

Computational Complexity of One-Step Methods for Systems of Differential Equations*

By Arthur G. Werschulz**

Abstract. The problem is to calculate an approximate solution of an initial value problem for an autonomous system of N ordinary differential equations. Using fast power series techniques, we exhibit an algorithm for the p th-order Taylor series method requiring only $O(p^N \ln p)$ arithmetic operations per step as $p \rightarrow +\infty$. (Moreover, we show that any such algorithm requires at least $O(p^N)$ operations per step.) We compute the order which minimizes the complexity bounds for Taylor series and linear Runge-Kutta methods and show that in all cases this optimal order increases as the error criterion ϵ decreases, tending to infinity as ϵ tends to zero. Finally, we show that if certain derivatives are easy to evaluate, then Taylor series methods are asymptotically better than linear Runge-Kutta methods for problems of small dimension N .

1. Introduction. Let \mathcal{D} be a set of points in the real N -dimensional linear space \mathbf{R}^N , and let \mathcal{V} be a set of operators on \mathbf{R}^N , such that the *initial value problem* of finding a function $x: [0, 1] \rightarrow \mathbf{R}^N$ satisfying

$$(1.1) \quad \begin{aligned} \dot{x}(t) &= v(x(t)) \quad \text{if } 0 < t < 1, \\ x(0) &= x_0, \end{aligned}$$

has a unique solution for every $(x_0, v) \in \mathcal{D} \times \mathcal{V}$; we assume that x is analytic on $[0, 1]$. (The autonomous form of this system is no restriction, since any nonautonomous system may be made autonomous by increasing the dimension of the system by one; however, see the comment after Theorem 3.1.)

In Werschulz [1976], we looked at the computational complexity of using *one-step methods* to generate an approximate solution to (1.1) on an *equidistant grid* in the sense of Stetter [1973]; that is, the methods considered computed approximations x_i to $x(ih)$ by the recursion

$$(1.2) \quad x_{i+1} = x_i + h\varphi(x_i, h) \quad (0 \leq i \leq n-1, n = h^{-1}),$$

where $h = n^{-1}$ is the *step-size* of a grid with n points, and φ is the *increment function* (Henrici [1962]) for the method. (To be brief, we will refer to "the method φ .") In

Received January 4, 1977; revised December 5, 1978.

1980 *mathematics subject classification*. Primary 65L05, 68C05, 68C25; Secondary 34-04, 68C20.

*This research was supported in part by the National Science Foundation under Grant MCS75-222-55 and the Office of Naval Research under Contract N00014-76-C-0370, NR 044-422.

***Author's present address*: Department of Mathematics, University of Maryland, Baltimore County, Baltimore, Maryland 21228.

that paper, we discussed the problem of optimal order and minimal complexity for rather general classes of one-step methods.

In this paper, we will use the techniques and results of Werschulz [1976] to analyze the complexity of using Taylor series methods and linear Runge-Kutta methods to generate approximate solutions whose error does not exceed ϵ . The model of computation, error measure, and complexity measure to be used are described in Section 2, as well as the relevant results from Werschulz [1976].

We discuss the complexity of Taylor series methods in Section 3. Using the fast power series techniques of Brent and Kung [1978], we show that $O(p^N \ln p)$ arithmetic operations suffice to compute the p th-order Taylor series approximation; moreover, we show that $O(p^N)$ operations are necessary. In Section 4, we discuss the complexity of linear Runge-Kutta methods. In both sections, we compute lower and upper bounds on the complexity using a fixed method of given order; these results are then used to compute optimal orders which minimize these complexity bounds. We show that in all cases, the optimal order increases as ϵ decreases, tending to infinity as ϵ tends to zero.

Finally, we compare these two classes of methods in Section 5, where we show that if the partial derivatives of v are easy to evaluate, then Taylor series methods are asymptotically better (as ϵ tends to zero) than linear Runge-Kutta methods for problems of small dimension N .

2. Preliminary Results. Before proceeding any further, we will establish some notational conventions. Let X be an ordered ring; then X^+ and X^{++} respectively denote the nonnegative and positive elements of X . (This is used in the cases $X = \mathbf{R}$, the real numbers, and $X = \mathbf{Z}$, the integers.) The symbol “:=” means “is defined to be.” We use “ I ” to denote the unit interval $[0, 1]$. The symbol “ ∇ ” is used to denote the gradient of a mapping. The notations “ $x \downarrow a$ ” and “ $x \uparrow a$ ” are used to indicate one-sided limits.

We next describe the model of computation to be used. We assume only that all arithmetic operations are performed exactly in \mathbf{R} (i.e., infinite-precision arithmetic) and that for any algorithm to be considered for the solution of (1.1), a set of procedures is given for the computation of any information about v required by that algorithm. (For instance, with Runge-Kutta methods, we must be able to compute v at any point in its domain.)

In addition, we must pick an error measure, so that we may measure the discrepancy between the approximate solution produced by φ (via (1.2)) and the true solution. For the sake of definiteness, we use the *global error*

$$(2.1) \quad \sigma(\varphi, h) := \max_{0 \leq i \leq n} \|x(ih) - x_i\|,$$

where $\|\cdot\|$ is a norm on \mathbf{R}^N . Other error measures may be used such as the *local error per step* and the *local error per unit step* (see Henrici [1962] and Stetter [1973] for definitions); this would involve only a slight modification of the results contained in the sequel.

We now describe the complexity measure to be used. Let $\Phi = \{\varphi_p : p \in \mathbf{Z}^{++}\}$ be a *basic sequence* in the sense of Werschulz [1976]; that is, we may write

$$(2.2) \quad \sigma(\varphi_p, h) = \kappa(p, h)h^p \quad \text{for } h \in I \text{ and } p \in \mathbf{Z}^{++},$$

where

$$(2.3) \quad 0 < \kappa_1(p) \leq \kappa(p, h) \leq \kappa_2(p) < +\infty \quad \text{for } h \in I.$$

We say that φ_p has *order* p . This is a slight extension of the definition of order given in Cooper and Verner [1972]; the function κ_1 introduced here is necessary and sufficient for the “order” of a method to be unique. *Then we will be interested in the total number of arithmetic operations $C(p, \alpha)$ required to guarantee that*

$$(2.4) \quad \sigma(\varphi_p, h) \leq \epsilon := e^{-\alpha},$$

for a given p and a given α . (Here e is the base of the natural logarithms.) We suppose that $0 < \epsilon < 1$, so that α is positive. Clearly, α increases as ϵ decreases, and α tends to infinity as ϵ tends to zero.

In the methods we consider, we may write

$$(2.5) \quad C(p, \alpha) = nc(p),$$

where n is the minimal number of steps required and the *cost per step* $c(p)$ is the number of arithmetic operations required for the method of order p . As in Traub and Woźniakowski [1976], we shall express the cost per step associated with φ_p in the form

$$(2.6) \quad c(p) := u(\mathfrak{N}_p(v)) + d(p).$$

Here $\mathfrak{N}_p(v)$ is the *information* about v required to perform one step of φ_p , and we write $u(\mathfrak{N}_p(v))$ for the *informational cost* of φ_p ; we call $d(p)$ the *combinatory cost* of φ_p .

Note that we explicitly indicate the dependence of \mathfrak{N}_p on v , so that we may compare the cost of (say) an evaluation of v with a scalar arithmetic operation. Basically, $u(\mathfrak{N}_p(v))$ measures the cost of getting new data about v required by φ_p , while $d(p)$ measures the cost of combining this new data to get an approximate value of the solution at a new point. For example, Euler’s method in \mathbf{R}^N ,

$$x_{i+1} = x_i + hv(x_i),$$

has informational cost $\sum_{i=1}^N u(v_i)$, where v_1, \dots, v_N are the components of v , and for any function $\omega: \mathbf{R}^N \rightarrow \mathbf{R}$, we define

$$(2.7) \quad u(\omega) := \text{cost of evaluating } \omega \text{ at one point},$$

which we assume to be independent of the point of evaluation. The combinatory cost is $2N$ arithmetic operations, i.e., one scalar multiplication and one scalar addition for each of the N components.

We must now face a problem that occurs in almost all areas of complexity

theory. The number of operations $c(p)$ required for one step of a p th-order method is usually unknown per se; we only have bounds of the form

$$(2.8) \quad c_1(p) \leq c(p) \leq c_2(p).$$

That is, $c_1(p)$ is a *lower bound* on the number of operations required per step, usually derived via *theoretical* considerations, and $c_2(p)$ is an *upper bound* on the number of operations required per step, which is derived by exhibiting an *algorithm* for computing the p th-order method. (In what follows, we shall assume that the functions c_i , $\kappa_i: \mathbf{R}^+ \rightarrow \mathbf{R}^+$ ($i = 1, 2$) are analytic and that $\lim_{p \rightarrow 0} \kappa_i(p)^{1/p} > 0$, although this requirement may be greatly weakened. However, this assumption holds for all examples that we consider.)

From the discussion in Section 3 of Werschulz [1976], we find that the step-size h must satisfy

$$(2.9) \quad h_2(p, \alpha) \leq h \leq h_1(p, \alpha),$$

where

$$(2.10) \quad h_i(p, \alpha) := \kappa_i(p)^{1/p} e^{-\alpha/p} \quad (i = 1, 2).$$

Using (2.5), (2.8), (2.9), and (2.10), we may find bounds on the complexity $C(p, \alpha)$.

THEOREM 2.1. *Define (for $i = 1, 2$)*

$$C_i(p, \alpha) := f_i(p) e^{\alpha/p}, \quad \text{where } f_i(p) := \kappa_i(p)^{1/p} c_i(p).$$

Then

$$(2.11) \quad C_1(p, \alpha) \leq C(p, \alpha) \leq C_2(p, \alpha).$$

Proof. See Theorem 3.1 of Werschulz [1976]. \square

Thus, we have bounds on the complexity of using φ_p to compute an approximate solution satisfying (2.4). We now wish to consider the problem of optimality. Define

$$(2.12) \quad C^*(\alpha) := \inf\{C(p, \alpha): \varphi_p \in \Phi\}.$$

We are interested in bounds for $C^*(\alpha)$ under reasonable assumptions about f_1 and f_2 .

We first suppose that (for $i = 1, 2$)

$$(2.13) \quad f_i(p) > 0 \quad \text{if } p > 0$$

and

$$(2.14) \quad \lim_{p \uparrow \infty} f_i(p) = +\infty.$$

Assumption (2.13) is that there is no method whose cost per step is zero, while (2.14) essentially means that the "better" a method is (i.e., the higher its order is), the more we should expect to pay for its use.

Using the techniques of elementary calculus, we find that a necessary condition

for p to minimize $C_i(\cdot, \alpha)$ is that

$$(2.15) \quad \alpha = G_i(p) := p^2 f_i'(p) / f_i(p).$$

Sufficient conditions for the existence and uniqueness of a solution to (2.15) which actually minimizes $C_i(\cdot, \alpha)$ are given in

LEMMA 2.1. *Let f_i satisfy (2.13), (2.14), and*

$$(2.16) \quad G_i'(p) > 0 \quad \text{if } G_i(p) > 0.$$

Then (2.15) has a unique solution $p_i^(\alpha)$, and*

$$(2.17) \quad C_i^*(\alpha) := C_i(p_i^*(\alpha), \alpha) \leq C_i(p, \alpha) \quad \text{for } p \in \mathbf{R}^{++},$$

with equality in (2.18) or (2.19) if and only if $p = p_i^(\alpha)$.*

Proof. See Theorem 2.1 and Lemma 3.1 of Werschulz [1976]. \square

We call $p_1^*(\alpha)$ (respectively, $p_2^*(\alpha)$) the *lower (upper) optimal order*, $C_1^*(\alpha)$ (respectively, $C_2^*(\alpha)$) the *lower (upper) optimal complexity*, and

$$(2.18) \quad h_1^*(\alpha) := h_1^*(p_1^*(\alpha), \alpha) \quad (\text{respectively, } h_2^*(\alpha) := h_2(p_2^*(\alpha), \alpha))$$

the *lower (upper) optimal step-size*. Combining (2.11), (2.12), and Lemma 2.1, we have

THEOREM 2.2.

$$C_1^*(\alpha) \leq C^*(\alpha) \leq C_2^*(\alpha). \quad \square$$

We next describe the behavior of these quantities as α increases and tends to infinity.

THEOREM 2.3. *Let f_i be as in Lemma 2.1. Then $p_i^*(\alpha)$ and $C_i^*(\alpha)$ all increase monotonically and tend to infinity with α .*

Proof. See Theorems 2.2 and 3.3 of Werschulz [1976]. \square

Finally, we need a restriction of the problem class $\mathcal{D} \times V$ to "sufficiently difficult" problems; this will allow us to determine κ_1 and thus establish lower bounds. We will assume that

$$(2.19) \quad \sigma(\varphi_p, h) \geq (M_1 h)^p \quad \text{if } h \in I \text{ and } p \in \mathbf{Z}^{++}$$

for some $M_1 > 0$ independent of h and p . In the methods we study, (2.23) holds provided all sharp upper bounds are attained.

3. Taylor Series Methods. The class Φ_T of *Taylor series methods* is defined by taking a truncated Taylor series as an approximate solution. From (1.1), we see that the increment function φ_p is given by

$$(3.1) \quad \varphi_p(x_p, h) := \sum_{k=0}^{p-1} v^{(k)}(x_i) \frac{h^k}{(k+1)!},$$

where (using the notation of Henrici [1962])

$$(3.2) \quad v^{(k)}(x_i) := \left(\frac{d}{dt} \right)^k v(x(t)) \Big|_{x(t)=x_i}.$$

The usual method of computing (3.2), as described in "classical" numerical analysis texts invokes the chain rule. This quickly leads to expressions of horrifying complexity; for this reason, most texts quickly abandon the discussion of high-order Taylor series methods.

We are interested in faster algorithms for computing φ_p . First, we address the problem of a lower bound for the combinatory cost $d(p)$ of φ_p .

THEOREM 3.1. *There exists $a_1 > 0$ such that*

$$(3.3) \quad d(p) \geq a_1 p^N \quad \text{for any } p \in \mathbf{Z}^{++}.$$

Proof. Any algorithm for computing φ_p requires the information

$$\mathfrak{N}_p(v) := \{D^\beta v : 0 \leq |\beta| \leq p-1\}.$$

We use the standard multi-index notation found in Friedman [1969]; that is, for $\beta = (\beta_1, \dots, \beta_N) \in (\mathbf{Z}^+)^N$

$$D^\beta v := \left(\frac{\partial}{\partial x_1} \right)^{\beta_1} \cdots \left(\frac{\partial}{\partial x_N} \right)^{\beta_N} v \quad \text{and} \quad |\beta| := \sum_{i=1}^N \beta_i.$$

It is then easy to see that the above set has $O(p^N)$ (as $p \uparrow \infty$) distinct elements, which are (generally) independent; this is an immediate consequence of Problem 11 in Chapter 1 of Pólya and Szegő [1925]. Thus, (3.3) gives a lower bound which is linear in the amount of information required. \square

Note that a_1 in (3.3) depends on N . Since we are treating the case where N is fixed and p is allowed to vary, we will not indicate this dependence explicitly. Also note that Theorem 3.1 need not hold if we cannot assume the independence of the derivatives of $v \in V$, i.e., if V satisfies some special property. In particular, the theorem need not hold for N -dimensional problems which arise by changing an $(N-1)$ -dimensional nonautonomous problem to an N -dimensional autonomous problem.

We now see how close we can get to an optimum value for $d(p)$.

THEOREM 3.2. *There exists a constant $a_2 > 0$ such that the combinatory cost $d(p)$ of computing $\varphi_p \in \Phi_T$ satisfies the bound*

$$(3.4) \quad d(p) \leq a_2 p^N \ln(p + e).$$

Proof. We first consider the case $N = 1$. Note that $x(h)$ is the zero of

$$(3.5) \quad F(z) := \int_{x_0}^z d\xi/v(\xi) - h.$$

As in Brent and Kung [1978], we consider the formal power series

$$P(s) := F(x_0 + s) - F(x_0),$$

where s is an indeterminate. Let V be the power series reversion of P . Adopting the notation of Brent and Kung [1978], we see that

$$x(s) = x_0 + V(s) = x_0 + V_p(s) + O(s^{p+1}).$$

By the uniqueness of the Taylor coefficients of an analytic function, we see that

$$\varphi_p(x_0, h) = h^{-1} V_p(h).$$

Since the number $V_p(h)$ can be computed in $O(p \ln p)$ operations from the Taylor coefficients of v (by Theorem 6.2 of Brent and Kung [1978]), the result for $N = 1$ follows.

For $N \geq 2$, we use Newton's method (Rall [1969]) applied to the formal power series operator P given by

$$(Py)(s) := y(s) - x_0 - \int_0^s v(y(\tau)) d\tau;$$

clearly, the formal power series $x(s)$ is the zero of P . The algorithm itself is defined recursively. Let a formal power series $x_{(p)}(s)$ satisfying $x_{(p)}(s) = x(s) + O(s^{p+1})$ be given. Precompute

$$(3.6) \quad w(s) := \int_0^s v(x_{(p)}(\tau)) d\tau - x_0 - x_{(p)}(s) + O(s^{2p+2}),$$

$$(3.7) \quad Q(s) := \nabla v(x_{(p)}(s)) + O(s^{2p+2}),$$

and let $u_{(0)}(s) := 0$. Then set

$$x_{(2p+1)}(s) := x_{(p)}(s) + u_{(p+1)}(s),$$

where

$$(3.8) \quad u_{(k+1)}(s) := \int_0^s Q(\tau) u_{(k)}(\tau) d\tau + w(s) + O(s^{2p+2}), \quad 0 \leq k \leq p.$$

Following the proof given in Rall [1969], we find that $x_{(2p+1)}(s) = x(s) + O(s^{2p+2})$.

We need only consider the cost $T(p, N)$ of computing the series $x_{(p)}(s)$ in determining $d(p)$, since $x(h)$ may be recovered from the formal power series in $O(p)$ operations. Clearly, we have the recursion

$$(3.9) \quad T(2p + 1, N) \leq T(p, N) + T_6 + T_7 + T_8,$$

where T_m is the cost of step (3.m) for $m = 6, 7, 8$. Let $\text{COMP}(p, N)$ be the time required to find the first p terms of the formal power series $f(y_1(s), \dots, y_N(s))$, where f, y_1, \dots, y_N are formal power series, and y_1, \dots, y_N have zero constant term. Theorem 7.1 of Brent and Kung [1978] states that

$$\text{COMP}(p, 2) = O(p^2 \ln p),$$

and it is easy to show that for any $N \in \mathbf{Z}^{++}$,

$$\text{COMP}(p, N + 1) = O(p \text{ COMP}(p, N)).$$

Thus, for $N \geq 2$ we have

$$(3.10) \quad \text{COMP}(p, N) = O(p^N \ln p);$$

and so, we see that $T_6 + T_7 = O((2p + 1)^N \ln p)$. Finally, let $\text{MULT}(p)$ be as in Brent and Kung [1978]; we see that

$$T_8 = (p + 1)[N^2 \text{MULT}(2p + 1) + O(p)] = O((2p + 1)^2 \ln p)$$

if Fast Fourier Transform multiplication (Borodin and Munro [1975]) is used. Since $N \geq 2$, we have

$$(3.11) \quad T_6 + T_7 + T_8 = O((2p + 1)^N \ln p),$$

and so (3.9) and (3.11) imply that $T(p, N) = O(p^N \ln p)$, which completes the proof. \square

(Note that the second algorithm is inferior to the first algorithm when applied to the scalar case $N = 1$, where we find that the second algorithm requires $O(p^2 \ln p)$ arithmetic operations.)

We now determine bounds on $C(p, \alpha)$. First, consider lower bounds. Clearly, there exists $u_1(v) \geq 0$ such that

$$(3.12) \quad u(D^\beta v_i) \geq u_1(v) \quad (1 \leq i \leq n, |\beta| \in \mathbf{Z}^+).$$

Since $\mathfrak{R}_p(v)$ has $O(p^N)$ elements, there exists a constant $b_1 > 0$ such that

$$(3.13) \quad u(\mathfrak{R}_p(v)) \geq b_1 u_1(v) p^N.$$

From (3.3) and (3.13), we have a lower-bound cost per step of

$$(3.14) \quad c_1(p) = [a_1 + b_1 e_1(v)] p^N.$$

This leads to

$$\text{THEOREM 3.3.} \quad C_1(p, \alpha) = M_1 [a_1 + b_1 e_1(v)] p^N e^{\alpha/p}.$$

Proof. This is an immediate consequence of (2.19) and (3.14). \square

Note that $f_1(p) := M_1 c_1(p)$ satisfies the conditions of Lemma 2.1. Thus, the optimality theory of Section 2 holds. In particular, we have

$$\text{THEOREM 3.4.} \quad C_1^*(\alpha) = M_1 [a_1 + b_1 e_1(v)] (e/N)^N \alpha^N.$$

Proof. From (2.17) and (3.14), we find that $G_1(p) = Np$, so that

$$p_1^*(\alpha) = \alpha/N \quad \text{and} \quad h_1^*(\alpha) = (M_1 e^N)^{-1}.$$

The result follows by letting $p = p_1^*(\alpha)$ in the definition of $C_1(p, \alpha)$. \square

Next, we turn to upper bounds on the complexity. Theorem 3.2 tells us how to combine the necessary information to get the solution at a new grid-point; we need only measure the cost of getting the information. So, let

$$u^{(k)}(v) = \max\{u(D^\beta v_i): 1 \leq i \leq N, |\beta| = k\}.$$

Using the result in Pólya and Szegő [1925], we see that

$$(3.15) \quad u(\mathfrak{N}_p(v)) \leq N \sum_{k=0}^{p-1} u^{(k)}(v) \binom{N+k-1}{k}.$$

Unfortunately, the right-hand side of (3.15) does not fit our general model, so we must assume that we know how $u^{(k)}(v)$ changes as k increases. *We will consider the case where the cost of derivative evaluation is bounded; that is, we will assume that*

$$(3.16) \quad u^{(k)}(v) \leq u_2(v)$$

for some $u_2(v)$ independent of k . Other cases (e.g., $u^{(k)}(v) = O(k^m)$ for some $m > 0$) may be analyzed in a similar manner; of course, they will give different results. By (3.15) and (3.16), there is a $b_2 > 0$ such that

$$(3.17) \quad u(\mathfrak{N}_p(v)) \leq b_2 u_2(v) p^N.$$

From (3.4) and (3.17), we have an upper-bound cost per step of

$$(3.18) \quad c_2(p) = a_2 p^N \ln(p + e) + b_2 u_2(v) p^N.$$

This leads to

THEOREM 3.5. *There exists an $M_2 > 0$ such that*

$$C_2(p, \alpha) = M_2 [a_2 p^N \ln(p + e) + b_2 u_2(v) p^N] e^{\alpha/p}.$$

Proof. By Cauchy's Integral Theorem there exists a $B > 0$ such that

$$\|x^{(k+1)}\| / (k + 1)! \leq B^k,$$

where we define

$$(3.19) \quad \|y\| := \max_{t \in I} \|y(t)\|$$

for any $y: I \rightarrow \mathbf{R}^N$. Thus by Section 3.3-3 of Henrici [1962], we see that a Lipschitz constant for φ_p in Φ_T is given by

$$\sum_{k=0}^{p-1} \|x^{(k+1)}\| h^k / (k + 1)! \leq \sum_{k=0}^{p-1} (Bh)^k \leq L := (1 - Bh_0)^{-1},$$

provided that $h \leq h_0 < B^{-1}$. By Sections 3.3-2 and 3.3-4 of Henrici [1962], there exists an $M_2 > 0$ such that $\sigma(\varphi_p, h) \leq (M_2 h)^p$. The result now follows from Theorem 4.1 and (3.18). \square

We are now ready to consider the optimal p for $C_2(p, \alpha)$.

THEOREM 3.6. (1) *For all $\alpha > 0$, there exists $p_2^*(\alpha)$ such that (2.17) holds.*

(2) *$p_2^*(\alpha)$ increases monotonically with α , and $p_2^*(\alpha) \sim \alpha/N$ as $\alpha \uparrow \infty$.*

(3) *$C_2^*(\alpha)$ increases monotonically with α , and $C_2^*(\alpha) \sim M_2 \alpha_2 (e/N)^N \alpha^N \ln \alpha$ as $\alpha \uparrow \infty$.*

(4) *$h_2^*(\alpha) \sim (M_2 e^N)^{-1}$ as $\alpha \uparrow \infty$.*

Proof. Clearly, c_2 satisfies (2.13) and (2.14). Now write

$$G_2(p) = G_{21}(p) + G_{22}(p),$$

where

$$G_{21}(p) = Np \quad \text{and} \quad G_{22}(p) = \nu p^2 / D_2(p);$$

here we set

$$D_2(p) := (p + e)[(p + e) \ln(p + e) + 1] \quad \text{and} \quad \nu := a_2 / [b_2 e_2(v)].$$

We see immediately that G_{21} satisfies (2.17); a straightforward calculation shows that

$$G'_{22}(p) = \nu \frac{vp [\ln(p + e) - 1] + 2e [\nu \ln(p + e) + 1]}{D(p)^2},$$

so that $G'_{22}(p) > 0$ for $p > 0$. Thus, G_{22} satisfies (2.16), which shows that G_2 satisfies (2.16). Hence, p_2^* and C_2^* behave as described in Theorem 2.3.

Since $p_2^*(\alpha)$ goes to infinity with α , we see that

$$\alpha = G_2(p_2^*(\alpha)) \sim Np_2^*(\alpha) + p_2^*(\alpha) / \ln p_2^*(\alpha) \sim Np_2^*(\alpha),$$

which gives the asymptotic estimate in (2). The rest of the theorem follows from this estimate. \square

Unfortunately, the estimates given above are only asymptotic as $\alpha \uparrow \infty$; this is to be expected, since many of the equations to be solved involve products of logarithmic and polynomial terms, and thus cannot be solved in closed form. On the other hand, these asymptotic expressions are sufficient for our purposes, since they describe how quickly $p_2^*(\alpha)$ and $C_2^*(\alpha)$ increase with α .

Note that as α tends to infinity, $C_2^*(\alpha)$ becomes independent of $u_2(v)$, which measures how hard it is to evaluate the derivatives of v ; this is because the combinatory cost eventually overwhelms the informational cost. This kind of behavior will be typical of the complexity analyses in this paper. Finally, note that the bound

$$(3.20) \quad C_1^*(\alpha) = O(\alpha^N) \leq C^*(\alpha) \leq O(\alpha^N \ln \alpha) = C_2^*(\alpha) \quad \text{as } \alpha \uparrow \infty$$

implies that

$$C_2^*(\alpha) / C_1^*(\alpha) = O(\ln \alpha) \quad \text{as } \alpha \uparrow \infty;$$

this indicates the gap in our knowledge of the complexity of solving (1.1) via Taylor series methods.

4. Linear Runge-Kutta Methods. For many functions v , calculation of the derivatives required by Taylor series methods is prohibitively expensive. For this reason, we are interested in methods which use information that is somewhat more readily available. In particular, we will consider methods that use only evaluation of v , combined in a highly structured manner. We say that Φ is a class of *linear Runge-Kutta methods* (abbreviated, “LRK methods”) if each increment function φ_p may be written

in the form

$$(4.1) \quad \varphi_p(x_i, h) := \sum_{l=0}^{s-1} \lambda_{sl} k_l,$$

where

$$(4.2) \quad k_l := v \left(x_i + h \sum_{j=0}^{l-1} \lambda_{lj} k_j \right) \quad \text{for } 0 \leq l \leq s-1,$$

the integer $s = s(p)$ is said to be the number of *stages* of φ_p ; the number of stages is equal to the number of times the vector function v must be evaluated. (In order to simplify notation, we will not explicitly indicate the dependence of λ_{lj} and k_j on p .) The method φ_p defined by (4.1) and (4.2) is *explicit* in that k_l depends only on k_0, \dots, k_{l-1} ; see Butcher [1964] for a discussion of *semiexplicit* and *implicit* methods. (We use the adjective "linear" to distinguish these methods from "nonlinear Runge-Kutta methods," which were first proposed in Brent [1974].) In what follows, we assume that φ_p is evaluated using the algorithm suggested by its definition.

In order to compute lower bounds on complexity, we consider the problem of finding the smallest value of $s(p)$ such that φ_p has order p . This minimal value is given by

$$(4.3) \quad s(p) = \begin{cases} p, & p = 1, 2, 3, 4, \\ p + 1, & p = 5, 6, \\ p + 2, & p = 7, \\ \text{unknown,} & p \geq 8. \end{cases}$$

For methods of order greater than seven, a gap develops. For instance, eighth-order methods with eleven stages exist, and it is known that any eighth-order method requires at least ten stages. For arbitrary $p \geq 8$, the best bounds known for the optimum value of $s(p)$ are

$$(4.4) \quad p + \theta(p) \leq s(p) \leq (p^2 - 7p + 14)/2,$$

where $\theta(p) \geq c \ln p$ for all sufficiently large p (for some $c > 0$). The lower bound is given in Butcher [1975]; the proof is quite involved, and the result is not much better than the "trivial" lower bound $s(p) \geq p$ (Hindmarsh [1974], p. 84). A class Φ_{CV} of methods such that φ_p requires only $(p^2 - 7p + 14)/2$ stages is given in Cooper and Verner [1972].

We first consider lower bounds on the complexity $C(p, \alpha)$ using LRK methods. The "trivial" lower bound $s(p) \geq p$ will be used, since the term $\theta(p)$ will be small when p is small and will not affect the asymptotic behavior of optimal order and complexity for p large. It is known (Butcher [1964]) that at least $O(p^2)$ of the subdiagonal elements of the matrix Λ (whose elements are the λ_{lj} in (4.2)) must be nonzero in order for Λ to define a p th-order method. Thus, there exists $a_1 > 0$ such that

$$(4.5) \quad d(p) \geq a_1 p^2;$$

since $s(p) \geq p$, we see that

$$(4.6) \quad u(\mathfrak{N}_p(v)) \geq u_1(v)p,$$

where we now write

$$u_1(v) := \min_{1 \leq i \leq N} u(v_i).$$

Thus, (4.5) and (4.6) show that a lower bound on the cost per step for φ_p is given by

$$(4.7) \quad c_1(p) = a_1 p^2 + Nu_1(v)p.$$

THEOREM 4.1.

$$C_1(p, \alpha) = M_1 [a_1 p^2 + Nu_1(v)p] e^{\alpha/p}.$$

Proof. This follows immediately from (2.19) and (4.7). \square

It is clear that $f_1(p) := M_1 [a_1 p^2 + Nu_1(v)p] e^{\alpha/p}$ satisfies (2.13) and (2.14).

We claim that f_1 yields a G_1 satisfying (2.17). Indeed, write

$$f_1(p) = f_{11}(p)f_{12}(p), \quad \text{where } f_{11}(p) := M_1 a_1 p$$

and

$$f_{12}(p) := p + \nu, \quad \text{where } \nu = \frac{N}{a_1} u_1(v).$$

Clearly, f_{11} yields a G_{11} satisfying (2.16). Since f_{12} is a linear polynomial with a negative zero, it may be shown that f_{12} yields a G_{12} satisfying (2.16). Thus, f_1 yields a G_1 satisfying (2.16); in fact, we have

$$(4.8) \quad G_1(p) + G_{12}(p) = \frac{p}{1 + (1 + \nu/p)^{-1}}.$$

This leads us to

THEOREM 4.2.

$$C_1^*(\alpha) \sim [M_1 a_1 e^2 / 4] \alpha^2 \quad \text{as } \alpha \uparrow \infty.$$

Proof. From (4.8), we see that $G_1(p) \sim 2p$ as $p \uparrow \infty$. Since (2.13), (2.14), and (2.16) hold, $p_1^*(\alpha)$ tends to infinity with α . Thus

$$\alpha = G_1(p_1^*(\alpha)) \sim 2p_1^*(\alpha) \quad \text{as } \alpha \uparrow \infty;$$

i.e., $p_1^*(\alpha) \sim \alpha/2$ as $\alpha \uparrow \infty$. The result now follows from Theorem 4.1. \square

We now turn to upper bounds on complexity. The class Φ_{CV} derived in Cooper and Verner [1972] has two deficiencies, the first of which is that no uniform upper bound on the local error per unit step is known for Φ_{CV} ; in addition, the combinatorial cost for this class of methods is $O(p^4)$ as $p \uparrow \infty$. Instead, we turn to the basic sequence Φ discussed in Theorem A.2 of the Appendix. There, we prove that there is an $M_2 > 0$ such that

$$(4.9) \quad \sigma(\varphi_p, h) \leq (M_2 \ln(p + e)h)^p$$

provided $h \leq h_p$, where $h_p = O((\ln p)^{-1})$ as $p \uparrow \infty$. Furthermore, there are a large number of extra zeros in the matrix Λ for $\varphi_p \in \Phi$. Using the notation of the Appendix, we see that the number of nonzero entries in Λ is

$$\sum_{i=0}^s \xi_i = \sum_{i=1}^{p-1} i^2 + p = p^3/3 - p^2/2 + 7p/6 \leq p^3/3 + 2p^2/3$$

for $p \in \mathbf{Z}^{++}$. Finally, note that the number of stages $s(p)$ required for $\varphi_p \in \Phi$ is

$$(4.10) \quad s(p) = \lfloor (p^2 - 2p + 4)/2 \rfloor \leq p^2/2 + p$$

for $p \in \mathbf{Z}^{++}$, which shows that the number of stages required for a p th-order method in Φ asymptotically equals the number required for a p th-order method in Φ_{CV} . Thus (considering the combinatory costs), the class Φ_{CV} actually costs more per step than does Φ ; ignoring the combinatory costs would have caused us to reach the opposite conclusion.

First, we look at the cost per step. By (4.10), we see that

$$(4.11) \quad u(\mathfrak{N}_p(v)) \leq \frac{1}{2} (p^2 + p)Nu_2(v),$$

where

$$u_2(v) := \max_{1 \leq i \leq N} u(v_i).$$

Since we are using Φ , it is easy to see that there is a $b_2 \geq 2/3$ such that

$$(4.12) \quad d(p) \leq (p^3/3 + b_2p^2)2N.$$

Combining (4.11) and (4.12), we see that the total combinatory cost per step is bounded by

$$(4.13) \quad c_2(p) = N[2p^3/3 + \beta_2p^2 + \gamma_2p],$$

where

$$\beta_2 := u_2(v)/2 + 2b_2 \quad \text{and} \quad \gamma_2 := u_2(v)/2.$$

Using (4.9) and (4.13) gives

THEOREM 4.3.

$$C_2(p, \alpha) = M_2N[2p^3/3 + \beta_2p^2 + \gamma_2p] \ln(p + e)e^{\alpha/p}. \quad \square$$

Now we look at the optimality theory for the upper bound.

THEOREM 4.4. (1) For all $\alpha > 0$, there exists $p_2^*(\alpha)$ such that (2.17) holds.

(2) $p_2^*(\alpha)$ increases monotonically with α , and $p_2^*(\alpha) \sim \alpha/3$ as $\alpha \uparrow \infty$.

(3) $C_2^*(\alpha)$ increases monotonically with α , and $C_2^*(\alpha) \sim [2M_2Ne^3/81]\alpha^3 \ln \alpha$ as $\alpha \uparrow \infty$.

(4) $h_2^*(\alpha) \sim (M_2e^3 \ln \alpha)^{-1}$ as $\alpha \uparrow \infty$.

Proof. We write

$$f_2(p) := M_2 \ln(p + e)c_2(p)$$

in the form

$$f_2(p) = f_{21}(p)f_{22}(p),$$

where

$$f_{21}(p) = M_2 N p \ln(p + e) \quad \text{and} \quad f_{22}(p) = 2p^2/3 + \beta_2 p + \gamma_2.$$

It is clear that f_{21} satisfies the hypotheses of Lemma 2.1. Now we consider f_{22} . Clearly, f_{22} has no positive zeros; it may be seen that the condition $b_2 \geq 2/3$ implies that f_{22} has a positive discriminant and, hence, has no complex roots. Thus, f_{22} has only negative roots; one may show that this guarantees that f_{22} satisfies the hypotheses of Lemma 2.1. Thus, the same may be said for $f_2 = f_{21}f_{22}$.

Thus p_2^* and C_2^* behave as described in Theorem 2.3. We also see that $G_2(p) \sim p$ as $p \uparrow \infty$. Thus, the estimate in (2.) holds, from which we get the estimates in (3.) and (4.). \square

So in the class of linear Runge-Kutta methods, we find that

$$(4.14) \quad C_1^*(\alpha) = O(\alpha^2) \leq C^*(\alpha) \leq C_2^*(\alpha) = O(\alpha^3 \ln \alpha)$$

as α tends to infinity; hence, the ratio

$$C_2^*(\alpha)/C_1^*(\alpha) = O(\alpha \ln \alpha)$$

indicates the gap in our knowledge of the complexity of linear Runge-Kutta methods.

5. Comparison of the Methods. We now wish to compare the classes of Taylor series methods and LRK methods. Let $C_{L,T}$ denote the lower bound and $C_{U,T}$ denote the upper bound on complexity of Taylor series methods, and let C_T denote the inherent complexity of Taylor series methods; then $C_{L,T}(\alpha) \leq C_T(\alpha) \leq C_{U,T}(\alpha)$. Similarly, we write $C_{L,LRK}$, $C_{U,LRK}$, and C_{LRK} for linear Runge-Kutta methods. Since we have only asymptotic expressions for these quantities, we are forced to use an asymptotic comparison. If $f, g: \mathbf{R}^{++} \rightarrow \mathbf{R}^{++}$ satisfy $\lim_{\alpha \uparrow \infty} f(\alpha) = \lim_{\alpha \uparrow \infty} g(\alpha) = +\infty$, we will write

$$(5.1) \quad f < g \quad \text{iff} \quad f(\alpha) = o(g(\alpha)) \quad \text{as} \quad \alpha \uparrow \infty;$$

we say that f is *asymptotically less than* g . If $f < g$, there is an $\alpha_0 > 0$ such that $f(\alpha) < g(\alpha)$ for $\alpha > \alpha_0$, so there is a nonasymptotic interpretation of the order relation $<$. Thus, if f and g are cost functions, the statement " $f < g$ " implies that the method whose cost is given by f is "better" (i.e., cheaper) than the method whose cost is given by g , for ϵ sufficiently small. Using the results of (3.20) and (4.14), we then have the following

THEOREM 5.1. *Suppose that (3.16) holds.*

- (1) *If $N = 1$, then $C_{U,T} < C_{L,LRK}$.*
- (2) *If $N = 2$, then $C_{U,T} < C_{U,LRK}$.*

(3) If $N = 3$, then $C_{U,T}(\alpha) = O(C_{U,LRK}(\alpha))$ and $C_{U,LRK}(\alpha) = O(C_{U,T}(\alpha))$ as $\alpha \uparrow \infty$.

(4) If $N \geq 4$, then $C_{U,LRK} < C_{L,T}$. \square

If (3.16) does *not* hold, then (1), (2), and (3) may be false, but (4) will certainly be true. As an immediate corollary to the above theorem, we have

THEOREM 5.2. (1) If $N = 1$ and (3.16) holds, then $C_T < C_{LRK}$.

(2) If $N \geq 4$, then $C_{LRK} < C_T$. \square

So if the derivatives of v are cheap to evaluate, we see that the best Taylor series method known is better than the best linear Runge-Kutta method possible for the scalar case $N = 1$; but if $N \geq 4$, the best linear Runge-Kutta method known is better than the best Taylor series method possible.

We hasten to point out that these comparisons are asymptotic as the error criterion tends to zero. This leads us to ask whether these results (especially the first statement of Theorem 5.2) have meaning for "practical" error criteria, e.g., five to twenty decimal places. We suspect that the fast Taylor series methods may be better than the linear Runge-Kutta methods for ϵ in this range (with $N = 1$), but more work has to be done along the lines of good implementations of the fast methods. As a matter of fact, even if we were to assume that the asymptotic formulas in Theorems 3.5 and 4.2 were exact, computing the crossover value of α for a given function or class of functions would still be difficult, in view of the fact that the asymptotic constants will be difficult to compute in practice.

Appendix: Error Bounds for a Sequence of LRK Methods. In this Appendix, we describe a subclass of a class of linear Runge-Kutta ("LRK") methods due to Cooper [1969]. We shall first prove the following

THEOREM A.1. *There is a basic sequence Φ of LRK methods such that*

(1) *Each $\varphi_p \in \Phi$ requires $s(p) := (p^2 - p + 2)/2$ evaluations of v per step.*

(2) *There exists an $M > 0$ such that*

$$(A.1) \quad \sigma(\varphi_p, h) \leq (M \ln(p + e)h)^p$$

for $h \leq h_p = O((\ln p)^{-1})$.

We use the notation of Cooper and Verner [1972]. Let $p \in \mathbf{Z}^{++}$ be given; define $\rho: \mathbf{Z}^+ \cap [0, p] \rightarrow \mathbf{Z}^+$ by

$$(A.2) \quad \rho(j) := \begin{cases} \sum_{k=0}^j k \equiv j(j+1)/2 & \text{if } j \neq p, \\ s & \text{if } j = p, \end{cases}$$

where we write "s" for " $s(p)$ " as defined above. Next, a set $\{\xi_0, \dots, \xi_s\}$ of integers is defined by picking $\xi_0 := p$, and setting ξ_i ($i \neq 0$) to be the unique integer in $[1, p]$

satisfying

$$(A.3) \quad \rho(\xi_i - 1) < i \leq \rho(\xi_i).$$

We now pick $u_0, \dots, u_s \in I$ satisfying

$$(A.4) \quad u_0 = 0, \quad u_s = 1, \quad u_i \neq 0 \quad \text{if } i \neq 0$$

and

$$(A.5) \quad (\xi_i = \xi_j \text{ and } i \neq j) \text{ implies } u_i \neq u_j.$$

Finally, we pick a matrix of coefficients $\Lambda := \{\lambda_{ij}: 0 \leq j \leq i-1, 1 \leq i \leq s\}$ such that

$$(A.6) \quad \lambda_{ij} = 0 \quad \text{if } \xi_i < \xi_j - 1 \quad (1 \leq i, j \leq s)$$

and

$$(A.7) \quad \sum_{j=0}^{i-1} \lambda_{ij} u_j^\tau = (\tau + 1)^{-1} u_i^{\tau+1} \quad (0 \leq \tau \leq \xi_i - 1, 1 \leq i \leq s).$$

Cooper and Verner [1972] point out that these conditions may always be fulfilled; the resulting Λ defines a p th-order LRK method with s stages.

We are interested in a choice of u_0, \dots, u_s which will give a small error coefficient. To this end, we will choose

$$(A.8) \quad \{u_j: \xi_j = n\} = \{(1 + x_{kn})/2: 1 \leq k \leq n\} \quad (1 \leq n \leq p-1),$$

where x_{1n}, \dots, x_{nn} are the zeros of the Jacobi polynomial $P_n := P_n^{(1,1)}$ (see Szegő [1959]). Since these zeros are distinct and lie in $[-1, 1]$, conditions (A.4) and (A.5) may be satisfied.

Now we are able to exhibit a solution to the i th system in (A.7). First, note that the equation for $\tau = 0$ may be separated from the others, since $u_0 = 0$. Setting $n := \xi_i - 1$, we see that

$$(A.9) \quad \lambda_{i0} = u_i - \sum_{j=1}^{i-1} \lambda_{ij} = u_i - \sum \{\lambda_{ij}: j < i \text{ and } \xi_j \geq n\},$$

the last by (A.6). We wish to determine the nonzero λ_{ij} , i.e., those λ_{ij} for which $\xi_j \geq n$ and $j < i$. So setting $\lambda_{ij} = 0$ unless $j \in \{j_1, \dots, j_n\}$, we see that the remaining λ_{ij} are the solution of the system

$$\sum_{k=1}^n u_{jk}^\tau \lambda_{ijk} = (\tau + 1)^{-1} u_i^{\tau+1} \quad (1 \leq \tau \leq n).$$

Thus, the λ_{ijk} are the weights for an interpolatory quadrature formula on $[0, u_i]$ with abscissae u_{j_1}, \dots, u_{j_n} . From the usual expression for such weights and (A.6), we see that

$$\lambda_{ijk} = \frac{1}{2P_n'(\cos \theta_{kn})} \int_{\theta_{i,n+1}}^{\pi} \frac{P_n(\cos \theta)}{\cos \theta - \cos \theta_{kn}} d\theta,$$

where $x_{kn} = \cos \theta_{kn}$ ($1 \leq k \leq n$).

LEMMA A.1. $\mu_{ikn} = O(n^{-1} \ln n)$ as $n \uparrow \infty$.

Proof. Since the zeros of P_n are symmetric about the origin, we may assume that $0 < \theta_{kn} \leq \pi/2$. Using (8.9.2) of Szegő [1959], we then find

$$\mu_{ikn} = O(k^{5/2}n^{-3}) \int_{\theta_{i,n+1}}^{\pi} \frac{P_n(\cos \theta)}{\cos \theta - \cos \theta_{kn}} \sin \theta \, d\theta.$$

Case 1. $\theta_{1,n+1} \leq \theta_{i,n+1} \leq \theta_{k,n+1}/2$. We consider the integral over $[\theta_{1n}/2, \theta_{i,n+1}]$, since Theorem 15.4 of Szegő [1959] proves that

$$O(k^{5/2}n^{-3}) \left[\left| \int_0^{\pi} \right| + \left| \int_0^{\theta_{1n}/2} \right| \right] = O(n^{-1}).$$

(Here the integrand is the same as in the preceding integral.) But the proof of (15.4.12) in Szegő [1959] extends almost immediately to a proof that the remaining integral is $O(k^{-2}n)$, since (15.4.12) is proved by order-of-magnitude estimates. Thus, $\mu_{ikn} = O(n^{-1}) = O(n^{-1} \ln n)$ for Case 1.

Case 2. $\theta_{k,n+1}/2 \leq \theta_{i,n+1} \leq 3\theta_{k,n+1}/2$. We consider the integral over $[\theta_{kn}/2, \theta_{i,n+1}]$, since Szegő [1959] shows that

$$O(k^{5/2}n^{-3}) \left| \int_{\theta_{kn}/2}^{\pi} \right| = O(n^{-1}).$$

As in (15.4.13) of Szegő [1959], we have

$$\int_{\theta_{kn}/2}^{\theta_{i,n+1}} = O(nk^{-3/2})I_1 + I_2.$$

Here

$$I_1 := \int_{\theta_{kn}/2}^{\theta_{i,n+1}} D(\theta) \sin \theta \, d\theta$$

with

$$D(\theta) := \frac{\cos(N\theta + \gamma) - \cos(N\theta_{kn} + \gamma)}{\cos \theta - \cos \theta_{kn}},$$

where $N := n + 3/2$ and $\gamma := -3\pi/4$, and

$$I_2 := \int_{\theta_{kn}/2}^{\theta_{i,n+1}} R_n(\theta, \theta_{kn}) \sin \theta \, d\theta = O(nk^{-3/2}),$$

with R_n the remainder term in (8.8.2) of Szegő [1959]. Unfortunately, the proof that (15.4.14) of Szegő [1959] is bounded does *not* extend to a proof that I_1 is bounded, since the proof of the former requires that the interval of integration be symmetric about θ_{kn} . However, it is straightforward to verify that

$$I_1 = O(1) \int_0^{\pi/4} \left| \frac{\sin N\theta}{\theta} \right| d\theta = O(\ln n).$$

Thus, $\mu_{ikn} = O(n^{-2}k \ln n) = O(n^{-1} \ln n)$ for Case 2.

Case 3. $3\theta_{k,n+1} \leq \theta_{i,n+1} \leq 3\pi/4$. We consider the integral over $[3\theta_{kn}/2, \theta_{i,n+1}]$, since Szegő [1959] proves that

$$O(k^{5/2}n^{-3}) \left| \int_{3\theta_{kn/2}}^{\pi} \right| = O(n^{-1}).$$

But the proof of (15.4.19) in Szegő [1959] extends to prove that the remaining integral is $O(k^{-5/2}n)$ (as in Case 1). Thus, $\mu_{ikn} = O(n^{-1}) = O(n^{-1} \ln n)$ for Case 3.

Case 4. $3\pi/4 \leq \theta_{i,n+1} \leq \theta_{n+1,n+1}$. We consider the integral over $[3\pi/4, \theta_{i,n+1}]$, since Szegő [1959] shows that

$$O(k^{5/2}n^{-3}) \left| \int_{3\pi/4}^{\pi} \right| = O(n^{-1}).$$

As in Cases 1 and 3, the proof of the above may be extended to prove a similar bound on the integral of interest. Thus, $\mu_{ikn} = O(n^{-1}) = O(n^{-1} \ln n)$ in Case 4, completing the proof of the lemma. \square

Thus, (A.9) and Lemma A.1 show the existence of a $\lambda > 0$ such that

$$(A.10) \quad \sum_{j=0}^{i-1} |\lambda_{ij}| \leq \lambda \ln(\xi_i + e);$$

here λ is independent of p . Moreover, the result for the case $i = s$ may be sharpened. We see that $\lambda_{sj} \geq 0$, since the u_j for the s th system in (A.7) are the abscissae for Lobatto quadrature. Thus

$$(A.11) \quad \sum_{j=0}^{s-1} |\lambda_{sj}| = \sum_{j=0}^{s-1} \lambda_{sj} = 1,$$

the consistency condition in the last equality being a consequence of (A.7) with $\tau = 0$.

Proof of Theorem A.1. As in Cooper and Verner [1972], we define

$$\epsilon_i := \dot{x}(u_i h) - k_i \quad \text{and} \quad \delta_i := \int_0^{u_i} \dot{x}(uh) du - \sum_{j=0}^{i-1} \lambda_{ij} \dot{x}(u_j h)$$

for $0 \leq i \leq s$; note that $\delta_0 = \epsilon_0 = 0$. Let $z(h)$ be the computed approximation to $x(h)$; then

$$(A.12) \quad \begin{aligned} h^{-1} \|x(h) - z(h)\| &= \left\| h^{-1} [x(h) - x(0)] - \sum_{i=0}^{s-1} \lambda_{si} k_i \right\| \leq \|\delta_s\| + \left\| \sum_{i=0}^{s-1} \lambda_{si} \epsilon_i \right\| \\ &\leq \|\delta_s\| + \max_{\xi_i=p-1} \|\epsilon_i\|, \end{aligned}$$

the last by (A.6) and (A.11). By the analyticity of x , there is an $A_1 > 0$ such that

$$\beta_i := h^{-1} \left\| x(u_i h) - \sum_{\tau=0}^{\xi_i} (u_i h)^\tau x^{(\tau)}(0)/\tau! \right\| \leq (A_1 h)^{\xi_i}$$

and

$$\gamma_{ij} := \left\| \dot{x}(u_j h) - \sum_{\tau=0}^{\xi_j-1} (u_j h)^\tau \dot{x}^{(\tau)}(0)/\tau! \right\| \leq (A_1 h)^{\xi_i},$$

so that the definition of δ_i gives

$$(A.13) \quad \|\delta_i\| \leq \beta_i + \sum_{j=0}^{i-1} |\lambda_{ij}| \gamma_{ij} \leq (A_1 h)^{\xi_i} + \sum_{j=0}^{i-1} |\lambda_{ij}| (A_1 h)^{\xi_i} \leq (A_2 h)^{\xi_i}$$

for a suitable $A_2 > 0$. Thus, (A.12) becomes

$$(A.14) \quad h^{-1} \|x(h) - z(h)\| \leq (A_2 h)^p + \max_{\xi_i = p-1} \|\epsilon_i\|.$$

We now use Lemma 1.1 of Cooper and Verner [1972] and (A.6) to find that if L is a Lipschitz constant for v , then there exists $A_3 > 0$ such that

$$\begin{aligned} \|\epsilon_i\| &\leq hL \|\delta_i\| + hL \sum_{j=0}^{i-1} |\lambda_{ij}| \max_j \|\epsilon_j\| \\ &\leq (A_3 h)^{\xi_i+1} + (A_3 h) \ln(\xi_i + e) \max_j \|\epsilon_j\|, \end{aligned}$$

the last by (A.10) and (A.13); here, the maximum is taken over all $j < i$ such that $\xi_j \geq \xi_i - 1$. A straightforward induction shows that if $(1 + \ln 2)A_3 h < 1$, then

$$\|\epsilon_i\| \leq (A_4 \ln(\xi_i + e)h)^{\xi_i+1}$$

for a suitable $A_4 > 0$. Combining this with (A.14), we find

$$(A.15) \quad h^{-1} \|x(h) - z(h)\| \leq (A_5 \ln(p + e)h)^p,$$

the desired bound for the local error for a single unit step.

To extend (A.15) to a global error result, we must look at the Lipschitz constants for the increment functions. Let L be a bound on $\|\nabla v\|$, and write “ $\nabla\varphi_p(y, h)$ ” to indicate gradient with respect to the vector variable y . Now

$$\|\nabla\varphi_p(y, h)\| \leq \sum_{i=0}^{s-1} |\lambda_{si}| \max_{0 \leq i \leq s-1} \|\nabla k_i(y, h)\| = \max_{0 \leq i \leq s-1} \|\nabla k_i(y, h)\|,$$

where we write “ $k_i(y, h)$ ” to indicate the dependence of k_i upon y and h . By the definition of $k_i(y, h)$, we find

$$\nabla k_i(y, h) = \nabla v(u) \left[1_{N \times N} + h \sum_{j=0}^{i-1} \lambda_{ij} \nabla k_j(y, h) \right],$$

where $u := y + h \sum_{j=0}^{i-1} \lambda_{ij} k_j(y, h)$ and $1_{N \times N}$ is an $N \times N$ identity matrix. Taking norms in the above gives the result

$$\zeta_i \leq L\lambda + hL\lambda [\ln(\xi_i + e) \max\{\zeta_j: j < i \text{ and } \xi_j \geq \xi_i - 1\}],$$

where $\zeta_j := \|\nabla k_j(y, h)\|$. Writing λ_p for the Lipschitz constant for φ_p , it is easy to see that (A.16) and the above inequality imply

$$\lambda_p \leq \sum_{j=0}^{p-1} (hL\lambda)^j \prod_{k=1}^{j-2} \ln(p + e - k),$$

which is bounded for all p , provided that $h \leq h_p < (L\lambda \ln(p + e))^{-1}$. Thus (A.1) follows from this result, (A.15), and Theorem 3.3 of Henrici [1962]. \square

The value for $s(p)$ indicated in Theorem A.1 may be improved somewhat by noting that since we are using a Lobatto quadrature, higher order may be expected with fewer steps. Indeed, if we use the strategy outlined in the comments following Theorem 4 of Cooper and Verner [1972], we have

THEOREM A.2. *There exists a basic sequence Φ of LRK methods such that (A.1) holds and φ_p requires*

$$s(p) := \lfloor (p^2 - 2p + 4)/2 \rfloor$$

evaluations of v per step. \square

Acknowledgements. I would like to thank Professor R. P. Brent of the Australian National University, Professors H. T. Kung and J. F. Traub of Carnegie-Mellon University, and Professor H. Woźniakowski of the University of Warsaw for their comments and suggestions on the results reported in this paper. In addition, a number of the results in Sections 3 and 4 were obtained by using the MACSYMA system developed by the Matlab group at Massachusetts Institute of Technology, which is supported by the Defense Advanced Research Projects Agency work order 2095, under Office of Naval Research Contract N00014-75-C-0661.

Department of Mathematics
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

A. BORODIN & I. MUNRO [1975], *The Computational Complexity of Algebraic and Numeric Problems*, American Elsevier, New York, 1975.

R. P. BRENT [1974], *Efficient Methods for Finding Zeros of Functions Whose Derivatives are Easy to Evaluate*, Computer Science Department Report, Carnegie-Mellon University, 1974.

R. P. BRENT & H. T. KUNG [1978], "Fast algorithms for manipulating formal power series," *J. Assoc. Comput. Mach.*, v. 25, 1978, pp. 581–595.

J. C. BUTCHER [1964], "Implicit Runge-Kutta processes," *Math. Comp.*, v. 18, 1964, pp. 50–64.

J. C. BUTCHER [1975], "An order bound for Runge-Kutta methods," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 304–315.

G. J. COOPER [1969], "Error bounds for some single-step methods," *Conf. on the Numerical Solution of Differential Equations*, Lecture Notes in Math., vol. 109, Springer-Verlag, Berlin, 1969, pp. 140–147.

G. J. COOPER & J. H. VERNER [1972], "Some explicit Runge-Kutta methods of high order," *SIAM J. Numer. Anal.*, v. 9, 1972, pp. 389–405.

A. FRIEDMAN [1969], *Partial Differential Equations*, Holt, Rinehart, and Winston, New York, 1969.

P. HENRICI [1962], *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York, 1962.

A. C. HINDMARSH [1974], *Numerical Solution of Ordinary Differential Equations: Lecture Notes*, Lawrence Livermore Laboratory Report No. UCID-16588, June, 1974.

G. PÓLYA & G. SZEGÖ [1925], *Aufgaben und Lehrsätze der Analysis*, Vol. I, Springer-Verlag, Berlin, 1925.

L. B. RALL [1969], *Computational Solution of Nonlinear Operator Equations*, Wiley, New York, 1969.

H. J. STETTER [1973], *Analysis of Discretization Methods for Ordinary Differential Equations*, Springer-Verlag, Berlin, 1972.

G. SZEGÖ [1959], *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ., Vol. 23, Amer. Math. Soc., Providence, R. I., 1959.

J. F. TRAUB & H. WOŹNIAKOWSKI [1976], "Strict lower and upper bounds on iterative complexity," in *Analytic Computational Complexity* (J. F. Traub, Ed.), Academic Press, New York, 1976.

A. G. WERSCHULZ [1976], *Optimal Order and Minimal Complexity of One-Step Methods for Initial Value Problems*, Computer Science Department Report, Carnegie-Mellon University, 1976.