MTH 4422 Midterm Study Guide - Solutions

Spring 2023

Pat Rossi Name ______________________

Instructions. Answer the following questions thoroughly.

1. Explain the idea behind Newton's Method for solving the equation $f(x) = 0$, using the Taylor's Series approach.

Newton's Method is used to solve equations of the form $f(x) = 0$, where $f(x)$ is a differentiable function, (i.e., We're looking for the value of x such that $f(x) = 0$.) Suppose that x_0 is our initial approximation to x. The Taylor Series expansion of $f(x)$ with center x_0 is given by:

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

Since x is such that $f(x) = 0$, this becomes:

$$0 = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2!} + f'''(x_0)\frac{(x - x_0)^3}{3!} + \ldots$$

If our first approximation is close to the solution x, then $(x - x_0)$ will be small, and consequently, $(x - x_0)^2$ and $(x - x_0)^3$ and $(x - x_0)^n$ will be very small. This means that we can ignore the terms of higher degree. Therefore we have:

$0 \approx f(x_0) + f'(x_0)(x - x_0)$

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

$\Rightarrow -\frac{f(x_0)}{f'(x_0)} \approx (x - x_0)$

$\Rightarrow x_0 - \frac{f(x_0)}{f'(x_0)} \approx x$

Or, $x \approx x_0 - \frac{f(x_0)}{f'(x_0)}$

Since the right side of the equation is an *approximation* of x, we give this approximation of x the name x_1.(our initial approximation of x was x_0).

i.e., $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$

Note that this approximation of x is formulated in terms of x_0 the previous approximation of x.

Continuing in this fashion, if x_n is our current approximation of x, then our next approximation of x will be given by:

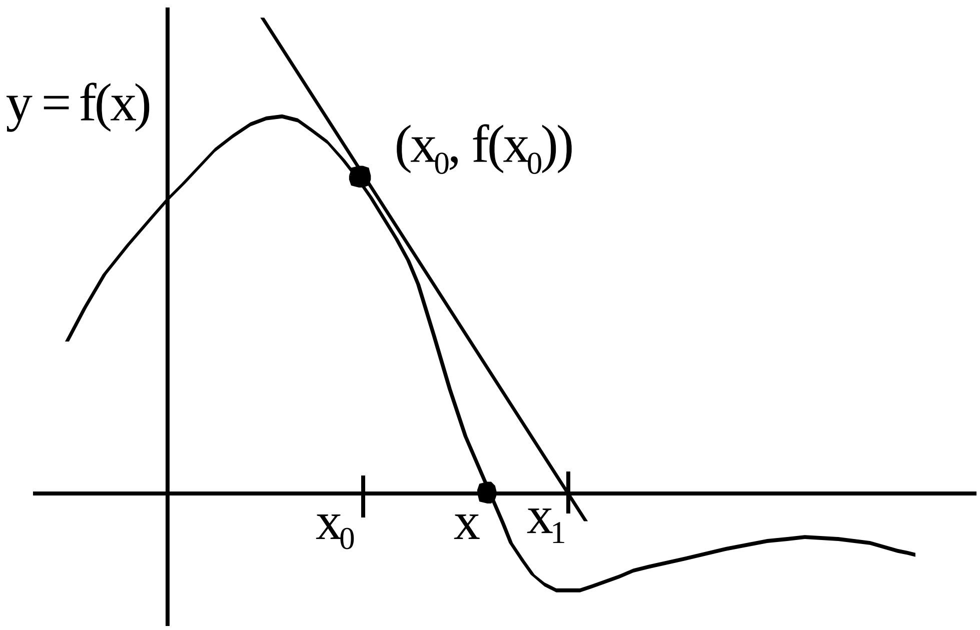
$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

We repeat this iterative process until two successive approximations differ by less than our tolerance, TOL.

2. Explain the idea behind Newton's Method for solving the equation $f(x) = 0$, using the Geometric approach.

Newton's Method solves equations of the form: $f(x) = 0$, where $f(x)$ is a differentiable function. (Note that x is the solution to the equation $f(x) = 0$.)

Consider the line tangent to the graph of $y = f(x)$, at the point $(x_0, f(x_0))$. Geometrically, Newton's Method is based on the supposition that if x_0 is a reasonably good initial approximation of x, then x_1, the point at which this tangent crosses the x-axis, is closer to x than the original approximation, x_0 is to x. The situation is depicted below:



Since x_1, the point of intersection of the line tangent and the x-axis, is closer to x than is the current approximation x_0, we use x_1 as the next approximation of x.

To solve for x_1, we observe that the points $(x_1, 0)$ and $(x_0, f(x_0))$ are points on a line whose slope is $f'(x_0)$. Using the slope formula we get:

$$f'(x_0) = \frac{0 - f(x_0)}{x_1 - x_0} \Rightarrow (x_1 - x_0) = -\frac{f(x_0)}{f'(x_0)} \Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

This formula can be generalized so that given the n^{th} approximation x_n, the $(n+1)^{\text{st}}$ approximation is given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

We repeat this iterative process until two successive approximations differ by less than our tolerance, TOL.

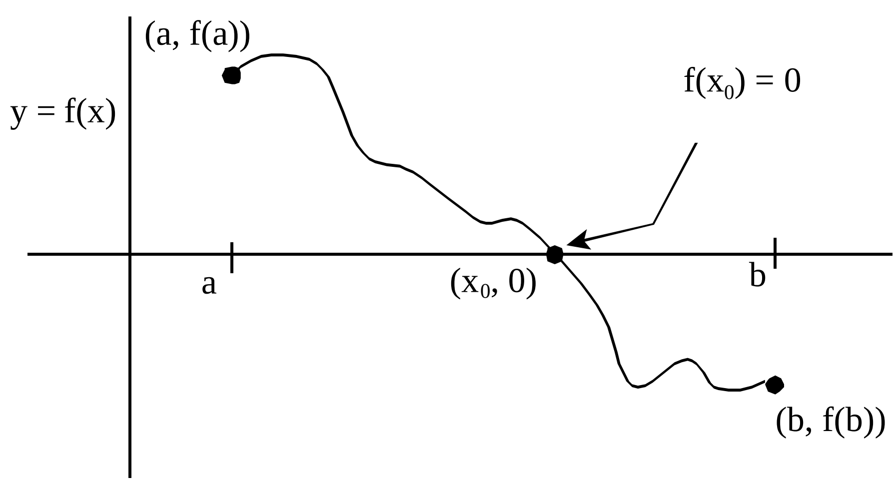
3. Explain how and why the Bisection Algorithm (for solving $f(x) = 0$) works.

The Bisection Algorithm solves equations of the form: $f(x) = 0$, where $f(x)$ is any continuous function which has at least one (real) root. As input, the algorithm requires two values of x ($x = a$ and $x = b$ with $a < b$) such that $f(a)$ and $f(b)$ are of opposite sign. If these conditions are satisfied, the *Intermediate Value Theorem* guarantees the existence of a value, $x = x_0$ with $a < x_0 < b$, such that $f(x_0) = 0$.

(In the context of this explanation, x_0 is the solution of the equation $f(x) = 0$ that we seek.)

Said another way: If $f(x)$ is continuous, then the graph of $f(x)$ is one connected piece. Furthermore, if $f(a)$ and $f(b)$ are of opposite sign, then in order for that graph (which is one connected piece) to get from the point $(a, f(a))$ on one side of the x-axis to the point $(b, f(b))$ on the other side of the x-axis, it has to cross the x-axis at some point $(x_0, 0)$ in between. This is the point at which $f(x) = 0$.

The situation is shown below:



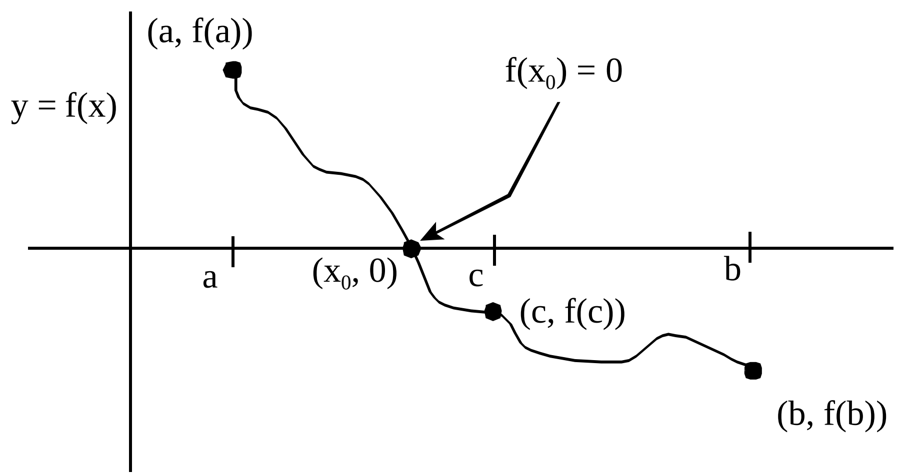
(By the way, in order to determine whether $f(a)$ and $f(b)$ are of opposite sign, we compute the product $f(a) \cdot f(b)$. The values $f(a)$ and $f(b)$ are of opposite sign exactly when their product is negative.)

Given that we have a continuous function, $f(x)$, and x-values, $x = a$ and $x = b$ such that $f(a)$ and $f(b)$ are of opposite sign, we proceed by computing the midpoint, $c = \frac{a+b}{2}$, of the interval $[a, b]$.

Exactly one of the following three statements must be true.

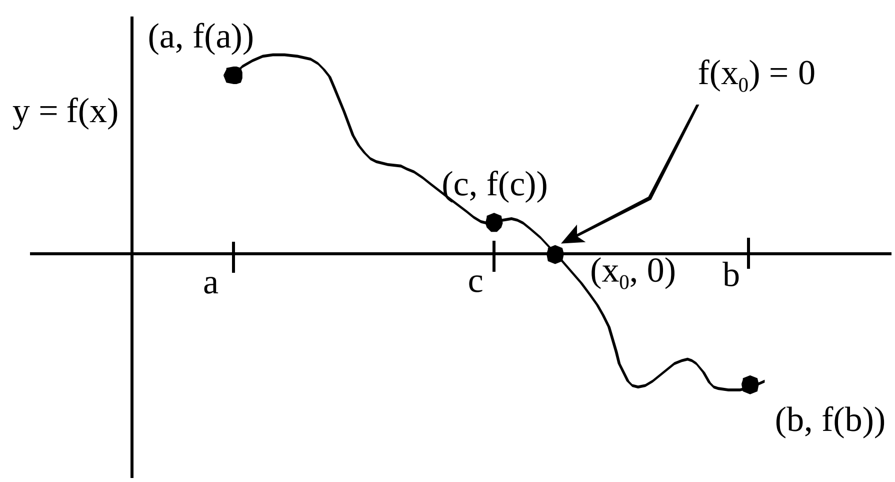
i. $f(c) = 0$.
In this case, c is the value that we seek (i.e. $c = x_{0.}$) So if this occurs, we print the solution and terminate the program.

ii. $f(a)$ and $f(c)$ are of opposite sign.
In this case, the solution x_0, is in the interval $[a, c]$ (shown below). Discard the interval $[c, b]$ by making the following assignments: $a_1 = a$ and $b_1 = c$.



iii. $f(a)$ and $f(c)$ are of like sign.

In this case, the solution x_0, is in the interval $[c, b]$ (shown below). Discard the interval $[a, c]$ by making the following assignments: $a_1 = c$ and $b_1 = b$.



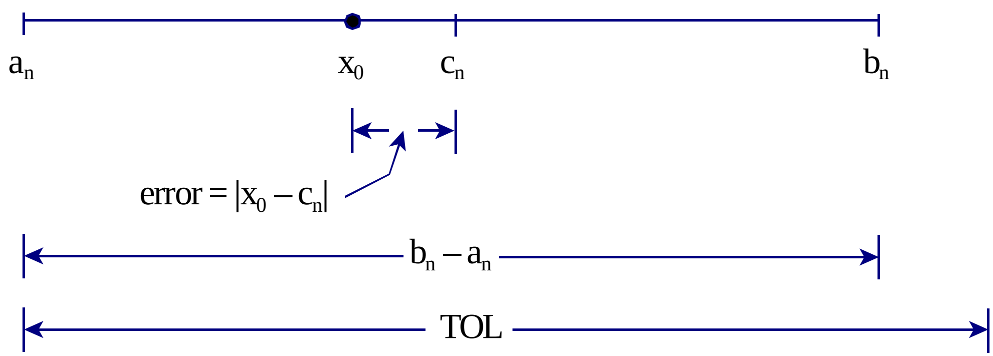
After forming the interval $[a_1, b_1]$, compute the midpoint $c_1 = \frac{a_1+b_1}{2}$. We follow the same "three step" procedure with this new interval, testing to see whether

i. $f(c_1) = 0$

ii. $f(a_1)$ and $f(c_1)$ are of opposite sign

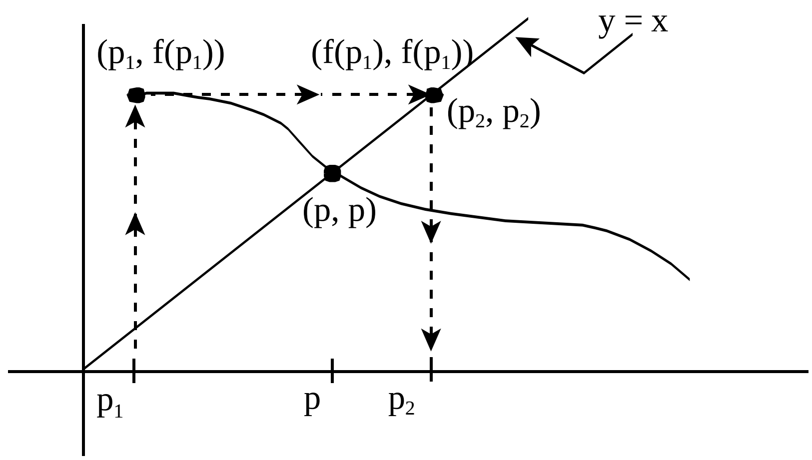
iii. $f(a_1)$ and $f(c_1)$ are of like sign

This iterative process is continued, with each new interval $[a_n, b_n]$ being one half the length of its predecessor. The procedure stops when either $f(c_n) = 0$, or when $|b_n - a_n|$ is less than the maximum error. When this stage is reached (i.e., $|b_n - a_n| < TOL$), we somewhat arbitrarily let our solution be given by $c_n = \frac{a_n+b_n}{2}$. We can do this because when $|b_n - a_n| < TOL$, ANY point in the interval is a good approximation of x_0. The reason for this is that since the interval contains x_0, the distance between x_0 and any point in the interval is less than the length of the interval, which in turn, is less than maximum error, TOL.

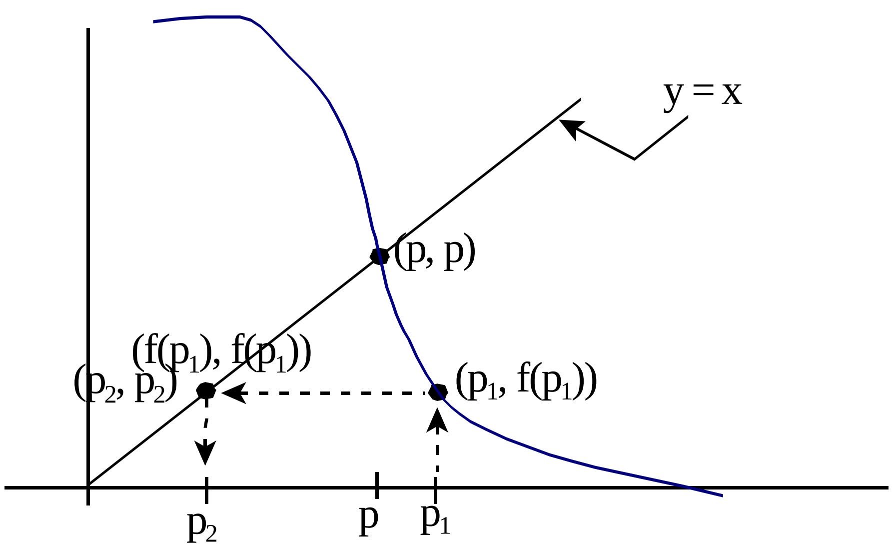


4. Explain how and why the Fixed Point Algorithm (for solving $f(p) = p$) works.

The Fixed Point Algorithm solves equations of the form $f(p) = p$, where $f(x)$ is a differentiable function whose graph intersects the line $y = x$. The algorithm is based on the assumption that if p_n is an approximation of the fixed point, then $p_{n+1} = f(p_n)$ will be a better approximation of the fixed point. If $|f'(x)| < 1$ throughout an interval containing the fixed point, p, then this assumption is valid, as the diagram below shows.



If $|f'(x_1)| \geq 1$ throughout an interval containing the fixed point, p, then the sequence of successive approximations diverges, (i.e., each approximation tends to be farther from the fixed point than its predecessor. See below)



If $|f'(x)| < 1$ throughout a region containing the fixed point, p, then we generate a sequence of approximations $p_{n+1} = f(p_n)$ until a pair of successive approximations differs by less than TOL. We let p_{n+1} be our approximation to the fixed point p.

5. Explain how LaGrange Polynomials are constructed to approximate the function whose data points include $(x_0, f(x_0)), (x_1, f(x_1)), \ldots, (x_n, f(x_n))$, and explain why this works.

Given the data points $(x_0, f(x_0)), (x_1, f(x_1)), \ldots, (x_n, f(x_n))$, from an unknown function $f(x)$, with $x_0 < x_1 < \ldots < x_n$, our goal is to create a polynomial $P(x)$ which closely approximates $f(x)$. (Specifically, we want $P(x) \approx f(x)$ for all values of x between $x = x_0$ and $x = x_n$.)

In our attempt to force $P(x)$ to closely approximate $f(x)$, we will construct $P(x)$ in such a way that $P(x)$ agrees with $f(x)$ at all of the data points. (i.e., $P(x_i) = f(x_i)$ for $i = 0, 1, 2, \ldots, n$.) Here's how we do it.

Given the data points $(x_0, f(x_0)), (x_1, f(x_1)), \ldots, (x_n, f(x_n))$, we define the n^{th} degree LaGrange polynomial $P(x)$ to be given by:

$$P(x) = L_{n,0}(x) f(x_0) + L_{n,1}(x) f(x_1) + \ldots + L_{n,n}(x) f(x_n)$$

where $L_{n,i}(x) = \frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}$.

Since $L_{n,i}(x)$ is an n^{th} degree polynomial and $f(x_i)$ is a constant for $i = 0, 1, 2, \ldots, n$, the term $L_{n,i}(x) f(x_i)$ is an n^{th} degree polynomial. Hence, $P(x)$ is the sum of n^{th} degree polynomials, and is therefore and n^{th} degree polynomial itself.

Note that $L_{n,i}(x) = 1$ when $x = x_i$. This is because for every factor $(x - x_j)$ in the numerator, there is a corresponding factor $(x_i - x_j)$ in the denominator, and vice versa. When $L_{n,i}(x)$ is evaluated at $x = x_i$, the factors in numerator and denominator are identical.

Note also that $L_{n,i}(x_j) = 0$, when $j \neq i$. This is because $(x - x_j)$ is a factor of $L_{n,i}(x)$.

Consequently,

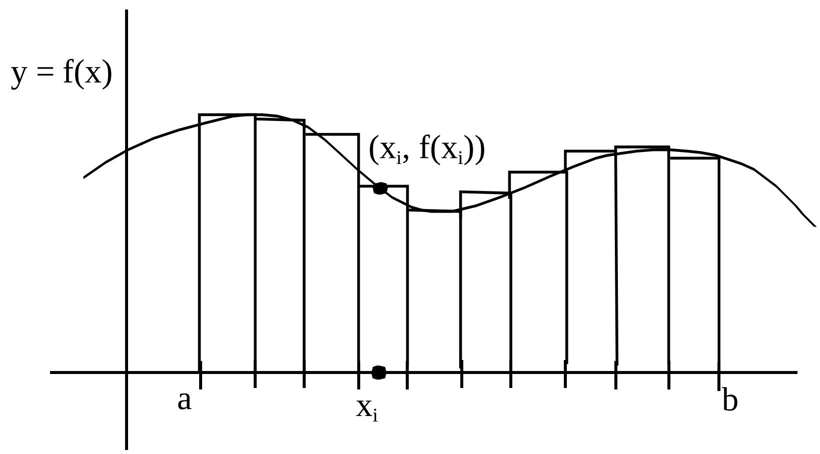
$$P(x_i) = \underbrace{L_{n,0}(x_i)}_{0} f(x_0) + \underbrace{L_{n,1}(x_i)}_{0} f(x_1) + \ldots + \underbrace{L_{n,1}(x_i)}_{1} f(x_i) + \ldots + \underbrace{L_{n,n}(x_i)}_{0} f(x_n)$$

That is, $P(x_i) = f(x_i)$

The polynomial agrees with $f(x)$ at all of the data points.

6. Describe the Trapezoidal Method for integration, and explain why it works.

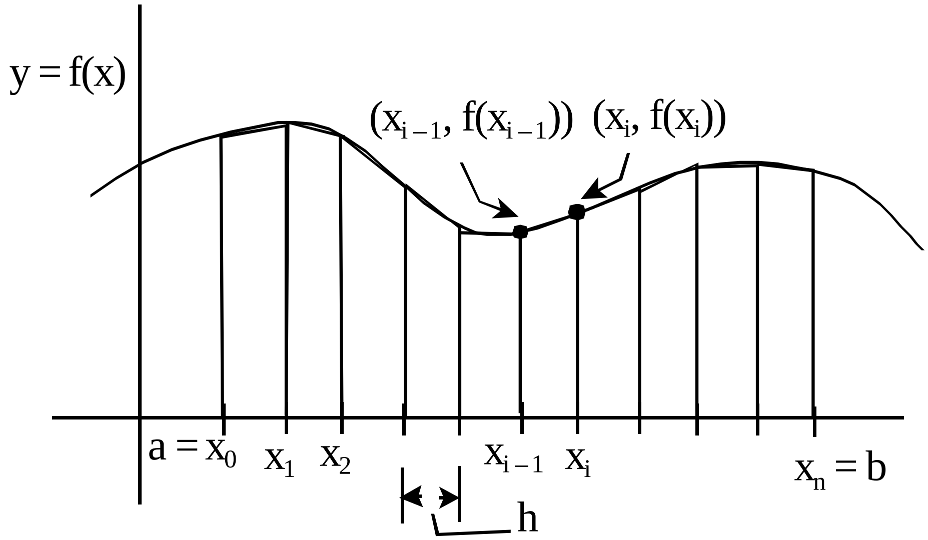
Essentially, we can think of $\int_a^b f(x)\,dx$ as the area between the graph of $f(x)$ and the x-axis, over the interval $[a,b]$. As a first approximation to this area, rectangles can be inscribed inside the region. To do this, we partition the interval $[a,b]$ into n subintervals of width $h = \frac{b-a}{n}$. Above each subinterval we create a rectangle whose base rests on the x-axis, and whose top intersects the graph of $f(x)$. This is shown below. To approximate the area, and hence the integral, we add up the areas of the rectangles. Thus, $\int_a^b f(x)\,dx \approx \sum_{i=1}^n f(x_i)\,h$. There will be "considerable" error associated with this method, since the tops of the rectangles don't conform precisely to the contour of the graph. (See the picture below.)



To improve on the accuracy, we inscribe trapezoids, instead of rectangles, inside the region. We do this as follows: For $i = 0, 1, 2, \ldots, n$, we define $x_i = a + ih$ and plot the points $(x_i, f(x_i))$. We connect adjacent points, $(x_i, f(x_i))$ and $(x_{i+1}, f(x_{i+1}))$, for $i = 0, 1, 2, \ldots, n-1$, with line segments. These line segments will be the tops of the trapezoids. (See the picture below.) For $i = 1, 2, \ldots, n$,the area of the i^{th} trapezoid is $\frac{f(x_i-1)+f(x_i)}{2}h$ (the average of the heights times the base). To approximate the area between the graph of $f(x)$ and the x-axis, and hence the value of $\int_a^b f(x)\,dx$, we add up the areas of the trapezoids. Thus we have:

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^n \underbrace{\frac{f(x_{i-1}) + f(x_i)}{2}h}_{\text{area of } i^{\text{th}} \text{ trapezoid}}$$

Since the tops of the trapezoids conform more closely to the contour of the graph of $f(x)$ than do the tops of the rectangles, we get a better approximation. (See the picture below.)



We would suspect that if we simulate the limit process by letting $h \to 0$, that we would eventually get two successive approximations that differ by less than TOL, and we would let the last approximation be our approximation to the integral. However, this is will not be the case.

Initially, our approximations will improve as we decrease the value of h. But eventually, we will reach the point where h assumes an optimal value. If we make h smaller than its optimal value, the accuracy of our approximations will actually become worse. (The accuracy gained by making h smaller will be more than offset by the roundoff and truncation error associated with the increased number of arithmetic operations that goes hand in hand with a smaller value of h.)

Consequently, we will approximate $\int_a^b f(x)\,dx$ by computing $\sum_{i=1}^{n} \frac{f(x_{i-1})+f(x_i)}{2} h$, using the optimal value of h.

Finally, in order to help reduce roundoff and truncation error, we minimize the total number of arithmetic operations as follows:

Instead of approximating $\int_a^b f(x)\,dx$ by:

$\sum_{i=1}^{n} \frac{f(x_{i-1})+f(x_i)}{2} h$

(i.e., multiplying $[f(x_{i-1}) + f(x_i)]$ by h and then dividing by 2 for $i = 0, 1, 2, \ldots, n$),

we approximate $\int_a^b f(x)\,dx$ by:

$\left(\sum_{i=1}^{n} [f(x_{i-1}) + f(x_i)]\right) \frac{h}{2}$

(i.e., we compute the entire sum $\sum_{i=1}^{n} [f(x_{i-1}) + f(x_i)]$ and then multiply once by $\frac{h}{2}$).

7. Describe Simpson's Method for approximating $\int_a^b f(x)\,dx$, and explain why it works.

The idea behind Simpson's Method is that instead of using line segments to conform to the contour of the graph of $f(x)$, as in the Trapezoidal Method, we use portions of a parabola (actually portions of the graph of a polynomial of degree three or less). An important fact used in this method, is that given any polynomial $p(x)$ of degree three or less, over any interval $[a, a+2h]$, the integral $\int_a^{a+2h} f(x)\,dx$ can be computed in terms of the value of $p(x)$ at the value of the endpoints, a and $a+2h$, and the midpoint, $a+h$, as follows:

$$\int_a^{a+2h} f(x)\,dx = \frac{h}{3}\left[f(a) + 4f(a+h) + f(a+2h)\right].$$

To approximate $\int_a^b f(x)\,dx$ using Simpson's Method, we divide the interval $[a,b]$ into $2n$ subintervals of length $h = \frac{b-a}{2n}$. For $i = 0, 1, 2, \ldots, 2n$, we define $x_i = a + ih$ and group adjacent subintervals into pairs, as follows:

$$\underbrace{[x_0, x_1], [x_1, x_2]}_{\text{first pair}};\underbrace{[x_2, x_3], [x_3, x_4]}_{\text{second pair}}; \ldots; \underbrace{[x_{2n-2}, x_{2n-1}], [x_{2n-1}, x_{2n}]}_{n^{\text{th}}\text{pair}}.$$

Over each pair of subintervals, we approximate the integral $\int_{x_{2(i-1)}}^{x_{2i}} f(x)\,dx$, by computing the integral of $p(x)$ (where $p(x)$ is a polynomial of degree 3 or less that contains the points $\left(x_{2(i-1)}, f\left(x_{2(i-1)}\right)\right)$, $(x_{2i-1}, f(x_{2i-1}))$, and $(x_{2i}, f(x_{2i}))$, using the formula:

$$\int_{x_{2(i-1)}}^{x_{2i}} p(x)\,dx = \frac{h}{3}\left[f\left(x_{2(i-1)}\right) + 4f(x_{2i-1}) + f(x_{2i})\right].$$

(A subtle point here, is that we're setting $p(x_i) = f(x_i)$ for each x_i, $i = 0, 1, 2, \ldots, 2n$.)

Thus:

$$\begin{aligned}
\int_a^b f(x)\,dx &\approx \int_{x_0}^{x_2} p(x)\,dx + \int_{x_2}^{x_4} p(x)\,dx + \ldots + \int_{x_{2n-2}}^{x_{2n}} p(x)\,dx \\
&= \tfrac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2)\right] + \tfrac{h}{3}\left[f(x_2) + 4f(x_3) + f(x_4)\right] + \ldots \\
&\quad \ldots + \tfrac{h}{3}\left[f\left(x_{2(n-1)}\right) + 4f(x_{2n-1}) + f(x_{2n})\right].
\end{aligned}$$

i.e., $\int_a^b f(x)\,dx \approx \sum_{i=1}^{n} \frac{h}{3}\left[f\left(x_{2(i-1)}\right) + 4f(x_{2i-1}) + f(x_{2i})\right]$

We do not compute successive approximations of $\int_a^b f(x)\,dx$, using increasingly smaller values of h. Instead, we approximate $\int_a^b f(x)\,dx =$ using the optimal value of h. Our reason for doing this is as follows:

Initially, our approximations will improve as we decrease the value of h. But eventually, we will reach the point where h assumes an optimal value. If we make h smaller than its optimal value, the accuracy of our approximations will actually become worse. (The accuracy gained by making h smaller will be more than offset by the roundoff and truncation error associated with the increased number of arithmetic operations that goes hand in hand with a smaller value of h.)

Finally, in order to help reduce roundoff and truncation error, we minimize the total number of arithmetic operations as follows:

Instead of approximating $\int_a^b f(x)\,dx$ by:

$$\sum_{i=1}^{n} \frac{h}{3}\left[f\left(x_{2(i-1)}\right)+4f\left(x_{2i-1}\right)+f\left(x_{2i}\right)\right]$$

(i.e., multiplying $\left[f\left(x_{2(i-1)}\right)+4f\left(x_{2i-1}\right)+f\left(x_{2i}\right)\right]$ by $\frac{h}{3}$ for $i=0,1,2,\ldots n$),

we approximate $\int_a^b f(x)\,dx$ by:

$$\left(\sum_{i=1}^{n}\left[f\left(x_{2(i-1)}\right)+4f\left(x_{2i-1}\right)+f\left(x_{2i}\right)\right]\right)\frac{h}{3}$$

(i.e., we compute the entire sum $\displaystyle\sum_{i=1}^{n}\left[f\left(x_{2(i-1)}\right)+4f\left(x_{2i-1}\right)+f\left(x_{2i}\right)\right]$ and then multiply once by $\frac{h}{3}$).

8. Given the data points $(-1,-2),(1,2),(2,7)$, compute the LaGrange Polynomial that agrees with the data points.

   We have 3 data points, so we should have a polynomial of degree 2.

   The LaGrange Polynomial is of the form:

$$P(x) = L_{2,0}(x) f(x_0) + L_{2,1}(x) f(x_1) + L_{2,2}(x) f(x_2),$$

   where $(x_i, f(x_i))$ is the i^{th} data point and $L_{2,i}(x)$ is the polynomial of degree 2 that is the cofactor of $f(x_i)$.

   Thus,

$$L_{2,0}(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} = \frac{(x-(1))(x-(2))}{((-1)-(1))((-1)-(2))} = \frac{1}{6}x^2 - \frac{1}{2}x + \frac{1}{3}$$

$$L_{2,1}(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} = \frac{(x-(-1))(x-(2))}{((1)-(-1))((1)-(2))} = -\frac{1}{2}x^2 + \frac{1}{2}x + 1$$

$$L_{2,2}(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} = \frac{(x-(-1))(x-(1))}{((2)-(-1))((2)-(1))} = \frac{1}{3}x^2 - \frac{1}{3}$$

   The LaGrange Polynomial is:

$$P(x) = L_{2,0}(x) f(x_0) + L_{2,1}(x) f(x_1) + L_{2,2}(x) f(x_2)$$

$$= \left(\frac{1}{6}x^2 - \frac{1}{2}x + \frac{1}{3}\right)(-2) + \left(-\frac{1}{2}x^2 + \frac{1}{2}x + 1\right)(2) + \left(\frac{1}{3}x^2 - \frac{1}{3}\right)(7) = x^2 + 2x - 1$$

   i.e., $P(x) = x^2 + 2x - 1$

   Check:

   $P(-1) = (-1)^2 + 2(-1) - 1 = -2$

   $P(1) = (1)^2 + 2(1) - 1 = 2$

   $P(2) = (2)^2 + 2(2) - 1 = 7$

9. Given the data points $(-2,-46),(-1,-14),(0,-4),(1,2),(2,22)$ compute the LaGrange Polynomial that agrees with the data points.

   We have 5 data points, so we should have a polynomial of degree 4.

   The LaGrange Polynomial is of the form:

   $$P(x) = L_{4,0}(x) f(x_0) + L_{4,1}(x) f(x_1) + L_{4,2}(x) f(x_2) + L_{4,3}(x) f(x_3) + L_{4,4}(x) f(x_4),$$

   where $(x_i, f(x_i))$ is the i^{th} data point and $L_{4,i}(x)$ is the polynomial of degree 4 that is the cofactor of $f(x_i)$.

   Thus,

   $$L_{4,0}(x) = \frac{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)(x_0-x_4)} = \frac{(x-(-1))(x-0)(x-1)(x-2)}{((-2)-(-1))((-2)-0)((-2)-1)((-2)-2)}$$

   $$= \tfrac{1}{24}x^4 - \tfrac{1}{12}x^3 - \tfrac{1}{24}x^2 + \tfrac{1}{12}x$$

   $$L_{4,1}(x) = \frac{(x-x_0)(x-x_2)(x-x_3)(x-x_4)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)(x_1-x_4)} = \frac{(x-(-2))(x-0)(x-1)(x-2)}{((-1)-(-2))((-1)-0)((-1)-1)((-1)-2)}$$

   $$= -\tfrac{1}{6}x^4 + \tfrac{1}{6}x^3 + \tfrac{2}{3}x^2 - \tfrac{2}{3}x$$

   $$L_{4,2}(x) = \frac{(x-x_0)(x-x_1)(x-x_3)(x-x_4)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)(x_2-x_4)} = \frac{(x-(-2))(x-(-1))(x-1)(x-2)}{(0-(-2))(0-(-1))(0-1)(0-2)} = \tfrac{1}{4}x^4 - \tfrac{5}{4}x^2 + 1$$

   $$L_{4,3}(x) = \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_4)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)(x_3-x_4)} = \frac{(x-(-2))(x-(-1))(x-0)(x-2)}{(1-(-2))(1-(-1))(1-0)(1-2)} = -\tfrac{1}{6}x^4 - \tfrac{1}{6}x^3 + \tfrac{2}{3}x^2 + \tfrac{2}{3}x$$

   $$L_{4,4}(x) = \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_3)}{(x_4-x_0)(x_4-x_1)(x_4-x_2)(x_4-x_3)} = \frac{(x-(-2))(x-(-1))(x-0)(x-1)}{(2-(-2))(2-(-1))(2-0)(2-1)}$$

   $$= \tfrac{1}{24}x^4 + \tfrac{1}{12}x^3 - \tfrac{1}{24}x^2 - \tfrac{1}{12}x$$

   The LaGrange Polynomial is:

   $$P(x) = L_{4,0}(x) f(x_0) + L_{4,1}(x) f(x_1) + L_{4,2}(x) f(x_2) + L_{4,3}(x) f(x_3) + L_{4,4}(x) f(x_4)$$

   $$= \left(\tfrac{1}{24}x^4 - \tfrac{1}{12}x^3 - \tfrac{1}{24}x^2 + \tfrac{1}{12}x\right)(-46) + \left(-\tfrac{1}{6}x^4 + \tfrac{1}{6}x^3 + \tfrac{2}{3}x^2 - \tfrac{2}{3}x\right)(-14)$$

   $$+ \left(\tfrac{1}{4}x^4 - \tfrac{5}{4}x^2 + 1\right)(-4) + \left(-\tfrac{1}{6}x^4 - \tfrac{1}{6}x^3 + \tfrac{2}{3}x^2 + \tfrac{2}{3}x\right)(2)$$

   $$+ \left(\tfrac{1}{24}x^4 + \tfrac{1}{12}x^3 - \tfrac{1}{24}x^2 - \tfrac{1}{12}x\right)(22)$$

   $$= 3x^3 - 2x^2 + 5x - 4$$

   i.e., $P(x) = 3x^3 - 2x^2 + 5x - 4$

Check:

$P(-2) = 3(-2)^3 - 2(-2)^2 + 5(-2) - 4 = -46$

$P(-1) = 3(-1)^3 - 2(-1)^2 + 5(-1) - 4 = -14$

$P(0) = 3(0)^3 - 2(0)^2 + 5(0) - 4 = -4$

$P(1) = 3(1)^3 - 2(1)^2 + 5(1) - 4 = 2$

$P(2) = 3(2)^3 - 2(2)^2 + 5(2) - 4 = 22$

10. With reference to the preceding exercise, we had 5 data points and yet the LaGrange Polynomial was only of degree 3. (We would expect that the LaGrange Polynomial that fits all **five** data points would have degree 4.) How can we explain this?

    Our LaGrange Polynomial $P(x) = 3x^3 - 2x^2 + 5x - 4$ of degree 3 agrees with the unknown function at all 5 data points. Under normal circumstances, we would need a LaGrange Polynomial for this to be true. So apparently, the fifth data point provides superfluous information about the unknown function $f(x)$. A quick check using **any** four of the data points will reveal that the LaGrange Polynomial $P(x) = 3x^3 - 2x^2 + 5x - 4$ can be obtained from the four data points that we have randomly chosen. This also implies that the fifth "cofactor" is a linear combination of the other four "cofactors." So again, the fifth data point provides superfluous information about the unknown function $f(x)$. So the five data points are all points on the graph of the same 3^{rd} degree polynomial, $P(x) = 3x^3 - 2x^2 + 5x - 4$.