

# CONVERGENCE AND STABILITY IN THE NUMERICAL INTEGRATION OF ORDINARY DIFFERENTIAL EQUATIONS

GERMUND DAHLQUIST

## 1. Introduction and summary

**1.1. Statement of the problem.** Consider a class of difference equations

$$(1.1) \quad \alpha_k y_{n+k} + \alpha_{k-1} y_{n+k-1} + \dots + \alpha_0 y_n = h(\beta_k f_{n+k} + \dots + \beta_0 f_n),$$

where  $h$  is a parameter and

$$f_0, f_1, f_2, \dots$$

are the values of a given function  $f(x, y)$  at equidistant arguments  $x_n = a + nh$ , i.e.

$$f_n = f(x_n, y_n).$$

We shall investigate the use of such equations for the numerical solution of the initial value problem for an ordinary differential equation

$$(1.2) \quad y' = f(x, y), \quad y(a) = y_0.$$

The behaviour of the solutions of (1.1) when  $h$  is small and  $n$  is large, is then of particular interest.

Some particular formulas of this class have been in practical use for a long time, the simplest example being the point-slope formula

$$(1.3) \quad y_{n+1} - y_n = h f_n.$$

It is well known in this case that  $y_n$  tends to the solution  $y(x)$  of (1.2) when  $h \rightarrow 0$  and  $n \rightarrow \infty$ , so that  $a + nh = x$ , provided that  $f(x, y)$  is continuous and satisfies a Lipschitz condition.

Two slightly more sophisticated formulas which have been utilized are

$$(1.4) \quad y_{n+1} - y_n = \frac{1}{2} h (f_{n+1} + f_n),$$

$$(1.5) \quad y_{n+2} - y_n = \frac{1}{3} h (f_{n+2} + 4f_{n+1} + f_n).$$

Equation (1.4) is based on a quadrature formula, known as the trapezoidal rule, and similarly (1.5) is based on Simpson's rule. We may note

Received October 23, 1955.

that (1.3) defines a straightforward computational scheme, whereas in (1.4)  $y_{n+1}$  is found implicitly on the right hand side so that there arises in each step an algebraic (or transcendental) equation, which is usually solved approximately by some iterative method. For brevity, this equation will always be called "algebraic" in accordance with the practice of Hartree and others. Such a formula is called a *closed* formula, while a formula like (1.3) (where  $y_{n+1}$  is not found on the right hand side) is *open*. Formula (1.5) is also closed. Note that it is a *second order* difference equation. A special procedure is therefore required in the beginning to give the value  $y_1$ . Such complications are accepted because of the greater accuracy of (1.4) or (1.5) when compared with (1.3).

One special feature is common to these particular cases: they contain the variable  $y$ , explicitly in two points only. This is the case for almost all practically used methods of this kind. More general formulas of the type (1.1) have recently been studied by Frei [4] and Quade [9]. However one may ask whether it would not be possible to make use of the information from the preceding points in a much more efficient manner by choosing a more complicated formula of the type (1.1). This question was the starting point of the writer's investigations. The main result is rather negative (Theorem 4), but there are new formulas of this general class which are at least comparable to the classical numerical methods when high accuracy is wanted, e.g. (1.10) and (1.11).

In order that the difference equation (1.1) should be useful for numerical integration, it is necessary that (1.1) be satisfied with good accuracy by the solution of (1.2), when  $h$  is small, for an arbitrary function  $f(x, y)$ . The exact definition of the word "arbitrary" will be given in Section 2.1. It follows from this that the value of the expression  $L[y(x)]$  defined by

$$(1.6) \quad L[y(x)] \equiv \sum_{v=0}^k \alpha_v y(x + vh) - h \beta_v y'(x + vh)$$

should be small when  $h$  is small, for all sufficiently regular functions  $y(x)$ . This imposes restrictions on the coefficients  $\alpha_v, \beta_v$  in (1.1). By expanding the terms in  $L[y(x)]$  into powers of  $h$ , we see that

$$L[y(x)] = O(h^{p+1})$$

for arbitrary  $y(x)$  if and only if the following  $p+1$  linear relations hold:

$$(1.7) \quad \sum_{v=0}^k \alpha_v = 0,$$

$$(1.8) \quad \sum_{v=0}^k (\alpha_v v^s / s! - \beta_v v^{s-1} / (s-1)!) = 0 \quad (s = 1, 2, \dots, p).$$

The largest value of  $p$  for which this holds will be called the *degree of the operator  $L$* , whereas the integer  $k$  (i.e. the number of preceding points occurring, explicitly or implicitly, in (1.1)) will be called the *order of the operator  $L$*  or the order of the difference equation (1.1). Since the number of coefficients is equal to  $2k + 2$ , one may expect that they can be chosen so that (1.7) and  $2k$  relations of the type (1.8) are satisfied, in which case  $p$  is equal to  $2k$ . A proof of this and an explicit expression for the corresponding operator will be found in Section 2.4.

By the expansion of (1.6) into powers of  $h$ , we see that, if  $p$  is the degree of  $L$ , then

$$(1.9) \quad L[y(x)] \sim -C h^{p+1} y^{(p+1)}(x) \quad (h \rightarrow 0),$$

where  $C$  is a non-zero constant independent of  $y(x)$ . In particular,  $L[y(x)]$  vanishes identically when  $y(x)$  is a polynomial whose degree is less than or equal to  $p$ .

A simple calculation gives for instance the following values for the three methods discussed above.

Point-slope method:  $k = 1, p = 1, C = 1/2$ .

Trapezoidal rule:  $k = 1, p = 2, C = 1/12$ .

Simpson's rule:  $k = 2, p = 4, C = 1/90$ .

The value of  $\alpha_k^{-1} L[y(x)]$  is essentially equal to the error in the estimation of  $y_{n+k}$  on the assumption that  $y_{n+k-1}, y_{n+k-2}, \dots, y_n$  are exact. More accurately, it gives the error in  $y_{n+k} - h\beta_k \alpha_k^{-1} f_{n+k}$  on that assumption. It is therefore called the *local truncation error*. For a good formula this should be small, and hence  $p$  should be large. This is however not sufficient, since one has also to consider the *inherited error*, i.e. the error produced in  $y_{n+k}$  by the errors occurring in the previous steps. It will be proved (Section 2.6) that those methods for which  $p$  is high show a very unpleasant error growth, which is in fact so strong that *it is necessary to reject all formulas for which  $p > k + 2$* . An "explanation" may be that the sensitivity to delicate features in the sequence  $y_n, y_{n+1}, \dots, y_{n+k}$ ;  $f_n, f_{n+1}, \dots, f_{n+k}$ , which is expressed by a high value of  $p$ , is indistinguishable from a sensitivity to perturbations. Some perturbations, e.g. *round-off errors*, are unavoidable in numerical computations, so that not even the *difference equation* is solved exactly. Another source of perturbations is the algebraic equation occurring in each step, which is usually not solved exactly. We must also observe that the difference equation requires  $k$  initial values. Only one of them is given in the original problem, and the accuracy of the final results depends on the choice of procedure used to obtain the estimates of the values  $y_1, y_2, \dots, y_{k-1}$ . From

a practical point of view one is led to require that small changes of the extra initial values must not cause large changes in the final results. A numerical example in the next section will probably make this so-called *strong numerical instability* more clear.

The instability of formulas of degree  $p > k + 2$  is the worse, the smaller  $h$  has been chosen. All formulas of degree  $p = k + 2$  give a different type of numerical instability, whose effect in a given range of integration can become arbitrarily small, provided that  $h$  has been chosen small enough. We may call this phenomenon *weak instability*. It is not so serious that the formulas of degree  $p = k + 2$  must be rejected. On the contrary, the writer is of the opinion that some of them are the best ones for practical use when high accuracy is wanted. Some care is however necessary, in particular when dealing with problems where there are transients which are damped out quickly compared to the timescale of the phenomena under consideration. In fact the difference equations in this case possess oscillating, "parasitic", solutions whose amplitudes increase the more quickly the quicker the transients are damped out in the correct solution. When these formulas are used, it is, however, important that the extra initial values  $y_1, y_2, \dots, y_{k-1}$  should be determined with an accuracy comparable to the local truncation error in the following computations, and the round-off errors kept smaller than the local truncation errors. In particular cases, the weak instability has been discussed in the literature, for instance by Todd, Dahlquist, Rutishauser, Löwdin, Craggs and Mitchell. See also Milne [7] and Collatz [2]. These problems will be dealt with in another paper under preparation, where some quantities, called *growth parameters*, will be introduced for the study of this *weak instability*.

The growth parameters have been used in a heuristic discussion of formulas of low order which results in the suggestion of two formulas of order 4 and degree 6.

$$(1.10) \quad y_{n+4} = -y_{n+3} + y_{n+1} + y_n + 3h(f_{n+3} + f_{n+1} + \frac{1}{10}\delta^4 f_{n+2}),$$

$$(1.11) \quad y_{n+4} = +y_{n+3} - y_{n+1} + y_n + h(f_{n+3} + f_{n+1} + \frac{4}{3}\delta^2 f_{n+2} + \frac{29}{90}\delta^4 f_{n+2}).$$

These may be compared to Newton-Cotes' formula, which can be written

$$(1.12) \quad y_{n+4} = y_n + 2h(f_{n+3} + f_{n+1} + \frac{1}{3}\delta^2 f_{n+2} + \frac{7}{45}\delta^4 f_{n+2}).$$

The first one has, roughly speaking, a smaller regular error component, but it is less stable, while the second one has a larger regular error component, but is more stable.

A more detailed error analysis and some numerical examples will be presented in the paper under preparation. Moreover, an attempt will be made to generalize to other types of differential equations and numerical methods.

**1.2. A numerical example.** Apply the formula

$$y_{n+2} = -4y_{n+1} + 5y_n + h(4f_{n+1} + 2f_n)$$

to the differential equation

$$y' = y, \quad y(0) = 1.$$

The local truncation error is small:

$$Ly \sim \frac{1}{6}h^4y^{(4)}(x).$$

This formula has in fact the smallest truncation error among all open formulas of order 2. If we take  $h=0,1$  we get the *difference equation*

$$(1.13) \quad y_{n+2} = -3,6 y_{n+1} + 5,2 y_n, \quad y_0 = 1$$

with the *exact* solution

$$y_n = (1-A)\zeta_1^n + A\zeta_2^n,$$

where

$$\zeta_1 \approx 1,105168, \quad \zeta_2 \approx -4,705168$$

are the roots of the quadratic equation

$$\zeta^2 + 3,6 \zeta - 5,2 = 0$$

and  $A$  is a constant determined by the choice of  $y_1$ . The solution of the *differential equation* is  $e^{0,1n} \approx 1,105171^n$ .

Note that the term  $A\zeta_2^n$  has a very strong growth, e.g.  $\zeta_2^{10} \approx 5,3 \cdot 10^6$ . We may call this the *parasitic solution*. We see below the results of two computations using (1.13) where  $y_{n+1}$  has been rounded off correctly to six decimals in each step before the computation of  $y_{n+2}$ .

CASE I: Take  $y_1 = 1,105171$ , i.e.  $e^{0,1}$  correct to six decimal places. In this case  $A \approx -5,16 \cdot 10^{-7}$ . Hence, although the initial value is good, we obtain a component of the parasitic solution.

CASE II: Take  $y_1 = \zeta_1$ . Hence  $A = 0$ . The error of the theoretical solution of the difference equation is then less than  $8 \cdot 10^{-5}$ . (See Case IIb in the table). But in a recursive numerical computation to six decimals, the round-off errors inevitably introduce the parasitic solution, and the actual error is larger than 0,2 for  $n = 10$ . See Case IIa in the table.

$n$	CASE I (numerical solution)		CASE IIa (numerical solution)		CASE IIb $\zeta_1^n$ with six correct dec.	
	$y_n$	$10^6 \cdot \text{error}$	$y_n$	$10^6 \cdot \text{error}$	$10^6 \cdot \text{error}$	
0	1,000000	0	1,000000	0	1	
1	1,105171	0	1,105168	3	1,10516781	3
2	1,221384	19	1,221395	8	1,221396	7
3	1,349907	-48	1,349852	7	1,349847	12
4	1,491532	293	1,491787	38	1,491808	17
5	1,650001	-1280	1,648797	-76	1,648698	23
6	1,815963	6156	1,821623	496	1,822088	31
7	2,042538	-28785	2,015902	-2149	2,013713	40
8	2,089871	135670	2,215192	10349	2,225491	50
9	3,097662	-638059	2,507999	-48396	2,459541	62
10	-0,284254	3,002536	2,490202	228080	2,718205	77

In spite of the favourable local truncation error, the results are very bad. Case II with the "wrong"  $y_1$  is better than Case I. The results would be still worse with a smaller value of  $h$ , since the number of steps in a given range would increase.  $|\zeta_2|$  increases also a little, but that is of minor importance. The instability is in this case a consequence of Theorem 5 below. Here  $k=2$ ,  $p=3$ .

The simple trapezoidal rule gives an error less than 0,0023 for  $n=10$ . Simpson's rule is able to give more than five decimals accuracy.

## 2. A theory of strong instability.

**2.1. The class of differential equations under consideration.** We assume that a solution  $y(x)$  to the equation

$$(2.1) \quad dy/dx = f(x, y), \quad y(a) = y_0$$

exists for  $a \leq x \leq b$ , and that  $y(x)$  has continuous derivatives in every point of this interval up to the order  $p+1$ , where  $p$  is the degree of the operator  $L$  under consideration. It is also assumed that  $f(x, y)$  possesses continuous derivatives up to the second order at every point  $(x, y)$  of a region  $R$  defined by the inequalities

$$|y - y(x)| \leq r, \quad a \leq x \leq b.$$

$r$  shall be so large that an approximate solution which does not stay in  $R$  is of no interest. Hence we may put

$$(2.2) \quad |y^{(p+1)}(x)| < K, \quad a \leq x \leq b,$$

$$(2.3a) \quad |f(x, y) - f(x, \tilde{y})| < M|y - \tilde{y}|$$

for arbitrary points  $(x, y) \in R$ ,  $(x, \tilde{y}) \in R$ .

The following theory will apply also to  $s$  simultaneous equations, unless otherwise stated. The quantities  $y, f$  etc. may be interpreted as vectors in  $s$  dimensions,  $|y|$  being the norm of the vector  $y$  in some suitable sense.

We shall give another formulation of the Lipschitz condition which is more convenient when dealing with integration formulas of the closed type. Since in this case the dependent variable has to be solved from an algebraic equation in each step, the meaning of the solution of the *difference equation* is not completely clear a priori. In order to define this, the following lemma is needed. (Cf. the proof of Theorem 1 below.) The lemma is not proved here, since it follows immediately from the general theorem on iteration, given for instance in the book of Collatz [2, p. 35 ff.]

**LEMMA.** *Let  $\tilde{u} = \tilde{u}(x)$  be a given vector in  $s$  dimensions,  $s \geq 1$ , defined for  $a \leq x \leq b$  and let  $\lambda$  be a given constant. Put*

$$y(x) - \lambda f(x, y(x)) = u(x).$$

*Suppose that  $|\lambda M| < 1$  and that  $|u(x) - \tilde{u}(x)| < r(1 - |\lambda M|)$ .*

*Then, for each  $x$ ,  $a \leq x \leq b$ , the equation*

$$\tilde{y}(x) - \lambda f(x, \tilde{y}(x)) = \tilde{u}(x)$$

*has exactly one solution  $\tilde{y}(x)$  in  $R$ .*

Let  $\tilde{u}$  and  $u$  be two vectors satisfying the conditions of the lemma, and let  $\tilde{y}$  and  $y$  be the corresponding solutions of the equation mentioned there. Then

$$|y - \tilde{y}| \leq |u - \tilde{u}| + |\lambda| |f(x, y) - f(x, \tilde{y})| \leq |u - \tilde{u}| + |\lambda M| |y - \tilde{y}|.$$

Hence

$$(2.4) \quad |y - \tilde{y}| \leq (1 - |\lambda M|)^{-1} |u - \tilde{u}|,$$

and the new formulation of the Lipschitz condition reads

$$(2.3b) \quad |f(x, y) - f(x, \tilde{y})| \leq M' |u - \tilde{u}|$$

where

$$M' = M(1 - |\lambda M|)^{-1}.$$

**2.2. The class of operators.** It will be convenient to discuss the operators  $L$  in terms of their *generating polynomials*

$$\begin{aligned}\varrho(\zeta) &\equiv \alpha_k \zeta^k + \alpha_{k-1} \zeta^{k-1} + \dots + \alpha_0, \\ \sigma(\zeta) &\equiv \beta_k \zeta^k + \beta_{k-1} \zeta^{k-1} + \dots + \beta_0.\end{aligned}$$

The roots of  $\varrho(\zeta)$  will be denoted by

$$\zeta_{10}, \zeta_{20}, \dots, \zeta_{k0}.$$

Let  $E$  be the operator defined by

$$E y_n = y_{n+1} \quad \text{or} \quad E y(x) = y(x+h).$$

We shall *always* put  $n = [(x-a)/h]$ . Then

$$L y(x) \equiv \varrho(E) y(x) - h \sigma(E) y'(x),$$

and the difference equation (1.1) reads

$$(2.5) \quad \varrho(E) y_n - h \sigma(E) f_n = 0.$$

The following assumptions will always be made:

- A. The coefficients  $\alpha_k, \beta_k$  are real,  $\alpha_k \neq 0$ .
- B. The polynomials  $\varrho(\zeta)$  and  $\sigma(\zeta)$  have no common factor.
- C. The degree  $p$  of the operator  $L$  is at least equal to 1.

Assumption A needs hardly any comment. Assumption B is made for the sake of convenience in the theory, but nothing essential is lost. Note for instance, that if  $\varphi(\zeta)$  were a common factor, then (2.5) would be equivalent to an equation of lower order

$$\varrho_1(E) y_n - h \sigma_1(E) f_n = \psi_n,$$

where

$$\begin{aligned}\varrho_1(\zeta) &\equiv \varrho(\zeta)/\varphi(\zeta), \\ \sigma_1(\zeta) &\equiv \sigma(\zeta)/\varphi(\zeta)\end{aligned}$$

and  $\psi_n$  is a solution to the equation  $\varphi(E)\psi_n = 0$ , completely determined by the method of start. The use of (2.5) is in the long run not much better than the use of the lower order equation

$$\varrho_1(E) y_n - h \sigma_1(E) f_n = 0.$$

Assumption C is more interesting. We see by (1.7) and (1.8) that it is equivalent to the two equations

$$(2.6) \quad \varrho(1) = 0, \quad \varrho'(1) = \sigma(1).$$

The first equation is *necessary* for the existence of a solution  $y_n$  to the difference equation which converges uniformly to a continuous function  $y(x) \neq 0$  when  $h \rightarrow 0, n \rightarrow \infty$  in such a way that  $n = [(x-a)/h]$ . For then



$$|y_{n+\nu} - y(x)| < \varepsilon \quad (\nu = 0, 1, \dots, k)$$

if  $h$  is small enough, and hence, by (2.5)

$$y(x) \sum_{\nu=0}^k \alpha_\nu = \varepsilon \sum_{\nu=0}^k |\alpha_\nu| + O(h).$$

Since  $y(x) \neq 0$  in general, it follows that  $\sum \alpha_\nu = 0$ . Hence  $\varrho(1) = 0$ .

If  $y_n$  converges uniformly to a continuous function  $y(x)$ , then the second equation (2.6) makes sure that  $y(x)$  satisfies the correct differential equation. For, put

$$\varrho(\zeta) \equiv (\zeta - 1) \varrho_1(\zeta).$$

Then, by (2.5)

$$\varrho_1(E) (y_{\nu+1} - y_\nu) - h \sigma(E) f_\nu = 0.$$

Sum from 0 to  $n$ , and put  $F_n = \sum_{\nu=0}^n h f_\nu$ :

$$\varrho_1(E) (y_{n+1} - y_0) = \sigma(E) F_n.$$

But

$$y_{n+\nu} \rightarrow y(x), \quad y_\nu \rightarrow y(a),$$

$$F_{n+\nu} \rightarrow \int_a^x (f(t, y(t)) dt \quad (\nu = 0, 1, \dots, k).$$

Hence

$$\varrho_1(1) (y(x) - y(a)) = \sigma(1) \int_a^x f(t, y(t)) dt.$$

According to Assumption B,  $\sigma(1)$  is different from zero. Hence, by differentiation we get

$$\varrho'(1) dy/dx = \sigma(1) f(x, y)$$

since  $\varrho'(1) = \varrho_1(1)$ . This differential equation is equivalent to (2.1) if and only if  $\varrho'(1) = \sigma(1)$ .

**2.3. Stable convergence and stable operators.** Besides the equation (2.5), the following perturbed equations will be considered

$$(2.7) \quad \varrho(E) \tilde{y}_n = h \sigma(E) \tilde{f}_n + \eta_n,$$

$$\tilde{y}_\kappa = y_\kappa + \vartheta_\kappa \quad (\kappa = 0, 1, \dots, k - 1).$$

$\tilde{y}_n$  should be interpreted as the estimate of  $y_n$ , with a prescribed number of digits or decimal places, obtained in a numerical solution, and  $\tilde{f}_n$  is defined as the exact value of  $f(x_n, \tilde{y}_n)$ .  $\vartheta_\kappa$  is the error in the initial value  $\tilde{y}_\kappa$ , and  $\eta_n$  is called the local perturbation in step  $n$ . In order to get an

idea of  $\eta_n$ , the reader may compare Case IIa and Case IIb in Section 1.2.  $\eta_n$  represents the contribution at step  $n$  to the deviation of the numerical solution from the theoretical solution of the *difference* equation (1.1) with initial values  $\tilde{y}_0, \tilde{y}_1, \dots, \tilde{y}_{k-1}$ . It does not include the interest of the "error capital" produced by round-off errors and other approximations in previous steps. Nor does it include the local truncation error. The local perturbation  $\eta_n$  arises from two sources:

a. The values of the derivatives actually used in the computations in step  $n$ , say,  $f_n^*, f_{n+1}^*, \dots, f_{n+k-1}^*$ , are usually not identical with  $\tilde{f}_n, \tilde{f}_{n+1}, \dots, \tilde{f}_{n+k-1}$  due to round-off and to approximations in the solution of the algebraic equation for  $\tilde{y}_n$  etc.

b. There are round-off errors and approximations of the kind mentioned in a, when  $y_{n+k}$  and  $f_{n+k}^*$  are computed from the given values

$$\tilde{y}_n, \tilde{y}_{n+1}, \dots, \tilde{y}_{n+k-1}, f_n^*, f_{n+1}^*, \dots, f_{n+k-1}^* .$$

The actual value of  $\eta_n$  is different in different arrangements of the computations, even if the method of integration is the same. In view of this, it is natural to require that the solution of (2.7) should be close to the solution of the differential equation (2.1) for *all* small perturbations. This requirement will now be formulated more precisely with the aid of the notion of *stable convergence*.

Let  $P(h, \varepsilon)$  be the class of all perturbations (2.7) to a given individual equation of the form (2.5) for which

$$\sum_{v \leq \frac{b-a}{h}} |\eta_v| + \max_{x \leq k-1} |\vartheta_x| \leq \varepsilon .$$

We shall say that  $\tilde{y}_n$  *converges stably* ( $P$ ) to  $y(x)$  in the interval  $(a, b)$  if

$$\overline{\lim}_{h \rightarrow 0} \sup_{P(h, \varepsilon)} \sup_{a \leq x \leq b} |\tilde{y}_n - y(x)| \leq K\varepsilon ,$$

where  $K$  is independent of  $h$  and  $\varepsilon$  ( $0 < \varepsilon < \varepsilon_0$ ), but it may depend on the interval and the function  $f(x, y)$  considered.

One may think of alternative definitions of this notion. From a practical point of view the limit process  $h \rightarrow 0$  stands essentially for the consideration of values of  $h$  which are sufficiently small in some sense. Since in actual computations the perturbations are usually kept smaller than the local truncation error, or at least of the same of order of magnitude, one may for instance consider a class  $P'(h, \alpha, \beta)$  or more briefly  $P'(h)$ , defined by the inequalities

$$|\eta_n| < \alpha h^\beta, \quad |\vartheta_x| < \alpha h^\beta ,$$

where  $\alpha > 0, \beta > 1$  are some fixed numbers. Then  $y_n$  converges stably ( $P'$ ) to  $y(x)$ , if

$$\sup_{P'(h)} |\tilde{y}_n - y(x)| \rightarrow 0$$

uniformly for  $a \leq x \leq b$ , when  $h \rightarrow 0$ . In connexion with this definition we shall put

$$\varepsilon = \sum_{r \leq \frac{b-a}{h}} |\eta_r| + \max_{\kappa \leq k-1} |\vartheta_\kappa| .$$

It is seen that  $\varepsilon \rightarrow 0$ , if  $h \rightarrow 0$ . The theorem below applies to both definitions, showing that they are essentially equivalent.

An operator  $L$  is said to be a *stable operator*, if the following conditions are satisfied in addition to the assumptions A, B, C stated above:

D. The roots of  $\varrho(\zeta)$  are located within or on the unit circle.

E. The roots on the unit circle are distinct.

This definition is motivated by the following theorem:

**THEOREM 1.** *A necessary and sufficient condition that  $\tilde{y}_n$  should converge stably to the solution of (2.1) for all differential equations of the class considered is that the operator  $L$  be stable.*

The necessity can be proved by the consideration of some special differential equation. Take for instance

$$dy/dx = qy, \quad y(0) = 1$$

in the interval  $(0,1)$ ;  $q$  is a non-zero constant.

The characteristic polynomial of the difference equation is then

$$\varrho(\zeta) - qh\sigma(\zeta) .$$

If  $h$  is small but different from zero, this polynomial has distinct roots  $\zeta_j = |\zeta_j| e^{i\varphi_j}$  given by the formula

$$(2.8) \quad \zeta_j = \zeta_{j0} + \left( \frac{m! \sigma(\zeta_{j0})}{\varrho^{(m)}(\zeta_{j0})} qh \right)^{1/m} + O(h^{2/m}) ,$$

provided that  $\zeta_{j0}$  is a root of  $\varrho(\zeta)$  of multiplicity  $m \geq 1$ . (Note that  $\sigma(\zeta_{j0}) \neq 0$ , by assumption B). The solution of the difference equation (2.5) is thus

$$c_1 \zeta_1^n + c_2 \zeta_2^n + \dots + c_k \zeta_k^n .$$

It is now easily seen that there are two difference equations belonging to the class  $P(h, \varepsilon)$  whose solutions differ by  $A |\zeta_j|^n \cos n\varphi_j$  if for instance  $A = \varepsilon \min \{ |\zeta_j|^{1-k}, 1 \}$ . Similarly if

$$A = \alpha h^\beta \min \{ |\zeta_j|^{1-k}, 1 \} ,$$

the same is true for the class  $P'(h)$ . Since  $n = [(x - a)/h]$ , it is seen in both cases that the condition

$$\overline{\lim}_{h \rightarrow 0} |\zeta_j|^{1/h} < \infty$$

is necessary for stable convergence. Condition  $D$  is an immediate consequence.

Next, assume that  $|\zeta_{j0}| = 1$ , and that  $c$  is the greatest number among the real parts of the  $m$  determinations of the  $m$ -th root of

$$m! q \sigma(\zeta_{j0}) / (\zeta_{j0} \varrho^{(m)}(\zeta_{j0})).$$

Then

$$\overline{\lim}_{h \rightarrow 0} |\zeta_j|^{1/h} = \overline{\lim}_{h \rightarrow 0} (1 + ch^{1/m} + O(h^{2/m}))^{1/h}$$

for at least one of the roots. Hence

$$\overline{\lim}_{h \rightarrow 0} |\zeta_j|^{1/h} = \infty$$

if  $m \geq 2$  and  $c > 0$ . Now if  $m \geq 3$ , at least one of the  $m$ -th roots of a non-zero complex number has a positive real part so that  $c > 0$ . The same is true for  $m = 2$ , unless  $q \sigma(\zeta_{j0}) / \zeta_{j0} \varrho^{(m)}(\zeta_{j0})$  is real and negative. But this limitation is of no importance, because stable convergence is required both for positive and for negative values of  $q$ . We may then conclude that condition  $E$  is necessary.

In order to prove the sufficiency, assume that  $L$  is stable. The proof is rather long. It is similar in principle to some proofs in the stability theory of differential equations [1, p. 82 ff.], but the method of successive approximations is replaced by a step by step method.

Let  $y(x)$  be the solution of the differential equation. Put

$$\begin{aligned} x_v &= a + v h, & y_v &= y(x_v), & f_v &= f(x_v, y_v), \\ \lambda &= h \beta_k \alpha_k^{-1}, & u_v &= y_v - \lambda f_v, & \tilde{u}_v &= \tilde{y}_v - \lambda \tilde{f}_v. \end{aligned}$$

Then, by (1.9) and (2.2),

$$\begin{aligned} \varrho(E) y_n &= h \sigma(E) f_n - l_n, \\ |l_n| &< K C' h^{p+1}, \end{aligned}$$

where  $C'$  depends on  $h$ . If  $h \rightarrow 0$ , then  $C' \rightarrow C$ . Subtract this from (2.7), and introduce  $u_n$  and  $\tilde{u}_n$ , instead of  $y_n$  and  $\tilde{y}_n$ :

$$\begin{aligned} (2.9) \quad \varrho(E) (\tilde{u}_n - u_n) &= q_n, \\ q_n &\equiv \sum_{j=0}^{k-1} (h \beta_j - \lambda \alpha_j) E^j (\tilde{f}_n - f_n) + \eta_n + l_n. \end{aligned}$$

Observe that  $\tilde{f}_{n+k} - f_{n+k}$  does not appear in  $q_n$ , so that (2.9) determines  $\tilde{u}_{n+k} - u_{n+k}$  explicitly. For this reason the variables  $u_n, \tilde{u}_n$  will be used instead of  $y_n, \tilde{y}_n$ .

It is easily verified that the difference equation (2.9) is equivalent to the sum equation

$$(2.10) \quad \tilde{u}_{n+k} - u_{n+k} = \sum_{\nu=0}^n g_{n+k-\nu} q_\nu + \vartheta_{n+k}'$$

where  $g_n$  and  $\vartheta_n'$  are solutions to the homogeneous difference equation  $\varrho(E)g_n = 0$ . The former is determined for  $n \geq 1$  by the initial conditions

$$g_1 = g_2 = \dots = g_{k-1} = 0, \quad g_k = \alpha_k^{-1},$$

and we extend the definition by putting  $g_n = 0$ , for  $n \leq 0$ , so that actually

$$\varrho(E)g_n = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

$\vartheta_n'$  is determined by the initial values

$$\vartheta_0' = 0, \quad \vartheta_\alpha' = \tilde{u}_\alpha - u_\alpha \quad (\alpha = 1, 2, \dots, k - 1).$$

Put  $\max|\tilde{y}_\alpha - y_\alpha| = \Theta$ . Then  $\max|\vartheta_\alpha'| \leq \Theta(1 + |\lambda M|)$ . For a stable operator the solutions of the homogeneous equation  $\varrho(E) = 0$  are bounded, since they are combinations of terms like  $\zeta_j^n, |\zeta_j| \leq 1$  or  $n^m \zeta_j^n, |\zeta_j| < 1$ . Hence, for a given operator, there exists a constant  $G$ , independent of  $h$ , such that

$$|g_n| < G, \quad |\vartheta_n'| < G\Theta.$$

Assume that  $h$  and  $\varepsilon$  are so small that

$$(2.11) \quad |\lambda M| < 1, \quad \varepsilon + KC'h^p(b-a) < r(1 - |\lambda M|)G^{-1}e^{-GM''(b-a)}$$

( $r$  was defined in Section 2.1, and  $M''$  is defined below). Provided that the following inequality holds,

$$|\tilde{u}_\nu - u_\nu| < r(1 - |\lambda M|),$$

the lemma in Section 2.1 defines  $\tilde{y}_\nu$  uniquely in the region  $R$  where the modified Lipschitz condition (2.3b) can be applied. Suppose that this inequality holds for  $\nu \leq n+k-1$ , where  $n$  is a certain given integer, such that  $a + (n+k)h \leq b$ . Now we shall apply mathematical induction. It is to be shown that

$$|\tilde{u}_{n+k} - u_{n+k}| < r(1 - |\lambda M|).$$

The assumption is true for  $n = 0$ , since

$$|\tilde{u}_x - u_x| < G\Theta \leq G\varepsilon < r(1 - |\lambda M|),$$

by (2.11).

We get, by the application of (2.3b) to the definition of  $q_v$ ,

$$|q_v| \leq hM'' \max_{\mu \leq v+k-1} |\tilde{u}_\mu - u_\mu| + |\eta_v| + KC'h^{p+1} \quad (v \leq n),$$

where

$$M'' = M (1 - |\lambda M|)^{-1} \sum_{j=0}^{k-1} |\beta_j - \alpha_j \beta_k \alpha_k^{-1}|.$$

The sum equation (2.10) then gives

$$\begin{aligned} |\tilde{u}_{n+k} - u_{n+k}| &\leq G \sum_{v=0}^n (hM'' \max_{\mu \leq v+k-1} |\tilde{u}_\mu - u_\mu| + |\eta_v| + KC'h^{p+1}) + G\Theta \\ &< hGM'' \sum_{v=0}^n \max_{\mu \leq v+k-1} |\tilde{u}_\mu - u_\mu| + G(\varepsilon + KC'h^p(b-a)). \end{aligned}$$

It follows that  $\max_{\mu \leq v} |\tilde{u}_\mu - u_\mu| < z_v$ , where  $z_v$  is the increasing sequence defined by the recurrence relation

$$z_{n+k} = hGM'' \sum_{v=k-1}^{n+k-1} z_v + G(\varepsilon + KC'h^p(b-a)) \quad (n \geq 0)$$

which is equivalent to the difference equation

$$\begin{aligned} z_{n+k} - z_{n+k-1} &= hGM'' z_{n+k-1} \quad (n \geq 0), \\ z_{k-1} &= G(\varepsilon + KC'h^p(b-a)). \end{aligned}$$

Hence

$$\begin{aligned} z_{n+k} &= (1 + hGM'')^{n+1} z_{k-1} < e^{GM''(b-a)} z_{k-1}, \\ |\tilde{u}_{n+k} - u_{n+k}| &< G(\varepsilon + KC'(b-a)h^p) e^{GM''(b-a)} \end{aligned}$$

and thus by (2.11)

$$|\tilde{u}_{n+k} - u_{n+k}| < r(1 - |\lambda M|).$$

By induction the preceding formulas are seen to be valid for every  $n$ , if  $a + (n+k)h \leq b$ . It now follows from (2.4) that

$$|\tilde{y}_{n+k} - y_{n+k}| < G(\varepsilon + KC'(b-a)h^p) e^{GM''(b-a)} (1 - |\lambda M|)^{-1}.$$

The stable convergence follows immediately from this inequality, which also gives an upper bound for the error, which is, however, in general rather poor.

**2.4. The existence of an operator of degree  $2k$ .** We shall obtain an expression for the operator for which  $p$  is as large as possible for a given  $k$ . We may put  $h=1$  without loss of generality. Put

$$\begin{aligned} \varphi(x) &= x(x-1)(x-2) \dots (x-k), \\ \varphi_\xi(x) &= \varphi(x)/(x-\xi) \quad (\xi = 0, 1, \dots, k). \end{aligned}$$

Let  $y_0, y_1, \dots, y_k, f_0, f_1, \dots, f_k$  be  $2k+2$  given numbers. There is a unique polynomial  $y(x)$  of degree  $2k+1$  for which

$$y(\xi) = y_\xi, \quad y'(\xi) = f_\xi, \quad (\xi = 0, 1, \dots, k).$$

It has the form

$$y(x) = \sum_{\xi=0}^k \left\{ \left( \frac{\varphi_\xi(x)}{\varphi_\xi(\xi)} \right)^2 y_\xi + \frac{\varphi_\xi(x)\varphi(x)}{(\varphi_\xi(\xi))^2} \left( f_\xi - \frac{2\varphi_\xi'(\xi)}{\varphi_\xi(\xi)} y_\xi \right) \right\}.$$

This formula is a confluent form of Lagrange's interpolation formula, which can be derived from the expansion of  $y(x)(\varphi(x))^{-2}$  into partial fractions. See e.g. Steffensen [11, p. 33]. In special cases the degree of  $y(x)$  becomes less than  $2k+1$ . We see by equating the coefficient of  $x^{2k+1}$  to zero in the expression for  $y(x)$  that the necessary and sufficient condition for this is

$$(2.12) \quad \sum_{\xi=0}^k (\varphi_\xi(\xi))^{-2} (f_\xi - 2y_\xi\varphi_\xi'(\xi)/\varphi_\xi(\xi)) = 0.$$

In other words: Equation (2.12) gives a relation between the values of a function and its derivative in  $k+1$  equidistant points which is valid for any polynomial of degree  $2k$  (or less), and it is the only linear relation between these quantities which is satisfied for *all* polynomials of degree  $2k$ .

We obtain from the definition of  $\varphi(x)$  etc. after a short calculation

$$\begin{aligned} (\varphi_\xi(\xi))^{-2} &= ((k-\xi)! \xi!)^{-2} = \binom{k}{\xi}^2 (k!)^{-2}, \\ \varphi_\xi'(\xi)/\varphi_\xi(\xi) &= h_\xi - h_{k-\xi}, \end{aligned}$$

where

$$h_m = 1 + 2^{-1} + 3^{-1} + \dots + m^{-1}, \quad h_0 = 0.$$

We put these expressions into (2.12) and divide by the coefficient of  $y_k$ . The result is:

**THEOREM 2.** *There is an operator of order  $k$  for which the degree is as high as  $2k$ . This operator is unique, apart from a trivial factor, and the coefficients are*

$$\alpha_j = h_k^{-1} (h_j - h_{k-j}) \binom{k}{j}^2, \quad \beta_j = \frac{1}{2} h_k^{-1} \binom{k}{j}^2.$$

The generating functions of this operator are thus

$$(2.13) \quad \begin{aligned} \rho_k(\zeta) &= h_k^{-1} \sum_{j=0}^k (h_j - h_{k-j}) \binom{k}{j}^2 \zeta^j, \\ \sigma_k(\zeta) &= \frac{1}{2} h_k^{-1} \sum_{j=0}^k \binom{k}{j}^2 \zeta^j. \end{aligned}$$

The polynomial  $\sigma_k(\zeta)$  can be expressed in terms of the classical Legendre polynomials. In fact, the relation

$$(2.14) \quad P_k(z) = \left(\frac{1}{2}(z-1)\right)^k \sum_{j=0}^k \binom{k}{j}^2 \left(\frac{z+1}{z-1}\right)^j \\ \equiv 2h_k \left(\frac{1}{2}(z-1)\right)^k \sigma_k\left(\frac{z+1}{z-1}\right)$$

is easily derived by the application of Leibniz' rule for the  $n$ -th derivative to Rodrigues' formula

$$P_k(z) = (k!)^{-1} \left(\frac{1}{2}d/dz\right)^k \left((z-1)^k (z+1)^k\right)$$

which is sometimes taken as the definition of the Legendre polynomial.

The corresponding operators are however in general unstable (cf. 2.6). For instance, if  $k=3$ , we get

$$\varrho_3(\zeta) = \zeta^3 + \frac{27}{11}(\zeta^2 - \zeta) - 1,$$

one root of which is equal to  $(-19 - (240)^{1/2})/11 \approx -3,136$ .

**2.5. The construction of operators from one arbitrary generating polynomial.** The constants  $C$  and  $p$  in the fundamental relation (1.9)

$$L[y(x)] \sim -C h^{p+1} y^{(p+1)}(x) \quad (h \rightarrow 0)$$

are independent of  $y(x)$ . They can thus be determined by a special example. Take  $y=e^x$  and put  $e^h=\zeta$ . Hence

$$(2.15) \quad \varrho(\zeta) - \sigma(\zeta) \log \zeta \sim -C (\zeta - 1)^{p+1} \quad (\zeta \rightarrow 1).$$

The problems under discussion are therefore related to the problem of approximating  $\log \zeta$  by rational functions in the neighbourhood of  $\zeta=1$  and to the study of the distribution of the zeros of the denominator.  $\log \zeta$  is here the branch defined by the condition  $\log 1=0$ . We shall in the following consider the  $\zeta$ -plane cut along the negative real axis. Part of the problem can therefore be treated by the analytic theory of continued fractions, but the writer prefers the following more direct way. It is convenient to perform the transformations

$$\zeta = (z+1)/(z-1), \quad z = (\zeta+1)/(\zeta-1),$$

$$(2.16) \quad R(z) = \left(\frac{1}{2}(z-1)\right)^k \varrho(\zeta) \equiv \sum_{j=0}^k a_j z^j, \\ S(z) = \left(\frac{1}{2}(z-1)\right)^k \sigma(\zeta) \equiv \sum_{j=0}^k b_j z^j.$$

The polynomials  $R(z)$  and  $S(z)$  will be called the first and second *asso-*



ciated polynomials of the operator  $L$ , respectively. Since  $\varrho(1)=0$ , according to (2.6), we get the condition  $a_k=0$ , so that  $R(z)$  is actually a polynomial of degree  $k-1$ .

In these notations the relation (2.15) transforms into

$$(2.17) \quad R(z) - S(z) \log \frac{z+1}{z-1} \sim -C \left(\frac{2}{z}\right)^{p-k+1} \quad (z \rightarrow \infty),$$

$$(2.18) \quad R(z) \left(\log \frac{z+1}{z-1}\right)^{-1} - S(z) \sim -C \left(\frac{2}{z}\right)^{p-k} \quad (z \rightarrow \infty).$$

The  $z$ -plane is cut along the segment from  $-1$  to  $1$  of the real axis. The logarithm is then uniquely determined by the condition that it vanishes for  $z = \infty$ .

**THEOREM 3.** *It is possible to construct an operator  $L$  of degree  $p \geq k+1$ , starting from any polynomial  $\varrho(\zeta)$  which vanishes for  $\zeta=1$ . For a given  $\varrho(\zeta)$  the construction is unique.*

For define  $S(z)$  as the principal part of  $R(z) \left(\log((z+1)/(z-1))\right)^{-1}$  at infinity. Since  $R(z)$  is a polynomial of degree  $(k-1)$ , this is a polynomial of degree  $k$ . Then

$$R(z) \left(\log \frac{z+1}{z-1}\right)^{-1} - S(z) = O(z^{-1})$$

and hence, by (2.18),  $p-k \geq 1$ , and the theorem is proved.

If an arbitrary constant  $c \neq 0$  is added to the polynomial  $S(z)$  thus defined, it follows from the same argument that the degree will be equal to  $k$ . If  $c$  is added to  $S(z)$ , then  $c(\zeta-1)^k$  is added to  $\sigma(\zeta)$ . An open operator is obtained by taking  $c = -\beta_k$ . Hence, it is possible to construct an open operator  $L$  of degree  $p \geq k$  starting from any polynomial  $\varrho(\zeta)$  vanishing for  $\zeta=1$ . Since  $(\zeta-1)^k$  is the generating function of the  $k$ -th difference  $\Delta^k f_n$ , we may also say that an arbitrary operator is equal to the sum of an open operator of degree  $p \geq k$  and a difference correction  $\beta_k \Delta^k \tilde{f}_n$ . This may be a convenient way of looking at the formulas in practice. One can obtain the first approximation to  $\tilde{y}_{n+k}$  by applying the open operator together with an extrapolated estimate of  $\beta_k \Delta^k \tilde{f}_n$ .

It may be proved that it is possible to construct an operator  $L$  of degree  $p \geq k$  from an arbitrary polynomial  $\sigma(\zeta)$ . It is only necessary to apply to (2.17) the argument in the proof of Theorem 3.

**2.6. The maximum degree for a stable operator.** The requirement of stability leads to inequalities for the associated polynomials, from which

interesting conclusions may be drawn. Since the unit circle in the  $\zeta$ -plane is mapped onto the half-plane  $\operatorname{Re} z \leq 0$  by the transformation (2.16), we may formulate the stability conditions in the following manner.

D'. The roots of  $R(z)$  have non-positive real parts.

E'. There are no multiple roots of  $R(z)$  on the imaginary axis. The coefficient  $a_{k-1} \neq 0$ .

The reason for the last condition is the simplicity of the root  $\zeta = 1$ , which corresponds to  $z = \infty$ . It follows from (2.16) that

$$a_{k-1} = 2^{1-k} \varrho'(1) \neq 0.$$

Consider the factorization

$$R(z) = a_{k-1} \prod (z + \alpha_\nu) \prod ((z + \alpha_\nu)^2 + \beta_\nu^2) \quad (\alpha_\nu \geq 0),$$

where the first product is taken over the real roots and the second product over the conjugate pairs of complex roots. The following condition is seen immediately.

LEMMA. *If  $L$  is a stable operator, all non-vanishing coefficients of  $R(z)$  have the same sign.*

We shall normalize the operators, so that

$$a_{k-1} = 1, \quad a_j \geq 0.$$

Put the expansion

$$(2.19) \quad \left( \log \frac{z+1}{z-1} \right)^{-1} = \frac{z}{2} - \sum_{\nu=0}^{\infty} \mu_{2\nu+1} z^{-(2\nu+1)}$$

into (2.18). Then

$$R(z) \left( \log \frac{z+1}{z-1} \right)^{-1} - S(z) = - \sum_{\nu=1}^{\infty} c_\nu z^{-\nu}$$

where

$$(2.20) \quad \begin{aligned} c_{2\lambda} &= \sum_{\nu} \mu_{2\nu+1} a_{2\nu-2\lambda+1}, \\ c_{2\lambda+1} &= \sum_{\nu} \mu_{2\nu+1} a_{2\nu-2\lambda}. \end{aligned}$$

The coefficients  $\mu_{2\nu+1}$  will play a major role in the following. The first values are

$$\mu_1 = 1/6, \quad \mu_3 = 2/45, \quad \mu_5 = 22/945, \quad \mu_7 = 214/14175.$$

The most essential fact is that *the coefficients  $\mu_{2\nu+1}$  are positive*, since by Cauchy's formula

$$\begin{aligned}
 \mu_{2\nu+1} &= -\frac{1}{2\pi i} \int_C z^{2\nu} \left( \log \frac{z+1}{z-1} \right)^{-1} dz \\
 &= -\frac{1}{2\pi i} \int_{-1}^1 x^{2\nu} \left( \pi^2 + \log^2 \frac{1+x}{1-x} \right)^{-1} \cdot \\
 (2.21) \quad &\quad \cdot \left( \left( -\pi i + \log \frac{1+x}{1-x} \right) - \left( \pi i + \log \frac{1+x}{1-x} \right) \right) dx \\
 &= \int_{-1}^1 x^{2\nu} \left( \pi^2 + \log^2 \frac{1+x}{1-x} \right)^{-1} dx > 0.
 \end{aligned}$$

Here  $C$  is an arbitrary curve enclosing the segment  $(-1, 1)$  of the real axis. It then follows from (2.20) that  $c_1 > 0$ , unless  $R(z)$  is an odd polynomial, in which case  $c_2 > 0$ . Hence, by (2.18),  $p-k$  cannot exceed 2, and we obtain the following theorem, which is one of the main results of the present investigation.

**THEOREM 4a.** *The degree  $p$  of a stable operator of order  $k$  can never exceed  $k+2$ . If an operator is stable, then the condition that  $R(z)$  is an odd function is necessary and sufficient for the degree to be equal to  $k+2$ . All roots of  $R(z)$  are then located on the imaginary axis. If  $k$  is odd, the degree of a stable operator cannot exceed  $k+1$ .*

The last statement follows from the assumption E'. For because  $a_{k-1} \neq 0$ ,  $R(z)$  cannot be an odd function when  $k$  is odd. The theorem is "the best possible", since if  $k$  is even, we can always find odd polynomials satisfying conditions D' and E'.

The following formulation of this theorem in terms of the *generating* polynomials is obtained by a straightforward application of the transformation formulas (2.16).

**THEOREM 4b.** *If an operator of even order  $k$  is stable, then the conditions*

$$(2.22) \quad \alpha_\nu = -\alpha_{k-\nu}, \quad \beta_\nu = \beta_{k-\nu}$$

*are necessary and sufficient in order that it should be of maximum degree  $k+2$ . All roots of  $\rho(\zeta)$  then have unit modulus.*

It is now easy to write down all stable operators of maximum degree for a given order  $k$ . For instance, when  $k=2$  there is only one stable operator of degree 4, apart from a trivial factor. This is determined by the polynomial  $R(z)=z$ , and it follows from the transformation formulas (2.16), that it corresponds to the operator

$$L[y] = y_{n+2} - y_n - \frac{1}{3}h(y_{n+2}' + 4y_{n+1}' + y_n'),$$

i.e. Simpson's rule.

For  $k=4$ , the general stable operator of maximum degree is determined by the polynomial

$$R(z) = z^3 + \lambda^2 z,$$

where  $\lambda$  is a real non-zero constant. The corresponding operator can be written

$$L[y] = y_{n+4} + \mu(y_{n+3} - y_{n+1}) - y_n - h \left\{ (4 + 2\mu)y_{n+2}' + \frac{1}{3}(8 + \mu)\delta^2 y_{n+2}' + \frac{1}{90}(28 - \mu)\delta^4 y_{n+2}' \right\},$$

where

$$\mu = 2(1 - \lambda^2)/(1 + \lambda^2) \quad |\mu| < 2.$$

This includes as particular cases ( $\mu=1$ ,  $\mu=-1$ ,  $\mu=0$ ) the formulas (1.10), (1.11), (1.12) mentioned in the introduction.

We shall prove one more theorem of the same type.

**THEOREM 5.** *If  $p > k$  for a stable operator, this must be closed. Actually,  $\beta_k/\alpha_k > 0$ . It is possible to construct an open, stable operator of degree  $p=k$ , starting from any polynomial  $\rho(\zeta)$ , satisfying the conditions A-E.*

An operator is open if  $\beta_k=0$ . But

$$\beta_k = \lim_{\zeta \rightarrow \infty} \zeta^{-k} \sigma(\zeta) = \lim_{z \rightarrow 1} ((z-1)/(z+1))^k (\frac{1}{2}(z-1))^{-k} S(z) = S(1).$$

Assume that  $p \geq k+1$ . Then  $S(z)$  is equal to the principal part of  $R(z)\{\log((z+1)/(z-1))\}^{-1}$ , and we may write (cf. (2.19) and (2.18)):

$$(2.23) \quad S(z) = R(z) \left( \frac{1}{2}z - \sum_{\nu=1}^{k-1} \mu^\nu z^{-\nu} \right) + \sum_{\nu=1}^{k-1} \mu'_\nu z^{-\nu},$$

where

$$\mu'_\nu = \sum_{\sigma=\nu}^{k-1} \mu_\sigma a_{\sigma-\nu}.$$

Hence

$$\beta_k = S(1) = R(1) \left( \frac{1}{2} - \sum_{\nu=1}^{k-1} \mu_\nu \right) + \sum_{\nu=1}^{k-1} \mu'_\nu.$$

Here  $R(1)$  and  $\sum_{\nu=1}^{k-1} \mu'_\nu = \sum_{\sigma=1}^{k-1} \sum_{\lambda < \sigma} \mu_\sigma a_\lambda$  are non-negative if the operator is stable. Furthermore, by Abel's theorem

$$\frac{1}{2} - \sum_{\nu=1}^{k-1} \mu_\nu > \frac{1}{2} - \sum_{\nu=1}^{\infty} \mu_\nu = \lim_{z \rightarrow 1+0} \left\{ \log((z+1)/(z-1)) \right\}^{-1} = 0.$$

It follows that  $\beta_k > 0$  for a stable operator if  $p > k$ . Hence  $p \leq k$  for an open stable operator. It was shown in Section 2.5 that an open operator

whose degree is at least equal to  $k$  can always be obtained with a given  $\varrho(\zeta)$ . The theorem is thus proved.

REFERENCES

1. R. Bellman, *Stability theory of differential equations*, New York, 1953.
2. L. Collatz, *Numerische Behandlung von Differentialgleichungen*, 2 Aufl., Berlin, 1955.
3. G. Dahlquist, *Fehlerabschätzungen bei Differenzenmethoden zur numerischen Integration gewöhnlicher Differentialgleichungen*, Z. Angew. Math. Mech. 31 (1951), 239–240.
4. T. Frei, *Anwendung der Momente der Integralkurven zur numerischen Lösung von Differentialgleichungen*, Magyar Tud. Akad. Alkalm. Int. Köz. 2 (1953), 395–414 (1954).
5. P.-O. Löwdin, *On the numerical integration of ordinary differential equations of the first order*, Quart. Appl. Math. 10 (1952), 97–111.
6. P.-O. Löwdin, *Några synpunkter på problemet om den numeriska integrationen av ett system av ordinära differentialekvationer med begynnelsevillkor. Del I–III. Rapport till Försvarets Forskningsanstalt*, Stockholm, 1953.
7. W. E. Milne, *Numerical solution of Differential Equations*, New York, 1953.
8. A. R. Mitchell and T. W. Craggs, *Stability of Difference Relations in the Solution of Ordinary Differential Equations*, Mathematical Tables and Other Aids to Computations 7 (1953), 127–129.
9. W. Quade, *Numerische Integration gewöhnlicher Differentialgleichungen durch Interpolation nach Hermite*, Proc. Int. Congr. Mathematicians 1954, p. 372, Amsterdam, 1954.
10. H. Rutishauser, *Über die Instabilität von Methoden zur Integration gewöhnlicher Differentialgleichungen*, Z. Angew. Math. Physik 3 (1952), 65–74.
11. J. F. Steffensen, *Interpolation*, Baltimore, 1927.
12. J. Todd, *Solution of Differential Equations by Recurrence Relations*, Mathematical Tables and Other Aids to Computation 4 (1950), 39–44.