1. Problem statement.
Most functions do not have antiderivatives in form of easily evaluable function such as polynomials, roots, exponentials, logarithms or their compositions. Consider for example
\[ \int_0^1 e^{x^2} \, dx. \]
This integral is some real number. In order to find it we can only approximate it numerically. Other examples are the error function \[ \int_0^t e^{-x^2} \, dx, \] important in probability, \[ \int_0^t \sin(x^2) \, dx, \] \[ \int_{-1}^1 \sqrt{1+x^3} \, dx. \] (Can you sketch the area that is represented by the integrals, in each case?)

Let’s go back to the original definition of an integral: We started with the problem of trying to find the area between the x-axis, the lines \( x = a \) and \( x = b \), and the graph of a nonnegative function \( y = f(x) \geq 0 \). We obtained this by first approximating the area by a sum of rectangles. In the limit as the number of rectangles goes to infinity, the area of the rectangles should approach the area under the graph of the function. This limit is defined to be the definite integral \( \int_a^b f(x) \, dx. \)

2. Example.
In class we approximated the simple integral (for which we happen to know the exact answer)
\[ \int_0^2 x^2 \, dx \quad \text{(} = \frac{8}{3} \approx 2.66666 \text{)} \]
using the Left-endpoint Rule, Right-endpoint rule, the Midpoint Rule, and the Trapezoid rule, using a range of values of \( n \). You need to be able to do this by hand for simple examples, for sufficiently small \( n \) that the calculations are manageable.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( L_n )</th>
<th>( R_n )</th>
<th>( T_n )</th>
<th>( M_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.00</td>
<td>5.00</td>
<td>3.00</td>
<td>2.50</td>
</tr>
<tr>
<td>4</td>
<td>1.75</td>
<td>3.75</td>
<td>2.75</td>
<td>2.626</td>
</tr>
<tr>
<td>8</td>
<td>2.1875</td>
<td>3.1875</td>
<td>2.6875</td>
<td>2.6562</td>
</tr>
<tr>
<td>16</td>
<td>2.4219</td>
<td>2.9219</td>
<td>2.6719</td>
<td>2.6641</td>
</tr>
<tr>
<td>32</td>
<td>2.5430</td>
<td>2.7930</td>
<td>2.6680</td>
<td>2.6660</td>
</tr>
<tr>
<td>64</td>
<td>2.6045</td>
<td>2.7295</td>
<td>2.6670</td>
<td>2.6665</td>
</tr>
</tbody>
</table>

We see that the left and right endpoint rules where the worst estimates. The trapezoid and midpoint where approximately equally good.

Note: Trapezoid rule \( T_n \) is the average of \( R_n \) and \( L_n \). Why?

Note: made in class, on the relation between \( R_n, L_n, T_n \) and \( M_n \) for increasing, concave up or increasing, concave down functions.

We now find general formulas for \( R_n, T_n, M_n, \ldots \), and will learn more about the approximation errors.

3. The Left and Right Endpoint Rules.
The goal is to approximate the definite integral \( I = \int_a^b f(x) \, dx \). In all our numerical approximations we first partition \( [a,b] \) into \( n \) intervals
\[ [x_{j-1}, x_j], \quad j = 1, \ldots, n \]
of equal width \( h = \Delta x = x_j - x_{j-1} \). This involves \( n + 1 \) equally spaced points \( x_j, j = 0, \ldots, n \). On each interval we approximate the area under the graph of \( f \) using the functions values at the points, \( f(x_j) \). For convenience, let
\[ f_j = f(x_j), \quad j = 0, \ldots, n \]
The left-endpoint rule consists of approximating the area on the $j$th interval $[x_{j-1}, x_j]$ by a rectangle of height $f_{j-1}$, with area

$$f_{j-1}h$$

A sum over all intervals yields the total approximation of the integral

$$I \approx L_n = \sum_{j=1}^{n} f_{j-1} h = (f_0 + f_1 + f_2 + \ldots + f_{n-1})h$$

Similarly, the right-endpoint rule is the approximation

$$I \approx R_n = \sum_{j=1}^{n} f_j h = (f_1 + f_2 + \ldots + f_n)h$$

4. The Trapezoid Rule.

The trapezoid rule consists of approximating the area on the $j$th interval $[x_{j-1}, x_j]$ by a trapezoid. See figure. The area of a trapezoid is the average height times the width of the base :

$$\frac{f_{j-1} + f_j}{2} h$$

A sum over all intervals yields the total approximation of the integral

$$I \approx T_n = \sum_{j=1}^{n} \frac{f_{j-1} + f_j}{2} h$$

$$= \frac{f_0 + f_1}{2} h + \frac{f_1 + f_2}{2} h + \frac{f_2 + f_3}{2} h + \ldots + \frac{f_{n-1} + f_n}{2} h$$

$$= \left( \frac{f_0}{2} + f_1 + \frac{f_2}{2} + f_3 + \ldots + \frac{f_n}{2} \right) h$$

**Example 1:** Approximate $\int_1^2 \sin(x^2)\,dx$ using the trapezoid rule with $n = 4$ intervals. Solution: $h = 1/4$.

$$T_4 = \frac{1}{4} \left( \frac{\sin(1^2)}{2} + \sin(1.25^2) + \sin(1.5^2) + \sin(1.75^2) + \frac{\sin(2^2)}{2} \right) \approx 0.4748$$

4. The Midpoint Rule.

The midpoint rule consists of approximating the area on the $j$th interval $[x_{j-1}, x_j]$ by a rectangle whose height is the value of $f$ at the midpoint $x_j^m = (x_{j-1} + x_j)/2$. Again, for convenience, let

$$f_j^m = f\left(\frac{x_{j-1} + x_j}{2}\right)$$

Then the midpoint approximation is

$$I \approx M_n = \sum_{j=1}^{n} f_j^m h = \left( f_1^m + f_2^m + f_3^m + \ldots + f_n^m \right) h$$

See Figure.
Example 3: Approximate \( \int_1^2 \sin(x^2) \, dx \) using the midpoint rule with \( n = 4 \) intervals. Solution: \( h = 1/4 \).

\[
M_4 = \frac{1}{4} \left( \sin 1.125^2 + \sin 1.375^2 + \sin 1.625^2 + \sin 1.875^2 \right) \approx 0.5045
\]

5. Simpson’s Rule.

For Simpson’s rule the interval \([a, b]\) is broken up into an **even** number of intervals of **equal width** \( h \), this is important. The intervals are then grouped into groups of 2 consecutive intervals. Any two consecutive intervals have three points \((x_j, y_j), (x_{j+1}, y_{j+1}), (x_{j+2}, y_{j+2})\) associated with them. There exists a unique quadratic function \( p(x) = a_0 + a_1 x + a_2 x^2 \) (with 3 coefficients \( a_0, a_1, a_2 \)) that interpolates the three points. Simpson’s rule consists of approximating the integral of the function by the integral of this interpolant. See figure. The quadratic interpolant over the first two intervals in this figure is hard to distinguish from the actual function \( y = f(x) \). The last two intervals show more clearly a difference between the two.

Note that the trapezoid rule consists of approximating the integral of the function by the integral of a linear interpolant of every two points \((x_{j-1}, y_{j-1}), (x_j, y_j)\), on the \( j \)th interval. Generally, a quadratic interpolant is a better approximation to a function than two linear pieces, so we would expect Simpson’s rule to be better than the trapezoid rule.

The integral of the quadratic interpolant over two consecutive subintervals \([x_j, x_{j+2}]\) is

\[
(\frac{1}{3} f_j + \frac{4}{3} f_{j+1} + \frac{1}{3} f_{j+2}) \, h
\]

(which we will not derive here). By combining the integral over all pairs of subintervals we obtain the resulting Simpson’s approximation of the integral of the function over \([a, b]\), given by

\[
I \approx S_n = \left( \frac{1}{3} f_0 + \frac{4}{3} f_1 + \frac{2}{3} f_2 + \frac{4}{3} f_3 + \ldots + \frac{2}{3} f_{n-2} + \frac{4}{3} f_{n-1} + \frac{1}{3} f_n \right) \, h
\]

Notice that the weights associated with \( f(x_j) \) are \( \frac{1}{3}, \frac{4}{3}, \frac{2}{3}, \frac{4}{3}, \frac{2}{3}, \ldots, \frac{4}{3}, \frac{1}{3} \). It is important that the first and last weights are \( \frac{1}{3} \), the second is \( \frac{4}{3} \), alternating with \( \frac{2}{3} \). If \( n \) is even, the second to last weight will be \( \frac{4}{3} \). If \( n \) is not even, Simpson’s rule does not work. By construction, it requires \( n \) to be even.

Example 4: Approximate \( \int_1^2 \sin(x^2) \, dx \) using Simpson’s rule with \( n = 4 \) intervals. Solution: \( h = 1/4 \).

\[
S_4 = \frac{1}{12} \left( \sin(1^2) + 4\sin(1.25^2) + 2\sin(1.5^2) + 4\sin(1.75^2) + \sin(2^2) \right) \approx 0.4964
\]
6. Comparison of these rules.

Is any one rule better than any other? If so, in what sense? It turns out that the Trapezoid and the Midpoint rule are equally good (even though the answers are different), whereas Simpson’s rule is better, provided the integrand \( f \) is sufficiently nice. Let’s look at the approximations to

\[
I = \int_{1}^{2} \sin(x^2) \, dx
\]

with increasing numbers of intervals \( n \). Ideally, as \( n \) increases, the approximations get closer and closer to the exact value. The following is a table of results using the right-rectangle rule \( R_n \), the trapezoid rule \( T_n \), the midpoint rule \( M_n \) and Simpson’s rule \( S_n \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( R_n )</th>
<th>( T_n )</th>
<th>( M_n )</th>
<th>( S_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.275061626001</td>
<td>0.474845811016</td>
<td>0.504496837340</td>
<td>0.496393174415</td>
</tr>
<tr>
<td>8</td>
<td>0.389779231671</td>
<td>0.499671324178</td>
<td>0.495110999463</td>
<td>0.494613161899</td>
</tr>
<tr>
<td>16</td>
<td>0.463357703402</td>
<td>0.493303749655</td>
<td>0.495110999463</td>
<td>0.494514558148</td>
</tr>
<tr>
<td>32</td>
<td>0.469234351432</td>
<td>0.494207374559</td>
<td>0.494658631137</td>
<td>0.494508582861</td>
</tr>
<tr>
<td>64</td>
<td>0.481946912856</td>
<td>0.494430028484</td>
<td>0.4945782317</td>
<td>0.494508212278</td>
</tr>
<tr>
<td>128</td>
<td>0.488246136801</td>
<td>0.494489325853</td>
<td>0.49517585284</td>
<td>0.494508199161</td>
</tr>
<tr>
<td>256</td>
<td>0.491381861042</td>
<td>0.494503489933</td>
<td>0.4950536977</td>
<td>0.494508177171</td>
</tr>
<tr>
<td>512</td>
<td>0.492946199008</td>
<td>0.494507012953</td>
<td>0.49508774955</td>
<td>0.49450817626</td>
</tr>
<tr>
<td>1024</td>
<td>0.493724869811</td>
<td>0.494507893954</td>
<td>0.49450834454</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>2048</td>
<td>0.494117910717</td>
<td>0.494508114204</td>
<td>0.494508224329</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>4096</td>
<td>0.494313067523</td>
<td>0.494508169266</td>
<td>0.494508196797</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>8192</td>
<td>0.494410632160</td>
<td>0.494508183032</td>
<td>0.494508199915</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>16384</td>
<td>0.494459411037</td>
<td>0.494508186473</td>
<td>0.494508188194</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>32768</td>
<td>0.494483799616</td>
<td>0.494508187334</td>
<td>0.49450817764</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>65536</td>
<td>0.494495959369</td>
<td>0.494508187549</td>
<td>0.49450817656</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>131072</td>
<td>0.494502090673</td>
<td>0.494508187602</td>
<td>0.49450817629</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>262144</td>
<td>0.494505139151</td>
<td>0.49450818716</td>
<td>0.49450817623</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>524288</td>
<td>0.494506663387</td>
<td>0.49450818619</td>
<td>0.49450817621</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>1048576</td>
<td>0.494507425504</td>
<td>0.49450818762</td>
<td>0.49450817621</td>
<td>0.49450817621</td>
</tr>
<tr>
<td>2097152</td>
<td>0.494507806562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4194304</td>
<td>0.494507997091</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8388608</td>
<td>0.494508092356</td>
<td>0.4945081876</td>
<td>0.4945081876</td>
<td>0.4945081876</td>
</tr>
<tr>
<td>16777216</td>
<td>0.494508139988</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33554432</td>
<td>0.494508163804</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The arrows indicate the values that the sequence of approximations appears to converge to as \( n \) increases, where only those significant digits are included which appear to have converged. We see that the Trapezoid, Midpoint and Simpsons approximations have converged to roughly the same number of digits (9-10), but Simpsons rule gets there much faster (with much fewer values of \( n \)), whereas Trapez and Midpoint take about equal number of \( n \) for equal number of digits. The rectangle rule has only converged to within 7 digits, even though the value of \( n \) is much larger.

Let’s look at an example for which we know the answer, then we can compute the error. The following lists the errors \( e_n \) in the approximations of

\[
I = \int_{0}^{1} \frac{dx}{1+x^2} = \frac{\pi}{4}
\]

| \( n \) | \(|I - R_n|\)   | \(|I - T_n|\)   | \(|I - M_n|\)   | \(|I - S_n|\)   |
|-------|------------|------------|------------|------------|
| 4     | 0.466994134256 | 0.003594101038 | 0.004722623380 | 0.039543153864 |
| 8     | 0.235858378818 | 0.000564261171 | 0.000276097301 | 0.001950381907 |
We see that as \( n \) is doubled the error decreases, but it seems to decrease faster for Simpson’s rule, than for the rectangle rule, for example. To get a better grip of how fast the error is decreasing, the following table plots the error \( e_n/e_{n+1} \) for each method.

<table>
<thead>
<tr>
<th>( n )</th>
<th>Rectangle</th>
<th>Trapez</th>
<th>Midpoint</th>
<th>Simpson</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.97997668308</td>
<td>6.369570019802</td>
<td>17.104924124569</td>
<td>20.274569671547</td>
</tr>
<tr>
<td>8</td>
<td>2.002343956421</td>
<td>3.916252030409</td>
<td>3.99660622713</td>
<td>484.905482998266</td>
</tr>
<tr>
<td>16</td>
<td>2.001223351887</td>
<td>3.998057341985</td>
<td>3.99650470273</td>
<td>172.356218822915</td>
</tr>
<tr>
<td>32</td>
<td>2.000612383664</td>
<td>3.999513707491</td>
<td>3.999149022028</td>
<td>15.9977381513781</td>
</tr>
<tr>
<td>64</td>
<td>2.000306303890</td>
<td>3.999873699988</td>
<td>3.999787152341</td>
<td>15.992108410888</td>
</tr>
<tr>
<td>128</td>
<td>2.000153171825</td>
<td>3.999968889796</td>
<td>3.999946785688</td>
<td>15.996733169992</td>
</tr>
<tr>
<td>256</td>
<td>2.000076589865</td>
<td>3.999992403279</td>
<td>3.999998309377</td>
<td>15.9990671641791</td>
</tr>
<tr>
<td>512</td>
<td>2.000038295795</td>
<td>3.9999998309387</td>
<td>3.9999997552396</td>
<td>15.028037383178</td>
</tr>
</tbody>
</table>

We see that as \( n \) is doubled (\( h \) decreases by a factor of 2),
- the error in the rectangle rule decreases by a factor of that approaches 2 as \( h \to 0 \),
- the error in the trapezoid and midpoint rule decreases by a factor that approaches 4 as \( h \to 0 \),
- the error in Simpson’s rule decreases by a factor that approaches 16 as \( h \to 0 \) (until the effect of roundoff error breaks this pattern).

Thus, for this example, the errors look like:

\[
\begin{align*}
|I - R_n| & \approx Ch \\
|I - T_n| & \approx Ch^2 \\
|I - M_n| & \approx Ch^2 \\
|I - S_n| & \approx Ch^4
\end{align*}
\]

as long as \( h \) is sufficiently small, where \( C \) are some constants. This can be shown to be true in general, as long as the integrand is sufficiently nice. For example, if the 4th derivative of \( f(x) \) is not bounded, Simpson’s rule will not converge as \( O(h^4) \). (More on this in section 8 below.)

7. MATLAB.

The results we looked at above in form of a table we will now look at graphically. In class we used the following matlab script, together with the functions below, to illustrate the different behaviour of the different numerical approximations.

MATLAB script:
```matlab
clear
f=inline('1./(1+x.^2)'); a=0; b=4; ex=atan(4);
for i=1:9
    n(i)=2^i;
    approx=trapez(f,a,b,n(i)); errt(i) = abs(approx-ex);
    approx=simpsons(f,a,b,n(i)); errs(i) = abs(approx-ex);
    approx=midpt(f,a,b,n(i)); errm(i) = abs(approx-ex);
    approx=leftpt(f,a,b,n(i)); errl(i) = abs(approx-ex);
end
err=errt;
```
m=length(err); ratio=err(1:m-1)./err(2:m); mat=[err', [0,ratio']]  

h=(b-a)/n;  

figure(1), clf, set(gca,'FontSize',20)  

loglog(h,errt,'*-b',h,errs,'+-r',h,errm,'-*g',h,errl,'-*m') legend('trapez','simpson','midpoint','leftpoint',4)  

xlabel('h'), ylabel('error')  

axis([0.7*10^(-2),2,10^(-12),1])  

set(gca,'YTick',10.^[-12:1])  

set(gca,'XTick',10.^[-2:1])

MATLAB functions:

function z = leftpt(f,a,b,n)  
h=(b-a)/n;  
x=a+(0:n)*h; y=f(x);  
z = sum(y(1:n)); z = z*h;  

function z = midpt(fname,a,b,n)  
h=(b-a)/n;  
x=a+h/2:h:b-h/2; y=feval(fname,x);  
z=sum(y)*h;  

function z = trapez(f,a,b,n)  
h=(b-a)/n;  
x=a+(0:n)*h; y=f(x);  
z = (y(1)+y(n+1))/2 + sum(y(2:n)); z = z*h;  

function z = simpsons(f,a,b,n)  
h=(b-a)/n;  
x=a+(0:n)*h; y=f(x);  
w([1,n+1])=1; w(2:2:n)=4; w(3:2:n-1)=2;  
z=h/3*w*y';

MATLAB output:
run the code above to get the following figure
8. Why do the errors plotted look linear?

For \( h \) sufficiently small, all of the curves plotted in the Matlab figure look linear. Why is this? What is their slope?

The curves are plotted on a log-log scale. From the tables in section 6 above we deduced that for \( h \) sufficiently small, all the errors look like

\[
\text{error} \approx Ch^p
\]

for some \( p \). (\( p=1 \) for left or right-endpoint rule, \( p=2 \) for trapezoid and midpoint rule, \( p=4 \) for simpsons)

Take the log of both sides of this equation to get

\[
\log(\text{error}) \approx \log(Ch^p) = \log(C) + \log(h^p) = \tilde{c} + p \times \log(h)
\]

So \( \log(\text{error}) \) is a linear function of \( \log(h) \)!! This is exactly what we see in the figure. The slope of the line should be \( p \).

You can check that if \( \log(h) \) changes by one unit, for example on the interval \( h \in [10^{-2}, 10^{-1}] \), where \( \log(h) \in [-2, -1] \), then the error changes by 4 units in the log scale for simpsons, from \( 10^{-12} \) to \( 10^{-8} \), giving a slope of \( 4/1=4 \), as expected from the table. Similarly, can confirm \( p = 2, 2, 1 \) for the other 3 cases.

Notice that the linear behaviour which shows that simpsons error decreases faster than trapezoid error only holds if \( h \) is sufficiently small. For large \( h \), simpsons is actually worse than trapezoid and does not decrease like \( h^4 \).

9. Analytical results

Here are some results that can be proven using Taylor series, which we will study later this semester. If \( E_T, E_M, E_S \) denote the errors in approximating a given integral \( \int_a^b f(x) \, dx \), then

\[
|E_T| \leq \frac{b-a}{12} h^2 \max_{x \in [a,b]} |f''(x)|
\]

\[
|E_M| \leq \frac{b-a}{24} h^2 \max_{x \in [a,b]} |f''(x)|
\]

\[
|E_S| \leq \frac{b-a}{180} h^4 \max_{x \in [a,b]} |f^{(4)}(x)|
\]

These are the same formulas as written on §7.7, page 534, in Stewart, but replacing

\[
h = \frac{b-a}{n}
\]

To use these formulas to estimate upper bounds for the errors you need to find bounds on the corresponding derivatives of \( f \). For example, if \( f(x) = \sin(kx) \), then

\[
|f''(x)| = | -k^2 \sin(kx)| \leq k^2
\]

For this you often need to use the Triangle Inequality, given in the margin of the book on page 78, and also in the appendix, page A8,

\[
|a + b| \leq |a| + |b|
\]

for and numbers \( a \) and \( b \). Using the triangle inequality, you can bound, for example \( |20 \sin(x^2) - 3x^2 \cos(x)| \) on \( x \in [0, 4] \) by

\[
|20 \sin(x^2) - 3x^2 \cos(x)| \leq 20 \cdot 1 + 3 \cdot 4^2 \cdot 1
\]

And yes, as Cristina correctly noted, to obtain these upper bounds you find the maximum of each term by plugging in possibly different values of \( x \).