# Physics 115/242

## Romberg Integration

Peter Young

In this handout we will see how, starting from the trapezium rule, we can obtain much more accurate values for the integral by repeatedly eliminating the leading contribution to the error. This is known as Romberg integration.

As already discussed in class, the trapezium rule is a simple numerical approximation to the integral

$$\boxed{I = \int_a^b f(x)\, dx.} \tag{1}$$

For reasons that will become clear later we will write the trapezium rule for $n$ intervals as $I_n^{(0)}$, i.e.

$$\boxed{I_n^{(0)} \equiv T_n = h\left[\tfrac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{n-1} + \tfrac{1}{2}f_n\right],} \tag{2}$$

where

$$h = \frac{b-a}{n} \tag{3}$$

is the width of one interval, $f_i \equiv f(x_i)$ with $x_i = x_0 + i\,h$, and $x_0 = a, x_n = b$. We have seen in class that, for small $h$, the error is proportional to $h^2$. It turns out that if one writes an expression for the error as a power series in $h$ then only even powers of $h$ appear (assuming that the integrand doesn't have any singularities). In other words

$$I = I_n^{(0)} + Ah^2 + Bh^4 + Ch^6 + \cdots , \tag{4}$$

where $A, B, C, \cdots$ are numbers which are related to derivatives of $f(x)$ at the endpoints (which of course we don't know in the numerics). We already showed that the leading error in the trapezium rule is of order $h^2$, i.e. there is no term linear in $h$. Equation (4), with explicit expressions for $A, B, C, \cdots$, is called the Euler-Maclaurin formula. Derivations, which are moderately complicated and involve "Bernoulli polynomials", are given in some of the more advanced books, and from a Google search I found derivations on the web at

`http://en.wikipedia.org/wiki/Euler-Maclaurin_formula`

and

`http://planetmath.org/encyclopedia/ProofOfEulerMaclaurinSummationFormula.html`.

A someone different derivation which avoids Bernoulli polynomials is given by Howie Haber in his notes for Physics 116A:
`http://scipp.ucsc.edu/ haber/ph116A/esum.pdf`.
Here, we don't need the expressions for $A, B, C$, etc. just the result that *only even powers* of $h$ arise in Eq. (4). I feel there should be a *simple* derivation of just this, but I haven't been able to find one.

To obtain a more accurate estimate for $I$ we will *eliminate the leading contribution to the error* in Eq. (4), the term of order $h^2$, by taking $n$ to be even and determining the trapezium rule for $n/2$ intervals as well as for $n$ intervals. Since the width of one interval is now $2h$ we have

$$I_{n/2}^{(0)} = 2h\left[\tfrac{1}{2}f_0 + f_2 + f_4 \cdots + f_{n-2} + \tfrac{1}{2}f_n\right]. \tag{5}$$

The error in this expression is of the same form as Eq. (4) but with $h$ replaced by $2h$, i.e.

$$I = I_{n/2}^{(0)} + A(2h)^2 + B(2h)^4 + C(2h)^6 + \cdots. \tag{6}$$

We can eliminate the leading contribution to the error by subtracting 4 times Eq. (4) from Eq. (6). Dividing by 3 we get

$$I = \frac{4I_n^{(0)} - I_{n/2}^{(0)}}{3} - 4Bh^4 - 20Ch^6 + \cdots. \tag{7}$$

Hence we obtain the estimate $I_n^{(1)}$, where

$$\boxed{I_n^{(1)} = \frac{4I_n^{(0)} - I_{n/2}^{(0)}}{3}.} \tag{8}$$

From Eq. (7) this is related to the exact value of the integral $I$ by

$$I = I_n^{(1)} + B'h^4 + C'h^6 + \cdots, \tag{9}$$

where $B' = -4B$ and $C' = -20C$.

From Eqs.(2), (5) and (8) we find that $I_n^{(1)}$ is given in terms of the function values by

$$I_n^{(1)} = \frac{h}{3}\left[f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \cdots 2f_{n-2} + 4f_{n-1} + f_n\right], \tag{10}$$

which is just *Simpson's rule* for $n$ intervals, i.e.

$$\boxed{I_n^{(1)} \equiv S_n} \tag{11}$$

Hence the first level of error elimination in Romberg integration gives Simpson's rule. However, the higher levels of error elimination that we discuss next will not produce familiar approximations[1].

If $n$ is a multiple of 4, we can clearly repeat the trick of eliminating the leading contribution to the error, which is now the $B'h^4$ term in Eq. (9), by forming the appropriate linear combination of $I_n^{(1)}$ and $I_{n/2}^{(1)}$. Hence the second level of error extrapolation gives $I_n^{(2)}$ where

$$I_n^{(2)} = \frac{16 I_n^{(1)} - I_{n/2}^{(1)}}{15}, \tag{12}$$

which has an error of order $h^6$. In general, if $n$ is a multiple of $2^k$ we can form iteratively a sequence of higher order approximations, $I_n^{(k)}$, where

$$\boxed{I_n^{(k)} = \frac{4^k I_n^{(k-1)} - I_{n/2}^{(k-1)}}{4^k - 1},} \tag{13}$$

for $k = 1, 2, 3, \cdots$, which have an error of order $h^{2k+2}$.

Romberg integration is a convenient way of carrying out this procedure of successively eliminating terms in the expression for the error. One proceeds as follows. Start with 1 interval and compute $I_1^{(0)}$. Then do two intervals and compute $I_2^{(0)}$ from which $I_2^{(1)}$ can be determined using Eq. (8). Then double the number of intervals again to 4 and calculate $I_4^{(0)}$, from which $I_4^{(1)}$ can be obtained using Eq. (8), and hence $I_4^{(2)}$ obtained from Eq. (12). The sequence of quantities thus obtained can be conveniently represented as a table,

| | $T_n$ | $S_n$ | | | |
|---|---|---|---|---|---|
| n k $\rightarrow$ | 0 | 1 | 2 | 3 | $\ldots$ |
| $\downarrow$ | | | | | |
| 1 | $I_1^{(0)}$ | | | | |
| 2 | $I_2^{(0)}$ | $I_2^{(1)}$ | | | |
| 4 | $I_4^{(0)}$ | $I_4^{(1)}$ | $I_4^{(2)}$ | | |
| 8 | $I_8^{(0)}$ | $I_8^{(1)}$ | $I_8^{(2)}$ | $I_8^{(3)}$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ |

.

[1] Note quite true; the next level, $I_n^{(2)}$ in Eq. (12), is sometimes called Bode's rule (see Numerical Recipes Sec. 4.1), which is a "Newton-Cotes" formula with 5 points. In Newton-Cotes formulae one fits a polynomial of order $n = N - 1$ through $N$ equally spaced points and integrates the polynomial. For $n = 1$ this gives the trapezoidal rule, for $n = 2$ it gives Simpson's rule, and for $n = 4$ it gives Bode's rule. Newton-Cotes formulae for larger $n$ are not very useful (for example some of the coefficients are negative for $n \geq 8$ which enhances roundoff errors), and do not correspond to higher order Romberg approximations.

Recall that $n$ is the number of intervals used in the trapezium rule, and $k$ is the number of levels of error elimination. The left hand column of estimates for the integral, labeled 0, contains results for the trapezium rule determined by evaluating the function at appropriate points. Each subsequent column is obtained, not from any more function evaluations, but rather by manipulating the data in the column to the left of it. These columns contain results for a higher order approximation so, for example, the $k = 1$ column is Simpson's rule. Note that the function values obtained for the trapezium rule with $n$ intervals also occur in the trapezium rule for $2n$ intervals, and so would *not be recomputed*, in a well written code when evaluating $T_{2n}$ ($\equiv I_{2n}^{(0)}$).

The number of number of intervals is given by $n = 2^k$ and the most accurate result is generally the one with the most error extrapolations, i.e. the diagonal entry $I_{2^k}^{(k)}$. We therefore define successive Romberg estimates by

$$\boxed{R_k = I_{2^k}^{(k)}.} \tag{14}$$

If the desired accuracy is $\epsilon$ then, as an empirical rule, we keep doubling the number of intervals until

$$\boxed{|R_k - R_{k-1}| < \epsilon,} \tag{15}$$

and take $R_k$ to be the estimate for the integral $I$. Frequently the error, $|I - R_k|$, turns out to be *much* smaller than $\epsilon$.

As an example consider

$$I = \int_1^2 \frac{1}{x^2}\, dx = 1/2. \tag{16}$$

Using Romberg integration I obtain the following table of values. Remember, the $k = 0$ column contains the numerical results from the trapezium rule and the other columns are obtained by manipulating the data in the $k = 0$ column. The results in the $k = 1$ column, obtained in this way, happen to correspond to Simpson's rule.

| $n$ | $k = 0$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | 0.62500000000 | | | | | |
| 2 | 0.53472222222 | 0.50462962963 | | | | |
| 4 | 0.50899376417 | 0.50041761149 | 0.50013681028 | | | |
| 8 | 0.50227085033 | 0.50002987904 | 0.50000403021 | 0.50000192259 | | |
| 16 | 0.50056917013 | 0.50000194339 | 0.50000008102 | 0.50000001833 | 0.50000001086 | |
| 32 | 0.50014238459 | 0.50000012275 | 0.50000000137 | 0.50000000010 | 0.50000000003 | 0.50000000002 |

As stated above, for a given value of $n$ (i.e. for a given number of function evaluations), the most accurate value is for the largest value of $k$, i.e. the diagonal entry. These are the successive Romberg estimates $R_k$, which in the present example are

| $k$ | $R_k$ | $\lvert R_k - R_{k-1}\rvert$ |
|---|---|---|
| 0 | 0.62500000000 | |
| 1 | 0.50462962963 | 0.12037037037 |
| 2 | 0.50013681028 | 0.00449281935 |
| 3 | 0.50000192259 | 0.00013488769 |
| 4 | 0.50000001086 | 0.00000191173 |
| 5 | 0.50000000002 | 0.00000001084 |

If we had wanted an accuracy of $\epsilon = 10^{-5}$ we would have stopped at $k = 4$ since $|R_4 - R_3| = 1.9\times10^{-6}$ ($< 10^{-5}$), whereas $|R_4 - R_3| = 1.3\times10^{-4}$ ($> 10^{-5}$). The resulting value of 0.50000001086 is indeed 1/2 to 5 decimal places. In fact it differs from the exact result by about $1.1 \times 10^{-8}$, quite a lot smaller than the desired accuracy.

Figure 1 shows the errors for the trapezium rule, Simpson's rule, and Romberg integration graphically (one more iteration has been done than in the table). The dashed lines indicate the leading error for the trapezium rule

$$-\frac{h^2}{12}(f'(b) - f'(a)) = \frac{7}{48}h^2, \tag{17}$$

and Simpson's rule

$$-\frac{h^4}{180}(f'''(b) - f'''(a)) = \frac{31}{240}h^4. \tag{18}$$

From Fig. 1 and these expressions for the error, we deduce that to reach a precision of $10^{-14}$, close to machine accuracy with double precision, Romberg integration needs 64 intervals, while Simpson's rule would need about 1900 intervals, and the trapezium rule would need no less than $3.8 \times 10^6$ intervals. Clearly, where high precision work is required, Romberg integration is to be preferred over Simpson's rule and the trapezium rule.

All in all, Romberg integration is a powerful but quite simple method, which I recommend for general use for integration of a smooth function, especially if high precision is needed. For a given number of intervals, it is more accurate than the trapezium rule and Simpson's rule but does not need any more function evaluations.

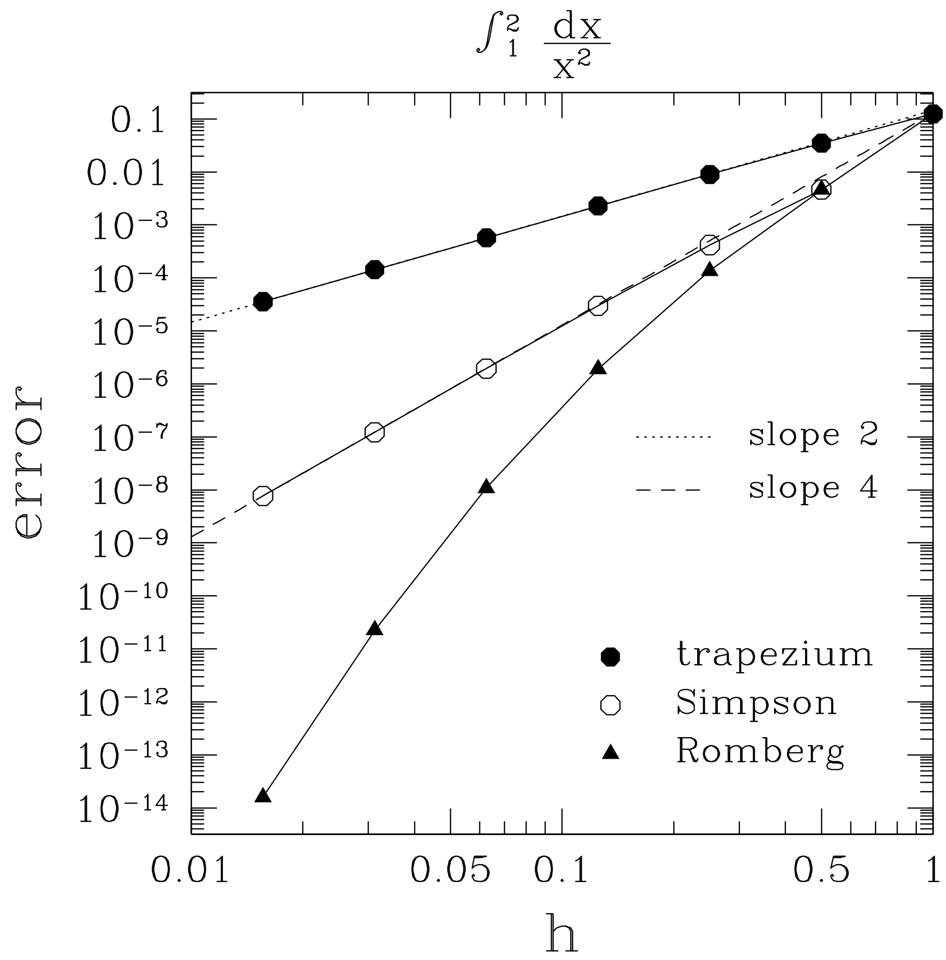


FIG. 1: Error in the integral $\int_1^2 dx/x$, using trapezium rule, Simpson's rule and Romberg integration. the number of intervals is $2^k$ with $k = 0, 1, 2, \cdots, 6$, and $h = (b - a)/n$ $(= 1/n$ here).