CALCULUS A, LAB 6 SOLUTIONS

1. The results of the *Sage* calculations are shown in the tables below.

Approximation to the area

n	Right rectangles	Left rectangles	Trapezoids	Midpoints
1	1.57078	0	0.78540	1.11071
10	1.076482	0.919403	0.9979430	1.0010288
100	1.0078334	0.9921255	0.999979438	1.000010281
1000	1.00078519	0.99921440	0.99999979438	1.00000010281
$10,000$	1.000078538	0.999921458	0.9999999979438	1.0000000010281
$100,000$	1.0000078540	0.9999921460	0.999999999979438	1.000000000010281

Error in the approximation

n	Right rectangles	Left rectangles	Trapezoids	Midpoints
1	0.57078	-1	-0.21460	0.11072
10	0.076482	-0.080597	-0.0020570	0.0010288
100	0.0078334	-0.0078745	-0.000020562	0.000010281
1000	0.00078519	-0.00078560	-0.00000020562	0.00000010281
$10,000$	0.000078538	-0.000078542	-0.0000000020562	0.0000000010281
$100,000$	0.0000078540	-0.0000078540	-0.000000000020562	0.000000000010281

2. The sine function is increasing and concave down everywhere in the interval $[0, \pi/2]$. This means that the right-hand rectangles will always lie above the curve, and that the left-hand rectangles and trapezoids will always lie below it. The right-hand sum therefore is larger than the actual area under the curve, and the left-hand and trapezoid sums are smaller.

To deal with the midpoint sums, look at a picture of a typical slice for a concave-down function, shown in Figure 1. The area under the midpoint rectangle (dotted line) is equal to the area under the trapezoid whose top is formed by the tangent line to the curve at the midpoint (dashed line). Because the curve is concave down, this tangent line always lies above the curve. The midpoint sum is therefore larger than the actual area.

3. Multiplying n by 10 seems to decrease the error for the right-hand or left-hand sum by a factor of 10, which means that the error for these methods is proportional to $1/n$. Getting 100 digit accuracy (an error of 10^{-100}) would take on the order of 10^{100} rectangles, a completely impossible number to evaluate.

For the trapezoid and midpoint methods, multiplying n by 10 seems to decrease the error by a factor of 10^2, which means that the error for these methods is proportional to $1/n^2$. Getting 100 digit accuracy (an error of 10^{-100}) would therefore take on the order of 10^{50} rectangles, a gigantic improvement, but still impossible number to evaluate.

4. Surprisingly, the midpoint method worked better than the trapezoid method. A geometric justification for this can be seen in Figure 1, which shows one slice of

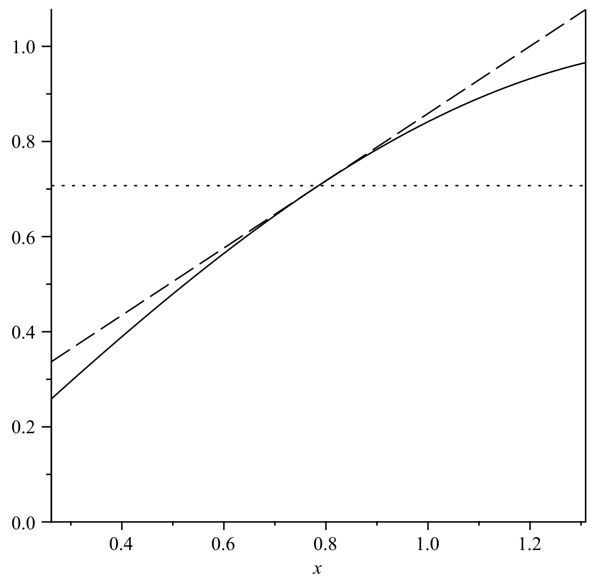

FIGURE 1. Midpoint rectangle and tangent trapezoid.

the function. The midpoint rule approximates the area under the curve by taking the area below the dotted horizontal line meeting the curve at the midpoint of the interval. It's easy to see geometrically, though, that the area under that dotted line is equal to the area under the dashed line tangent to the curve at its midpoint. Now ask yourself which straight line better approximates the function: this tangent line or the chord joining the points at the left and right endpoints (which is what the trapezoid method would use)? It seems obvious that the tangent line is the better choice. It also seems clear that at least for functions that don't have inflection points in the particular slice we are looking at, the exact area will lie between the midpoint and tangent approximations. This is borne out by looking at the numerical data, which also make it appear as though the absolute value of the error of the midpoint method is about half that of the trapezoid method.

5. The insight in the previous problem makes it appear that a good approach might be to start at the value of the midpoint sum M and to go $\frac{1}{3}$ of the way toward the trapezoid sum T. That is, we would compute

$$M + \tfrac{1}{3}(T - M) = \frac{2M + T}{3}.$$

This approach is called Simpson's Rule. The values and errors of computations with this method are shown in the table below.

n	Simpson's Rule Value	Simpson's Rule Error
1	1.0022799	0.0022799
10	1.00000021155	0.00000021155
100	1.000000000021139	0.000000000021139
1000	1.0000000000000021139	0.0000000000000021139
$10,000$	1.00000000000000000021139	0.00000000000000000021139
$100,000$	1.000000000000000000000021139	0.000000000000000000000021139

The error in Simpson's Rule seems to be proportional to $1/n^4$, since multiplying n by 10 decreases the error by a factor of 10,000. This means that we'd need around 10^{25} slices to get 100 digit accuracy with Simpson. This is still enormous, but it is almost within reach of a multi-year global distributed effort.

Suppose we asked instead for 48 digit accuracy. The right-hand and left-hand approaches would take a number of slices equal to about 10^{14} times the number of nanoseconds since the Big Bang; so if a hundred trillion computers had started on the problem at that time, we'd be just about done now if we could handle one slice per computer per nanosecond, which is still highly optimistic. The midpoint and trapezoid approaches would take about 10^{24} slices, about 30 million times the number of nanoseconds in a year. Using the wasted cycles of most of the world's computers could probably do the problem with these methods in around a year. Finally, Simpson's method would require only about 10^{12} slices, the number of nanoseconds in 20 minutes. Earlham's computers could solve the problem with this method in an hour.

Another way to compare Simpson with our previous approaches is to say that even with 1 slice, Simpson is as accurate as the right-hand method with about 40 slices. With 1000 slices, Simpson is as good as the right-hand method with 4×10^{14} slices or the trapezoid method with 10^7 slices. It's a gigantic improvement, and we got it just by messing about.

6. What you discover if you apply Simpson's Rule to a quadratic (or, indeed, to a cubic) polynomial is that Simpson gives the exact answer, regardless of the value of n. This is sort of like the observation that the Midpoint and Trapezoid methods give exactly correct answers for linear functions, and that the Right and Left Endpoint methods are exact for constants.

One way to argue for this fact is to show that it works for a single slice, since if Simpson is exact for every single slice, then it is exact for the whole interval. To make the arithmetic easier, let's suppose the slice we are looking at is $[-1, 1]$. Surely if Simpson is exact on this interval for every parabola, however thin or fat and centered wherever, then it will be exact for every parabola on every interval.

So we're trying to approximate the area under $f(x) = ax^2 + bx + c$ on $[-1, 1]$. The left hand rectangle has area

$$f(-1) \cdot 2 = 2(a - b + c),$$

and the right hand rectangle has area

$$f(1) \cdot 2 = 2(a + b + c);$$

so the Trapezoid Rule gives us the average of these two areas,

$$T = 2a + 2c.$$

The Midpoint Rule gives an area

$$M = f(0) \cdot 2 = 2c.$$

Simpson therefore gives the approximation

$$S = \frac{2M + T}{3} = \frac{2a + 6c}{3} = \tfrac{2}{3}a + 2c.$$

The exact value of the integral over this slice is

$$\begin{aligned}\int_{-1}^{1} (ax^2 + bx + c)\, dx &= \left[\tfrac{1}{3}ax^3 + \tfrac{1}{2}bx^2 + cx\right]_{-1}^{1} \\ &= \left[\tfrac{1}{3}a + \tfrac{1}{2}b + c\right] - \left[-\tfrac{1}{3}a + \tfrac{1}{2}b - c\right] = \tfrac{2}{3}a + 2c.\end{aligned}$$

Thus, Simpson gives the exact value for the integral of a quadratic.

This gives us a useful and more theoretically grounded way to think about Simpson's Rule. The Right and Left Endpoint Rules tell us to approximate f on each slice by a constant. They will give exact values for constants, but will give an error if f departs from being a constant. The Trapezoid and Midpoint Rules tell us to approximate f on each slice by a straight line of any slope—the line joining the endpoints of the interval for the Trapezoid Rule, and the tangent line at the midpoint for the Midpoint Rule. They will give exact values for every line, but will give an error if the graph of f departs from being a line. Normally this results in much smaller errors in each slice than the Endpoint Rules give, since most functions are closer to some line than they are to being constant.

Since Simpson is exact for parabolas, Simpson is equivalent to saying that in each slice, we fit a parabola through the three points on the graph at the ends an the middle of that slice, and we then compute the exact area under the parabola. Adding these up gives Simpson's Rule. Simpson will therefore be exact for every parabola, but will give an error if the graph of f departs from being a parabola. Normally this results in even smaller errors in each slice than the Midpoint and Trapezoid Rules, since parabolas can approximate functions more closely than lines can.

This suggests where one might go if one wanted even more accurate numerical methods of integration than Simpson's Rule. We could pick more points in each slice, fit a polynomial of higher degree than 2 to the curve using these points, take the exact area under the polynomial, and add these up for every slice. This is what *Maple* does when computing numerical integrals.

Once again, I'm impressed at how much we have been able to learn by just messing about and thinking about what we can learn by just messing about and paying attention to what we see.