

## NUMERICAL METHODS COURSEWORK

### INFORMAL NOTES ON NUMERICAL INTEGRATION COURSEWORK

For this piece of coursework students must use the methods for numerical integration they meet in the Numerical Methods module to produce approximations to an area which cannot be calculated exactly using other techniques they know. Students need good knowledge of the methods themselves and the way that they behave in relation to error. By using a spreadsheet program students should be able to produce very accurate approximations to their chosen area and be able to judge how reliable their approximations are.

If students are doing coursework on a topic other than numerical integration this document should still prove useful

- as a guide for how to write the report,
- as a guide for ideas about error,
- to give you some idea of the level of detail required in this piece of coursework.

**IMPORTANT:** Students' chosen problem should not have an *analytical* solution. This means that it shouldn't be something that they can solve using the techniques they have learnt/will learn in pure maths modules. For example, they should/will know how to evaluate the following integrals from standard A-level pure maths work, so they shouldn't pick anything similar for your Numerical Methods coursework.

$$\int_1^2 x^2 dx, \int_3^8 (x^5 + 3x^2 - 1) dx, \int_0^{\pi/2} \sin(x) dx, \int_0^{\pi/2} (x \cos(x)) dx, \int_0^1 (e^x \sin x) dx$$

### MAKE SURE THAT STUDENTS CHECK WITH THEIR TUTOR THAT THEIR PROBLEM IS APPROPRIATE BEFORE THEY EMBARK ON THEIR COURSEWORK

This document contains

- details of the report student should write on their coursework
- a recap of key ideas from the Numerical Methods module relating to the behaviour of error in the various methods of numerical integration that students are likely to use in their coursework.

# The Report

### Phase 1 – Problem Specification

Make sure that you tell the reader *exactly* what you are attempting to do in your coursework. For example, if you are trying to approximate an integral make sure that you tell the reader between which values you are integrating – show this on a diagram. You should briefly explain why your problem cannot be solved using normal analytical pure maths methods. See above. This phase should be fairly short.

### Phase 2 – Strategy

Say what you are going to do to solve the problem. State which Numerical Methods you will use and the formulae involved in them. Where appropriate, you may wish to include a diagram to explain the formulae, but do not derive any of the standard formulae – you will earn no marks for deriving standard results.

### Phase 3 – Formula Application

Do this on a spreadsheet. Try to set it out as clearly as you possibly can, so that the reader can see how all of the cells relate to one another. Remember to include the version of the spreadsheet where formulae are displayed somewhere in your report (perhaps in an appendix or maybe next to the spreadsheet itself). Label cells where appropriate. Try to stick to a standard notation.

### Phase 4 – Use of technology

You will probably use Excel to devise your spreadsheets. You should say what its limitations are (e.g how many decimal places can it handle?). You should say how many decimal places your calculations are being made to.

### Phase 5 – Error Analysis

This is a very important section. You should try to give an estimate of how much error there is in your solution to the problem. You may wish to analyse differences between successive approximations and then the ratio of differences, in order to see how fast your series of approximations in tending to the true value; you can use this to give an estimate of the error involved (see notes below).

### Phase 6 – Interpretation

Relate the solution you have obtained back to the original problem. Maybe you will want to discuss how you could improve your estimate and how long this would take in terms the number of extra iterations required. If you think your results are as good as the calculating power of the spreadsheet will allow, you should say why you think this.

### Phase 7 – Oral Communication

You should be completely familiar with your work and able to discuss it with your teacher.

## Numerical Integration Coursework - Ideas

### A Reminder about Error

The following table looks at the midpoint rule, the trapezium law and Simpson's Law being used to estimate the integral  $\int_0^1 5x^4 dx$ . The exact value of this integral can be easily calculated to be 1.

The shaded columns are almost constant. This suggests that the error in the trapezium law and the midpoint rule is proportional to  $h^2$  and the error in Simpson's rule is proportional to  $h^4$ . In fact it can be proved that this is the case for most of the functions you will ever encounter (certainly for all the ones you will encounter in your A-level course).

x	f(x)
0	0
0.0625	7.6294E-05
0.125	0.0012207
0.1875	0.00617981
0.25	0.01953125
0.3125	0.04768372
0.375	0.09887695
0.4375	0.18318176
0.5	0.3125
0.5625	0.50056458
0.625	0.76293945
0.6875	1.11701965
0.75	1.58203125
0.8125	2.17903137
0.875	2.9309082
0.9375	3.86238098
1	5

		$h$	Error	Error /h	Error /h <sup>2</sup>	Error /h <sup>3</sup>	Error /h <sup>4</sup>
$M_2$	0.800781	0.5	0.19921875	0.398438	0.796875	1.59375	3.1875
$M_4$	0.948486	0.25	0.051513672	0.206055	0.824219	3.296876	13.1875
$M_8$	0.987015	0.125	0.012985229	0.103882	0.831055	6.64844	53.18752
$T_2$	1.40625	0.5	0.40625	0.8125	1.625	3.25	6.5
$T_4$	1.103516	0.25	0.103515625	0.414063	1.65625	6.625	26.5
$T_8$	1.026001	0.125	0.026000977	0.208008	1.664063	13.3125	106.5
$T_{16}$	1.006508	0.0625	0.006507874	0.104126	1.666016	26.65625	426.5
$S_2$	1.002604	0.5	0.002604167	0.005208	0.010417	0.020833	0.041667
$S_4$	1.000163	0.25	0.00016276	0.000651	0.002604	0.010417	0.041667
$S_8$	1.00001	0.125	0.000010172	0.0000814	0.000651	0.005208	0.041667

### Things to Notice

- the midpoint rule estimates are underestimates and the trapezium rule estimates are overestimates. There is a good reason for this, although sometimes it can be the other way around. This is very useful because it allows us to say with certainty what the true value of the integral is to a number of decimal places.
- Notice also that the error in a midpoint rule estimate is about half the size of the corresponding trapezium rule estimate.

## Coursework Notes



### Error in the Midpoint Rule

In the midpoint rule, error is proportional to  $h^2$ . In other words there is a constant  $k$  such that

$$\text{error} = kh^2$$

In general, for a given integral, if  $\text{error}_n$  is the absolute error in  $M_n$ , then there is a constant,  $k$ , such that  $\text{error}_n \approx kh^2$  (where  $h$  is the strip width corresponding to  $n$  strips). Therefore, if the midpoint rule with  $n$  strips has a strip width of  $h$ , then the midpoint rule with  $2n$  strips has a strip width of  $\frac{h}{2}$ , and

$$\frac{\text{error}_{2n}}{\text{error}_n} \approx \frac{k\left(\frac{h}{2}\right)^2}{kh^2} = \frac{1}{4}.$$

This means that halving  $h$ , or, equivalently, doubling  $n$  will reduce the error by a factor of  $\frac{1}{4} = 0.25$ . The "error multiplier" is 0.25. We say that the midpoint rule is a "second order method".

### Error in the Trapezium Law

Similarly, in the trapezium law, error is proportional to  $h^2$ . In other words there is a constant  $k$  such that

$$\text{error} = kh^2$$

This means that halving  $h$ , or, equivalently, doubling  $n$  will reduce the error by a factor of  $\frac{1}{4} = 0.25$ . The "error multiplier" is 0.25. We say that the trapezium law is a "second order method".

### Error in Simpson's Rule

Finally, in Simpson's Rule, error is proportional to  $h^4$ . In other words there is a constant  $k$  such that

$$\text{error} = kh^4$$

This means that halving  $h$ , or, equivalently, doubling  $n$  will reduce the error by a factor of  $\frac{1}{16} = 0.0625$ . The "error multiplier" is 0.0625. We say that Simpson's Rule is a "fourth order method".

## Viewing Error in Terms of Differences and Ratio of Differences

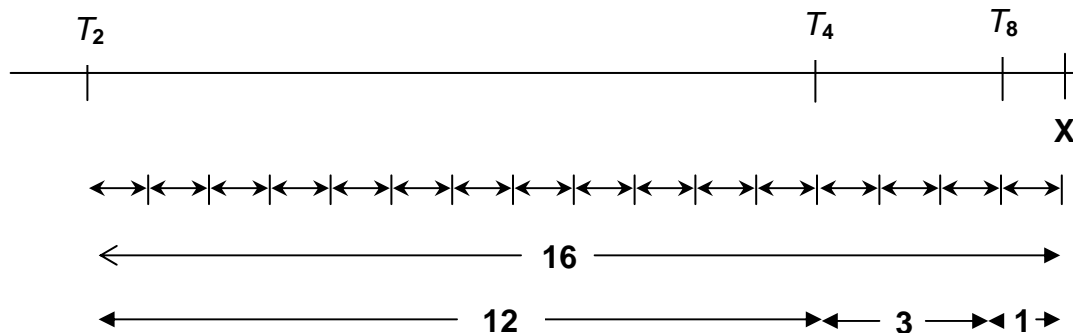
With doubling values of  $n$ , the "error multiplier" in each case above is the same as the "ratio of differences" between successive estimates. The table below confirms this for the particular integral we have been looking at.

$M_2$	0.800781	Differences	Ratio of
$M_4$	0.948486	0.147705078	Differences
$M_8$	0.987015	0.038528442	0.260847107
$T_2$	1.40625		
$T_4$	1.103516	-0.302734375	
$T_8$	1.026001	-0.077514648	0.256048387
$T_{16}$	1.006508	-0.019493103	0.251476378
$S_2$	1.002604		
$S_4$	1.000163	-0.002441406	
$S_8$	1.00001	-0.000152588	0.0625

### Why is this the case?

The following diagram shows why this is in the case of the trapezium law. Each successive approximation (doubling the number of strips) is four times as close to the true value  $X$  as the previous one. So the distance between  $T_8$  and  $X$  is one sixteenth of the distance between  $T_2$  and  $X$ . The diagram below illustrates this and also shows that the ratio  $T_8 - T_4 : T_4 - T_2$  is 3:12 or 1:4. So

$$T_8 - T_4 = 0.25 \times (T_4 - T_2)$$



So the ratio of differences is the same as the "error multiplier". We can use this to our advantage..... Let's change integral and look at  $\int_1^2 \sqrt{\sin x} dx$ .

# Coursework Notes



## Getting "better estimates" by extrapolation using the ratio of differences.

Here we have assumed that the "ratio of differences" would have continued to be 0.25 had we calculated further trapezium or midpoint estimates. We have worked backwards to see what the differences and then the estimates themselves would have had to have been. Can you see what we have done? The normal font for T and M represents actual calculated values, the italic font and grey background represents extrapolated values.

x	f(x)
1	0.917317
2	0.953571

1	0.917317
1.5	0.998747
2	0.953571

1	0.917317
1.25	0.974158
1.5	0.998747
1.75	0.991961
2	0.953571

1	0.917317
1.125	0.949878
1.25	0.974158
1.375	0.9904
1.5	0.998747
1.625	0.999265
1.75	0.991961
1.875	0.976773
2	0.953571

1	0.917317
1.0625	0.934652
1.125	0.949878
1.1875	0.963035
1.25	0.974158
1.3125	0.983273
1.375	0.9904
1.4375	0.995555
1.5	0.998747
1.5625	0.999983
1.625	0.999265
1.6875	0.996593
1.75	0.991961
1.8125	0.985359
1.875	0.976773
1.9375	0.966185
2	0.953571

	Trapezium Estimates	Differences $T_{2i}-T_i$	Ratio of Differences
$T_1$	0.9354441		
$T_2$	0.9670954	0.03165131	
$T_4$	0.9750775	0.00798207	0.25218769
$T_8$	0.9770783	0.00200085	0.25066821
$T_{16}$	0.9775789	0.00050057	0.25017771
" $T_{32}$ "	0.977704	0.00012514	0.25
" $T_{64}$ "	0.9777353	3.1286E-05	0.25
" $T_{128}$ "	0.9777431	7.8214E-06	0.25
" $T_{256}$ "	0.9777451	1.9553E-06	0.25
" $T_{512}$ "	0.9777456	4.8884E-07	0.25
" $T_{1024}$ "	0.9777457	1.2221E-07	0.25

	Midpoint Estimates	Differences $M_{2i}-M_i$	Ratio of Differences
$M_1$	0.9987467		
$M_2$	0.9830595	-0.0156872	
$M_4$	0.9790792	-0.0039804	0.25373399
$M_8$	0.9780795	-0.0009997	0.25116135
" $M_{16}$ "	0.9778295	-0.0002499	0.25
" $M_{32}$ "	0.977767	-6.248E-05	0.25
" $M_{64}$ "	0.9777514	-1.562E-05	0.25
" $M_{128}$ "	0.9777475	-3.905E-06	0.25
" $M_{256}$ "	0.9777465	-9.763E-07	0.25
" $M_{512}$ "	0.9777463	-2.441E-07	0.25
" $M_{1024}$ "	0.9777462	-6.102E-08	0.25
" $M_{2056}$ "	0.9777462	-1.525E-08	0.25

Here the midpoint rule and trapezium rule are being used to estimate

$$\int_1^2 \sqrt{\sin x} dx .$$

### Notice

- The strip width is halving each time and the ratio of differences is close to 0.25. This illustrates further that these are both second order methods.

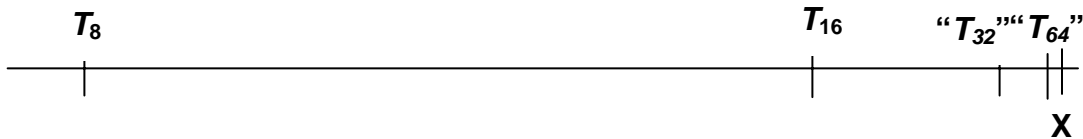
# Coursework Notes



The picture below shows some of the formulae entered to generate the spreadsheet above.

	A	B	D	E	F	G	H
1					Trapezium	Differences	Ratio of
2	x	f(x)			Estimates	$T_{2i} - T_i$	Differences
3	1	=SQRT(SIN(A3))		$T_1 =$	=0.5*(B1+B2)		
4	2	=SQRT(SIN(A4))		$T_2 =$	=0.25*(B6+2*B7+F6+F6+B8)	=F4-F3	
5				$T_4 =$	=0.125*(B10+2*(B11+B12+B13)+B14)	=F5-F4	=G5/G4
6	1	=SQRT(SIN(A6))		$T_8 =$	=0.0625*(B16+2*SUM(B17:B23)+B24)	=F6-F5	=G6/G5
7	1.5	=SQRT(SIN(A7))		$T_{16} =$	=0.03125*(B26+2*SUM(B27:B41)+B42)	=F7-F6	=G7/G6
8	2	=SQRT(SIN(A8))		" $T_{32}$ "	=F7+G8	=G7*H8	0.25
9				" $T_{64}$ "	=F8+G9	=G8*H9	0.25
10	1	=SQRT(SIN(A10))		" $T_{128}$ "	=F9+G10	=G9*H10	0.25
11	1.25	=SQRT(SIN(A11))		" $T_{256}$ "	=F10+G11	=G10*H11	0.25
12	1.5	=SQRT(SIN(A12))		" $T_{512}$ "	=F11+G12	=G11*H12	0.25
13	1.75	=SQRT(SIN(A13))		" $T_{1024}$ "	=F12+G13	=G12*H13	0.25
14	2	=SQRT(SIN(A14))					
15					Midpoint	Differences	Ratio of
16	1	=SQRT(SIN(A16))			Estimates	$M_{2i} - M_i$	Differences
17	1.125	=SQRT(SIN(A17))		$M_1 =$	=B7		

## Extrapolation in terms of a diagram and geometric progressions



According to the theory derived earlier  $T_{32} \approx T_{16} + \frac{1}{4}(T_{16} - T_8)$

This gives us the so called "extrapolated" value " $T_{32}$ " =  $T_{16} + \frac{1}{4}(T_{16} - T_8)$ .

Note, this is exactly how " $T_{32}$ " was calculated on the previous page.

And then

$$"T_{64}" = "T_{32}" + \left(\frac{1}{4}\right)^2 (T_{16} - T_8) = T_{16} + \frac{1}{4}(T_{16} - T_8) + \left(\frac{1}{4}\right)^2 (T_{16} - T_8)$$

This is exactly how " $T_{64}$ " was calculated on the previous page. We can obviously continue to do this as many times as we like.

In fact if we do this several times (infinitely many times in fact) we are reminded of the expression below which involves a geometric progression.

## Coursework Notes



$$T_{16} + \frac{1}{4}(T_{16} - T_8) + \left(\frac{1}{4}\right)^2 (T_{16} - T_8) + \left(\frac{1}{4}\right)^3 (T_{16} - T_8) + \dots$$

This a geometric progression with  $a = \frac{T_{16} - T_8}{4}$  and

$r = \frac{1}{4}$ . Its sum is  $\frac{a}{1-r}$  which in this case is

$$\frac{\frac{T_{16} - T_8}{4}}{\frac{3}{4}} = \frac{T_{16} - T_8}{3}$$

Therefore the value above is  $T_{16} + \frac{T_{16} - T_8}{3} = \frac{4T_{16} - T_8}{3}$ . Similar formulae can be derived using midpoint rule and Simpson's rule estimates. These are all given on the following page along with all the other numerical integration formulae you'll need.

# Coursework Notes



## NUMERICAL METHODS – INTEGRATION FORMULAE

Formula	When to use
<p>Midpoint Rule:</p> $M_n = h(f(m_1) + f(m_2) + \dots + f(m_n)), \text{ where } h = \frac{b-a}{n}.$	Obvious
<p>Trapezium Law:</p> $T_n = \frac{h}{2}[f_0 + 2(f_1 + f_2 + f_3 + \dots + f_{n-1}) + f_n], \text{ where } h = \frac{b-a}{n}.$	Obvious
<p>Simpson's Law:</p> $S_n = \frac{2M_n + T_n}{3}$	When you have $M_n$ and $T_n$ this is the quickest way to get $S_n$
<p>Simpson's Law:</p> $S_n = \frac{h}{3}[f_0 + 4(f_1 + f_3 + f_5 + \dots + f_{2n-1}) + 2(f_2 + f_4 + \dots + f_{2n-2}) + f_{2n}] \text{ where}$ $h = \frac{b-a}{2n}.$	Not often used, this is another way to get $S_n$ , but often it is quicker to calculate $M_n$ and $T_n$ and then use the previous formulae
$T_{2n} = \frac{T_n + M_n}{2} \text{ for example, } T_8 = \frac{T_4 + M_4}{2}$	A quick way to get $T_{2n}$ when you have $M_n$ and $T_n$
$S_n = \frac{4T_{2n} - T_n}{3}, \text{ for example } S_{16} = \frac{4T_{32} - T_{16}}{3}$	A quick way to get $S_n$ when you have $T_{2n}$ and $T_n$
<p>Best estimate based on calculated values <math>M_n</math> and <math>M_{2n}</math> obtained from a G.P.</p> <p>"Best midpoint" = <math>\frac{4M_{2n} - M_n}{3}</math>, e.g. <math>\frac{4M_{16} - M_8}{3}</math></p>	Use this for a quick best estimate when you have calculated two successive midpoint values like $M_8$ and $M_{16}$ for example.
<p>Best estimate based on calculated values <math>T_n</math> and <math>T_{2n}</math> obtained from a G.P.</p> <p>"Best trapezium" = <math>\frac{4T_{2n} - T_n}{3}</math>, e.g. <math>\frac{4T_8 - T_4}{3}</math></p>	Use this for a quick best estimate when you have calculated two successive trapezium values like $T_8$ and $T_{16}$ for example.
<p>Best estimate based on calculated values <math>S_n</math> and <math>S_{2n}</math> obtained from a G.P.</p> <p>"Best Simpson's" = <math>\frac{16S_{2n} - S_n}{15}</math>, e.g. <math>\frac{16S_8 - S_4}{15}</math></p>	Use this for an absolute best estimate when you have calculated two successive Simpson values like $S_8$ and $S_{16}$ for example.

# Coursework Notes

