RUNGE-KUTTA METHODS AND MINIMIZATION OF TRUNCATION ERROR

by

LAURENCE RAY NEISES

B.A., St. Mary of the Plains College, 1965

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1968

Approved by:

J. Thomas Parker
Major Professor

TABLE OF CONTENTS

INTRODUCTION

In solving an ordinary differential equation, one desires to find a "solution", $y = F(x)$, which satisfies the differential equation and whatever boundary conditions are imposed. In problems which occur in practice, it very frequently is not possible to obtain an explicit $F(x)$. In some cases, even when explicit representations are possible, the calculation of y for a series of values x_i involves a prohibitive amount of work including usually a series of approximations. Numerical solutions involve the calculation of values of y for $x = x_0, x_1, \ldots, x_n$ by means of approximation methods. Usually the tabulated values of y are desired for the x_i's which are, more often than not, equally spaced. In addition, it is customary to discuss the accuracy of each derived approximation.

Numerical methods for the solution of ordinary differential equations are generally put into two categories: predictor-corrector methods and Runge-Kutta (one step) methods. The advantages of the former methods are their greater accuracy and error estimating ability, especially in systems of any complexity. Runge-Kutta methods have the advantage of being self-starting and easy to program for the computer. Neither of these reasons is very compelling when subroutines can be written to handle systems of ordinary differential equations, nor do they overcome their disadvantages in error-estimating ability and speed relative to predictor-corrector methods. Runge-Kutta methods, however, find application in starting the computation and in changing the interval size.

Considering, then, Runge-Kutta methods for starting the computation and, perhaps, changing the interval, matters concerning stability and minimization of round-off errors are not significant. Also, on modern computers, minimiza-

tion of storage is becoming less critical. In fact, the only criterion of significance in judging Runge-Kutta methods in this context is minimization of the truncation error. It is the purpose of this study to derive Runge-Kutta methods of the second, third and fourth orders which have minimum truncation error bounds.

For the derivation, we first consider the case of integrating a single first-order differential equation and later extend the method to a system of n simultaneous first-order equations. We then derive, for the single first-order differential equation, bounds which give the least truncation error. It seems reasonable, then, to assume that methods which are best in terms of truncation error for one equation of the first-order will be at least nearly best for most systems of first-order equations.

We are concerned with the solution of the first-order differential equation problem given by

$$\frac{dy}{dx} = f(x,y) \tag{1}$$

with starting value, $y(x_0) = y_0$.

In addition we assume that $f(x,y)$ satisfies the conditions stated in the following theorem:

Theorem I — If

i. $f(x,y)$ is defined and continuous in the strip

$$a \leq x \leq b, \quad -\infty < y < \infty \quad \text{with a and b finite;}$$

ii. there exists a constant L such that for any $x \in [a,b]$ and any two numbers y and y^*,

$$\left|f(x,y) - f(x,y^*)\right| \leq L \left|y - y^*\right|$$

i.e. the Lipschitz condition of order 1 is satisfied for all x;
then, $F(x)$ is continuous and differentiable for all $x \in [a,b]$,
$F'(x) = f(x,F(x))$ and $F(x_0) = y_0$ (i.e. the initial value problem
(1) has a unique solution $y = F(x)$ for all $x \in [a,b]$).

We omit the proof of this theorem. (See Henrici [5,65].)

CLASSICAL METHODS

Derivation by Taylor Series:

The basis of all Runge-Kutta methods is the expression of the difference between the values of y at x_{n+1} and x_n by

$$y_{n+1} - y_n = \sum_{i=1}^{m} \omega_i k_i \quad , \tag{2}$$

where the ω_i are constant, and $k_i = h_n f(x_n + \alpha_i h_n, y_n + \sum_{j=1}^{i-1} \beta_{ij} k_j)$, (2a)

with $\alpha_1 = 0$, $h_n = x_{n+1} - x_n$. Given the weights ω_i , the parameters α_i , β_{ij} and using equation (2) we can solve equation (1).

We desire to determine the ω_i , α_i , β_{ij} so that the coefficients of in the Taylor series expansion of both sides of equation (2) about the point (x_n, y_n) are identical for r = 1,2...,m for some fixed m.

Expanding $F(x_{n+1})$ we have,

$$F(x_{n+1}) = F(x_n) + F'(x_n)(x_{n+1} - x_n) + \frac{F''(x_n)(x_{n+1} - x_n)^2}{2!} + \cdots .$$

Hence, $y_{n+1} - y_n = F(x_{n+1}) - F(x_n)$

$$= (x_{n+1} - x_n) F'(x_n) + \frac{(x_{n+1} - x_n)^2 F''(x_n)}{2!} + \cdots .$$

But, $h_n = x_{n+1} - x_n$,

so that

$$y_{n+1} - y_n = h_n F'(x_n) + \frac{h_n^2 F''(x_n)}{2!} + \frac{h_n^3 F'''(x_n)}{3!} + \cdots ,$$

or more generally,

$$y_{n+1} - y_n = \sum_{t=1}^{\infty} \frac{h_n^t y_n^{(t)}}{t!} \quad . \tag{3}$$

In addition, since $f(x_n, y_n) = y_n'$, assuming differentiability, it follows that

$$y_n^{(n)} = \frac{d^{n-1}}{dx^{n-1}} f(x_n, y_n) ,$$

$$y_n^{(n)} = \left(\frac{\partial}{\partial x} + \frac{dy}{dx}\frac{\partial}{\partial y}\right)^{n-1} f(x_n, y_n) ,$$

and

$$y_n^{(n)} = \left(\frac{\partial}{\partial x} + f\frac{\partial}{\partial y}\right)^{n-1} f(x_n, y_n) , \tag{4}$$

where $\frac{dy}{dx} = f(x,y) \equiv f = y'$.

Thus we have,

$$y_{n+1} - y_n = \sum_{k=0}^{\infty} \frac{h_n^{k+1}}{(k+1)!} \left(\frac{\partial}{\partial x} + f\frac{\partial}{\partial y}\right)^k f(x_n, y_n) . \tag{5}$$

Now define $D \equiv \frac{\partial}{\partial x} + f_n \frac{\partial}{\partial y}$ where $f_n = f(x_n, y_n)$; (6)

then

$$\left(\frac{\partial}{\partial x} + f\frac{\partial}{\partial y}\right)^2 f(x_n, y_n)\Big|_n = D^2 f + f_y Df\Big|_n \tag{7}$$

In general,

$$D^n = \sum_{k=0}^{n} \binom{n}{k} f^k \frac{\partial^n}{\partial x^{n-k} \partial y^k} ,$$

$$D(D^n u) = D^{n+1} u + n(Df) D^{n+1} u_y ,$$

$$D^0 u = u .$$

The expansion of equation (5) gives,

$$y_{n+1} - y_n = hf + \frac{h^2}{2!}\left(\frac{\partial}{\partial x} + f\frac{\partial}{\partial y}\right)f + \frac{h^3}{3!}\left(\frac{\partial}{\partial x} + f\frac{\partial}{\partial y}\right)^2 f + \cdots ,$$

and, using equation (7), we have

$$\begin{aligned} y_{n+1} - y_n = \Big[& hf + \frac{h^2}{2!}(Df) + \frac{h^3}{3!}\left(D^2 f + f_y Df\right) \\ & + \frac{h^4}{4!}\left(D^3 f + f_y D^2 f + f_y^2 Df + 3Df\,Df_y\right) + \frac{h^5}{5!}\left(D^4 f \right. \\ & \left. + 6Df\,D^2 f_y + 4D^2 f\,Df_y + D^2 f f_y^2 + Df f_y^3 + 3(Df)^2 f_{yy} + D^3 f f_y \right. \\ & \left. + 7 f_y Df\,Df_y)\right]\Big|_n + O(h^6) . \end{aligned} \tag{8}$$

To obtain the expansion of the right hand side of equation (2), we use the Taylor series expansion for two variables [16,227]

$$\begin{aligned} f\left[x_n + a_i h_n ,\ y_n + \left(\textstyle\sum \beta_{ij}\right) h_n f_n\right] = {} & f(x_n, y_n) + \left(a_i h_n \frac{\partial}{\partial x} + \left(\textstyle\sum \beta_{ij}\right) h_n f_n \frac{\partial}{\partial y}\right) f(x_n, y_n) \\ & + \frac{1}{2!}\left(a_i h_n \frac{\partial}{\partial x} + \left(\textstyle\sum \beta_{ij}\right) h_n f_n \frac{\partial}{\partial y}\right)^2 f(x_n, y_n) + \cdots \\ & + \frac{1}{n!}\left(a_i h_n \frac{\partial}{\partial x} + \left(\textstyle\sum \beta_{ij}\right) h_n f_n \frac{\partial}{\partial y}\right)^n f(x_n, y_n) + \cdots , \end{aligned}$$

or

$$f\left[x_n + a_i h_n ,\ y_n + \left(\textstyle\sum \beta_{ij}\right) h_n f_n\right] = \sum_{i=0}^{\infty} \frac{\left(a_i h_n \frac{\partial}{\partial x} + \left(\sum \beta_{ij}\right) h_n f_n \frac{\partial}{\partial y}\right)^i}{i!} f(x_n, y_n) . \tag{9}$$

Let $D_i = \left(a_i \frac{\partial}{\partial x} + \left(\sum \beta_{ij}\right) f_n \frac{\partial}{\partial y}\right)$; on factoring out the h_n^i , (10)

we have

$$f\left[x_n + a_i h_n ,\ y_n + \left(\textstyle\sum \beta_{ij}\right) h_n f_n\right] = \sum_{i=0}^{\infty} \frac{h_n^i D_i^i f(x_n, y_n)}{i!} . \tag{11}$$

Now, using equations (10) and (11), we can expand each k_i on the right hand side of equation (2):

$$k_1 = h_n f(x_n + \alpha_1 h_n ,\ y_n + \sum_{j=1}^{0} \beta_{1j} k_j) \ ;$$

$$k_1 = h_n f(x_n, y_n) \ . \tag{12}$$

$$k_2 = h_n f(x_n + \alpha_2 h_n ,\ y_n + \beta_{21} h_n f_n) \ . \tag{13}$$

$$k_3 = h_n f(x_n + \alpha_3 h_n ,\ y_n + \sum_{j=1}^{2} \beta_{3j} k_j) \ ,$$

$$k_3 = h_n f\left[x_n + \alpha_3 h_n ,\ y_n + (\beta_{31} + \beta_{32}) h_n f_n + \beta_{32} (k_2 - h_n f_n)\right] \ . \tag{14}$$

or, generalizing, we have,

$$k_i = h_n \sum_{t=0}^{m} \left[h_n D_i + \sum_{j=1}^{i-1} \beta_{ij} (k_j - h_n f_n) \frac{\partial}{\partial y}\right]^t \frac{f(x_n, y_n)}{t!} \ . \tag{15}$$

Using the results for k_1 , $j<i$, to write k_i as an expression in powers of h_n , we have,

$$k_1 = h_n f_n = hf\Big|_n \ , \tag{16}$$

$$k_2 = h_n \sum_{t=0}^{m} \frac{h_n^t D_i^t f(x_n, y_n)}{t!} \ ,$$

$$k_2 = h_n f(x_n, y_n) + h_n^2 D_1 f(x_n, y_n) + \frac{h_n^3 D_1^2 f(x_n, y_n)}{2!} + \frac{h_n^4 D_1^3 f(x_n, y_n)}{3!} + \frac{h_n^5 D_1^4 f(x_n, y_n)}{4!} + O(h_n^6) \ ,$$

$$k_2 = hf + h^2 D_1 f + \frac{h^3 D_1^2 f}{2!} + \frac{h^4 D_1^3 f}{3!} + \frac{h^5 D_1^4 f}{4!}\Big|_n \ . \tag{17}$$

Similarly, the derived equations for k_3 and k_4 are,

$$k_3 = hf + h^2 D_2 f + h^3 \left(\tfrac{1}{2} D_2^2 f + \beta_{32} f_y D_1 f\right) + h^4 \left[\tfrac{1}{6} D_2^3 f + \left(\tfrac{\beta_{32}}{2}\right) f_y D_2^2 f + \beta_{32} D_1 f D_2 f_y\right] + h^5 \left[\tfrac{1}{24} D_2^4 f + \left(\tfrac{\beta_{32}}{6}\right) f_y D_2^3 f + \left(\tfrac{\beta_{32}}{2}\right) D_2^2 f D_2 f_y + \left(\tfrac{\beta_{32}}{2}\right) D_1 f D_2^2 f_y\right]\Big|_n + O(h_n^6) \ . \tag{18}$$

$$k_4 = hf + h^2 D_x f + h^3\left(\tfrac{1}{2} D_x^2 f + \rho_{42} f_y D_x f + \rho_{43} f_y D_y f\right) + h^4\left[\tfrac{1}{6} D_x^3 f + \tfrac{1}{2}\rho_{42} f_y D_x^2 f + \rho_{32}\rho_{43} f_y^2 D_x f + \tfrac{1}{2}\rho_{43} f_y D_y^2 f + \rho_{42} D_x f D_x f_y + \rho_{43} D_y f D_x f_y\right] + h^5\left[\tfrac{1}{24} D_x^4 f + \tfrac{1}{6}\rho_{42} f_y D_x^3 f + \tfrac{1}{2}\rho_{42} D_x f_y D_x^2 f + \tfrac{1}{2}\rho_{43} D_x f_y D_y^2 f + \tfrac{1}{2}\rho_{43}^2 f_{yy} D_x^2 f + \rho_{42}\rho_{43} f_{yy} D_x f D_y f + \tfrac{1}{2}\rho_{43}^2 f_{yy} D_y^2 f + \tfrac{1}{2}\rho_{42} D_x f D_x^2 f_y + \tfrac{1}{2}\rho_{43} D_y f D_x^2 f_y + \rho_{43}\rho_{42} f_y D_y f D_x f_y\right]\Big|_n + O(h_n^6) . \tag{19}$$

Equations (16) through (19) will enable us to develop all Runge-Kutta methods through order four, and the terms involving h_n^5 will facilitate the discussion of the error terms.

Substituting the expressions from equations (16) through (19) and equation (6) into equation (2) and equating powers of h_n through h_n^4 we have,

$$h_n f = \omega_1 h_n f + \omega_2 h_n f + \omega_3 h_n f + \omega_4 h_n f ,$$

$$\therefore \quad \omega_1 + \omega_2 + \omega_3 + \omega_4 = 1 , \tag{20}$$

$$D h_n^2 f = h_n^2 \omega_2 D_x f + h_n^2 \omega_3 D_y f + h_n^2 \omega_4 D_x f ,$$

$$\therefore \quad \omega_2 D_x f + \omega_3 D_y f + \omega_4 D_x f = \frac{Df}{2!} , \tag{21}$$

$$\tfrac{1}{2}\left[\omega_2 D_x^2 f + \omega_3 D_y^2 f + \omega_4 D_x^2 f\right] + f_y\left[\omega_3\rho_{32} D_x f + \omega_4(\rho_{42} D_x f + \rho_{43} D_y f)\right] = \frac{1}{3!}\left(D^2 f + f_y Df\right) ,$$

$$\tfrac{1}{6}\left[\omega_2 D_x^2 f + \omega_3 D_y^2 f + \omega_4 D_x^2 f\right] + \tfrac{1}{2} f_y\left[\omega_3\rho_{32} D_x^2 f + \omega_4(\rho_{42} D_x^2 f + \rho_{43} D_y^2 f)\right] + \left[\omega_3\rho_{32} D_x f D_y f_y + \omega_4(\rho_{42} D_x f D_x f_y + \rho_{43} D_y f D_x f_y)\right] + \left[\omega_4\rho_{32}\rho_{43} f_y^2 D_x f\right] \tag{22}$$

$$= \frac{1}{4!}\left[D^3 f + f_y D^2 f + 3 Df Df_y + f_y^2 Df\right] . \tag{23}$$

Since the α_i , β_{ij} and ω_i are to be independent of $f(x,y)$, equations (20) through (23) actually represent eight equations. Therefore, the expressions

in brackets in these equations, which are homogeneous in the operators, must equal the corresponding terms on the right. Also if these eight equations are to be independent of $f(x,y)$ then the ratios,

$$\frac{D_j f}{D f} \quad , \; j = 2,3,4 \qquad \text{and} \qquad \frac{D_j f_y}{D f_y} \quad , \; j = 3,4 \quad , \tag{24}$$

must be constant. This will be true if

$$\alpha_i = \sum_{j=1}^{j=i} \beta_{ij} \quad , \quad i = 2,3,4 \; . \tag{25}$$

Therefore,

$$D_i = \alpha_i D \; . \tag{26}$$

Finally, these eight equations become,

$$\begin{aligned}
\omega_1 + \omega_2 + \omega_3 + \omega_4 &= 1 \quad , \\
\omega_2 \alpha_2 + \omega_3 \alpha_3 + \omega_4 \alpha_4 &= \tfrac{1}{2} \quad , \\
\omega_2 \alpha_2^2 + \omega_3 \alpha_3^2 + \omega_4 \alpha_4^2 &= \tfrac{1}{3} \quad , \\
\omega_3 \alpha_2 \beta_{32} + \omega_4 (\alpha_2 \beta_{42} + \alpha_3 \beta_{43}) &= \tfrac{1}{6} \quad , \\
\omega_2 \alpha_2^3 + \omega_3 \alpha_3^3 + \omega_4 \alpha_4^3 &= \tfrac{1}{4} \quad , \\
\omega_3 \alpha_2^2 \beta_{32} + \omega_4 (\alpha_2^2 \beta_{42} + \alpha_3^2 \beta_{43}) &= \tfrac{1}{12} \quad , \\
\omega_3 \alpha_2 \beta_{32} + \omega_4 (\alpha_2 \beta_{42} + \alpha_3 \beta_{43}) &= \tfrac{1}{8} \quad , \\
\omega_4 \alpha_2 \beta_{32} \beta_{43} &= \tfrac{1}{24} \; .
\end{aligned} \tag{27}$$

In the equations above, the first corresponds to equation (20), the second te equation (21), the third and fourth to equation (22) and the last four to equation (23).

Considering then, equations (25) and (27), we have eleven equations in 13 unknowns, which will generally be sufficient to determine the parameters

on assigning values to the two free parameters. We will now consider the Runge-Kutta methods of orders 2, 3, and 4.

For the case when m = 2, the system of equations (27) retains only the equations pertaining to h_n^2 . These, with equation (25) for i = 2 are:

$$\begin{aligned} \omega_1 + \omega_2 &= 1 \ , \\ \alpha_2 \omega_2 &= \tfrac{1}{2} \ , \\ \alpha_2 &= \beta_{21} \ . \end{aligned} \tag{28}$$

Three second order equations of interest correspond to the following values:

$$\alpha_2 = 1/2,\ 2/3,\ 1 \ .$$

Substituting in equation (1) we have, respectively,

$$y_{n+1} - y_n = h_n f\left(x_n + \tfrac{1}{2} h_n \ ,\ y_n + \tfrac{1}{2} h_n f_n\right) \ , \tag{29}$$

$$y_{n+1} - y_n = \tfrac{1}{4} h_n \left[f(x_n, y_n) + 3 f\left(x_n + \tfrac{2}{3} h_n \ ,\ y_n + \tfrac{2}{3} h_n f_n\right)\right] \ , \tag{30}$$

$$y_{n+1} - y_n = \tfrac{1}{2} h_n \left[f(x_n, y_n) + f(x_n + h_n \ ,\ y_n + h_n f_n)\right] \ . \tag{31}$$

It is interesting to note in the above equations that equation (29) is the Newton-Cotes open type formula (or the improved polygon method or the modified Euler method) and if f(x,y) is a function of x only, it reduces to the mid-point rule of numerical integration. Also, equation (31) is like the familiar trapezoidal rule when f(x,y) is a function of x only, otherwise it is the familiar Runge-Kutta second-order method (or the improved Euler method or the Heun method) most commonly seen.

For the case when $n = 3$, we have,

$$\begin{aligned} \omega_1 + \omega_2 + \omega_3 &= 1 \ , \\ \alpha_2\omega_2 + \alpha_3\omega_3 &= \tfrac{1}{2} \ , \\ \alpha_2^2\omega_2 + \alpha_3^2\omega_3 &= \tfrac{1}{3} \ , \\ \alpha_2\beta_{32}\omega_3 &= \tfrac{1}{6} \ , \\ \alpha_2 &= \beta_{21} \ , \\ \alpha_3 &= \beta_{31} + \beta_{32} \ . \end{aligned} \tag{32}$$

This is a two parameter family which can be rewritten as,

$$\begin{aligned} \omega_1 &= 1 + \frac{2 - 3(\alpha_2 + \alpha_3)}{6\alpha_2\alpha_3} \ , \\ \omega_2 &= \frac{3\alpha_3 - 2}{6\alpha_2(\alpha_3 - \alpha_2)} \ , \\ \omega_3 &= \frac{2 - 3\alpha_2}{6\alpha_3(\alpha_3 - \alpha_2)} \ , \\ \beta_{21} &= \alpha_2 \ , \\ \beta_{31} &= \frac{3\alpha_2\alpha_3(1 - \alpha_2) - \alpha_3^2}{\alpha_2(2 - 3\alpha_2)} \ , \\ \beta_{32} &= \frac{\alpha_3(\alpha_3 - \alpha_2)}{\alpha_2(2 - 3\alpha_2)} \ , \end{aligned} \tag{33}$$

where $\alpha_2 \neq \alpha_3$, $\alpha_2, \alpha_3 \neq 0$, $\alpha_2 \neq \frac{2}{3}$.

Two common third order systems are:

$$y_{n+1} - y_n = \frac{2}{9}k_1 + \frac{1}{3}k_2 + \frac{4}{9}k_3 \quad ,$$

where $k_1 = h_n f(x_n, y_n)$,

$$k_2 = h_n f(x_n + \tfrac{1}{2}h_n \, , \, y_n + \tfrac{1}{2}k_1) \, , \tag{34}$$

$$k_3 = h_n f(x_n + \tfrac{3}{4}h_n \, , \, y_n + \tfrac{3}{4}k_2) \, ,$$

and,

$$y_{n+1} - y_n = \frac{1}{6}(k_1 + 4k_2 + k_3) \, ,$$

where $k_1 = h_n f(x_n, y_n)$,

$$k_2 = h_n f(x_n + \tfrac{1}{2}h_n \, , \, y_n + \tfrac{1}{2}k_1) \, , \tag{35}$$

$$k_3 = h_n f(x_n + h_n \, , \, y_n - k_1 + 2k_2) \, .$$

Equation (35) above is similar to the common Simpson's rule when f(x,y) is a function of x only.

For the case when $n = 4$, we again have a two parameter family of equations which can be solved to give the following:

$$\omega_1 = \frac{1}{2} + \frac{1 - 2(\alpha_2 + \alpha_3)}{12\alpha_2\alpha_3} \quad ,$$

$$\omega_2 = \frac{2\alpha_3 - 1}{12\alpha_2(\alpha_3 - \alpha_2)(1 - \alpha_2)} \quad ,$$

$$\omega_3 = \frac{1 - 2\alpha_2}{12\alpha_3(\alpha_3 - \alpha_2)(1 - \alpha_3)} \quad ,$$

$$\omega_4 = \frac{1}{2} + \frac{2(\alpha_2 + \alpha_3) - 3}{12(1 - \alpha_2)(1 - \alpha_3)} \quad , \qquad (36)$$

$$\beta_{32} = \frac{\alpha_3(\alpha_3 - \alpha_2)}{2\alpha_2(1 - 2\alpha_2)} \quad ,$$

$$\beta_{42} = \frac{(1 - \alpha_2)[\alpha_2 + \alpha_3 - 1 - (2\alpha_3 - 1)^2]}{2\alpha_2(\alpha_3 - \alpha_2)[6\alpha_2\alpha_3 - 4(\alpha_2 + \alpha_3) + 3]} \quad ,$$

$$\beta_{43} = \frac{(1 - 2\alpha_2)[1 - \alpha_2](1 - \alpha_3)}{\alpha_3(\alpha_3 - \alpha_2)[6\alpha_2\alpha_3 - 4(\alpha_2 + \alpha_3) + 3]} \quad ,$$

$\alpha_4 = 1$, where $\alpha_1, \alpha_3 \neq 0$, $\alpha_2, \alpha_3 \neq 1$, $\alpha_2 \neq \alpha_3$ and the denominators of the β's do not vanish.

The most common fourth order method is obtained when we let $\alpha_2 = \alpha_3 = 1/2$,

$$y_{n+1} - y_n = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \quad ,$$

where
$$\begin{aligned} k_1 &= h_n f(x_n, y_n) \quad , \\ k_2 &= h_n f(x_n + \tfrac{1}{2}h_n, y_n + \tfrac{1}{2}k_1) \quad , \\ k_3 &= h_n f(x_n + \tfrac{1}{2}h_n, y_n + \tfrac{1}{2}k_2) \quad , \\ k_4 &= h_n f(x_n + h_n, y_n + k_3) \quad . \end{aligned} \qquad (38)$$

The Use of Padé Approximants in the Derivation of the Second-Order Method:

We will discuss the use of Padé approximants in the construction of difference equations which will lead to a familiar numerical solution of differential equations. As before, we are given the following

$$\frac{dy}{dx} = f(x,y), \quad \text{with } y(x_0) = y_0. \tag{39}$$

We define Padé approximants after the following discussion which motivates the definition: Suppose we have a function P(x) for which the Taylor series can be written,

$$P(x) = c_0 + c_1 x + c_2 x^2 + \ldots = \sum_{k=0}^{\infty} c_k x^k .$$

Let

$$D_p(x) = t_0 + t_1 x + t_2 x^2 + \ldots + t_p x^p , \quad p > 0 .$$

Then $D_p(x)$ is a polynomial of degree p in x. We form the product $P(x)D_p(x)$ which, after simplification, yields:

$$P(x)D_p(x) = c_0 t_0 + x(c_1 t_0 + c_0 t_1) + x^2(t_2 c_0 + c_1 t_1 + c_2 t_0) + \ldots . \tag{40}$$

We have (p + 1) parameters t_k, (k = 0,1, . . . ,p) which can be chosen so that the coefficients of x^{q+r} will vanish identically, for r = 1,2, ... ,p. Let $N_q(x)$ be the polynomial of degree less than (q + 1) formed by all terms of degree less than (q + 1) on the right hand side of equation (40). Then,

$$P(x)D_p(x) = N_q(x) + \sum_{k=q+p+1}^{\infty} a_k x^k ,$$

or

$$P(x) = \frac{N_q(x)}{D_p(x)} + \frac{\sum_{k=q+p+1}^{\infty} a_k x^k}{D_p(x)} , \quad D_p(x) \neq 0 .$$

The rational fraction, $f(x) = \frac{N_q(x)}{D_p(x)}$ is the $[p,q]$ Padé approximant to the function $P(x)$, where in the notation $[p,q]$, p denotes the degree of the denominator and q denotes the degree of the numerator.

This Padé approximant enjoys the following basic properties:

1. It is a uniquely determined rational fraction approximation to $P(x)$.

2. If the Padé approximant $[p,q]$ is expanded in a Taylor series, the first (p+q+1) terms will be identical with the first (p+q+1) terms in the Taylor series expansion of the original function.

3. An estimate of the error involved in a given Padé approximant can be calculated from the remainder term in the Taylor series expansion.

Kopal [9,163] regards property 2. as being fundamental and from it is seen that the first (p+q+1) terms of the Taylor series expansion form a Padé approximant, namely the $[o, p+q]$ approximant.

It can be shown that the Padé approximants $[p,p]$ to $\ln(1 - x)$ evaluated near $x = -1$ are more accurate than the $[o, 2p]$ Taylor series.

We denote $\frac{d}{dx}$ by the operator D, and h as the interval step size; then $x_i = x_o + hi$, $y_i = y(x_i)$. Define the forward and backward differences, respectively as follows:

$$\Delta y_i = y_{i+1} - y_i ,$$

$$\nabla y_i = y_i - y_{i-1} .$$

Using this notation, we can derive the following symbolic relationship:

$$-hD = ln(1-\nabla) \cong \frac{2\nabla}{\nabla - 2} ,$$

$$D \cong \frac{1}{h}\left[\frac{2\nabla}{2-\nabla}\right] . \qquad (41)$$

Now, using the operator, as defined, in the original equation (39) and substituting this result in equation (41), we have,

$$\frac{dy}{dx} = Dy = f(x,y) ,$$

$$\frac{1}{h}\left[\frac{2\nabla}{2-\nabla}\right]y \cong f(x,y) . \qquad (42)$$

For a particular x and y, equation (42) becomes

$$2\nabla y_i = h(2-\nabla)f_i ,$$

where $f_i = f(x_i, y_i)$

$$2(y_i - y_{i-1}) = 2hf_i - h\nabla f_i ,$$

$$y_i = y_{i-1} + h(1 - \tfrac{1}{2}\nabla)f_i .$$

In terms of f_1, we have

$$y_i = y_{i-1} + h\left[f_i - \tfrac{1}{2}(f_i - f_{i-1})\right] ,$$

$$y_i = y_{i-1} + \frac{h}{2}\left(f_i + f_{i-1}\right) ,$$

or

$$y_{n+1} = y_n + \frac{h}{2}\left(f_n + f_{n+1}\right) . \qquad (43)$$

Equation (43) is precisely the simplified Runge-Kutta second-order formula with an error $O(h^3)$.

Geometric Interpretation:

Consider again the differential equation,

$$\frac{dy}{dx} = f(x,y) \quad , \quad y(x_0) = y_0 \quad , \tag{44}$$

and use the ordinary fourth order method,

$$y_{n+1} = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) + y_n \quad ,$$

where,

$$\begin{aligned} k_1 &= hf(x_n, y_n) \quad , \\ k_2 &= hf(x_n + \tfrac{1}{2}h \, , \, y_n + \tfrac{1}{2}k_1) \quad , \\ k_3 &= hf(x_n + \tfrac{1}{2}h \, , \, y_n + \tfrac{1}{2}k_2) \quad , \\ k_4 &= hf(x_n + h \, , \, y_n + k_3) \quad . \end{aligned} \tag{45}$$

The geometric significance of these formulae is shown in Figure 1. There it is seen that

$$f(x,y)\Big|_{P_1} = \frac{dy}{dx}\Big|_{P_1}$$

is the slope at P_1 of the curve y = F(x) which satisfies the differential equation in equation (44). Thus from equation(45)we have,

$\frac{k_1}{h} = f(x_n, y_n)$ $\equiv$ the slope of the curve y = f(x) at (x_n, y_n),

$\frac{k_2}{h} = f(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1)$ $\equiv$ the slope of the approximating function at $(x_n + \frac{1}{2}h \, , \, y_n + \frac{1}{2}k_1)$,

$\frac{k_3}{h} = f(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2)$ $\equiv$ the slope of the approximating function at $(x_n + \frac{1}{2}h \, , \, y_n + \frac{1}{2}k_2)$,

$\frac{k_4}{h} = f(x_n + h \, , y_n + k_3)$ $\equiv$ the slope of the approximating function at $(x_n + h \, , \, y_n + k_3)$.

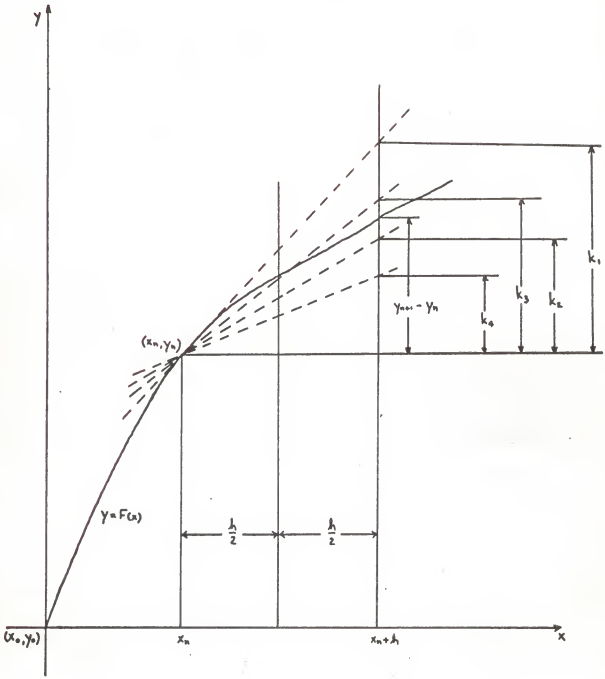

Figure 1

Each of these straight lines when drawn at (x_n, y_n) illustrates the significance of the corresponding k_i; i = 1, 2, 3, 4.

TRUNCATION ERROR

As was stated before, when using the Runge-Kutta method for starting a solution and/or changing the interval, matters such as stability (propagated error), and round-off error, are not significant. We would like then to minimize, as best we can, the truncation (or discretization) error. To do this we first look at the error terms for the three methods mentioned earlier.

Equation (2) is to be exact for powers of h_n, up to h_n^m; hence the truncation error T_m can be expressed as,

$$T_m = \gamma_m h_n^{m+1} + O(h_n^{m+1}) , \tag{46}$$

where both γ_m and T_m are dependent upon the function $f(x,y)$. In order to estimate T_m, we have to consider only γ_m since the bounds that we will give γ_m are so conservative (i.e. the true magnitude of γ_m will generally be much less than the bound we give) that the term $O(h_n^{m+1})$ will be very small compared with $\gamma_m h_n^{m+1}$ (which should be the case if h_n is small). Then the bound on $\gamma_m h_n^{m+1}$ will usually bound the entire error term.

Using equations (8), (16) to (19) and (26) we can calculate the terms γ_2, γ_3 and γ_4. For $m = 2$, from equations (16) to (19) and equation (8) we have,

$$h^3\gamma_2 = \frac{h^3}{3!}\left(D^2 f + f_y Df\right) - \omega_2 \frac{h^3}{2!} D_2^2 f ,$$

$$\gamma_2 = \frac{D^2 f}{6} + \frac{f_y Df}{6} - \frac{\omega_2 D_2^2 f}{2} .$$

and, using equation (26),

$$D_2 = \alpha_2 D .$$

Hence,

$$\gamma_2 = \left(\frac{1}{6} - \frac{\omega_2 \alpha_2^2}{2}\right) D^2 f + \frac{f_y Df}{6} . \tag{47}$$

Similar results can be derived for γ_3 and γ_4. The algebraic manipulation, however, is too tedious to be given here. The results are:

$$\gamma_3 = \left[\frac{1}{4!} - \frac{1}{3!}(\alpha_1^3\omega_2 + \alpha_3^3\omega_3)\right] D^3 f + \left(\frac{1}{4!} - \frac{1}{2!}(\alpha_1^2\beta_{32}\omega_3)\right) f_y D^2 f$$
$$+ \left[\frac{3}{4!} - \alpha_2\alpha_3\beta_{32}\omega_3\right] Df\, Df_y + \frac{1}{4!} f_y^3 Df \,, \tag{48}$$

$$\gamma_4 = \left(\frac{1}{120} - \frac{\omega_2\alpha_2^4 + \omega_3\alpha_3^4 + \omega_4\alpha_4^4}{24}\right) D^4 f$$
$$+ \left(\frac{1}{120} - \frac{\omega_3\alpha_2\alpha_3^2\beta_{32}}{2} + \frac{\omega_4\alpha_4^2(\alpha_2\beta_{42} + \alpha_3\beta_{43})}{2}\right) D^3 f_y Df$$
$$+ \left(\frac{1}{30} - \frac{\omega_2\beta_{32}\alpha_2^2\alpha_3}{2} + \frac{\omega_4\alpha_4(\beta_{42}\alpha_2^2 + \beta_{43}\alpha_3^2)}{2}\right) Df_y D^2 f$$
$$+ \left(\frac{1}{120} - \frac{\omega_4\beta_{43}\beta_{32}\alpha_2^2}{2}\right) f_y^2 D^2 f \,. \tag{49}$$
$$+ \left(\frac{1}{40} - \frac{\omega_3\beta_{32}^2\alpha_2^2 + \omega_4(\beta_{42}\alpha_2 + \beta_{43}\alpha_3)^2}{2}\right) f_{yy} D^2 f$$
$$+ \left(\frac{1}{120} - \frac{\omega_3\beta_{32}\alpha_3^3 + \omega_4(\beta_{42}\alpha_2^3 + \beta_{43}\alpha_3^3)}{6}\right) f_y D^2 f$$
$$+ \left(\frac{7}{120} - \omega_4\beta_{43}\beta_{32}\alpha_2(\alpha_3 + \alpha_4)\right) f_y Df_y Df + \frac{1}{120} f_y^3 Df \,.$$

In order now to bound γ_m, we assume the following bounds on $f(x,y)$ and its derivatives in a neighborhood of (x_n, y_n):

$$|f(x,y)| \le M \,,$$
$$\left|\frac{\partial^{(i+j)} f}{\partial x^i \partial y^j}\right| < \frac{L^{(i+j)}}{M^{(i-1)}} \,, \tag{50}$$

where $i + j \le m$ and M and L are constants, such that $M \ge 1$.

Hence the bounds for γ_2 can be determined by,

$$\gamma_L = \left(\frac{1}{c} - \frac{\omega_1\omega_2^2}{2}\right) D^2 f + \frac{1}{c} f_y Df \quad .$$

Now from equation (6),

$$Df = \frac{\partial f}{\partial x} + f_n \frac{\partial f}{\partial y} \quad ,$$

$$D^2 f = \frac{\partial^2 f}{\partial x^2} + 2 f_n \frac{\partial^2 f}{\partial x \partial y} + f_n^2 \frac{\partial^2 f}{\partial y^2} \quad ,$$

and, using equation (50),

$$\left|\frac{\partial f}{\partial x}\right| < ML \quad , \quad |f_n| < M \quad ,$$

$$\left|\frac{\partial f}{\partial y}\right| < L \quad , \quad |Df| < 2ML$$

Similarly,

$$\left|\frac{\partial^2 f}{\partial x^2}\right| < ML^2 \quad , \quad \left|\frac{\partial^2 f}{\partial x \partial y}\right| < L^2 \quad ,$$

$$\left|\frac{\partial^2 f}{\partial y^2}\right| < \frac{L^2}{M} \quad .$$

Hence,

$$|D^2 f| < ML^2 + 2ML^2 + ML^2 = 4ML^2 \quad ,$$

and

$$|\gamma_1| < \left(4\left|\frac{1}{c} - \frac{\omega_1^2\omega_2}{2}\right| + \frac{1}{3}\right) ML^2 \quad . \tag{51}$$

In a similar way, we can determine bounds for γ_3 and γ_4. The derivation, however, will be omitted.

$$|\gamma_3| < \left[8\left|\tfrac{1}{24} - \tfrac{1}{6}(\alpha_1^2\omega_1 + \alpha_3^2\omega_3)\right| + 4\left|\tfrac{1}{24} - \tfrac{1}{2}\alpha_1^2\beta_{31}\omega_3\right| + 4\left|\tfrac{1}{8} - \alpha_1\alpha_3\beta_{31}\omega_3\right| + \tfrac{1}{12}\right] ML^3 , \tag{52}$$

$$|\gamma_4| < \left[16|b_1| + 4|b_2| + |b_2+3b_3| + |2b_2+3b_3| + |b_2+b_3| + |b_3| + 8|b_4| + |2b_5+b_7| + |b_5+b_6+b_7| + |b_6| + |2b_6+b_7| + |b_7| + 2|b_8| \right] ML^4 ,$$

where

$$b_1 = \frac{1}{120} - \frac{1}{24}(\alpha_1^4\omega_1 + \alpha_3^4\omega_3 + \omega_4) ,$$

$$b_2 = \frac{1}{120} - \frac{1}{2}\left[\alpha_1\alpha_3^3\beta_{31}\omega_3 + (\alpha_1\beta_{41} + \alpha_3\beta_{43})\omega_4\right] ,$$

$$b_3 = \frac{1}{120} - \frac{1}{6}\left[\alpha_1^3\beta_{31}\omega_3 + (\alpha_1^3\beta_{41} + \alpha_3^3\beta_{43})\omega_4\right] ,$$

$$b_4 = \frac{1}{30} - \frac{1}{2}\left[\alpha_1^2\alpha_3\beta_{31}\omega_3 + (\alpha_1^2\beta_{41} + \alpha_3^2\beta_{43})\omega_4\right] , \tag{53}$$

$$b_5 = \frac{1}{120} - \frac{1}{2}(\alpha_1^2\beta_{31}\beta_{43}\omega_4) ,$$

$$b_6 = \frac{1}{40} - \frac{1}{2}\left[\alpha_1^2\beta_{31}\omega_3 + (\alpha_1\beta_{41} + \alpha_3\beta_{43})^2\omega_4\right] ,$$

$$b_7 = \frac{7}{120} - \alpha_1(1-\alpha_3)\beta_{31}\beta_{43}\omega_4 ,$$

$$b_8 = \frac{1}{120} .$$

Referring to equations (27), (28) and (33), for which the parameters are underdetermined, in all three cases the underdetermined parameters will be assigned values so as to minimize the bounds we have just set on the γ_m.

Now let us consider the second order system (i.e. m = 2) and the one parameter family in equation (28) resulting from equating like powers of h_n up through h_n^2 :

$$\omega_1 + \omega_2 = 1 \quad ,$$

$$\alpha_2 \omega_2 = \frac{1}{2} \quad ,$$

$$\alpha_2 = \beta_{21} \quad .$$

We have then,

$$\begin{aligned} \omega_1 &= 1 - \frac{1}{2\alpha_2} \quad , \\ \omega_2 &= \frac{1}{2\alpha_2} \quad , \\ \beta_{21} &= \alpha_2 \quad . \end{aligned} \tag{54}$$

Since,

$$|\gamma_2| < \left(4\left|\frac{1}{6} - \frac{\alpha_2^2\omega_2}{2}\right| + \frac{1}{3}\right)ML^2 \quad ,$$

we have, after substituting $\omega_2 = \frac{1}{2\alpha_2}$,

$$|\gamma_2| < \left(4\left|\frac{1}{6} - \frac{\alpha_2}{4}\right| + \frac{1}{3}\right)ML^2 \quad . \tag{55}$$

It is clear from relation (55) that if $\alpha_2 = 2/3$, γ_2 is minimized and, moreover, γ_2 is strictly less than $\frac{ML^2}{3}$. It is seen that for $\alpha_2 = 2/3$ we have the following,

$$y_{n+1} - y_n = \frac{1}{4}h_n f(x_n, y_n) + \frac{3}{4}h_n f\left(x_n + \frac{2}{3}h_n,\ y_n + \frac{2}{3}h_n f_n\right),$$

which was one of the second order systems given in equation (30).

For $n = 3$, consider the two parameter family given in equation (33). After solving for w_1 and w_2 in terms of w_3, let $\alpha_3 = 0$, $\alpha_2 = 2/3$ and obtain the following one parameter family of equations,

$$\begin{aligned} \alpha_{12} &= \tfrac{2}{3} \ , \\ \beta_{32} &= \frac{\omega_3}{4} \ , \\ \omega_1 &= \tfrac{1}{4} - \omega_3 \ , \\ \omega_2 &= \tfrac{3}{4} \ . \end{aligned} \tag{56}$$

Similarly we can let $\alpha_2 = \alpha_3 = 2/3$ with the resulting one parameter family of equations being given by

$$\begin{aligned} \omega_1 &= \tfrac{1}{4} \ , \\ \omega_2 &= \tfrac{3}{4} - \omega_3 \ , \\ \beta_{32} &= \frac{1}{4\omega_3} \ . \end{aligned} \tag{57}$$

We again follow the method used for $n = 2$ and see that the coefficient of ML^3 in relation (52) will be a minimum if $\alpha_2 = 1/2$, $\alpha_3 = 3/4$, which gives,

$$|\gamma_3| < \tfrac{1}{9} ML^3 \ . \tag{58}$$

Using the values given in the system of equations (56) we have the following bound for γ_3,

$$|\gamma_3| < \tfrac{2}{3} ML^3 \ , \tag{59}$$

or, by using those values in the system of equations (57) we have as a bound on γ_3,

$$|\gamma_3| < \tfrac{1}{4} ML^3 \ . \tag{60}$$

Thus relation (58) obviously will give the least bound on the truncation error. Using the values of the free parameters (i.e. $\alpha_2 = 1/2$, $\alpha_3 = 3/4$) which were used to arrive at relation (58) we have the following third order equation which will yield values of y which are best in the 'least' truncation error sense.

$$y_{m} - y_n = \tfrac{2}{9} k_1 + \tfrac{1}{3} k_2 + \tfrac{4}{9} k_3 \ ,$$

$$\text{where} \quad k_1 = h_n f(x_n, y_n) \ , \tag{61}$$

$$k_2 = h_n f(x_n + \tfrac{1}{2} h_n \ , \ y_n + \tfrac{1}{2} k_1) \ ,$$

$$k_3 = h_n f(x_n + \tfrac{3}{4} h_n \ , \ y_n + \tfrac{3}{4} k_2) \ .$$

For $n = 4$ we have the two parameter family given in equation (37) whose possible solutions are listed below for various (common) choices of the free parameters.

$$\alpha_2 = \alpha_3 = \alpha_4 = 1 \quad , \quad \omega_1 = \omega_4 = \tfrac{1}{6} \quad ,$$

$$\omega_2 = \tfrac{2}{3} - \omega_3 \quad , \quad \beta_{32} = \frac{1}{6\omega_3} \quad , \tag{62}$$

$$\beta_{42} = 1 - 3\omega_3 \quad , \quad \beta_{43} = 3\omega_3 \ .$$

$$\alpha_2 = \alpha_4 = 1 \quad , \quad \alpha_3 = \tfrac{1}{2} \quad ,$$

$$\omega_1 = \tfrac{1}{6} \quad , \quad \omega_2 = \tfrac{1}{6} - \omega_4 \quad , \tag{63}$$

$$\omega_3 = \tfrac{2}{3} \quad , \quad \beta_{32} = \tfrac{1}{2} \quad ,$$

$$\beta_{42} = -\frac{1}{12\omega_4} \quad , \quad \beta_{43} = \frac{1}{3\omega_4} \ .$$

$$\alpha_2 = \frac{1}{2} \quad , \quad \alpha_3 = 0 \quad , \quad \alpha_4 = 1 \quad ,$$

$$\omega_1 = \frac{1}{6} - \omega_3 \quad , \quad \omega_2 = \frac{2}{3} \quad , \quad \omega_4 = \frac{1}{6} \quad , \tag{64}$$

$$\beta_{32} = \frac{1}{12\omega_3} \quad , \quad \beta_{42} = \frac{3}{2} \quad , \quad \beta_{43} = 6\omega_3 \quad .$$

Using the bound on γ_4 as given in relation (53) with some lengthy computation, it can be shown that the coefficient of ML^4 will be minimized when $\alpha_2 = 0.4$, $\alpha_3 = 7/8 - 3/16\sqrt{15}$.

We have then,

$$|\gamma_4| < 5.46 \times 10^{-2}\, ML^4 \quad . \tag{65}$$

We can compare this bound with others for various values of the coefficients.

In the system of equations (62) with $w_3 = 5/8$, we have,

$$|\gamma_4| < 7.22 \times 10^{-2}\, ML^4 \quad . \tag{66}$$

In the system of equations (63) with $w_4 = 10/51$, we have,

$$|\gamma_4| < 19.72 \times 10^{-2}\, ML^4 \quad . \tag{67}$$

In the system of equations (64) with $w_3 = -5/78$, we have,

$$|\gamma_4| < 17.64 \times 10^{-2}\, ML^4 \quad .$$

A much earlier bound was found by Lotkin [9,130] for the system given in the system of equations (62) with $w_3 = 1/3$ and was

$$|\gamma_4| < 10.14 \times 10^{-2}\, ML^4 \quad .$$

For the sake of illustration, we will list some of the more common error bounds previously derived for the fourth order method. The bound for the

classical fourth order method with $\alpha_2 = 0.4$, $\alpha_3 = 0.6$ is given by,

$$|\gamma_4| < 7.70 \times 10^{-2} ML^4 , \qquad [13,435] .$$

For the method by Kutta with $\alpha_2 = 1/3$, $\alpha_3 = 2/3$ we have,

$$|\gamma_4| < 9.91 \times 10^{-2} ML^4 , \qquad [13,436] .$$

And for the method by Gill, with $\alpha_2 = 1/2$, $\alpha_3 = 1/2$, $w_3 = 1 + \frac{\sqrt{2}}{2}$ we have,

$$|\gamma_4| < 8.83 \times 10^{-2} ML^4 , \qquad [4,106] .$$

In all cases it is obvious that the minimum error bound is given by relation (65). This was first derived by Ralston [10,435] . Using the values of the free parameters which were used to arrive at relation (65) we have the following minimum fourth order Runge-Kutta method:

$$y_{n+1} - y_n = .17476028\ k_1 - .55148066\ k_2 + 1.20553560\ k_3 + .17118478\ k_4,$$

where

$$\begin{aligned}
k_1 &= h_n f(x_n, y_n) , \\
k_2 &= h_n f(x_n + .4 h_n , y_n + .4 k_1) , \\
k_3 &= h_n f(x_n + .45573725\, h_n , \\
&\qquad y_n + .29697761\, k_1 + .15875964\, k_2) , \\
k_4 &= h_n f(x_n + h_n , y_n + .21810040 k_1 \\
&\qquad - 3.05096516\, k_2 + 3.83286476\, k_3) .
\end{aligned} \tag{68}$$

Here y_{n+1} is an approximation to the solution of the given differential equation.

THE RUNGE-KUTTA METHOD AS APPLIED TO A SYSTEM OF FIRST ORDER DIFFERENTIAL EQUATIONS

A system of ordinary differential equations of the first order is a system of equations of the form

$$\begin{aligned} y^{1\prime} &= f^1(x, y^1, y^2, \cdots, y^s) , \\ y^{2\prime} &= f^2(x, y^1, y^2, \cdots, y^s) , \\ &\cdots \\ y^{s\prime} &= f^s(x, y^1, y^2, \cdots, y^s) , \end{aligned} \tag{69}$$

where $f^1, f^2, \ldots, f^s$ are given functions of (s+1) arguments, and the second superscript on the y's refers to the derivative with respect to x.

A set of functions $y^1(x), y^2(x), \ldots, y^s(x)$ which are defined and differentiable in an interval $[a,b]$ and satisfy identically in x the relation,

$$y^{i\prime} = f^i(x, y^1(x), y^2(x), \ldots, y^s(x)) , \qquad i = 1, 2, \ldots, s , \tag{70}$$

is called a solution to the system.

The problem which arises frequently in practice and which we will discuss here is to find a solution of the system of equations (69) which satisfies the initial conditions

$$y^i(x_0) = \eta_i , \qquad i = 1, 2, \ldots, s , \tag{71}$$

where the η_i are preassigned constants.

It should be noted in passing that other conditions, more complicated than those of equation (71), also arise in practice, but will not be discussed here.

Systems of ordinary differential equations arise in several ways; two general situations are given in the following pages:

(i) Theoretically, every ordinary differential equation of order higher than the first can be reduced to a system of first order equations. Consider the differential equation of order m given by

$$y^{(m)} = f(x, y, y', y'', \dots, y^{(m-1)}) , \tag{72}$$

where f is a given function of (m+1) arguments. The reduction of this equation to a system of equations (69) is accomplished by setting

$$y^1 = y \quad , \quad y^2 = y' \quad , \quad y^3 = y'' \quad , \dots , \quad y^m = y^{(m-1)} .$$

Now if the functions $y^1, y^2, \dots, y^m$ satisfy the system of equations,

$$\begin{aligned} y^{1\prime} &= y^2 , \\ y^{2\prime} &= y^3 , \\ &\dots \\ y^{m\prime} &= f(x, y^1, y^2, \dots, y^m) , \end{aligned} \tag{73}$$

then the function $y(x) = y^1(x)$ will satisfy equation (72). Thus the system of equations (73) is a special case of the system of equations (69).

Some authorities (Milne [12,82] and Gill [4,96]) recommend this reduction of higher order equations to a system of equations of the first order also for numerical purposes; others (such as Collatz [2,117])take the opposite position, arguing that reduction to a first order system increases both the error and the necessary number of operations. Methods for the direct integration of equations of higher order can be found in most advanced texts on the subject. However, the theoretical and experimental results presented dealing with one equation will be at least nearly best for most systems of

equations. Evidence indicates that it is possible to control the truncation error and it can be shown that the round-off error is frequently substantially decreased when an equation of higher order is first reduced to a first order system and then solved by an equivalent method for such a system. See Henrici [6,123] .

(ii) Systems of ordinary differential equations also arise in a natural way from many physical problems. Classical examples are electric circuits with more than one loop and mechanical problems with several degrees of freedom. More specific examples are the equations of motion of a gyroscope, the fundamental equations of exterior ballistics and the equations governing the flights of rockets and missiles.

Vector Notation:

It will be convenient for us at this time to simplify the subsequent analysis both conceptually and formally by considering the quantities y^i, $i = 1,2,\dots,s$ as components of the vector

$$\bar{y} = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^s \end{bmatrix} .$$

Consequently we write $f^i(x,y^1,y^2,\dots,y^s) = f^i(x,\bar{y})$.

Combine the s functions of $f^i(x,\bar{y})$ into another vector:

$$\bar{f}(x,\bar{y}) = \begin{bmatrix} f^1(x,\bar{y}) \\ f^2(x,\bar{y}) \\ \vdots \\ f^s(x,\bar{y}) \end{bmatrix} .$$

Write equation (69) in the more compact form as

$$\bar{y}' = \bar{f}(x,\bar{y}) \ . \tag{74}$$

If we define the vector $\bar{\eta}$ by

$$\bar{\eta} = \begin{bmatrix} \eta^1 \\ \eta^2 \\ \vdots \\ \eta^s \end{bmatrix} \ ,$$

then the initial condition in (3) is given by,

$$\bar{y}(x_0) = \bar{\eta} \ . \tag{75}$$

In addition, as in the case of one equation, we assume the following, that the vector-valued function $\bar{f}(x,\bar{y})$ of the scalar variable x and the vector $\bar{y} = (y^1, y^2, \ldots, y^s)$ satisfy the following two hypotheses:

(i) $\bar{f}(x,\bar{y})$ is defined and continuous in the region

$$a \le x \le b, \ -\infty < y^i < \infty \quad , \ i = 1,2, \ldots, s$$

(ii) there exists a constant L such that for some $x \in [a,b]$ and any two vectors $\bar{y}$ and $\bar{y}^*$,

$$\left\| \bar{f}(x,\bar{y}) - \bar{f}(x,\bar{y}^*) \right\| \le L \left\| \bar{y} - \bar{y}^* \right\| \ ,$$

where $\|\bar{v}\|$ indicates the norm of the vector $\bar{v}$.

We omit the lengthy proof of the following theorem [4,113].

Theorem: Let the function $\bar{f}(x,\bar{y})$ satisfy the conditions (i) and (ii) and let $\bar{\eta}$ be a given vector; then there exists exactly one function with the following three properties:

a. $\bar{y}(x)$ is continuous and continuously differentiable for $x \in [a,b]$,

b. $\bar{y}'(x) = \bar{f}(x, \bar{y}(x))$, $x \in [a,b]$,

c. $\bar{y}(x_0) = \bar{\eta}$.

(i.e. the initial value problem given by

$$\bar{y}' = \bar{f}(x,\bar{y}) \quad , \quad \bar{y}(x_0) = \bar{\eta} \quad , \tag{76}$$

has a unique solution.).

Let us now consider the Runge-Kutta method for our system of equations. We let $x \in [a,b]$, $\bar{y}$ be an arbitrary vector and $\bar{z}(t)$ denote the solution of the system of equations given by,

$$\bar{z}' = \bar{f}(t,\bar{z}) \quad , \quad \bar{z}(x) = \bar{y} \ ,$$

and set

$$\bar{\Delta}(x,\bar{y};h) = \begin{cases} \dfrac{\bar{z}(x+h) - \bar{z}(x)}{h} & , \ h \neq 0 \ . \\ \bar{f}(x,\bar{y}) & , \ h = 0 \ . \end{cases}$$

We call $\bar{\Delta}$ the exact relative increment of the solution of $\bar{z}' = \bar{f}(t,\bar{z})$. A one-step method (Runge-Kutta is the most sophisticated of these methods) for the solution of the initial value problem given in equation (8) is defined by,

$$\bar{y}^0 = \bar{\eta} \ .$$

$$\bar{y}^{n+1} = \bar{y}^n + h\,\bar{\Phi}(x, \bar{y}^n; h) \ , \quad n = 0,1,\ldots \ .$$

Here $\bar{\Phi}$ is called the increment function and is chosen so as to approximate $\bar{\Delta}$ as closely as possible.

In order to eliminate the special role played by the independent variable x, we augment the system of equations (69) by the differential equation

$$y^{0\prime}(x) = 1 \tag{77}$$

to be satisfied by a new function $y^0(x)$, subject to the initial condition

$$y^0(x_0) = x_0 \tag{78}$$

Equations (77) and (78) clearly imply that $y^0(x) = x$. Hence we can replace the system of equations (69) by the equivalent system for (s+1) functions,

$$y^{i\prime} = f^i(y^0, y^1, \dots, y^s) \quad , \quad i = 0, 1, \dots, s \quad , \tag{79}$$

where $f^0(y^0, y^1, \dots, y^s) = 1$

The system now given in equation (79) has the advantage that the variables entering into the functions f^i may all be considered dependent. For simplicity we can continue to denote the dependent variables by $y^0, y^1, \dots, y^s$ whether or not one of them is x. Thus we may write the initial value problem as

$$\bar{y}' = \bar{f}(\bar{y}) \quad , \quad \bar{y}(x_0) = \bar{\eta} \quad , \tag{80}$$

where $\bar{y}$, $\bar{f}$ and $\bar{\eta}$ are all vectors with s components.

If $\bar{f}(\bar{y})$ does not depend explicitly on x, neither does the function $\bar{\Delta}$, nor does the increment $\overline{\Phi}$. Hence

$$\bar{\Delta} = \bar{\Delta}(\bar{y}; h) \quad , \quad \text{and} \quad \overline{\Phi} = \overline{\Phi}(\bar{y}, h) \quad .$$

Now if the function $\bar{y}(x)$ is a solution of equation (80) and assuming the components of $\bar{f}$ are sufficiently differentiable then the higher deriva-

tives of $\bar{y}(x)$ can be expressed in terms of the function $\bar{f}$ and its derivatives. For example

$$\bar{y}'(x) = \frac{d}{dx}\bar{f}(\bar{y}(x)) = \sum_{j=1}^{s}\frac{\partial \bar{f}}{\partial y^j}\frac{dy^j}{dx},$$

$$\bar{y}''(x) = \sum_{j=1}^{s}\frac{\partial \bar{f}}{\partial y^j}f^j .$$

Generally, for k = 1,2, ..., assuming differentiability,

$$\bar{y}^{(k+1)}(x) = \frac{d^k}{dx^k}\bar{f}(\bar{y}(x)) = \bar{f}^{(k)}(\bar{y}(x)) .$$

Since the functions $\bar{f}^{(k)}(\bar{y})$ exist, we have

$$\bar{\Delta}(\bar{y};h) = \bar{f}(\bar{y}) + \frac{h}{2!}\bar{f}'(\bar{y}) + \frac{h^2}{3!}\bar{f}''(\bar{y}) + \cdots . \qquad (81)$$

As in the case of a single differential equation, a method for approximate integration can be based on a truncated Taylor series. The increment function for the Taylor series expansion of order p is

$$\bar{\Phi}(\bar{y};h) = \bar{f}(\bar{y}) + \frac{h}{2!}\bar{f}'(\bar{y}) + \cdots + \frac{h^{p-1}}{p!}\bar{f}^{(p-1)}(\bar{y}) . \qquad (82)$$

This method, being of no great practical interest since the evaluation of many derivatives is involved, is however of theoretical importance since the Runge-Kutta methods are based on the idea of approximating equation (82) by expressions which do not involve any functions other than $\bar{f}(\bar{y})$.

Let us look now in more detail at the derivatives and their structure for $\bar{f}^{(k)}(\bar{y})$. To simplify the notation, write for i,j,k = 1,2, ... ,s

$$\frac{\partial f^i}{\partial y^j} = f^i_j , \quad \frac{\partial^2 f^i}{\partial y^j \partial y^k} = f^i_{jk} , \quad \cdots , \qquad (83)$$

whereby $\bar{f}_i$, $\bar{f}_{jk}$, ... , is meant the vectors with the components f^i_j , f^i_{jk} , ... , $(i = 1,2, \ldots ,s)$. We also will find it appropriate at this time to adopt the summation convention which is used quite often in vector and tensor analysis (i.e. if an index occurs both as a subscript and a superscript, the terms should be summed with respect to this index from 1 to s). Thus we have,

$$f^i_j f^j = \sum_{j=1}^{s} \frac{\partial f^i}{\partial y^j} f^j \quad . \tag{84}$$

It is clear that sums of products of the form of equation (84) can be differentiated like ordinary products. Hence

$$\frac{d}{dx}\left(f^i_j f^j\right) = \left(\frac{d}{dx} f^i_j\right) f^j + f^i_j \left(\frac{d}{dx} f^j\right) = f^i_{jk} f^j f^k + f^i_j f^j_k f^k \quad .$$

Let

$$\begin{aligned} A^i &= f^i \quad , & B^i &= f^i_{km} f^k f^m \quad , \\ B^i &= f^i_j f^j \quad , & F^i &= f^i_{km} f^k_n f^n f^m \quad , \\ C^i &= f^i_{jk} f^j f^k \quad , & G^i &= f^i_j f^j_{kn} f^m f^n \quad , \\ D^i &= f^i_j f^j_k f^k \quad , & H^i &= f^i_j f^j_k f^k_m f^m \quad , \end{aligned} \tag{85}$$

where i,j,k,m = 1,2, ... ,s and the argument of every function is understood to by $\bar{y}$. Also denote by $\bar{A}$, $\bar{B}$, ... the vectors with the components A^i, B^i, ... $(i = 1,2, \ldots ,s)$ respectively.

With these conventions, we can write formally the first few derivatives in the following compact manner:

$$\begin{aligned} \bar{f}'(\bar{y}) &= \bar{A} \quad , \\ \bar{f}''(\bar{y}) &= \bar{B} \quad , \\ \bar{f}'''(\bar{y}) &= \bar{C} + \bar{D} \quad , \\ \bar{f}^{iv}(\bar{y}) &= \bar{E} + 3\bar{F} + \bar{G} + \bar{H} \quad . \end{aligned} \tag{86}$$

Since we will have to expand expressions of the form $\bar{f}(\bar{y} + h\bar{a})$ in powers of h, where $\bar{y}$ and $\bar{a}$ are fixed vectors, we write Taylor's series for functions of several variables [16,227] .

$$f^i(\bar{y} + h\bar{a}) = f^i + h f^i_j a^j + \frac{h^2}{2!} f^i_{jk} a^j a^k + \frac{h^3}{3!} f^i_{jkm} a^j a^k a^m + O(h^4) . \quad (87)$$

We want to combine values of the function which $\bar{f}$ takes at different points, in such a way that the resulting function $\bar{\Phi}(\bar{y}; h)$ agrees as closely as possible with

$$\bar{\Delta}(\bar{y}; h) = \bar{A} + \frac{h}{2!}\bar{B} + \frac{h^2}{3!}(\bar{C} + \bar{D}) + \frac{h^3}{4!}(\bar{E} + 3\bar{F} + \bar{G} + \bar{H}) + \cdots (88)$$

(Note that equation (88) is the same as equation (81).)

For the second order Runge-Kutta method, we put

$$\bar{\Phi}(\bar{y}; h) = a_1 \bar{f}(\bar{y}) + a_2 \bar{f}(\bar{y} + ph\bar{f}(\bar{y})) ,$$

where a_1, a_2 and p must be determined. Using equation (87), we have,

$$\begin{aligned} \bar{f}(\bar{y} + ph\bar{f}(\bar{y})) &= \bar{f}(\bar{y}) + hp \bar{f}_j(\bar{y}) f^j(\bar{y}) \\ &\quad + \frac{(hp)^2}{2} \bar{f}_{jk}(\bar{y}) f^j(\bar{y}) f^k(\bar{y}) + O(h^3) , \\ &= \bar{A} + hp\bar{B} + \frac{(hp)^2}{2}\bar{C} + O(h^3) . \end{aligned}$$

Hence

$$\bar{\Phi}(\bar{y}; h) = (a_1 + a_2)\bar{A} + a_2 h\bar{B} + \tfrac{1}{2} a_2 (hp)^2 \bar{C} + O(h^3) .$$

Now equating the constant and linear term in h (it is impossible to obtain agreement in h^2 since $\bar{D}$ is not present in the above equation) with the corresponding terms in equation (88), we have,

$$a_1 + a_2 = 1,$$
$$a_2 p = 1/2,$$

giving the general solution of the above system of equations as

$$a_2 = \alpha \quad ,$$
$$a_1 = 1 - \alpha \quad ,$$
$$p = 1/2\alpha \quad ,$$

where $\alpha \neq 0$.

Thus the increment function is given by,

$$\overline{\Phi}(\bar{y};h) = (1-\alpha)\bar{f}(\bar{y}) + \alpha \bar{f}\left(\bar{y} + \frac{h}{2\alpha}\bar{f}(\bar{y})\right) , \quad \alpha \neq 0 . \qquad (89)$$

It deviates from equation (88) by $O(h^2)$ and two evaluations of $\bar{f}(\bar{y})$ are required to compute $\overline{\Phi}$.

In a similar manner, as in the preceeding case and analogous to the single equation, we will arrive at the classical Runge-Kutta fourth order method.

The increment function is given by

$$\overline{\Phi}(\bar{y};h) = a_1\bar{k}_1 + a_2\bar{k}_2 + a_3\bar{k}_3 + a_4\bar{k}_4 \quad ,$$

and one set of appropriate choices of a_1, a_2, a_3 and a_4 is found to be

$$a_1 = a_4 = 1/6 \quad , \quad a_2 = a_3 = 1/3 \quad .$$

Thus,

$$\overline{\Phi}(\bar{y};h) = \frac{1}{6}\left(\bar{k}_1 + 2\bar{k}_2 + 2\bar{k}_3 + \bar{k}_4\right) , \qquad (90)$$

where $\bar{k}_1 = \bar{f}(\bar{y})$,

$\bar{k}_2 = \bar{f}(\bar{y} + \frac{1}{2}h\bar{k}_1)$,

$\bar{k}_3 = \bar{f}(\bar{y} + \frac{1}{2}h\bar{k}_2)$,

$\bar{k}_4 = \bar{f}(\bar{y} + h\bar{k}_3)$.

It is clear that equation (90) is a special case of the classical Runge-Kutta fourth order equation derived earlier for the case of the single differential equation.

NUMERICAL EXAMPLES

For the sake of illustration, we list five numerical examples (Table 1). In papers of this kind, numerical examples are desirable, especially those which illustrate how well the derived method compares with others. It is sometimes difficult to choose meaningful examples to illustrate Runge-Kutta methods, since the complicated nature of the error term makes it difficult to choose a function $f(x,y)$ which really serves as a test while at the same time yields a problem which can be solved analytically.

A FORTRAN program was written to compute both the classical and the minimum fourth-order solution for a given differential equation and to calculate the error in each method for the desired number of iterations (see Appendix). The results for these five differential equations are given in Table 2.

The first three of these examples show that the minimum method compares favorably with the classical method while in the fourth example, there is really no comparison and finally the fifth example is not nearly as favorable in comparison. We note that it is only a matter of a little ingenuity to find other examples to make the minimum method appear more or less favorable in comparison with the classical method or other methods.

In conclusion, we re-state the main point of this report. If Runge-Kutta methods are to be used to start the solution and/or to change the interval size, one is interested only in being able to limit the truncation error to as small a quantity as is possible. Hence we choose the method which puts the smallest bound on the error term in this sense. Therefore, when a fourth-order method is desired, equation (68) should be used, when a third-order method is employed equation (61) should be used, and equation (30) should be used when a second-order method is under consideration. In all

cases, if the method is best for a single equation, it is at least nearly best for a system of equations.

TABLE 1

Example	Differential Equation	Initial Condition	Solution
I	$\frac{dy}{dx} = \frac{x(x+1) + 2y}{x}$	$y(1) = 1$	$y = x^2 \log x + 2x^2 - x$
II	$\frac{dy}{dx} = -x - 2y$	$y(0) = -1$	$y = \frac{1 - 5e^{-2x} - 2x}{4}$
III	$\frac{dy}{dx} = \frac{1}{1 + \tan^2 y}$	$y(0) = 0$	$y = \arctan x$
IV	$\frac{dy}{dx} = 1 - y^2$	$y(0) = 0$	$y = \tanh x$
V	$\frac{dy}{dx} = \frac{e^x(y^3+xy^3+1)}{3y^2(xe^x - 6)}$	$y(0) = 1$	$y = \left(\frac{e^x + 5}{6 - xe^x}\right)^{1/3}$

In each of the above problems, we calculate the value of y when x = 4. We calculate the "exact" value from the solution given in Table 1 and then calculate the "approximate" solution using numerical methods. Thus in example I we calculate y for x = 1(.1)4 and also for x = 1(.2)4*. In example II we calculate y for x = 1(.1)4, etc. The results for x = 4 are compared with the "exact" values and these differences are noted in Table 2.

---

* Here the notation x = 1(.1)4 means that we let x take on successive values throughout the interval from x = 1 to x = 4 in steps of length .1 (i.e. $x_0 = 1$, $x_1 = 1.1$, $x_2 = 1.2$, ..., $x_{n-1} = 3.9$, $x_n = 4.0$).

TABLE 2

Example	Step Size	Number of Iterations	Error for Classical Fourth-Order Method	Error for Minimum Fourth-Order Method
I	.1	40	$-.338 \times 10^{-3}$	$-.265 \times 10^{-3}$
	.2	20	$-.433 \times 10^{-2}$	$-.328 \times 10^{-2}$
II	.1	40	$.200 \times 10^{-6}$	$.000 \times 10^{-99}$
	.2	20	$-.900 \times 10^{-6}$	$-.900 \times 10^{-6}$
III	.1	40	$-.800 \times 10^{-6}$	$-.700 \times 10^{-6}$
	.2	20	$-.180 \times 10^{-5}$	$-.150 \times 10^{-5}$
IV	.1	40	$.000 \times 10^{-99}$	$.000 \times 10^{-99}$
	.2	20	$.000 \times 10^{-99}$	$.000 \times 10^{-99}$
V	.1	10	$.140 \times 10^{-5}$	$.260 \times 10^{-5}$

ACKNOWLEDGEMENT

The author wishes to express his appreciation to Dr. S. T. Parker for his helpful suggestions and patient assistance during the preparation of this report and is grateful to the Kansas State University Computing Center for the use of the IBM 1410 in calculating the examples.

BIBLIOGRAPHY

1. Butcher, J. C. "On Runge-Kutta Processes of High Order", Journal of the Australian Mathematical Society, v. 4, 1964, pp. 179-194.

2. Collatz, L. The Numerical Treatment of Differential Equations, 3rd edition, Springer, Berlin, 1960.

3. Fox, Augustus H. Fundamentals of Numerical Analysis, The Ronald Press Company, New York, 1963.

4. Gill, S. "A Process for the Step-by-Step integration of Differential Equations in an Automatic Digital Computing Machine", Proceedings of the Cambridge Philosophical Socley, v. 47, pp. 96-108.

5. Henrici, Peter. Elements of Numerical Analysis, John Wiley and Sons, Inc., New York, 1964.

6. _____. Discrete Variable Methods in Ordinary Differential Equations, John Wiley and Sons, Inc., 1962.

7. Jennings, Walter. First Course in Numerical Methods, The Macmillan Company, New York, 1964.

8. King, Richard. "Runge-Kutta Methods with Constrained Minimum Error Bounds", Mathematics of Computation, v. 20, 1966, p. 386.

9. Kopal, Zdenek. "Operational Methods in Numerical Analysis Based on Rational Approximations", On Numerical Approximation, ed. by Rudolph E. Langer, Madison, Wisc., The Univ. of Wisconsin Press, 1959.

10. Levy, H. Numerical Solutions of Differential Equations, Dover Publications, Inc., New York, 1st American edition, 1950.

11. Lotkin, M. "On the Accuracy of Runge-Kutta Methods", MTAC, v. 5, 1951, pp. 128-132.

12. Milne, W. E. Numerical Solution of Differential Equations, John Wiley and Sons, Inc., New York, 1953.

13. Ralston, Anthony. "Runge-Kutta Methods With Minimum Error Bounds", Mathematics of Computation, v. 16, 1962, pp. 431-437.

14. _____. A First Course in Numerical Analysis, McGraw-Hill Book Company, New York, 1965.

15. Ralston, A. and Wilf, H. Mathematical Methods for Digital Computers, John Wiley and Sons, New York, 1962.

16. Taylor, A. Advanced Calculus, Ginn and Company, New York, 1955.

17. Wasg, Gerard and Reed, Georgia. Introduction to Numerical Analysis, Blaisdell Publishing Company (A Division of Ginn and Company), Waltham, Mass., 1966.

APPENDIX

```
      DIMENSIONU(41),W(41)
00001 FORMAT(1H ,7X,17HABSOLUTE ERROR = E14.8,5X,17HRELATIVE ERROR = E14
     1.8)
00002 FORMAT(1HT,9X,13HTRUE SOLUTION)
00003 FORMAT(1HS,7X,8HX VALUES8X,8HY VALUES)
00004 FORMAT(1H ,5X,F10.8,5X,F15.8)
00005 FORMAT(1H1,12X,20HRUNGE-KUTTA SOLUTION//)
00006 FORMAT(1HS,7X,8HX VALUES8X,8HY VALUES5X,9HSTEP SIZE25X,12HINTERMED
     1IATE10X,8HX VALUES5X,8HY VALUES)
00007 FORMAT(1H ,5X,F15.8,5X,F15.8,10X,F3.2)
00008 FORMAT(1H ,90X,F10.8,5X,F15.5)
00009 FORMAT(1H1,4X,37HRALSTONS MINIMUM RUNGE-KUTTA SOLUTION//)
00010 FORMAT(5E10.4)
      GIV(X,Y)=1./(1.+(SIN(Y)/COS(Y))*(SIN(Y)/COS(Y)))
      WRITE(3,2)
      WRITE(3,3)
      DO18K=1,40
      AN=K
      U(K)=AN*.1
      W(K)=ATAN(U(K))
00018 WRITE(3,4)U(K),W(K)
      CONTINUE
      WRITE(3,5)
      WRITE(3,6)
00022 READ(1,10)XA,YA,DA,XAND,YAND
      IF(DA.EQ.0.)GOTO60
00024 DO .5*DA
      JJ=0
00026 X=XA
      Y=YA
      J=1
00029 Z=GIV(X,Y)
      GOTO(31,36,40,45),J
00031 C1=DA*Z
      X=XA+DD
      Y=YA+C1/2.
      J=2
      GOTO29
00036 C2=DA*Z
      Y=YA+C2/2.
      J=3
      GOTO29
00040 C3=DA*Z
      X=XA+DA
      Y=YA+C3
      J=4
      GOTO29
00045 XA=X
      C4=DA*Z
00047 YA=YA+(C1+2.*(C2+C3)+C4)/6.
      JJ=JJ+1
      IF(JJ-LT.25)GOTO52
      WRITE(3,8)XA,YA
      JJ=0
00052 IF(XA.LT.XAND)GOTO26
00053 XAND=XA
      YAND=YA
      ABER=YAND-W(40)
      RELE=ABER/W(40)
      WRITE(3,7)XAND,YAND,DA
      WRITE(3,1)ABER,RELE
      GOTO22
00060 CONTINUE
      WRITE(3,9)
      WRITE(3,6)
00063 READ(1,10)XB,YB,DB,XBND,YBND
00064 IF(DB.EQ.0.)STOP
      KK=0
00066 X=XB
      Y=YB
      I=1
00069 Z=GIV(X,Y)
      GOTO(71,76,81,86),I
00071 C1=DB*Z
```

```
      X=XB+.4*DB
      Y=YB+.4*C1
      I=2
      GOTO69
00076 C2=DB*Z
      X=XB+.455737*DB
      Y=YB+.296978*C1+.158759*C2
      I=3
      GOTO69
00081 C3=DB*Z
      X=XB+DB
      Y=YB+.218100*C1-3.050965*C2+3.832864*C3
      I=4
      GOTO69
00086 XB=X
      C4=DB*Z
00088 YB=YB+(.174760*C1-.551481*C2+1.205536*C3+.171185*C4)
      KK=KK+1
      IF(KK.LT.25)GOTO93
      WRITE(3,8)XB,YB
      KK=0
00093 IF(XB.LT.XBND)GOTO66
00094 XBND=XB
      YBND=YB
      ABER=YBND-W(40)
      RELE=ABER/W(40)
      WRITE(3,7)XBND,YBND,DB
      WRITE(3,11)ABER,RELE
      GOTO63
      END
```

RUNGE-KUTTA METHODS AND MINIMIZATION OF TRUNCATION ERROR

by

LAURENCE RAY NEISES

B.A., St. Mary of the Plains College, 1965

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1968

Runge-Kutta methods are numerical means for solving a differential equation (or a system of differential equations) given initial values. They are one-step methods and require (for the most popular method) four iterations at each stage of the calculation. Compared with the more popular multi-step methods which require two iterations for each calculation, Runge-Kutta methods are more time consuming even on present day computers. Runge-Kutta methods are, however, self-starting and as such are used primarily to calculate starting values to be used then by the more stable predictor-corrector or multi-step methods.

Considering, then, Runge-Kutta methods only for starting the solution, we are concerned with being able to minimize the truncation or discretization error. In this report, we derive Runge-Kutta methods of second, third and fourth orders, and use this derivation, assuming certain bounds on the function and its partial derivatives, to arrive at expressions for the error term in each method. These are minimized by appropriate choices of the arbitrary parameters. With these choices for the arbitrary parameters, new coefficients are determined and used to write the minimum Runge-Kutta methods.

An analogous treatment of the derivation for a single differential equation is given for a system of differential equations.

Five differential equations are chosen as examples. Each differential equation was solved by the classical method and by the minimum method, and the error was calculated in each case after a particular number of iterations. A FORTRAN program was written and the 1410 computer was used to carry out the computations. Although some of the examples compare more favorably with the theory than others, it was pointed out that with a little foresight, one can find additional examples which either do or do not compare favorably with the theory.

If Runge-Kutta methods are to be used to start a solution, these minimum methods are generally better. Also if a system of differential equations is under consideration, these methods will be at least nearly best in the minimum truncation error sense.