Romberg integration for a class of singular integrands

By L. Fox*

Simpson's rule and the trapezoidal rule are members of a particular class of approximate quadrature formulae. They have previously been used in connection with the deferred approach to the limit, or Romberg integration, which take account of the correcting terms. New correcting expressions are here obtained, and used with the extrapolation process, for three quadrature formulae and for integrands with an infinity in function or first derivative at the limits of integration.

1. In a recent paper Hamming and Pinkham (1966) considered a class of quadrature formulae of the type

$$I = \int_{x_0}^{x_{2n}} f(x)dx = G(h) + E_G(h) + R_G, \tag{1}$$

where $G(h)$ is a "Lagrangian approximation" to the integral I, $E_G(h)$ is a series of "correcting terms", and R_G is the "remainder" associated with the truncation of $E_G(h)$ after some particular term.

Here we take the Lagrangian form

$$G(h) = h(af_0 + bf_1 + cf_2 + \dots \\ \dots + cf_{2n-2} + bf_{2n-1} + af_{2n}), \tag{2}$$

where $f_r = f(x_0 + rh)$. The weight factors have the pattern $(a, b, c, b, c, \dots, c, b, a)$, and for this we need a range of $2n + 1$ pivotal points as indicated in the limits of integration in (1). If the formula is to be meaningful, with weights independent of n, we must necessarily have

$$b = 2(1 - a), \quad c = 2a. \tag{3}$$

The case $a = \frac{1}{3}$ gives Simpson's rule, which we designate by $G(h) = S(h)$; $a = \frac{1}{2}$ the trapezium rule, $G(h) = T(h)$; and Hamming and Pinkham also consider in detail the case $a = 0$, which we here denote by $G(h) = U(h)$. The trapezium rule, of course, does not require an odd number of pivotal points.

In (1) the term $E_G(h)$, the "correction" to the approximate formula, is given in the literature for the trapezium rule, both in terms of differences and also of derivatives, under the names Gregory formula and Euler–Maclaurin formula, respectively. The correction involves differences or derivatives only at the end points x_0 and x_{2n} of the interval of integration.

2. The remainder term, R_G, is expressible in various ways. For example, for the Euler–Maclaurin formula we can write

$$R_T = \frac{2n}{(2m)!} B_{2m} h^{2m+1} f^{(2m)}(\xi), \tag{4}$$

where $x_0 < \xi < x_{2n}$, B_{2m} is a Bernoulli number, and the highest derivative retained in the correction term is of order $2m - 3$. Now for fixed h the remainder term will not generally tend to zero as m increases, but for fixed m there will usually be some sufficiently small value of h for which R_T is negligible for some required precision. This is the justification for the use of this kind of quadrature formula, which is generally asymptotic but, provided that h is not too large (and this must be watched), has the property that truncation at a small enough term will give rise to an error of order of magnitude of this term. We shall assume that this is so for functions considered in this paper and whose definite integrals exist, and shall make no further reference to the remainder term.

3. Corresponding to the Gregory formula, Hamming and Pinkham find corrections in terms of differences for the S and U formulae. In this note I give a different method of obtaining these results, and also produce the corresponding derivative form of the correcting terms. For some purposes the latter is more valuable, since it gives the nature of the error of the approximate quadrature formula as a function of h, which is necessary for the successful application of the technique of the "deferred approach to the limit", or "extrapolation to zero interval length".

Now for well-behaved functions the correcting terms always consist of a series of even powers of h. For other integrands, however, for example those with an infinite derivative at one terminal point, the form of the correction might be quite different, and I show how we can adapt the basic formulae to produce this information. An adaptation of the extrapolation technique can then give very good answers without using small intervals or special techniques to take account of the singularity.

We can even deal with an infinity of the function at a terminal point, and for this purpose the use of the U formula avoids the obvious embarrassments of the S and T formulae. Various numerical examples indicate the power of these methods. Finally, we can use our new devices in a systematic way, effectively an extension of the Romberg integration technique which has formerly been applied only to well-behaved functions.

Difference formulae for smooth integrands

4. In classical finite-difference notation we can write

$$E_G(h) = D^{-1}(E^{2n} - 1)f_0 - G(h), \tag{5}$$

and from (2) and (3) we easily find

$$h^{-1}G(h) = a(1 + E^{2n}) + bE(1 + E^2 + \dots + E^{2n-2}) \\ + cE^2(1 + E^2 + \dots + E^{2n-4}) \\ = \left\{\frac{a(E-1)^2 + 2E}{E^2 - 1}\right\}(E^{2n} - 1). \tag{6}$$

Oxford Computing Laboratory, 19 Parks Road, Oxford.

Then

$$E_G(h) = h\left\{(hD)^{-1} - \frac{a(E-1)^2 + 2E}{E^2 - 1}\right\}(f_{2n} - f_0). \quad (7)$$

We can, of course, express the correcting terms in any way we please. In terms of forward differences of f_0 we find that the contribution at the lower limit has the symbolic form

$$h\left\{\frac{a(E-1)^2 + 2E}{E^2 - 1} - (hD)^{-1}\right\}$$

$$= h\left\{\frac{a\Delta}{2+\Delta} + \frac{1}{\Delta} + \frac{1}{2+\Delta} - \frac{1}{\ln(1+\Delta)}\right\}$$

$$= h\,\{a\,(\tfrac{1}{2}\,\Delta - \tfrac{1}{4}\,\Delta^2 + \tfrac{1}{8}\,\Delta^3 - \ldots) + (-\tfrac{1}{6}\,\Delta + \tfrac{1}{12}\,\Delta^2 - \tfrac{13}{360}\,\Delta^3 + \tfrac{1}{80}\,\Delta^4 - \tfrac{41}{30240}\,\Delta^5 + \ldots)\}, \quad (8)$$

where we have used the relation $E = 1 + \Delta$. Corresponding treatment with backward differences at the upper limit produces the result

$$E_T(h) = h\,\{-\tfrac{1}{12}\,(\nabla f_{2n} - \Delta f_0) - \tfrac{1}{24}\,(\nabla^2 f_{2n} + \Delta^2 f_0) - \tfrac{19}{720}\,(\nabla^3 f_{2n} - \Delta^3 f_0) - \tfrac{3}{160}\,(\nabla^4 f_{2n} + \Delta^4 f_0) - \tfrac{863}{60480}\,(\nabla^5 f_{2n} - \Delta^5 f_0) - \ldots\}, \quad (9)$$

which is the Gregory correction to the trapezium rule ($a = \frac{1}{2}$), and

$$E_S(h) = h\,\{-\tfrac{1}{180}\,(\nabla^3 f_{2n} - \Delta^3 f_0) - \tfrac{1}{120}\,(\nabla^4 f_{2n} + \Delta^4 f_0) - \tfrac{137}{15120}\,(\nabla^5 f_{2n} - \Delta^5 f_0) - \ldots\}, \quad (10)$$

which is the corresponding correction to the Simpson rule, the choice $a = \frac{1}{3}$ having eliminated both the first-difference and also the second-difference terms. We also have

$$E_U(h) = h\,\{\tfrac{1}{6}\,(\nabla f_{2n} - \Delta f_0) + \tfrac{1}{12}\,(\nabla^2 f_{2n} + \Delta^2 f_0) + \tfrac{13}{360}\,(\nabla^3 f_{2n} - \Delta^3 f_0) + \ldots\}, \quad (11)$$

for the case $a = 0$, the coefficients coming from the last term in (8).

Hamming and Pinkham show that the Simpson coefficients are everywhere smaller than the Gregory coefficients, and that those of (11) are smaller than either for differences of order greater than four. They therefore recommend the U formula as the best of its class, and this has some justification if, as they seem to think and as I also think, corrections in terms of differences should be employed wherever possible and as a matter of course even in modern machine computation.

On the other hand we note that if the integrand is a polynomial of degree $2n$ all the formulae terminate with the differences Δ^{2n} or ∇^{2n}. They then all represent exactly the same linear combination of the same pivotal values $f_0, f_1, \ldots, f_{2n}$, differing only in the way in which this combination is shared between the Lagrangian terms and the correcting terms.

Derivative formulae for smooth integrands

5. For those who do not wish to make a correction, however, and ask that the correcting terms should be of high order in h and therefore negligible at a not too small interval, the Simpson rule is obviously the best. Writing

$$\frac{(E-1)^2}{E^2-1} = \frac{E-1}{E+1} = \tanh \tfrac{1}{2}hD, \quad \frac{2E}{E^2-1} = \operatorname{cosech} hD, \quad (12)$$

equation (7) becomes

$$E_G(h) = h\{(hD)^{-1} - a \tanh \tfrac{1}{2}hD - \operatorname{cosech} hD\}. \quad (13)$$

Then, with the aid of the expansions

$$\left.\begin{aligned} \tanh \tfrac{1}{2}hD &= \tfrac{1}{2}hD - \tfrac{1}{24}\,h^3D^3 + \tfrac{1}{240}\,h^5D^5 - \tfrac{17}{41320}\,h^7D^7 + \ldots \\ \operatorname{cosech} hD &= (hD)^{-1} - \tfrac{1}{6}\,hD + \tfrac{7}{360}\,h^3D^3 - \tfrac{31}{15120}\,h^5D^5 + \ldots \end{aligned}\right\} \quad (14)$$

we can produce the required formulae. For $a = \frac{1}{2}$ we have the Euler–Maclaurin correction

$$E_T(h) = -\tfrac{1}{12}\,h^2(f'_{2n} - f'_0) + \tfrac{1}{720}\,h^4(f'''_{2n} - f'''_0) - \tfrac{1}{30240}\,h^6(f^{(v)}_{2n} - f^{(v)}_0) + \ldots, \quad (15)$$

for $a = \frac{1}{3}$ we obtain the corresponding Simpson correction

$$E_S(h) = -\tfrac{1}{180}\,h^4(f'''_{2n} - f'''_0) + \tfrac{1}{1512}\,h^6(f^{(v)}_{2n} - f^{(v)}_0) - \ldots, \quad (16)$$

(which I gave in Fox, 1961) and for the U formula we find

$$E_U(h) = \tfrac{1}{6}\,h^2(f'_{2n} - f'_0) - \tfrac{7}{360}\,h^4(f'''_{2n} - f'''_0) + \tfrac{31}{15120}\,h^6(f^{(v)}_{2n} - f^{(v)}_0) - \ldots. \quad (17)$$

It is interesting to note that the early U coefficients are larger than the T coefficients and of opposite sign (as, of course, they are in (11) compared with (9)), so that the unrefined U formula will have a larger error than the T formula. Equation (17) is in fact a special case of the general Euler–Maclaurin formula, effectively with interval $2h$, and like (15) has coefficients proportional to the Bernoulli numbers.

If sufficient derivatives exist everywhere, these results imply that

$$\left.\begin{aligned} I - T(h) &= A_T h^2 + B_T h^4 + C_T h^6 + \ldots \\ I - S(h) &= A_S h^4 + B_S h^6 + C_S h^8 + \ldots \\ I - U(h) &= A_U h^2 + B_U h^4 + C_U h^6 + \ldots \end{aligned}\right\} \quad (18)$$

where I is the correct value of the integral and the $A, B, C, \ldots$, are constants. These series are asymptotic, but if the remainder term is negligible we can say that the error is $O(h^2)$ in $T(h)$ and $U(h)$ and $O(h^4)$ in $S(h)$, and use the extended forms (18) to get a better result from two or more approximate computations with different values of h.

Integrands with an integrable singularity at a limit of integration

6. In equations (15)–(17) the correcting terms involving f_0 and f_{2n} can be treated separately, and either can be

expressed, for example, in terms of derivatives at some other pivotal point. Formally this can be achieved by the classic device incorporated in the formula

$$\begin{aligned} E_T(h)f_0 &= (\tfrac{1}{12}h^2D - \tfrac{1}{720}h^4D^3 + \ldots)E^{-1}Ef_0 \\ &= (\tfrac{1}{12}h^2D - \tfrac{1}{720}h^4D^3 + \ldots)(e^{-hD})f_1 \\ &= (\tfrac{1}{12}h^2D - \tfrac{1}{12}h^3D^2 + \tfrac{29}{720}h^4D^3 - \ldots)f_1. \quad (19) \end{aligned}$$

The derivatives of f_1 are functions of h, and this would seem to imply that the correcting terms include all powers of h from the second onwards. If $f(x)$ is well-behaved, however, with a convergent Taylor's series at every point, reversal of the method which produced (19) shows that the coefficients of odd powers of h are identically zero.

If the Taylor's series does not exist at x_0, however, for example when $f'(x_0)$ is infinite, we would like to use a formula like (19) in which the derivatives do exist. We cannot now expect that odd powers of h will disappear, or even that the correcting terms are expressible as a power series in h. The unsuspecting use of (18), for purposes of extrapolating to the limit, will then have no validity.

7. We consider first the case in which $f'(x_0)$ is infinite, but assume that $f(x)$ is everywhere finite and has no other singularity in the interval x_0 to x_{2n}, that is its Taylor's series exists at every point other than x_0. To obtain the required formula we treat separately the contributions to the integral from the ranges x_0 to x_1 and x_1 to x_{2n}. For the latter we have, for the $T(h)$ formula,

$$\left.\begin{aligned} \int_{x_1}^{x_{2n}} f(x)dx &= h\{\tfrac{1}{2}f_1 + f_2 + \ldots + f_{2n-1} + \tfrac{1}{2}f_{2n}\} \\ &\qquad + E + R \\ E &= -\tfrac{1}{12}h^2(f'_{2n} - f'_1) + \tfrac{1}{720}h^4(f'''_{2n} - f'''_1) - \ldots \end{aligned}\right\} \quad (20)$$

For the interval (x_0, x_1) we use the Taylor's series at the point x_1, given by

$$f\{x_1 + (x - x_1)\} = f_1 + (x - x_1)f'_1 + \frac{(x - x_1)^2}{2!}f''_1 + \ldots, \quad (21)$$

in which all the derivatives exist. This, by our hypothesis, will converge on the negative side of x_1 as far as x_0 (but no further!), so that

$$f_0 = f_1 - hf'_1 + \frac{h^2}{2!}f''_1 - \ldots. \quad (22)$$

Then also

$$\int_{x_0}^{x_1} f(x)dx = hf_1 - \frac{h^2}{2!}f'_1 + \frac{h^3}{3!}f''_1 - \ldots, \quad (23)$$

and from (22) and (23) we deduce the required result

$$\int_{x_0}^{x_1} f(x)dx = \tfrac{1}{2}h(f_0 + f_1) - \frac{h^3}{2\cdot 3!}f''_1 + \frac{2h^4}{2\cdot 4!}f'''_1 - \frac{3h^5}{2\cdot 5!}f^{(iv)}_1 + \ldots. \quad (24)$$

The addition of (20) verifies that the formal result (19) holds in this case also.

Similar treatment of (16) and (17), either by the process which led to (19) or by the rigorous analysis which now needs separate treatment of the first two intervals (x_0, x_1) and (x_1, x_2), produces the formulae

$$\left.\begin{aligned} E_S(h)f_0 &= (\tfrac{1}{180}h^4D^3 - \tfrac{1}{180}h^5D^4 \\ &\qquad + \tfrac{2}{945}h^6D^5 - \ldots)f_1 \\ E_U(h)f_0 &= (-\tfrac{1}{6}h^2D + \tfrac{1}{6}h^3D^2 \\ &\qquad - \tfrac{23}{360}h^4D^3 + \ldots)f_1 \end{aligned}\right\}, \quad (25)$$

as the contributions to the correcting terms at the offending lower limit in the S and U formulae. [I gave the results (19) and the first of (25), in forms based on the rigorous approach, in Fox, 1961.]

The deferred approach to the limit

8. We propose to use (19) and (25) solely for the purpose of discovering the form of the correcting terms of the approximate quadrature formulae, for use with the deferred approach to the limit. The point is that the derivatives of f_1 are themselves functions of h, and their nature may give different expressions on the right-hand sides of the new equations corresponding to (18).

Consider, for example, the integraion between $x = 0$ and $x = 1$ of the functions

$$f_1(x) = x^{1/2}, f_2(x) = x \ln x, f_3(x) = x^{1/2}\ln x, \quad (26)$$

all of which have infinite first derivatives at the point $x = 0$. By inspecting successive derivatives, and remembering the contributions from the upper limit, we find the correcting terms shown in **Table 1**. In all cases the U formula has the same terms as the T formula, though in different magnitudes.

Table 1

Function	$I - T(h)$	$I - S(h)$
$x^{1/2}$	$h^{3/2}, h^2, h^4, \ldots$	$h^{3/2}, h^4, h^6, \ldots$
$x \ln x$	$h^2 \ln h, h^2, h^4, \ldots$	$h^2, h^4, h^6, \ldots$
$x^{1/2} \ln x$	$h^{3/2} \ln h, h^{3/2}, h^2, h^4, \ldots$	$h^{3/2} \ln h, h^{3/2}, h^4, h^6, \ldots$

The function

$$f_4(x) = \{x(1 - x)\}^{1/2} \quad (27)$$

has an infinite first derivative at both limits $x = 0$ and $x = 1$, and we find that all three formulae have correcting terms

$$h^{3/2}, h^{5/2}, h^{7/2}, \ldots, \quad (28)$$

with no even powers of h in sight.

9. The method of the deferred approach proceeds as follows. For the integration of $x^{1/2}$ with the T rule, for example, we write

$$I - T(h_r) = Ah_r^{3/2} + Bh_r^2 + Ch_r^4 + \ldots, \quad (29)$$

take two or more values of h_r and compute the corresponding $T(h_r)$, then eliminate one or more terms on the right-hand side of (29) to produce a better approximation to I. In what follows we shall assume successive halving of the interval, so that $h_2 = \frac{1}{2}h_1$, $h_3 = \frac{1}{2}h_2$, etc.

Eliminating A in (29), and neglecting the remaining terms, we find as a better approximation to the integral the quantity

$$T(h_1, h_2) = (2\sqrt{2}T_2 - T_1)/(2\sqrt{2} - 1), \qquad (30)$$

where $T_r = T(h_r)$. The corresponding formula holds (in all cases) for the U quadrature, and here also for Simpson's rule, though the application of (30) with Simpson's rule should give better results since the other neglected terms are smaller.

In fact for the T and U computation we might prefer to eliminate both A and B in (29) using three approximate computations, and with the neglect of later terms we find the better approximation

$$T(h_1, h_2, h_3) = \tfrac{1}{21}\{(32 + 8\sqrt{2})T_3 - (12 + 10\sqrt{2})T_2 + (1 + 2\sqrt{2})T_1\}, \qquad (31)$$

where $T_r = T(h_r)$ and $h_3 = \frac{1}{2}h_2 = \frac{1}{4}h_1$.

Numerical examples

10. We give some results for $\int_0^1 f(x)dx$ for the integrands of Table 1. In all cases we compute pivotal values correct to six decimals, and avoid further rounding errors by keeping a "guarding" figure in the evaluation of $T(h)$, etc.

For $\int_0^1 x^{1/2}dx$ we find

$$\left.\begin{array}{ll} S(\frac{1}{4}) = 0\cdot656526_2 & S(\frac{1}{8}) = 0\cdot663079 \\ T(\frac{1}{4}) = 0\cdot643283 & T(\frac{1}{8}) = 0\cdot658130 \\ U(\frac{1}{4}) = 0\cdot683012_5 & U(\frac{1}{8}) = 0\cdot672977 \\ \multicolumn{2}{c}{S(\frac{1}{16}) = 0\cdot665398_2} \\ \multicolumn{2}{c}{T(\frac{1}{16}) = 0\cdot663581_1} \\ \multicolumn{2}{c}{U(\frac{1}{16}) = 0\cdot669032_2} \end{array}\right\}, \qquad (32)$$

and with decreasing h all the formulae converge to $I = 2/3$.

Extrapolating, we find from (30), in obvious notation, the results

$$\left.\begin{array}{ll} T(\frac{1}{4}, \frac{1}{8}) = 0\cdot666250, & T(\frac{1}{8}, \frac{1}{16}) = 0\cdot666562 \\ U(\frac{1}{4}, \frac{1}{8}) = 0\cdot667488, & U(\frac{1}{8}, \frac{1}{16}) = 0\cdot666875 \end{array}\right\}, \qquad (33)$$

and (31) gives the still better results

$$T(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot666667, \quad U(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot666670. \qquad (34)$$

The formula (30) is already very good for the Simpson rule, giving

$$S(\tfrac{1}{4}, \tfrac{1}{8}) = 0\cdot666663, \quad S(\tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot666667. \qquad (35)$$

11. For $-\int_0^1 x \ln x\, dx$ the Simpson rule, interestingly enough, has "h^2-extrapolation" (compared with h^4-extrapolation in well-behaved cases), so that with $h_2 = \frac{1}{2}h_1$ we have

$$S(h_1, h_2) = S_2 + \tfrac{1}{3}(S_2 - S_1). \qquad (36)$$

For the corresponding T (and U) formulae the elimination of the first term of the correction gives

$$T(h_1, h_2) = \frac{T_2(4 \ln h_1) - T_1(\ln h_1 - \ln 2)}{3 \ln h_1 + \ln 2}. \qquad (37)$$

This can be used but is relatively unsatisfactory since it depends on the current absolute size of the interval.

In any case the second correcting term is of comparable size to the first, and we certainly do better to eliminate both $h^2 \ln h$ and h^2, obtaining the satisfactory formula

$$T(h_1, h_2, h_3) = \tfrac{1}{9}(16T_3 - 8T_2 + T_1). \qquad (38)$$

We find

$$\left.\begin{array}{ll} S(\frac{1}{4}) = 0\cdot245207_7 & S(\frac{1}{8}) = 0\cdot248798 \\ T(\frac{1}{4}) = 0\cdot227227_5 & T(\frac{1}{8}) = 0\cdot243405_4 \\ U(\frac{1}{4}) = 0\cdot281168 & U(\frac{1}{8}) = 0\cdot259583_2 \\ \multicolumn{2}{c}{S(\frac{1}{16}) = 0\cdot249699_4} \\ \multicolumn{2}{c}{T(\frac{1}{16}) = 0\cdot248125_9} \\ \multicolumn{2}{c}{U(\frac{1}{16}) = 0\cdot252846_4} \end{array}\right\}, \qquad (39)$$

which are converging to $I = \frac{1}{4}$.

Extrapolating from (36) we obtain

$$S(\tfrac{1}{4}, \tfrac{1}{8}) = 0\cdot249995, \quad S(\tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot250000. \qquad (40)$$

From (37), which for $h_1 = 2^{-n}$ reduces to

$$T(h_1, h_2) = \frac{4nT_2 - (n + 1)T_1}{3n - 1}, \qquad (41)$$

with a similar formula for $U(h_1, h_2)$, we compute

$$\left.\begin{array}{ll} T(\frac{1}{4}, \frac{1}{8}) = 0\cdot253112, & T(\frac{1}{8}, \frac{1}{16}) = 0\cdot250486 \\ U(\frac{1}{4}, \frac{1}{8}) = 0\cdot246632, & U(\frac{1}{8}, \frac{1}{16}) = 0\cdot249478 \end{array}\right\}, \qquad (42)$$

and (38) gives

$$T(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot250000, \quad U(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot250005. \qquad (43)$$

All these results confirm our expectations.

12. For $-\int_0^1 x^{1/2} \ln x\, dx$ we should expect, from Table 1, that in all cases we need to eliminate the first two correcting terms to get a good answer, and that the Simpson computation, with smaller neglected later terms, would have the advantage. In all cases the extrapolation formula is typified by

$$S(h_1, h_2, h_3) = \frac{S(h_3) - 2^{-1/2}S(h_2) + \frac{1}{8}S(h_1)}{1 - 2^{-1/2} + \frac{1}{8}}. \qquad (44)$$

We find

$$\left.\begin{array}{ll} S(\tfrac{1}{4}) = 0\cdot395784 & S(\tfrac{1}{8}) = 0\cdot424752 \\ T(\tfrac{1}{4}) = 0\cdot358104 & T(\tfrac{1}{8}) = 0\cdot408090 \\ U(\tfrac{1}{4}) = 0\cdot471143_5 & U(\tfrac{1}{8}) = 0\cdot458076 \\ \multicolumn{2}{c}{S(\tfrac{1}{16}) = 0\cdot436602_7} \\ \multicolumn{2}{c}{T(\tfrac{1}{16}) = 0\cdot429474_5} \\ \multicolumn{2}{c}{U(\tfrac{1}{16}) = 0\cdot450859} \end{array}\right\}, \quad (45)$$

which are converging to $I = \frac{4}{9}$. Extrapolation from (44) gives

$$S(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot444445, \quad T(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot444310, \quad U(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot444715, \quad (46)$$

which again confirm our predictions.

13. Finally, for $\int_0^1 \{x(1-x)\}^{1/2} dx$, the result (28) suggests that the Simpson rule, extrapolated or otherwise, is not an order of magnitude better than the other formulae. In all cases elimination of the first correcting term gives the formula typified by (30), and the elimination of the first two terms is effected with the extrapolation typified by

$$T(h_1, h_2, h_3) = \frac{16T_3 - 6\sqrt{2}T_2 + T_1}{17 - 6\sqrt{2}}. \quad (47)$$

Computation gives

$$\left.\begin{array}{ll} S(\tfrac{1}{4}) = 0\cdot372008 & S(\tfrac{1}{8}) = 0\cdot385448_7 \\ T(\tfrac{1}{4}) = 0\cdot341506 & T(\tfrac{1}{8}) = 0\cdot374463 \\ U(\tfrac{1}{4}) = 0\cdot433012 & U(\tfrac{1}{8}) = 0\cdot407420 \\ \multicolumn{2}{c}{S(\tfrac{1}{16}) = 0\cdot390148_5} \\ \multicolumn{2}{c}{T(\tfrac{1}{16}) = 0\cdot386227_1} \\ \multicolumn{2}{c}{U(\tfrac{1}{16}) = 0\cdot397991_2} \end{array}\right\}, \quad (48)$$

which are converging to $I = \frac{1}{8}\pi = 0\cdot392699...$. Extrapolation from (30) gives

$$S(\tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot392719, \quad T(\tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot392661, \quad U(\tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot392834, \quad (49)$$

and (47) produces

$$S(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot392702, \quad T(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot392698, \quad U(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 0\cdot392708. \quad (50)$$

Again our major expectations are confirmed.

The power of the method is here indicated by a remark of Hamming and Pinkham, that this function, using the U formula with corrections up to and including fourth differences, "integrates miserably with 128 points, the relative error being 0·0002".

Integrands with an infinity in function value

14. If we have $2n + 1$ pivotal points with the largest interval, and halve the interval once, both the S and T formulae require $4n + 1$ function evaluations, and a further halving gives a total of $8n + 1$ evaluations. For the U formula, which with the T formula has "easier" weights than Simpson, these figures are $3n$ and $7n$ respectively. This hardly implies, for the type of examples treated, that U is superior to S.

If the integrand has an infinity at the lower limit, however, the U formula comes into its own. The S and T formulae have an embarrassing infinity in f_0, and moreover an equation like (21) fails to converge at $x = 0$ so that formulae (19) and the first of (25) are no longer valid. For the U formula we do not use the value f_0, equation (21) is not required to converge at $x = 0$, and if the integral exists we are perfectly satisfied with the second of (25) to produce the contribution at the lower limit to the error in the approximate formula. We can then proceed to use the deferred approach in the standard manner.

15. Consider, for example, the computation of $\int_0^1 x^{-1/2} dx$, whose integrand is infinite at the lower limit. Consideration of its derivatives, at $x = h$, show that the correction terms are multiples of

$$h^{1/2}, h^2, h^4, \ldots . \quad (51)$$

Extrapolation to eliminate the first term, with $h_2 = \frac{1}{2}h_1$, produces

$$U(h_1, h_2) = (2^{1/2}U_2 - U_1)/(2^{1/2} - 1), \quad (52)$$

and if we also eliminate h^2 we find

$$U(h_1, h_2, h_3) = \frac{(16\sqrt{2} - 8)U_3 - 14U_2 + (4 - \sqrt{2})U_1}{15\sqrt{2} - 18}. \quad (53)$$

Computation gives

$$U(\tfrac{1}{4}) = 1\cdot577350, \quad U(\tfrac{1}{8}) = 1\cdot698844, \quad U(\tfrac{1}{16}) = 1\cdot786461, \quad (54)$$

which are converging at no great rate to $I = 2$. Extrapolation using (52) gives

$$U(\tfrac{1}{4}, \tfrac{1}{8}) = 1\cdot992156, \quad U(\tfrac{1}{8}, \tfrac{1}{16}) = 1\cdot997987, \quad (55)$$

and (53) gives the remarkably good result

$$U(\tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}) = 1\cdot999931. \quad (56)$$

This may be compared with a computation given by Davis and Rabinowitz (1965), in which the unrefined Simpson's rule and the trapezium rule (taking $f(0) = 0$) have an error of about 0·03 with more than 1000 pivotal values.

Analogy with Romberg integration

16. In Romberg's scheme of extrapolation (Romberg, 1955) the calculation is as shown in **Table 2**.

Here $G(h_r)$ is the value obtained by the approximate quadrature formula at interval h_r. The second column gives the results of eliminating the first correcting term from $G(h_r)$ and $G(h_{r+1})$. These quantities have errors

Table 2

$G(h_1)$		
	$G(h_1, h_2)$	
$G(h_2)$		$G(h_1, h_2, h_3)$
	$G(h_2, h_3)$	
$G(h_3)$		
⋮	⋮	⋮

dominated by the second term of the original correction formula, and this is now eliminated from $G(h_r, h_{r+1})$ and $G(h_{r+1}, h_{r+2})$ to give a quantity in the third column. Successive columns converge with increasing speed to the correct result.

Even with well-behaved functions the formulae in successive extrapolations are not the same. For example with Simpson's rule, with successive halving of the interval we shall normally have

$$\left.\begin{aligned} G(h_1, h_2) &= G(h_2) + \frac{1}{2^4 - 1}\{G(h_2) - G(h_1)\} \\ G(h_1, h_2, h_3) &= G(h_2, h_3) \\ &\quad + \frac{1}{2^6 - 1}\{G(h_2, h_3) - G(h_1, h_2)\} \end{aligned}\right\}, \quad (57)$$

and so on.

17. We can express our new computations in a very similar form. As a first example consider the evaluation of the integral in § 10 with the U formula. With the array of Table 2, the second column is obtained from equation (30), and the results have errors dominated by a constant multiple of h^2. The third column is then obtained by h^2-extrapolation, corresponding to the factor $(2^2 - 1)^{-1}$ in an equation similar to (57), the third column by h^4-extrapolation, and so on. The tableau is shown in **Table 3**. The results of the third column are, of course, identical with those obtained in § 10 from equation (31).

Table 3

$U(\frac{1}{2}) = 0\cdot707106_8$			
	$0\cdot669834_9$		
$U(\frac{1}{4}) = 0\cdot683012_5$		$0\cdot666706$	
	$0\cdot667488_4$		$0\cdot666668$
$U(\frac{1}{8}) = 0\cdot672977$		$0\cdot666670$	
	$0\cdot666874_7$		
$U(\frac{1}{16}) = 0\cdot669032_2$			

For the problem of § 13 elimination of successive error terms is accomplished by a formula similar to (30), the term $2\sqrt{2}$ in that equation being replaced by $2^{3/2}, 2^{5/2}, 2^{7/2}, \ldots$ in successive columns. We find the array of **Table 4**.

Likewise, without further description, we find for the problem of § 15 the array of **Table 5**.

Table 4

$U(\frac{1}{2}) = 0\cdot500000$			
	$0\cdot396375$		
$U(\frac{1}{4}) = 0\cdot433012$		$0\cdot392789_5$	
	$0\cdot393423_3$		$0\cdot392697$
$U(\frac{1}{8}) = 0\cdot407420$		$0\cdot392707_9$	
	$0\cdot392834_4$		
$U(\frac{1}{16}) = 0\cdot397991_2$			

Table 5

$U(\frac{1}{2}) = 1\cdot414214$			
	$1\cdot971195$		
$U(\frac{1}{4}) = 1\cdot577350$		$1\cdot999143$	
	$1\cdot992156$		$1\cdot999984$
$U(\frac{1}{8}) = 1\cdot698844$		$1\cdot999931$	
	$1\cdot997987$		
$U(\frac{1}{16}) = 1\cdot786461$			

18. In these examples the formulae varied from column to column, but were constant within each column. This cannot always be guaranteed, and indeed for the problems of §§ 11 and 12 the formulae for the second and third columns, but not for succeeding columns, depend on the position in the column. We can find the relevant formulae, but we might prefer to omit these "awkward" columns. In the problem of § 12, for example, we might go directly to the third column by way of equation (44) for the U formula, say, and then proceed with $h^2, h^4, \ldots$ extrapolation in succeeding columns. This gives the array of **Table 6**.

Table 6

$U(\frac{1}{2}) = 0\cdot490129$			
	—		
$U(\frac{1}{4}) = 0\cdot471143$		$0\cdot445552$	
	—		$0\cdot444436$
$U(\frac{1}{8}) = 0\cdot458076$		$0\cdot444715$	
	—		
$U(\frac{1}{16}) = 0\cdot450859$			

Extensions

19. The methods can obviously be extended to cover more elaborate integrals than those mentioned in § 8 and illustrated in subsequent sections. Consider, for example, the integrals

$$I_1 = \int_0^1 x^{1/2} g(x)dx, \quad I_2 = \int_0^1 x \ln x \, g(x)dx,$$
$$I_3 = \int_0^1 x^{1/2} \ln x \, g(x)dx, \quad (58)$$

in which $g(x)$ has no singularity and can therefore be represented by a convergent power series. It is then sufficient to consider the respective integrands

$$f_5(x) = x^{r+1/2}, \; f_6(x) = x^{r+1} \ln x, \; f_7(x) = x^{r+1/2} \ln x, \quad r = 0, 1, 2, \ldots \quad (59)$$

Table 7

$I_1 - T(h)$	$h^{3/2}, h^2, h^{5/2}, h^{7/2}, h^4, \ldots$
$I_1 - S(h)$	$h^{3/2}, h^{5/2}, h^{7/2}, h^4, h^{9/2}, h^{11/2}, h^6, \ldots$
$I_2 - T(h)$	$h^2 \ln h, h^2, h^3, h^4 \ln h, h^4, h^5 \ln h, h^5, \ldots$
$I_2 - S(h)$	$h^2, h^3, h^4 \ln h, h^4, h^5 \ln h, h^5, \ldots$
$I_3 - T(h)$	$h^{3/2} \ln h, h^{3/2}, h^2, h^{5/2} \ln h, h^{5/2}, h^{7/2} \ln h, h^{7/2}, h^4, \ldots$
$I_3 - S(h)$	$h^{3/2} \ln h, h^{3/2}, h^{5/2} \ln h, h^{5/2}, h^{7/2} \ln h, h^{7/2}, h^4, \ldots$

Analysis corresponding to that of § 8 reveals the correcting terms shown in **Table 7**.

We note the vanishing of a perhaps expected term $h^3 \ln h$ in $I_2 - T(h)$. Other terms might vanish fortuitously, for a $g(x)$ of the form

$$g(x) = a + bx + cx^2 + \ldots, \quad (60)$$

for some particular combinations of the constants $a, b, c, \ldots$. It is, of course, necessary to eliminate in the analogous Romberg process the terms which *are* present in the correcting expression, but it does not matter if we eliminate a term which is not present!

20. For

$$I_2 = -\int_0^1 \left(\frac{x}{1+x}\right) \ln x \, dx = 1 - \frac{\pi^2}{12} \sim 0\cdot177533, \quad (61)$$

we find for the $T(h)$ formula the array of **Table 8**.

Table 8

$T(1) = 0\cdot000000$			
$T(\frac{1}{2}) = 0\cdot115524_5$	$0\cdot178022_7$		
		$0\cdot177589_3$	
$T(\frac{1}{4}) = 0\cdot157900$	$0\cdot177643_5$		$0\cdot177531_3$
		$0\cdot177534_9$	
$T(\frac{1}{8}) = 0\cdot171654_2$	$0\cdot177548_5$		
$T(\frac{1}{16}) = 0\cdot175829_4$			

The second column is the result of eliminating simultaneously the correcting terms $h^2 \ln h$ and h^2 by means of (38). The third column is obtained by eliminating the next term h^3. (We note that even the term $T(1)$ can play an effective part in the computation. In fact if we use $T(1)$ and $T(\frac{1}{2})$ in the Romberg table corresponding to the examples of §§ 10–13, we find answers correct to within a rounding error without using $T(\frac{1}{16})$ at all!)

At this stage we find, with obvious notation,

$$\left. \begin{aligned} I - (T_{1,\,1/2,\,1/4,\,1/8}) &= D \ln 2 + E \\ I - (T_{1/2,\,1/4,\,1/8,\,1/16}) &= \tfrac{1}{40} D \ln 2 + \tfrac{1}{16} E \end{aligned} \right\}, \quad (62)$$

where D and E are constants involved in the respective correcting terms $h^4 \ln h$ and h^4. We would normally eliminate these simultaneously from *three* approximations, but in the absence of a third we do well to eliminate the E term at this stage, obtaining the single value in the fourth column.

Conclusion

21. We have demonstrated, in the main parts of the paper, a method for finding the dominant correcting terms in the integration of functions with an infinite value or infinite first derivative at a pivotal point. The method of the deferred approach to the limit, and the special form of this embodied in Romberg integration, have been adapted to cover this situation. Numerical examples reveal both the power and the simplicity of the method. Other cases, such as an infinity in a higher derivative, or singularities at some internal point in the range, can obviously be treated by similar processes.

We have not discussed the question of accumulation of rounding error, but this is clearly not a difficult problem. At each stage every number is expressible as a linear combination of pivotal values, and in our examples the resulting effect of errors of $0\cdot5$ in the last digit of the pivotal values is small. In these examples the maximum error can nowhere exceed two units in the last figure in the final results.

I am grateful to Dr. J. C. P. Miller for several important suggestions, which considerably improved the first draft of this paper.

References

DAVIS, P. J., and RABINOWITZ, P. (1965). "Ignoring the singularity in approximate integration", *J. Soc. Indust. Appl. Math. (Series B)*, Vol. 2, pp. 367–383.

FOX, L. (1961). "Some comments on the accuracy and convenience of finite-difference processes in automatic computation", *Premier Congrès de L'Association Française de Calcul*, Grenoble 1961. Paris: Gauthier-Villars.

HAMMING, R. W., and PINKHAM, R. S. (1966). "A class of integration formulas", *J. Ass. Comp. Mach.*, Vol. 13, pp. 430–438.

ROMBERG, W. (1955). "Vereinfachte numerische Integration", *Norske Vid. Selsk. Forh.* (Trondheim), Vol. 28, pp. 30–36.