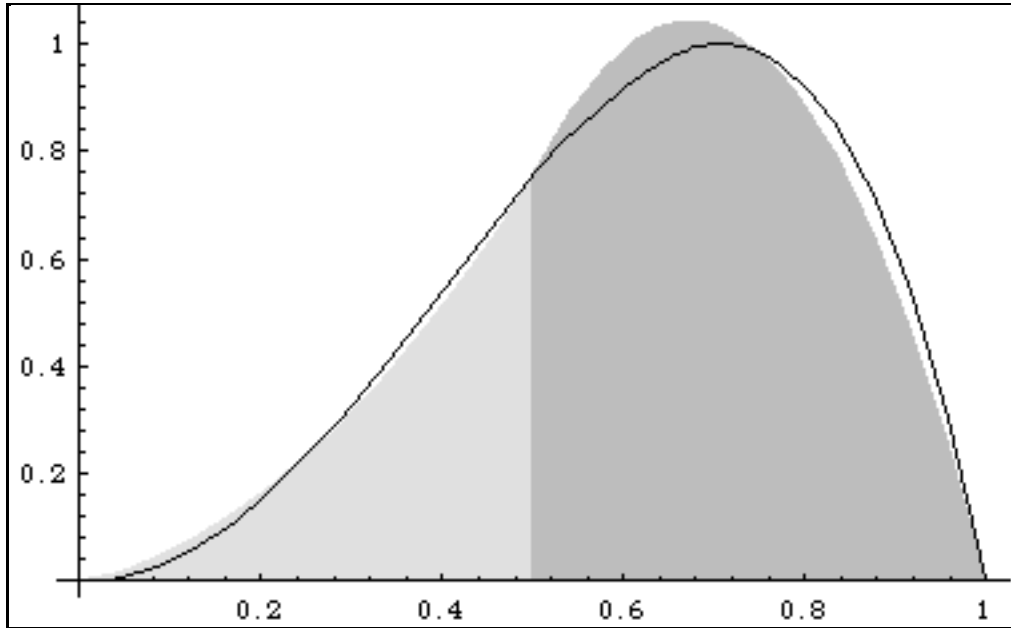


Integration 2



The METRIC Project, Imperial College.

© Imperial College of Science Technology and Medicine, 1996.

Instructions for Getting Started

Launch *Mathematica*.

Type

```
<<Mathetic`intpack`
```

hold down the “shift” key, and press the “return” key. Wait for *Mathematica*’s response. (Note: be sure to use the ` symbol rather than the '. You may need to hunt for it on your keyboard: on most, it's in the top left corner.)

This essential first step sets up *Mathematica* for this module. If you omit this bit, the “special commands” (see below) will not work.

Mathematica Commands

The following *Mathematica* commands may be useful to you in this module.

Commands that come with *Mathematica*:

```
Integrate, Clear, Solve, Simplify
```

Special commands for this module:

```
IntegralPlot, TrapeziumRule, SimpsonsRule, GiveQuestion,  
LastAnswer
```

For information on, say, the **Integrate** command, type

```
?Integrate
```

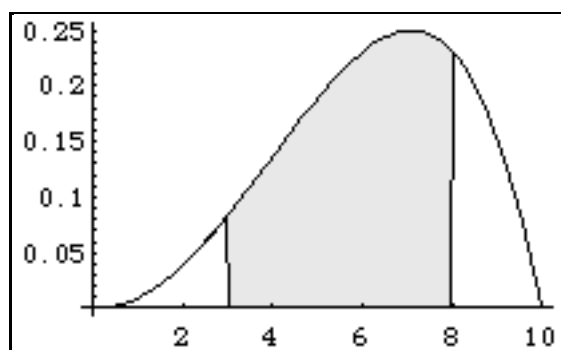
hold down “shift”, and press “return”.

Experiment 1: Definite integrals

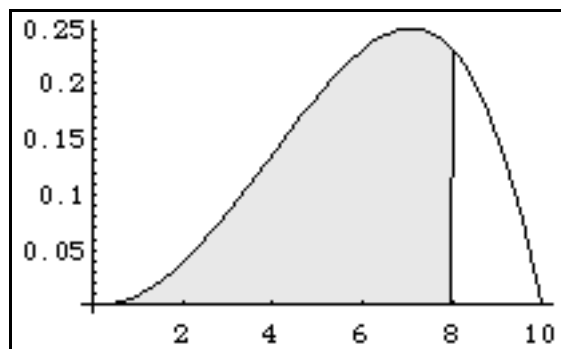
Preparatory reading

If you have looked at the *Integration 1* module you may recall the close connection between the idea of integration and the calculation of areas under curves. The first module's treatment of symbolic, indefinite integration will now be of use to us in devising a method — an *exact* one, not an approximation — for finding such areas. The key ideas are these:

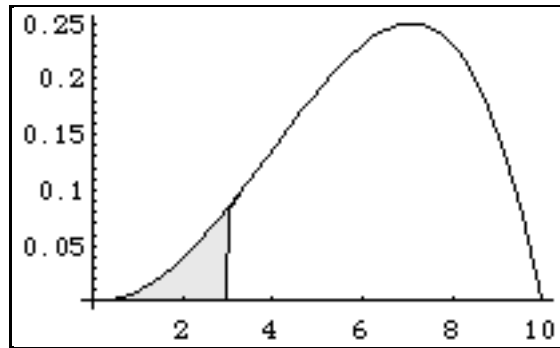
1. The indefinite integral is really an “area so far” function, which tells us the area under the curve “up to” a certain value of x . (It includes an arbitrary constant whose value depends on what point we start measuring from.)
2. We might typically be interested in finding the area under a certain curve (what's known as the **definite integral** of the function) between, say, $x = 3$ and $x = 8$.



3. We can get, from the indefinite integral, an expression for the area “up to” $x = 8$...



...and for the area “up to” $x = 3$:



A way of using these two expressions may already have occurred to you.

- 1) Type the following input to set up an “area so far” function (see the Preparatory Reading):

```
integf[x_] = Integrate[E^(2x), x]
```

Evaluate **integf[1.2]** and **integf[2.7]**. Use your answers to find the area under the graph of the function $y = e^{2x}$ between $x = 1.2$ and $x = 2.7$.

Using a similar technique, find the area under the graph of the function $y = \cos x$ between $x = 0$ and $x = \pi/3$, and the area under the graph of $y = 1/x$ between $x = 1$ and $x = 4$. Try some of your own devising.

- 2) Compare your findings in part 1 with the output from the following commands:

```
Integrate[E^(2x), {x, 1.2, 2.7}]
```

```
Integrate[Cos[x], {x, 0, Pi/3}]
```

```
Integrate[1/x, {x, 1, 4}]
```

Describe the effect of the **Integrate** command used in this way.

- 3) This module incorporates a special function called **IntegralPlot**. Type the following input:

```
IntegralPlot[E^(2x), {x, 1.2, 2.7}]
```

Experiment further, then write a brief description of the **IntegralPlot** command and what it does.

*You can use most of the options you associate with the **Plot** command when using **IntegralPlot**. For example, try adding the option **PlotRange -> {{0, 3}, {0, 200}}** to the previous command.*

- 4) Write down what you expect to be the output from the command

```
Integrate[Sin[x], {x, 0, 2 Pi}]
```

Test your answer.

Generate an **IntegralPlot** for this calculation.

Write a brief explanation of what you have observed in part 4 of this experiment.

5) Calculate the value of the total area enclosed by the graph of $y = \sin x$ and the x -axis between $x = 0$ and $x = 2\pi$. Explain why this is *not* the same as your result from part 4.

6) Type the following:

```
Integrate[1/x, {x, -2, 2}]
```

Write a brief explanation of why you get the output you do.

Note: this section uses this module's special functions. If they fail to work, try going back to the Instructions for Getting Started at the beginning.

Post-experiment reading

A definite integral calculation is set out formally as follows:

$$\begin{aligned}\int_1^2 e^{2x} dx &= \left[\frac{1}{2} e^{2x} \right]_1^2 \\ &= \frac{1}{2} e^4 - \frac{1}{2} e^2 \\ &= \frac{1}{2} e^2 (e^2 - 1).\end{aligned}$$

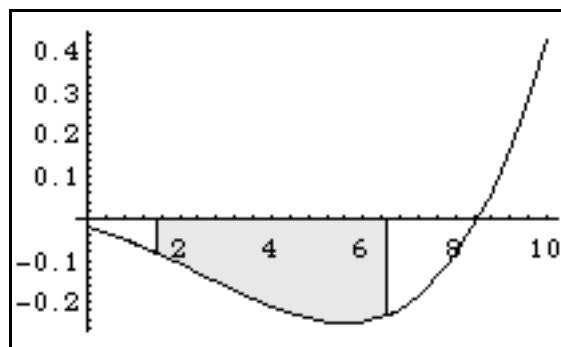
It's important to note the following things.

(i) The definite integral notation: $\int_a^b f(x) dx$ means “the definite integral of $f(x)$ between $x = a$ and $x = b$ ”. The numbers a and b are called the **limits of integration**.

(ii) The square bracket notation: $[F(x)]_a^b$ is shorthand for “ $F(b) - F(a)$ ”. We just evaluate what's inside the square brackets at each of the limits of integration, then subtract one from the other.

(iii) We don't include the constant of integration (if we did, it would cancel anyway).

(iv) Where the curve dips below the x -axis (in other words, in those regions for which the function is negative), the integral regards enclosed areas above the x -axis as positive and below as negative. (So for sections of the curve below the x -axis, it is incorrect to speak of the area ‘under’ the curve.)



This explains the result you probably got when you integrated $\sin x$ between 0 and 2π .

It's not at first obvious why *Mathematica* had so much trouble integrating $1/x$ between $x = -2$ and $x = 2$. It looks OK to reason as follows:

$$\begin{aligned}\int_{-2}^2 \frac{1}{x} dx &= [\ln|x|]_{-2}^2 \\ &= \ln|2| - \ln|-2| \\ &= \ln 2 - \ln 2 \\ &= 0.\end{aligned}$$

Here's the problem, though: the **domain of integration** — the set of numbers between -2 and 2 — clearly includes the number zero, where the functions $1/x$ and $\ln x$ are not defined. The number zero is called a **singularity** of the integral.

Singularities aren't always disastrous, but they can be, as here. The integral $\int_{-2}^2 \frac{1}{x} dx$ is said to be **undefined**.

Practice Questions

We've included a feature which allows you to get *Mathematica* to generate *practice questions* and their answers. To generate a question on definite integration, type

```
GiveQuestion["definite integration"]
```

not forgetting to “shift-return”. To generate the answer for checking, type

```
LastAnswer["definite integration"]
```

You can do this as often as you want: the questions are randomly generated, and repetitions should be rare. [Note: you do not need to retype the commands. Simply click on what you have already typed, and then “shift-return”.]

Note: this section uses this module's special functions. If they fail to work, try going back to the Instructions for Getting Started at the beginning.

Experiment 2: The trapezium rule

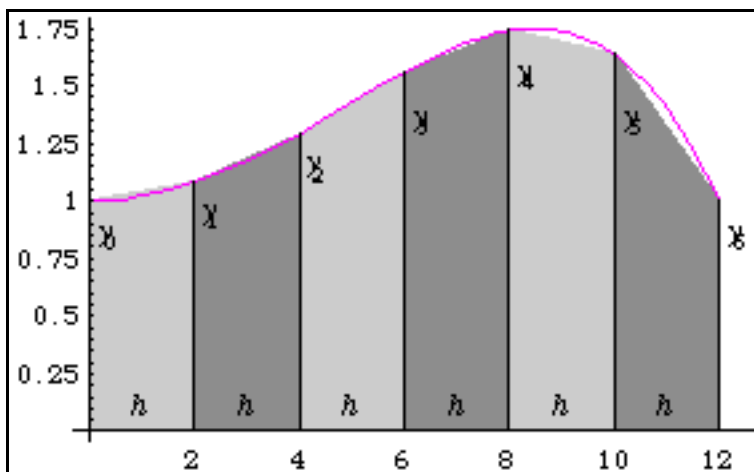
Preparatory reading

Using the indefinite, symbolic, integral to deduce actual areas (definite integrals) is the ideal method when it works, because the areas we get are exact. Symbolic integration is notoriously hard, though; functions don't have to get all that complicated before we run out of techniques for integrating them symbolically.

Instead, we often fall back on the approximate methods, such as the Riemann approximation and the trapezium rule, that you have briefly met if you have worked in the *Integration 1* module. The Riemann approximation is, as we've said elsewhere, of theoretical rather than practical interest; in practice, we generally want the trapezium rule or something better.

In this experiment you study the trapezium rule in more detail. As you'll recall, we split the area we want to find into a number of strips, each of which we think of as being a trapezium (they're not: that's where the approximation comes in).

Suppose the width of each trapezium is h , and the y -values (**ordinates**) at each trapezium boundary are y_0, y_1, y_2 , etc, as shown in the diagram below.



Then the area of the first trapezium is, by the standard formula, $\frac{h}{2}(y_0 + y_1)$.

That of the second is $\frac{h}{2}(y_1 + y_2)$, that of the third $\frac{h}{2}(y_2 + y_3)$, and so on.

So, if there are n trapezia (and therefore $n + 1$ ordinates, from y_0 up to y_n) then the total area of all the trapezia is

$$\begin{aligned} & \frac{h}{2}(y_0 + y_1) + \frac{h}{2}(y_1 + y_2) + \frac{h}{2}(y_2 + y_3) + \cdots + \frac{h}{2}(y_{n-1} + y_n) \\ &= \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-2} + 2y_{n-1} + y_n) \end{aligned}$$

This serves as an approximation for the true area.

It's a good idea, when doing trapezium rule calculations, to add up the numbers y_1 to y_{n-1} *before* doubling: that way, you only have to double once. Here's a worked example.

To estimate, using eleven ordinates (and therefore ten strips), the integral $\int_1^2 \frac{1}{x} dx$:

$$h = 0.1$$

x	$y = 1/x$	
1.0	1.000000	
1.1		0.909091
1.2		0.833333
1.3		0.769231
1.4		0.714286
1.5		0.666667
1.6		0.625000
1.7		0.588235
1.8		0.555556
1.9		0.526316
2.0	0.500000	
TOTALS:	1.500000	6.18771

$$\begin{aligned} \text{Area} &= \frac{0.1}{2}(1.5 + 2 \times 6.18771) \\ &= 0.693771 \end{aligned}$$

You will have already met the special command **TrapeziumRule** if you have worked on the *Integration 1* module. By using the option setting **TabulateQ->True**, we can generate, as well as the diagram and the output, a tabulated calculation of the type shown in the Preparatory Reading for this experiment. For example:

```
TrapeziumRule[1/x, {x, 1, 2}, 0.1,
  TabulateQ->True]
```

- 1) Use the above *Mathematica* command to estimate the area under the curve $y = 1/x$ between $x = 1$ and $x = 2$. Compare the printed table with the one that appears in the Preparatory Reading.
- 2) Based on the diagram, would you expect your trapezium rule approximation to be an overestimate or an underestimate of the true area? Write a brief explanation. Use **Integrate** to find this true value (remember, to express *Mathematica* output as a decimal number, you can use the **N** command).
- 3) Use **Integrate** to try to find the value of $\int_{-1}^1 e^{x^3-x^2} dx$.

Write a brief description of what has occurred, speculating about why.

Note: the above calculation may cause acute memory problems. Be prepared to quit and restart Mathematica afterwards if necessary.

- 4) Without the aid of *Mathematica*, use the trapezium rule with six ordinates (that is, with a strip width of 0.4) to estimate the value of the integral in part 3. Compare your calculations and your result with those produced by the **TrapeziumRule** command with the **TabulateQ** option set to **True**. Reduce the strip width and obtain a better approximation.
- 5) Try to find another function that *Mathematica* cannot integrate exactly. Use the trapezium rule, by hand, to estimate the definite integral of this function over a suitable domain. Check using the **TrapeziumRule** command. Reduce the strip width and obtain a better approximation.

Experiment 3: Parabolic segments

Preparatory reading

We can, of course, make the trapezium rule as accurate as we like by making h sufficiently small. However, it's possible, without changing the value of h , to make the same data work harder for us. The argument (and it's a clever one) goes like this.

The trapezium rule depends on our taking segments of the graph, and pretending that each is simpler than it is. In fact, we pretend that each is as simple as it possibly could be: a straight line segment. We'd lose some simplicity, but possibly gain some accuracy, if we pretended instead that each segment was the second simplest thing it could possibly be.

The next simplest function after a linear one is a quadratic, and in the same way the next simplest curve after a straight line segment is a segment of a quadratic curve — what's called a *parabolic segment*.

That's the strategy, then: approximate the curve, not by a sequence of straight line segments, but by a sequence of parabolic ones.

- 1) By typing the following input, define a quadratic function $p(x)$, with all its coefficients unknown:

```
Clear[a, b, c, p, x]
p[x_] = a x^2 + b x + c
```

- 2) Suppose you are informed that the graph of $p(x)$ passes through the points $(-1, 5)$ and $(0, 1)$. Is this sufficient information to determine the values of a , b , and c ? Briefly note down your thoughts on this issue. Test your ideas by typing the following input:

```
Solve[ {p[-1] == 5, p[0] == 1}, {a, b, c}]
```

- 3) Suppose you are now informed that the graph of $p(x)$ passes through the points $(-1, 5)$, $(0, 1)$ and $(1, -1)$. Do you now have sufficient information to determine the values of a , b , and c ? Use the **Solve** command to test your ideas.

- 4) By typing the following, find, in terms of h , y_0 , y_1 and y_2 the values of a , b and c given that $p(x)$ passes through $(-h, y_0)$, $(0, y_1)$ and (h, y_2) :

```
solutions =
```

```
Solve[{p[-h] == y0, p[0] == y1, p[h] == y2},
      {a, b, c}]
```

Find simplified forms of these solutions by typing

```
solutions = Simplify[solutions]
```

Then remove the extra pair of curly brackets by typing

```
solutions = Flatten[solutions]
```

5) Set a , b and c equal to these values by typing

```
{a, b, c} = {a, b, c} /. solutions
```

Now integrate $p(x)$ between $x = -h$ and $x = h$, and simplify your result using *Mathematica*.

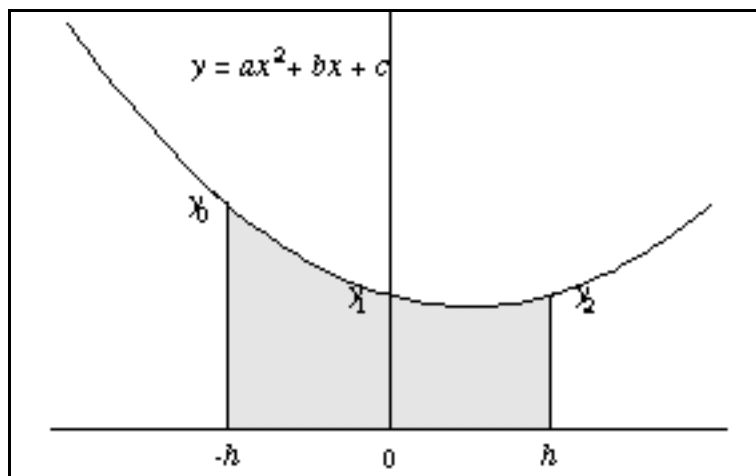
Make a note concerning the area under the parabola which passes through $(-h, y_0)$, $(0, y_1)$ and (h, y_2) between $x = -h$ and $x = h$.

Post-experiment reading

In this experiment you encountered the main drawback of trying to approximate a curve by a set of parabolic segments rather than a set of straight lines: namely, that each segment needs *three* points to specify it rather than two.

This leaves us with two options if we wish to use the “parabolic segments” idea to find approximate integrals. Either we sample the function in more places than we would for the trapezium rule, or we sample in the same number of places but settle for fewer segments. In practice, even if we take the second option we usually get a better approximation than the trapezium rule would give us.

This experiment seems to indicate that the area under the parabola that passes through $(-h, y_0)$, $(0, y_1)$ and (h, y_2) between $x = -h$ and $x = h$ is $\frac{h}{3}(y_0 + 4y_1 + y_2)$. This is illustrated on the next page:

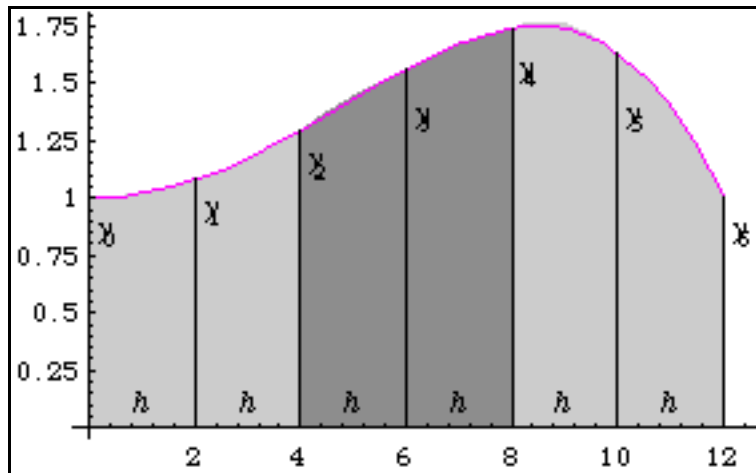


Now, the fact that, in the case we examined, the y -axis happened to be exactly in the middle is irrelevant. $\frac{h}{3}(y_0 + 4y_1 + y_2)$ is also the area under the parabola that passes through the points $(r - h, y_0)$, (r, y_1) and $(r + h, y_2)$:

Experiment 4: Simpson's rule

Preparatory reading

Presented with a function for which we wish to find an approximate direct integral, we divide the range of integration into an even number of strips, each of width h . This gives us half that number of parabolic segments, each of width $2h$, as shown:



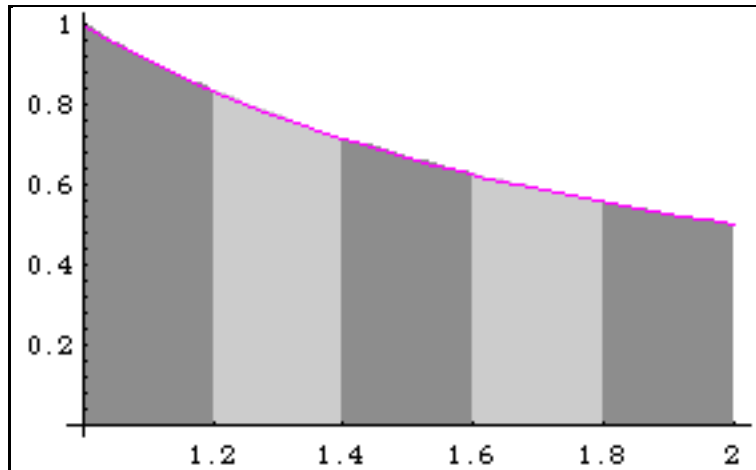
The area of the first pair of strips put together is approximately equal to that under the first parabolic segment, namely $\frac{h}{3}(y_0 + 4y_1 + y_2)$. The area of the second pair is, similarly, approximately $\frac{h}{3}(y_2 + 4y_3 + y_4)$. That of the third pair is approximately $\frac{h}{3}(y_4 + 4y_5 + y_6)$, and so on.

The integral, then, is approximately

$$\begin{aligned} & \frac{h}{3}(y_0 + 4y_1 + y_2) + \frac{h}{3}(y_2 + 4y_3 + y_4) \\ & + \frac{h}{3}(y_4 + 4y_5 + y_6) + \cdots + \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n) \\ & = \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n). \end{aligned}$$

This is known as **Simpson's rule**. As with the trapezium rule, it's a good idea to add those ordinates which need to be multiplied by 2 or 4 first, and then perform the multiplication.

Here is a worked example. To estimate the integral $\int_1^2 \frac{1}{x} dx$, using *eleven* ordinates (and therefore ten strips, and a strip width of 0.1):



$$h = 0.1$$

x	$y = 1/x$		
1.0	1.000000		
1.1			0.909091
1.2		0.833333	
1.3			0.769231
1.4		0.714286	
1.5			0.666667
1.6		0.625000	
1.7			0.588235
1.8		0.555556	
1.9			0.526316
2.0	0.500000		
TOTALS:	1.500000	2.72817	3.45954

$$\begin{aligned} \text{Area} &= \frac{0.1}{3} (1.5 + 2 \times 2.72817 + 4 \times 3.45954) \\ &= 0.69315 \end{aligned}$$

The special command **SimpsonsRule** is in every respect similar to the **TrapeziumRule** command you met if you did Experiment 2. For example:

```
SimpsonsRule[1/x, {x, 1, 2}, 0.1,
TabulateQ->True]
```

- 1) Use the above *Mathematica* command to estimate the area under the curve $y = 1/x$ between $x = 1$ and $x = 2$. Compare the printed table with the one that appears in the Preparatory Reading.
- 2) Use **Integrate** to find the true value of the area (remember, to express *Mathematica* output as a decimal number, you can use the **N** command). Compare this true value with those given by (i) Simpson's rule and (ii) the trapezium rule with the same strip width.

3) Without the aid of *Mathematica*, use Simpson's rule with five ordinates (that is, with a strip width of 0.5) to estimate the value of $\int_{-1}^1 e^{x^3-x^2} dx$. Compare your calculations and your result with those produced by the **SimpsonsRule** command with the **TabulateQ** option set to **True**. Reduce the strip width and obtain a better approximation.

4) Calculate an alternative approximation to the above integral using **TrapeziumRule**. Compare your two approximate results with the output from the following command:

```
NIntegrate[Exp[x3-x2], {x, -1, 1}]
```

Write a brief account of what you believe the function of the built-in command **NIntegrate** to be.

5) Carry out some investigations of your own into the comparative performance of the trapezium rule, Simpson's rule and *Mathematica*'s **NIntegrate** command in calculating approximate integrals.

Briefly write up your findings. Include an account of the number of decimal places to which **NIntegrate** appears to be accurate, based on comparisons with **Integrate**. Include also a description of under what conditions the trapezium rule and Simpson's rule can be made to attain this accuracy.

*You may find that you want to inspect more digits than Mathematica usually shows you. If so, recall that you can use the **N** command to increase the number of digits in your output. Try, for example*

```
N[ NIntegrate[1/x, {x, 1, 2}], 20]
```

and

```
N[ Integrate[1/x, {x, 1, 2}], 20]
```

Post-experiment reading

The two approximate integration methods you've met can be applied to a far wider range of functions than can any symbolic technique, including the more sophisticated ones in the next module. There's a drawback, though, and it's an obvious one: because these methods are approximate, our results may not be entirely trustworthy. There will be, in each case, an **error**: a difference between our estimate and the true value of the integral.

Often — though not always — this error will be quite small. In almost all cases of practical importance, Simpson's rule outperforms the trapezium rule for the same number of ordinates, provided the strip width is small enough.

Mathematica's **NIntegrate** function uses a highly sophisticated range of numerical integration techniques. It usually outperforms both the trapezium rule and Simpson's rule, except where the strip width is extremely small. (A very small strip width can give rise to other kinds of accuracy problems, though, as well as making the execution time for **TrapeziumRule** and **SimpsonsRule** excessive.) **NIntegrate** is usually accurate to about fifteen decimal places. This makes it quite as good as **Integrate** for most practical applications of definite integration (with the advantage of working for a far larger class of functions and often using far less memory). Furthermore, it is possible to set options within **NIntegrate** which make it as accurate as you choose.

Practice Questions

There are two sets of questions on numerical integration, accessible by the commands

```
GiveQuestion["trapezium rule"]
```

and

```
GiveQuestion["Simpson's rule"]
```

Use **LastAnswer** in the normal way for the answers.