

# A Convergence Criterion for a Class of Integration Methods

By J. C. Butcher

**Abstract.** A criterion is established for the convergence of sequences of a general type of Runge-Kutta method. The criterion is expressed in terms of "weak convergence", a property defined in the paper.

**1. Introduction.** While much is known about local properties of large classes of numerical methods for ordinary differential equations, attention to global questions is mainly confined to the consideration of the behaviour of the result obtained by iterated applications of a basic numerical method. The main results of this sort can be found in the work of Dahlquist [1] and of Henrici [2]. Dahlquist found necessary and sufficient conditions for convergence in the case of linear multi-step methods and Henrici gave an exposition of this theory and also dealt in detail with the case of a general class of one-step method. With a view to generalising and unifying these theories, the present author [3] dealt with a class of methods characterised by a pair of matrices; it was possible to express conditions for convergence of a method in terms of algebraic properties of these matrices.

The present paper may be regarded as a sequel to [4]. In that paper, certain properties of a generalised type of Runge-Kutta method were investigated. In particular, it was shown that numerical properties of a Runge-Kutta method  $m$  are characterised by  $w(m) \in G$  where  $G$  is the set of real valued functions on  $T$ , the set of (rooted) trees, and  $w(m)$  is defined by  $w(m)(t) = v_a(t)(h_1)$  for all  $t \in T$  where  $m = (H, a, h_1)$ . The various notations and the function  $v_a$  are defined in [4].

In this paper, a characterisation is given for convergent sequences of the type of generalised Runge-Kutta method described in [4]. This formulation is general enough to be applicable not only to the usual sort of global procedure in which a given method is used repeatedly but also to a wider class of methods.

**2. Preliminary Remarks.** In this section, certain developments of the theory in [4] are presented. They will be made use of later in the paper.

Let  $H$  be an arbitrary set,  $B(H)$  the set of bounded real valued functions on  $H$  and  $A(H)$  the set of continuous linear functions  $B(H) \rightarrow B(H)$ . If  $a \in A(H)$ ,  $b \in A(K)$ , where  $H, K$  are finite sets, then it is possible to define the direct product  $a \otimes b \in A(H \times K)$  by the following formulae

$$(a \otimes b)(x)(h, k) = a(u(\cdot, k))(h), \quad u(h, \cdot) = b(x(h, \cdot)),$$

where  $x \in B(H \times K)$ ,  $h \in H$ ,  $k \in K$  and  $x(h, \cdot) \in B(K)$  is defined by  $x(h, \cdot)(k) =$

Received April 17, 1969; revised February 10, 1971.

AMS 1970 subject classifications. Primary 65L05; Secondary 34A45, 34A50.

Key words and phrases. Runge-Kutta methods, initial value problems, convergence.

Copyright © 1972, American Mathematical Society

$x(h, k)$  with similar definitions for  $y(\cdot, k)$ ,  $y(h, \cdot)$ , etc. Alternately,  $a \otimes b$  can be defined in an analogous manner but with  $a$  operating before  $b$ . We shall show in Lemma 2.1 that there is no difficulty in extending this definition to the case where one of  $H, K$  is infinite.

LEMMA 2.1. *If one of  $H, K$  is finite,  $a \in A(H), b \in A(K)$ , then there exists  $a \otimes b \in A(H \times K)$  with the property that if  $x \in B(H \times K)$  and  $(a \otimes b)(x) = z$ , then there are  $u, v \in B(H \times K)$  such that*

$$\begin{aligned} z(\cdot, k) &= a(u(\cdot, k)), & u(h, \cdot) &= b(x(h, \cdot)), \\ z(h, \cdot) &= b(v(h, \cdot)), & v(\cdot, k) &= a(x(\cdot, k)), \end{aligned}$$

for all  $h \in H, k \in K$ .

*Proof.* Suppose, for example, that  $H$  is finite with elements  $h_1, h_2, \dots, h_n$ . For  $x \in B(H \times K)$  and  $i = 1, 2, \dots, n$ , let  $x_i \in B(K)$  be defined by  $x_i(k) = x(h_i, k)$  for all  $k \in K$ . Let  $a_{i,j}$  ( $i, j = 1, 2, \dots, n$ ) be defined by

$$\sum_{i=1}^n a_{i,j} y(h_i) = a(y)(h_j), \quad i = 1, 2, \dots, n,$$

for  $y \in B(H)$ . Then, the lemma can be immediately verified with  $u, v, z$  defined by

$$u(h_i, \cdot) = b(x_i), \quad v(\cdot, k) = a(x(\cdot, k)), \quad z(h_i, k) = \sum_i a_{i,j} b(x_i)(k).$$

It is clear that  $a \otimes b$  as defined is linear and bounded.

In [4] for a given  $a \in A(H)$ , the functions  $\mu_a, \nu_a: T \rightarrow B(H)$  were defined as follows:

(2.1) 
$$\mu_a(\tau) = e_H,$$

(2.2) 
$$\nu_a(t) = a(\mu_a(t)), \quad t \in T,$$

(2.3) 
$$\mu_a(tu) = \mu_a(t)\nu_a(u), \quad t, u \in T.$$

Here,  $e_H$  denotes the function taking each  $h \in H$  to 1. The product on the right-hand side of (2.3) is pointwise and the product on the left of elements of  $T$  is defined in [4]. We will need the following result:

LEMMA 2.2. *If one of  $H, K$  is finite and  $a \in A(H), b \in A(K)$ , then for all  $t \in T$ ,*

$$\mu_{a \otimes b}(t) = \mu_a(t) \otimes \mu_b(t), \quad \nu_{a \otimes b}(t) = \nu_a(t) \otimes \nu_b(t),$$

*Proof.* Note that the product  $\otimes: B(H) \times B(K) \rightarrow B(H \times K)$  is defined by  $(\alpha \otimes \beta)(h, k) = \alpha(h) \cdot \beta(k)$  for all  $\alpha \in B(H), \beta \in B(K), h \in H, k \in K$ . Because of the way  $\mu_a, \nu_a$  are defined in (2.1), (2.2), (2.3), it is necessary only to show that for all  $t, u \in T$ ,

(2.4) 
$$e_{H \times K} = e_H \otimes e_K,$$

(2.5) 
$$\nu_a(t) \otimes \nu_b(t) = (a \otimes b)(\mu_a(t) \otimes \mu_b(t)),$$

(2.6) 
$$\mu_a(tu) \otimes \mu_b(tu) = (\mu_a(t) \otimes \mu_b(t)) \cdot (\nu_a(u) \otimes \nu_b(u))$$

and of these (2.4), (2.5) are easily verified. To verify (2.6), we use an obvious property of pointwise products to show that the right-hand side is  $(\mu_a(t)\nu_a(u)) \otimes (\mu_b(t)\nu_b(u))$  which equals the left-hand side.

The application of Lemma 2.2 will occur in Section 5. In that section also, we will need to consider an equation of the form

$$(2.7) \quad y(\theta) = \theta f(y(\theta)),$$

where for a given  $\theta$ ,  $y(\theta)$  is a member of a Banach space  $X$  and  $f: X \rightarrow X$  is a function satisfying the condition that for each  $x \in X$  and for each positive integer  $n$ , the  $n$ th derivative  $f^{(n)}(x)$  can be defined. By this, we shall mean that  $f^{(n)}(x)$  is a bounded linear function on  $X^n$  to  $X$  such that for given  $x_1, x_2, \dots, x_n \in X$  and any positive number  $\epsilon$ , there is a positive number  $\delta$  such that for all  $z$  satisfying  $\|z - x\| < \delta$  it holds that

$$\begin{aligned} & \|f^{(n)}(z)(x_1, x_2, \dots, x_n) - f^{(n)}(x)(x_1, x_2, \dots, x_n) \\ & \quad - f^{(n+1)}(x)(x_1, x_2, \dots, x_n, z - x)\| < \epsilon \|z - x\|. \end{aligned}$$

It will also be assumed that for each  $n$ , the set of values of  $\|f^{(n)}(x)\|$  for  $x \in X$  is bounded. We shall denote  $\sup_{x \in X} \|f^{(1)}(x)\|$  by  $L$ . Let  $E$  denote the open interval  $(-1/L, 1/L)$ .

**THEOREM 2.3.** *If  $f$  satisfies the conditions described above, then there is a unique function  $y: E \rightarrow X$  satisfying (2.7). Furthermore,  $y$  is differentiable arbitrarily often on  $E$ .*

*Proof.* It is easy to see that  $f$  satisfies the Lipschitz condition

$$\|f(z) - f(x)\| \leq L \|z - x\|$$

for all  $z, x \in X$ . Hence, the existence and uniqueness of  $y(\theta)$  for each  $\theta \in E$  satisfying (2.7) follows from the contraction mapping theorem. This defines the function  $y: E \rightarrow X$ . To prove that  $y$  is differentiable at each  $\theta \in E$ , consider  $z(\theta)$  satisfying

$$(2.8) \quad z(\theta) = f(y(\theta)) + \theta f'(y(\theta))(z(\theta)).$$

Since  $f'(y(\theta))$  is a linear operator satisfying  $\|f'(y(\theta))\| \leq L$ , we see that for each  $\theta \in E$ ,  $z(\theta)$  is defined uniquely by (2.8), again using the contraction mapping theorem. For  $\theta, \theta' \in E$ , we find from (2.7) that

$$y(\theta') - y(\theta) = \theta(f(y(\theta')) - f(y(\theta))) + (\theta' - \theta)f(y(\theta')).$$

Hence,

$$(2.9) \quad \|y(\theta') - y(\theta)\| \leq \frac{c}{1 - |\theta|L} |\theta' - \theta|,$$

where  $c = \sup_{x \in X} \|f(x)\|$ . We now compute  $y(\theta') - y(\theta) - (\theta' - \theta)z(\theta)$  making use of (2.7) and (2.8). We find

$$\begin{aligned} y(\theta') - y(\theta) - (\theta' - \theta)z(\theta) &= \theta'(f(y(\theta')) - f(y(\theta))) - f'(y(\theta))(y(\theta') - y(\theta)) \\ &\quad + \theta f'(y(\theta))(y(\theta') - y(\theta)) - (\theta' - \theta)z(\theta) \\ &\quad + (\theta' - \theta)f'(y(\theta))(y(\theta') - y(\theta)). \end{aligned}$$

Hence, it follows that

$$\begin{aligned} & \|y(\theta') - y(\theta) - (\theta' - \theta)z(\theta)\| \\ & \leq \frac{1}{1 - |\theta|L} \left\{ \frac{1}{L} \|f(y(\theta')) - f(y(\theta)) - f'(y(\theta))(y(\theta') - y(\theta))\| \right. \\ & \quad \left. + L |\theta' - \theta| \cdot \|y(\theta') - y(\theta)\| \right\}. \end{aligned}$$

By the definition of  $f'$ , for any positive  $\epsilon_1$ , we can choose a positive  $\epsilon_2$  so that as long as  $\|y(\theta') - y(\theta)\| < \epsilon_2$ ,

$$\|f(y(\theta')) - f(y(\theta)) - f'(y(\theta))(y(\theta') - y(\theta))\| \leq L \epsilon_1 \|y(\theta') - y(\theta)\|.$$

By (2.9), there is  $\delta_1$  such that for  $\theta' \in E$  satisfying  $|\theta' - \theta| < \delta_1$  we have  $\|y(\theta') - y(\theta)\| < \epsilon_2$ .

Choose  $\delta = \min(\delta_1, \epsilon_1/L)$  so that for  $\theta' \in E$  satisfying  $|\theta' - \theta| < \delta$  we have

$$\begin{aligned} \|y(\theta') - y(\theta) - (\theta' - \theta)z(\theta)\| &\leq \frac{2\epsilon_1}{1 - |\theta|L} \|y(\theta') - y(\theta)\| \\ &\leq \frac{2c\epsilon_1}{(1 - |\theta|L)^2} |\theta' - \theta|. \end{aligned}$$

Given  $\epsilon > 0$ , choose  $\epsilon_1$  so that  $2c\epsilon_1/(1 - |\theta|L)^2 < \epsilon$  and then choose  $\delta$  in the way that has been described above. It is seen that for  $\theta' \in E$  satisfying  $|\theta' - \theta| < \delta$ , we have  $\|y(\theta') - y(\theta) - (\theta' - \theta)z(\theta)\| \leq \epsilon |\theta' - \theta|$ . That is,  $y$  is differentiable at  $\theta$  and  $z(\theta)$  is its derivative.

In a similar way, we can prove that  $z$  (which we will now write as  $y^{(1)}$ ) is differentiable. We obtain a sequence of equations for  $y^{(1)}, y^{(2)}, \dots$  of the form

$$(2.10) \quad y^{(n)}(\theta) = g_n(\theta) + \theta f'(y(\theta))(y^{(n)}(\theta)),$$

where  $g_1(\theta) = f(y(\theta))$  and, for  $n = 1, 2, \dots$ ,

$$(2.11) \quad g_{n+1}(\theta) = g'_n(\theta) + f'(y(\theta))(y^{(n)}(\theta)) + \theta f''(y(\theta))(y^{(n)}(\theta), y^{(1)}(\theta)).$$

By a tedious calculation of the type carried out above, we can show that  $y^{(2)}, y^{(3)}, \dots$  are the higher derivatives of  $y$ . Apart from these details, which we omit, the proof is complete.

We shall now obtain a convenient expression for  $y^{(n)}(\theta)$ .

LEMMA 2.4. Let  $k_1, k_2, \dots : E \rightarrow X$  be defined by  $k_1(\theta) = f(y(\theta))$ ,  $k_2 = k'_1$ ,  $k_3 = k'_2$ ,  $\dots$ . Then, for  $n = 1, 2, \dots$ ,

$$y^{(n)}(\theta) = nk_n(\theta) + \theta k_{n+1}(\theta).$$

*Proof.* In the case  $n = 1$ , the result is identical to (2.8). We complete the proof by induction on  $n$ . Assuming its truth for lower values we verify the result in the case when  $n$  is replaced by  $n + 1$ .

By (2.10) and the induction hypothesis, we have

$$g_n(\theta) = nk_n(\theta) + \theta k_{n+1}(\theta) - \theta f'(y(\theta))(y^{(n)}(\theta)).$$

Differentiating this result, substituting into (2.11) and simplifying we find

$$g_{n+1}(\theta) = (n + 1)k_{n+1}(\theta) + \theta k_{n+2}(\theta) - \theta f'(y(\theta))(y^{(n+1)}(\theta))$$

so that using (2.10) with  $n$  replaced by  $n + 1$  we obtain

$$y^{(n+1)}(\theta) = (n + 1)k_{n+1}(\theta) + \theta k_{n+2}(\theta)$$

and the proof is complete.

We will wish to have formulae for  $k_1, k_2, \dots$  and it turns out to be convenient to describe these functions making use of partitions. By a partition of  $n$ , we will mean a positive integral valued function  $\pi$  on  $\{1, 2, \dots, \sigma(\pi)\}$ , where  $\sigma(\pi)$  depends

on  $\pi$ , such that  $\pi$  is nondecreasing and such that  $\sum_{i=1}^{\sigma(\pi)} \pi(i) = n$ . Associated with  $\pi$ , there is a function  $\bar{\pi}$  on  $\{1, 2, \dots\}$  to  $\{0, 1, 2, \dots\}$  such that  $\bar{\pi}(i)$  is zero except for a finite number of values of  $i$  and such that  $\sum_{i=1}^{\infty} \bar{\pi}(i)i = n$ . In fact,  $\bar{\pi}(i)$  is defined as the number of members of  $\{1, 2, \dots, \sigma(\pi)\}$  which map onto  $i$  under  $\pi$ . For  $\pi$  a partition of  $n$ , let  $C(\pi)$  denote the integer

$$C(\pi) = n! / \prod_{i=1}^{\infty} ((i!)^{\bar{\pi}(i)} \bar{\pi}(i)!)$$

and let  $f_{\pi}(y)$  denote the function  $E \rightarrow X$  defined by

$$f_{\pi}(y)(\theta) = f^{(\sigma(\pi))}(y(\theta))(y^{(\pi(1))}(\theta), y^{(\pi(2))}(\theta), \dots, y^{(\pi(\sigma(\pi)))}(\theta)).$$

Let  $P_n$  denote the set of partitions of  $n$ . We will now prove the following result.

LEMMA 2.5. For  $n = 1, 2, \dots$ ,

$$k_{n+1} = \sum_{\pi \in P_n} C(\pi) f_{\pi}(y).$$

*Proof.* The result is clear in the case  $n = 1$ . For  $n > 1$  and  $\pi \in P_n$  such that  $\pi(1) = c_1, \pi(2) = c_2, \dots, \pi(\sigma(\pi)) = c_{\sigma(\pi)}$  define  $\pi^0 \in P_{n+1}$  such that  $\sigma(\pi^0) = \sigma(\pi) + 1$  and such that  $\pi^0(1) = 1, \pi^0(2) = c_1, \dots, \pi^0(\sigma(\pi^0)) = c_{\sigma(\pi)}$ . Also, for  $i = 1, 2, \dots, \sigma(\pi)$ , define  $\pi^i \in P_{n+1}$  such that  $\sigma(\pi^i) = \sigma(\pi)$  and such that  $\pi^i$  maps  $1, 2, \dots, \sigma(\pi)$ , onto  $c_1, c_2, \dots, c_i + 1, \dots, c_{\sigma(\pi)}$  (possibly rearranged into nondecreasing order). Computing the derivative of  $f_{\pi}(y)$  we find

$$\begin{aligned} f_{\pi}(y)'(\theta) &= f^{(\sigma(\pi)+1)}(y(\theta))(y^{(\pi(1))}(\theta), \dots, y^{(1)}(\theta)) \\ &+ f^{(\sigma(\pi))}(y(\theta))(y^{(\pi(1)+1)}(\theta), \dots, y^{(\pi(\sigma(\pi)))}(\theta)) \\ &+ \dots + f^{(\sigma(\pi))}(y(\theta))(y^{(\pi(1))}(\theta), \dots, y^{(\pi(\sigma(\pi))+1)}(\theta)) \end{aligned}$$

so that

$$f_{\pi}(y)' = \sum_{i=0}^{\sigma(\pi)} f_{\pi^i}(y).$$

It now follows that  $k_{n+1}$  is indeed a linear combination of functions of the form  $f_{\pi}(y)$  with  $\pi \in P_n$ . It remains to verify that the coefficients are as the lemma states. Assuming (we are again using an inductive argument) that the result of the lemma is true for  $k_1, k_2, \dots, k_n$  we find

$$k_{n+1} = \sum_{p \in P_{n-1}} C(p) \sum_{i=0}^{\sigma(p)} f_{p^i}(y).$$

If  $\pi \in P_n$  and  $p, i$  are such that  $p^i = \pi$ , then either  $i = 0$  and  $\bar{p}(j) = \bar{\pi}(j)$  for  $j \neq 1, \bar{p}(1) + 1 = \bar{\pi}(1)$  or else, for some positive  $l, \bar{p}(j) = \bar{\pi}(j)$  for  $j$  not equal to either  $l$  or  $l + 1, \bar{p}(l) = \bar{\pi}(l) + 1, \bar{p}(l + 1) + 1 = \bar{\pi}(l + 1)$ .

In the second case, there will be  $\bar{p}(l)$   $i$  values for which  $p^i = \pi$  and

$$\begin{aligned} C(p) &= (n - 1)! / \prod_{i=1}^{\infty} (i!^{\bar{p}(i)} \bar{p}(i)!) \\ &= \frac{\bar{\pi}(l + 1) \cdot (l + 1)}{\bar{p}(l) \cdot n} C(\pi). \end{aligned}$$

The total contribution from terms derived from  $p$  will thus be

$$\frac{\bar{\pi}(l + 1) \cdot (l + 1)}{n} C(\pi).$$

A similar calculation for the first case can be made and we find, for the total coefficient of  $f_{\pi}(y)$  in  $k_{n+1}$ ,

$$C(\pi) \left( \frac{\bar{\pi}(1)}{n} + \frac{2\bar{\pi}(2)}{n} + \dots \right) = C(\pi).$$

This completes the inductive argument and the proof of the lemma.

We have proved that, under certain assumptions on  $f$ ,  $y: E \rightarrow X$  is differentiable arbitrarily often. We now prove that on compact subsets of  $E$ , the various derivatives are bounded.

LEMMA 2.6. *Let  $I = [-\Theta, \Theta] \subset E$  be a closed interval, then there are constants  $c_1, c_2, \dots$ , such that, for  $n = 1, 2, \dots$  and  $\theta \in I$ ,*

$$\|y^{(n)}(\theta)\| \leq c_n.$$

*Proof.* For the case  $n = 1$ , we have

$$y^{(1)}(\theta) = f(y(\theta)) + \theta f'(y(\theta))(y^{(1)}(\theta)).$$

Hence,  $\|y^{(1)}(\theta)\| \leq \|f\| + |\theta| \cdot \|f'(y(\theta))\| \cdot \|y^{(1)}(\theta)\|$  so that

$$\|y^{(1)}(\theta)\| \leq \|f\| / (1 - \Theta L),$$

for  $\theta \in I$ , where we have written  $\|f\| = \sup_{x \in X} \|f(x)\|$ . (We shall, of course, also write  $\|f^{(1)}\|, \dots$  for  $\sup_{x \in X} \|f^{(1)}(x)\|, \dots$ .)

Having established the result for  $n = 1$ , we again use an inductive argument and assume its truth for  $i = 1, 2, \dots, n - 1$ , where  $n > 1$ . Let  $\pi_0 \in P_n$  be defined by  $\sigma(\pi_0) = 1, \pi_0(1) = n$ . Using Lemmas 2.4, 2.5, we find

$$(2.12) \quad y^{(n)}(\theta) = n \sum_{\pi \in P_{n-1}} C(\pi) f_{\pi}(y)(\theta) + \theta \sum_{\pi \in P_n \setminus \{\pi_0\}} C(\pi) f_{\pi}(y)(\theta) + \theta f'(y(\theta))(y^{(n)}(\theta)).$$

For  $\pi \in P_{n-1} \cup (P_n \setminus \{\pi_0\})$ , we have

$$\begin{aligned} \|f_{\pi}(y)(\theta)\| &= \|f^{(\sigma(\pi))}(y(\theta))(y^{(\pi(1))}(\theta), \dots, y^{(\pi(\sigma(\pi)))}(\theta))\| \\ &\leq \|f^{(\sigma(\pi))}\| c_{\pi(1)} c_{\pi(2)} \dots c_{\pi(\sigma(\pi))} = d_{\pi}, \quad \text{say.} \end{aligned}$$

Substituting this result into (2.12), we find

$$\|y^{(n)}(\theta)\| \leq n \sum_{\pi \in P_{n-1}} C(\pi) d_{\pi} + |\theta| \sum_{\pi \in P_n \setminus \{\pi_0\}} C(\pi) d_{\pi} + |\theta| L \|y^{(n)}(\theta)\|$$

so that

$$\|y^{(n)}(\theta)\| < \frac{1}{1 - \Theta L} \left( n \sum_{\pi \in P_{n-1}} C(\pi) d_{\pi} + \Theta \sum_{\pi \in P_n \setminus \{\pi_0\}} C(\pi) d_{\pi} \right).$$

It is possible to transform the formula for  $y^{(n)}(\theta)$  in the special case when  $\theta = 0$  into a much simpler form than that given by Lemmas 2.4, 2.5. In doing this, it will be convenient to make use of some of the terminology concerning trees that was

introduced in [4] and we now extend this terminology by introducing a notational device convenient for dealing with certain iterated products of trees. If  $(t_1, t_2, \dots, t_s)$  is an  $s$ -tuple of trees and  $t_0$  is any tree, then we define the product

$$t_0(t_1, t_2, \dots, t_s) \text{ as } (\dots ((t_0 t_1)t_2)t_3 \dots)t_s.$$

That is, this new product satisfies the recurrence relation

$$t_0(t_1, t_2, \dots, t_s) = (t_0 t_1)(t_2, \dots, t_s)$$

with  $t_0(t_1) = t_0 t_1$ . Because of the rule  $(t_0 t_1)t_2 = (t_0 t_2)t_1$  (see [4]), it is clear that

$$t_0(t_1, t_2, \dots, t_s) = t_0(t'_1, t'_2, \dots, t'_s),$$

where  $t'_1, t'_2, \dots, t'_s$  is any permutation of  $t_1, t_2, \dots, t_s$ .

For  $t \in T$  we now define  $F(t) \in X$  by the recurrence formulae

$$F(\tau) = f(0_X),$$

$$F(\tau(t_1, t_2, \dots, t_s)) = f^{(s)}(0_X)(F(t_1), F(t_2), \dots, F(t_s)),$$

where  $0_X$  is the zero element of  $X$ . Next we define an integer valued function  $\delta$  on  $T$  as follows:

$$\delta(\tau) = 1,$$

$$\delta(\tau(t_1, t_2, \dots, t_s)) = n_1! n_2! \dots n_s! \delta(t_1) \delta(t_2) \dots \delta(t_s),$$

where it is supposed that  $t_1, t_2, \dots, t_s$  are distinct but that, for  $i = 1, 2, \dots, s$ ,  $t_i$  occurs  $n_i$  times in  $t_1, t_2, \dots, t_s$ . We are now in a position to prove the promised simple formula for  $y^{(n)}(0)$ .

**THEOREM 2.7.** For  $n = 1, 2, \dots$ ,

$$y^{(n)}(0) = n! \sum_{r(t)=n} (F(t)/\delta(t)).$$

*Proof.* Note that  $r(t)$  denotes the order of  $t$  and  $\sum_{r(t)=n}$  denotes a summation over all  $t \in T$  such that  $r(t) = n$ .

For  $n = 1$ , the result is trivial. We assume now that  $n > 1$ , that the result is true for  $y^{(1)}(0), \dots, y^{(n-1)}(0)$  and we prove the result for  $y^{(n)}(0)$  to complete an inductive argument. We have

$$\begin{aligned} y^{(n)}(0) &= n \sum_{\pi \in P_{n-1}} C(\pi) f_{\pi}(y)(0) \\ &= n! \sum_{\pi \in P_{n-1}} \left( \prod_{i=1}^{\infty} (i!)^{\bar{\pi}(i)} \bar{\pi}(i)! \right)^{-1} f^{(\sigma(\pi))}(0_X)(y^{(\pi(1))}(0), \dots, y^{(\pi(\sigma(\pi)))}(0)). \end{aligned}$$

That is,

$$\begin{aligned} (2.13) \quad y^{(n)}(0) &= n! \sum_{\pi \in P_{n-1}} \sum_1 \sum_2 \dots \sum_{\sigma(\pi)} \left( \prod_{i=1}^{\infty} \bar{\pi}(i)! \right)^{-1} \\ &\quad \cdot f^{(\sigma(\pi))}(0_X) \left( \frac{F(t_1)}{\delta(t_1)}, \frac{F(t_2)}{\delta(t_2)}, \dots, \frac{F(t_{\sigma(\pi)})}{\delta(t_{\sigma(\pi)})} \right), \end{aligned}$$

where for  $j = 1, 2, \dots, \sigma(\pi)$ ,  $\sum_j$  denotes a summation of  $t_j$  over all trees such that  $r(t_j) = \pi(j)$ .

This is a linear combination of the  $F(t)$  with  $r(t) = n$ . To find the coefficient of  $F(t)$ , suppose that  $t = \tau(\bar{i}_1, \bar{i}_2, \dots, \bar{i}_s)$  where amongst  $\bar{i}_1, \bar{i}_2, \dots, \bar{i}_s$ , there are  $n_1$  trees with order 1,  $n_2$  with order 2 and so on. Also suppose that amongst the  $n_i$  trees of order  $i$  ( $i = 1, 2, \dots$ ),  $n_{i,1}$  are the same of one kind,  $n_{i,2}$  are the same of a second kind and so on.  $F(t)$  is equal to

$$f^{(\sigma(\pi))}(0_X)(F(t_1), F(t_2), \dots, F(t_{\sigma(\pi)}))$$

if and only if  $\sigma(\pi) = s$  and  $t_1, t_2, \dots, t_s$  are the same as  $\bar{i}_1, \bar{i}_2, \dots, \bar{i}_s$  in some order. Thus,  $\pi$  in (2.13) must be such that  $\bar{\pi}(i) = n_i$  for  $i = 1, 2, \dots$ . The number of choices of the  $n_i$  trees of order  $i$  amongst  $t_1, t_2, \dots, t_s$  which correspond to the trees of order  $i$  amongst  $\bar{i}_1, \bar{i}_2, \dots, \bar{i}_s$  is given by the multinomial coefficient  $n_i!/(n_{i,1}!n_{i,2}!\dots)$ . Hence, the coefficient of  $F(t)$  in (2.13) is equal to

$$n! \prod_{i=1}^{\infty} \frac{1}{n_i!} \prod_{i=1}^{\infty} \frac{n_i!}{n_{i,1}! n_{i,2}! \dots} \prod_{i=1}^s \frac{1}{\delta(t_i)} = \frac{n!}{\delta(t)}$$

and the inductive proof is complete.

**3. Convergence Criteria.** Let  $m_0 = (H_0, a_0, h_{01})$  be a given integration method and let  $M$  be a sequence of methods given by

$$\begin{aligned} M(1) &= m_1 = (H_1, a_1, h_{11}), \\ M(2) &= m_2 = (H_2, a_2, h_{21}), \\ &\vdots \qquad \qquad \qquad \vdots \end{aligned}$$

When we consider such collections of methods as  $m_0, m_1, m_2, \dots$ , there is no loss of generality in supposing that any pair of sets from  $H_0, H_1, \dots$  is disjoint. We shall always assume this is the case.

*Definition 3.1.*  $M$  is a bounded sequence if  $\{\|a_1\|, \|a_2\|, \dots\}$  is a bounded set. In this paper, we will consider only bounded sequences.

As in [4], we will consider, for a given method  $m = (H, a, h_1)$ , equations of the form

$$(3.1) \qquad y = \eta_0 e_H + a \times (f \circ y),$$

$$(3.2) \qquad \eta_1 = y(h_1).$$

In (3.1),  $f$  is a Lipschitz continuous function on  $X$  to  $X$  (where  $X$  is a finite-dimensional real vector space),  $a$  is a continuous linear operator on  $B(H)$  to  $B(H)$  and the ‘‘unknown’’  $y$  is a bounded function on  $H$  to  $X$ . By the contraction mapping theorem,  $y$  is defined by (3.1) if  $\|a\| L < 1$  where  $L$  is the Lipschitz constant for  $f$ .

A triple  $(X, f, \eta_0)$  will specify a ‘‘problem’’ in the same way as  $(H, a, h_1)$  specifies a method. For a method  $m$  and a problem  $(X, f, \eta_0)$ , we will write  $m \cdot (X, f, \eta_0)$  for the numerical solution obtained, that is  $\eta_1$ .

*Definition 3.2.* A bounded sequence  $M$  converges strongly to a method  $m_0$  if for all problems  $(X, f, \eta_0)$  such that  $f$  satisfies a Lipschitz condition with constant  $L$  satisfying  $\sup\{\|a_i\| : i = 0, 1, \dots\} L < 1$  then

$$\|m_i \cdot (X, f, \eta_0) - m_0 \cdot (X, f, \eta_0)\| \rightarrow 0 \quad \text{as } i \rightarrow \infty.$$

*Definition 3.3.* A bounded sequence  $M$  converges weakly to a method  $m_0$  if for all  $t \in T$ ,



$$w(m_i)(t) \rightarrow w(m_0)(t)$$

as  $i \rightarrow \infty$ . That is, for all  $t \in T$ ,

$$(3.3) \quad \nu_{a_i}(t)(h_{i1}) \rightarrow \nu_{a_0}(t)(h_{01}) \quad \text{as } i \rightarrow \infty.$$

LEMMA 3.4. *A bounded sequence  $M$  converges weakly to  $m_0$  if and only if for all  $t \in T$ ,*

$$(3.4) \quad \mu_{a_i}(t)(h_{i1}) \rightarrow \mu_{a_0}(t)(h_{01}) \quad \text{as } i \rightarrow \infty.$$

*Proof.* To prove sufficiency, we note that  $\nu_{a_i}(t) = \mu_{a_i}(\tau t)$ . To prove necessity we note that if  $t = \tau(t_1, t_2, \dots, t_s)$  then  $\mu_{a_i}(t) = \nu_{a_i}(t_1) \cdot \nu_{a_i}(t_2) \cdot \dots \cdot \nu_{a_i}(t_s)$ .

We now come to the main purpose of the paper which is to prove the following result.

THEOREM 3.5. *A bounded sequence  $M$  converges strongly to a method  $m_0$  if and only if  $M$  converges weakly to  $m_0$ .*

The proof will be given in the next two sections. Throughout, we will be concerned with a fixed sequence  $M$  and a fixed method  $m_0$ . Since we are dealing with a bounded sequence, we can define  $a \in A(H)$ , where  $H = H_0 \cup H_1 \cup H_2 \cup \dots$ , by the equations

$$a(x) | H_i = a_i(x | H_i)$$

for all  $x \in B(H)$  and all  $i = 0, 1, 2, \dots$ .

It is trivial that  $(H, a, h_i)$  is equivalent (in the sense of [4]) to  $(H_i, a_i, h_{i1})$  for all  $i = 0, 1, 2, \dots$ . For convenience, we will make use of the representation involving  $H$  and  $a$  in the proof which follows.

**4. Proof of Sufficiency.** In this section, we will show that weak convergence of  $M$  to  $m_0$  implies strong convergence. Suppose, on the contrary, that there exists a problem  $(X, f, \eta_0)$  where  $f$  has Lipschitz constant  $L$  satisfying  $\|a\|L < 1$ , such that there is a positive  $\epsilon$  with the property that the set

$$\{i: \|m_i \cdot (X, f, \eta_0) - m_0 \cdot (X, f, \eta_0)\| \geq \epsilon\}$$

is infinite.

Let  $y$  satisfy (3.1) and denote  $y(h_{i1})$  by  $y_i$  so that  $m_i \cdot (X, f, \eta_0) = y_i$  for all  $i = 0, 1, 2, \dots$ . Making use of Lemma 4.2 and Theorem 5.4 of [4], we can choose  $c: U \rightarrow R$ , where  $U$  is a finite subset of  $T$  such that

$$\left\| y - \sum_{t \in U} \mu_a(t)c(t) \right\| < \epsilon/4.$$

Hence, for all  $i = 1, 2, \dots$ , we have

$$\begin{aligned} \|y_i - y_0\| &< \frac{\epsilon}{2} + \left\| \sum_{t \in U} (\mu_a(t)(h_{i1}) - \mu_a(t)(h_{01}))c(t) \right\| \\ &\leq \frac{\epsilon}{2} + \sum_{t \in U} \|c(t)\| \cdot \max_{t \in U} \|\mu_a(t)(h_{i1}) - \mu_a(t)(h_{01})\|. \end{aligned}$$

If  $i$  is such that  $\|y_i - y_0\| \geq \epsilon$ , we therefore have

$$(4.1) \quad d \max_{t \in U} \|\mu_a(t)(h_{i1}) - \mu_a(t)(h_{01})\| > \frac{\epsilon}{2}$$

where  $d = \sum_{t \in U} \|c(t)\|$ . Thus, (4.1) holds for an infinite number of values of  $i$ . By Lemma 3.4, for each  $t \in U$  we can find  $n_t$  so that  $d \|\mu_a(t)(h_{i1}) - \mu_a(t)(h_{01})\| < \epsilon/2$  for all  $i > n_t$ . If  $n = \max_{t \in U} n_t$ , we thus see that, for all  $i > n$ ,

$$(4.2) \quad d \max_{t \in U} \|\mu_a(t)(h_{i1}) - \mu_a(t)(h_{01})\| < \frac{\epsilon}{2}.$$

Hence, the infinite set for which (4.1) holds is a subset of the finite set  $\{1, 2, \dots, n\}$ . This contradiction completes the proof.

**5. Proof of Necessity.** Conversely, we will prove in this section that weak convergence of  $M$  to  $m_0$  is implied by strong convergence.

Suppose on the contrary that  $M$  converges strongly to  $m_0$  but that for some  $t_0 \in T$ ,  $\nu_a(t_0)(h_{i1})$  does not converge to  $\nu_a(t_0)(h_{01})$  as  $i \rightarrow \infty$ . Let  $n$  be the order of  $t_0$ . We will now construct a triple  $(X, f, \eta_0)$  for which the condition of strong convergence is not satisfied. Let  $K$  be a finite set and  $(K, b, k_1)$  be a method such that for all  $t$  of order not exceeding  $n$  we have

$$\begin{aligned} \nu_b(t)(k_1) &= 0 & (t \neq t_0), \\ \nu_b(t_0)(k_1) &= 1. \end{aligned}$$

Note that the existence of  $(K, b, k_1)$  follows from Theorem 6.9 of [4]. The assumption that  $K$  is finite can be made here since the constructive proof of that theorem made use only of finite sets.

Let  $\varphi: R \rightarrow R$  be a bounded function such that all its derivatives exist and are bounded and such that

$$\varphi(0) = \varphi'(0) = \varphi''(0) = \dots = \varphi^{(n-1)}(0) = 1.$$

For example, we can choose  $\varphi$  as defined by

$$\varphi(x) = P(x)e^{-x^2},$$

where  $P$  is a suitably chosen polynomial.

We now choose  $X = B(K)$ ,  $\eta_0 = 0_x$  and we define  $f$  by  $f(x) = \theta_0 b(\varphi \circ x)$  for all  $x \in X$ . The real number  $\theta_0$  which occurs in the formula for  $f$  will be determined later. It is easy to see that  $f$  is infinitely differentiable with bounded derivatives of all orders. In fact, for  $m = 1, 2, \dots$ , the  $m$ th derivative  $f^{(m)}: X \rightarrow L(X^m, X)$  is defined by

$$f^{(m)}(x)(x_1, x_2, \dots, x_m) = \theta_0 b((\varphi^{(m)} \circ x) \cdot x_1 \cdot x_2 \cdot \dots \cdot x_m),$$

where the products in the argument of  $b$  are pointwise.

Let  $L$  be defined as the maximum of 1 and  $\|a\| \cdot \|b\| \cdot \sup_{z \in R} |\varphi'(z)|$  and let  $E$  be the open interval  $(-1/L, 1/L)$ .

Consider  $z: E \rightarrow B(H) \otimes B(K)$  defined by

$$(5.1) \quad z(\theta) = \theta(a \otimes b)(\varphi \circ z(\theta)).$$

To see that  $z$  is defined by (5.1), note that for any  $\theta \in E$  the function taking  $\zeta \in B(H) \otimes B(K)$  to  $\theta(a \otimes b)(\varphi \circ \zeta)$  is a contraction.

If  $\theta_0 \in E$  and  $y \in B(H, X)$  is defined by  $y(h)(k) = z(\theta_0)(h, k)$ , we observe that  $y$  satisfies (2.7) with  $\theta$  replaced by  $\theta_0$ . Also,  $z$  satisfies the conditions of  $y$  in Theorem 2.3. Hence,  $z$  is infinitely differentiable and its derivatives at zero can be found by

making use of Theorem 2.7. We have, for  $m = 1, 2, \dots$ ,

$$z^{(m)}(0) = m! \sum_{r(t)=m} F(t)/\delta(t)$$

where  $F(\tau) = (a \otimes b)(\varphi \circ 0_{B(H \times K)}) = (a \otimes b)(e_{H \times K}) = \nu_a(\tau) \otimes \nu_b(\tau)$  and if  $t = \tau(t_1, t_2, \dots, t_s)$  then

$$F(t) = (a \otimes b)((\varphi^{(s)} \circ 0_{B(H \times K)}) \cdot F(t_1) \cdot F(t_2) \cdot \dots \cdot F(t_s)) = \nu_a(t) \otimes \nu_b(t),$$

where  $s < m = r(t) \leq n$ .

We now compute  $z^{(m)}(0)(h, k_1)$  for all  $h \in H$  and  $m \leq n$  making use of the special property for which  $b$  was chosen. We have

$$z^{(m)}(0)(h, k_1) = 0 \quad (m < n),$$

$$z^{(n)}(0)(h, k_1) = n! \nu_a(t_0)(h)/\delta(t_0).$$

Let  $\Theta = 1/2L$ ,  $I = [-\Theta, \Theta] \subset E$ . By Lemma 2.6,  $\|z^{(n+1)}(\theta)(h, k_1)\|$  is bounded for  $\theta \in I$  and  $h \in H$ . Let  $(n + 1)!c$  bound this quantity.

For  $\theta \in I$  we have by Taylor's theorem

$$|z(\theta)(h, k_1) - \theta^n \nu_a(t_0)(h)/\delta(t_0)| \leq c |\theta|^{n+1}$$

for all  $h \in H$ . Hence,

$$|z(\theta)(h_{01}, k_1) - z(\theta)(h_{i1}, k_1)| \geq |\theta|^n |\nu_a(t_0)(h_{01}) - \nu_a(t_0)(h_{i1})| - 2c |\theta|^{n+1}.$$

Since  $\nu_a(t_0)(h_{i1})$  does not converge to  $\nu_a(t_0)(h_{01})$  there is an infinite set  $\{i_1, i_2, \dots\} = S$  say such that

$$\inf_{i \in S} |\nu_a(t_0)(h_{01}) - \nu_a(t_0)(h_{i1})| > 0.$$

Let  $d$  denote this infimum. We now select  $\theta_0 \in I$  such that  $2c |\theta_0| < d$  and  $|\theta_0| > 0$  so that for all  $i \in S$ ,

$$\begin{aligned} |z(\theta_0)(h_{01}, k_1) - z(\theta_0)(h_{i1}, k_1)| &\geq |\theta_0|^n d - 2c |\theta_0|^{n+1} \\ &= C, \text{ say} \\ &> 0. \end{aligned}$$

That is, for  $i \in S$ ,

$$\|y(h_{i1}) - y(h_{01})\| \geq |y(h_{i1})(k_1) - y(h_{01})(k_1)| \geq C:$$

Since  $S$  is infinite, this contradicts the assumption that  $y(h_{i1}) \rightarrow y(h_{01})$  as  $i \rightarrow \infty$ . This completes the proof.

Mathematics Department  
The University of Auckland  
Auckland, New Zealand

1. G. DAHLQUIST, "Convergence and stability in the numerical integration of ordinary differential equations," *Math. Scand.*, v. 4, 1956, pp. 33-53. MR 18, 338.
2. P. HENRICI, *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York, 1962. MR 24 #B1772.
3. J. C. BUTCHER, "On the convergence of numerical solutions to ordinary differential equations," *Math. Comp.*, v. 20, 1966, pp. 1-10. MR 32 #6678.
4. J. C. BUTCHER, "An algebraic theory of integration methods," *Math. Comp.*, v. 26, 1972, pp. 79-106.