Comparison of numerical approximation methods for the solution of first order differential equations
by Leon J D Rouge

A THESIS Submitted to the Graduate Faculty In partial fulfillment of the requirements for the degree of Master of Science In Applied Mathematics
Montana State University

Abstract:
Since the solution of an nth order differential equation can be reduced to the solution of a system of first order differential equations, we shall concern ourselves only with the solution of the latter. In our discussion we shall consider the basic conditions needed to assure a solution by Picard's Method, a short description of the method, determination of the error inherent in the method, extensions of the method within the region of convergence, procedures to minimize errors, and an illustration of its application.

Parallelling Picard's Method, we shall analyze the method of Taylor's series. In similar manner the difference methods are presented, pointing out in particular that, although these methods are more accurate than the analytic methods such as Picard's and Taylor's, they are step-by-step numerical approximations and, unless the results are fitted into an expression such as Newton's formula, they cannot produce a solution in analytic form.

An analytic discussion follows, outlining recommended procedures to be followed in the solution of first order differential equations, emphasizing the need for careful preliminary analysis of error terms, spacings, and the inherent characteristics of each method in order to lead most efficiently to desired results.

As an illustration of the problems encountered and the accuracy obtained by each method, an example is presented and results compared.

COMPARISON OF NUMERICAL APPROXIMATION METHODS FOR THE SOLUTION OF FIRST ORDER DIFFERENTIAL EQUATIONS

by

LEON J. D. ROUGE

A THESIS

Submitted to the Graduate Faculty

in

partial fulfillment of the requirements

for the degree of

Master of Science in Applied Mathematics

at

Montana State College

Approved:

John W. Hurst
Head, Major Department

Chairman, Examining Committee

J. A. Nelson
Dean, Graduate Division

Bozeman, Montana
June, 1952

TABLE OF CONTENTS

ABSTRACT

Since the solution of an nth order differential equation can be reduced to the solution of a system of first order differential equations, we shall concern ourselves only with the solution of the latter. In our discussion we shall consider the basic conditions needed to assure a solution by Picard's Method, a short description of the method, determination of the error inherent in the method, extensions of the method within the region of convergence, procedures to minimize errors, and an illustration of its application.

Parallelling Picard's Method, we shall analyze the method of Taylor's series. In similar manner the difference methods are presented, pointing out in particular that, although these methods are more accurate than the analytic methods such as Picard's and Taylor's, they are step-by-step numerical approximations and, unless the results are fitted into an expression such as Newton's formula, they cannot produce a solution in analytic form.

An analytic discussion follows, outlining recommended procedures to be followed in the solution of first order differential equations, emphasizing the need for careful preliminary analysis of error terms, spacings, and the inherent characteristics of each method in order to lead most efficiently to desired results.

As an illustration of the problems encountered and the accuracy obtained by each method, an example is presented and results compared.

I. INTRODUCTION

Given an nth order differential equation of the form

$$y^{(n)}(x) = f(x,y,y', \ldots y^{(n-1)})$$

subject to initial conditions $y(x_o) = y_o$, $y'(x_o) = y_o', \ldots y^{(n-1)}(x_o) = y_o^{(n-1)}$. By introduction of parameters $y_1, y_2, \ldots y_{n-1}$, the equation can be reduced to the system

$$y' = y_1,\ y_1' = y_2, \ldots y_{n-1}' = f(x,y,y_1, \ldots y_{n-1})$$

each of which is a differential equation of first order.

For example, the second order differential equation

$$y'' = f(x,y,y')$$

with initial conditions $y(x_o) = y_o$, $y'(x_o) = y_o'$, by the substitution $y'(x) = p(x)$, reduces to the system

$$p'(x) = f(x,y,p)$$

$$y'(x) = p(x)$$

Thus, our basic problem is the solution of first order differential equations of the form

$$y'(x) = f(x,y)$$

with initial condition $y(x_o) = y_o$. In our discussion we shall consider the basic conditions needed to assure a solution by each method presented, a short description of each method, an analysis of the accuracy obtained, and an analytic comparison of the various methods. Finally, as an illustration of the application of each method, a problem is presented.

II. PICARD'S METHOD

1. <u>Existence of solution</u>. The existence and uniqueness of a solution y(x) of the differential equation

(1) $$y' = f(x,y)$$

with initial condition $y(x_0) = y_0$ are guaranteed if the following three basic assumptions are satisfied:

(a) f(x,y) is singlevalued in a region R of the xy-plane which includes the point (x_0,y_0).

(b) f(x,y) is continuous in R, hence

(2) $$|f(x,y)| \leq M$$

for all (x,y) in R.

(c) f(x,y) satisfies a Lipschitz condition for any two points (x,y_1) and (x,y_2) in R:

(3) $$|f(x,y_2) - f(x,y_1)| < K$$

where K is a constant dependent on f(x,y) but independent of (x,y_1) and (x,y_2).

We define the sequence of functions:

(4) $$y_1(x) = y_0$$
$$y_{n+1}(x) = y_0 + \int_{x_0}^{x} f(x, y_n(x))\,dx \qquad n = 1, 2, 3, \ldots$$

We will show that for x within a certain interval this sequence has a limit as n → ∞, which satisfies (1).

From (4) it follows:

$$(5)\qquad |y_1(x)-y_0| = \left|\int_{x_0}^{x} f(x,y_0)\,dx\right| \leq \int_{x_0}^{x} |f(x,y_0)|\,dx \leq \int_{x_0}^{x} M\,dx \leq M|x-x_0|$$

Hence if we choose a constant α small enough, then

$$(6)\qquad |x-x_0| < \alpha, \quad |y_1(x)-y_0| < M\alpha$$

defines a new region R* about (x_0, y_0) which will be contained in R.

By induction we then obtain from (4)

$$(7)\qquad |y_n(x)-y_0| \leq \int_{x_0}^{x} |f(x,y_{n-1}(x))|\,dx \leq M|x-x_0| < \beta$$

so that $y_n(x)$ will remain in R* for $|x - x_0| < \alpha$.

To prove $\lim_{n\to\infty} y_n(x) = y(x)$ for all x in $|x - x_0| < \alpha$, observe that

$$(8)\qquad |y_2(x)-y_1(x)| \leq \int_{x_0}^{x} |f(x,y_1(x))-f(x,y_0)|\,dx \leq \int_{x_0}^{x} K|y_1(x)-y_0|\,dx$$
$$\leq \int_{x_0}^{x} KM|x-x_0|\,dx = KM\frac{|x-x_0|^2}{2!}$$

Similarly

$$(9)\qquad |y_n(x)-y_{n-1}(x)| \leq K^{n-1}M\frac{|x-x_0|^n}{n!}$$

Since:

$$(10)\qquad |y_n(x)-y_0| \leq |y_n(x)-y_{n-1}(x)| + |y_{n-1}(x)-y_{n-2}(x)| + \cdots\cdots + |y_1(x)-y_0|$$

we evidently have

$$(11)\qquad |y_n(x)-y_0| \leq \sum_{k=1}^{n} MK^{k-1}\frac{|x-x_0|^k}{k!}$$

Consequently, the sequence

$$y_n(x) = y_0 + \sum_{k=1}^{n} \left(y_k(x) - y_{k-1}(x) \right) \tag{12}$$

converges as n → ∞, and its convergence is uniform for all x in $|x - x_0| < \alpha$ since it is majorized by the exponential series. Thus we have

$$\lim_{n \to \infty} y_n(x) = y(x) \tag{13}$$

and y(x) clearly is a solution of (1) over the interval $|x - x_0| < \alpha$ with $y(x_0) = y_0$. The uniqueness of y(x) is proved by standard methods.

2. Error determination. Let us assume the exact solution of $y' = f(x,y)$ is y(x) and let $y_k(x) = Y(x)$ be an approximate solution obtained by K applications (steps) of Picard's Method. Both y(x) and Y(x) satisfy the initial conditions $y(x_0) = y_0$, $Y(x_0) = y_0$, however,

$$\frac{dY(x)}{dx} = f(x, Y(x)) + A(x, Y(x)) \tag{14}$$

where A(x,Y(x)) is the error term, introduced by the fact that Y(x) is an approximate solution.

Let us apply Picard's method of successive approximations to the difference y(x) - Y(x). Then

$$y_1(x) - Y(x) = \int_{x_0}^{x} \left[f(x,y_0) - f(x,y_0) - A(x,y_0) \right] dx = -\int_{x_0}^{x} A(x,y_0)\, dx \tag{15}$$

Since A(x,Y) is bounded in R, $|A(x,Y)| \leq N$, we obtain

$$|y_1(x) - Y(x)| \leq \int_{x_0}^{x} |A(x,y_0)|\, dx \leq N |x - x_0| \tag{16}$$

Similarly

$$|y_2(x) - Y(x)| \leq \int_{x_0}^{x} \left[|f(x, y_1(x)) - f(x, Y)| + |A(x, Y)| \right] dx \tag{17}$$

$$\leq \int_{x_0}^{x} \left[K|y_1(x) - Y| + |A(x, Y)| \right] dx$$

$$\leq KN \frac{|x-x_0|^2}{2!} + N|x-x_0|$$

and

$$|y_n(x) - Y(x)| \leq K^{n-1} N \frac{|x-x_0|^n}{n!} + \cdots + N|x-x_0| \tag{18}$$

$$\leq \sum_{k=1}^{n} N K^{k-1} \frac{|x-x_0|^k}{k!}$$

Taking limits we find

$$|\epsilon| = |y(x) - Y(x)| = \lim_{n \to \infty} |y_n(x) - Y(x)| \leq N e^{K|x-x_0|} \tag{19}$$

where ϵ is the error in the k-th approximation $y_k(x) \doteq Y(x)$. Therefore, if we wish to determine the "maximum error"* due to k applications of Picard's Method over an interval $|x - x_0| < \alpha$, we must first find an upper bound N for A(x,y) in R*, and then solve (19) for ϵ for the interval in question. The value of N, of course, depends upon k, the number of steps used in determining $y_k(x)$ from (4).

If we wish to obtain an approximate solution of a certain accuracy, we can decrease N by increasing k, or, more easily, reduce the interval $|x - x_0|$ to the extent necessary.

3. Example. The following will illustrate Picard's Method of

---

*The term "maximum error" is used here and thereafter in the sense of an upper bound for the error.

successive approximations.

Consider the equation

$$y' = y + x \tag{20}$$

which we propose to solve under the initial condition $y(0) = 0$. Clearly, $f(x,y) = y + x$ satisfies conditions (a), (b), and (c) in the entire xy-plane.

Let us take a region $R[2,-2,1,-1]$; then $M = 3$ is the maximum value of $f(x,y)$ in R. From (6) it follows that x must be restricted to the interval $|x| \leq 1/3$ in order that y remain within R. Thus the new region R^* is defined as $R^*[1/3, -1/3, 1, -1]$.

Solving the equations (4) successively, we obtain

$$\begin{aligned} &y_1(x) = \frac{x^2}{2!} \\ &y_2(x) = \frac{x^2}{2!} + \frac{x^3}{3!} \\ &y_n(x) = \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\cdots + \frac{x^n}{n!} = \sum_{k=2}^{n} \frac{x^k}{k!} = \sum_{k=0}^{n} \frac{x^k}{k!} - 1 - x \end{aligned} \tag{21}$$

whence, by taking limits we find

$$y(x) = \lim_{n\to\infty} y_n(x) = e^x - 1 - x \tag{22}$$

This is the exact solution of (20) satisfying $y(0) = 0$.

For the error approximation after two steps ($k = 2$) we calculate

$$y(0.2) = 0.02133 \tag{23}$$

and

$$|A(x,Y)| = \left| -\frac{x^3}{3!} \right| \leq \frac{(0.2)^3}{6} = \frac{0.008}{6} = N_2 \tag{24}$$

Then, since $K = |f_y(x,y)| = 1$ in R (and R*), we find

(25) $$|\epsilon| \leq \frac{0.008}{6} e^{0.2} = 0.00163$$

Thus, the maximum error obtained in this case is ± 0.00163 while the actual error in (23) is only - 0.00005.

4. Extensions of the method within R*. If we wish to evaluate y(x) at several points within the interval of convergence, two methods are available:

(a) Using $y_k(x)$, the k-th approximation to y(x), determine its numerical value for $x_1, x_2, \ldots x_n$, with consequent increase in error due to increase in $|x - x_0|$.

(b) Determine a new $\bar{y}_k(x)$, using the approximate value $y_k(x_1)$ as initial condition. Here we incorporate an error due to approximate initial value, but reduce the error due to the method, as $|x - x_1|$ is less than $|x - x_0|$.

The effect of an error ϵ in the initial condition can be determined as follows: let y(x) be the exact solution of (1) with $y(x_0) = y_0$ and let Y(x) be the exact solution of (1) which satisfies the initial condition $Y(x_0) = y_0 + \epsilon$. By applying Picard's successive approximations to the difference y(x) - Y(x) we obtain, as in (18),

(26) $$|y_n(x) - Y_n(x)| \leq \sum_{k=0}^{n} |\epsilon| \, K^k \frac{|x-x_0|^k}{k!}$$

whence, by taking limits,

(27) $$|y(x) - Y(x)| = \lim_{n\to\infty} |y_n(x) - Y(x)| \le |\epsilon| e^{K|x-x_0|}$$

Therefore, the maximum error due to use of approximate initial condition is directly proportional to ϵ, the error in the initial condition.

Summarizing, we see that by extending Picard's Method over two intervals, we have accumulated three errors:

(a) An error ϵ_1 due to the first application of Picard's Method with k steps and given (correct) initial condition:

(28) $$|\epsilon_1| \le N_1 e^{K|x-x_0|}$$

(b) An error ϵ_2 due to second application of Picard's Method with k steps and approximate initial condition:

(29) $$|\epsilon_2| \le N_2 e^{K|x-x_0|}$$

(c) An error ϵ_3 due to approximate initial condition in step (b) above:

(30) $$|\epsilon_3| \le |\epsilon_1| e^{K|x-x_0|} \le N_1 e^{2K|x-x_0|}$$

This error incorporates the error described in (a).

Hence, combining these errors, we obtain for an interval of length 2h:

(31) $$|\epsilon| \le |\epsilon_2| + |\epsilon_3| \le N_1 e^{2Kh} + N_2 e^{Kh}$$

5. Optional procedures to minimize errors in extensions of data.

In view of the above error approximations we have two principal methods of procedure for the calculation of $y(x)$ at several points within R^*:

(a) The continued use of the initial approximation $y_k(x)$ over R^*. The consequent need for greater accuracy in $y_k(x)$ eventually will necessitate more steps in the calculation of further values in order to reduce the effect of a larger exponent in (19) or (28). We can reduce the value of N, but after m steps, the reduction in N is not appreciable; on the contrary, the effect of an increase in $|x - x_0|$ greatly outweighs any further refinements of N.

(b) The extension by new approximations $\bar{y}_k(x)$ with initial conditions obtained from $y_k(x)$. In this case, the exponent in (19) or (28) can be made to remain constant by use of equal spacing and N_k can be made very small by increasing the number of steps in each approximation, but these refinements in order to be effective must outweigh the additional errors (30) due to approximate initial conditions. A natural limitation of this extension is the error we decide to accept. However, an advantage to this procedure accrues in that we are no longer limited in our extensions to the bounds of R^*, but may proceed to the bounds of R.

In a given problem, we must therefore weigh the work involved by either method of procedure with consequent accuracy in order to determine the best approach. It can be shown that a "most accurate" or "best possible" method exists for each problem, where combinations of spacings and re-applications of the method could be developed in order to extend the data over the greatest possible interval within a previously designated maximum error.

6. Illustration. In order to illustrate the two methods of procedure discussed above, we propose to calculate $y(0,2)$ if

$$y' = y + x \tag{20}$$

and $y(0) = 0$.

Using the first method, we directly obtain

$$y_2(x) = \frac{x^2}{2!} + \frac{x^3}{3!} \tag{32}$$

whence $y_2(0.2) = 0.02133$.

Using the second method, we have to calculate first $y_2(0.1) = 0.00517$ whence, after substitution into (20) and subsequent integration, we find

$$\bar{y}_2(0.2) = 0.02137 \tag{33}$$

The maximum error in (32) is found to be ± 0.00163, the maximum error in (33) is ± 0.00199. Hence, the maximum error due to two applications is slightly larger than that due to one application; however, the exact value of $y(0.2)$ is 0.02138, demonstrating that by both methods far greater tolerances have been allowed than needed.

III. METHOD OF TAYLOR'S SERIES

1. Introduction. The method of approximation by Taylor's Series is based on Taylor's series expansion of a solution $y(x)$ of (1) about $x = x_0$, and results in approximate solution satisfying the initial condition $y(x_0) = y_0$. Sufficient conditions which permit use of the method are (a) singlevaluedness of $f(x,y)$ in R and (b) existence of all derivatives of $f(x,y)$ in R.

From Taylor's theorem we obtain

$$(34) \quad y(x) = y(x_0) + y'(x_0)(x-x_0) + y''(x_0)\frac{(x-x_0)^2}{2!} + \cdots + y^{(n)}(x_0)\frac{(x-x_0)^n}{n!} + R_n$$

where the derivatives of $y^{(k)}(x_0)$ are calculated according to (1) and R_n is the remainder term,

$$(35) \quad R_n = y^{(n+1)}\big(x_0 + \theta(x-x_0)\big)\frac{(x-x_0)^{n+1}}{(n+1)!} \qquad 0 < \theta < 1$$

2. Error estimates. The maximum error incurred by using only the first n+1 terms in (34) can be determined from the maximum value of R_n over the interval considered.

The error due to approximate initial conditions is calculated in the same manner as for Picard's Method.

As an illustration let us again consider the equation

$$(20) \quad y' = y + x$$

with initial condition $y(0) = 0$. We have

$$(36)\quad \begin{aligned} y'' &= y'+1 = y+x+1, & y''(0) &= 1 \\ y''' &= y'' = y+x+1, & y'''(0) &= 1 \\ y^{(k+1)} &= y^{(k)} = y+x+1, & y^{(k+1)}(0) &= 1, \qquad k = 2, 3, \ldots \end{aligned}$$

Thus we obtain from (34)

$$(37)\quad y(x) = \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\cdots$$

and $y(0.1) = 0.00517$. Similarly,

$$(38)\quad y'(0.1) = 0.10517,\ y''(0.1) = 1.10517,\ \ldots\ y^{(k+1)}(0.1) = 1.10517,\ k=2,3,\ldots$$

whence

$$(39)\quad y(x) = 0.00517 + 0.10517(x-0.1) + \frac{1.10517}{2}(x-0.1)^2 + \frac{1.10517}{6}(x-0.1)^3 + \cdots$$

and $y(0.2) = 0.02139$.

Tocalculate the error in (37) we observe that $y^{(iv)}(x) < 1.5$ over the interval $0 \leq x \leq 0.1$. Hence the maximum error in (37) is less than ± 0.00001.

The maximum error in (39) is found by adding the error due to the approximate value of $y(0.1)$ to the error due to the use of four terms of Taylor's series in (39). This maximum error turns out to be ± 0.000035, and is thus slightly larger than the error that would have been obtained by calculating $y(0.2)$ directly from (37).

3. Modified method of Taylor's series. In the standard method of approximation by Taylor's series, we obtain a series for $y(x)$, which is

evaluated at $x = x_1$. This value $y(x_1) = y_1$ is used as initial value in determining $y(x_2)$ by reapplication of the method:

$$(40)\qquad y_2 = y(x_2) = \sum_{k=0}^{n} y^{(k)}(x_1)\,\frac{(x_2-x_1)^k}{k!} + R_{2,n}$$

or, in general:

$$(41)\qquad y_m = y(x_m) = \sum_{k=0}^{n} y^{(k)}(x_{m-1})\,\frac{(x_m-x_{m-1})^k}{k!} + R_{m,n}$$

Thus, we can approximate the values y_1, y_2, ...y_m within the interval of convergence in R. This process, however, requires evaluation of n derivatives for each application of the method, an operation which is sometimes quite complicated.

As an approximation to the standard procedure, a possible alternative is to differentiate

$$(42)\qquad y(x) = \sum_{k=0}^{n} y^{(k)}(x_1)\,\frac{(x-x_1)^k}{k!}$$

successively and calculate the derivatives $y'(x_2)$, $y''(x_2)$, ...$y^{(n)}(x_2)$ from the resulting expressions. Then another approximation can be set up

$$(43)\qquad y(x) = \sum_{k=0}^{n} y^{(k)}(x_2)\,\frac{(x-x_2)^k}{k!}$$

which may be treated similarly to obtain $y(x_3)$ and $y'(x_3)$, $y''(x_3)$, ...$y^{(n)}(x_3)$. Proceeding this way, we can approximate $y(x_4)$,...$y(x_m)$.

To illustrate the difference between the two methods of attack, let us again consider the equation

$$(20)\qquad y' = y + x$$

Above we found $y(0.2) = 0.02139$ if $y(0) = 0$. Here we use as approximation (37),

$$(44) \quad y(x) = \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!}$$

which we differentiate to obtain

$$(45) \quad \begin{aligned} y'(x) &= x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}, & y'(0.1) &= 0.10517 \\ y''(x) &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}, & y''(0.1) &= 1.10517 \\ y'''(x) &= 1 + x + \frac{x^2}{2!}, & y'''(0.1) &= 1.105 \\ y^{iv}(x) &= 1 + x, & y^{iv}(0.1) &= 1.1 \\ y^{v}(x) &= 1, & y^{v}(0.1) &= 1 \end{aligned}$$

Then from (43) we find

$$(46) \quad y(x) = 0.00517 + 0.10517(x - 0.1) + \cdots$$

whence $y(0.2) = 0.02139$, which is exactly the value calculated before.

Clearly, the second method of procedure is to be preferred if the calculation of the successive derivatives becomes extremely difficult.

IV. DIFFERENCE METHODS

1. Introduction. In the difference methods we are primarily concerned with step-by-step approximations of the numerical values of $y(x)$ of (1) without deriving an analytic expression for $y(x)$. Although the previously described methods are suitable for calculation of approximate values of $y(x)$ within the interval of convergence, greater accuracy is usually obtained by the difference methods. We shall consider two, the Adams Method and the Runge-Kutta Method, together with their allied difference formulae (Simpson's Rule, Milne's Rule, etc.).

For the derivation of the difference methods a sufficient condition is that $f(x,y)$ be of class C^n in R. For actual calculations it is only necessary to have tabulated data.

The difference methods are extremely useful when great difficulty is involved in carrying out the required operations in Picard's Method or in the method of Taylor's series. The principal disadvantage to the Adams Method is that several values of (x,y) must be known in order to secure accurate extrapolations or interpolations. To overcome this disadvantage, it is necessary to employ other methods as "starter" methods and then to correct these approximations by difference formulae. Once sufficient data are available, the difference methods generally furnish more accurate approximations over a larger interval. On the other hand, Runge-Kutta Method requires only one initial condition, but the necessary calculations are somewhat more involved.

2. Adams Method. In the derivation of the Adams Method we approximate

$y(x)$ in (1) at the point (x_{n+h}, y_{n+h}) by two or more terms of a Taylor's series about (x_n, y_n)

(47) $y(x_{n+h}) = y(x_n) + h\, y'(x_n) + \cdots\cdot$

Denoting $y(x_{n+h})$ by y_{n+1}, $y(x_n)$ by y_n, $y'(x_n)$ by f_n, etc., we obtain for the first two terms

(48) $y_{n+1} - y_n \cong h f_n$

or

(49) $f_n \cong \dfrac{y_{n+1} - y_n}{h}$

which represents the average slope from y_n to y_{n+1}.

Similarly, we obtain for the first three terms of the Taylor's series

(50) $y_{n+1} - y_n \cong h f_n + \frac{h^2}{2} f_n'$

Here $f_n' = y''(x_n)$ which can be further expressed as

(51) $f_n' \cong \dfrac{f_n - f_{n-1}}{h}$,

the average rate of change of slope from y_{n-1} to y_n. The numerator is defined as the "first backward difference" or ∇f_n.

Using this nomemclature, we obtain the formula of first differences:

(52) $y_{n+1} - y_n \cong h f_n + \frac{1}{2} h \nabla f_n = h(f_n + \frac{1}{2} \nabla f_n)$

Similarly we obtain the formula of fourth differences:

(53) $y_{n+1} - y_n \cong h f_n + \frac{1}{2} h \nabla f_n + \frac{5}{12} h \nabla^2 f_n + \frac{3}{8} h \nabla^3 f_n + \frac{251}{720} h \nabla^4 f_n$

This method requires knowledge of the values of n+1 points, where n is the order of differences desired.

In similar manner we can derive the four basic formulae leading to the standard difference formulae, as indicated below:

$$y_{n+1} \cong y_n + h\left(f_n + \tfrac{1}{2}\nabla f_n + \tfrac{5}{12}\nabla^2 f_n + \tfrac{3}{8}\nabla^3 f_n + \tfrac{251}{720}\nabla^4 f_n + \tfrac{475}{1440}\nabla^5 f_n\right) \tag{54}$$

$$y_n \cong y_{n-1} + h\left(f_n - \tfrac{1}{2}\nabla f_n - \tfrac{1}{12}\nabla^2 f_n - \tfrac{1}{24}\nabla^3 f_n - \tfrac{19}{720}\nabla^4 f_n - \tfrac{27}{1440}\nabla^5 f_n\right) \tag{55}$$

$$y_{n-1} \cong y_{n-2} + h\left(f_n - \tfrac{3}{2}\nabla f_n + \tfrac{5}{12}\nabla^2 f_n + \tfrac{1}{24}\nabla^3 f_n + \tfrac{11}{720}\nabla^4 f_n + \tfrac{11}{1440}\nabla^5 f_n\right) \tag{56}$$

$$y_{n-2} \cong y_{n-3} + h\left(f_n - \tfrac{5}{2}\nabla f_n - \tfrac{23}{12}\nabla^2 f_n - \tfrac{3}{8}\nabla^3 f_n - \tfrac{19}{720}\nabla^4 f_n - \tfrac{11}{1440}\nabla^5 f_n\right) \tag{57}$$

Errors attributed to the use of the difference formulae can be approximated by the first neglected term. The error approximation will be particularly accurate if the difference term immediately preceding the first neglected term has zero coefficient. Even without knowledge of the difference term used for the error approximation, we can closely estimate the error by evaluation of the corresponding derivative of y at x_n. Correctness of this estimate is based on the relationship between differences and derivatives when the spacing h is small. It can be shown that the nth derivative evaluated at a point is the limit of the quotient of the nth difference of the function by the nth power of h when h tends to zero.*

Examination of (54) through (57) reveals that

(a) (54) alone is the "backward difference" formula as derived from

---

*<u>Cours d'Analyse Infinitesimale</u>, Ch. J. de la Valle Pouissin.

Adams Method. Using fourth differences only, the error approximation if $(475/1440)h\,\nabla^5 f_n$ or $(475/1440)\,h^6 y^{vi}(x_n)$. Equation (54) may also be written

$$(58)\quad y_{n+1} \cong y_n + \frac{h}{12}(23f_n - 16f_{n-1} + 5f_{n-2}) + \frac{3}{8}h^4 y^{iv}(x_n)$$

where second differences are used, the error being the third difference converted to derivative form as described above. This formula is used in one-step extrapolation from y_n to y_{n+1}.

(b) adding (54) to (55) produces:

$$(59)\quad y_{n+1} \cong y_{n-1} + h\left(2f_n + \frac{1}{3}\nabla^2 f_n + \frac{1}{3}\nabla^3 f_n + \frac{232}{720}\nabla^4 f_n + \frac{448}{1440}\nabla^5 f_n\right)$$

Absence of the first difference leads to the reasonably accurate "predictor" formula:

$$(60)\quad y_{n+1} \cong y_{n-1} + 2hf_n + \frac{1}{3}h^3 y'''(x_n)$$

Another useful extrapolation formula can be derived from (59) by using only third order differences:

$$(61)\quad y_{n+1} \cong y_{n-1} + \frac{h}{3}(7f_n - 2f_{n-1} + f_{n-2}) + \frac{1}{3}h^4 y^{iv}(x_n)$$

(c) Adding (54), (55), (56), and (57), we obtain

$$(62)\quad y_{n+1} \cong y_{n-3} + h\left(4f_n - 4\nabla f_n + \frac{8}{3}\nabla^2 f_n + \frac{14}{45}\nabla^4 f_n - \frac{14}{45}\nabla^5 f_n\right)$$

from which we derive Milne's Rule:

$$(63)\quad y_{n+1} \cong y_{n-3} + \frac{4h}{3}(2f_n - f_{n-1} + 2f_{n-2}) + \frac{14}{45}h^5 y^{v}(x_n)$$

(d) Simpson's Rule results from the addition of (55) and (56)

$$(64) \quad y_n \cong y_{n-2} + h\left(2f_n - 2\nabla f_n + \frac{1}{3}\nabla^2 f_n - \frac{8}{720}\nabla^4 f_n - \frac{16}{1440}\nabla^5 f_n\right)$$

producing:

$$(65) \quad y_n \cong y_{n-2} + \frac{h}{3}\left(f_n + 4f_{n-1} + f_{n-2}\right) - \frac{h^5}{90} y^{v}(x_n)$$

Absence of third differences makes Simpson's Rule an accurate corrector of previously obtained approximations. It can also be used as an interpolation formula for f_{n-1}.

(e) Adding (56) and (57), we obtain

$$(66) \quad y_{n-1} \cong y_{n-3} + \frac{h}{90}\left(-f_n + 34f_{n-1} + 114f_{n-2} + 34f_{n-3} - f_{n-4}\right)$$

which can be used as a corrector.

(f) Adding the corresponding difference formula for y_{n+2} to (54), we obtain

$$(67) \quad y_{n+2} \cong y_n + \frac{h}{3}\left(19f_n - 20f_{n-1} + 7f_{n-2}\right) + \frac{8}{3} h^4 y^{iv}(x_n)$$

which can be used as a two-step predictor. It is particularly useful if the value of the fourth derivative is small.

We may determine the maximum error due to reapplication of difference formulae by use of methods previously described. Since, in these formulae, we combine several previously determined values of y in order to extrapolate further values of y, this has a "bridging" effect and thereby reduces the actual error considerably.

To illustrate Adams Method, we again consider the equation

(20) $y' = y + x$

with initial conditions $y(0) = 0$ and $y_1(0.1) = 0.00517$. Using (52), we obtain $y(0.2) = 0.02095$, with an error of ± 0.00046.

Using two steps of Picard's approximation we obtained $y(0.2) =$ 0.02133, which is considerably more accurate.

Starting with correct values for three points, we may approximate the fourth by (61). We obtain $y(0.3) = 0.04981$, with an error of ± 0.00004. The accuracy is considerably greater than that obtained by use of first difference formula (52).

Extrapolating by Milne's Rule (63) to $y(0.4)$ by use of the above data, we obtain $y(0.4) = 0.09181$, with an error of ± 0.000004, plus error due to error in one initial value. If we use correct initial values, we obtain $y(0.4) = 0.09182$, not an appreciable improvement. The correct value is 0.09182.

If we use values of $y(0.1)$, $y(0.2)$, and $y(0.3)$ obtained by two steps of Picard's Method, the approximate value of $y(0.4)$ by Milne's Rule is 0.09173, whereas by Picard's Method it is only 0.09067. This serves to emphasize the greater accuracy of the difference formulae.

4. Runge-Kutta Method. The Runge-Kutta Method, based on Taylor's series, but requiring no evaluation of derivatives, has the advantage of requiring only one set of initial values for (x,y) in addition to (1). Furthermore, in the absence of an analytic expression for $y'(x)$, solutions may be obtained by use of tabulated data.

Formulae employed in the method are:

(a) Second order accuracy:

$$y_{n+1} = y_n + \frac{h}{2}\left(f(x_n,y_n) + f(x_n+h, y_n+hf_n)\right) \tag{68}$$

(b) Third order accuracy:

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{6}(a_1 + 4a_2 + a_3) \\ a_1 &= hf(x_n,y_n) \\ a_2 &= hf(x_n+\tfrac{1}{2}h, y_n+\tfrac{1}{2}a_1) \\ a_3 &= hf(x_n+h, y_n+2a_2-a_1) \end{aligned} \tag{69}$$

(c) Fourth order accuracy:

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{6}(b_1 + 2b_2 + 2b_3 + b_4) \\ b_1 &= hf(x_n,y_n) \\ b_2 &= hf(x_n+\tfrac{1}{2}h, y_n+\tfrac{1}{2}b_1) \\ b_3 &= hf(x_n+\tfrac{1}{2}h, y_n+\tfrac{1}{2}b_2) \\ b_4 &= hf(x_n+h, y_n+b_3) \end{aligned} \tag{70}$$

Errors due to the method are normally computed by use of a comparison test: let the error associated with using a procedure of nth order accuracy k times with spacing h be expressed in the form $K\,kh^{n+1}$. It may be assumed that the correct result is obtained by adding this error to the calculated result. Further, assuming that K is not strongly dependent on the other variables, we can compare this error ϵ_h with that obtained by doubling the spacing h over the interval in question, ϵ_{2h} :

$$\frac{\epsilon_h}{\epsilon_{2h}} = \frac{Kkh^{n+1}}{K\frac{k}{2}(2h)^{n+1}} = \frac{1}{2^n} \tag{71}$$

If we subtract the ordinate ($y^{(2)}$), obtained by double spacing, from that obtained by single spacing ($y^{(1)}$)

$$y^{(1)} - y^{(2)} = Kkh^{n+1}(2^n - 1) \tag{72}$$

which is (2^n-1) times the error in $y^{(1)}$. Therefore, we can approximate the error in $y^{(1)}$ by

$$\epsilon_h = \frac{y^{(1)} - y^{(2)}}{2^n - 1} \tag{73}$$

and add this error to $y^{(1)}$ to obtain a corrected value.

Errors due to reapplication of the method are evaluated in the same manner as for other methods.

To illustrate the Runge-Kutta Method, let us again consider

$$y' = y + x \tag{20}$$

with initial condition $y(0) = 0$.

Using third order accuracy, we obtain

$$y(0.1) = \frac{1}{6}(0.02 + 0.011) = 0.00517 \tag{74}$$

The correct value is $y(0.1) = 0.00517$. Similarly,

(75) $y(0.2) = 0.02140$ (0.02140)

The error is found to be $\epsilon(0.2) < 0.000005$. Likewise,

(76) $y(0.3) = 0.04985$ (0.04985)

$y(0.4) = 0.09181$ (0.09182)

where $\epsilon(0.4) < 0.000005$.

V. ANALYSIS OF METHODS

Given a differential equation of first order with initial condition $y(x_0) = y_0$, the first problem is to determine the region R* in which we can expect to find a solution. Examination of $f(x,y)$ to determine its integrability and differentiability will disclose the feasibility of applying either Picard's Method or the Method of Taylor's series. Should $f(x,y)$ be of a type difficult to integrate or differentiate, we may turn immediately to the Runge-Kutta Method either completely or long enough for evaluation of sufficient approximations of $y(x_k)$ in order to make possible application of difference formulae.

If we decide to apply Picard's Method or to construct a Taylor's series, the accuracy desired in the last extrapolation will determine the number of applications of the method, each involving a certain number of steps. If more than one application is required, the spacing must be determined. Here, we must reach a compromise between large spacing with fairly rapid, but more inaccurate results, and very small spacing with slower, more accurate results if the errors due to approximated initial conditions do not completely offset the advantages of small spacings.

In the event we are unable to secure the accuracy desired, we must abandon these methods either completely or after enough approximations (within accuracy limits) to apply difference formulae.

For each reapplication of Picard's or Taylor's methods, we may select a new K in the Lipschitz condition to be satisfied by $f(x,y)$, thereby decreasing the maximum error estimate for all intervals except possibly those in which K is near its upper bound over the entire region.

Once we have determined two or more points (x,y) beyond (x_o,y_o), we can begin to apply the difference formulae with proper order accuracy, such as (58) with its relatively high accuracy, (59) and (67) which possess fair accuracy with additional advantage of serving as two step predictors.

With a total of four points (x,y) including (x_o,y_o), we may apply Milne's Rule with its high degree of accuracy as a one-step predictor, followed by Simpson's Rule as a corrector. This procedure is carried out to the final point (x,y) if the desired accuracy is attainable.

Should the desired accuracy at (x_k,y_k) not be realized by the difference formulae, it will become necessary to apply the Runge-Kutta Method. Here we must determine the largest spacing consistent with the desired accuracy. After each step of the method, we will determine the error term as previously described, adding this to the approximated values, and continuing to the next point.

If an analytic expression is desired to approximate the solution between (x_o,y_o) and (x_n,y_n), this may be obtained by use of Newton's interpolation formula:

$$(77) \qquad y(x) = a_0 + a_1(x-x_n) + a_2(x-x_n)(x-x_{n-1}) + \cdots\cdots$$

In general, careful preliminary analysis of error terms, spacings, and the inherent characteristics of each method will lead most efficiently to results within the desired degree of accuracy.

VI. COMPARISON OF METHODS BY ILLUSTRATIVE PROBLEM

Let us consider the first order non-linear differential equation

$$y' = (y + x)^2 \qquad (78)$$

with initial condition $y(0) = 0$. We wish to determine $y(0.4)$ and $y(0.8)$ correct to three decimal places.

In our preliminary analysis we see that $f(x,y) = (y + x)^2$ is continuous throughout the xy-plane, therefore we may select any R about (0,0). To secure a minimum error by analytic methods, we select an R* with a small M which will contain at least one of the desired points, or we shall be faced with more involved methods of successive approximations.

Selection of a least upper bound for $f(x,y)$ containing $y(0.4)$ results in a region R*(0.5,-0.5,0.5,-0.5) in which $M = 1$ and $K = 2$. This will produce an error approximation at $y(0.4)$

$$\epsilon(0.4) \le N_k e^{0.8} \qquad (79)$$

By minimizing N through use of several steps of Picard's Method we can theoretically obtain the desired accuracy. In proceeding to $y(0.8)$ we must use two or more overlapping regions beyond R*, in each of which we evaluate a new K.

By Picard's Method the following results are obtained:

$$\begin{aligned} y_1(0.4) &= 0.02133 & \quad |\epsilon| &\le 0.03809 \\ y_2(0.4) &= 0.022789 & \quad |\epsilon| &\le 0.00267 \\ y_3(0.4) &= 0.022789 & \quad |\epsilon| &\le 0.00118 \end{aligned} \qquad (80)$$

Since R* does not contain y(0.8), we establish a new R** containing R* and, using Picard's Method, we obtain for y(0.8),

$$\begin{aligned} y_1(0.8) &= 0.17067 \\ y_2(0.8) &= 0.22529 \\ y_3(0.8) &= 0.22769 \end{aligned} \tag{81}$$

The error in this case is extremely difficult to evaluate, but it can be seen that $\lim_{n\to\infty} y_n(0.8) = y(0.8)$.

Exact values for y(0.4) and y(0.8) are 0.02279 and 0.22964, respectively.

Since the maximum error for $y_3(0.4)$ is too large to meet the requirements and further steps in the method are quite involved, we continue the method only to tabulate data for comparison purposes.

Using the formulae

$$\begin{aligned} y_1(x) &= \frac{x^3}{3} \\ y_2(x) &= \frac{x^3}{3} + \frac{2x^5}{15} + \frac{x^7}{63} \\ y_3(x) &= \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{114x^9}{8505} + \frac{1306x^{11}}{11\cdot189\cdot275} + \frac{4x^{13}}{13\cdot15\cdot63} + \frac{x^{15}}{15(63)^2} \end{aligned} \tag{82}$$

we obtain results as shown in Table I.

By use of $y_3(x)$ in (82), we extend our data from y(0.4) to y(0.8) by a reapplication of Picard's Method obtaining

$$\begin{aligned} \bar{y}_1(0.8) &= 0.18327 \\ \bar{y}_2(0.8) &= 0.22300 \\ \bar{y}_3(0.8) &= 0.22916 \end{aligned} \tag{83}$$

TABLE I

COMPARISONS OF VALUES BY PICARD'S METHOD

x	y_1	y_2	y_3	y
0.0	0	0	0	0
0.1	0.00033	0.00033	0.00033	0.00033
0.2	00267	00271	00271	00271
0.3	00900	00932	00933	00933
0.4	02133	02279	02279	02279
0.5	04167	04596	04629	04626
0.6	07200	08281	08401	08414
0.7	11423	13805	14177	14229
0.8	17067	22529	22769	22964

By the method of Taylor's series we derive

(84) $y(x) = \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62}{2835}x^9$

Whereby, using the first three terms,

(85) $y(0.4) = 0.02279, \qquad |\epsilon| \leq 0.00147$

By using four terms,

(86) $y(0.4) = 0.022792, \qquad |\epsilon| \leq 0.00049$

Since the latter error is very close to maximum tolerance, we proceed to y(0.8) by direct calculation from (84) only for comparative purposes,

(87) $y(0.8) = 0.22820$

and, again for comparative purposes, we apply an extension of the method over a new interval about y(0.4), with the results that:

(88) $y(0.8) = 0.22952$

from which it can be seen that apparently reapplication of the method leads to more accurate results since the true value of y(0.8) is 0.22964.

For the purpose of comparison we obtain, by use of the modified method of Taylor's series (four terms),

(89) $y(0.4) = 0.02279$

$y(0.8) = 0.22852$

Since the analytic methods do not yield the desired results without considerable difficulty, we proceed to the difference methods. First, by Milne's (63), using the data from Picard's Method for y(0.1), y(0.2), y(0.3), we obtain (see Table II):

$$(90) \quad y(0.4) = 0.02279, \qquad |\epsilon_1| \leq 0.00007$$
$$y(0.5) = 0.04618, \qquad |\epsilon_2| \leq 0.00010$$

The error ϵ_2 added to ϵ_1 produces an error $|\epsilon_2'| \leq 0.00018$. Continuing:

$$(91) \quad y(0.6) = 0.08396, \qquad |\epsilon_3| \leq 0.00018 \qquad |\epsilon_3'| \leq 0.00036$$
$$y(0.7) = 0.14197, \qquad |\epsilon_4| \leq 0.00025 \qquad |\epsilon_4'| \leq 0.00061$$

Whereby we are exceeding the error tolerance. To complete our data,

$$(92) \quad y(0.8) = 0.22895, \qquad |\epsilon_5| \leq 0.00049, \qquad |\epsilon_5'| \leq 0.00110$$

Inasmuch as the desired accuracy was not obtained by Milne's Rule, we proceed to the Runge-Kutta Method.

In solving the problem by the Runge-Kutta Method, we have several options:

(a) Approach y(0.4) and y(0.8) by spacing of 0.4, checking y(0.8) by a double spacing 0.8.

(b) Approach y(0.4) and y(0.8) by spacing of 0.2, checking with double spacing 0.4.

(c) Approach y(0.4) and y(0.8) by spacing of 0.1, checking with double spacing 0.2.

TABLE II

DIFFERENCE TABLE — MILNE'S RULE

x	y	f(x,y)	$\Delta f(x,y)$	$\Delta^2 f(x,y)$	$\Delta^3 f(x,y)$	$\Delta^4 f(x,y)$
0.0	0.0	0.0	0.01007	0.02095	0.00263	0.00219
0.1	00033	01007	03102	02358	00482	00314
0.2	00271	04109	05460	02840	00796	00559
0.3	00934	09569	08300	03634	01355	00813
0.4	02272	17869	11964	04989	02168	01554
0.5	04618	29833	16953	07157	03722	
0.6	08396	46786	24110	10879		
0.7	14197	70896	34989			
0.8	22895	1.05885				

(d) Combinations of the above.

By the first method we obtain, for third order accuracy,

(93) $y(0.4) = 0.02155$

and for fourth order accuracy,

(94) $y(0.4) = 0.02271$

Solving for y(0.8) from y(0.4),

(95) $y(0.8) = 0.22958$

Checking results by obtaining y(0.8) in one step:

(96) $y(0.8) = 0.20478$

which gives an error of 0.00165, which when added to (95), gives y(0.8) a corrected value of 0.23125.

Since we have not secured the desired accuracy, we proceed to the second method (b). Determination of y(0.2) from y(0) yields:

(97) $y(0.2) = 0.00271$

Similarly, we obtain y(0.4) from y(0.2):

(98) $y(0.4) = 0.02279$

Then, in order to correct our results, we use the value of y(0.4) in (90), which indicates an error of less than 0.00001.

Proceeding in a similar manner, we obtain

$$\begin{aligned} &(99) \quad y(0.6) = 0.08415 \\ & \quad y(0.8) = 0.22968 \end{aligned}$$

and calculating y(0.8) from y(0.4) as in (95), we obtain an error less than 0.00001, whereby our final corrected value for y(0.8) is 0.22968.

This method obviously gives the desired accuracy.

In order to compare the accuracy obtained by option (c), we list the following data:

$$\begin{aligned} &(100) \quad y(0.1) = 0.00033 \qquad && y(0.5) = 0.04631 \\ & \quad y(0.2) = 0.00271 && y(0.6) = 0.08415 \\ & \quad y(0.3) = 0.00934 && y(0.7) = 0.14231 \\ & \quad y(0.4) = 0.02280 && y(0.8) = 0.22968 \end{aligned}$$

As a more compact comparison of the results obtained, Table III has been prepared.

In conclusion, it is desired to point out that by variations of the above described methods, as outlined in Section V, we can also obtain similar results with the desired accuracy.

TABLE III

COMPARISON OF RESULTS

x	True y	Picard's $y_3(x)$	Taylor's (4 terms)	Modified* (4 terms)	Milne's	Runge-Kutta h=0.4	Runge-Kutta h=0.2	Runge-Kutta h=0.1
0.1	0.00033	0.00033	0.00033					0.00033
0.2	00271	00271	00271				0.00271	00271
0.3	00933	00933	00933					00934
0.4	02279	02279	02279	0.02279	0.02279	02271	02279	02280
0.5	04626	04629	04626		04618			04631
0.6	08414	08401	08407		08396		08415	08415
0.7	14229	14177	14194		14197			14231
0.8	0.22964	0.22916* 0.22769	0.22952* 0.22820	0.22852	0.22895	0.23125	0.22968	0.22968

*By two applications.

VII. CONCLUSION

In this paper we have confined ourselves to the basic applications of the analytic and difference methods of numerical approximation. However, many of the side aspects of these methods which were not considered herein should furnish ample material for a further more thorough discussion. Other subjects for further investigation are:

(a) Analysis of "type" equations, with the possibility of developing methods of attack most favorably suited to each. This might easily take the form of a compendium similar to the Differentialgleichungen, Losungsmethoden und Losungen by E. Kamke.

(b) Analysis of errors inherent to each method, with the possibility of further reducing the maximum error by more precise analysis of the influencing factors. This would lead to greater extensions of data within predesignated tolerances.

Literature Consulted

Hildebrand, F. B. 1950. ADVANCED CALCULUS FOR ENGINEERS.
594 pp., Prentice-Hall, Inc., New York.

Kamke, E. 1948. DIFFERENTIALGLEICHUNGEN, LOSUNGSMETHODEN UND LOSUNGEN.
666 pp., Chelsea Publ. Co., New York.

Knopp, Konrad. 1928. THEORY AND APPLICATION OF INFINITE SERIES.
571 pp., Blackie & Son Limited, London and Glasgow.

Levy, H. and Baggott, E. A. 1934. NUMERICAL STUDIES IN DIFFERENTIAL EQUATIONS. (Vol. I)
238 pp., Watts & Co., London.

Milne, W. E. 1949. NUMERICAL CALCULUS.
393 pp., Princeton Univ. Press, Princeton, N.J.

Widder, David V. 1947. ADVANCED CALCULUS.
432 pp., Prentice--Hall, Inc., New York.

Vallee Poussin, Ch.-J. de la. 1946. COURS D'ANALYSE INFINITESIMALE (Vol. I).
460 pp., Dover Publications, New York.

[redacted]

[redacted]

Rouge, L.J.D.
Comparison of numerical approximation methods...differential equations.

DATE	ISSUED TO
MAY 28 '54	Glenn Ingram Math. Dept.
3/14/[illegible]	Ted Koch[illegible] 503 W. College
	David Seibel 120 S. Tracy
4-30-66	[illegible]
JUN 22 1961	Frazi[illegible]

[redacted]