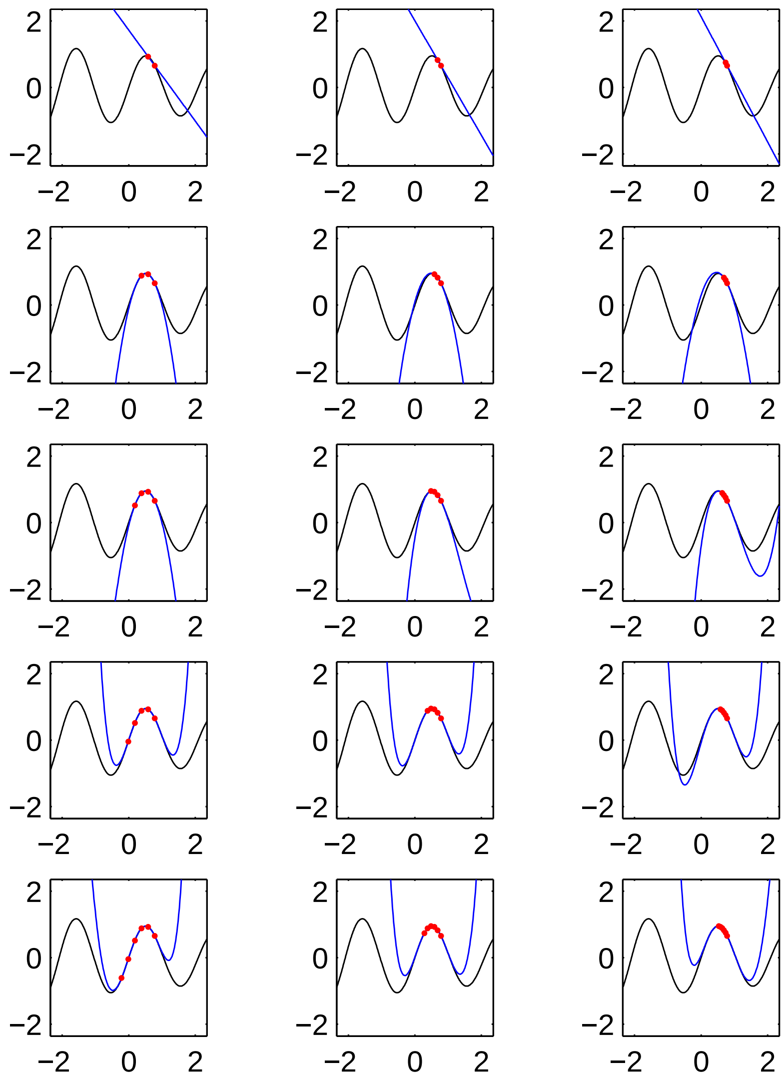

Figure 4: Polynomial approximation of $y = e^{-x/10}\sin(3\,x)$ at $x_0 = \pi/4$

Polynomial Approximation

In the previous handout we reviewed the tangent line approximation and investigated its error. This led us to some conjectures about the quadratic approximation. Here we will formalize these conjectures and derive general formulas for polynomial approximation of any order. This is known as Taylor's theory.

Tangent line approximation

It will be helpful to review yet again the tangent line approximation. Let $T_1 = a_0 + a_1\,x$ be the linear (tangent line) approximation of $y = f(x)$ at a given point x_0. In order to find the coefficients a_0 and a_1 we reason as follows.

Firstly, we would like the graph of T_1 to pass through the point $(x_0, f(x_0))$:

$$a_0 + a_1\,x_0 = f(x_0).$$

This gives us one equation involving two unknowns. To get another equation, we pick an *auxiliary point* on the graph of $y = f(x)$ some distance h to the left from the *base point* $(x_0, f(x_0))$. With two points at our disposal, we can construct the *secant* line by solving:

$$\begin{aligned} a_0 + a_1\,x_0 &= f(x_0) \\ a_0 + a_1\,(x_0 - h) &= f(x_0 - h) \end{aligned}$$

Then elimination of variables leads to the familiar equation of the secant:

$$\widetilde{T}_1 = f(x_0) + \frac{f(x_0) - f(x_0 - h)}{h}\,(x - x_0).$$

Notice that we have not really done any Calculus yet—the construction of the secant is a purely algebraic process. However, we do need Calculus to pass to the equation of the tangent. To this end we let the auxiliary point get closer and closer to the base point: that is, we define the tangent T_1 as the limit of the secant $\widetilde{T}_1$ as h goes to zero:

$$T_1 = \lim_{h\to 0} \widetilde{T}_1 = f(x_0) + \left(\lim_{h\to 0} \frac{f(x_0) - f(x_0 - h)}{h}\right)\,(x - x_0) \qquad (17)$$

Of course, the limit inside the parentheses is nothing but the derivative of f evaluated at x_0. Denoting the latter with the customary prime, we can rewrite Equation (17) as:

$$T_1(x) = f(x_0) + f'(x_0)\,(x - x_0). \tag{18}$$

Let us take a pause and reflect before moving on. It may seem that Equations (17) and (18) say one and the same thing using different notation. However, these equations are conceptually very different. Equation (17) encodes the *process* of construction of the tangent which is illustrated in the top row of Figure 4—it is this equation that in Calculus I motivated, or, perhaps, should have motivated the introduction of the derivative as the limit of the difference quotient:

$$\lim_{h\to 0}\left(\frac{f(x_0) - f(x_0 - h)}{h}\right) = f'(x_0).$$

Equation (18), on the other hand, is a Calculus "black box": it just tells you how to get the equation of the tangent quickly by using various differentiation rules. The "prime" notation completely obscures the analytic process. Indeed, there is nothing in Equation (18) that suggests that an auxiliary point is involved. The moral is:

> Use Equation (18) for rote computations but think of tangent lines and the first derivative in terms of Equation (17).

Quadratic approximations

For most functions the tangent line approximation works well only in a very small neighborhood of the base point. For instance, the damped harmonic in Figure 4 looks much more like a parabola near $x_0 = \pi/4$ than a straight line. This suggests approximating $y = f(x)$ with a quadratic $T_2 = a_0 + a_1\,x + a_2\,x^2$. As the second row in Figure 4 shows, the quadratic approximation is clearly superior to the linear one.

The process for finding the coefficients of T_2 is exactly the same as in the case of T_1. The difference is that instead of one auxiliary point we must pick two auxiliary points. In Figure 4 the auxiliary points are taken h and $2\,h$ units to the left from the base point. Their coordinates are therefore

$(x_0-h, f(x_0-h))$ and $(x_0-2\,h, f(x_0-2\,h))$. This choice of auxiliary points leads to the following three equations for determining a_0, a_1, and a_2:

$$\begin{aligned}
a_0 + a_1\,x_0 + a_2\,x_0^2 &= f(x_0)\\
a_0 + a_1\,(x_0-h) + a_2\,(x_0-h)^2 &= f(x_0-h)\\
a_0 + a_1\,(x_0-2\,h) + a_2\,(x_0-2\,h)^2 &= f(x_0-2\,h)
\end{aligned}$$

Straightforward, albeit tedious, elimination of variables leads to:

$$\begin{aligned}
a_0 &= f(x_0) - x_0\,\frac{3\,f(x_0) - 4\,f(x_0-h) + f(x_0-2\,h)}{2\,h} + x_0^2\,\frac{f(x_0) - 2\,f(x_0-h) + f(x_0-2\,h)}{2\,h^2}\\
a_1 &= \frac{3\,f(x_0) - 4\,f(x_0-h) + f(x_0-2\,h)}{2\,h} - x_0\,\frac{f(x_0) - 2\,f(x_0-h) + f(x_0-2\,h)}{h^2}\\
a_2 &= \frac{f(x_0) - 2\,f(x_0-h) + f(x_0-2\,h)}{2\,h^2}
\end{aligned}$$

The 'secant parabola', which is properly called the *interpolating* quadratic, is therefore given by:

$$\begin{aligned}
\widetilde{T}_2 = f(x_0) &+ \frac{3\,f(x_0) - 4\,f(x_0-h) + f(x_0-2\,h)}{2\,h}\,(x-x_0)\\
&+ \frac{f(x_0) - 2\,f(x_0-h) + f(x_0-2\,h)}{2\,h^2}\,(x-x_0)^2.
\end{aligned} \tag{19}$$

Reasoning as before, we define T_2, the Taylor polynomial of order two, as it is called, to be the limit:

$$T_2 = \lim_{h\to 0} \widetilde{T}_2.$$

Evidently, in order to find T_2, we must compute:

$$\lim_{h\to 0} \frac{3\,f(x_0) - 4\,f(x_0-h) + f(x_0-2\,h)}{2\,h}$$

and

$$\lim_{h\to 0} \frac{f(x_0) - 2\,f(x_0-h) + f(x_0-2\,h)}{2\,h^2}$$

This prompts the first interlude.

Limits of difference quotients

In principle, all limits such as the two above can be unraveled using L'Hôspital's rule. However, in the spirit of avoiding Calculus black boxes, we will follow a more circuitous route. Let us denote

$$\lim_{h\to 0} \frac{3\,f(x) - 4\,f(x-h) + f(x-2\,h)}{2\,h} = L(f).$$

Think of L as a mystery operation which when applied to a function f produces another function as its output. Our goal is to relate f to familiar Calculus operations, like the derivative. Now, it is easy to compute the limit defining L for small powers x^n, e.g.:

$$\begin{aligned} L(x^2) &= \lim_{h\to 0} \frac{3\,x^2 - 4\,(x-h)^2 + (x-2\,h)^2}{2\,h} \\ &= \lim_{h\to 0} \frac{3\,x^2 - 4\,(x^2 - 2\,x\,h + h^2) + (x^2 - 4\,x\,h + h^2)}{2\,h} \\ &= \lim_{h\to 0} \frac{4\,x\,h - 3\,h^2}{2\,h} = \lim_{h\to 0} \left(2\,x - \frac{3\,h}{2}\right) = 2\,x. \end{aligned}$$

The table below gives the values of $L(x^n)$ for $n = 0, 1, 2, 3, 4$ and 5.

f	1	x	x^2	x^3	x^4	x^5
$L(f)$	0	1	$2\,x$	$3\,x^2$	$4\,x^3$	$5\,x^4$

The results are identical to the power rule of differentiation. We conclude that $L(f) = f'(x)$.

Similarly, let us denote

$$\lim_{h\to 0} \frac{f(x) - 2\,f(x-h) + f(x-2\,h)}{2\,h^2} = M(f)$$

and compute $M(x^n)$ for small values of n. This time we get:

f	1	x	x^2	x^3	x^4	x^5
$M(f)$	0	0	1	$3\,x$	$6\,x^2$	$10\,x^3$

Since M reduces powers by two, it is reasonable to assume that it is related to the second derivative. In fact, $M(f) = f''(x)/2$.

Taylor polynomial of order two

We now know that

$$\lim_{h\to 0} \frac{3\,f(x_0) - 4\,f(x_0 - h) + f(x_0 - 2\,h)}{2\,h} = f'(x_0)$$
$$\lim_{h\to 0} \frac{f(x_0) - 2\,f(x_0 - h) + f(x_0 - 2\,h)}{2\,h^2} = f''(x_0)/2$$

Consequently, the limit of Equation (19) is

$$T_2(x) = f(x_0) + f'(x_0)\,(x - x_0) + \frac{f''(x_0)}{2}\,(x - x_0)^2. \tag{20}$$

Equation (20) gives a general formula for computing the Taylor polynomial of order two centered at x_0 for a function f; the latter is, obviously, required to be twice-differentiable at x_0. We hasten to add that, just as Equation (18), Equation (20) is a Calculus black box because it does not reflect the process of interpolation using three points on the graph. Think of quadratic approximations, which is what 'Taylor' is synonymous to, in terms of the limit of Equation (19). However, use Equation (20) for rote computations.

Taylor polynomials of higher order

It is easy to extend the procedure for finding T_2 to a general algorithm for finding T_N. To compute the Taylor polynomial of degree N centered at x_0, pick N auxiliary points on the graph. Of the many possible choices, one is to pick auxiliary points to the left of the base point spacing them h units apart:

$$(x_0 - h, f(x_0 - h)), (x_0 - 2\,h, f(x_0 - 2\,h)), \ldots, (x_0 - N\,h, f(x_0 - N\,h)).$$

Any collection of $(N + 1)$ distinct points lying on the graph of $y = f(x)$ defines a unique interpolating polynomial of degree N whose graph passes through those points. It is easy to compute that polynomial using software. For instance, in `Matlab`, one can use the command `polyfit`:

```
...
x = x0 - h*(0:N);
y = f(x);
p = polyfit(x,y,N);
...
```

The variable `p` in this code snippet will contain the coefficients of the interpolating polynomial in descending order (starting with a_N). Unfortunately, the manual computation of the interpolating polynomial can quickly get cumbersome. Therefore, having found T_1 and T_2 constructively, let us try to deduce the general formula for T_N without interpolation.

Let us begin by comparing the 'prime' formulas for T_1 and T_2 given by Equations (18) and (20). What do you notice? Make an observation before reading on.

Did you notice that T_2 differs from T_1 only by one term—the quadratic term? Symbolically,

$$T_2 = T_1 + \frac{f''(x_0)}{2}\,(x - x_0)^3.$$

If we define the constant approximation $T_0 = f(x_0)$, we can also say that

$$T_1 = T_0 + f'(x_0)\,(x - x_0).$$

This suggests that the Taylor cubic can be obtained from the Taylor quadratic with addition of a cubic term:

$$T_3 = T_2 + a_3\,(x - x_0)^3 = f(x_0) + f'(x_0)\,(x - x_0) + \frac{f''(x_0)}{2}\,(x - x_0)^2 + a_3\,(x - x_0)^3.$$

It is also reasonable to assume that a_3 should be a multiple of $f'''(x_0)$:

$$T_3 = f(x_0) + f'(x_0)\,(x - x_0) + \frac{f''(x_0)}{2}\,(x - x_0)^2 + k\,f'''(x_0)\,(x - x_0)^3.$$

In essence, we reduced the problem of finding T_3 to determination of a single number k. Since the quadratic term in T_2 has a factor of $\frac{1}{2}$, a popular guess is that $k = \frac{1}{3}$. Yet it is wrong! To see that, let us return to the theme of this section which is approximation.

As you undoubtedly can recall from Calculus I, the tangent line approximation T_1 is "best" in the sense that T_1 matches the value and the slope of f at the center of expansion:

$$T_1(x_0) = f(x_0) \quad T_1'(x_0) = f'(x_0).$$

It is not hard to see that T_2 also matches the value and the slope. Yet it also matches the value of the second derivative of the function at x_0:

$$T_2(x_0) = f(x_0) \quad T_1'(x_0) = f'(x_0), \quad T_2''(x_0) = f''(x_0).$$

For this reason T_2 will, generally, better agree with f than T_1, unless $f''(x_0) = 0$ in which case $T_2 = T_1$. Now, since T_3 subsumes T_2, so to speak, it matches the value of the function and its first two derivatives at x_0 the way T_2 does. However, T_3 has a cubic term that can be used to match—what else—the third derivative. If we set

$$T_3'''(x_0) = f'''(x_0)$$

then we must choose $k = \frac{1}{6}$. Hence:

$$T_3 = f(x_0) + f'(x_0)\,(x - x_0) + \frac{f''(x_0)}{2}\,(x - x_0)^2 + \frac{f'''(x_0)}{6}\,(x - x_0)^3. \quad (21)$$

Quite generally, given a smooth function $y = f(x)$ and a point x_0, we can define the Taylor polynomial $T_N(x)$ to be the unique polynomial which at x_0 matches the value of the function and the value of its first N derivatives: $T_N^{(n)}(x_0) = f^{(n)}(x_0)$, $n = 0, \ldots, N$. Enclosing a superscript in parentheses, as in $f^{(n)}$, will henceforth be our preferred notation for the derivative of order n; the derivative of order zero is just the function itself. Another useful notation that we are going to adopt is that of a *factorial*. A factorial of a positive integer is the product of all smaller positive integers: $n! = 1 \times 2 \times 3 \times \ldots \times n$. By convention $0! = 1$. The use of superscripts to denote order of differentiation, and factorials, is common in conjunction with the sigma notation for sums. For instance, we can write T_3 compactly as:

$$T_3(x) = \sum_{n=0}^{3} \frac{f^{(n)}(x_0)}{n!}\,(x - x_0)^n.$$

As you may have guessed, the general formula for T_N is just:

$$T_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(x_0)}{n!}\,(x - x_0)^n. \quad (22)$$

Applications of Taylor polynomials

In some sense, everything in differential Calculus is an application of Taylor's theory. Think of it this way: a difficult question about a function can be recast as an equivalent question about a polynomial. Now, polynomials are easy to work with, are not they? That is the philosophy of differential Calculus.

Optimization

Recall that in order to find maxima and minima of some function you would first find critical points by setting the first derivative to zero. You would then investigate critical points by examining the sign of the second derivative. Where do all these recipes come from?

Consider Figure 5. It is evident that at a critical point the graph of T_1—the blue line—is parallel to the x-axis. This means that T_1 has zero slope which translates into $f'(x_0) = 0$. Now examine the plot of T_2, shown in red. The parabola opens downward, so the critical point is a maximum. For the parabola to open downward the highest coefficient must be negative. For T_2 the highest coefficient is $f''(x_0)/2$. So, $f''(x_0) < 0$ at a critical point implies local maximum.

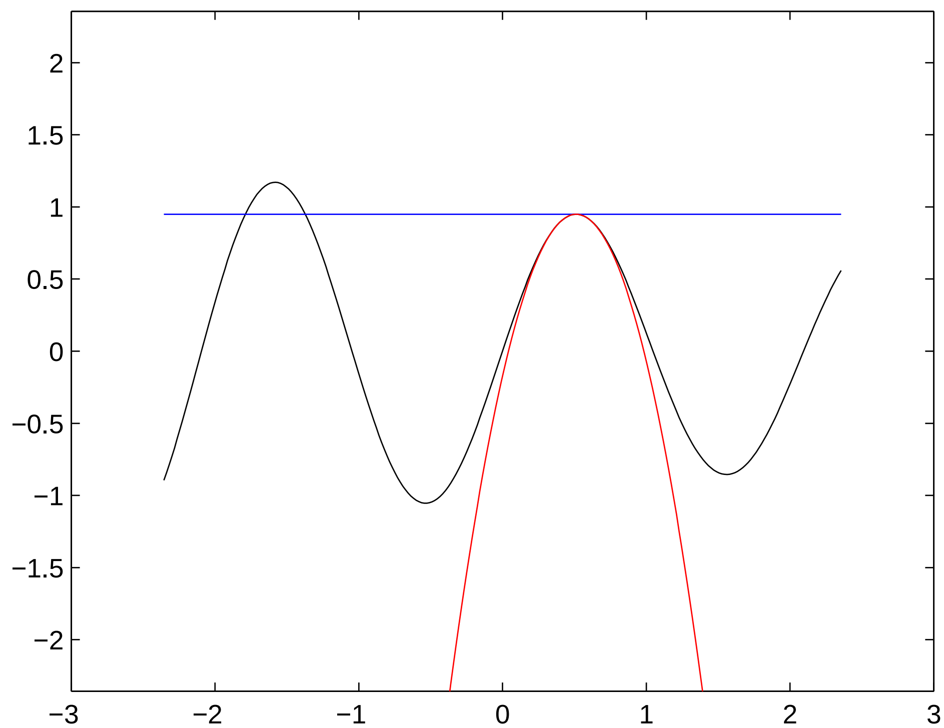

Figure 5: Polynomial approximation of $y = e^{-x/10} \sin(3\,x)$ at a critical point $x_0 = \frac{1}{3} \tan^{-1}(30) = .5124917769\ldots$. The blue line is T_1; the red is T_2.

You will recall from Calculus I that the second derivative test may be inconclusive. Indeed, if $f''(x_0) = 0$ at a critical point then more investigation is needed. You can probably guess now that the additional investigation consists of examining higher order derivatives. Let N be the number of the first nonzero derivative at x_0. Note that $N > 1$ since x_0 is, by assumption, a critical point. Locally, the shape of $y = f(x)$ at x_0 is that of

$$T_N = f(x_0) + \frac{f^{(N)}(x_0)}{N!}\,(x - x_0)^N,$$

which is to say, the graph of f is shaped like $k\,x^N$ where $k = f^{(N)}(x_0)/N!$. If N is odd, the graph looks like a cubic and the critical point is neither a minimum nor a maximum—it is the point of inflection. If N is even, the critical point is a minimum or a maximum depending on whether $f^N(x_0)$ is positive or negative.

Indeterminate forms

Of all the black box type of formulas in Calculus, L'Hôspital's Rule is one of the blackest. Recall that is says that if $f(a) = g(a) = 0$ then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

We can now explain this remarkable fact using polynomial approximation. Suppose that $g'(a) \neq 0$. Approximate f and g with Taylor polynomials of order one. Since $f(a) = 0$, the tangent line approximation for f at $x = a$ is, simply, $f'(a)\,(x-a)$; similarly, the tangent line approximation for g at $x = a$ is $g'(a)\,(x - a)$. Therefore,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(a)\,(x-a)}{g'(a)\,(x-a)} = \frac{f'(a)}{g'(a)},$$

since, by assumption, $g'(a) \neq 0$. If $g'(a) = 0$ then two possibilities exist. If $f'(a) \neq 0$ then tangent line approximation leads to bad division by zero and the limit fails to exist. If $f'(a) = 0$ then the limit can be found using Taylor approximation of order two. Assuming that $g''(a) \neq 0$, we replace f and g with their Taylor quadratics centered at a. This leads to:

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{\frac{1}{2}\,f''(a)\,(x-a)^2}{\frac{1}{2}\,g''(a)\,(x-a)^2} = \frac{f''(a)}{g''(a)}.$$

Arguing in this manner, one concludes that the limit of an indeterminate form of type $0/0$ either does not exist or is given by the ratios of derivatives of some order evaluated at a.

Functional values

Say we wish to approximate $\sqrt{4.1}$. It is clear that the value is slightly greater than $\sqrt{4} = 2$. Yet suppose we need more accuracy, say, three digits after the decimal point. In Calculus I you would use tangent approximation as follows. Set $f(x) = \sqrt{x}$. The tangent line approximation at $x = 4$ is given by:

$$f(x) \approx T_1(x) = 2 + \frac{1}{4}(x-4).$$

Therefore,

$$\sqrt{4.1} = f(4.1) \approx 2 + \frac{1}{4}(.1) = 2.025$$

It is reasonable to assume that the first digit after the decimal point is zero while the second is either two or three. However the third digit is unclear at this point. The good news is that we can now approximate to any accuracy using higher degree polynomials. Using T_2, we get

$$f(x) \approx T_2 = 2 + \frac{1}{4}(x-4) - \frac{1}{64}(x-4)^2$$

and, consequently,

$$\sqrt{4.1} \approx 2 + \frac{1}{4}(.1) - \frac{1}{64}(.1)^2 = 2.024843750$$

The second digit after the decimal point is surely two and the third one is either four or five. Using T_3, we obtain:

$$\sqrt{4.1} \approx 2 + \frac{1}{4}(.1) - \frac{1}{64}(.1)^2 + \frac{1}{512}(.1)^3 = 2.024845703$$

We can now claim with certainty that $\sqrt{4.1} = 2.024\ldots$ In fact, the cubic approximation is correct to six decimals.

You may find it very strange that at this day and age one would want to know how to approximate functional values when they can be computed exactly using calculators. Two things are to be said here. Firstly, calculators and computers typically do not compute functional values exactly. They simply cannot do that using finite arithmetics. Secondly, the way that calculators and computers often find approximate values of common functions is precisely by using Taylor approximation! So, you need to understand how Taylor approximation is used to find banal things like square roots in order to have some idea how modern computing works.

Definite integrals

This, perhaps, is the most relevant application from the point of view of Calculus II. Consider the following definite integral:

$$I = \int_0^{\frac{1}{2}} \cos(x)\, e^{-x^2}\, dx.$$

This integral cannot be computed using any of the symbolic techniques available to us (try it!) However, it can be approximated. Figure 6 shows the graph of the integrand $f = \cos(x)\, e^{-x^2}$ and its Taylor approximation $T_2 = 1 - \frac{3}{2}\, x^2$ centered at zero. There is some discrepancy, of course, but altogether the agreement between f and T_2 is good. Hence:

$$I = \int_0^{\frac{1}{2}} f(x)\, dx \approx \int_0^{\frac{1}{2}} T_2(x)\, dx = \int_0^{\frac{1}{2}} \left(1 - \frac{3}{2}\, x^2\right)\, dx = \frac{7}{16}.$$

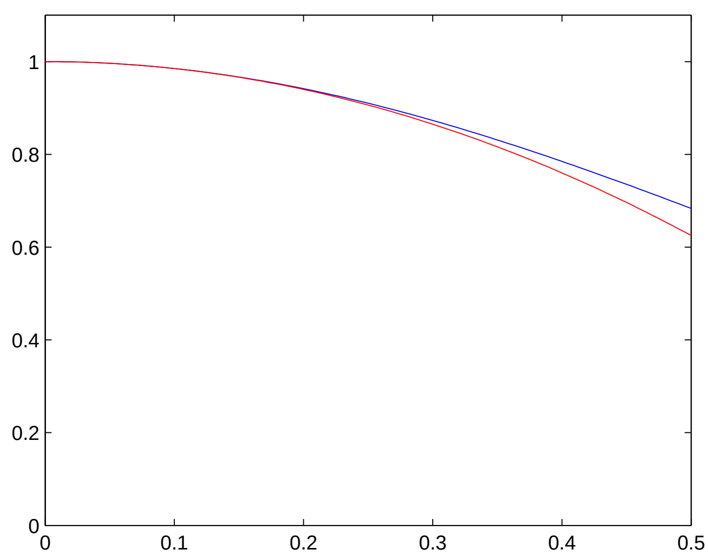

Figure 6: Polynomial approximation of $y = \cos(x)\, e^{-x^2}$ with T_2 centered at zero

In fact, the figure suggests that T_2 produces an underestimate. Thus we can say that $I > 7/16 = .4375$. Since for this integrand $T_3 = T_2$, the next approximation is obtained using T_4:

$$I \approx \int_0^{\frac{1}{2}} T_4(x)\,dx = \int_0^{\frac{1}{2}} \left(1 - \frac{3}{2}\,x^2 + \frac{25}{24}\,x^4\right)\,dx = \frac{341}{768} = .4440104167\ldots$$

It is reasonable to claim that $I = .44\ldots$ In fact, the correct digits are $I = .4435277218\ldots$ and all of them can be obtained by integrating T_{18}.

Exercises

1. In class we estimated $\ln(2)$ using linear approximation. Use Taylor polynomials of degrees 2 through 7 to refine the estimate. In each case compute the error. How quickly does the error approach zero?

2. Construct Taylor polynomials T_n for $n = 1, \ldots, 7$ for $f(x) = e^x$ at $x = 0$. Plot the function and its Taylor polynomials on the interval $[-1, 1]$ and comment on the picture: what does it show.

3. Use the results of the previous problem to approximate

$$\int_{-1}^{1} e^x\,dx \approx \int_{-1}^{1} T_n(x)\,dx$$

for $n = 1, \ldots, 7$. In each case find the error and try to gauge the rate at which it decreases.

4. In this handout we constructed the Taylor polynomial of degree two at x_0 by using two auxiliary points with abscissas $x_0 - h$ and $x_0 - 2\,h$. Review this construction and then attempt it with a more symmetric choice of auxiliary points: $x_0 - h$ and $x_0 + h$.