

tions showed: If $t < .001$ and $|\Delta x_k| < .02$, then $|x_{k+1} - x^*| < .00035$ and, as $|x_{k+1} - x^*|$ and $|x^* - x|$ depend on t and assume their maxima at different points, $|x_{k+1} - x| < .0004$ (even $< .00037$). Thus, the iteration may be ended as soon as $|\Delta x_k| < .02$, if an accuracy of $.0004$ is sufficient. In the same way, if $t < .001$ and $|\Delta x_k| < .0035$, then $|x_{k+1} - x^*| < .00052$, $|x_{k+1} - x| < .0001$. As initial values, x_0 , of the iteration process are recommended in the first case ($\epsilon_M < .0004$) $\psi(.000962) - .020 = -3.122$ and in the second case ($\epsilon_M < .0001$) $\psi(.0003624) - .0035 = -3.3835$ in order to cover the greatest possible t -interval with a single iteration. In the first case, 3 iterations are needed for $t = 0.0001$ and 4 for $t = 0.00001$.

A program to compute $\psi(t)$ was written for the Siemens 2002 computer, using linear approximations, rational functions $R_i(t)$ and the iteration process, $\epsilon_M = 0.0004$. One iteration needed about the time of 40 multiplications. However, in most values of t , the iteration process is not involved, and the average computation time was approximately that of four and a half multiplications. This program was part of a multidimensional integration problem [1], where more than 900 000 normal deviates were computed.

University of Connecticut
Storrs, Connecticut

1. FRIEDRICH GEBHARDT, "On the risk of some strategies for outlying observations," submitted to *Ann. Math. Statist.*

2. MERVIN E. MULLER, "An inverse method for the generation of random normal deviates on large-scale computers," *MTAC*, v. 12, 1958, p. 167-174.

3. MERVIN E. MULLER, "A comparison of methods for generating normal deviates on digital computers," *J. Assoc. Comput. Mach.*, v. 6, 1959, p. 376-383.

Optimum Runge-Kutta Methods

By T. E. Hull and R. L. Johnston

Abstract. The optimum Runge-Kutta method of a particular order is the one whose truncation error is a minimum. Various measures of the size of the truncation error are considered. The optimum method is practically independent of the measure being used. Moreover, among methods of the same order which one might consider using the difference in size of the estimated error is not more than a factor of 2 or 3. These results are confirmed in practice insofar as the choice of optimum method is concerned, but they underestimate the variation in error between different methods.

1. **Introduction.** For the solution of

$$(1) \quad y' = f(x, y), \quad y(x_0) = y_0$$

the general Runge-Kutta method of order m uses the formula

$$(2) \quad y_{n+1} = y_n + \sum_{i=1}^m w_i k_i$$

Received September 24, 1963, revised November 19, 1963. This research was supported in part by the Defence Research Board of Canada.

to obtain y_r , which is an approximation to $y(x_r)$. The w 's are constants and

$$k_i = hf \left(x_n + \alpha_i h, y_n + \sum_{j=1}^{i-1} \beta_{ij} k_j \right)$$

with $h = x_{n+1} - x_n$ and $\alpha_1 = 0$.

The w 's, α 's and β 's are constrained by requiring the true solution of (1) to satisfy (2) except for a truncation error which is $O(h^p)$ as $h \rightarrow 0$ for as large a power of p as possible.

Typically this still leaves one or two parameters to be determined. Originally these parameters seem to have been chosen only for the sake of simplicity in hand computation. They have also been chosen in such a way that round off error and storage requirements are minimized [1]. On modern computers it is more important to choose these parameters so that the truncation error is a minimum. The corresponding methods will be called optimum, but of course what is optimum will depend on what is used as a measure of the size of the truncation error.

2. Criteria for Optimum. Complete solutions of the constraining equations are known for the cases $m = 2, 3, 4$. For example they are given by Ralston [5] whose notation will be followed here and whose results will be compared with ours.

For purposes of illustration we will consider the two parameter family of fourth order methods. The truncation error in this case turns out to be $T = bh^5 + O(h^6)$ where

$$b = \frac{1}{1440} [b_1 D^4 f + b_2 Df D^2 f_x + b_3 f_x D^3 f + b_4 Df_x D^2 f + b_5 f_x^2 D^2 f + b_6 f_{xx} (Df)^2 + b_7 f_x Df Df_x + b_8 f_x^3 Df],$$

with

$$D = \frac{\partial}{\partial x} + f \frac{\partial}{\partial y},$$

and

$$b_1 = 5u + 5v - 10uv - 3,$$

$$b_2 = 6(5v - 3),$$

$$b_3 = 4(3 + 10uv - 5u - 5v),$$

$$b_4 = 6(5u - 2),$$

$$b_5 = 6(2 - 5u),$$

$$b_6 = 3 \left\{ 12 - 5 \left[\frac{v(v-u)}{(1-v)(1-2u)} + \frac{(1-u)(3-4v)^2}{(1-v)(6uv-4u-4v+3)} \right] \right\},$$

$$b_7 = 12(2 - 5v),$$

$$b_8 = 12.$$

These expressions cannot be used when $u = \frac{1}{2}$ or $v = 1$, or when $6uw + 3 = 4u + 4v$. Moreover the corresponding expressions for the α 's, β 's, and w 's (which are needed in (2)) cannot be used when $u = 0, 1$ or $v = 0, 1$ or when $u = v$. However there are three one parameter families of methods which correspond to $u = v = \frac{1}{2}$, $u = 1$ and $v = \frac{1}{2}$, and $u = \frac{1}{2}$, $v = 1$ respectively. Further details of these special cases are given by Ralston [5].

The expressions given above can be obtained from (5.1) and (5.6) of Ralston's paper by making use of results given by Romanelli [6, p. 114]. To avoid some subscripts we have put Ralston's $\alpha_2 = u$ and $\alpha_3 = v$.

The function $f(x, y)$ is evaluated at points where $x = x_n, x_n + uh, x_n + vh$, and $x_n + h$. Thus we would expect $0 < u, v < 1$. In fact we would expect the best values of u and v to be not close to 0 or 1, and probably with $u < v$.

Of course it is impossible to choose u and v to minimize T without regard to the function f . But for a general purpose method it is natural to consider choosing u and v so as to minimize some measure of T which is independent of f . For example, it is natural to try to minimize $\sum |b_i|$ or $\sum (b_i)^2$. Ralston minimized

$$B = 16|b_1| + 4|b_2| + |b_2 + 3b_3| + |2b_2 + 3b_3| + |b_2 + b_3| + |b_3| + 8|b_4| \\ + |b_5| + |2b_5 + b_7| + |b_5 + b_6 + b_7| + |b_6| + |2b_6 + b_7| + |b_7| + 2|b_8|.$$

This expression is obtained when the bounds given by Lotkin [4] are applied to b after the terms in b have been grouped in a certain way.

The terms in the expression for b can also be grouped differently. Moreover, other measures of the size of b could be used. But we decided to consider only the above three criteria and to find out if they led to significantly different optimum formulas. The sensitivity of each measure to changes in u and v is also of interest. Finally, our main objective is to compare the dependence of each measure on u and v with the way in which the errors produced in practice by the corresponding methods depend on u and v . It is to be hoped that one or more of these measures will reflect at least the average behaviour of the corresponding methods.

One special point should be noted. For certain values of u and v the value of b_6 is much larger than all the other b_i 's. But the term involving b_6 does not appear in T for the special case of linear differential equations because then $f_{yy} = 0$. A measure containing b_6 or b_6^2 would therefore be most inappropriate for linear differential equations, at least if any values of u and v which make b_6 large turn out to provide accurate methods for such equations. We decided to consider separately the same three measures, but with b_6 replaced by zero.

3. Computations. Programs were written to evaluate the various measures of size of truncation error for all pairs of values of u and v from the set .05, .15, \dots , .95, except that pairs with $u = v$ were omitted. The resulting 90 pairs avoided all the singularities in the above expressions for the b 's, as well as in the expressions for the parameters defining (2).

Programs were also written to solve known differential equations and to find the error associated with the method determined by each of these pairs of values of u and v .

The following problems were used:

$$\begin{aligned}
 y' &= -y + \sin 2x, & x_0 &= 0, & y_0 &= -0.4, & \left(\text{here } \frac{\partial f}{\partial y} < 0 \right), \\
 y' &= y + \sin 2x, & x_0 &= 0, & y_0 &= -0.4, & \left(\text{here } \frac{\partial f}{\partial y} > 0 \right), \\
 y' &= y + \cos x, & x_0 &= 0, & y_0 &= 1.0, & \left(\text{here } \frac{\partial f}{\partial y} \text{ oscillates} \right), \\
 y' &= y - 2x/y, & x_0 &= 0, & y_0 &= 1.0, & (\text{non-linear}).
 \end{aligned}$$

In each case we took $h = .1$, and printed out the error of largest magnitude which was observed in the interval $0 < x \leq 5$.

4. Results and Conclusions. The results for all six measures of size of the truncation error were the same in two essential respects. Their smallest values all occurred with the same pair $u = .35, v = .45$. However their values were not very sensitive to changes in u and v ; in each case, for all pairs in $.3 < u < v < .7$ they were not more than about twice the smallest value. The only significant difference between the various measures was that those involving b_6 were extremely large near where $6uv + 3 = 4u + 4v$, as would be expected.

The obvious conclusion is that all criteria lead to approximately the same optimum formula, namely one corresponding to values of u and v near to $.35$ and $.45$ respectively. (This result is consistent with the result $u = .4, v = .45573725$, which Ralston gives for the exact position of the optimum according to his criterion.)

What is equally significant is the conclusion that none of the criteria are particularly sensitive to the choice of parameter values, as long as these values are near to the optimum ones.

The results for the differential equations were less consistent. Nevertheless, in each case their errors had minima for methods corresponding to values of u and v which were within $.1$ of $u = .35, v = .45$. However, the errors were quite a bit more sensitive to changes in u and v than were the measures of size. For some of the pairs in $.3 < u < v < .7$ they were as much as 25 times the smallest value.

One exceptional result occurred with the first equation. The errors in this case had a second minimum for $u = .45, v = .15$, which was slightly smaller than the other minimum.

Our conclusion is that the optimum formulas suggested by the earlier criteria would in general be best in practice. However, the measures of size given by these criteria underestimate how much the errors depend on u and v .

Of course these criteria did not take account of the possibility that different terms in the truncation error might be of opposite sign and thus tend to cancel each other. With all terms being considered positive it is not surprising that the minima given by the criteria were flatter than those appearing in practice. We observed that the errors in practice were of opposite sign on opposite sides of the minima, either for the same value of u and different values of v , or vice versa.

Ralston also considered the one parameter families of fourth order methods.

He found that the associated optimum formulas had error bounds which were larger by a factor of $1\frac{1}{2}$ or 2 than the optimum for the two parameter family we have been considering.

Our computed results indicate further that the minima in these cases are again consistent, and relatively insensitive to changes in the parameter. In practice the different differential equations had errors which corresponded to these results about as they did in the two parameter case.

Analogous results were obtained by Johnston [2] for the second and third order methods. A few special cases had been considered earlier by Kuntzmann [3].

5. Acknowledgments. We wish to thank Charles Warden who wrote some of the original programs. We are also grateful to Lockheed Aircraft and to Case Institute of Technology where the calculations were performed.

University of British Columbia, now at
University of Toronto,
Toronto 5, Ontario.

University of British Columbia, now at
Case Institute of Technology,
Cleveland, Ohio.

1. S. GILL, "A process for the step-by-step integration of differential equations in an automatic digital computing machine," *Proc. Cambridge Philos. Soc.*, v. 47, 1951, p. 96-108.

2. R. L. JOHNSTON, "On optimum Runge-Kutta methods for the numerical solution of ordinary differential equations," M.A. Thesis, University of British Columbia, 1961.

3. J. KUNTZMANN, "Deux formules optimales du type de Runge-Kutta," *Chiffres*, v. 2, 1959, p. 21-26.

4. M. LOTKIN, "On the accuracy of Runge-Kutta's method," *MTAC*, v. 5, 1951, p. 128-132.

5. A. RALSTON, "Runge-Kutta methods with minimum error bounds," *Math. Comp.*, v. 16, 1962, p. 431-437; Corrigendum, v. 17, 1963, p. 488.

6. M. J. ROMANELLI, "Runge-Kutta methods for the solution of ordinary differential equations," *Mathematical Methods for Digital Computers*, A. Ralston and H. S. Wilf, editors, Wiley, N. Y., 1960.

Helpful Formulas for Integrating Polynomials in Three Dimensions

By Gilbert C. Best

In this note formulas are given for integrating the expression $x^a y^b z^c$ along a line between points P_1, P_2 , or over the surface of a plane triangle with corners at points P_1, P_2, P_3 , or over the volume of a tetrahedron with corners at points P_1, P_2, P_3, P_4 , where $P_i = (x_i, y_i, z_i)$. Since these geometric figures are of dimensions $d = 1, 2, 3$ respectively, the writing of these formulas can be facilitated by defining "generalized volume" V to be successively the length of the line, the area of the triangle or the volume of the tetrahedron in question. Then the three formulas mentioned can be given by the single equation

$$(1) \int_V x^a y^b z^c dV = V \frac{a!b!c!d!}{(a+b+c+d)!} \sum_{i=1}^{d+1} \frac{(p_i + q_i + r_i)!}{p_i!q_i!r_i!} x_i^{p_i} y_i^{q_i} z_i^{r_i} \quad d = 1, 2, 3,$$