The evaluation of definite integrals using high-order formulae

J. Oliver

Computing Centre, University of Essex, Colchester, Essex

The use of high-order integration formulae in general-purpose library routines is widely discouraged in the literature. The reasons advanced for the recommended preference for the trapezoidal, mid-point and Simpson's rules are here analysed, and found to be either irrelevant to modern computation, or highly inconclusive. Attainable error bounds are presented which help to make high-order formulae equally attractive in problems for which they were formerly regarded as inefficient.

(Received September 1969, Revised July 1970)

1. Introduction

When evaluating a definite integral of the form

$$I = \int_a^b f(x)\,dx \qquad (1)$$

in which $f(x)$ is defined and calculable over $[a, b]$, we can in theoretical evaluations utilise special properties of the integrand, while in numerical calculations we can often use these to select an optimum formula. Nevertheless, the usual requirement is for a general-purpose 'library' procedure or program, universal in application, which will 'fail-safe', giving an appropriate report, in circumstances that become beyond its scope; such failures will be detected by an examination of calculated error estimates. Multiple access systems offer some potential for the computer user to select a suitable formula in conversation with the system, but since this has yet to be realised we shall be concerned here with quadrature formulae used blindly, without regard to any special characteristics of the integrand.

Now, for any given formula or algorithm a pathological problem can always be devised for which an arbitrarily small accuracy cannot be attained; we can therefore never argue the universality of any particular method, and we do not attempt this. Instead we examine the converse—an apparently widespread assumption, which has not been fully analysed, that all high order formulae are by their nature inferior, as a basis for 'universal' routines, to the best low-order formulae; this has been given—with various reasons—as a basis for recommending that they *never* be used.

We are, incidentally, here taking the order of a quadrature formula to be the maximum integer m such that the formula integrates exactly all polynomials of degree at most m. This definition is adequate for all the methods under consideration here, and could readily be extended to include non-unit weight functions if required, although formulae involving them might well be regarded as special-purpose rather than general-purpose ones.

The disadvantages which have, at various times, been associated with some or all high-order formulae are as follows:

1. Their possible non-convergence for continuous integrands.
2. The occurrence of negative weights.
3. The difficulty of estimating errors directly using expressions involving high-order derivatives.
4. The (supposed) inevitability of a large discretisation error when some low-order derivative of the integrand is non-existent or discontinuous in $[a, b]$ or attains a very large magnitude.
5. The computational effort required to achieve a given accuracy in comparison with low-order formulae.

We shall argue that the first three of these objections are no longer relevant to current algorithms, while the remaining two have not been fully substantiated in the literature. Indeed, by extending the error bounds developed by Stroud and Secrest, we show that on this basis of comparison there is little to choose between formulae of high and low order. Thus we shall conclude that there seems no valid case for excluding absolutely either high- or low-order formulae from such algorithms. Finally, having hitherto implicitly assumed that the various formulae considered are applied uniformly over $[a, b]$, we end by considering adaptive integration methods. Although many of those currently in use employ only low-order formulae, we shall argue that this again does not warrant the exclusion of high-order formulae from a general-purpose quadrature algorithm, and indeed the optimum order should be selected by the algorithm itself.

Note that we are not advocating the exclusive use of either high-order or low-order formulae; rather we are drawing attention to the lack of any substantial theoretical or empirical evidence for preferring either, in the hope of stimulating further research into this question.

2. Possible non-convergence

One apparently major defect of the high-order Newton-Cotes rules which has been widely quoted is the possible non-convergence of the estimates as the order increases, even when $f(x)$ is continuous throughout $[a, b]$ (see, for more recent examples, Bauer, Rutishauser and Stiefel, 1963; Davis and Rabinowitz, 1967, p. 31). Although this may certainly occur for functions which are analytic on $[a, b]$, as Kusmin (1931) and Pólya (1933) showed, Davis (1955) proved convergence for analytic integrands which are regular in an ellipse centred at $\frac{1}{2}(a + b)$, with semimajor axis on the x-axis and of length $\frac{5}{8}(b - a)$, and semiminor axis of length $\frac{3}{8}(b - a)$.

Unfortunately it seems that the results of Kusmin and Pólya may have wrongly contributed to a mistrust of other high-order formulae, such as Gauss-Legendre, Clenshaw-Curtis and the more efficient of the polynomial extrapolation methods, even though all these are known to converge as the order increases for bounded integrands which are continuous on $[a, b]$, and even under such weaker conditions as Riemann integrability. A more important point, however, and one which is not universally appreciated, is that the question of convergence in the limit is irrelevant in practice since we simply wish to achieve a specified, non-infinitesimal precision together with an estimated error which is sufficiently small.

3. Negative weights

A more practical objection which has been levelled against the high-order Newton-Cotes formulae is the possible magnification, resulting from negative weights, of any rounding errors in the integrand values. In fact the order must be high for this magnification to be potentially serious, and Grant (1964, Ch. 3) has given a simple modification of the Newton-Cotes formulae, involving two additional integrand values, which approximately doubles the minimum order of formula for which negative weights occur. Most of the other possible general-purpose formulae do not have negative weights at all, while one class which does, namely the polynomial extrapolation methods favoured by Bulirsch and Stoer, does not suffer from serious magnification (Bauer *et al.*, 1963; Bulirsch, 1964; Bulirsch and Stoer, 1964; Oliver, 1971a). As with convergence, therefore, a former possible disadvantage of high-order formulae arising with one particular class of methods is largely irrelevant.

4. Estimation of derivatives

The classical error expressions for the $2m + 1$ and $2m + 2$ point Newton-Cotes formulae, valid when $f(x)$ is $2m + 2$ times continuously differentiable over $[a, b]$, are of the form

$$I-(b-a)\sum_{i=0}^{2m} A_i^{(2m)} f\left\{a+\frac{i(b-a)}{2m}\right\} = C_{2m}(b-a)^{2m+3} f^{(2m+2)}(\xi) \quad (a \leq \xi \leq b) \quad (2a)$$

$$I-(b-a)\sum_{i=0}^{2m+1} A_i^{(2m+1)} f\left\{a+\frac{i(b-a)}{2m+1}\right\} = C_{2m+1}(b-a)^{2m+3} f^{(2m+2)}(\xi) \quad (a \leq \xi \leq b) \quad (2b)$$

where the A_i are appropriate weights and the C_i are constants independent of $f(x)$ (Krylov, 1962). While a numerical analyst may sometimes be able to establish a strict error bound by utilising these expressions and bounding the derivative, this course is rarely followed by the general user.

One possible alternative, for which algorithms have been devised (Legras, 1967), is to estimate numerically the approximate derivative at the equally-spaced pivotal points by differentiating an interpolating polynomial, but Ralston (1960, p. 244) has remarked that, 'From a practical point of view higher derivatives are generally very difficult to estimate so that the use of high order Newton-Cotes formulas is not recommended.' He also levelled the same objection (Ralston, 1960, p. 246) against the Gauss-Legendre N-point formula, for which the remainder term is (Krylov, 1962, p. 109),

$$I-\tfrac{1}{2}(b-a)\sum_{i=1}^{N} W_i^{(N)} f\{\tfrac{1}{2}(a+b)+\tfrac{1}{2}x_i^{(N)}(b-a)\} = \frac{(N!)^4}{(2N+1)(2N!)^3}(b-a)^{2N+1} f^{(2N)}(\xi)\,(a \leqq \xi \leqq b)\ , \quad (3)$$

and repeated it more recently (Ralston, 1965, p. 114): 'Since rapid growth of derivatives may start for quite low derivatives and because of the difficulty in estimating high derivatives, high-order [Gaussian] quadrature formulas are seldom used.'

The fact that one possible method of error estimation is unsatisfactory is, however, surely no reason for condemning the quadrature formula; rather it means that alternative techniques should if possible be devised, and since this has indeed been done (see, for example, the algorithm for Gaussian quadrature of Tompa, 1967, or that for extrapolation of Bulirsch and Stoer, 1967), Ralston's argument is no longer valid.

5. Error expressions involving derivatives

Note that in the above quotation, Ralston mentions a second explanation for the supposed infrequent use of high-order Gaussian formulae, and since it would appear to be widely held and to apply also to other high-order formulae, it merits close study. Essentially the reasoning is that because high-order derivatives may assume large magnitudes in $[a, b]$, which is particularly likely to occur when $f(z)$ is singular at a point near the real interval $[a, b]$, the error term (3) may be much larger than would be the case for the commonly-used low-order formulae with the same number of abscissae. Although Ralston's later summary (1965, p. 128) of the disadvantages of using high-order formulae for continuous integrands indicates that he attaches more importance to the practical difficulty of establishing a realistic error estimate, O'Hara and Smith (1969) reiterate this other implication of high-order derivatives appearing in error expressions.

Certainly Stern (1967) develops this argument explicitly; after giving the remainders based on (2) and (3) for the composite trapezoidal and Simpson's rules and for Gaussian quadrature over [0, 1], he states: 'From these error bounds we see that the trapezoid rule is going to be better than the others if $f''(x)$ is bounded in [0, 1], but $f^{(iv)}(x)$ and higher derivatives are not. Similarly we expect Simpson's rule to be better if $f^{(iv)}(x)$ is bounded and higher derivatives are not. In any case if the 'size' of the $2m$th derivative increases too rapidly with m, we should use Simpson's rule rather than Gaussian quadrature.' The same philosophy is, to a much lesser extent, implicit in the definition of 'best' quadrature formulae adopted by Fraser and Wilson (1966) and Sard (1949).

However, this conclusion is clearly a non-sequitur, since the fact that the appropriate derivative is of large or unbounded magnitude over part of the range $[a, b]$ does not necessarily imply that the same is true at the particular point $x = \xi$ occurring in an error expression such as (2) or (3), although certainly the result would follow from the existence of a large *lower* bound on the magnitude of the derivative over $[a, b]$. In the absence of such a lower bound, one can simply say that such a classical error expression yields a large upper bound on the error, but with no indication of the sharpness of this bound.

6. Bounds based on low-order derivatives

Such an unsatisfactory conclusion indicates that perhaps alternative error expressions or bounds should be sought in order to compare the effectiveness of various formulae for such integrands, and this is even more true, as we shall see below, if a singularity in a derivative of some order renders the usual error expression involving that derivative invalid. One such approach, for example, is that of Davis and Rabinowitz (1954); while another, primarily motivated by the desire to apply polynomial extrapolation when the integrand or its derivatives of low order are singular at some point, is that of Fox (1967), Fox and Hayes (1970), and Miller (1968). We shall concentrate here, however, on certain error bounds developed by Stroud and Secrest which involve only low-order derivatives, even for high-order interpolatory quadrature formulae, since these bounds facilitate the comparison of formulae of different orders when the integrand is such that the usual error expressions discussed above are unhelpful.

These classical error expressions (2) and (3) for the Newton-Cotes and Gaussian formulae, and the corresponding expression (Bulirsch, 1964) for polynomial extrapolation of the trapezoidal rule are derived on the assumption that $f(x)$ is sufficiently many times continuously differentiable over $[a, b]$. Thus, as Rabinowitz (1968) remarked in connection with the Gauss rule, 'this seems to indicate that such a rule is not efficient for integrating functions of low order continuity, i.e. functions which have only a few derivatives in the entire interval of integration.' Similarly, when considering the merits of interval bisection in the Gauss and Clenshaw-Curtis methods, Wright (1966) suggested that for non-analytic functions, such as functions with discontinuities in some low-order derivative, a simpler formula is usually beneficial. However, Rabinowitz (1968) continued with the reminder that Stroud and Secrest (1966) have shown that Gaussian integration is efficient even in these circumstances, and it is the implications of this work which we now explore.

These authors, and also Stroud (1965, 1966), recalled that the Taylor expansion for functions $f(x)$ which are $k-1$ times continuously differentiable on $[a, b]$ and possess a piecewise continuous derivative of order $k \geq 1$ satisfying

$$|f^{(k)}(x)| \leq M_k \qquad (a \leq x \leq b) \tag{4}$$

can be used to show (Peano's theorem) that the remainder $R(f)$ in any N-point quadrature formula,

$$\int_a^b f(x)\, dx = \tfrac{1}{2}(b-a) \sum_{i=1}^{N} A_i f\{\tfrac{1}{2}(a+b)+\tfrac{1}{2}x_i(b-a)\}+R(f), \tag{5}$$

which is exact for polynomials of degree up to $m \geq k-1$ satisfies

$$R(f) = \tfrac{1}{2}(b-a)\int_{-1}^{1} F_k(t)\frac{d^k}{dt^k} f\{\tfrac{1}{2}(a+b)+\tfrac{1}{2}t(b-a)\}\, dt \tag{6}$$

for a known function $F_k(t)$ depending only on A_i and x_i. This leads to the attainable upper bound,

$$|R(f)| \leq \left(\frac{b-a}{2}\right)^{k+1} M_k e_k , \tag{7}$$

$$e_k \equiv \int_{-1}^{1} |F_k(t)|\, dt , \tag{8}$$

and the minimum possible values of e_1 and e_2 are (Krylov, 1962, Ch. 8):

$$e_1 = 1/N, \qquad e_2 = 1/\{2(N-1)+\sqrt{3}\}^2 . \tag{9}$$

e_1 and e_2 are known explicitly (Krylov, 1962; Stroud, 1965, 1966; Stroud and Secrest, 1966) for the composite trapezoidal rule,

$$e_1 = 1/(N-1) , \qquad e_2 = 2/3(N-1)^2 ; \tag{10}$$

the composite mid-point rule,

$$e_1 = 1/N , \qquad e_2 = 1/3N^2 ; \tag{11}$$

and the composite Simpson's rule,

$$e_1 = 10/9(N-1) , \qquad e_2 = 32/81(N-1)^2 ; \tag{12}$$

while Stroud (1965) and Stroud and Secrest (1966) have calculated e_k for representative values of k and N for the Gauss and Romberg formulae. In particular, they found that for the Gauss N-point formula, the ratio of e_1 to the 'best' value (9) appears to approach a limit $<1{\cdot}5$ as N increases and the ratio for e_2 a limit $<2{\cdot}0$, with similar, though in some cases weaker, bounds for Romberg. Alternative error bounds for Gaussian quadrature have since been computed by Rabinowitz (1968), but although sharper in some cases they are not quite so well-suited to our present purpose.

It is assumed, though they have not been calculated, that the values of e_1 and e_2 for the polynomial extrapolation methods recommended by Bulirsch and Stoer will lie between those for Gauss and Romberg, though in any case these methods produce estimates of low order as well, commencing with the trapezoidal values; this may well be one of their principal advantages, provided that in a practical algorithm *every* estimate is regarded as constituting a potential approximation to I, and not just those of high order.

For the Clenshaw-Curtis N-point formula (Clenshaw and Curtis, 1960), for which the abscissae and weights in (5) are

$$x_i = \cos (i-1)\,\pi/(N-1) \qquad (i = 1, 2, \ldots, N) , \tag{13}$$

$$\left.\begin{matrix} 2A_i\,(i = 1, N) \\ A_i\,(i = 2, \ldots, N-1)\end{matrix}\right\} = -\frac{2}{N-1}\sum_{r=0}^{N-1}{}'' \frac{\{1+(-1)^r\}}{(r^2-1)} T_r(x_i), \tag{14}$$

we have computed values of e_1 and e_2 using the 'exact' method of Stroud and Secrest (1966, p. 65) applied to

Table 1 e_1 **for various** N**-point formulae**

$N-1$	TRAPEZOIDAL	SIMPSON	BEST	GAUSS	C − C	C − C:BEST
4	$2{\cdot}500 \times 10^{-1}$	$2{\cdot}778 \times 10^{-1}$	$2{\cdot}000 \times 10^{-1}$	$2{\cdot}252 \times 10^{-1}$	$3{\cdot}099 \times 10^{-1}$	1·550
6	$1{\cdot}667 \times 10^{-1}$	$1{\cdot}852 \times 10^{-1}$	$1{\cdot}429 \times 10^{-1}$	$1{\cdot}648 \times 10^{-1}$	$2{\cdot}065 \times 10^{-1}$	1·445
8	$1{\cdot}250 \times 10^{-1}$	$1{\cdot}389 \times 10^{-1}$	$1{\cdot}111 \times 10^{-1}$	$1{\cdot}300 \times 10^{-1}$	$1{\cdot}547 \times 10^{-1}$	1·392
12	$8{\cdot}333 \times 10^{-2}$	$9{\cdot}259 \times 10^{-2}$	$7{\cdot}692 \times 10^{-2}$	$9{\cdot}142 \times 10^{-2}$	$1{\cdot}030 \times 10^{-1}$	1·339
16	$6{\cdot}250 \times 10^{-2}$	$6{\cdot}944 \times 10^{-2}$	$5{\cdot}882 \times 10^{-2}$	$7{\cdot}051 \times 10^{-2}$	$7{\cdot}718 \times 10^{-2}$	1·312
32	$3{\cdot}125 \times 10^{-2}$	$3{\cdot}472 \times 10^{-2}$	$3{\cdot}030 \times 10^{-2}$	$3{\cdot}683 \times 10^{-2}$	$3{\cdot}856 \times 10^{-2}$	1·273
64	$1{\cdot}562 \times 10^{-2}$	$1{\cdot}736 \times 10^{-2}$	$1{\cdot}538 \times 10^{-2}$	$1{\cdot}913 \times 10^{-2}$ ($N = 64$)	$1{\cdot}928 \times 10^{-2}$	1·253
128	$7{\cdot}812 \times 10^{-3}$	$8{\cdot}681 \times 10^{-3}$	$7{\cdot}752 \times 10^{-3}$	$9{\cdot}601 \times 10^{-3}$ ($N = 128$)	$9{\cdot}638 \times 10^{-3}$	1·243

Table 2 e_2 **for various** N**-point formulae**

$N-1$	TRAPEZOIDAL	SIMPSON	BEST	GAUSS	C − C	C − C:BEST
4	$4{\cdot}167 \times 10^{-2}$	$2{\cdot}469 \times 10^{-2}$	$1{\cdot}056 \times 10^{-2}$	$1{\cdot}439 \times 10^{-2}$	$2{\cdot}847 \times 10^{-2}$	2·696
6	$1{\cdot}852 \times 10^{-2}$	$1{\cdot}097 \times 10^{-2}$	$5{\cdot}303 \times 10^{-3}$	$7{\cdot}625 \times 10^{-3}$	$1{\cdot}214 \times 10^{-2}$	2·288
8	$1{\cdot}042 \times 10^{-2}$	$6{\cdot}173 \times 10^{-3}$	$3{\cdot}180 \times 10^{-3}$	$4{\cdot}723 \times 10^{-3}$	$6{\cdot}724 \times 10^{-3}$	2·114
12	$4{\cdot}630 \times 10^{-3}$	$2{\cdot}743 \times 10^{-3}$	$1{\cdot}510 \times 10^{-3}$	$2{\cdot}327 \times 10^{-3}$	$2{\cdot}956 \times 10^{-3}$	1·958
16	$2{\cdot}604 \times 10^{-3}$	$1{\cdot}543 \times 10^{-3}$	$8{\cdot}788 \times 10^{-4}$	$1{\cdot}382 \times 10^{-3}$	$1{\cdot}657 \times 10^{-3}$	1·885
32	$6{\cdot}510 \times 10^{-4}$	$3{\cdot}858 \times 10^{-4}$	$2{\cdot}314 \times 10^{-4}$	$3{\cdot}764 \times 10^{-4}$	$4{\cdot}127 \times 10^{-4}$	1·783
64	$1{\cdot}628 \times 10^{-4}$	$9{\cdot}645 \times 10^{-5}$	$5{\cdot}942 \times 10^{-5}$	$1{\cdot}015 \times 10^{-4}$ ($N = 64$)	$1{\cdot}031 \times 10^{-4}$	1·735
128	$4{\cdot}069 \times 10^{-5}$	$2{\cdot}411 \times 10^{-5}$	$1{\cdot}505 \times 10^{-5}$	$2{\cdot}556 \times 10^{-5}$ ($N = 128$)	$2{\cdot}576 \times 10^{-5}$	1·711

$$k!\,F_k(t) = (1-t)^k - k \sum_{j=i}^{N} A_j(x_j-t)^{k-1} \quad (x_{i-1} \leqq t \leqq x_i) \ . \qquad (15)$$

As **Tables 1 and 2** show, the resulting bounds are weaker than for Gauss, but not much weaker than the corresponding ones for the composite trapezoidal and Simpson's rules, and the factors by which they differ are insignificant in terms of error bounds. It is interesting to note that unlike Gauss, the ratios of the Clenshaw-Curtis values to the best values of e_1 and e_2 appear to decrease as N increases, at least for the values of N in the tables.

If the integrand does not even possess a piecewise continuous first-order derivative, then it may not be possible to derive error expressions or bounds of the above type, even for the trapezoidal or mid-point rules, although Baker (1968) has shown that a weaker Lipschitz condition may suffice. In the absence of any theoretical basis for selection in this special case, the simplicity of the trapezoidal or mid-point rules may well be the only significant criterion, as Stroud (1965) suggests.

For integrands which do have continuous low-order derivatives, however, the error is thus subject to very similar attainable bounds in terms of these derivatives for both high-order and low-order methods of the above types, irrespective of whether high-order derivatives of the integrand are discontinuous, attain large magnitudes over $[a, b]$, or are unbounded. We have shown, therefore, that the presence in the classical remainder term for some quadrature formulae of a high-order derivative or, in the case of the Clenshaw-Curtis formula, the fact that it is exact for high-degree polynomial integrands, does not necessarily mean that the error will be relatively large for such badly-behaved integrands, and so this particular argument against the use of high-order quadrature formulae is again invalid.

There is, however, the related question of whether the non-existence of a high-order derivative or its attainment of a large magnitude will affect the error estimation procedure. If the procedure is based on an error expression involving this derivative then certainly it may be unsatisfactory, but as mentioned earlier, this points to the need for a more reliable error estimation process, rather than to any inherent disadvantage of high-order quadrature methods.

7. Relative efficiency

Although the quoted objections to high-order formulae on the grounds of their inefficiency were primarily concerned with integrands which are 'badly-behaved' in the sense considered above, their relative computational efficiency for more general integrands also merits careful study.

The fact that certain quadrature formulae are better suited than others to specific classes of problem has little relevance to this discussion in the present state of the art, since in practice the general user is often unable or unwilling to classify his problem. This is of course unfortunate, since some knowledge of relevant characteristics of the problem, such as the existence and nature of singularities in the integrand or its low-order derivatives, might suggest one of the specially designed approaches of Bulirsch (1964), Davis and Rabinowitz (1965), Eisner (1967), Fox (1967), Fox and Hayes (1970), Hunter (1967), Miller (1968), or Smith and Lyness (1969), while it would be possible to avoid the ineffective exercise of extrapolating the trapezoidal values in the case of an integrand whose derivatives were known to be periodic over $[a, b]$ (Davis, 1959; Bauer *et al.*, 1963; Thacher, 1964).

Were general-purpose routines to be developed in which the optimum method is selected on the basis of interrogation of the user, or of an analytical or numerical investigation of relevant properties of the integrand, then a comparison of special-purpose methods would certainly be relevant, but as Lyness (1969) agrees, this is not yet the case. Consequently we desire a single method which minimises the total computation over all the problems to which it will be applied. Difficult though this is to study, some pertinent observations can be made.

The principal measure of computational effort is either the number of distinct integrand evaluations required to achieve the specified accuracy, usually in those cases where the appropriate formula for doing this is somehow known beforehand, or, and more probably, the number required to yield an error estimate of sufficiently small magnitude. Note incidentally that, since the accuracy required may occasionally be low, it is desirable that a general-purpose method should be capable of producing low accuracy with a small number of integrand evaluations when appropriate, as well as offering high accuracy when required. The usual algorithms based on the Clenshaw-Curtis and Gauss-Legendre formulae offer this facility, since the order of the formula is progressively increased; though with the latter the fact that formulae of different order have no abscissae in common must be remembered when assessing the number of integrand evaluations. Extrapolation algorithms based on low-order formulae such as the trapezoidal rule are also very satisfactory from this standpoint.

Some empirical comparisons of relative efficiency have been made on the above basis with selected and inevitably unrepresentative problems, and although no definite consensus has yet emerged, there is certainly no evidence that low-order methods are essentially superior (Bauer *et al.*, 1963; Bulirsch and Stoer, 1967; Oliver, 1971a; Rabinowitz, 1966; Stroud, 1965; Tompa, 1967). Indeed the experience with extrapolation methods, in which approximations of both high and low order are obtained from the same integrand values, tends to indicate the reverse if anything. The number of distinct evaluations is not, however, the sole factor in assessing computational effort; in the Gauss formulae the abscissae and weights must be provided explicitly, while the Clenshaw-Curtis formulae require their provision or calculation. The extrapolation methods do not of course suffer this disadvantage.

Even taking these factors into account, no substantial published evidence of an empirical nature is known to support the belief that quadrature methods in which the order is restricted to be small consistently require less computational effort for a given accuracy; again, therefore, one of the possible disadvantages of high-order methods has no apparent foundation in the literature.

8. Adaptive integration

We have hitherto implicitly assumed that the problem involves the selection of a particular integration formula to be applied uniformly over the fixed interval $[a, b]$. This includes algorithms in which satisfaction of the error criterion is sought by using a class of related formulae of progressively higher order, or by using a composite rule of a specific order and progressively more abscissae. The composite trapezoidal and Simpson's rules fall into the latter category, while the formulae of Gauss-Legendre and Clenshaw-Curtis can be used in either way, and extrapolation processes encompass both approaches in a single algorithm.

Some problems are 'badly-behaved' with respect to such a quadrature method only over a small part of $[a, b]$, in the sense that the contribution to the discretisation error from that part of the interval is very much larger, proportional to the sub-interval length, than from the remainder of $[a, b]$. If this characteristic were known beforehand, the error criterion might well be satisfied with fewer integrand evaluations by suitably subdividing the interval, and using different formulae in each with pro rata maximum errors. Since the optimum combination of interval subdivisions and integration formulae are rarely known in advance, however, the decision to apply integration formulae non-uniformly over $[a, b]$ means that the algorithm itself should determine both the partitioning and the formula for each sub-interval.

No such algorithms are yet known; instead the current adaptive algorithms adopt the less ambitious target of determining the optimum interval subdivision for a specified integration rule, perhaps the trapezoidal or Simpson's rule (Kuncir, 1962; Lyness, 1970; McKeeman, 1962, 1963; McKeeman and Tesler, 1963; Schweikert, 1970) or a Clenshaw-Curtis formula (O'Hara and Smith, 1969). In order to minimise the loss of integrand values when a sub-interval is found to be unacceptable in terms of its proportional error contribution, all such algorithms use quadrature formulae with a small number of abscissae, often equally spaced to allow re-use if the interval is subdivided equally, and these formulae are thus necessarily of low order.

While such algorithms may be more efficient than the uniform application of a high-order formula for the special class of integrands described above, there is no evidence to indicate that this will be true for other classes of problem, and indeed it will clearly not be the case if the integrand behaves uniformly over $[a, b]$—consider, for example, a polynomial integrand of slightly higher degree than the order of the adaptive formula. The existence of such adaptive methods is therefore irrelevant to our discussion of high-order formulae in the sense that they are special-purpose rather than general-purpose methods, though this is not entirely true since an adaptive method might be considered for a universal algorithm on the grounds that, though inefficient for 'well-behaved' integrands, it will cope successfully with certain difficult problems.

Much more useful, however, would be a more flexible algorithm which not only subdivides the interval of integration selectively, but also chooses the most efficient order of formula for each sub-interval on the evidence only of computed information concerning the integrand. Such a general-purpose algorithm, which might perhaps be based on an examination of the computed coefficients in a Chebyshev series approximation to the integrand in each sub-interval (Oliver, 1971b), would of course make it unnecessary to base a universal adaptive algorithm on one, or a small number as with O'Hara and Smith (1969), of formulae of particular, pre-chosen orders. In the interim, however, formulae of high order should certainly not be excluded from consideration, and might, perhaps, be best combined with low-order approximations as in a non-adaptive extrapolation method.

9. Conclusions

We have discussed the various possible arguments against the use of high-order quadrature formulae, and shown that they are either irrelevant to the modern usage of such formulae, or are unsupported by published evidence. Consequently there is no reason why a universal general-purpose algorithm of the non-adaptive type should not be based on high-order formulae, combined with low-order formulae if this should be thought desirable, as in an extrapolation method. The ultimate requirement is, however, for more ambitious adaptive algorithms which can utilise available analytical or numerical information to optimise the type and order of the integration formula as well as the partitioning of the range of integration.

Acknowledgements

The author gratefully acknowledges the many helpful comments of Dr. J. C. P. Miller of the Cambridge University Computer Laboratory on the original draft of this paper.

References

Baker, C. T. H. (1968). On the nature of certain quadrature formulas and their errors, *SIAM J. Numer. Anal.*, Vol. 5, pp. 783-804.
Bauer, F. L., Rutishauser, H., and Stiefel, E. (1963). New aspects in numerical quadrature, *Proc. Symp. Appl. Math.*, *Amer. Math. Soc.*, Vol. 15, pp. 199-218.
Bulirsch, R. (1964). Bemerkungen zur Romberg-Integration, *Numer. Math.*, Vol. 6, pp. 6-16.
Bulirsch, R., and Stoer, J. (1964). Fehlerabschätzungen und Extrapolation mit rationalen Funktionen bei Verfahren vom Richardson-Typus, *Numer. Math.*, Vol. 6, pp. 413-427.
Bulirsch, R., and Stoer, J. (1967). Numerical quadrature by extrapolation, *Numer. Math.*, Vol. 9, pp. 271-278.
Clenshaw, C. W., and Curtis, A. R. (1960). A method for numerical integration on an automatic computer, *Numer. Math.*, Vol. 2, pp. 197-205.
Davis, P. J. (1955). On a problem in the theory of mechanical quadratures, *Pacific J. Math.*, Vol. 5, pp. 669-674.
Davis, P. J. (1959). On the numerical integration of periodic analytic functions. In: *On Numerical Approximation*, R. Langer (Ed). Madison: University of Wisconsin Press, pp. 45-59.
Davis, P. J., and Rabinowitz, P. (1954). On the estimation of quadrature errors for analytic functions, *M.T.A.C.*, Vol. 8, pp. 193-203.
Davis, P. J., and Rabinowitz, P. (1965). Ignoring the singularity in approximate integration, *SIAM J. Numer. Anal.*, Vol 2, pp. 367-383.
Davis, P. J., and Rabinowitz, P. (1967). *Numerical Integration*. London: Blaisdell.
Eisner, E. (1967). Numerical integration of a function that has a pole, *CACM*, Vol. 10, pp. 239-243.
Fox, L. (1967). Romberg integration for a class of singular integrands, *The Computer Journal*, Vol. 10, pp. 87-93.
Fox, L., and Hayes, L. (1970). On the definite integration of singular integrands, *SIAM Rev.*, Vol. 12, pp. 449-457.
Fraser, W., and Wilson, M. W. (1966). Remarks on the Clenshaw-Curtis quadrature scheme, *SIAM Rev.*, Vol. 8, pp. 322-327.
Grant, J. A. (1964). *Topics in Numerical Quadrature*. Ph.D. Thesis, University of Cambridge. (Unpublished).
Hunter, D. B. (1967). Romberg's method for certain integrals involving a singularity, *BIT*, Vol. 7, pp. 200-205.
Krylov, V. I. (1962). *Approximate Calculation of Integrals*. New York: Macmillan.
Kuncir, G. F. (1962). Algorithm 103: Simpson's rule integrator, *CACM*, Vol. 5, p. 347.
Kusmin, R. O. (1931). Zur Theorie der mechanischen Quadraturen, *Nachr. Polyt. Inst. Leningrad*, Vol. 33, pp. 5-14.
Legras, J. (1967). Calcul des quadratures. In: *Procédures ALGOL en Analyse Numerique*. Paris: Centre National de la Recherche Scientifique, pp. 209-244.
Lyness, J. N. (1969). The effect of inadequate convergence criteria in automatic routines, *The Computer Journal*, Vol. 12, pp. 279-281.
Lyness, J. N. (1970). Algorithm 379: Squank (Simpson quadrature used adaptively—noise killed), *CACM*, Vol. 13, pp. 260-263.
McKeeman, W. M. (1962). Algorithm 145: Adaptive numerical integration by Simpson's rule, *CACM*, Vol. 5, p. 604.
McKeeman, W. M. (1963). Algorithm 198: Adaptive integration and multiple integration, *CACM*, Vol. 6, pp. 443-444.
McKeeman, W. M., and Tesler, L. (1963). Algorithm 182: Nonrecursive adaptive integration, *CACM*, Vol. 6, p. 315.
Miller, J. C. P. (1968). Neville's and Romberg's processes: a fresh appraisal with extensions, *Phil. Trans. Roy. Soc.*, Vol. 263A, pp. 525-562.
O'Hara, H., and Smith, F. J. (1969). The evaluation of definite integrals by interval subdivision, *The Computer Journal*, Vol. 12, pp. 179-182.
Oliver, J. (1971a). The efficiency of extrapolation methods for numerical integration, *Numer. Math.*, Vol. 17, pp. 17-32.
Oliver, J. (1971b). A practical strategy for the Clenshaw-Curtis quadrature method, *J.I.M.A.*, (In press).
Pólya, G. (1933). Über die Konvergenz von Quadraturverfahren, *Math. Z.*, Vol. 37, pp. 264-286.
Rabinowitz, P. (1966). Automatic integration of a function with a parameter, *CACM*, Vol. 9, pp. 804-806.
Rabinowitz, P. (1968). Error bounds in Gaussian integration of functions of low-order continuity, *Math. Comp.*, Vol. 22, pp. 431-434.

RALSTON, A. (1960). Methods for numerical quadrature. In: *Mathematical Methods for Digital Computers*, A. Ralston and H. S. Wilf (Eds) New York: John Wiley, pp. 242-248.
RALSTON, A. (1965). *A First Course in Numerical Analysis*. New York: McGraw-Hill.
SARD, A. (1949). Best approximate integration formulas, *Amer. J. Math.*, Vol. 71, pp. 80-91.
SCHWEIKERT, D. G. (1970). A comparison of error improvement estimates for adaptive trapezoid integration, *CACM*, Vol. 13, pp. 163-166.
SMITH, W. E., and LYNESS, J. N. (1969). Applications of Hilbert transform theory to numerical quadrature, *Math. Comp.*, Vol. 23, pp. 231-252.
STERN, M. D. (1967). Optimal quadrature formulae, *The Computer Journal*, Vol. 9, pp. 396-403.
STROUD, A. H. (1965). Error estimates for Romberg quadrature, *SIAM J. Numer. Anal.*, Vol. 2, pp. 480-488.
STROUD, A. H. (1966). Estimating quadrature errors for functions with low continuity, *SIAM J. Numer. Anal.*, Vol. 3, pp. 420-424.
STROUD, A. H., and SECREST, D. (1966). *Gaussian quadrature formulas*. Englewood Cliffs, N.J.: Prentice-Hall.
THACHER, H. C. (1964). Remark on Algorithm 60: Romberg integration. *CACM*, Vol. 7, pp. 420-421.
TOMPA, H. (1967). Gaussian numerical integration of a function depending on a parameter, *The Computer Journal*, Vol. 10, pp. 204-205.
WRIGHT, K. (1966). Series methods for integration, *The Computer Journal*, Vol. 9, pp. 191-199.