REPORT No.4

February, 1947

THE COLLEGE OF AERONAUTICS

CRANFIELD

Technique of the Step-by-Step Integration of Ordinary Differential Equations

-by-

W.J. Duncan, D.Sc., M.I.Mech.E., F.R.Ae.S.,
Professor of Aerodynamics at the
College of Aeronautics, Cranfield

------oOo------

-SUMMARY-

In Part I step-by-step methods are examined critically and emphasis is placed on the dependence of the error on the number n of steps used for a given range of the independent variable. The index of a process is defined and it is shown that the errors can be assessed and partially corrected when the index is known and results obtained for two or more values of n. Attention is drawn to the advantages in certain cases of a part-analytical process.

In Part II methods of numerical integration in general are classified and briefly reviewed. The chart, Table 2.2.1, summarises the classification.

WJD/DMS

/Contents.....

- CONTENTS -

/0. Introduction..

0. Introduction

Processes for the step-by-step integration of ordinary differential equations are amongst the most valuable of the methods for the numerical solution of such equations with one-point boundary conditions. Some of the advantages of these processes are:-

(a) They are universally applicable to ordinary differential equations of any order and to simultaneous sets of such equations, provided only that there are no singularities in the range of integration.

(b) Results of any required degree of accuracy can be obtained.

(c) Some of the most generally useful of the processes are intuitively obvious. They are therefore easy to learn and to remember.

In the present paper emphasis is placed on the simpler techniques of step-by-step integration and only brief reference is made to those depending on the use of finite differences, such as the Adams-Bashforth process. It is stressed that the error in the approximate solution for a given range of the independent variable is a function of that variable and of the number n of equal steps used. Provided that exactly the same process is used throughout and the same numerical accuracy maintained at all times, it becomes possible to estimate the error when the calculation is made with two or more values of n ; it even becomes possible to improve the solution obtained with the greatest number of steps. Accordingly it is suggested that it should be a normal procedure in serious work to make the computation with two different numbers of intervals; if it is anticipated that $2m$ steps will give the required accuracy a preliminary calculation with say m steps should be made.

In Part I the technique of step-by-step integration of ordinary differential equations or sets of such equations is considered in a general way and the index of the process is defined. Part II contains a brief review of some other methods for the numerical solution of ordinary differential equations. It is concluded that with the development of digital calculating machines of great power and rapidity the numerical methods of integration will oust all others; including notably those depending on physical or kinematical analogies (differential analysers etc).

/1.1 .....

Part I

Step-by-Step Integration

1.1 Reduction of Any Equation or Set to a First Order Set

It is well known that an ordinary differential equation of any order or a simultaneous set of such equations can always be reduced to a set of first order equations by the use of auxiliary variables. The auxiliary variables can be chosen in an indefinitely great number of ways, but the most obvious choice consists of all the differential coefficients of the dependent variables, excepting only the one of highest order for each variable. For example, the pair of second order equations

$$\left.\begin{aligned} \frac{d^2 x_1}{dt^2} &= f_1\left(t, x_1, x_2, \frac{dx_1}{dt}, \frac{dx_2}{dt}\right) \\ \frac{d^2 x_2}{dt^2} &= f_2\left(t, x_1, x_2, \frac{dx_1}{dt}, \frac{dx_2}{dt}\right) \end{aligned}\right\} \quad \ldots\ldots\ldots (1.1:1)$$

can be replaced by the following set of four first order equations:-

$$\left.\begin{aligned} \frac{dx_1}{dt} &= x_3 \\ \frac{dx_2}{dt} &= x_4 \\ \frac{dx_3}{dt} &= f_1\left(t, x_1, x_2, x_3, x_4\right) \\ \frac{dx_4}{dt} &= f_2\left(t, x_1, x_2, x_3, x_4\right). \end{aligned}\right\} \quad \ldots\ldots\ldots (1.1:2)$$

In the remainder of this paper, it will be supposed that the equation or equations under consideration have been reduced to a first order set.

1.2 Formulation of the Problem and Outline of Method

At present we confine attention to one-point boundary problems. The independent variable is t and the dependent variables are $x_1, x_2, \ldots\ldots x_m$.

/The.....

The data are the differential equations themselves and the values for t equal to a of each dependent variable or information from which these initial values can be found. It is required to construct an approximate numerical solution for the range a to b of t.

We shall suppose that the general march of the solution within the range is known independently, for example from physical intuition, or from a rough preliminary exploration, that singularities are absent and all the differential coefficient finite. Valuable hints about the preliminary exploration and suggestions for overcoming difficulties are given in the treatise by Runge[1]. As an example, an infinite differential coefficient can be avoided by taking a new independent variable, which is a linear combination of the original independent and dependent variables, throughout the whole or a part of the range of integration.

We divide the range a to b of t into n equal parts h, so

$$h = \frac{b - a}{n}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1.2:1)$$

and write

$$\left.\begin{aligned} t_0 &= a \\ t_r &= a + rh = \frac{(n - r)\,a + r\,b}{n}. \end{aligned}\right\} \ldots\ldots (1.2:2)$$

The approximate value of a dependent variable x_j corresponding to t_r will be written $x_j(r)$.

The first step in the solution consists in finding $x_j(1)$ $(j = 1, 2, \ldots\ldots m)$. To do this we calculate, by a process characteristic of the particular method adopted, a mean slope $s(j, 1)$ for the interval; in this symbol the first number corresponds to the dependent variable and the second to the interval. Then

$$x_j(1) = x_j(0) + h\,s(j, 1). \ldots\ldots\ldots\ldots (1.2:3)$$

The second and further steps are carried out in the same manner.
Thus

$$x_j(r+1) = x_j(r) + h\,s(j, r+1). \ldots\ldots (1.2:4)$$

We regard it as part of the definition of a step-by-step method that $s(j, r+1)$ is derived from the values $x_j(r)$ by <u>exactly</u> the same process as $s(j, 1)$ is

/derived

derived from the values x_j (o). Without this proviso the process of solution would lack uniformity and the difficulty of the discussion would be increased.

In the original method given by Euler in the 18th century s (j, r) is simply $\frac{dx_j}{dt}$ at the beginning of the rth interval as given by the differential equation in terms of t_{r-1} and the already calculated corresponding values of the dependent variables. This method, although still much used on account of its simplicity, is relatively crude and various more refined methods have been devised.

1.3 Definition of the Index of a Step-by-Step Process

For the present we shall regard the following as fixed:-

(a) The differential equation or equations
(b) The initial conditions
(c) The range of the independent variable
(d) The uniform process of step-by-step integration.

The only variable left is the number n of equal steps used. We propose to study the manner in which the errors in the approximate solution at the end of the range of integration (t = b) depend on n .

In the first place we shall suppose that the actual computations have been carried out with perfect accuracy, so that the only errors are those inherent in the process adopted. Consequently the errors are, in the conditions postulated, functions of the number n only. In the simple process of Euler (see § 1.2) only the first two terms in the Taylor expansion are retained. Hence the error in each interval will be an infinite power series in h beginning with a term in h^2 or, what amounts to the same thing, an infinite power series in n^{-1} beginning with a term in n^{-2}. Evidently here the error when t = b at the end of n intervals can be expanded as an infinite power series in n^{-1} beginning with the first power, for the leading term will be the sum of n terms each proportional to n^{-2}. We shall accordingly say that the Euler process has the index 1. For other processes the lowest power may be higher and we define the index of a process to be the index of the lowest power of n^{-1} which appears in the expressions for the errors at the end of the total interval of integration when these are expanded as infinite power series in the reciprocal of n.

The process of Runge and Kutta[6,7,1,2] has the index 4, but is not on that account to be assumed to be necessarily superior to processes having lower indices; the relative merits of processes are briefly considered in §1.7.

A very simple and useful process of index 2 is as follows. Use the intital slopes for the (r+1)th interval as given by the differential equations and the already computed values of the dependent variables at t_r to compute approximate values of the dependent

/variables...

variables for

$$t_{r+\frac{1}{2}} = t_r + \frac{h}{2} ,$$

i.e., at the middle of the $(r+1)$th interval. Then take the slope $s(j, r+1)$ to be the value of $\frac{dx_j}{dt}$ as computed from the differential equation for the foregoing mean values of t and of the dependent variables. Examples illustrating this process and that of Euler are given in § 1.5.

1.4 Use of the Index in Assessment of Errors and Partial Correction

In accordance with §1.3 we shall assume that the error $\epsilon_r(t)$ in the value of $x_r(t)$ can be expanded as

$$\epsilon_r(t) = n^{-k}\left(e_0 + e_1 n^{-1} + e_2 n^{-2} + \text{etc.}\right) \quad \ldots(1.4:1)$$

where k is the index of the process, and will be supposed known. Then, provided the interval h is sufficiently small, the first term in (1.4:1) will be dominant. Hence we shall have approximately

$$\epsilon_r(t) = e_0 n^{-k} . \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1.4:2)$$

Now let the process be carried out with two different numbers n_1 and n_2 of intervals and let the corresponding approximations to $x_r(t)$ be $x_{r1}(t)$ and $x_{r2}(t)$. Then we shall have with close approximation

$$\left.\begin{aligned} x_r(t) &= x_{r1}(t) + e_0 n_1^{-k} \\ &= x_{r2}(t) + e_0 n_2^{-k} . \end{aligned}\right\} \quad \ldots\ldots\ldots(1.4:3)$$

$$\text{Hence } e_0 = \frac{x_{r2}(t) - x_{r1}(t)}{n_1^{-k} - n_2^{-k}} \quad \ldots\ldots\ldots\ldots(1.4:4)$$

from which e_0 can be computed since all the terms on the right hand side of the equation are known. We can then calculate $x_r(t)$ from (1.4:3) and so obtain an improved approximation to $x_r(t)$.

/The.....

The index of the corrected value will be raised by 1.

If three different numbers n_1 , n_2 and n_3 of intervals be used it becomes possible to calculate both e_o and e_1 if it be assumed that the higher terms in the ^{1} expansion (1.4:1) can be neglected. When this is done a very helpful comparison can be made between the approximations to $x_r(t)$ obtained by taking all possible pairs and finally all three together.

The following procedure is suggested in the practical use of this method. First, estimate (or guess) the number n_1 of intervals required to give an approximate solution of the required accuracy. Begin by taking a smaller number n_2 of intervals, say half the number, and carry through the process. Repeat the process using n_1 intervals and compute the corrections as explained above; quite apart from the use of the corrections, the repetition of the process with a different number of intervals is a most valuable check. If it appears that the errors in the corrected results are likely to be too large, the process must be carried out a third time with a number n_3 of intervals greater than n_1 and final corrections applied.

The regularity in the computed results required for the application of the process of correction will be somewhat impaired by accidental errors of calculation. This will not give rise to trouble provided that a suitable degree of numerical accuracy is maintained throughout the work as is sufficiently demonstrated in the examples given in §1.5.

1.5 Illustrative Examples

Examples of great simplicity have been selected but they suffice to illustrate the methods already described.

Example (a)

The differential equation is

$$\frac{dx}{dt} = 1 + 0.2\,t - 0.5\,x \qquad \ldots\ldots\ldots\ldots (1.5:1)$$

and the initial condition is

$$x = 1 \quad \text{when} \quad t = 0.$$

It is required to compute x when $t = 1$.

We shall use Euler's process, (see §1.2) for which the index is 1, and we shall begin by using 3 intervals, which is obviously too small to yield anything but a rough approximation. The calculation proceeds as shown in Table 1.5.1.

/Table....

Table 1.5.1

Step-by-Step Integration of a Differential Equation by Euler's Process. Index 1.

Number of Intervals 3. h = 1/3.

Column Number					
1	2	3	4	5	6
r	t_r	x_r	$\left(\frac{dx}{dt}\right)_r$	Δx_r = h X col.4	x_{r+1} = col. 3 + col. 5
0	0	1	$\frac{1}{2}$	1/6	7/6
1	1/3	1 1/6	29/60	29/180	239/180
2	2/3	1 59/180	169/360	169/1080	$\frac{1603}{1080}$ = 1.4843

The computation has been repeated with 5 and 10 intervals in use and the results are given in Table 1.5.2.

Table 1.5.2

Step-by-Step Integration of a Differential Equation by Euler's Process. Index 1..

Number of intervals : 5 and 10.

t	x	
	5 intervals used	10 intervals used
0	1.0	1.0
0.1		1.05
0.2	1.1	1.0995
0.3		1.1485
0.4	1.198	1.1971
0.5		1.2453
0.6	1.2942	1.2930
0.7		1.3403
0.8	1.3888	1.3873
0.9		1.4340
1.0	1.4819	1.4802

/Now..

Now let us derive a corrected value of x for $t = 1$ from the results obtained with 3 and 5 intervals. Since the index is 1 we get

$$\left.\begin{aligned} 1.4843 &= x(1) + \frac{e_o}{3} \\ 1.4819 &= x(1) + \frac{e_o}{5} \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots (1.5:2)$$

Hence by difference

$$0.0024 = \frac{2\,e_o}{15},$$

$$e_o = 0.0180$$

and

$$\frac{e_o}{5} = 0.0036.$$

Hence $x(1) = 1.4819 - 0.0036$

$= 1.4783.$

From the results for $n = 5$ and 10 we obtain

$$\left.\begin{aligned} 1.4819 &= x(1) + \frac{e_o}{5} \\ 1.4802 &= x(1) + \frac{e_o}{10} \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots (1.5:3)$$

Hence $e_o = 0.0170$

and $x(1) = 1.4785.$

Similarly from the results for $n = 3$ and 10 we derive

$$e_o = 0.0176$$

and $x(1) = 1.4784.$

Lastly, if we take all three results and retain the first two terms in the expansion for the error we have

$$\left.\begin{aligned} 1.4843 &= x(1) + \frac{e_o}{3} + \frac{e_1}{9} \\ 1.4819 &= x(1) + \frac{e_o}{5} + \frac{e_1}{25} \\ 1.4802 &= x(1) + \frac{e_o}{10} + \frac{e_1}{100} \end{aligned}\right\} \quad \ldots\ldots\ldots (1.5:4)$$

/These.....

These equations yield

$$e_0 = 0.0157$$
$$e_1 = 0.0043$$

$$x(1) = 1.4786.$$

An inspection of these results would lead us to assert with considerable confidence that $x(1)$ is 1.4786 with an error probably not exceeding 1×10^{-4} and almost certainly not exceeding 2×10^{-4}. The true solution is in fact

$$x(t) = 0.4t + 1.2 - 0.2\, e^{-0.5t} \quad \ldots\ldots\ldots (1.5:5)$$

which yields

$$x(1) = 1.4787.$$

Since the whole computation was made with only 4 places of decimals in use, a more accurate result could only have been obtained by a favourable accident. The assessment of the error is confirmed.

From each pair of corresponding results contained in Table 1.5.2 we can compute a corrected value of x, and the corrections to be applied to the intermediate results can be obtained by interpolation. In this way a corrected table can be constructed giving x for every 0.1 increment of t.

In the practical working of this example three different numbers of intervals would not be used, but one of the pairs 3, 5 or 5, 10 adopted according to the accuracy desired.

Example (b)

The differential equation is

$$\frac{dx}{dt} = \sqrt{(1 - x^2)}, \quad \ldots\ldots\ldots\ldots\ldots (1.5:6)$$

and the boundary condition $x = 0$ when $t = 0$. Find x when $t = 1$. The true solution is

$$x = \sin t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1.5:7)$$

and $x = 0.8415$ when $t = 1$. Here the solution can be obtained by interchange of the dependent and independent variables, quadrature for a range of values of x and interpolation. However, the example provides a good illustration of the methods at present being applied.

We begin by adopting Euler's method (index 1) and shall use 3, 5 and 10 intervals. Only the final results for $x(1)$ will be quoted, namely

/$x(1)_3$....

$$x\,(1)_3 = 0.9016$$
$$x\,(1)_5 = 0.8766$$
$$x\,(1)_{10} = 0.8586.$$

The corrected results obtained in the manner given under Example (a) are:

From 3 and 5 intervals	$x\,(1)$	$= 0.8391$
" 3 " 10 "	$x\,(1)$	$= 0.8402$
" 5 " 10 "	$x\,(1)$	$= 0.8406$
" 3, 5 and 10 "	$x\,(1)$	$= 0.8413.$
True value	$x\,(1)$	$= 0.8415$ (to 4 places).

Example (c)

The problem is the same as in Example (b) but we shall now adopt the process of index 2 briefly described in § 1.3. The tabular computation with 5 intervals is given in Table 1.5.3.

Table 1.5.3.

Step-by-Step Integration of a Differential Equation by a Process of Index 2.

Number of intervals 5. $h = 0.2$.

Column Number						
1	2	3	4	5	6	7
r	t_r	x_r	$\left(\frac{dx}{dt}\right)_r$	$x_{r+\frac{1}{2}}$	$\left(\frac{dx}{dt}\right)_{r+\frac{1}{2}}$	x_{r+1}.
0	0	0	1.0	0.1	$\sqrt{0.99} = 0.9950$	0.1990
1	0.2	0.1990	$\sqrt{0.9604}=0.9800$	0.2970	$\sqrt{0.9118} = 0.9549$	0.3900
2	0.4	0.3900	$\sqrt{0.8479}=0.9208$	0.4821	$\sqrt{0.7676} = 0.8761$	0.5652
3	0.6	0.5652	$\sqrt{0.6805}=0.8249$	0.6477	$\sqrt{0.5805} = 0.7619$	0.7176
4	0.8	0.7176	$\sqrt{0.4851}=0.6965$	0.7872	$\sqrt{0.3803} = 0.6167$	0.8409

In the Table the number under the radical in column 4 is 1 minus the square of the entry in column 3, and the final entry in column 4 is the initial slope for the interval. This is used to find an approximate value of x for the middle of the interval (column 5); the increment of x is $0.1 \times$ entry in column 4. Column 6 is derived from column 5 in the same manner as 4 from 3; the entry is a mean slope for the interval. The entries in column 5 are much less accurate than those in column 7, and

/are.....

are to be discarded when the calculation is complete. Column 3 is not really necessary since all the entries, except the first, already appear in column 7.

The great superiority of this process over the Euler process in accuracy is shown by comparison of the error in the last entry of Table 1.5.3 with that obtained before correction in Example (b) with 5 intervals; the errors are 0.0006 and 0.0351 respectively. Since the additional labour in the present process is small, it must be regarded as much preferable to that of Euler.

A comparison of the results obtained by the method of index 2 with 5 and 10 intervals is provided in Table 1.5.4. (page 13a)..

Table 1.5.4

Step-by-Step Integration of a Differential Equation

by a Process of Index 2.

Numbers of Intervals : 5 and 10

t	Uncorrected approximations to x(t).		Correction to value for 10 intervals deduced from corresponding values.	Interpolated correction	Final corrected approximation	True value of x(t). (To 4 figures)	Error in final approximation
	5 intervals	10 intervals					
0	0	0	0	—	0	0	0
0.1	-	0.0999	-	0	0.0999	0.0998	-0.0001
0.2	0.1990	0.1988	-0.0001	-	0.1987	0.1987	0
0.3	-	0.2957	-	-0.0001	0.2956	0.2955	-0.0001
0.4	0.3900	0.3896	-0.0001	-	0.3895	0.3894	-0.0001
0.5	-	0.4796	-	-0.0001	0.4795	0.4794	-0.0001
0.6	0.5652	0.5648	-0.0001	-	0.5647	0.5646	-0.0001
0.7	-	0.6443	-	-0.0001	0.6442	0.6442	0
0.8	0.7176	0.7174	-0.0001	-	0.7173	0.7174	+0.0001
0.9	-	0.7833	-	0	0.7833	0.7833	0
1.0	0.8409	0.8413	+ 0.0001	-	0.8414	0.8415	+0.0001

The process of correction will now be applied to the two results obtained for x(1). Since the index is 2 and the numbers of intervals 5 and 10, we have

$$\left.\begin{aligned} 0.8409 &= x(1) + \frac{e_o}{25} \\ 0.8413 &= x(1) + \frac{e_o}{100} \end{aligned}\right\} \quad \ldots\ldots\ldots (1.5:8)$$

Hence $e_o = -0.0133$

and $x(1) = 0.8414.$

Since the correction to the value of x(1) obtained with 10 intervals is only 0.0001, we should conclude that the error probably does not exceed this. The error (in the four figure answer) is just 0.0001. A more thorough test would have involved making the computations to five or six places of decimals throughout.

Note on the Correction Process

In all the examples we have taken equal intervals throughout the range of integration, but the same process of assessment of error and partial correction can be used with unequal intervals, provided that each of the original intervals is evenly subdivided into the same number of smaller intervals.

1.6 Use of an Analytical Solution in Each Interval of a Step-by-Step Integration

It sometimes happens that over a limited range of the independent variable the equation or equations to be solved approximate closely to an equation or set having a closed analytical solution, and that the soluble equations can be adjusted to the set to be integrated in each interval by suitable changes of parameters. Such a process of substitution is likely to be advantageous when the interval throughout which the parameters can be kept constant is much larger than the interval which must be used in the ordinary step-by-step integration. This can perhaps be made clearer by considering dynamical applications. In the ordinary step-by-step solution of dynamical equations, the time intervals must be so small that the velocities and displacements have only small percentage changes in each interval; for oscillatory motions the time intervals must be much smaller than the periods. However, a single analytical approximation can be used so long as the percentage changes of the parameters are small, and this may be true for intervals covering many periods.

/An.....

An example illustrating what has been said is provided by R.A. Frazer's method of mean coefficients[6] for the solution of sets of linear ordinary differential equations with variable coefficients. Here throughout each interval the coefficients are treated as constants and given values equal to their arithmetical mean values for the interval. The initial values of the dependent variables for the rth interval are obtained from the initial values for the first interval by continued matrix multiplication in Frazer's process; the introduction of matrices is convenient but not essential. It should be emphasised that the method described in the present section applies to non-linear as well as to linear equations.

An aeronautical application of the analytical step-by-step solution is contained in an investigation of the free motion of a stable glider in an atmosphere of variable density[9]. Here the rate of change of each dependent variable is expressed as the product of a linear function of these variables and of the square root of the relative air density. Since the relative density is a function of the dependent variables, the equations are non-linear. However, as it varies only relatively slowly it can be treated as constant during a time interval of many seconds. In the numerical example worked out in detail it was shown that satisfactory accuracy was obtained with an interval of 50 seconds whereas the periodic time of the phugoid oscillation was 36 seconds.

A final and obvious remark is that the method described here is of no value unless the closed analytical solution is suitable for reasonably easy computation.

1.7 Choice of a Step-by-Step Process

The circumstances in which the analytical step-by-step method can be used with advantage have been indicated in §1.6. Further discussion relates to the numerical step-by-step processes and is restricted to one-point boundary conditions.

Provided that all the methods under consideration are reliable and of reasonable simplicity, the criterion for selection should be the attainment of given accuracy with minimum labour. In accordance with §1.4 we shall assume that the dominant term in the error is $e_0 n^{-k}$ and that this represents the error with enough accuracy for the purpose of comparison of processes. Now we can easily ascertain the index k for any process but to set bounds to e_0 would require detailed investigation in each instance. Manifestly we have no right to assume that e_0 has like values for processes of differing indices. Nevertheless, we may fairly assume that a method of high index has been so designed that a given accuracy is attained with considerably wider intervals than would be required with processes of smaller indices. This advantage is, however, bought at the cost of greater labour per interval. For instance, the labour per interval in the Kutta process of index 4 is very much greater than in the Euler process of index 1.

/An.....

An important consideration is that the interval used in the integration should not be larger than the interval of the tabulation required in the result unless this is very small, for it would almost certainly be wasteful to use elaborate methods to integrate accurately with wide intervals and then to use interpolation to fill out the tabulation.

The present discussion is inevitably inconclusive, but the writer's belief is that processes of low index are usually to be preferred. On the whole he would favour a process of index 2, such as that described in §1.3. This is much more accurate than the Euler process when the interval is the same while the extra labour is small.

1.8 Two-Point Boundary Conditions

Step-by-Step methods are not well adapted to the solution of differential equations with two-point boundary conditions, but they can be used. To take first the simplest case, suppose that x_1 is given for $t = a$ and that x_2 is given for $t = b$, these being the only dependent variables (see § 1.1). Then we may start a step-by-step integration at $t = a$ with the known value of x_1 and some trial value λ of x_2, which should be guessed or estimated on the best evidence available. The integration is carried through, and the value x_2 (b) found. Unless by a fortunate accident, this will not be the assigned value and the calculation must be repeated with another value for λ. The process must be continued until the true value of λ can be found by interpolation with the required accuracy. Clearly the process is relatively laborious.

The same general method could be used for 2-point boundary problems with more dependent variables and for n-point boundary problems.

1.9 Aeronautical Applications

Step-by-step integration of differential equations or sets of them has often been used in aeronautical investigations and some representative examples are included in the list of references.[10,11,12,13] The authors of these papers have not always described the process of integration used, but it appears that these were usually of the simplest. Jones and Trevelyan[11] state "The initial forces acting upon the aeroplane are supposed constant for one tenth of a second; the change of the motion under the action of these forces during this time is then calculated, and taken to be the initial motion for the next tenth of a second, and so on".

Step-by-step integration has been most frequently applied to problems of looping, pull-outs from dives and stalled flight. In general the accuracy of

/the.....

the data has been so low that much refinement in the integration would not have been justified.

Very extensive applications of step-by-step integration have been made to exterior ballistics (motion of projectiles, bombs and guided missiles) but further consideration of these is beyond the scope of this paper.

/PART II.....

Part II

Classification and Review of Methods for the Numerical Integration of Ordinary Differential Equations

2.1 Introduction

The aim is to classify methods of numerical solution of differential equations and sets and to review existing methods very briefly.

2.2 Classification of Methods

The first and broadest division is into purely numerical or digital methods and those which are not entirely numerical, though capable of giving numerical results (see the classification chart - Table 2.2.1). The methods which are not purely digital can be divided into analogic and graphical. Analogic methods depend on the use of a physical or kinematic analogy, and include the use of such instruments as the differential analyser of Bush. Graphical methods are, strictly, analogic and this is indicated by the dotted line in the chart.

The digital methods can be divided broadly into progressive and holic or unitary. In the progressive methods the process of solution begins at one point and passes to other points in regular sequence; in some the whole process may be repeated until the required accuracy is obtained. The characteristic of the holic methods is that the range of integration is treated as a whole throughout the process of integration.

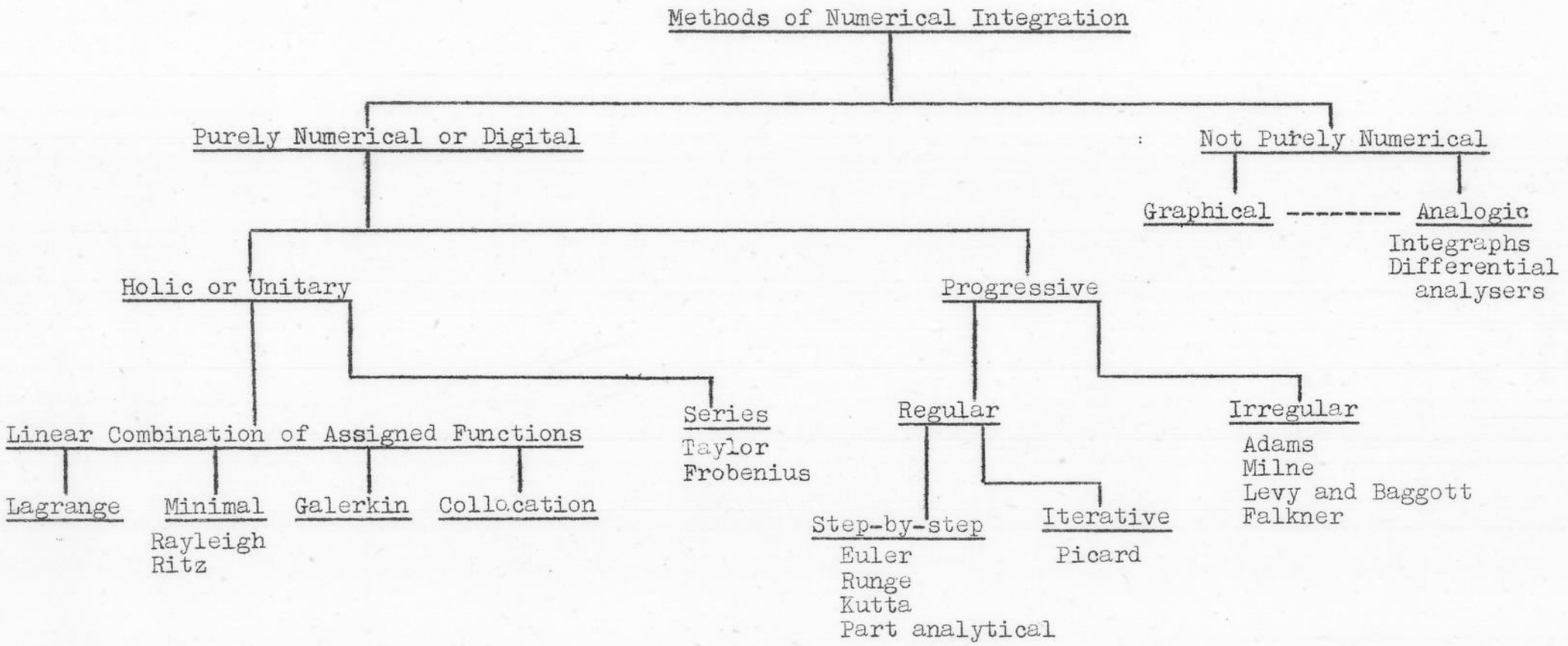
Progressive methods may be classified as regular, when the process of integration is the same throughout the range, and irregular. The irregular methods usually start with a very accurate computation of the solution over a short initial part of the range and then proceed step-by-step. The ordinary step-by-step and iterative methods are regular.

The holic methods, so far as they are known to the writer, depend either on approximate analytical solutions (series etc.) or representation of the solution as a linear combination of known functions which satisfies the boundary conditions for all values of the parameters. The latter are so determined that an approximate solution is obtained.

/Table 2.2.1.....

Table 2.2.1

Classification Chart for Methods of Numerical Integration of Ordinary Differential Equations



2.3 Brief Review of Methods

2.3.1 Graphical Methods. These may be of value in rough work or in preliminary explorations.[1,3] The attainable accuracy is low.

2.3.2 Analogic Methods. These are probably of most value for rather complicated equations or systems and when the range of integration is great. The attainable accuracy is not high.

2.3.3 Regular Progressive Methods. These include the step-by-step methods, numerical or part analytical, discussed in Part I of this paper, and the iterative method of Picard. The latter is discussed at length by Scarborough.[2]

2.3.4 Irregular Progressive Methods. The prototype is the method of Adams[4] where the highly accurate solution in the first part of the range is obtained from Taylor's series. The solution proceeds by extrapolation and the notation of finite differences is adopted.

Adams' method is the only one described by Whittaker and Robinson[5] who declare roundly "The best method of integrating differential equations numerically is one devised by J.C. Adams---" On the other hand Scarborough[2] makes the following critical comments "The success of the method in starting a solution depends upon (1) the ease with which successive derivatives of the unknown function can be calculated and (2) the rapidity with which the Taylor series converges. If the successive derivatives are easily calculated and the Taylor series converges rapidly, the method furnishes the best means of starting a solution and should be used in preference to any other. But if, on the other hand, the successive derivatives are not easily calculated, or if the Taylor series is such that the interval of convergence is not easily determined, the method should not be used." "Because of the difficulty of calculating the successive derivatives of the air-resistance function the method of Adams is not suitable for starting the computation of trajectories." A refinement of Adams' method is described by Falkner.[14]

The method of W.E. Milne[15] starts with a relatively accurate solution in an initial part of the range, obtained by the method of Adams, Runge-Kutta or other suitable method. The solution is extended by using two integrated forms of Newton's interpolation formula. An interesting feature of the method is that each value of a dependent variable is obtained by two processes and if the results do not agree sufficiently closely the interval must be reduced.

/2.3.5.....

2.3.5 Unitary or Holic Methods. These can be exemplified by minimal methods in which use is made of some stationary property of the solution. The approximate solution may be expressed as a linear combination of a finite number of known linearly independent functions so selected that the expression satisfies the boundary conditions for all values of the coefficients of some (or all) of the functions; these coefficients are then determined by the stationary property. The methods of Rayleigh and Ritz can be included here.

In the method of Galerkin[16,17] the solution is expressed in the same way, but the coefficients are determined by making n linearly independent weighted means of the residual error in the differential equation zero, where n is the number of functions with unknown coefficients.

The method here called Lagrangian depends on finding a dynamical analogy of the differential equation or equations; when the equations are derived from a mechanical problem the analogy is already provided. The solution is expressed in the same way as before and the unknown coefficients of the selected functions are regarded as generalised coordinates and are found from the Lagrangian dynamical equations. The minimal method may be regarded as a special case of the Lagrangian method and the Galerkin method becomes identical with the Lagrangian when the "weight functions" used in forming the mean errors are chosen in a particular manner.

In the method of collocation[18] the representation of the solution is the same as before but the unknown coefficients are found by making the residual error exactly zero at n selected points within the range of integration.

2.3.6 Relaxation. The relaxation technique can be applied to aid in the solution of equations obtained by various methods of approximation, e.g., the minimal. It is described by Southwell in two books[19,20] and in numerous papers.

/List of References....

- List of References -

Ref. No.	Authors	Title, etc.
1	Runge, C.	"Graphical Methods". New York : Columbia University Press, 1912.
2	Scarborough, J.B.	"Numerical Mathematical Analysis". Baltimore : The Johns Hopkins Press, 1930.
3	Levy, H and Baggott, E.A.	"Numerical Studies in Differential Equations, Part I". London : Watts and Co., 1934
4	Bashforth, F. and Adams, J.C.	"Capillary Action". Cambridge University Press, 1883.
5	Whittaker, E.T. and Robinson, G.	"The Calculus of Observations". London : Blackie, 1924.
6	Frazer, R.A. Duncan, W.J. and Collar, A.R.	"Elementary Matrices" Cambridge University Press, 1938.
7	Runge, C.	Mathematische Annalen, Vol.46, 1895.
8	Kutta, W.	Zeitschrift für Mathematik and Physik, Vol.46, p.443, 1901.
9	Duncan, W.J. and Scull, W.J.	"Free Motion of a Stable Glider in an Atmosphere of Variable Density". R. and M. 2081. March, 1941.
10	Bryant, L.W.	"On the Possible Loading of the Wings and Body of an Aeroplane in Flight". R. and M. 496, May, 1917.
11	Jones, B.M. and Trevelyan, A.	"Step-by-Step Calculations upon the Asymmetric Movements of Stalled Aeroplanes". R. and M. 999. Oct. 1925,
12	Bryant, L.W. and Williams, D.H.	Appendix 16 to "Report on the Puss Moth Accidents". R. and M.1699, July, 1936.
13	Newmark, S. Young, A.D. and Young, E.	"Results of a Step-by-Step Calculation of the Recovery of the Typhoon from a Terminal Velocity Dive". R.A.E. Report No. Aero.1903, A.R.C. 7422, January, 1944.

/14....

Ref. No.	Authors	Title, etc.
14	Falkner, V.M.	"A Method for the Numerical Solution of Differential Equations". Phil. Mag., Ser.7, vol.XXI, p.624, March, 1936.
15	Milne, W.E.	"Numerical Integration of Ordinary Differential Equations". American Mathematical Monthly, vol.33, p.455, 1926.
16	Duncan, W.J.	"Galerkin's Method in Mechanics and Differential Equations". R. and M. 1798, August, 1937.
17	Duncan, W.J.	"The Principles of the Galerkin Method". R. and M. 1848, September, 1938.
18	Frazer, R.A. Jones, W.P. and Skan, S.W.	"Approximations to Functions and to the Solutions of Differential Equations". R. and M. 1799, March, 1937.
19	Southwell, R.V.	"Relaxation Methods in Engineering Science". Oxford University Press, 1940.
20	Southwell, R.V.	"Relaxation Methods in Theoretical Physics". Oxford University Press, 1946.

-----oOo-----