THESIS

on

A NEW METHOD OF NUMERICAL INTEGRATION OF

DIFFERENTIAL EQUATIONS OF THE THIRD ORDER.

Submitted to the

OREGON STATE AGRICULTURAL COLLEGE

In partial fulfillment of
the requirements for the
Degree of

MASTER OF SCIENCE

BY

RALPH CORNELIUS CONRAD

MAY 1, 1933

APPROVED:

Redacted for privacy

Professor of Mathematics

In Charge of Major

Redacted for privacy

Chairman of Committee on Graduate Study.

7 Aug. 33 Gaither est. 3.30 + $.75 bind.

TABLE OF CONTENTS

INTRODUCTION

The subject of numerical integration is one of considerable interest today in the fields of physics, astronomy, and ballistics. Although the foundation of such work is due to Euler, it is only rather recently that its development has gone forward to any great extent. In 1890 Picard established the method of successive approximations on sound bases. Among others who have contributed, one might mention Moulton whose work is largely in connection with celestial mechanics. Antedating the work of Moulton by a few years is the method of solution devised by J. C. Adams. In 1894, Runge gave his method of solution which was extended in 1901 by Kutta. Among recent publications are those of W. E. Milne, who has devised a method of solution applicable to both first and second order equations. In his latest study he has given a solution for a special type of second order equation. Nystrom has given a very excellent review of some special forms of numerical integration and supplemented it with his own work. Through the work of these and many others, the subject of numerical integration is emerging as a specialized field of mathematics of such a practical nature as to demand constantly increasing attention.

The next few pages give a brief resume of a few of the methods mentioned above and their application to some simple examples. The last pages of this paper are given to a discussion of a new method of integration for a special type of third order equation. It has been worked out in more detail than previously discussed methods as it represents the initial attack on this type of equation. It is hoped that it will contribute its bit to the common cause--Numerical Integration.

THE METHOD OF DIFFERENCES

This method is dependent on the summing of a rapidly converging series. It has the advantage that it will approximate to any desired degree of accuracy and is the easiest to perform when all the work is done by hand. It makes use of Newton's interpolation formula,

$$P_n(x) = u_0 + s\Delta u_0 + \frac{s(s-1)}{2!}\Delta^2 u_0$$

$$+ \frac{s(s-1)(s-2)}{3!}\Delta^3 u_0 + \frac{s(s-1)(s-2)(s-3)}{4!}\Delta^4 u_0 \cdots\cdots$$

integrated over an interval and put in usable form for backward differences. Thus it becomes

$$(1) \quad \int_{x_0}^{x_1} P_n(x)\,dx = h\left[u_0 - \tfrac{1}{2}\Delta u_0 - \tfrac{1}{12}\Delta^2 u_0 - \tfrac{1}{24}\Delta^3 u_0 - \tfrac{19}{720}\Delta^4 u_0 - \tfrac{3}{160}\Delta^5 u_0 \cdots\cdots\right]$$

where h is the interval over which the integration is performed. This is known as Gregory's Formula.

1. Let us look more in detail to its application to a first order equation. To begin we shall assume we have four or five values of y and y'. This will give us a set of differences complete to the third or fourth. To facilitate rapid and easy calculation we shall arrange the work in columns in the following scheme.

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$	A	B

Besides the table of differences we may wish to record other items of the work depending on the particular problem at hand. These may be placed in the column at the right hand side.

To continue with the computation, we first find a trial value of y'. To do this we shall assume for our next row of work that our highest order difference is equal to the previous highest order difference immediately above. We shall now write in the trial values for y' and its differences making use of the equation,

$$(2) \qquad \Delta^{n} y'_{n} = \Delta^{n} y'_{n+1} + \Delta^{n+1} y'_{n}$$

Our next step is to find a trial value of y by applying Gregory's Formula (1) to these values. The value of y may now be found by using (2) where the primes are dropped. We should bear in mind at this point that all these values are dependent on the assumption that our highest order differences are constant. In general this is not absolutely true, so we shall correct in the following manner. Using our new found value of y, let us solve the equation itself for y'. We may now if necessary correct our y' differences by applying (2) in the slightly modified form,

$$(2a) \qquad \Delta^{n+1} y_{n} = \Delta^{n} y'_{n} - \Delta^{n} y'_{n-1}$$

At this point we shall resort to Gregory's Formula again and continue until our values check.

A few moments inspection of an example will make the work clear. As our equation let us take

$$\frac{dy}{dx} + \frac{x}{y} + \frac{y}{2} + 1 = 0$$

with the initial conditions that y= 3.75 when x = -10. For the sake of simplicity we shall omit the work of starting and begin with the following row of work.

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$	$\frac{x}{y}$	$-(\frac{y}{2}+1)$
-7.2	3.11496	-9721	-232	-16	6	-.24606	-598	-38	-11	-3	-2.31142	-2.55748

Let us choose h equal to .4 and our next value of x will be -6.8. We shall set y' equal to -3 and using (2) we arrive at the following values;

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$	$\frac{x}{y}$	$-(\frac{y}{2}+1)$
-7.2	3.11496	-9721	-232	-16	-6	-.24606	-598	-38	-11	-3	-2.31142	-2.55748
-6.8	3.01528	-9970				-.25256	-650	-52	-14	-3		

Using (1) we find Δy equal to .09970 and y equal to 3.01528. Solving for x/y and (y/2) + 1 we find them to be -2.25518 and 2.50764 respectively. Substituting in the equation we get a corrected y' equal to -.25246. We shall correct our y' differences with (2a) and continue until values check. Proceeding in this manner we secure the next few lines of work:

x	y	Δy	$\Delta^2 y'$	$\Delta^3 y$	$\Delta^4 y$	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$	$\frac{x}{y}$	$-(\frac{y}{2}+1)$
-7.2	3.11496	-9721	-232	-16	-6	-.24606	-598	-38	-11	-3	-2.31142	-2.55748
-6.8	3.01527	-9969	-248	-16	0	-.25244	-638	-40	-2	9	-2.25519	-2.50763
-6.4	2.91293	-10234	-265	-17	1	-.25936	-692	-54	-14	-12	-2.19710	-2.45646
-6.0	2.80770	-10523	-289	-24	-7	-.26687	-751	-61	-7	7	-2.13698	-2.40385
-5.6	2.69933	-10837	-314	-25	-1	-.27507	-820	-69	-8	-1	-2.07459	-2.34966
-5.2	2.58752	-11181	-344	-30	-5	-.28412	-905	-85	-16	-8	-2.00964	-2.29376
-4.8	2.47190	-11562	-381	-37	-7	-.29413	-1001	-96	-11	5	-1.94182	-2.23595
-4.4	2.35205	-11985	-423	-42	-5	-.30532	-1119	-118	-22	-11	-1.87071	-2.17603
-4.0	2.22745	-12460	-475	-52	-10	-.31794	-1262	-143	-25	-3	-1.79578	-2.11372

2. The method of differences will apply equally well to equations of the second order. The computation is, of course, somewhat more extended. In working out this type of problem the solution is best arranged in the same manner as the first order equation with the addition of columns to accommodate y" and its differences. Presuming we have the solution started together with the necessary differences, we continue as before except that y" differences are attacked first. After we get y" and its differences, Gregory's Formula (1) gives a trial $\Delta y'$. Formulas (2) and (2a) give y' and the other necessary differences of y'. From this point on the work is identical to that of the first order equation. Of course each recheck must be carried completely through both rows of dif-

ferences.

3. There are one or two points that should be noted in addition to the ones of the two previous paragraphs. In the actual work the y differences play little part except as a check. Any large fluctuation in the fourth order differences is usually indicative of an error. That is, they should vary fairly regularly. The same may be said for the y' differences. Another caution to observe is the rapid increase in y and y'. When the function curve becomes rather steep, accuracy falls off. The only remedy for this is to shorten the interval. Of course the whole process fails at a point where a vertical tangent exists.

METHOD OF ORDINATES

When a calculating machine is available the method of ordinates is a more elegant type of solution then the relatively clumsy method of differences. Two forms* are exhibited in the next three paragraphs. Types of equations which may be integrated by this scheme are

$$\text{(a)}\quad \frac{dy}{dx} = u(x,y),$$

$$\text{(b)}\quad \frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = 0,$$

$$\text{(c)}\quad \frac{d^2y}{dx^2} = u(x,y),$$

where P and Q are constants or functions of x. The solution of (b) is merely an extension of the method used for (a). Of course the last mentioned type is a particular form of (b) but to eliminate needless work a special method is employed. Both types of solution are characterized by neatness in execution.

1. Let us consider again Newton's interpolation formula,

$$P_n(x) = u_0 + s\Delta u_0 + \frac{s(s-1)}{2!}\Delta^2 u_0 + \frac{s(s-1)(s-2)}{3!}\Delta^3 u_0 + \cdots\cdots$$

If this is integrated over intervals of 2h and 4h, put in the form of backward differences, and simplified, we secure the two formulas,

$$\int_{x_2}^{x_4} P_n(x)dx = \frac{h}{3}\left\{u_4 + 4u_3 + u_2\right\} - \frac{8h\,\Delta^4 u_0}{720} \quad + \text{ higher differences,}$$

$$\int_{x_0}^{x_4} P_n(x)dx = \frac{4h}{3}\left\{2u_3 - u_2 + 2u_1\right\} + \frac{224h\,\Delta^4 u_0}{720} \quad + \text{ higher differences.}$$

Reducing these to ordinate form and neglecting fourth and higher order differences we find

$$(3)\quad y_n = y_{n+1} + \frac{4h}{3}\left\{2y'_{n-1} - y'_{n-2} + 2y'_{n-3}\right\}$$

$$(4)\quad y_n = y_{n-2} + \frac{h}{3}\left\{y'_n + 4y'_{n-1} + y'_{n-2}\right\}$$

The first of these is for prediction while the second is for checking.

Suppose, for example, we have the equation, $\frac{dy}{dx} = xy$

* See Bibliography II(1) and II(2).

with the initial condition that y equals one when x equals zero. Further, suppose we have the first four values of y and y' which we might get by one of several methods discussed later. If we choose h equal .1 we get the following values:

x	y	ϵ	y'
.0	1.0000		.0000
.1	1.0050		.1005
.2	1.0202		.2040
.3	1.0460		.3138

Applying (3) we find a trial value of the next y to be 1.0830, with this value of y the equation is satisfied for y' equal to .4332. Checking with (4) we correct y to 1.0833. Substituted in the equation this gives y' equal to .4333. As further computation leaves these values unchanged we take them as correct. In similar manner we continue and arrive at the following results:

x	y	ϵ	y'
.4	1.0833	.0003	.4333
.5	1.1332	1	.5666
.6	1.1972	0	.7183
.7	1.2777	2	.8944
.8	1.3771	0	1.1017
.9	1.4994	1	1.3495
1.0	1.6487	0	1.6487

The column adjoining y represents the change in y due to recalculation. It provides a constant check on the accuracy of solution. The error due to the neglected fourth order differences in the use of (4) is approximately 1/29 of ϵ. Consequently if ϵ increases to the point where $\epsilon/29$ will effect the value of y we should choose a smaller interval.

2. The type of solution of the preceding paragraph may be extended to equations of the second order. The only change necessary is the additional calculation of the second derivative. Assuming we have four values of y, y', and y", the first approximation to y_4' is given by

$$(3a) \qquad y_4' = y_0' + \frac{4h}{3}\left\{2y_3'' - y_2'' + 2y_1''\right\}$$

This is of course formula (3) with a change of primes. Using this value of y', we approximate y with (3). The equation is then solved for y". We change (4) to

$$(4a) \qquad y_4' = y_2' + \frac{h}{3}\left\{y_4'' + 4y_3'' + y_2''\right\}$$

and check y'. Rechecking is continued until no corrections are necessary. The following illustrates a solution of this kind. Here the first four values were determined by successive approximations. The initial evaluation of coth x $\frac{dy}{dx}$ was by the usual method of differentiating numerator and denominator of $\frac{dy}{dx}$/tanh x separately. Our example is

$$\frac{d^2y}{dx^2} + \coth x \frac{dy}{dx} + \left[\cosh x - .225\cosh^2 x + .075\right]y = 0$$

with the conditions that x_0, y_0, y_0' are respectively equal to zero, one, and zero.

x	y	y'	y"	(.......)	cothx	(........)y
.2	.992	.085	-.423			
.0	1.000	.000	-.425	.850	-.425	.850
.2	.992	-.085	-.423	.861	-.431	.854
.4	.966	-.169	-.418	.893	-.445	.863
.6	.924	-.252	-.403	.944	-.469	.872
.8	.866	-.330	-.378	1.010	-.497	.875
1.0	.792	-.402	-.329	1.082	-.528	.857
1.2	.706	-.459	-.238	1.117	-.551	.789
1.4	.610	-.498	-.161	1.185	-.562	.723
1.6	.508	-.519	-.022	1.158	-.563	.588
1.8	.298	-.507	.133	1.010	-.535	.402
2.0	.307	-.466	.283	.653	-.483	.200
2.2	.214	-.393	.416	-.052	-.405	-.011
2.4	.150	-.303	.505	-1.316	-.308	-.197
2.6	.094	-.198	.526	-3.465	-.200	-.326
2.8	.071	-.090	.588	-6.996	-.091	-.497
3.0	.060	.041	.713	-12.664	.047	-.760

3.

Another equation of rather frequent occurence is (c); that is, the second order equation in which the first derivative is missing. Special methods have been utilized in its solution to avoid needless computation and inherent errors. The formulas used are similar to (3) and (4) in some respects but have been developed from combinations of integrals of Newton's interpolation formula over various intervals, grouped to eliminate desirable differences. In addition to this the terms corresponding to y' have been eliminated. The following formulas are used:

$$y_{n+1} = y_n + y_{n-2} - y_{n-3} + \frac{h^2}{4}\left\{5u_n + 2u_{n-1} + 5u_{n-2}\right\}, \tag{5}$$

$$y_n = 2y_{n-1} + 2y_{n-2} + \frac{h^2}{12}\left\{u_n + 10u_{n-1} + u_{n-2}\right\}. \tag{6}$$

The first of these is for integrating ahead and the second for checking. Where these fail to give desired accuracy a pair involving five ordinates may be used. They are

$$y_{n+1} = y_n + y_{n-4} - y_{n-5} + \frac{h^2}{48}\left\{67u_n - 8u_{n-1} + 122u_{n-2} - 8u_{n-3} + 67u_{n-4}\right\},$$

$$y_n = y_{n-1} + y_{n-3} - y_{n-4} + \frac{h^2}{240}\left\{17u_n + 232u_{n-1} + 222u_{n-2} + 232u_{n-3} + 17u_{n-4}\right\}.$$

The solution following is obtained by the three-ordinate formulas. Here the first six values were computed by a series. Our equation is

$$\frac{d^2y}{dx^2} + \operatorname{sech}^2 x\left[-.072 + .216\tanh^2 x\right]y = 0,$$

with the initial conditions that x_0, y_0, and y_0' are respectively equal to zero, one, and zero.

x	y	y"
.0	1.000	.072
.2	1.001	.062
.4	1.005	.035
.6	1.010	.007
.8	1.016	-.013
1.0	1.021	-.023
1.2	1.025	-.024
1.4	1.028	-.022
1.6	1.031	-.017
1.8	1.032	-.013
2.0	1.033	-.009

2.2	1.034	-.007
2.4	1.035	-.005
2.6	1.035	-.003
2.8	1.035	-.003
3.0	1.036	-.002
3.2	1.036	-.001
3.4	1.036	-.001
3.6	1.036	+.000

METHODS OF STARTING

Practically every method of integrating numerically is dependent upon formulas which call for a few starting values of the variables. Consequently special devises must be pursued to start computation.

1. One practice commonly used is that of series. The first necessary values of y are computed by a Taylor's Series:

$$y = y_0 + y'(x-x_0) + \frac{y''(x-x_0)^2}{2!} + \frac{y'''(x-x_0)^3}{3!} + \frac{y^{IV}(x-x_0)^4}{4!} + \dots$$

The series usually converges very rapidly after two or three terms and gives a fairly good approximation for starting values. This method will fail when the function is infinite and in certain problems becomes rather difficult to apply. In cases where x equals zero, the series becomes even simpler, reducing to a Maclaurin Series. This is probably the most satisfactory method of starting when the equation is not too involved.

2. Another mode of starting which can be used quite readily in problems solved by differences is merely an extension of this method to starting. That is, instead of assuming that fourth order differences are negligible we start by letting first order differences be dropped. From this we get the first estimated y and we then calculate the first y' approximation. This gives our first difference for the second row of work. This is recorded and y is rechecked with Gregory's Formula (1) applied to y' and $\Delta y'$. Several recalculations are usually necessary especially at the start but there is the advantage of simplicity in principle to recommend this way of starting. At each step we add, of course, one more difference so that after four or five rows of work we have all the necessary orders to continue and the process has tightened down to much less labor.

Let us look in some detail to an example of this method of starting. A simple problem to answer our purpose is

$$\frac{dy}{dx} = xy \tag{7}$$

with initial conditions that y equals two when x equals zero. If we take h equal to .1 and carry the work to four decimal places we may start the table:

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$
0.0	2.0000					0.0000				
0.1	2.0000					0.2000	.20000			

The first value of y' is derived directly from the equation (7). If we bluntly assume that y' differences are negligible, y is also equal to two and is recorded above. This value of y substituted in (7) gives y' equal to .2000 which in turn provids our first $\Delta y'$ also equal to .2000.

As we now have a y' and a first difference we may apply Gregory's Formula (1) to secure Δy and check y . This gives Δy equal to .0100. It is obvious that this Δy necessitates a change in y which calls for a correction of y' and $\Delta y'$. Making these necessary changes and rechecking until values are unchanged, we write out the table:

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	Δy	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$
.0	2.0000					.0000				
.1	2.0000	.0100				.2000	.2000			
	(2.0100)	(.0100				(.2010)	(.2010)			

The number in parentheses is the corrected value which replaces the number above. Proceeding in this manner we get the following table where only corrected values appear:

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	Δy	y'	$\Delta y'$	$\Delta^2 y'$	$\Delta^3 y'$	$\Delta^4 y'$
.0	2.0000					.0000				
.1	2.0100	.0100				.2010	.2010			
.2	2.0404	.0304	.0204			.4081	.2071	.0061		
.3	2.0921	.0517	.0213	.0009		.6276	.2195	.0124	.0063	
.4	2.1665	.0744	.0227	.0014	.0005	.8666	.2395	.0200	.0076	.0013

The computation has now reached the stage where it runs along rather smoothly. Differences of a higher order than the fourth are in general of little value to increase accuracy.

A slight modification of this method which works out very well with certain types of equations is that of working backward as well as forward from our initial values. The work is practically the same except we have a more complete set of differences for the fore part of the problem.

3. Another manner of commencing numerical integration which is often satisfactory is that of successive approximations. It gives us four values with which to continue. Here we assume

$$y_{-1} = y_0 - hy'_0,$$

$$y_1 = y_0 + hy'_0,$$

and

$$y_2 = y_0 + 2hy'_0$$

From the differential equation we find values (trial) of y'_{-1}, y'_1, and y'_0 using the y values of these three formulas. The next step is the checking of y_{-1}, y_1, and y_2 with

$$y_{-1} = y_1 - \frac{h}{3}\left\{ y'_1 + 4y'_0 + y'_{-1} \right\},$$

$$y_{-1} = y_0 + \frac{h}{24}\left\{ -y'_2 + 13y'_1 + 13y'_0 - y'_{-1} \right\},$$

$$y_2 = y_0 + \frac{h}{3}\left\{ y'_2 + 4y'_1 + y'_0 \right\}$$

We again turn to the differential equation rechecking for y'_{-1} , y'_1 , and y'_2 , and continue until both y and y' values remain unchanged.

CHOICE OF INTERVAL

The choice of interval in numerical integration must be guided by two factors. The desired degree of accuracy is, of course, the item of prime importance. Hand in hand with this must be considered the nature of the function itself in various parts of the range of integration. If the function has a gradually increasing derivative, it is quite possible that we may need a shorter interval to maintain accuracy. On the other hand, if the derivative decreases and remains small, work may be reduced by increasing the size fo the interval. We shall note briefly simple ways of changing h to suit our needs in the next two paragraphs.

1. Let us suppose for the moment that we have a problem under way, and further, that y' has decreased to such an extent that we wish to increase h. This is done by doubling and if necessary redoubling h. To accomplish this when using differences the following simple scheme may be used. Taking only every other value of y and y' we may make a new set of differences which will fit an interval of 2h. We are now ready to continue the solution with the new interval. This may be repeated to further enlarge the interval.

2. In the matter of shortening the interval it is often convenient to recalculate starting values. However, the various values may be secured in this way. Taking Newton's interpolation formula in the symbolic form:

$$P_n(x) = (1+\Delta)^t u_0 \quad \text{where} \quad t = \frac{(x - x_0)}{h}$$

Now if we let Δ_1 and Δ_2 be the difference symbols for h and 2h respectively, we have

$$(1 + \Delta_1)^{2t} u_0 = (1 + \Delta_2)^t u_0$$

That is,

$$(1 + \Delta_1)^2 = 1 + \Delta_2$$

From this quadratic,

$$\Delta_1 = -1 + \sqrt{1 + \Delta_2}$$

Continuing, we may compute by the binomial theorem

$$\Delta_1 = \frac{1}{2}\Delta_2 - \frac{1}{8}\Delta_2^2 + \frac{1}{16}\Delta_2^3 + - - - - ,$$

$$\Delta_1^2 = \frac{1}{4}\Delta_2^2 - \frac{1}{8}\Delta_2^3 + \frac{5}{64}\Delta_2^4 - - - - - ,$$

$$\Delta_1^3 = \frac{1}{8}\Delta_2^3 - \frac{3}{32}\Delta_2^4 - - - - - - - ,$$

$$\Delta_1^4 = \frac{1}{16}\Delta_2^4 - \frac{1}{16}\Delta_2^5 - - - - - .$$

These latter formulas may be used for halving h. It is to be expected when a change is made in h that we very probably will need several recheckings before our computation is running smooth again.

When the method of ordinates is used, greater accuracy is achieved by shifting from the three to the five ordinate formulas. When h must still be shortened in spite of this change, it is usually more convenient to compute new starting values by series or successive approximations.

NUMERICAL INTEGRATION OF THE THIRD ORDER EQUATION IN WHICH THE FIRST AND SECOND DERIVATIVES ARE ABSENT

The third order equation of the form,

$$\frac{d^3 y}{dx^3} = u(x, y)$$

requires considerable work in its numerical integration unless special methods are employed which provide for the omission of the calculation of the first and second derivatives. Another drawback to the usual methods of integration is the likelihood of magnifying errors through additional computation. An advantage in favor of the solution herein discussed is the rapidity with which it converges to the required degree of accuracy, and this necessitates a minimim of recalculation.

The principle involved in this type of approximation is rather simple. It consists in the main of setting up a polynomial which is equal at regular intervals to the equation. The true solution of the equation is secured by integrating three times and hence the polynomial is integrated three times. The accuracy of the approximation is, of course, dependent on a number of factors, more important of which are the number of points of equality and the distance between them.

1. As a basis we shall use Newton's interpolation formula,

$$P_n(x) = u_0 + s\Delta u_0 + \frac{s(s-1)}{2!}\Delta^2 u_0 + \frac{s(s-1)(s-2)}{3!}\Delta^3 u_0 + \cdots$$

For convenience this must be put in the form for backward differences and integrated three times. If we define a set of functions $A_i(s)$, where i = 0, 1, 2, etc., by the equations,

$$A_i(s) = (-1)^i \int_0^s ds \int_0^s ds \int_0^s \frac{s(s+1)(s+2)\cdots(s+i-1)}{i!}\, ds,$$

we shall reduce the work of integration to a minimum.

In this notation,

$$s = (x-x_n)/h,$$

$$h = x_n - x_{n-1},$$

$$dx = hds.$$

The general expression for the approximating integral is

$$(8) \qquad \int_{x_n}^{x} dx \int_{x_n}^{x} dx \int_{x_n}^{x} u\,dx = h^3 \sum_i A_i(s)\, \Delta^i u_n .$$

The following table gives a few of the values of $A_i(s)$:

s	$A_0(s)$	$A_1(s)$	$A_2(s)$	$A_3(s)$	$A_4(s)$	$A_5(s)$
+1	+1	-1	7	-17	205	-731
0	0	0	0	0	0	0
-1	-1	-1	3	-5	47	-139
-2	-8	-16	16	-32	304	-896
-3	-27	-81	-81	-81	729	-2187
-4	-64	-256	-768	-512	1280	-4096
-5	-125	-625	-3125	-3125	-3125	-6875
common denominator	6	24	240	720	10080	40320

Let us now turn our attention to the differential equation:

$$\frac{d^3y}{dx^3} = u(x, y).$$

Integrating the first time,

$$\frac{d^2y}{dx^2} = \int_{x_n}^{x} u(x,y)\,dx + C_1 .$$

Setting $x = x_n$ gives $C_1 = y_n''$.

Our equation now becomes

$$\frac{d^2y}{dx^2} = \int_{x_n}^{x} u(x,y)\,dx + y_n''.$$

Integrating again and solving for C_2 we find

$$\frac{dy}{dx} = y_n' + y_n''(x - x_n) + \int_{x_n}^{x} dx \int_{x_n}^{x} u(x,y)\,dx .$$

In similar manner we arrive at the general integral equation,

$$y = y_n + y_n'(x - x_n) + \frac{y_n}{2}(x - x_n)^2 + \int_{x_n}^{x} dx \int_{x_n}^{x} dx \int_{x_n}^{x} u(x,y)\,dx .$$

If we now replace the triple integral by its approximate value from (8);

$$y = y_n + y_n'(x - x_n) + \frac{y_n''}{2}(x - x_n)^2 + h^3 \sum_i A_i(s)\, \Delta^i u_n .$$

Let us set $\qquad x = x_{n-k}$

and our equation becomes

$$y_{n-k} = y_n + y_n'(x_{n-k} - x_n) + \frac{y''}{2}(x_{n-k} - x_n)^2 + h^3 \sum_i A_i(-k)\Delta^i u_n .$$

For k = -1, 0, 1, 2, 3, 4, and 5, we get a number of useful formulas:

$$(9) \quad y_{n+1} = y_n + hy_n' + \tfrac{1}{2}h^2 y_n'' + h^3 \sum A_i(+1)\Delta^i u_n,$$

$$(10) \quad y_n = y_n,$$

$$(11) \quad y_{n-1} = y_n - hy_n' + \tfrac{1}{2}h^2 y_n'' + h^3 \sum A_i(-1)\Delta^i u_n,$$

$$(12) \quad y_{n-2} = y_n - 2hy_n' + 2h^2 y_n'' + h^3 \sum A_i(-2)\Delta^i u_n,$$

$$(13) \quad y_{n-3} = y_n - 3hy_n' + \tfrac{9}{2}h^2 y_n'' + h^3 \sum A_i(-3)\Delta^i u_n,$$

$$(14) \quad y_{n-4} = y_n - 4hy_n' + 8h^2 y_n'' + h^3 \sum A_i(-4)\Delta^i u_n,$$

$$(15) \quad y_{n-5} = y_n - 5hy_n' + \tfrac{25}{2}h^2 y_n'' + h^3 \sum A_i(-5)\Delta^i u_n.$$

These equations may be combined in various ways to eliminate y' and y". Let us choose these so as to rid the series on the right of designated differences. In this case we select third order differences to vanish.

If we select (9) + 3(11) - (12) we find

$$(16) \quad y_{n+1} = 3y_n - 3y_{n-1} + y_{n-2} + h^3\left[1 + \tfrac{1}{2}\Delta + \tfrac{1}{240}\Delta^4 - \tfrac{1}{160}\Delta^5 \cdots\right]u_n .$$

Again choosing -3(11) + 3(12) -(13) gives us

$$(17) \quad y_n = 3y_{n-1} - 3y_{n-2} + y_{n-3} + h^3\left[1 + \tfrac{3}{2}\Delta + \tfrac{1}{2}\Delta^2 + \tfrac{1}{240}\Delta^4 - \tfrac{1}{480}\Delta^5 \cdots\right]u_n$$

These two equations give us a formula for integrating ahead, (16), and a recalculation formula (17). Others may be found such as (9) + 2(12)

-(13), both of which give rise to prediction formulas. The first is undesirable because its large coefficients will introduce large errors. The latter is probably more accurate than (16) but as we have no suitable recalculation formulas to pair with it at present we shall pass it by.

Returning to (16) and (17), let us put them in ordinate form. Thus we write

$$(18) \quad y_{n+1} = 3y_n - 3y_{n-1} + y_{n-2} + \frac{h^3}{2}\left\{u_n + u_{n-1}\right\} + \frac{h^3}{240}\Delta^4 u_n + \text{higher differences.}$$

$$(19) \quad y_n = 3y_{n-1} - 3y_{n-2} + y_{n-3} + \frac{h^3}{2}\left\{u_{n-1} + u_{n-2}\right\} + \frac{h^3}{240}\Delta^4 u_n + \text{higher differences.}$$

The remarkable feature of these two formulas is, of course, their almost identical form. When terms on the right of the parenthesis are neglected as they are in practice, they do become identical.

Let us look again at equations (9) to (15) and see what they offer in the way of more accurate five ordinate formulas. If we group 2(9) + 3(14) - 2(15) and -2(11) + 2(13) - (14) we find

$$(20) \quad 2y_{n+1} = 3y_n - 3y_{n-4} + 2y_{n-5} + h^3\Big[10 + 20\Delta + \frac{33}{2}\Delta^2 + \frac{13}{2}\Delta^3 + \frac{25}{4}\Delta^4 + \frac{463}{60480}\Delta^6 \cdots\cdots\Big]u_n,$$

$$(21) \quad y_n = 2y_{n-1} - 2y_{n-3} + y_{n-4} + h^3\Big[2 + 4\Delta + \frac{5}{2}\Delta^2 + \frac{1}{2}\Delta^3 + \frac{1}{120}\Delta^4 + \frac{1}{30240}\Delta^6 \cdots\cdots\Big]u_n,$$

in which fifth order differences are lacking. As before recasting these in ordinate form we derive

$$(22) \quad 2y_{n+1} = 3y_n - 3y_{n-4} + 2y_{n-5} + \frac{h^3}{24}\Big\{25u_n + 56u_{n-1} + 78u_{n-3} + 56u_{n-4} + 25u_{n-5}\Big\} + \frac{463}{60480}\Delta^6 u_n + \text{higher differences.}$$

$$(23) \quad y_n = 2y_{n-1} - 2y_{n-3} + y_{n-4} + \frac{h^3}{120}\left\{u_n + 56u_{n-1} + 126u_{n-2} + 56u_{n-3} + u_{n-4}\right\} + \frac{h^3}{30240}\Delta^6 u_0 + \text{higher differences.}$$

These two provide a rather accurate means of numerical integration. As in the case of (5) and (6) the recalculation formula (23) is in general more accurate then (22). Dispensing with any proof of this point at present, this fact is borne out roughly by the relative sizes of the coefficients in the first neglected terms. As may be noted later the largest error in the application of (22) and (23) is in the necessity of terminating our figures at a practical number of decimal places.

2. Let us turn our attention to the solution of an example by the various formulas and see how well they adapt themselves to practical work. For the benefit of comparison let us integrate the same equation with both sets of formulas. If we choose an equation of such a nature that we may secure its analytic solution we have a valuable check. For our equation then let us select

$$(24) \qquad \frac{d^3y}{dx^3} = y$$

with the initial conditions: $x_0 = 0$; $y_0 = 1$; $y_0' = 0$; $y_0'' = 1$.

Choosing $h = .1$ we proceed to calculate a few starting values by Maclaurin's series. We record them in the table:

x	y	u
0.	1.0000000	1.0000000
.1	1.0051668	1.0051668
.2	1.0213361	1.0213361

These are sufficient for (18) and 19). Predicting with (18) we find for $x = .3$, $y_3 = 1.0495213$. This is entered inboth the y and u columns as in the particular problem used y equals u. There is no checking formula to apply in this case as we recall that (18) and (19) are identical. Continuing in this way we find

x	y	C
0.	1.0000000	
.1	1.0051668	
.2	1.0213361	

x	y	C
.3	1.0495213	
.4	1.0907578	
.5	1.1461157	
.6	1.2167134	
.7	1.3037323	
.8	1.4084326	
.9	1.5321704	
1.0	1.6764160	1.6764164
1.1	1.8427737	
1.2	2.0330031	
1.3	2.2490421	
1.4	2.4930317	
1.5	2.7673429	
1.6	3.0746059	
1.7	3.4177417	
1.8	3.7999964	
1.9	4.2249789	
2.0	4.6967017	4.6967091

Here we have omitted the u column as needless repetition. The numbers in the C column represent the analytic solution for the values of x which they follow.

To integrate (24) with (22) and (23′ we shall need six values of y. Computing them in the same way as before we find them to be

x	y
0.	1.0000000
.1	1.0051668

x	y
.2	1.0213361
.3	1.0495213
.4	1.0907577
.5	1.1461156

Applying (22), we find y_0 equal to 1.2167132. Checking with (23) we find it should be 1.2167133. Besides recording this it is well to note the correction in y in a separate column. Rechecking leaves this latter value unchanged so we proceed to y. The results of the integration from zero to two are given. The column headed δy lists the correction in y in rechecking with (23) while C has the same function as before.

x	y	δy	C
0.	1.0000000		
.1	1.0051668		
.2	1.0213361		
.3	1.0907577		
.4	1.0907577		
.5	1.1461156		
.6	1.2167133	.0000001	
.7	1.3037323	0	
.8	1.4084327	0	
.9	1.5321707	-1	
1.0	1.6764165	0	1.6764164
1.1	1.8427745	0	
1.2	2.0330042	0	
1.3	2.2490436	0	
1.4	2.4930336	0	

x	y	δy	C
1.5	2.7673454	.0000001	
1.6	3.0746090	0	
1.7	3.4177456	0	
1.8	3.8000012	1	
1.9	4.2249849	1	
2.0	4.6967090	1	4.6967091

In this computation δy gives the correction of (23) on (22). In general this is rather small and recheck is seldom necessary. The accuracy may be verified in column C at two points as in the previous example.

ACCURACY OF THE PROCESS

Let us consider the following function:

$$(25)\quad Q(x) = u(x) - P_n(x) - R(x-x_0)(x-x_1)\cdots\cdots(x-x_n),$$

where $u(x)$ and $P_n(x)$ are respectively the values of the right hand member of our differential equation and the approximating polynomial. Then $R(x-x_0)(x-x_1)(x-x_2)\cdots\cdots(x-x_n)$ gives us the remainder or error of $P_n(x)$ in approximating $u(x)$ at any point in the interval of integration. Now $Q(x)$ must vanish at n+1 points but we may choose R, an arbitrary constant, so that $Q(x)$ will vanish at some one other point not coinciding with $x_0, x_1, \cdots\cdots x_n$.

Then $Q(x)$ has at least n+2 zeros. Now if we assume $u(x)$ possesses continuous derivatives up to and including the (n+1)st, we may differentiate and find

$$Q'(x) = u'(x) - P_n'(x) - \frac{d}{dx}\left[R(x-x_0)(x-x_1)\cdots(x-x_n)\right],$$

$$Q''(x) = u''(x) - P_n''(x) - \frac{d^2}{dx^2}\left[R(x-x_0)(x-x_1)\cdots(x-x_n)\right].$$

By Rolle's Theorem, $Q'(x)$ has at least n+1 zeros and the number of zeros decreases by unity at each differentiation. If we differentiate n+1 times,

$$Q^{(n+1)}(x) = u^{(n+1)}(x) - R(n+1)!$$

As $P_n(x)$ is a polynomial of degree n, it vanishes in this expression. Now $Q^{(n+1)}(x)$ has as a minimum one zero at, let us say, z. Therefore

$$R = \frac{u^{(n+1)}(z)}{(n+1)!} \qquad x_0 \leqq z \leqq x_n$$

If we substitute in (25), we find

$$u(\bar{x}) = P_n(\bar{x}) + \frac{u^{n+1}(z)}{(n+1)!}(\bar{x}-x_0)(\bar{x}-x_1)\cdots(\bar{x}-x_n)$$

In general, this gives us, as z is a value of $\bar{x}$ dependent on an arbitrary R,

$$u(x) = P_n(x) + \frac{u^{(n+1)}(x)}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n).$$

Setting $x-x_n = hs$, we arrive at

$$u(x) = P_n(x) + \frac{u^{(n+1)}(x)h^{n+1}}{(n+1)!} s(s+1)(s+2) \cdots (s+n).$$

Let us designate

$$\epsilon = u(x) - P_n(x)$$

and we may write

$$(26) \qquad \epsilon = u^{(n+1)}(x) h^{n+1} s(s+1)(s+2) \cdots (s+n).$$

We must recall that ϵ is the error in one of the formulas (9) to (15). Formulas (18), (19), (22), and (23) are various combinations of these. The error of any of these latter formulas may be secured by integrating (26) over the correct intervals. Final results are given by (18) and (23) so we shall examine them in some detail. The general expression for the error of (18) is given by

$$E = \sum_k h^3 \int ds \int ds \int \frac{u^{(4)}(x) h^4}{4!} s(s+1)(s+2)(s+3) ds,$$

where k gives the sum of the absolute values of each integral over separate intervals of integration. If we choose M_3 and N_3 respectively to denote the absolute maximum values of $u^{(4)}(x)$ and the triple integral, we see

$$(27) \qquad E_3 = \frac{h^7 M_3 N_3}{24}$$

Again choosing n = 5 and allowing M and N to be analogous items, we have the corresponding error of (23) to be

$$(28) \qquad E_5 = \frac{h^9 M_5 N_5}{720}.$$

It is evident that these are only errors over particular in intervals. As the error of either (18) or (23) is carried over into the coefficients of the formulas we see that it grows from interval to interval. If we consider only the maximum value of the error over one interval we may conveniently represent its growth by the series, for (18)

1, 4, 10, 20, 35, 56, 84, 120, ...

where the terms are the respective coefficients of E throughout the ran range of integration. This is a recurring series of the third order whose generating function is $(n^3 + 6n^2 + 11n + 6)/6$. Denoting the range of integration by r, we have

$$r = nh,$$

$$n = r/h.$$

The generating function is now by substitution

$$(r^3 + 6r^2h + 11rh^2 + 6h^3)/ 6h^3.$$

For an arbitrarily small h this may be made to approach

$$r^3/6h^3,$$

And the error for any particular range, i, is

$$\frac{r^3}{6h^3} E_3.$$

We may now write the untrue equation

$$E_{18} = \frac{h^7 M_3 N_3}{4!}$$

In final form this becomes

(29) $$E_{18} = \frac{65}{6048} M_3 r^3 h^4,$$

where h, r, and M_3 are dependent on the particular problem.

In a similar manner the error growth of (23) over a succession of intervals may be shown to be represented by the series,

1, 3, 7, 13, 22, 34, 50, 70, 95, 125, 161, 203,........

This responds rather nicely to treatment if we break it up into two series

1, 7, 22, 50, 95, 161,............

3, 13, 34, 70, 125, 203,..........

whose generating functions are $(4n^3 + 15n^2 + 17n + 6)/6$ and $(4n^3 + 21n^2 + 35n + 18)/6$ respectively. As before n equals r/h, and we see that either of these functions for an arbitrarily small h may be made to approach

$$\frac{2r^3}{3h^3}.$$

Finally, we have from this expression and (28)

$$E_{23} = \frac{2r^3}{3h^3}\left(\frac{51749}{302400}\right) M_5 h^9,$$

where the fraction in parenthesis gives the value of N_5. This on reduction and substitution of an approximate value for the large fraction becomes

$$E_{23} = \frac{7}{60} r^3 h^6 M_5. \tag{30}$$

It should be noted here that neither (29) or (30) give the exact error. They have been built upon the assumption of cumulative error on either the positive or negative side. We are interested primarily in the fact that the expression for the error involves h to a power. This fact guarantees preassigned accuracy for a suitable small h. In general it will be noted that the error introduced by dropping decimal places in the simpler equations may account for larger errors than those introduced by the approximating process.

BIBLIOGRAPHY

I. Books

1. Scarbourough, J. B. — Numerical Mathematical Analysis

2. Milne, W. E. — Numerical Calculus

3. Steffensen, J. F. — Interpolation

4. Moulton, F. R. — Differential Equations

5. Taylor, J. M. — College Algebra

II. Papers, Publications, etc.

1. Milne, W. E. — Numerical Integration of Ordinary Differential Equations-- American Mathematical Monthly 1926

2. Milne, W. E. — On Numerical Integration of Certain Differential Equations of the Second Order--- American Mathematical Monthly 1933