.2. *Suppose $f(x) = e^x$. Calculate the first few Taylor polynomials centered at $a = -1$. Derivatives of the exponential are easy enough to calculate; $f'(x) = f''(x) = f'''(x) = e^x$ therefore we find*

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

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



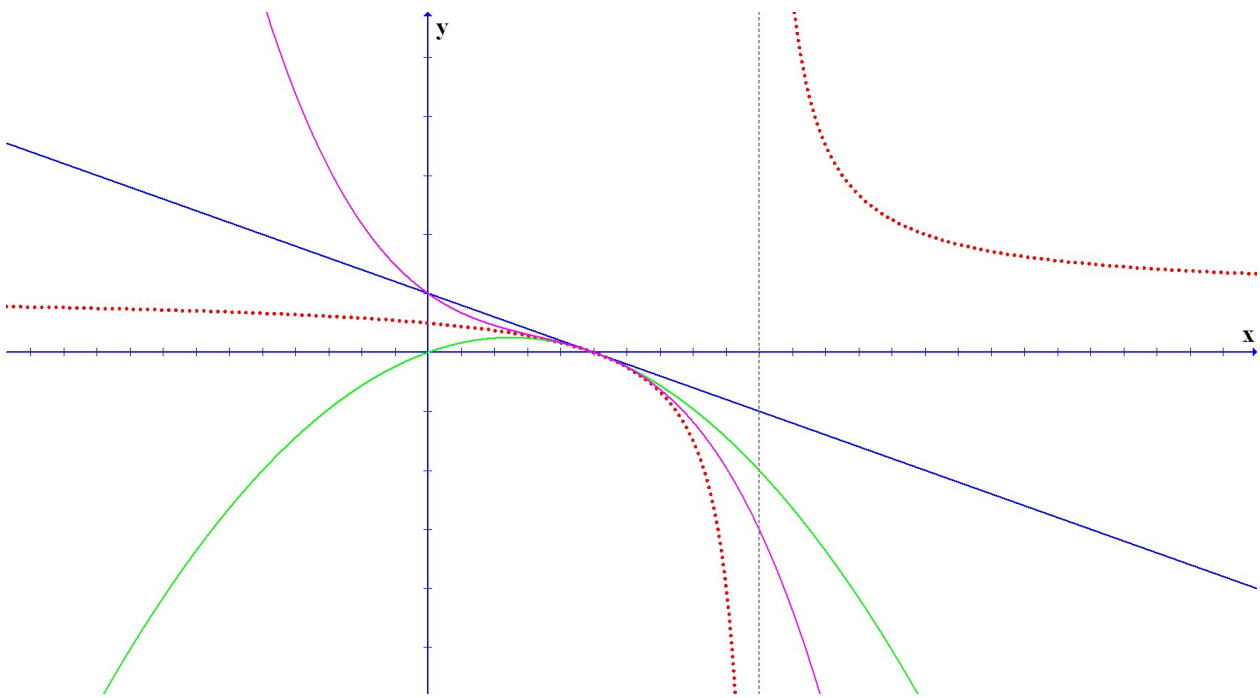
Example 0.1.3. *Suppose* $f(x) = \frac{1}{x-2} + 1$. *Calculate the first few Taylor polynomials centered at* $a = 1$. *Observe*

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

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

$$\begin{aligned} T_1(x) &= -(x-1) \\ T_2(x) &= -(x-1) + (x-1)^2 \\ T_3(x) &= -(x-1) + (x-1)^2 - (x-1)^3 \end{aligned}$$

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

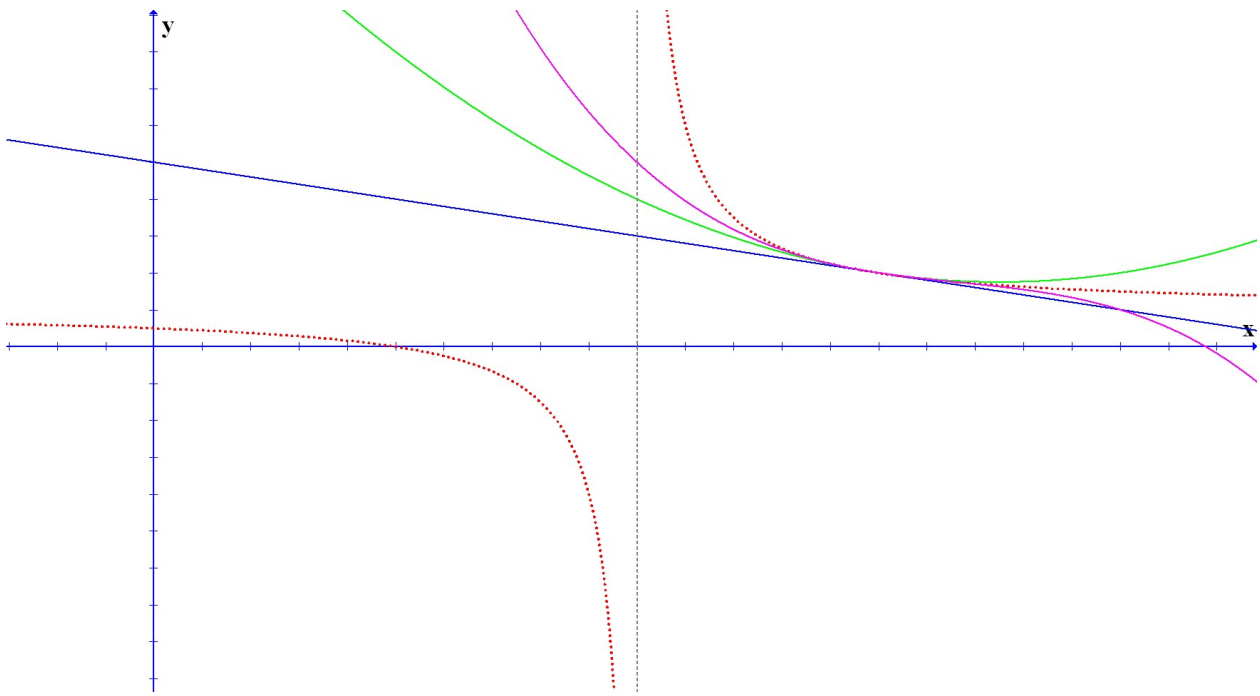


On the next page we examine the same function approximated at a different center point. In other words, for a given $f(x)$ *we can consider its Taylor polynomial approximation at different points. Generally, these approximations differ (can you think of a case where they would be the same ?)*

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

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

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



Example 0.1.4. *Let* $f(x) = \sin(x)$. *It follows that*

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

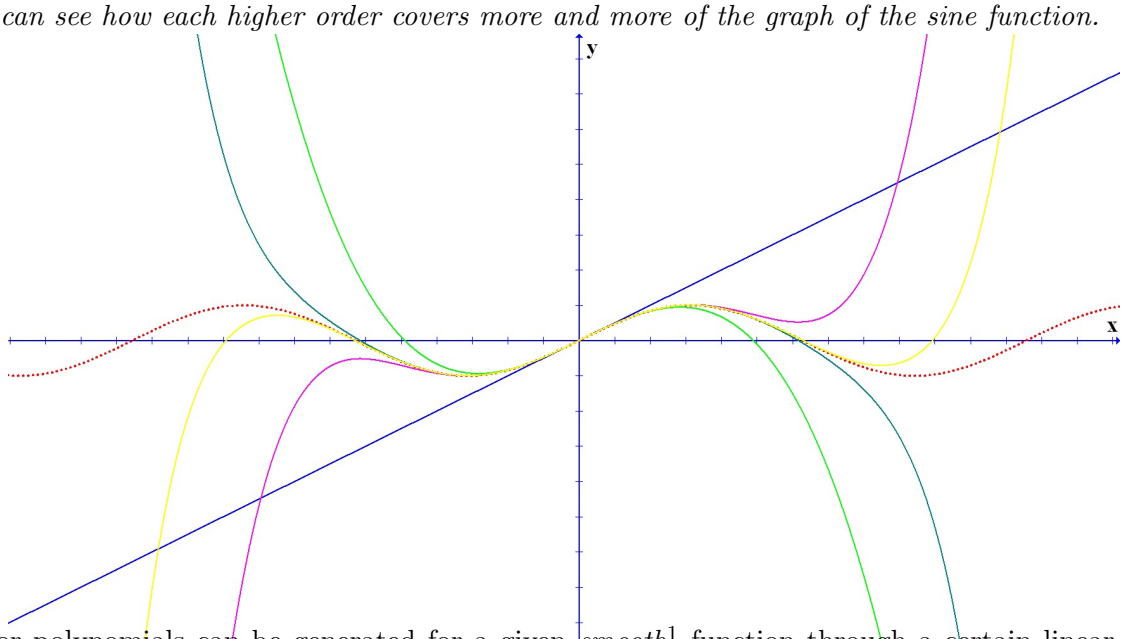
Hence, $f(0) = 0, f'(0) = 1, f''(0) = 0, f'''(0) = -1, f^{(4)}(0) = 0, f^{(5)}(0) = 1$. *Therefore the Taylor polynomials of orders* $1, 3, 5$ *are*

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

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

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

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



Taylor polynomials can be generated for a given *smooth*[1] function through a certain linear combination of its derivatives. The idea is that we can approximate a function by a polynomial[2], at least locally. We discussed the tangent line approximation to a function. We found that the linearization of a function gives a good approximation for points close to the point of tangency. If we calculate second derivatives we can similarly find a quadratic approximation for the function. Third derivatives go to finding a cubic approximation about some point. I should emphasize from the outset that a Taylor polynomial is just a polynomial, it will not be able to exactly represent a function which is not a polynomial. In order to exactly represent an analytic function we'll need to take infinitely many terms, we'll need a power series. Finally, let me show you an example of how Taylor polynomials can be of fundamental importance in physics.

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

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

Example 0.1.5. *The relativistic energy E of a free particle of rest mass m_o is a function of its velocity v:*

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

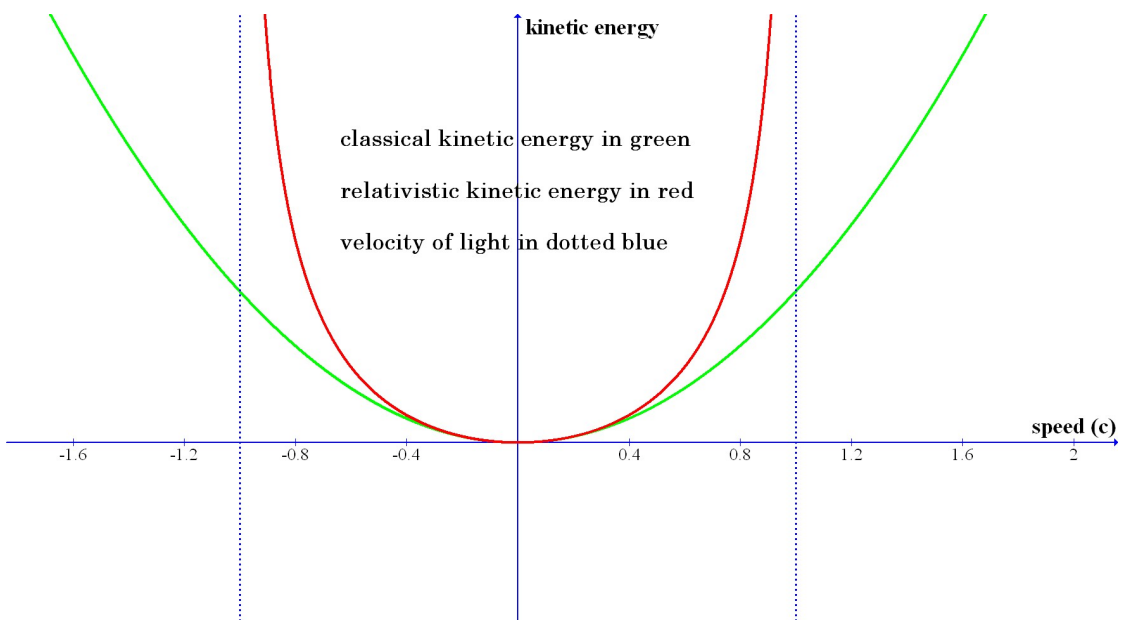
for $-c < v < c$ where c is the speed of light in the space. We calculate,

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

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

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

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



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

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

0.1.6 error in Taylor approximations

We've seen a few examples of how Taylor's polynomials will locally mimic a function. Now we turn to the question of extrema. Think about this, if a function is locally modeled by a Taylor polynomial centered at a critical point then what does that say about the nature of a critical point? To be precise we need to know some measure of how far off a given Taylor polynomial is from the function. This is what Taylor's theorem tells us. There are many different formulations of Taylor's theorem[3], the one below is partially due to Lagrange.

Theorem 0.1.6. *Taylor's theorem with Lagrange's form of the remainder.*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Corollary 0.1.7. *error bound for $T_k(x)$.*

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

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

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

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

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

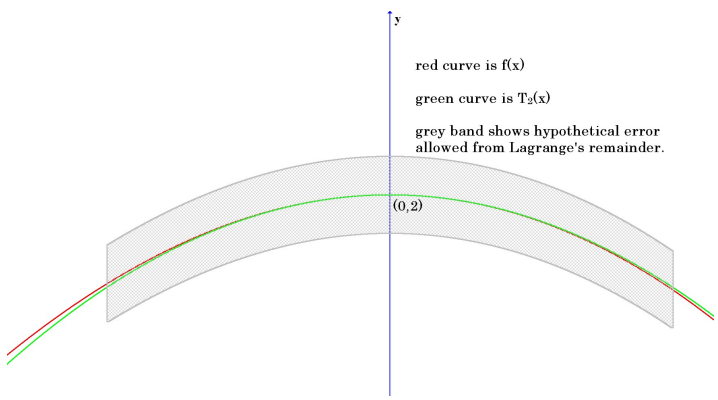
$$|f(x) - f(c) - \frac{1}{2}f''(c)(x-c)^2| = |\frac{1}{6}f'''(\zeta_x)(x-c)^3| \leq \frac{4M\delta^3}{3}$$

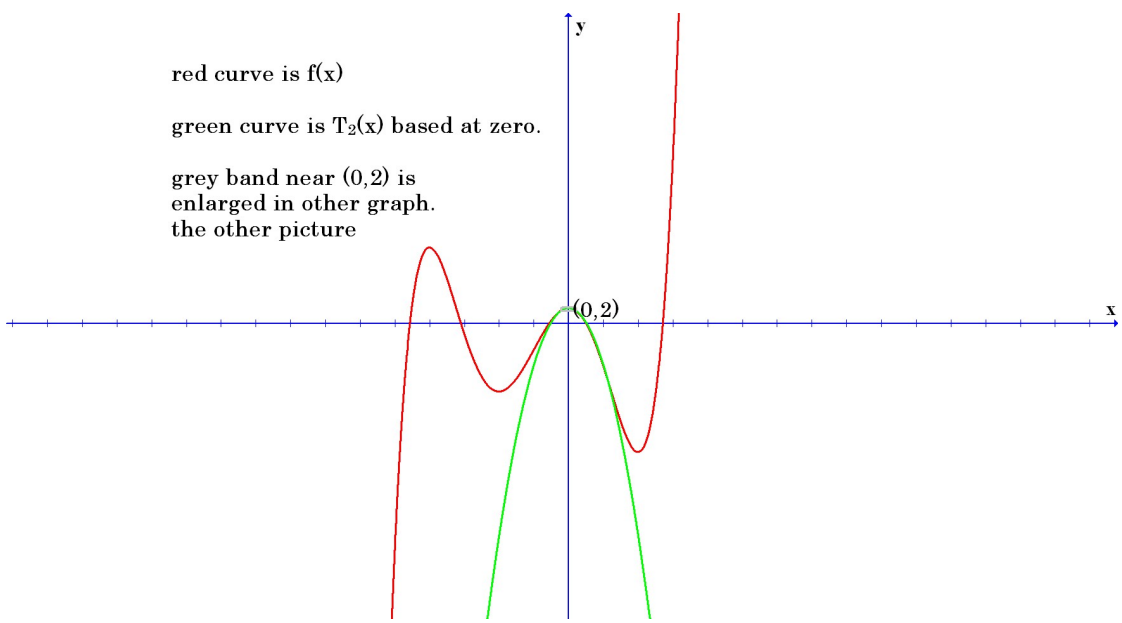
for all $x \in [c-\delta, c+\delta]$. This inequality reveals that we have $f(x) \approx f(c) + \frac{1}{2}f''(c)(x-c)^2$ as $\delta \to 0$. Therefore, locally the graph of the function resembles a parabola which either opens up or down at the critical point. If it opens up ($f''(c) > 0$) then $f(c)$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

Example 0.1.8. *Suppose* $f(x) = 6x^5 + 15x^4 - 10x^3 - 30x^2 + 2$. *We can calculate that* $f'(x) = 30x^4 + 60x^3 - 30x^2 - 60x$ *therefore clearly* $(0, 2)$ *is a critical point of* f. *Moreover,* $f''(x) = 120x^3 + 180x^2 - 60x - 60$ *shows* $f''(0) = -60$. *I aim to show how the quadratic Taylor polynomial* $T_2(x) = f(2) + f'(2)x + \frac{1}{2}f''(2)x^2 = 2 - 30x^2$ *gives a good approximation for* $f(x)$ *in the sense that the maximum error is essentially bounded by the size of Lagrange's term. Note that*

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

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





Example 0.1.9. *Here we examine Taylor polynomials for* $f(x) = \sin(x)$ *on the interval* $(-1, 1)$ *and second on* $(-2, 2)$*. In each case we use sufficiently many terms to guarantee an error of less than* $\epsilon = 0.1$*. Notice that* $f^{(2k-1)}(x) = \pm\sin(x)$ *whereas* $f^{(2k-2)}(x) = \pm\cos(x)$ *for all* $k \in \mathbb{N}$ *therefore* $|f^{(n)}(x)| \leq 1$ *for all* $x \in \mathbb{R}$*.*

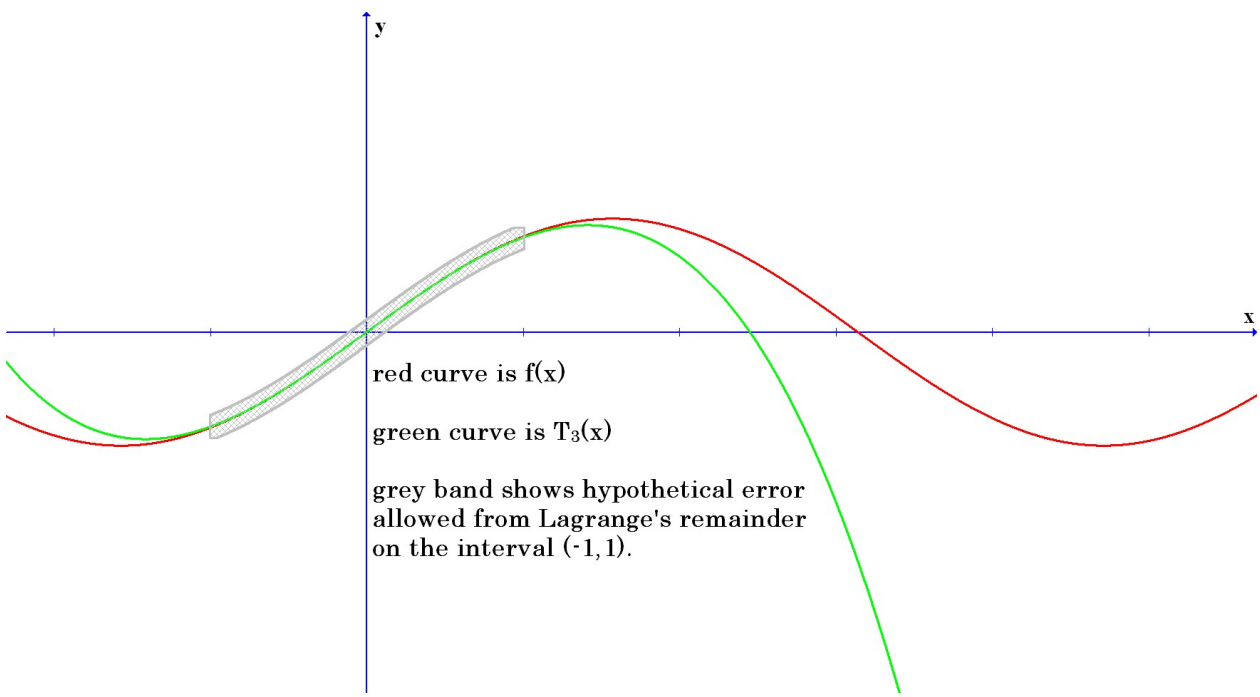
If we wish to bound the error to 0.1 *on* $-1 < x < 1$ *then we to bound the remainder term as follows: (note* $-1 < x < 1$ *implies* $l = 2$ *and we just argued* $M = 1$ *is a good bound for any* k*)*

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

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

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

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



On the next page I determine what is needed to mimic the sine function on the larger interval $-2 < x < 2$*,*

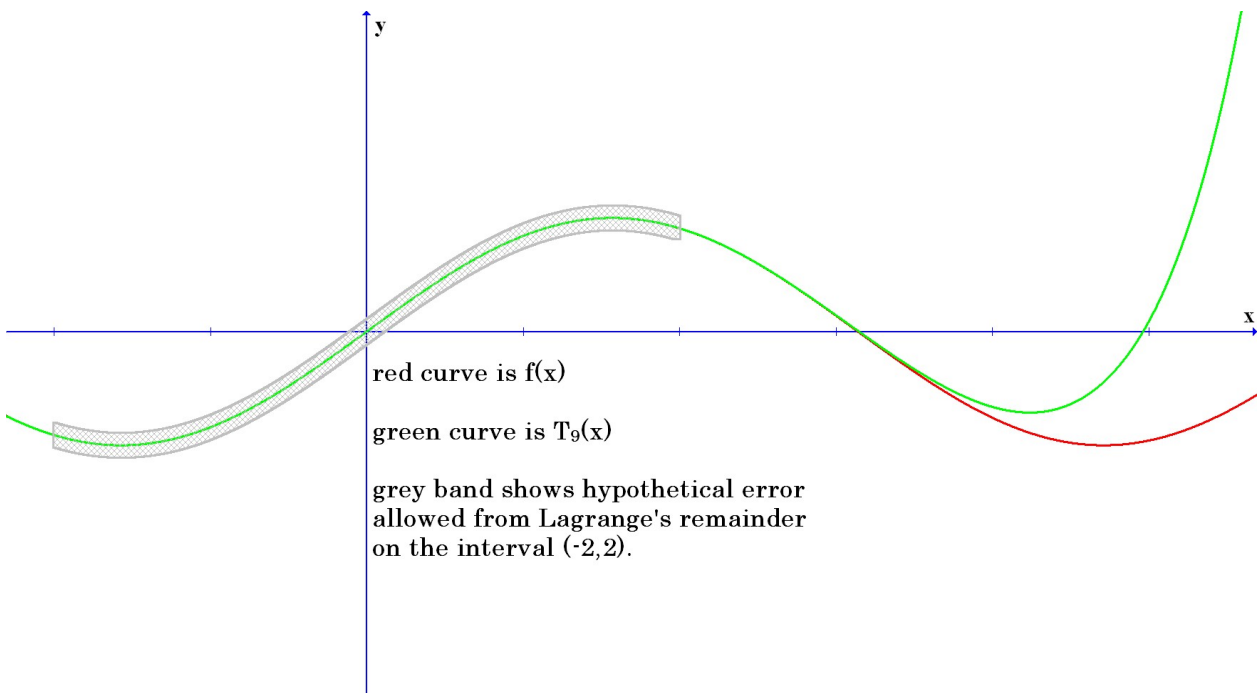
If we wish to bound the error to 0.1 *on* $-2 < x < 2$ *then we to bound the remainder term as follows: (note* $-2 < x < 2$ *implies* $l = 4$ *)*

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

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

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

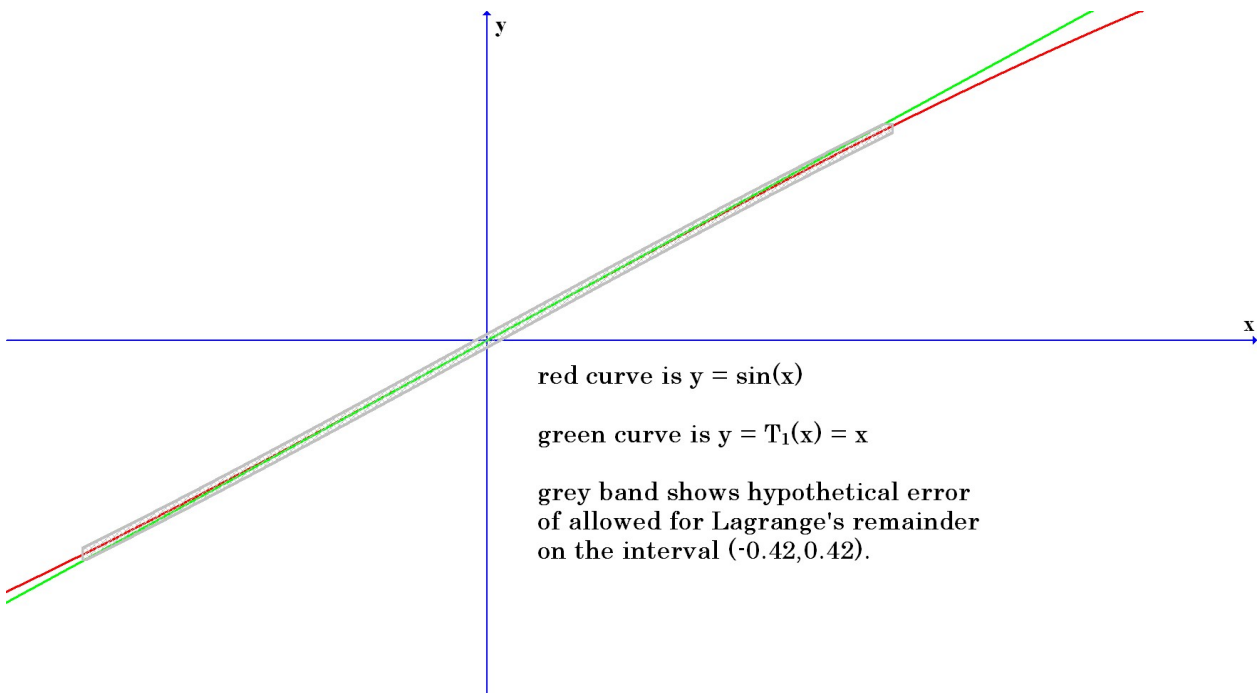
This shows that $T_10(x)$ *will provide the desired accuracy. But, it just so happens that* $T_9 = T_{10}$ *in this case so we find* $T_9(x) = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \frac{1}{362880}x^9$ *will suffice. In fact, as you can see below it fits the* ± 0.1 *tolerance band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



Example 0.1.10. *Let's think about* $f(x) = \sin(x)$ *again. This time, answer the following question: for what domain* $-\delta < x < \delta$ *will* $f(x) \approx x$ *to within* ± 0.01 *? We can use* $M = 1$ *and* $l = 2\delta$*. Furthermore,* $T_1(x) = T_2(x) = x$ *therefore we want*

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

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



Example 0.1.11. *Suppose we are faced with the task of calculating $\sqrt{4.03}$ to an accuracy of 5-decimals. For the purposes of this example assume all calculators are evil. It's after the robot holocaust so they can't be trusted. What to do? We use the Taylor polynomial up to quadratic order: we have $f(x) = \sqrt{x}$ and $f'(x) = \frac{1}{2\sqrt{x}}$ and $f''(x) = \frac{-1}{4(\sqrt{x})^3}$. Apply Taylor's theorem,*

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

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

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

Therefore, $\boxed{\sqrt{4.03} = 2.007486 \pm 0.000001}$. As far as I know my TI-89 is still benevolent so we can check our answer; the calculator says $\sqrt{4.03} = 2.\underline{00748}598999$.

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

Example 0.1.12. *Newton postulated that the gravitational force between masses m and M separated by a distance of r is*

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

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

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

We calculate,

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

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

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

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

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

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

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

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