VARIABLE ORDER ADAMS CODES

LAWRENCE F. SHAMPINE*

Abstract. Variable step size, variable order (VSVO) Adams codes are very effective for solving initial value problems for first order systems of ordinary differential equations. The theory of fixed order codes is classical, but when the order is varied, there is no theory explaining fundamental issues. With realistic assumptions about order and step size selection, we prove convergence, approximate locally the behavior of the error, and justify standard error estimators.

Key words. Adams codes, variable order, convergence, error estimation

AMS subject classifications. 65L05, 65L06

1. Introduction. The influential comparison of Hull et alia [1] showed that variable step size, variable order (VSVO) Adams codes are a very effective way to solve the initial value problem for a first order system of ordinary differential equations (ODEs). It directed attention to two codes of this kind, Krogh's DVDQ [2] and Gear's DIFSUB [3], that differ in important ways. Each came into very wide use and each spawned a line of successors. Despite the importance of VSVO Adams codes, there is no theory explaining issues of the most fundamental nature. This is entirely due to the variation of order because when the order is fixed, the theory of Adams methods is classical and complete enough to describe adequately what is seen in practice.

A wealth of computational experience says that for the DVDQ line of codes [2, 4, 5, 6], as a tolerance $\tau \to 0$, the error is $O(\tau)$. This has passed into the folklore of the subject, but it has never been proven. Stability results that allow the order and step size to be varied are found in [5, 7, 8], but neither these investigations nor the texts [3, 4, 5, 8, 9] that present VSVO Adams methods go on to prove this convergence result. That is because they do not take into account how order and step size are selected in practice. Here we prove that with a reasonable model of step size and order selection in the DVDQ line of VSVO Adams codes, the error is uniformly $O(\tau)$. Experience with the DIFSUB line of codes has shown that as $\tau \to 0$, the behavior of the error is similar, but less regular. We do not know how to characterize this behavior, but we prove that in the circumstances we investigate, it is not $O(\tau)$. For this reason we study here only the DVDQ line of codes.

Step size and order selection depend on estimating the local truncation error. This is more complicated and difficult than in fixed order codes because order selection algorithms require estimates of the local truncation error that would have been made if formulas of other orders had been used. An estimate of the error that would have been made with a formula of higher order is especially hard to justify. The classical theory of error estimators requires that the order is constant and the step sizes are given by a step size selection function. Neither of these requirements is satisfied when integrating with a VSVO Adams code, so we must take a completely different approach. We approximate the local behavior of the error in the course of an integration with specific tolerance. This approximation and our result about the global behavior of the error are then used to justify standard estimates of the local truncation error, including the error of a formula of higher order.

*Math. Dept., Southern Methodist University, Dallas, TX 75275. (lshampin@mail.smu.edu)

2. Preliminaries. We approximate numerically the solution $y(x)$ of a first order system of ODEs

$$y' = f(x, y) \tag{2.1}$$

for $a \le x \le b$ with given initial value $y(a)$. The function $f(x, y)$ is assumed to be as smooth as necessary. In particular, it is continuous and satisfies a Lipschitz condition with respect to y with constant L. It is convenient to write $J(x) = f_y(x, y(x))$ and $Y^{(m)} = \max \|y^{(m)}(x)\|$.

We consider a sequence of integrations as a tolerance τ tends to zero. For each τ, the solver determines a mesh $a = x_0 < x_1 < \ldots < x_N = b$ and computes $y_n \approx y(x_n)$. The step size $h_n = x_{n+1} - x_n$. The integration begins with $y_0 = y(a)$. In this paper we study algorithms characteristic of the DVDQ line of VSVO Adams codes. In particular, the integration is performed with a fully variable step size implementation of an Adams-Bashforth-Moulton PECE scheme with local extrapolation. Details about Adams formulas and their use in such codes can be found in [2, 4, 5, 9, 10]. We assume there is a constant $\Gamma > 0$ such that

$$h_{n+1} \le \Gamma h_n \tag{2.2}$$

for all n and all integrations.

When we speak of a step taken at order k, we mean that the Adams-Bashforth formula of order k is used as a predictor and the Adams-Moulton formula of order $k + 1$ is used as corrector. Such a step has the form of a prediction

$$p_{n+1} = y_n + h_n \sum_{j=1}^{k} \alpha_{k,j} f(x_{n+1-j}, y_{n+1-j}) \tag{2.3}$$

followed by a correction

$$y_{n+1} = y_n + h_n \sum_{j=1}^{k} \alpha^*_{k+1,j} f(x_{n+1-j}, y_{n+1-j}) + h_n \alpha^*_{k+1,0} f(x_{n+1}, p_{n+1})$$

In a fully variable step size implementation, the coefficients here depend on the step sizes $h_n, \ldots, h_{n+1-k}$. Often it is convenient to write $f_m = f(x_m, y_m)$ and $f^p_m = f(x_m, p_m)$. We suppose that formulas of orders $1 \le k \le K$ are available. Any order for which the formula can be evaluated might be used at any step. This limits the order in the first few steps of the integration and in particular, requires that the formula of order 1 be used for the first step. When the order k is not clear from context, we include it in the notation. For instance, the Adams-Bashforth formula of order $k - 1$ forms $p_{n+1}(k - 1)$.

All the numerical examples of this paper involve the A3 problem of [1]

$$y' = cos(x)y, \quad 0 \le x \le 20, \quad y(0) = 1 \tag{2.4}$$

This is a standard test problem and it is simple enough that we can work out some analytical expressions needed to illustrate the results of §5. Unless stated otherwise, the `ode113` solver of the MATLAB ODE Suite [6] was used for the examples. The solver was given an absolute error tolerance of τ and a relative error tolerance of 100 units of roundoff, which for the A3 problem is effectively a pure absolute error control with tolerance τ.

3. Convergence. Many years of experience with the DVDQ line of VSVO Adams codes have shown that in a sequence of integrations with tolerance τ tending to zero, the error is uniformly $O(\tau)$ for smooth problems, see for example the tables of Krogh [11] and Shampine and Gordon [5, Chap. 11]. This behavior is clear in Fig. 3.1 where we display $\max_n |y_n - y(x_n)|$ when integrating equation (2.4) with `ode113`. The assumptions of §2 are realistic for the DVDQ line of codes. With an additional, realistic assumption about how the step size is related to the tolerance and the order of the formula, we prove that the error is uniformly $O(\tau)$.

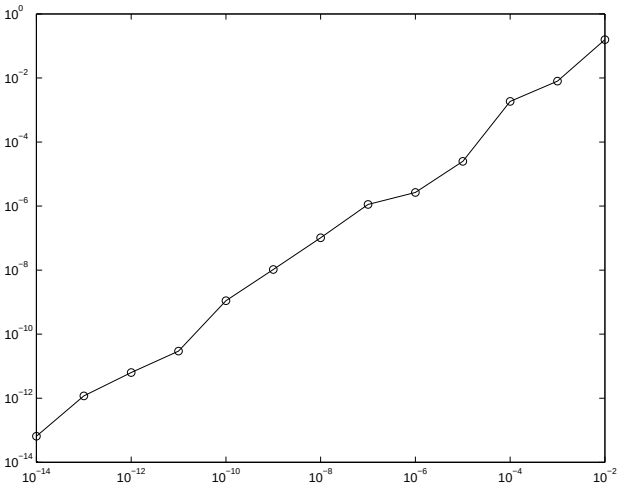

FIG. 3.1. *Log-log plot of maximum error as a function of* τ.

The local truncation error of the Adams-Bashforth formula of order k, lte_k, is defined by

$$y(x_{n+1}) = y(x_n) + h_n \sum_{j=1}^{k} \alpha_{k,j} f(x_{n+1-j}, y(x_{n+1-j})) + lte_k \tag{3.1}$$

and the local truncation error of the Adams-Moulton formula of order $k+1$, lte^*_{k+1}, is defined by

$$\begin{aligned} y(x_{n+1}) = y(x_n) &+ h_n \sum_{j=1}^{k} \alpha^*_{k+1,j} f(x_{n+1-j}, y(x_{n+1-j})) \\ &+ h_n \alpha^*_{k+1,0} f(x_{n+1}, y(x_{n+1})) + lte^*_{k+1} \end{aligned}$$

Standard results about Adams methods [10] state that with the assumptions of §2 and in particular, the upper bound (2.2) on the ratio of successive step sizes, there are constants C_1, C_2, C_3 such that

$$C_1 \geq \sum_{j=1}^{k} |\alpha_{k,j}|, \quad C_2 \geq \sum_{j=1}^{k} |\alpha^*_{k+1,j}|, \quad C_3 \geq |\alpha^*_{k+1,0}|$$

for all x_n and $1 \leq k \leq K$. Let $\mathcal{L} = LC_2 + LC_3(1 + (b-a)LC_1)$.

For each value m, we shall define E_m so that

$$E_m \geq \max_{0 \leq j \leq m} \|y_j - y(x_j)\|$$

We begin with $E_0 = 0$ because the integration starts with $y_0 = y(a)$. A little manipulation of the definitions of the formulas and the local truncation errors shows that

$$\|y_{n+1} - y(x_{n+1})\| \leq E_n(1 + h_n\mathcal{L}) + \|lte^*_{k+1}\| + h_n L C_3 \|lte_k\|$$

Suppose now that

$$\|lte^*_{k+1}\| + h_n L C_3 \|lte_k\| \leq h_n \sigma \tag{3.2}$$

for all steps. With this assumption it is easy to see that we can define

$$E_{n+1} = E_n(1 + h_n\mathcal{L}) + h_n\sigma$$

An induction argument then shows that

$$E_m \leq \sigma(x_m - a)e^{\mathcal{L}(x_m - a)}$$

for all m. This implies that for all $a \leq x_m \leq b$,

$$\|y_m - y(x_m)\| \leq \sigma(b - a)e^{\mathcal{L}(b-a)} \tag{3.3}$$

The conclusion (3.3) is essentially Theorem 4 of [5]. We have included the definitions and arguments for completeness and because we need most of them later. The question now is, how do codes select the step size so that the inequality (3.2) holds? Step size and order selection algorithms are complicated and involve many heuristics. We do not need here any details, just the fact that when controlling the local truncation error by a criterion of error per step (EPS), standard algorithms result in a step size that is $O(\tau^{1/(k+1)})$ for a formula of order k. We also suppose that there is a constant $G > 0$ such that

$$h_n \leq G h_{n+1} \tag{3.4}$$

for all n and all integrations. For smooth problems this bound and the bound (2.2) taken together amount to a mild assumption about the efficiency of the step size selection algorithm.

The local truncation error of the Adams-Bashforth formula of order k can be written as

$$lte_k = \frac{y^{(k+1)}(\xi)}{k!} \int_{x_n}^{x_{n+1}} \prod_{i=1}^{k} (t - x_{n+1-i})\, dt \tag{3.5}$$

for some $\xi \in [x_{n+1-k}, x_{n+1}]$. If we let $t = x_n + sh_n$, the integral becomes

$$h_n^{k+1} \int_0^1 s \prod_{i=2}^{k} \left(s + \frac{x_n - x_{n+1-i}}{h_n} \right) ds$$

With the assumption (3.4), we have

$$x_n - x_{n+1-i} = h_{n-1} + \ldots + h_{n+1-i} \leq (G + \ldots G^{i-1})h_n$$

With this observation and $\|y^{(k+1)}(\xi)\| \le Y^{(k+1)}$, it is clear that lte_k is $O(h_n^{k+1})$. Similarly, we can write

$$lte_{k+1}^* = \frac{y^{(k+2)}(\eta)}{(k+1)!} \int_{x_n}^{x_{n+1}} \prod_{i=0}^{k} (t - x_{n+1-i})\, dt \tag{3.6}$$

and prove it to be $O(h_n^{k+2})$. With our assumption that h_n is $O(\tau^{1/(k+1)})$, we find that the inequality (3.2) holds with a σ that is $O(\tau)$. In conjunction with the inequality (3.3) we conclude that the error is uniformly $O(\tau)$. We have now proven

THEOREM 3.1. *In addition to the assumptions of §2, suppose that the restriction (3.4) holds and that the step sizes used at order k are uniformly $O(\tau^{1/(k+1)})$. With these assumptions, the error is $O(\tau)$ uniformly on $[a, b]$.*

In the circumstances we investigate, the error is not $O(\tau)$ when using algorithms characteristic of the DIFSUB line of VSVO Adams codes. To see this, suppose that the formula of order 1 is used for all the steps. The asymptotic behavior of this one-step method is worked out in [12] for a variety of error controls with assumptions about the step size selection algorithm that correspond to the ones made here. With the control of the DIFSUB line of codes, namely EPS and no local extrapolation, we proved in [12] that $y_n - y(x_n) \sim e(x_n)\sqrt{\tau}$.

4. Behavior of the Error. In this section we show that for a given tolerance τ, if we begin using order k at x_n and let $H = \tau^{1/(k+1)}$, then for a few steps that are all taken at order k,

$$y_m = y(x_m) + e(x_m, \tau)\tau + O(H\tau) \tag{4.1}$$

VSVO Adams codes often take a number of steps at the same order. Indeed, they do not even consider increasing the order until $k + 1$ steps have been taken at order k. The equation (4.1) can be written in a form

$$y_m = y(x_m) + e(x_m, \tau)H^{k+1} + O(H^{k+2})$$

that resembles a classical asymptotic expansion

$$y_m = y(x_m) + e(x_m)H^{k+1} + O(H^{k+2})$$

but the situation is entirely different. The classical theory requires that the same order be used throughout the integration. The function $e(x)$ does not depend on H and the approximation is more accurate as H is reduced. In contrast, (4.1) describes the behavior of the error in the course of an integration with a specific tolerance τ. The function $e(x, \tau)$ depends on (x_n, y_n). Both x_n and y_n depend on the orders and step sizes used prior to x_n, and the algorithms for selecting them depend on τ. The approximation of (4.1) is exact at x_n and we expect it to become less accurate as the integration proceeds.

Using MATLAB [13] and `ode113` it is easy to study experimentally the behavior of the scaled error $(y_n - y(x_n))/\tau$. We modified the solver so that it would report the order k used at each step. Fig. 4.1 shows the scaled error when solving equation (2.4) with $\tau = 10^{-4}$. In this computation the solver took 101 steps and used orders k that ranged from 1 to 10. This solver can change the order and step size often in the course of an integration. A relatively long sequence of points computed with order 6 and similarly, a sequence computed with order 8, are distinguished in the

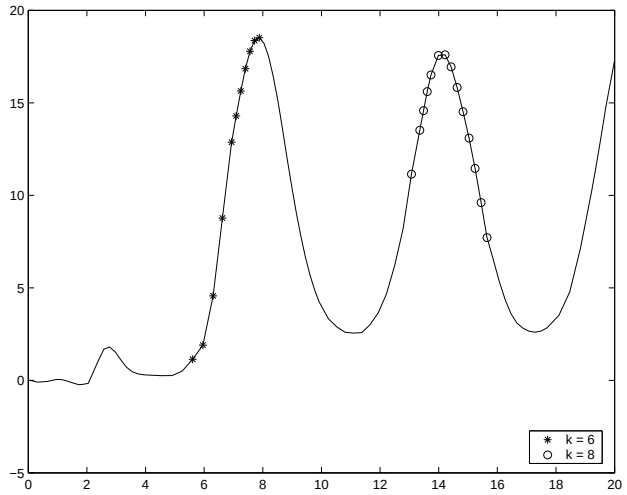

FIG. 4.1. *Scaled error* $(y_n - y(x_n))/\tau$ *with tolerance* $\tau = 10^{-4}$.

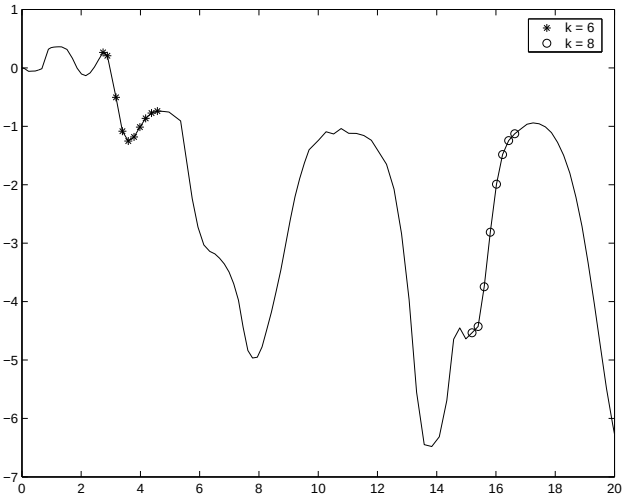

FIG. 4.2. *Scaled error* $(y_n - y(x_n))/\tau$ *with tolerance* $\tau = 5 \times 10^{-5}$.

figure. The solver changed step size twice in the course of the first sequence and four times in the second. It appears that despite prior changes of step size and order, the error at successive steps with the same order has a regular behavior. And, this is so immediately after changing order and in the presence of changes of step size. Fig. 4.2 shows what happens when the tolerance is reduced to $\tau = 5 \times 10^{-5}$. Comparison of the two figures shows that the behavior of the scaled error depends strongly on the tolerance. On the other hand, we again see a regular behavior of the error for successive steps taken at the same order. This is the behavior we aim to approximate.

The VSVO Adams codes control the local truncation error by a criterion of error per step (EPS). With this criterion, the optimal step size h_n at order k is defined by $\tau = \|lte_k^*\|$. When the step size is slowly varying, (3.6) tells us that to leading order this local truncation error is equal to $\gamma_k^* \, h_n^{k+1} \, y^{(k+1)}(x_n)$. Using this observation, the

selection of step size is idealized by defining h_n so that

$$\tau = h_n^{k+1} \, \|\gamma_k^* \, y^{(k+1)}(x_n)\|$$

This step size can be expressed as

$$h_n = \theta_k(x_n) \, \tau^{1/(k+1)}$$

with

$$\theta_k(x) = \left(\frac{1}{\|\gamma_k^* \, y^{(k+1)}(x)\|} \right)^{1/(k+1)}$$

At each step of an integration with tolerance τ, an ideal VSVO Adams code estimates the optimal step size for not only the current order k, but also for order $k-1$ and possibly order $k+1$. For smooth problems, these optimal step sizes are smooth functions of x. Ideally the algorithm changes the order where the optimal step size at order $k-1$ or $k+1$ matches and later exceeds the optimal step size at order k. In this idealization the step size is a continuous function of x.

We have described a scheme that idealizes certain aspects of practical schemes for the selection of step size and order. It is a reasonable description of practice, but for our analysis, we need only certain properties of the scheme, namely

VSVO Assumptions: *For each tolerance τ, the step size and order selection algorithms partition $[a, b]$ into subintervals on which a constant order is used. On a subinterval where order k is used, the step sizes are defined by $h_m = \theta_k(x_m) \, \tau^{1/(k+1)}$. Each $\theta_k(x)$ satisfies a Lipschitz condition and is bounded away from zero. The step size is a continuous function of x.*

Theorem 3.1 applies to an Adams code with step size and order selection algorithms that satisfy the VSVO Assumptions: Clearly the step sizes are uniformly $O(\tau^{1/(k+1)})$ on subintervals where order k is used. The assumptions about the step size imply the existence of constants for which the bounds (2.2) and (3.4) hold. Indeed, they imply that

$$\frac{h_{r+1}}{h_r} = 1 + O(h_r) \tag{4.2}$$

To see this, let λ_k be a Lipschitz constant for $\theta_k(x)$ and $\mu_k = \min \theta_k(x) > 0$. If both steps are taken at order k, equation (4.2) follows from

$$\left| \frac{h_{r+1}}{h_r} - 1 \right| = \left| \frac{\theta_k(x_{r+1}) - \theta_k(x_r)}{\theta_k(x_r)} \right| \leq \frac{\lambda_k}{\mu_k} \, |x_{r+1} - x_r| = \frac{\lambda_k}{\mu_k} \, h_r$$

Now suppose that h_r is taken at order k and h_{r+1} is taken at order $k-1$. At some intermediate point $x_r < \hat{x} \leq x_{r+1}$, the two orders are equally efficient, meaning that

$$\hat{h} = \theta_k(\hat{x}) \, \tau^{1/(k+1)} = \theta_{k-1}(\hat{x}) \, \tau^{1/k}$$

Regarding $\hat{h}$ as chosen at order k, we have

$$\left| \frac{\hat{h}}{h_r} - 1 \right| \leq \frac{\lambda_k}{\mu_k} \, |\hat{x} - x_r| \leq \frac{\lambda_k}{\mu_k} \, h_r$$

hence $\hat{h}/h_r = 1 + O(h_r)$. Now regarding $\hat{h}$ as chosen at order $k-1$, we have

$$\left| \frac{h_{r+1}}{\hat{h}} - 1 \right| \leq \frac{\lambda_{k-1}}{\mu_{k-1}} |x_{r+1} - \hat{x}| \leq \frac{\lambda_{k-1}}{\mu_{k-1}} h_r$$

hence $h_{r+1}/\hat{h} = 1 + O(h_r)$. Combining the two relations we find that equation (4.2) holds when the order is lowered. The argument is the same when the order is raised. We remark that the papers [7, 14] also make important use of the property (4.2), but they assume it directly rather than deriving it from more fundamental assumptions. With these preparations, we are now ready to prove

THEOREM 4.1. *Suppose that (2.1) is solved with tolerance τ by a code that satisfies the VSVO Assumptions. Suppose that order k is used on a subinterval that includes x_n. Then for $m = n, n+1, \ldots,$*

$$y_m = y(x_m) + e(x_m, \tau)\tau + O(H\tau)$$

Here, $H = \tau^{1/(k+1)}$ and $e(x,\tau)$ is the solution of

$$e' = J(x)e - \psi(x), \qquad e(x_n, \tau) = (y_n - y(x_n))/\tau \tag{4.3}$$

where

$$\psi(x) = \theta_k^{k+1}(x) \left[\gamma_{k+1}^* \, y^{(k+2)}(x) + A_{k+1}^* J(x) \gamma_k \, y^{(k+1)}(x) \right]$$

and the constant $A_{k+1}^* = \alpha_{k+1,0}^* + O(H)$.

We have already argued that with the VSVO Assumptions, the conditions of Theorem 3.1 are satisfied. It tells us that the error $\delta_m = y_m - y(x_m)$ is $O(\tau)$ for all m. The function $e(x,\tau)$ is continuous, so

$$\Delta_m = y_m - y(x_m) - e(x_m, \tau)\,\tau = \delta_m - e(x_m, \tau)\,\tau$$

is $O(\tau)$ for $n - k < m < n$. This is so regardless of the orders that were used, but starting with the step from x_n, we assume that the order is k. We aim to prove that Δ_m is $O(H\tau)$ for $m = n, n+1, \ldots$. We have defined $e(x_n, \tau)$ so that $\Delta_n = 0$. To show the result for subsequent steps, we begin by subtracting (3.1) from (2.3) to get

$$p_{m+1} - y(x_{m+1}) = \delta_m - lte_k + h_m \sum_{j=1}^{k} \alpha_{k,j} \left[f_{m+1-j} - f(x_{m+1-j}, y(x_{m+1-j})) \right]$$

From our assumption about the step sizes, each h_m is $O(H)$. Using this, the Lipschitz condition on f, and the order of the δ_{m+1-j}, we find that

$$p_{m+1} - y(x_{m+1}) = \delta_m - lte_k + O(H\tau) \tag{4.4}$$

A similar treatment of the corrector formula in which we expand terms in brackets rather than bounding them results in

$$\begin{aligned}
\delta_{m+1} &= \delta_m + h_m \sum_{j=1}^{k} \alpha_{k+1,j}^* \left[f_{m+1-j} - f(x_{m+1-j}, y(x_{m+1-j})) \right] \\
&\quad + h_m \, \alpha_{k+1,0}^* \left[f_{m+1}^p - f(x_{m+1}, y(x_{m+1})) \right] - lte_{k+1}^* \\
&= \delta_m + h_m \sum_{j=1}^{k} \alpha_{k+1,j}^* \, J(x_{m+1-j}) \, \delta_{m+1-j} \\
&\quad + h_m \, \alpha_{k+1,0}^* \, J(x_{m+1})(p_{m+1} - y(x_{m+1})) - lte_{k+1}^* + O(H\tau^2)
\end{aligned}$$

Substituting (4.4) into this last expression, we find that

$$\delta_{m+1} = \delta_m + h_m \sum_{j=1}^{k} \alpha^*_{k+1,j}\, J(x_{m+1-j})\, \delta_{m+1-j} \\ - h_m\, \alpha^*_{k+1,0}\, J(x_{m+1})\, lte_k - lte^*_{k+1} + O(H\tau) \tag{4.5}$$

The property (4.2) implies that to leading order, the step size is constant in the span of a formula. Using this observation in (3.5), we find that

$$lte_k = \gamma_k\, h_m^{k+1}\, y^{(k+1)}(x_m) + O(H^{k+2})$$

and similarly,

$$lte^*_{k+1} = \gamma^*_{k+1}\, h_m^{k+2}\, y^{(k+2)}(x_m) + O(H^{k+3})$$

For the same reason there is a constant $A^*_{k+1} = \alpha^*_{k+1,0} + O(H)$. Using these expansions and expanding $J(x_{m+1})$ about x_m in (4.5) results in

$$\delta_{m+1} = \delta_m + h_m \sum_{j=0}^{k} \alpha^*_{k+1,j}\, J(x_{m+1-j})\, \delta_{m+1-j} \\ - h_m^{k+2} \left[\gamma^*_{k+1}\, y^{(k+2)}(x_m) + \alpha^* J(x_m)\gamma_k\, y^{(k+1)}(x_m)\right] + O(H\tau) \tag{4.6}$$

Application of the Adams-Moulton formula of order $k+1$ to (4.3) yields

$$e(x_{m+1},\tau) = e(x_m,\tau) + h_m \sum_{j=0}^{k} \alpha^*_{k+1,j}\, J(x_{m+1-j})\, e(x_{m+1-j},\tau) \\ - h_m \sum_{j=0}^{k} \alpha^*_{k+1,j}\, \psi(x_{m+1-j}) + O(H^{k+2})$$

Simplifying the second sum by expanding the $\psi(x_{m+1-j})$ about x_m and using the fact that the coefficients of the formula sum to one results in

$$e(x_{m+1},\tau) = e(x_m,\tau) + h_m \sum_{j=0}^{k} \alpha^*_{k+1,j}\, J(x_{m+1-j})\, e(x_{m+1-j},\tau) - h_m\psi(x_m) + O(H^2)$$

Substituting the definition of $\psi(x_m)$ then leads to

$$e(x_{m+1},\tau) = e(x_m,\tau) + h_m \sum_{j=0}^{k} \alpha^*_{k+1,j}\, J(x_{m+1-j})\, e(x_{m+1-j},\tau) \\ - h_m\, \theta_k^{k+1}(x_m) \left[\gamma^*_{k+1}\, y^{(k+2)}(x_m) + A^*_{k+1} J(x_m)\gamma_k\, y^{(k+1)}(x_m)\right] + O(H^2)$$

Multiplying this expression by $\tau = H^{k+1}$ and subtracting it from (4.6) shows that

$$\Delta_{m+1} = \Delta_m + h_m \sum_{j=0}^{k} \alpha^*_{k+1,j}\, J(x_{m+1-j})\, \Delta_{m+1-j} + O(H\tau) \tag{4.7}$$

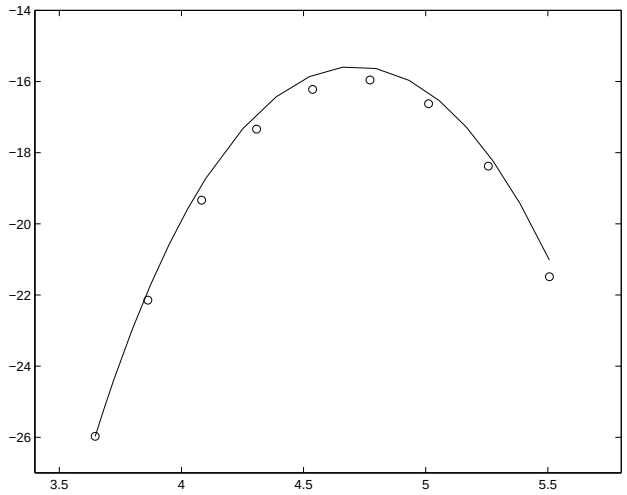

FIG. 4.3. *Scaled errors* $(y_m - y(x_m))/\tau$ *compared to* $e(x,\tau)$.

We have already argued that Δ_r is $O(\tau)$ for $n-k < r < n$ and further that $\Delta_n = 0$. It is then immediate from (4.7) that Δ_{n+1} is $O(H\tau)$. Repetition of the argument shows that Δ_m is $O(H\tau)$ for $m = n, n+1, \ldots$. We are interested only in what happens for a few steps past x_n, but if order k is used for as many as k successive steps, the fact that Δ_m is $O(H\tau)$ for these steps and a standard argument like that of Theorem 3.1 can be used to show that the same is true of all subsequent steps taken at order k.

For our experiments with the approximation of Theorem 4.1, we wrote a program that allows the step size and order to be specified at each step. We used it to integrate equation (2.4) with $\tau = 10^{-4}$. To illustrate the quality of the approximation immediately after changes of order and step size, we started the integration with $k = 1$ and increased the order at every step until reaching $k = 10$. We then decreased the order at every step until reaching $k = 6$. At each order k, we used the step size $\tau^{1/(k+1)}$. This means that we changed the order and step size at every step until reaching $\bar{x} \approx 3.6472$. The order was then changed to $k = 5$ and held at this value, but the step size was changed at every step according to $\theta_5(x) = 1 + 0.1(x - \bar{x})$. This resulted in step sizes that increased about 2% at each step, hence about 16% over the range shown in Fig. 5.2. We evaluated $e(x,\tau)$ for the figure by solving the initial value problem (4.3) with `ode113` and stringent error tolerances. Although the function ψ of (4.3) can be very complicated, the solution of (2.4) is relatively simple and its high order derivatives are obtained easily with Maple [15]. For these challenging circumstances, $e(x,\tau)$ agrees in Fig. 5.2 remarkably well with the scaled errors observed in the integration.

5. Error Estimation. Using our convergence result, it is not hard to justify standard estimates of local truncation error at orders lower than the current one. With local extrapolation, this is also true of the current order. However, even with local extrapolation, estimating the local truncation error of a formula of order higher than the current one is delicate because this error is smaller than the (global) error in the solution values used for the estimation. Indeed, if the error in the solution did not behave in a regular way, this would not be possible. The classical theory of error

estimation is based on an asymptotic expansion of the global error when the order is constant and the step sizes come from a step size selection function [16, 17, 18]. Again we must take an entirely different approach for VSVO Adams codes. Using the approximation developed in §4, we are able to justify a standard estimate of local truncation error at order one higher than the current order. Although it is natural to estimate and control lte_k, the codes estimate and control lte_k^*. The two procedures are closely related. For the sake of simplicity we justify the first and then explain the relationship.

In this section we work with approximations in the course of a specific integration. When we write, e.g., that a quantity is $O(\tau)$, this is generally to be understood as saying that the quantity is not greatly larger than τ in this specific integration rather than a statement about the behavior as $\tau \to 0$. We assume the hypotheses of Theorem 4.1 and again write $H = \tau^{1/(k+1)}$ when using order k.

We begin by using the convergence result of Theorem 3.1 and some intermediate results from the proof of Theorem 4.1 to justify estimation of the local truncation error of the predictor at the current and lower orders. Suppose that we have just taken a step of size h_m from x_m with the $(k, k+1)$ pair. From (4.4) we find that the error of the predicted solution is

$$p_{m+1}(k) - y(x_{m+1}) = \delta_m - lte_k + O(H^{k+2})$$

Because of convergence, each of the δ_{m+1-j} in (4.5) is $O(\tau)$, hence $O(H^{k+1})$. Because of local extrapolation, lte_{k+1}^* is $O(H^{k+2})$. With these observations, we see from (4.5) that the error of the corrected solution is

$$y_{m+1} - y(x_{m+1}) = \delta_{m+1} = \delta_m + O(H^{k+2})$$

From these two equations it is immediate that

$$lte_k = y_{m+1} - p_{m+1}(k) + O(H^{k+2})$$

Similar arguments justify estimates at lower orders, namely

$$lte_{k-1} \approx y_{m+1} - p_{m+1}(k-1)$$
$$lte_{k-2} \approx y_{m+1} - p_{m+1}(k-2)$$

In practice it is convenient to use identities that allow the estimates to be computed directly from stored function values. For instance, when the step size is a constant h,

$$lte_k \approx y_{m+1} - p_{m+1}(k) = \gamma_k h \nabla^k f_{m+1}^p$$

Because of local extrapolation, f_{m+1} has a higher order of accuracy than f_{m+1}^p. It is not used when estimating the local truncation errors at the current and lower orders because it is not necessary and if the step is a failure, the code does not form it, thereby halving the cost of a failed step.

The VSVO Adams codes consider increasing the order only when the step size and order are constant in the span of the formula. Gear [3, § 9.8] presents a number of reasons for this. An increase of order is considered only after the step has been accepted and f_{m+1} formed. When estimating the local truncation error at order $k+1$, we must use f_{m+1} because f_{m+1}^p is not sufficiently accurate. When the step

size is a constant h in the span of the formula and we are working at order k, our approximation to the behavior of the error tells us that

$$\begin{aligned}\nabla^{k+1} f_{m+1} &\approx \nabla^{k+1} f(x_{m+1}, y(x_{m+1}) + H^{k+1} e(x_{m+1}, \tau) + O(H^{k+2}))\\ &\approx \nabla^{k+1} f(x_{m+1}, y(x_{m+1})) + H^{k+1} \nabla^{k+1} J(x_{m+1}) e(x_{m+1}, \tau) + O(H^{k+2})\end{aligned}$$

A basic property of backward differences [5, Chap. 2] and the ODE (2.1) imply that

$$\nabla^{k+1} f(x_{m+1}, y(x_{m+1})) = \nabla^{k+1} y'(x_{m+1}) = h^{k+1} y^{(k+2)}(x_m) + O(h^{k+2})$$

Differentiability of $e(x, \tau)$ implies that $\nabla^{k+1} J(x_{m+1}) e(x_{m+1}, \tau)$ is at least $O(h)$. Recalling that h is $O(H)$ by the VSVO Assumptions, these observations imply that

$$\begin{aligned}lte_{k+1} &= \gamma_{k+1} h^{k+2} y^{(k+2)}(x_m) + O(h^{k+3})\\ &\approx \gamma_{k+1} h \nabla^{k+1} f_{m+1} + O(H^{k+3})\end{aligned}$$

This justifies the usual estimate of the local truncation error at order $k+1$.

We illustrate error estimation in VSVO Adams codes with a numerical experiment. The results of Table 5.1 were computed just as in the experiment illustrating Theorem 4.1 except that after changing to $k = 5$ at $\bar{x} \approx 3.6472$, both the order and the step size h were held constant. In the table we display the ratio of the estimated and true local truncation errors. The estimates are satisfactory, even in the first step from $\bar{x}$ involving values that were all computed with different orders and step sizes. The error is estimated at order $k+1$ only when the formula is applied to values that were all computed at order k and step size h. In our experience, the estimates of the table are rather good, so the incorrect sign and small ratio of one of the estimates at order $k-1$ is a useful reminder that we need robust step size and order selection algorithms to deal with estimates that can be quite poor. (Near a change of sign of the local truncation error, it is common that an estimated error has the wrong sign. Generally this does not have a noticeable effect because it is smaller than the estimated errors in other components, but it is exposed when solving scalar ODEs like (2.4).)

$k+1$						1.02	0.94	0.94
k	1.45	1.19	1.05	0.88	0.99	0.92	0.78	0.95
$k-1$	1.02	0.95	0.91	1.00	1.03	-0.23	0.88	0.87
$k-2$	0.98	1.04	0.99	1.00	1.00	1.00	1.00	0.99

TABLE 5.1
Ratios est/lte at steps after change to constant order and step size.

We now consider briefly the local truncation error actually estimated and controlled in the codes. When taking a step with the $(k, k+1)$ pair, the codes also take the step with the PECE pair consisting of the Adams-Bashforth predictor of order k and the Adams-Moulton corrector of order k. We have been working with the local truncation errors of the individual formulas, lte_k and lte^*_{k+1}. The local truncation error of the $(k, k+1)$ pair is obtained by first eliminating p_{n+1} in the formula for y_{n+1}. A little manipulation then shows that the local truncation error of the pair is

$$lte^*_{k+1} + h_n \alpha^*_{k+1,0} J(x_n)\, lte_k + O(H^{k+3})$$

The local truncation error of the (k, k) pair is qualitatively different. The same argument shows that it is $lte^*_k + O(H^{k+2})$. That is, to leading order the local truncation

error of the (k,k) pair is equal to that of the Adams-Moulton formula of order k. This is the local truncation error estimated and controlled by the codes. The estimators and their justifications are much like those of the predictor alone. Alternatively, from (3.5) it is immediate that

$$lte_k = \frac{y^{(k+1)}(x_n)}{k!} \int_{x_n}^{x_{n+1}} \prod_{i=1}^{k} (t - x_{n+1-i})\, dt + O(H^{k+2})$$

and from (3.6), that

$$lte_k^* = \frac{y^{(k+1)}(x_n)}{k!} \int_{x_n}^{x_{n+1}} \prod_{i=0}^{k-1} (t - x_{n+1-i})\, dt + O(H^{k+2})$$

These expressions show that to leading order, lte_k^* is a fraction of lte_k, a fraction that depends only on the mesh spacing. For instance, when the step size is constant in the span of the formulas, this fraction is γ_k^*/γ_k.

6. Conclusions. In this paper we have come to an understanding of some fundamental issues in the theory of the DVDQ line of VSVO Adams codes. With realistic assumptions, we proved that as a tolerance $\tau \to 0$, the error is uniformly $O(\tau)$. With additional realistic assumptions about order and step size selection algorithms, we developed an approximation to the behavior of the error for a few steps taken at constant order in the course of an integration with a specific tolerance. We used the convergence result and the approximation to justify standard error estimators, including estimation of the local truncation error of a formula of order one higher than the current formula.

7. Acknowledgment. A visit at the University of Western Ontario as the Ontario Research Chair in Computer Algebra made this investigation possible.

REFERENCES

[1] T.E. Hull, W.H. Enright, B.M. Fellen, and A.E. Sedgwick, Comparing numerical methods for ordinary differential equations, *SIAM J. Numer. Anal.* **9** 603–637 (1972).

[2] F.T. Krogh, Changing step size in the integration of differential equations using modified divided differences, in *Numerical Solution of Ordinary Differential Equations*, (Edited by D.G. Bettis), pp. 22–71, Springer, New York (1974).

[3] C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ (1971).

[4] J.R. Dormand, *Numerical Methods for Differential Equations*, CRC Press, Boca Raton, FL (1996).

[5] L.F. Shampine and M.K. Gordon, *Computer Solution of Ordinary Differential Equations*, W.H. Freeman, San Francisco (1975).

[6] L.F. Shampine and M.W. Reichelt, The MATLAB ODE suite, *SIAM J. Sci. Comput.* **18** 1–23 (1997).

[7] C.W. Gear and D.S. Watanabe, Stability and convergence of variable order multistep methods, *SIAM J. Numer. Anal.* **11** 1044–1058 (1974).

[8] E. Hairer, S.P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I*, Springer-Verlag, Berlin (1987).

[9] J.D. Lambert, *Numerical Methods for Ordinary Differential Systems*, J. Wiley & Sons, New York (1991).

[10] L.F. Shampine, *Numerical Solution of Ordinary Differential Equations*, Chapman & Hall, New York (1994).

[11] F.T. Krogh, On testing a subroutine for the numerical integration of ordinary differential equations, *JACM* **20** 545–562 (1973).

[12] L.F. Shampine, The step sizes used by one-step codes for ODEs, *Appl. Numer. Math.* **1** 95–106 (1985).
[13] MATLAB 5, The MathWorks, Inc., 3 Apple Hill Dr., Natick, MA 01760, 1998.
[14] C.W. Gear and K.W. Tu, The effect of variable mesh size on the stability of multistep methods, *SIAM J. Numer. Anal.* **11** 1025–1043 (1974).
[15] Maple V Release 5, Waterloo Maple Inc., 450 Phillip St., Waterloo, ON N2L 5J2, 1998.
[16] P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations*, J. Wiley & Sons, New York (1962).
[17] P. Henrici, *Error Propagation for Difference Methods, Krieger*, New York (1977).
[18] H.J. Stetter, *Analysis of Discretization Methods for Ordinary Differential Equations*, Springer, New York (1973).