Computational Complexity of One-Step Methods for a Scalar Autonomous Differential Equation

Arthur G. Werschulz

September, 1976

Department of Mathematics
Carnegie-Mellon University
Pittsburgh, PA 15213

Abstract: The problem is to calculate an approximate solution of an initial value problem for a scalar autonomous differential equation. A generalized notion of a nonlinear Runge-Kutta (NRK) method is defined. We show that the order of any s-stage NRK method cannot exceed 2s - 1; hence, the family of NRK methods due to Brent has the maximal order possible. Using this result, we derive complexity bounds on the problem of finding an approximate solution with error not exceeding ε. We also compute the order which minimizes these bounds, and show that this optimal order increases as ε decreases, tending to infinity as ε tends to zero.

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.

1. Introduction

Let $\mathcal{D}$ be a subset of the real numbers $\mathbb{R}$, and let $\mathcal{V} = \{v : \text{domain}(v) \subset \mathbb{R} \to \mathbb{R}\}$ be a set of functions, such that the initial value problem of finding a function $x : [0, 1] \to \mathbb{R}$ satisfying

$$\begin{aligned} \dot{x}(t) &= v(x(t)) \qquad 0 < t < 1 \\ x(0) &= x_0 \end{aligned} \tag{1.1}$$

has a unique solution for every $(x_0 , v) \in \mathcal{D}\times\mathcal{V}$. (The differential equation in (1.1) is said to be a scalar autonomous differential equation.) We are interested in the computational complexity of using one-step methods to generate an approximation to (1.1) on an equidistant grid (in the sense of Stetter [73]); that is, the methods considered give approximations x_i to $x(ih)$ by the recurrence

$$x_{i+1} = x_i + h\,\varphi(x_i,h) \qquad (0 \le i \le n - 1), \tag{1.2}$$

where $h = n^{-1}$ is the step-size of a grid with n points, and φ is the increment function (Henrici [62]) for the method. (For brevity, we will refer to "the method φ.")

In Werschulz [76a], we discussed the complexity of solving autonomous systems of differential equations; in this paper, we will consider only the case of a single scalar autonomous equation. Clearly, the results of Werschulz [76a] hold for problems of the form (1.1). However, in this paper we will discuss the complexity of solving (1.1) via nonlinear Runge-Kutta methods (abbreviated, "NRK methods"). We only consider the scalar case (1.1), since it is not known whether NRK methods exist for more general systems.

In Section 2, we give the formal definition of "NRK method," and show that no NRK method using s evaluations of v ("stages") can have order exceeding 2s - 1. Thus,

the set of s-stage methods of order 2s - 1 described in Brent [74] has maximal order in the class of NRK methods.

In Section 3, we use the results of Brent [74] and Section 2 to find upper and lower bounds on the complexity of finding an approximate solution whose error does not exceed ε, using a method of fixed order. These results are then used to calculate optimal orders which minimize these complexity bounds. We show that the optimal order increases as ε decreases, tending to infinity as ε tends to zero. Finally, we compare the complexities of NRK methods, Taylor series methods, and linear Runge-Kutta methods. We show that the best NRK methods known are asymptotically better (as ε tends to zero) than the best linear Runge-Kutta methods possible, but are asymptotically worse than the best Taylor series methods known if the cost of evaluating the k^{th} derivative of v is bounded for all k.

2. Maximal Order for NRK Methods

Before proceeding any further, we will review some basic notions from Werschulz [76b]. The following notational conventions will be used. Let $\mathfrak{X}$ be an ordered ring; then "$\mathfrak{X}^+$" and "$\mathfrak{X}^{++}$" respectively denote the nonnegative and positive elements of $\mathfrak{X}$. (This is used in the cases $\mathfrak{X} = \mathbb{R}$, the real numbers, and $\mathfrak{X} = \mathbb{Z}$, the integers.) The symbol ":=" means "is defined to be." We use "I" to denote the unit interval [0, 1]. The notations "$x \downarrow a$" and "$x \uparrow a$" are used to indicate one-sided limits, as in Buck [65]. Finally, if $\chi_1, \chi_2 : \mathbb{R} \to \mathbb{R}$ and $\omega : \mathbb{R}^2 \to \mathbb{R}$ are differentiable, then for $i = 1, 2$, we write

$$\partial_i \omega(\chi_1(t), \chi_2(t))$$

for the result of differentiating $\omega(\chi_1, \chi_2)$ with respect to χ_i, and then substituting $\chi_1 = \chi_1(t)$, $\chi_2 = \chi_2(t)$.

We next describe the model of computation to be used. We assume only that all arithmetic operations are performed exactly in $\mathbb{R}$ (i.e., infinite-precision arithmetic) and that for all $v \in \mathcal{V}$, we are able to compute the value of 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

$$\sigma_G(\varphi,h) := \max_{0 \le i \le n} |x(ih) - x_i| \quad . \tag{2.1}$$

Other error measures may be used, such as the local error per step and the local error per unit step (see Henrici [62] and Stetter [73] for definitions); this would involve only a slight modification of the results contained in the sequel.

Finally, we will say that $\Phi = \{\varphi_p : p \in \mathbb{Z}^{++}\}$ is a basic sequence of methods if there exist functions $\kappa : \mathbb{R}^+ \times I \to \mathbb{R}$ and $\kappa_L, \kappa_U : \mathbb{R}^+ \to \mathbb{R}^+$ such that

(2.2) $\sigma_G(\varphi_p,h) = \kappa(p,h)\, h^p$ for $h \in I$ and $p \in \mathbb{Z}^{++}$,

where

(2.3) $0 < \kappa_L(p) \le \kappa(p,h) \le \kappa_U(p) < +\infty$ 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 [72]; the function κ_L introduced here is necessary and sufficient for the "order" of a method to be unique. (Here we introduce the convention of attaching the subscripts "L" and "U" to quantities dealing with lower and upper bounds (respectively) on complexity.)

We now consider a generalization of the familiar linear Runge-Kutta methods which are found in standard texts such as Henrici [62]. A basic sequence Φ is said to be a sequence of nonlinear Runge-Kutta methods ("NRK methods") if each increment function $\varphi_p \in \Phi$ may be written in the form

(2.4) $\varphi_p(x_i,h) := \tau_s(x_0,h; k_0, \ldots, k_{s-1})$,

where

(2.5) $k_j := v(y_j)$, $y_j := \tau_j(x_i,h; k_0, \ldots, k_{j-1})$ $(0 \le j \le s-1)$

for suitable functions $\tau_j : \mathbb{R}\times\mathbb{R}\times\mathbb{R}^j \to \mathbb{R}$ $(0 \le j \le s)$. We say that φ_p has $s = s(p)$ stages, so that an s-stage NRK method uses s evaluations of v. Since the one-step method φ_p defined by (2.4) and (2.5) is stationary (i.e., does not change from step to step), we need only describe how x_1 is generated from x_0.

Brent [74], [76] considered the problem of finding a simple root ξ of a nonlinear function $F : \mathbb{R} \to \mathbb{R}$, using the Brent-information (Meersman [76])

(2.6) $\mathfrak{N}_{B,s}(F) := \{F(x_0), F'(x_0), F'(y_1), \ldots, F'(y_{s-1})\}$,

where x_0 is an initial approximation to ξ, and $y_1, \ldots, y_{s-1}$ are to be determined. Let x_1 be a sufficiently good approximation of the appropriate zero of the minimal-degree

polynomial interpolating the information $\mathfrak{N}_{B,s}(F)$. Then Brent [74] showed how to choose $y_1, \ldots, y_{s-1}$ so that

$$|x_1 - \xi| = O(|x_0 - \xi|^{2s}) \quad \text{as} \quad x_0 \to \xi \ . \tag{2.7}$$

This defines an iterative method of order 2s for finding ξ.

Let us now define a function F by setting

$$F(z) := \int_{x_0}^{z} d\xi \, / \, v(\xi) \ - \ h \ , \tag{2.8}$$

and note that x(h) is the zero of F. Recalling that order for iterations is defined differently than is order for one-step methods, (2.8) shows how an s-stage NRK method of order p may be derived from a $(p + 1)^{th}$-order iterative method for zero-finding which uses the Brent-information (2.6). Using this transformation and (2.7), Brent [74], [76] exhibited a sequence Φ_{MBRK} of "modified" Brent-Runge-Kutta methods ("BRK methods"), in which the s-stage method has order

$$p = 2s - 1 \ . \tag{2.9}$$

Furthermore, Meersman [76] proved that this order is the greatest possible in the class of all such BRK methods. We now extend Meersman's result to include all NRK methods.

Theorem 2.1: No s-stage NRK method can have order greater than 2s - 1.

Proof: Let φ be an s-stage method with order p. We will construct (from φ) an iterative method ψ of order $q := p + 1$ for finding a simple zero ξ of an arbitrary analytic function $F : \mathbb{R} \to \mathbb{R}$.

The method ψ is defined as follows. Let x_0 be an approximation to ξ such that F' is nonzero between x_0 and ξ. (Since $F'(\xi) \neq 0$, such an x_0 exists.) Write $t_0 := F(x_0)$; without loss of generality, assume $t_0 < 0$. Now apply one step of φ, using a step-size of $-t_0$, to the problem

$$\dot{x}(t) = F'(x(t))^{-1} \quad (t_0 < t < 0) \quad \text{with} \quad x(t_0) = x_0 \ ,$$

(whose solution is the functional inverse of F, so that $x(0) = F^{-1}(0) = \xi$); then ψ is given by

$$\psi(x_0) := x_0 - t_0 \varphi(x_0, t_0) .$$

By the definition of order for iterative methods, it is clear that ψ has order q; moreover, ψ uses the generalized Brent information (Definition II.3.8 of Meersman [76])

$$\mathfrak{I}_{GB,s}(F) := \{F(x_0), F'(y_0), F'(y_1), \ldots, F'(y_{s-1})\} .$$

Suppose that $y_0 \neq x_0$; then $q \leq 2s$ by Theorem II.3.3 of Meersman [76]. On the other hand, if $y_0 = x_0$, then ψ uses the Brent-information (2.6); by Theorem II.2.4 of Meersman [76] (also due to Woźniakowski), we have $q \leq 2s$ in this case also. Thus in either case, we find that

$$p + 1 = q \leq 2s ,$$

and the desired result follows. ∎

Thus Φ_{MBRK} is informationally-optimal in the class of NRK methods, in the sense that each φ_p in Φ_{MBRK} uses the minimum number of stages possible for a p^{th}-order NRK method.

3. Complexity Bounds for NRK Methods

In this Section, we will compute lower and upper bounds on the total number of arithmetic operations $C(p,\alpha)$ required to guarantee that if φ_p is a p^{th}-order NRK method, then

$$\sigma_G(\varphi_p,h) \leq \varepsilon := e^{-\alpha} \tag{3.1}$$

for a given $p \in \mathbb{Z}^{++}$ and $\alpha \in \mathbb{R}^{++}$. (Here e is the base of the natural logarithms.) Since $\alpha > 0$, we have $0 < \varepsilon < 1$; clearly α increases as ε decreases, and α tends to infinity as ε tends to zero.

In the methods we consider, we may write

$$C(p,\alpha) = n\, c(p) = h^{-1}\, c(p)\ , \tag{3.2}$$

where n is the minimal number of steps required (so that $h = n^{-1}$ is the maximal step-size permitted), and the cost per step c(p) is the number of arithmetic operations required for the execution of one step of a p^{th}-order NRK method. As in Traub and Woźniakowski [76], we shall express the cost per step in the form

$$c(p) := e(\mathfrak{N}_p(v)) + d(p)\ . \tag{3.3}$$

Here $\mathfrak{N}_p(v)$ is the information about v required to perform one step of a p^{th}-order NRK method φ_p, and we write $e(\mathfrak{N}_p(v))$ for the informational cost of φ_p; we call d(p) the combinatory cost of φ_p. For example, Euler's method

$$x_{i+1} = x_i + h\, v(x_i)$$

has informational cost

$$e(v) := \text{cost of evaluating } v \text{ at one point.} \tag{3.4}$$

The combinatory cost is two operations (i.e., one addition and one multiplication).

We now assume that the solution x of (1.1) is analytic on I. Thus Cauchy's Integral Theorem (Ahlfors [66], pg. 122) shows that there exists an $M > 0$ such that

$$|x^{(k)}(t)| / k! \le M^k \quad \text{for all } t \in I .$$

Finally, we shall restrict our attention to problems which are "sufficiently difficult," i.e., for which there exists an $M_L > 0$ independent of h and p so that

$$\sigma_G(\varphi_p,h) \ge (M_L h)^p \quad \text{if } h \in I \text{ and } p \in \mathbb{Z}^{++} . \tag{3.5}$$

(See Section 4 of Werschulz [76b].)

We will now derive a lower bound for the complexity $C(p,\alpha)$ via NRK methods. Clearly, Theorem 2.1 implies that for any p^{th}-order NRK method, we must have

$$e(\mathfrak{N}_p(v)) \ge e(v)\,(p + 1) / 2 , \tag{3.6}$$

and a linear lower bound on the combinatory cost states that

$$d(p) \ge a_L\, p \tag{3.7}$$

for some $a_L > 0$. By (3.6) and (3.7), a lower bound on the cost per step for φ_p is

$$c_L(p) = (a_L + e(v)/2)\, p + e(v)/2 , \tag{3.8}$$

which leads to

Theorem 3.1:

$$C(p,\alpha) \ge C_L(p,\alpha) := M_L\, [(a_L + e(v)/2)\, p + e(v)/2]\, e^{\alpha/p} .$$

Proof: From (3.5), we see that if (3.1) holds, then

$$h \le h_L(p,\alpha) := M_L^{-1}\, e^{-\alpha/p} ,$$

Using this result, (3.2), and (3.8), the theorem follows. ∎

Next, we consider upper bounds on the number of operations required. Instead of using Φ_{MBRK}, we will use the class Φ_{BRK} of "unmodified" BRK methods described in the Appendix, where it is shown that Φ_{BRK} is order-convergent in the sense of Werschulz [76b]. That is, there is an $M_U > 0$ such that

$$\sigma_G(\varphi_p,h) \le (M_U h)^p ; \tag{3.9}$$

no such bound is known for Φ_{MBRK}. In addition, Φ_{MBRK} requires the solution of p - 1

linear systems of equations, the i^{th} having $p - i$ unknowns, in order to perform a "reorthogonalization." So the smallest known combinatory cost for this class is about $O(p^{3.81})$ arithmetic operations; this is obtained by using Strassen's technique for linear systems (described in Borodin and Munro [75]). On the other hand, most of the combinatory cost for φ_p in Φ_{BRK} is involved in finding the coefficients of the polynomial p_{n+1} (see the Appendix); once these coefficients are known, the remaining combinatory cost is $O(p \ln p)$ as $p \uparrow \infty$. An estimate of how much work is required to compute these coefficients is given in

Lemma 3.1: Let $x_0, y_1, \ldots, y_r, w_0, z_0, \ldots, z_r$ be given, and let

$$Q(x) := \Sigma_{i=0}^{r+1} q_i x^i$$

be the unique polynomial of degree at most $r + 1$ satisfying

$$Q(x_0) = w_0, \quad Q'(x_0) = z_0, \quad \text{and} \quad Q'(y_i) = z_i \quad (1 \le i \le r) .$$

If $T(r)$ is the time required to compute $q_0, \ldots, q_{r+1}$, then

$$T(r) = O(r \ln^2 r) \text{ as } r \uparrow \infty .$$

Proof: The coefficients $q_1, 2q_2, \ldots, (r+1)q_{r+1}$ of Q' may be computed in time $O(r \ln^2 r)$ by using a fast algorithm for computing the coefficients of the Lagrange polynomial interpolating the points $(x_0, z_0), (y_1, z_1), \ldots, (y_r, z_r)$; see Borodin and Munro [75] for details. Then $O(r)$ operations yield $q_1, \ldots, q_{r+1}$, and Horner's rule gives q_0 with $O(r)$ additional operations. ∎

Thus there exists $a_U > 0$ such that

$$d(p) \le a_U p \ln^2(p+e) . \tag{3.10}$$

(We write "$\ln (p+e)$", where e is the base of the natural logarithms, rather than "$\ln p$" as a technical convenience. However, an expression of the form "$\ln (p+\gamma)$" with $\gamma > 0$ is necessary to guarantee that $d(1) > 0$.) In order to simplify matters a bit, note that Theorem A.1 of the Appendix implies that

(3.11) $$e(\mathfrak{N}_p(v)) \le e(v)\, p \quad .$$

Although the estimate above is not exact for $p > 2$, it is asymptotically equal to that in Theorem A.1. (If necessary, the sharper estimate given there may be used, but the calculation of optimal order (see below) involves considerably more detail; moreover, the asymptotic formulae for optimal complexity, order, and step-size are the same in either case.) Combining (3.10) and (3.11), we see that the cost per step is bounded by

(3.12) $$c_U(p) = e(v)\, p + a_U\, p\, \ln^2(p+e) \quad ,$$

which leads to

Theorem 3.2:

$$C(p,\alpha) \le C_U(p,\alpha) := M_U\, [e(v)\, p + a_U\, p\, \ln^2(p+e)]\, e^{\alpha/p} \quad .$$

Proof: If we set

$$h = h_U(p,\alpha) := M_U^{-1}\, e^{-\alpha/p} \quad ,$$

we find that (3.9) implies that (3.1) holds. Using this result, (3.2), and (3.12), the theorem follows. ∎

Thus we have found bounds

(3.13) $$C_L(p,\alpha) \le C(p,\alpha) \le C_U(p,\alpha)$$

on the number of operations required for a p^{th}-order NRK method to provide an approximate solution satisfying (3.1). We would like to compute

(3.14) $$C^*(\alpha) := \inf \{C(p,\alpha) : p \in \mathbb{Z}^{++}\} \quad .$$

This is not possible, since we only have bounds for $C(p,\alpha)$, and hence cannot compute $C(p,\alpha)$ exactly. However, we can pick optimal orders which minimize these bounds. First, we prove

Lemma 3.2: Define

$$G_L(p) := p^2\, c_L'(p) / c_L(p) \quad \text{and} \quad G_U(p) := p^2\, c_U'(p) / c_U(p) \quad .$$

Then for $p > 0$, we have $G_L'(p) > 0$ and $G_U'(p) > 0$.

Proof: Since c_L is a linear polynomial with a negative zero, the first part follows immediately. Now write $c_U(p) = c_1(p)\, c_2(p)$, where

$$c_1(p) := p \quad \text{and} \quad c_2(p) := 1 + \beta \ln^2(p + e),$$

with $\beta := a_U / e(v)$. Define

$$G_i(p) := p^2\, c_i'(p) / c_i(p) \qquad (i = 1, 2).$$

Clearly $G_1'(p) > 0$ if $p > 0$. Now

$$G_2(p) = 2\beta p^2 \ln (p+e) / D_2(p), \quad \text{where} \quad D_2(p) := (p+e) f_2(p),$$

so that

$$G_2'(p) = 2\beta p\, g_2(p) / D_2(p)^2,$$

where

$$g_2(p) := \beta p \ln^2(p+e) [\ln (p+e) - 1] + 2\beta e \ln^2(p+e) + (p + 2e) \ln (p+e) + p .$$

Thus $G_2'(p) > 0$ for $p > 0$. Since $G_U = G_1 + G_2$, the desired result follows. ∎

We now have the following

Theorem 3.3: For any $\alpha > 0$, there exist $p_L^*(\alpha)$ and $p_U^*(\alpha)$ such that

$$\alpha = G_L(p) \text{ iff } p = p_L^*(\alpha) \quad \text{and} \quad \alpha = G_U(p) \text{ iff } p = p_U^*(\alpha).$$

Moreover,

$$C_L^*(\alpha) := C_L(p_L^*(\alpha),\alpha) < C_L(p,\alpha) \quad \text{unless} \quad p = p_L^*(\alpha)$$

and

$$C_U^*(\alpha) := C_U(p_U^*(\alpha),\alpha) < C_U(p,\alpha) \quad \text{unless} \quad p = p_U^*(\alpha).$$

Proof: Using (3.5), (3.9), and Lemma 3.2, this follows immediately from Lemma 2.1 of Werschulz [76a]. ∎

From (3.13), (3.14), and the above Theorem, we have bounds

$$C_L^*(\alpha) \le C^*(\alpha) \le C_U^*(\alpha). \tag{3.15}$$

We call $p_L^*(\alpha)$ (respectively, $p_U^*(\alpha)$) the lower (upper) optimal order, $C_L^*(\alpha)$ (respectively, $C_U^*(\alpha)$) the lower (upper) optimal complexity, and

(3.16) $h_L^*(\alpha) := h_L(p_L^*(\alpha),\alpha)$ (respectively, $h_U^*(\alpha) := h_U(p_U^*(\alpha),\alpha)$)

the lower (upper) optimal step-size. We now examine how these quantities behave as α increases.

Theorem 3.4: $p_L^*(\alpha)$, $p_U^*(\alpha)$, $C_L^*(\alpha)$, and $C_U^*(\alpha)$ all increase monotonically and tend to infinity with α. Moreover, the following asymptotic formulae hold as α tends to infinity.

(1.) $p_L^*(\alpha) \sim \alpha$ and $p_U^*(\alpha) \sim \alpha$.

(2.) $C_L^*(\alpha) \sim M_L\, e\, [a_L + e(v)/2]\, \alpha$ and $C_U(\alpha) \sim M_U\, a_U\, e\, \alpha \ln^2\alpha$.

(3.) $h_L^*(\alpha) \sim (M_L\, e)^{-1}$ and $h_U^*(\alpha) \sim (M_U\, e)^{-1}$.

Proof: The first statement follows from Lemma 3.2 and from Theorem 2.3 of Werschulz [76b]. Now Lemma 3.2 implies that

$$G_L(p) \sim p \quad \text{and} \quad G_U(p) \sim p \quad \text{as } p \uparrow \infty .$$

Using this result and the fact that $\lim_{\alpha\uparrow\infty} p_L^*(\alpha) = \lim_{\alpha\uparrow\infty} p_U^*(\alpha) = +\infty$, (1.) follows. Finally, (2.) and (3.) follow from (1.), Theorem 3.1, and Theorem 3.2. ∎

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

(3.17) $$C_L^*(\alpha) = O(\alpha) \le C^*(\alpha) \le C_U^*(\alpha) = O(\alpha \ln^2\alpha)$$

as α tends to infinity; so, the ratio

$$C_U^*(\alpha) / C_L^*(\alpha) = O(\ln^2\alpha) \quad \text{as } \alpha \uparrow \infty$$

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

Finally, we wish to compare the complexities of NRK methods, Taylor series methods, and linear Runge-Kutta ("LRK") methods. We write $C_{U,NRK}^*$, $C_{U,LRK}^*$, $C_{U,T}^*$ for C_U^* in the class of NRK methods, LRK methods, and Taylor series methods; other notations ($C_{L,LRK}^*$, C_{LRK}^*, etc.) are formed in an analogous manner. Finaly, if $f, g : \mathbb{R}^{++} \to \mathbb{R}^{++}$ satisfy $\lim_{\alpha\uparrow\infty} f(\alpha) = \lim_{\alpha\uparrow\infty} g(\alpha) = +\infty$, we write

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

we say f is asymptotically less than g. (See Section 5 of Werschulz [76a].) We then have

Theorem 3.5:

(1.) $C_{U,NRK}^* < C_{L,LRK}^*$.

(2.) $C_{U,T}^* < C_{U,NRK}$ if the cost of evaluating the k^{th} derivative of v is bounded for all k.

Proof: Immediate from (3.20) and (4.14) of Werschulz [76a] and (3.17). ∎

As a corollary we see that $C_{NRK}^* < C_{LRK}^*$, so that the best NRK method known is better than the best LRK method possible. Moreover, if the derivatives of v are easy to evaluate, the best Taylor series method known is better than the best NRK method known. However, if the cost of evaluating the k^{th} derivative of v increases faster than O(ln k) as $k \uparrow \infty$, then it is easy to show that the opposite will be true.

Appendix: Order-Convergence of a Basic Sequence

In this Appendix, we describe a subclass of a class of iterative methods for the solution of scalar nonlinear equations. This subclass will then be used to generate an order-convergent basic sequence Φ_{BRK} of nonlinear Runge-Kutta methods.

Lemma A.1: Let $F: D \subset \mathbb{R} \to \mathbb{R}$ have a simple zero ξ, and suppose that F is analytic at ξ. Pick $k, m \in \mathbb{Z}^{++}$ with $m + 1 \geq k$. Then there is a sequence $\Psi_{km} := \{\psi_{kmn} : n \in \mathbb{Z}^{++}\}$ of stationary multipoint methods without memory such that the following hold:

(1.) The method ψ_{kmn} uses the information

$$\mathfrak{N}_{kmn}(F) := \{F(x_0), \ldots, F^{(m)}(x_0), F^{(k)}(y_1), \ldots, F^{(k)}(y_n)\}$$

(the points $y_1, \ldots, y_n$ being suitably chosen) to compute a new approximation x_1 to ξ from a given approximation x_0 by setting

$$x_1 := \psi_{kmn}(x_0) .$$

(2.) There exists a $B > 0$ and an $h_0 > 0$ such that if $|x_0 - \xi| \leq h_0$, then

$$|x_1 - \xi| \leq (B\,|x_0 - \xi|)^{\rho} \quad \text{for all } n \in \mathbb{Z}^{++},$$

where

$$\rho := \min (m + 2n + 1, 2m + n + 1) . \tag{A.1}$$

Before proving the Lemma, we describe how the method ψ_{kmn} computes an improved approximation x_1 from the old approximation x_0.

Algorithm for computing $x_1 := \psi_{kmn}(x_0)$.

(1.) Let $\delta := |F(x_0)/F'(x_0)|$.

(2.) Let z_1 be an approximate zero of

$$p_1(x) := \Sigma_{i=0}^{m} (x - x_0)^i F^{(i)}(x_0) / i!$$

satisfying

(A.2) $$z_1 = x_0 + O(\delta) \quad \text{and} \quad |p_1(z_1)| \le (A_1 \delta)^{m+1},$$

where A_1 is independent of n.

(3.) Let

$$y_i := x_0 + \alpha_{in} (z_1 - x_0) \quad (1 \le i \le n),$$

where

$$\alpha_{in} := (1 + x_{in}) / 2$$

and $x_{1n} > \ldots > x_{nn}$ are the zeros of the Jacobi polynomial

$$P_n(x) := P_n^{(k-1,\, m+1-k)}(x)$$

(see Szegő [59]).

(4.) Let p_{n+1} be the polynomial of degree at most m + n that interpolates the information $\mathfrak{N}_{kmn}(F)$, and let x_1 be an approximate zero of p_{n+1} satisfying

(A.3) $$x_1 = x_0 + O(\delta) \quad \text{and} \quad |p_{n+1}(x_1)| \le (A_2 \delta)^{\rho},$$

where A_2 is independent of n and ρ is given by (A.1).

Here we use the notation of Brent [74]. Clearly, $\psi_{kmn} \in C'(k, m, n)$, the only difference being that conditions (A.2) and (A.3) replace (2.2) and (2.4) of Brent [74]. It is easy to see that (A.2) and (A.3) may be realized by using $\lceil \log_2(m+1) \rceil - 1$ and $\lceil \log_2(\rho/(m+1)) \rceil$ iterations of Newton's method, with the respective starting approximations of $x_0 - F(x_0) / F'(x_0)$ and z_1.

Proof of Lemma A.1: Let x_1' be the exact zero of p_{n+1} near x_0. We then find that there is a ξ between x_1' and z_1 such that

$$|F(x_1')| \le |p_{n+1}(z_1) - F(z_1)| + |p_{n+1}'(\xi) - F'(\xi)|\,|x_1' - z_1| \quad (A.4)$$

Using (A.3), the analyticity of F, and standard techniques of interpolation theory (Traub [64]), it is easy to show that (2.9) and (2.10) of Brent [74] may be rewritten as

$$\begin{aligned} |p_{n+1}(x) - F(x)| &\le (A_3\,\delta)^{m+n+1} \quad \text{and} \\ |p_{n+1}'(x) - F'(x)| &\le (A_4\,\delta)^{m+n} \end{aligned} \quad (A.5)$$

for $|x - x_0| \le 4\delta$. (Here all constants A_r will be independent of n.) Similarly, we find that

$$|x_1' - \xi| \le (A_5\,\delta)^{m+n} \quad \text{and} \quad |z_1 - \xi| \le (A_6\,\delta)^{m+1},$$

so that the triangle inequality gives

$$|x_1' - z_1| \le (A_7\,\delta)^{m+1}. \quad (A.6)$$

Using (A.4), (A.5), and (A.6), we see that

$$\begin{aligned} |f(x_1')| &\le |p_{n+1}(z_1) - F(z_1)| + (A_8\,\delta)^{2m+n+1} \\ &\le |p_{n+1}(z_1) - F_1(z_1)| + |F_2(z_1)| + (A_8\,\delta)^{2m+n+1}, \end{aligned} \quad (A.7)$$

where

$$F_1(x) := \sum_{i=0}^{m+2n} (x - x_0)^i\, F^{(i)}(x_0)\,/\,i! \quad \text{and} \quad F_2(x) := F(x) - F_1(x).$$

Clearly $|F_2(x)| \le (A_9\,\delta)^{m+2n+1}$, so that (A.7) becomes

$$|F(x_1')| \le |p_{n+1}(z_1) - F_1(z_1)| + (A_{10}\,\delta)^{\rho}. \quad (A.8)$$

As in Brent [74], we now write

$$p_{n+1}(x) = r_1(x) + r_2(x),$$

where r_i $(i = 1, 2)$ is the polynomial of degree at most m + n satisfying

$$r_i^{(j)}(x_0) = F_i^{(j)}(x_0) \quad (0 \le j \le m)$$

and

$$r_i^{(k)}(y_j) = F_i^{(k)}(y_j) \quad (1 \le j \le n).$$

If we let

$$P(x) := r_1(x + x_0) - F_1(x + x_0) ,$$

and write $\varepsilon := z_1 - x_0$ (in this Appendix only), we find that

$$P^{(i)}(0) = 0 \quad (0 \le i \le m) \quad \text{and} \quad P^{(k)}(\alpha_{in}\, \varepsilon) = 0 \quad (1 \le i \le n) .$$

We may easily alter the proof of Lemma 4.3 in Brent [74] to show that

$$r_1(z_1) - F_1(z_1) = P(\varepsilon) = 0 .$$

Thus (A.8) becomes

$$|F(x_1')| \le |r_2(z_1)| + (A_{10}\, \delta)^{\rho} . \tag{A.9}$$

To bound the remaining term, let us write

$$r_2(x) = \Sigma_{j=1}^{m} a_{j+m} (x - x_0)^{j+m} ,$$

recalling that r_2 has a zero of multiplicity m at x_0. Using the notation of Stewart [73], we see that the nonzero coefficients of r_2 are given by the solution of the linear system

$$Wy = c ,$$

where

$$\omega_{ij} := \alpha_{in}^{j-1} \quad (1 \le i, j \le n) ,$$

$$\eta_j := a_{j+m}\, \varepsilon^{j+m}\, (j + m)! \,/\, (j + m - k)! \quad (1 \le j \le n) , \text{ and}$$

$$\gamma_i := \varepsilon^{k}\, F_2^{(k)}(y_i) \,/\, \alpha_{in}^{m-k+1} \quad (1 \le i \le n) .$$

Since W^T is a Vandermonde matrix, we find that the entries of $U = W^{-1}$ are given by

$$u_{ij} = \alpha_{jn}\, (-1)^{n-i}\, \sigma_{n-i,n-1,j} \,/\, \Pi_{r \ne j}\, (\alpha_{jn} - \alpha_{rn}) ,$$

where

$$\sigma_{\mu,n-1,j} := \Sigma\ \alpha_{p_1,n} \cdots \alpha_{p_\mu,n} ,$$

the sum being taken over all multi-indices $p_1 \cdots p_\mu$ not including j (Gregory and Karney [69]). Since there are fewer than 2^n summands, each of which lies in [0, 1], we see that $\sigma_{\mu,n-1,j} \le 2^n$, implying that

$$|v_{ij}| \le 2^n \alpha_{jn} / \Pi_{r \neq j} (\alpha_{jn} - \alpha_{rn}) .$$

So we have

$$\begin{aligned} |\eta_i| &= |\Sigma_{j=1}^n v_{ij} \gamma_j| \\ &\le n\, 2^n \max_{1 \le j \le n} | \epsilon^k F_2^{(k)}(y_j) / [\alpha_{jn}^{m-k} G_n'(\alpha_{jn})] | , \end{aligned} \tag{A.10}$$

where

$$G_n(x) := G_n(m+1, m+2-k, x) = \Pi_{r=1}^n (x - \alpha_{rn})$$

(see Abramowitz and Stegun [64]).

Now it is clear that

$$\max_{1 \le j \le n} 1 / \alpha_{jn}^{m-k} = 1 / \alpha_{nn}^{m-k} .$$

By Theorem 8.9.1 of Szegő [59], we may show that

$$\alpha_{nn} \ge A_{11}\, n^{-2} ;$$

using this result and (22.5.2) of Abramowitz and Stegun [64], we find that

$$\begin{aligned} &\max_{1 \le j \le n} [\alpha_{jn}^{m-k} G_n'(\alpha_{jn})]^{-1} \\ &\le A_{12}\, n^{2(m-k)} \binom{m+2n+1}{m} \max_{1 \le j \le n} |P_n'(x_{jn})|^{-1} , \end{aligned} \tag{A.11}$$

By the symmetry relation (4.1.3) of Szegő [59], we may assume that $0 \le x_{jn} < 1$. Using Theorem 8.9.1 of Szegő [59], we may show that

$$|P_n'(x_{jn})|^{-1} \le (A_{13})^n ,$$

and so (A.10), (A.11), the definition of F_2, and the above imply that

$$|\eta_i| \le (A_{14}\, \delta)^{m+2n+1} ,$$

yielding the result

$$|r_2(z_1)| \le \Sigma_{j=1}^n a_{j+m} \epsilon^{j+m} \le n \max_{1 \le i \le n} |\eta_i| \le (A_{15}\, \delta)^{m+2n+1} .$$

So (A.9) becomes

$$|F(x_1')| \le (A_{16}\, \delta)^{\rho} .$$

By Taylor's Theorem, this implies

$$|x_1' - \xi| \le (A_{17}\,\delta)^{\rho} .$$

The desired result then follows from (A.3) and from (2.5) of Brent [74]. ∎

We now describe the basic sequence Φ_{BRK}. The methods in this basic sequence are given by

$$\varphi_1(x_0 , h) := v(x_0) ,$$

$$\varphi_2(x_0 , h) := v(x_0 + h\, v(x_0) / 2) ,$$

and for $p \ge 2$,

$$\varphi_p(x_0 , h) := h^{-1} [\psi_{1,1,p-2}(x_0) - x_0] ,$$

with $\psi_{1,1,p-2}$ applied to the function F given by (2.8) and the approximation x_1 to x_1' being given by an appropriate number of iterations of Newton's method (as described above).

Theorem A.1: The basic sequence Φ_{BRK} is order-convergent with respect to the global error. Moreover, the number of stages $s(p)$ required by $\varphi_p \in \Phi_{BRK}$ is given by

$$s(p) = \begin{cases} p & \text{if } p \le 2 \\ p - 1 & \text{if } p > 2 \end{cases} .$$

Proof: We use the notation of Lemma A.1, writing $z(h)$ for the computed p^{th}-order approximation x_1 to $x(h)$ and $p_{n+1}(\cdot , x_0)$ for the polynomial p_{n+1}. The result of Lemma A.1 is that

$$h^{-1} |z(h) - x(h)| \le (B\,h)^p ,$$

the desired result for a single unit step. To prove the global result, we must consider the Lipschitz constants for Φ_{BRK}.

We implicitly differentiate the result $p_{n+1}(x_1', x_0) \equiv 0$ to find

$$\partial_1 \varphi_p(x_0 , h) = -h^{-1}\, Q_{n+1}(x_1', x_0) + s_p(x_0) ,$$

where

$$Q_{n+1}(x_1', x_0) = 1 + \partial_2 p_{n+1}(x_1', x_0) / \partial_1 p_{n+1}(x_1', x_0)$$

and

$$\delta_p(x_0) = h^{-1} (d/dx_0) [x_1 - x_1'] .$$

It is easy to see that x_1 and x_1' are analytic functions of x_0. Since their difference tends to zero uniformly on the domain of v as $p \uparrow \infty$, it follows that

$$\lim_{p \uparrow \infty} \delta_p(x_0) = 0 .$$

We claim that

$$Q_{n+1}(x_1', x_0) = O(h \ln n) \quad \text{as} \quad n \uparrow \infty ,$$

uniformly in x_0. To see this, note that we may write the interpolation polynomial p_{n+1} in terms of Jacobi polynomial P_n, finding that

$$p_{n+1}(x,x_0) = (-1)^n (h/2) \int_{-1}^{\xi(x)} P_n(t)\, dt + h\, v(x_0) \sum_{k=1}^{n} I_{kn} - h ,$$

where

$$\xi(x) := 2 (x - x_0) / [h\, v(x_0)] - 1$$

and

$$I_{kn} := [2 (1 + x_{kn})\, v(y_k)\, P_n'(x_{kn})]^{-1} \int_{-1}^{\xi(x)} (t + 1)\, P_n(t) / (t - x_{kn})\, dt .$$

Now

$$\partial_1 p_{n+1}(x_1', x_0) = (-1)^n P_n(\xi_1) / v(x_0) + (1 + \xi_1) \sum_{k=1}^{n} g(x_{kn})\, L_{kn}(\xi_1) ,$$

where

$$\xi_1 := \xi(x_1') ,$$

$$L_{kn}(x) := P_n(x) / [P_n'(x_{kn})\, (x - x_{kn})] , \text{ and}$$

$$g(t) := 1 / [(1 + t)\, v(x_0 + (1 + t)\, h\, v(x_0) / 2)] .$$

By (8.21.10) of Szegö [59], the first term in the expression for $\partial_1 p_{n+1}(x_1', x_0)$ goes to zero as $n \uparrow \infty$. A minor modification of the proof of Theorem 14.4 of Szegö [59] shows that the sum in the remaining term tends to $g(\xi(x(h)))$ as $n \uparrow \infty$. So

$$\partial_1 p_{n+1}(x_1', x_0) \sim v(x(h))^{-1} \quad \text{as} \quad n \uparrow \infty .$$

Using Lemma A.1 of the Appendix in Werschulz [76a] and techniques similar to those yielding the above estimate, we find

$$\partial_2 P_{n+1}(x_1', x_0) = O(h \ln n) - v(x(h))^{-1} \quad \text{as } n \uparrow \infty .$$

This gives the estimate claimed for $Q_{n+1}(x_1', x_0)$.

So the Lipschitz constant for $\varphi_p \in \Phi_{BRK}$ grows as the logarithm of p. By Proposition 4.3 of Werschulz [76b], Φ_{BRK} is order-convergent. ∎

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 Section 3 were obtained by using the MACSYMA system developed by the Mathlab 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.

References

Abramowitz and Stegun [64]:
Abramowitz, M., and I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Washington, D. C.: National Bureau of Standards, 1964.

Ahlfors [66]:
Ahlfors, L. V., Complex Analysis, Second Edition. New York: McGraw-Hill, 1966.

Borodin and Munro [75]:
Borodin, A. and I. Munro, The Computational Complexity of Algebraic and Numeric Problems. New York: American Elsevier, 1975.

Brent [74]:
Brent, R. P., "Efficient Methods for Finding Zeros of Functions Whose Derivatives are Easy to Evaluate," Report, Computer Science Department, Carnegie-Mellon University, 1974.

Brent [76]:
Brent, R. P., "A Class of Optimal-Order Zero Finding Methods Using Derivative Evaluations," in Analytic Computational Complexity, edited by J. F. Traub. New York: Academic Press, 1976.

Buck [65]:
Buck, R. C., Advanced Calculus, Second Edition. New York: McGraw-Hill, 1965.

Gregory and Karney [69]:
Gregory, R. T. and D. L. Karney, A Collection of Matrices for Testing Computational Algorithms. New York: Wiley-Interscience, 1969.

Henrici [62]:
Henrici, P., Discrete Variable Methods in Ordinary Differential Equations. New York: Wiley, 1962.

Meersman [76]:
Meersman, R., "On Maximal Order of Families of Iterations for Nonlinear Equations," Doctoral Thesis, Vrije Universiteit Brussel, Brussels, 1976.

Stetter [73]:
Stetter, H. J., Analysis of Discretization Methods for Ordinary Differential Equations. Berlin: Springer-Verlag, 1972.

Stewart [73]:
Stewart, G. W.., Introduction to Matrix Computations. New York: Academic Press, 1973.

Szegö [59]:
Szegö, G., Orthogonal Polynomials. Amer. Math. Soc. Colloquium Publications, Vol. XXIII. New York: Amer. Math. Soc., 1959.

Traub [64]:
Traub, J. F., Iterative Methods for the Solution of Equations. Englewood Cliffs: Prentice-Hall, 1964.

Traub and Woźniakowski [76]:
Traub, J. F. and H. Woźniakowski, "Strict Lower and Upper Bounds on Iterative Complexity," in Analytic Computational Complexity, edited by J. F. Traub. New York: Academic Press, 1976.

Werschulz [76a]:
Werschulz, A. G., "Computational Complexity of One-Step Methods for Systems of Differential Equations." Report, Computer Science Department, Carnegie-Mellon University, 1976.

Werschulz [76b]:
Werschulz, A. G., "Optimal Order and Minimal Complexity of One-Step Methods for Initial Value Problems." Report, Computer Science Department, Carnegie-Mellon University, 1976.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* COMPUTATIONAL COMPLEXITY OF ONE-STEP METHODS FOR A SCALAR AUTONOMOUS DIFFERENTIAL EQUATION		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR*(s)* Arthur G. Werschulz		8. CONTRACT OR GRANT NUMBER*(s)* N00014-76-C-0370, NR 044-422
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept. Pittsburgh, PA 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE September 1976 13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

20. *Abstract:* The problem is to calculate an approximate solution of an initial value problem for a scaiar autonomous differential equation. A generalized notion of a nonlinear Runge-Kutta (NRK) method is defined. We show that the order of any s-stage NRK method cannot exceed 2s - 1; hence, the family of NRK methods due to Brent has the maximal order possible. Using this result, we derive complexity bounds on the problem of finding an approximate solution with error not exceeding ε. We also compute the order which minimizes these bounds, and show that this optimal order increases as ε decreases, tending to infinity as ε tends to zero.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*