Numer. Math. 45, 173–182 (1984)

Numerische Mathematik
© Springer-Verlag 1984

Convergence of the Lambert-McLeod Trajectory Solver and of the Celf Method

J.M. Sanz-Serna

Departamento de Ecuaciones Funcionales, Facultad de Ciencias,
Universidad de Valladolid, Valladolid, Spain

Summary. A trajectory problem is an initial value problem $d\mathbf{y}/dt=\mathbf{f}(\mathbf{y})$, $\mathbf{y}(0)=\boldsymbol{\eta}$ where the interest lies in obtaining the curve traced by the solution (the trajectory), rather than in finding the actual correspondance between values of the parameter t and points on that curve. We prove the convergence of the Lambert-McLeod scheme for the numerical integration of trajectory problems. We also study the CELF method, an explicit procedure for the integration in time of semidiscretizations of PDEs which has some useful conservation properties. The proofs rely on the concept of restricted stability introduced by Stetter. In order to show the convergence of the methods, an idea of Strang is also employed, whereby the numerical solution is compared with a suitable perturbation of the theoretical solution, rather than with the theoretical solution itself.

Subject Classifications Primary 65L05, Secondary 65M10, CR: G1.7.

1. Introduction

We consider the initial value problem

$$\frac{d\mathbf{y}}{dt}=\mathbf{f}(\mathbf{y}), \qquad 0\leqq t\leqq T, \qquad \mathbf{y}(0)=\boldsymbol{\eta}, \tag{1.1}$$

where $\mathbf{y}$ takes values in $\mathbb{R}^d$. Following [8] we say that (1.1) is a *trajectory problem* if the interest lies in obtaining the curve traced by the solution $\mathbf{y}(.)$ (the trajectory), rather than in finding the actual correspondance between values of the parameter t and points on that curve. Trajectory problems arise in the computation of trajectories in mechanical problems, in the plotting of phase-planes of second order autonomous differential equations, etc... [8].

The Lambert-McLeod explicit method [6]

$$\mathbf{y}_0, \mathbf{y}_1 \text{ given}, \tag{1.2}$$

$$\mathbf{y}_{n+2}-\mathbf{y}_n=2[(\mathbf{y}_{n+1}-\mathbf{y}_n)^T\mathbf{F}_{n+1}]\,\mathbf{F}_{n+1}, \qquad n=0,1,\ldots$$

$$\mathbf{F}_{n+1}=(1/\|\mathbf{f}(\mathbf{y}_{n+1})\|)\,\mathbf{f}(\mathbf{y}_{n+1}),$$

was specifically introduced for the numerical integration of trajectory problems and possesses some remarkable properties, which will now be outlined. (In this paper $\|-\|$ denotes the Euclidean norm in $\mathbb{R}^d$.)

i) *Equispacing:* $\|\mathbf{y}_{n+1}-\mathbf{y}_n\|=\|\mathbf{y}_1-\mathbf{y}_0\|$, $n=0,1,\ldots$

This was first noted by Laurie [7]. The proof is elementary. Observe that no step-length appears explicitly in the scheme (1.2). The equispacing property entails that the method carries implicitly a 'step-length', namely the Euclidean distance between any two consecutively computed points. The existence of this 'hidden' step-length enables us to derive a convergence result for the Lambert-McLeod method: Under suitable hypotheses, we prove in Sect. 3 that if we denote by h the length $\|\mathbf{y}_1-\mathbf{y}_0\|$ then $\|\mathbf{Y}_n-\mathbf{y}_n\|=O(h^2)$, where $\mathbf{Y}_n$ is the point on the trajectory whose distance from $\boldsymbol{\eta}=\mathbf{y}(0)$ *along the curve* $\mathbf{y}(.)$ equals nh.

ii) *Circular exactness*: Lambert and McLeod [6] proved that if the trajectory of (1.1) is a circle in $\mathbb{R}^d$ and $\mathbf{y}_0, \mathbf{y}_1$ lie on it, then all the computed points $\mathbf{y}_{n+2}$, $n=0,1,\ldots$ also lie on the trajectory. More generally it is shown in [6], [8] how, at least in principle, one can construct schemes which are exact whenever the trajectory belongs to any given family of curves.

In i) and ii) we have assumed that no round-off error is present. We refer to [8] for additional material on trajectory problems.

The formula (1.2) can also be applied to the integration of *conventional* initial value problems, i.e. problems where the correspondance $t \to \mathbf{y}(t)$ is of interest. In order to recover this correspondance, which is not given by (1.2), an additional formula must be employed to generate a sequence t_n, $n=0,1,\ldots$ in such a way that $\mathbf{y}_n$ approximates $\mathbf{y}(t_n)$. An instance is provided by the method

$$\mathbf{y}_{n+2}-\mathbf{y}_n=2\tau_{n+1}\mathbf{f}(\mathbf{y}_{n+1}), \tag{1.3a}$$

$$\tau_{n+1}=(\mathbf{y}_{n+1}-\mathbf{y}_n)^T\mathbf{f}(\mathbf{y}_{n+1})/\|\mathbf{f}(\mathbf{y}_{n+1})\|^2, \tag{1.3b}$$

$$t_{n+2}-t_n=2\tau_{n+1}, \tag{1.3c}$$

which has been suggested and tested numerically in [10], [11]. Clearly (1.3a), (1.3b) are equivalent to the Lambert-McLeod formula. The scheme (1.3) can be viewed as a *variable step leap-frog* (mid-point) method, where the choice of step-length τ_{n+1} is determined by the previous computed points $\mathbf{y}_{n+1}, \mathbf{y}_n$ according to (1.3b). Hence the name CELF (circularly exact leap-frog) method employed to refer to (1.3). Note that in (1.3b) $\|\mathbf{f}\|^{-2}\,\mathbf{f}$ is the so-called Samelson inverse of $\mathbf{f}$ [15]. Samelson inverses have been used by Wambeck [15] in his rational Runge-Kutta methods. (More recent references to these methods can be seen in [1].) Of course the Lambert-McLeod formula can be regarded as a *rational two-step method.*

A useful property of the CELF method will now be described. Assume that $\mathbf{f}$ satisfies

$$\mathbf{v}^T\mathbf{f}(\mathbf{v})=0, \qquad \forall \mathbf{v}\in\mathbb{R}^d. \tag{1.4}$$

leading to the first integral $\|\mathbf{y}(t)\|=$constant. The property (1.4) is often found [10] in situations where the system (1.1) is a semidiscretization of an evolutionary PDE which conserves the L^2-norm. (Examples include Galerkin and finite-difference semidiscretizations of many nonlinear wave systems.) When (1.4) holds, the *explicit* CELF scheme possesses the conservation property $\|\mathbf{y}_{n+2}\| = \|\mathbf{y}_n\|$, $n=0,1,\ldots$ a potentially useful feature ensuring that the computed points will not blow up, as it is sometimes the case if conventional explicit methods (such as the standard leap-frog) are employed [9], [10]. An account of the benefits to be gained by using schemes with conservation properties is given by Morton in [9]. Recently, Dekker an Verwer [2] have pursued the idea behind the CELF method and constructed rational explicit Runge-Kutta methods with conservation properties. We also mention that the CELF method can be modified [11] in order to cater for conservation laws more general than $\|\mathbf{y}(t)\| = \text{constant}$.

We prove in Sect. 4 that the CELF method is convergent of the second order. (The proof does not require that (1.4) holds.) It should be pointed out that our analysis is complicated by the fact that the Lambert-McLeod formula is *not stable* (in the sense of [13, Chapter 1]). In order to see this, assume that (1.1) is one-dimensional ($d=1$). Then (1.2) reads simply $y_{n+2} = 2y_{n+1} - y_n$, an unstable recursion. Note however that in the *absence of round-off*, the points y_n generated by the recursion are identical to the 'exact' points $Y_n = nh$, in agreement with our convergence claim.

It turns out that the appropriate stability concept associated with (1.2) is similar to that of *restricted stability* introduced by Stetter [12] (and independently by Kuo Pen-Yu, see the references in [3]) (cf. also [4], [13, p. 79]). We recall that a restricted stable scheme is convergent provided that its order of consistency is *higher* than the so-called *stability index*. (The convergence is still preserved in the absence of round-off, provided that the latter is suitably small.) Unfortunately the method (1.2) possesses an order of consistency *lower* than its stability index, and in order to prove convergence one must resort to the construction of an asymptotic expansion of the global error, thus following an idea of Strang [4]. This point is explained in detail later.

2. A Stability Result

We begin by presenting the hypotheses on the IVP (1.1) that are required for our analysis to hold. We assume

(H1) The function $\mathbf{f}$ takes values in $\mathbb{R}^d$ and is defined in an open set $\Omega_0 \subset \mathbb{R}^d$.

(H2) The problem (1.1) has a unique solution $\mathbf{y}(.)$. Furthermore this solution belongs to $C^4([0,T])$.

(H3) An open set Ω and a positive constant μ can be found so that: i) $\{\mathbf{y}(t)\colon t \in [0,T]\} \subset \Omega \subset \Omega_0$. ii) $\|\mathbf{f}(\mathbf{v})\| > \mu$ if $\mathbf{v} \in \Omega$. iii) $\mathbf{f}$ has continuous, bounded first and second derivatives in Ω.

Note that condition ii) in (H3) is essentially equivalent to the requirement that constant (equilibrium) solutions of the system in (1.1) should not be considered. Observe in this connection the denominator in (1.2).

In the Lambert-McLeod method, the function $\mathbf{f}$ appears only in the combination $\|\mathbf{f}\|^{-1}\mathbf{f}$. This suggests the introduction of the function $\mathbf{F}: \Omega \to \mathbb{R}^d$ defined by $\mathbf{F} = \|\mathbf{f}\|^{-1}\mathbf{f}$ and of the initial value problem

$$\frac{d\mathbf{Y}}{ds} = \mathbf{F}(\mathbf{Y}), \qquad 0 \leqq s \leqq \gamma, \qquad \mathbf{Y}(0) = \boldsymbol{\eta}, \tag{2.1}$$

where γ is the length of the curve in $\mathbb{R}^d$ defined by $\mathbf{y}(t)$, $0 \leqq t \leqq T$. Note that $\|\mathbf{F}\| \equiv 1$ and therefore solutions of (2.1) are parameterized by arclength. The following properties are consequences of (H1)-(H3).

(P1) $\mathbf{F}$ *has continuous, bounded first and second derivatives in* Ω. This follows easily from (H3) ii)-iii). In particular $\mathbf{F}$ is Lipschitz continuous in Ω.

(P2) *The problem* (2.1) *has a unique solution* $\mathbf{Y}(.)$; *furthermore* $\mathbf{Y}(.) \in C^4([0, \gamma])$. In fact if $\mathbf{y}(.)$ is the unique solution to (1.1), then the function $s(t)$ defined by $s(0) = 0$, $ds/dt = \|d\mathbf{y}(t)/dt\|$ is a C^4-diffeomorphism (i.e. a C^4 function with C^4 inverse) of $[0, T]$ onto $[0, \gamma]$. Upon denoting by $t = t(s)$ the inverse diffeomorphism and introducing the composition $\mathbf{Y}(s) = \mathbf{y}(t(s))$ we can write

$$\frac{d\mathbf{Y}}{ds} = \frac{dt}{ds}\frac{d\mathbf{y}}{dt} = \frac{1}{\|\mathbf{f}(\mathbf{y}(t(s)))\|}\mathbf{f}(\mathbf{y}(t(s))) = \mathbf{F}(\mathbf{Y}(s)),$$

so that $\mathbf{Y}(s)$, $0 \leqq s \leqq \gamma$ is a solution of (2.1). We emphasize that $\mathbf{Y}(s)$, $0 \leqq s \leqq \gamma$ is the parameterization in terms of arclength of the curve described by $\mathbf{y}(t)$, $0 \leqq t \leqq T$ (i.e. the sought trajectory).

(P3) The identity

$$\|\mathbf{F}(\mathbf{v})\| \equiv 1, \qquad \mathbf{v} \in \Omega \tag{2.2}$$

implies, after differentiation,

$$(\mathbf{F}(\mathbf{v}))^T J(\mathbf{v}) \equiv \mathbf{0}^T, \qquad \mathbf{v} \in \Omega \tag{2.3}$$

$$(\dot{\mathbf{Y}}(s))^T J(\mathbf{Y}(s))\,\dot{\mathbf{Y}}(s) = (\dot{\mathbf{Y}}(s))^T \ddot{\mathbf{Y}}(s) \equiv 0, \qquad 0 \leqq s \leqq \gamma, \tag{2.4}$$

In this paper a dot represents d/ds and $J(\mathbf{v})$ denotes the Jacobian matrix of $\mathbf{F}$ evaluated at $\mathbf{v}$. In order to shorten the notation, we set, if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are in Ω

$$\boldsymbol{\Phi}(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} - \mathbf{c} - 2[(\mathbf{b} - \mathbf{c})^T \mathbf{F}(\mathbf{b})]\,\mathbf{F}(\mathbf{b}),$$

so that the formula (1.2) now reads $\Phi(\mathbf{y}_{n+2}, \mathbf{y}_{n+1}, \mathbf{y}_n) = \mathbf{0}$.

We are now in a position to investigate the stability properties of the scheme (1.2). Let h be a parameter taking values in $(0, h_0)$ and assume that for each value of $h, \mathbf{u}_n^h, \mathbf{v}_n^h$, $n = 0, 1, \ldots, [\gamma/h]$ are sequences in Ω. Set, for $n = 0, 1, \ldots, [\gamma/h] - 2$

$$h\rho_{n+2}^h = \boldsymbol{\Phi}(\mathbf{u}_{n+2}^h, \mathbf{u}_{n+1}^h, \mathbf{u}_n^h), \tag{2.5a}$$

$$h\sigma_{n+2}^h = \boldsymbol{\Phi}(\mathbf{v}_{n+2}^h, \mathbf{v}_{n+1}^h, \mathbf{v}_n^h). \tag{2.5b}$$

Our aim is to bound $\mathbf{u}_n^h - \mathbf{v}_n^h$, $n = 0, 1, \ldots, [\gamma/h]$ in terms of the 'perturbations' $\boldsymbol{\rho}_n^h, \boldsymbol{\sigma}_n^h, \mathbf{u}_0^h - \mathbf{v}_0^h, \mathbf{u}_1^h - \mathbf{v}_1^h$.

Theorem 1. *Let* (H1)–(H3) *hold. Assume that for* $h \in (0, h_0)$ *the sequences* $\mathbf{u}_n^h, \mathbf{v}_n^h$, $n = 0, 1, \ldots, [\gamma/h]$ *are such that*

i) $\mathbf{u}_n^h, \mathbf{v}_n^h \in \Omega$, $0 < h < h_0$, $n = 0, 1, \ldots, [\gamma/h]$.

ii) If $\boldsymbol{\rho}_{n+2}^h, \boldsymbol{\sigma}_{n+2}^h$ *are defined in* (2.5), *then there exist positive constants* C_1, C_2, *independent of h, such that, for* $n = 0, 1, \ldots, [\gamma/h] - 2$

$$\|\boldsymbol{\rho}_{n+2}^h\| \leqq C_1 h^3, \qquad \|\boldsymbol{\sigma}_{n+2}^h\| \leqq C_2 h^3.$$

iii) There exists a positive constant C_3, *independent of h* such that

$$\|\mathbf{u}_0^h - \mathbf{v}_0^h\| + \|\mathbf{u}_1^h - \mathbf{v}_1^h\|\,\| \leqq C_3 h^3$$

iv) There exist positive constants C_4, C_5, *independent of h such that, for* $0 < h < h_0$, $n = 0, 1, \ldots, [\gamma/h] - 1$

$$|\|\mathbf{u}_{n+1}^h - \mathbf{u}_n^h\| - h| \leqq C_4 h^4,$$
$$|\|\mathbf{v}_{n+1}^h - \mathbf{v}_n^h\| - h| \leqq C_5 h^4.$$

v) There exists a positive constant C_6, *independent of h, such that, for* $0 < h < h_0$, $n = 0, 1, \ldots, [\gamma/h] - 1$

$$\|h^{-1}(\mathbf{u}_{n+1}^h - \mathbf{u}_n^h) - \mathbf{F}(\mathbf{u}_{n+1}^h)\| \leqq C_6 h.$$

Then there exist positive constants C_7, C_8 *depending only on* C_i, $i = 1, \ldots, 6$, *on* γ *and on the Lipschitz constant L of* $\mathbf{F}$, *such that for* $0 < h < C_7$, $n = 0, 1, \ldots, [\gamma/h]$

$$\|\mathbf{u}_n^h - \mathbf{v}_n^h\| \leqq C_8 h^3. \tag{2.6}$$

Remarks. The hypothesis i) ensures that $\mathbf{F}(\mathbf{u}_n^h), \mathbf{F}(\mathbf{v}_n^h)$ are defined. Hypotheses ii) and iii) are basically equivalent to the requirement that the perturbations $\boldsymbol{\rho}_n^h, \boldsymbol{\sigma}_n^h$, $\mathbf{u}_0^h - \mathbf{v}_0^h$, $\mathbf{u}_1^h - \mathbf{v}_1^h$ should be small for the stability analysis to hold. In fact, in the framework of [13, Chapt. 1], they impose an $O(h^3)$ *stability threshold.* In this sense the method is 3-*restricted stable* at $\mathbf{u}_n^h$, [12]. The hypothesis iv) is related to the equispacing property considered in the introduction. It basically means that $\mathbf{u}_n^h, \mathbf{v}_n^h$ refer to the same 'steplength' h. Finally note that v) is satisfied if $\mathbf{u}_n^h = \mathbf{Y}(nh)$ (the 'theoretical solution').

Proof. Set $\mathbf{e}_n = \mathbf{u}_n - \mathbf{v}_n$, $n = 0, 1, \ldots, [\gamma/h]$. (The superscript h is dropped for simplicity.) Upon substracting in (2.5) and rearranging, we can write

$$\begin{aligned}\mathbf{e}_{n+2} = {} & \mathbf{e}_n + h\boldsymbol{\rho}_{n+2} - h\boldsymbol{\sigma}_{n+2} \\ & + 2((\mathbf{u}_{n+1} - \mathbf{u}_n) - (\mathbf{v}_{n+1} - \mathbf{v}_n))^T \mathbf{F}(\mathbf{u}_{n+1})\,\mathbf{F}(\mathbf{u}_{n+1}) \\ & + 2(\mathbf{v}_{n+1} - \mathbf{v}_n)^T (\mathbf{F}(\mathbf{u}_{n+1}) - \mathbf{F}(\mathbf{v}_{n+1}))\,\mathbf{F}(\mathbf{u}_{n+1}) \\ & + 2(\mathbf{v}_{n+1} - \mathbf{v}_n)^T \mathbf{F}(\mathbf{v}_{n+1})(\mathbf{F}(\mathbf{u}_{n+1}) - \mathbf{F}(\mathbf{v}_{n+1})),\end{aligned}$$

whence, according to (2.2), hypotheses ii) and iv) and the Schwartz inequality

$$\begin{aligned}\|\mathbf{e}_{n+2}\| \leqq {} & \|\mathbf{e}_n\| + C_1 h^4 + C_2 h^4 \\ & + 2|((\mathbf{u}_{n+1} - \mathbf{u}_n) - (\mathbf{v}_{n+1} - \mathbf{v}_n))^T \mathbf{F}(\mathbf{u}_{n+1})| \\ & + 4(h + C_5 h^4) L \|\mathbf{e}_{n+1}\|.\end{aligned} \tag{2.7}$$

We now bound the inner product in (2.7) as follows

$$\begin{aligned}((\mathbf{u}_{n+1}-\mathbf{u}_n)-(\mathbf{v}_{n+1}-\mathbf{v}_n))^T\mathbf{F}(\mathbf{u}_{n+1})&=(\mathbf{e}_{n+1}-\mathbf{e}_n)^T\mathbf{F}(\mathbf{u}_{n+1})\\ &=h^{-1}(\mathbf{e}_{n+1}-\mathbf{e}_n)^T(\mathbf{u}_{n+1}-\mathbf{u}_n)\\ &\quad-(\mathbf{e}_{n+1}-\mathbf{e}_n)^T(h^{-1}(\mathbf{u}_{n+1}-\mathbf{u}_n)-\mathbf{F}(\mathbf{u}_{n+1})).\end{aligned}$$

The second inner product in the righ-hand side is clearly less than $C_6h(\|\mathbf{e}_{n+1}\| + \|\mathbf{e}_n\|)$. Turning now to the first inner product we square both sides of the identity

$$\mathbf{v}_{n+1}-\mathbf{v}_n=(\mathbf{u}_{n+1}-\mathbf{u}_n)-(\mathbf{e}_{n+1}-\mathbf{e}_n)$$

to obtain

$$\begin{aligned}\|\mathbf{v}_{n+1}-\mathbf{v}_n\|^2=\|\mathbf{u}_{n+1}-\mathbf{u}_n\|^2&-2(\mathbf{u}_{n+1}-\mathbf{u}_n)^T(\mathbf{e}_{n+1}-\mathbf{e}_n)\\ &+\|\mathbf{e}_{n+1}-\mathbf{e}_n\|^2.\end{aligned}$$

Therefore

$$\begin{aligned}2(\mathbf{u}_{n+1}-\mathbf{u}_n)^T(\mathbf{e}_{n+1}-\mathbf{e}_n)&=\|\mathbf{e}_{n+1}-\mathbf{e}_n\|^2\\ &+(\|\mathbf{u}_{n+1}-\mathbf{u}_n\|-\|\mathbf{v}_{n+1}-\mathbf{v}_n\|)(\|\mathbf{u}_{n+1}-\mathbf{u}_n\|+\|\mathbf{v}_{n+1}-\mathbf{v}_n\|)\end{aligned}$$

and

$$\begin{aligned}2|(\mathbf{u}_{n+1}-\mathbf{u}_n)^T(\mathbf{e}_{n+1}-\mathbf{e}_n)|&\leqq(\|\mathbf{e}_{n+1}\|+\|\mathbf{e}_n\|)^2\\ &+(C_4+C_5)h^4(2h+C_4h^4+C_5h^4).\end{aligned}$$

Upon using these bounds in (2.7) the inequality

$$E_{n+1}\leqq E_n+ChE_n+Ch^{-1}E_n^2+Ch^4$$

is obtained, where $E_n=\|\mathbf{e}_{n+1}\|+\|\mathbf{e}_n\|$ and C is a positive constant which depends on C_i, $i=1,2,4,5,6$ and on L. The result is now a consequence of the generalized discrete Gronwall Lemma [5, lemma].

3. Convergence of the Lambert-McLeod Method

In order to prove the convergence of the Lambert-McLeod method one would like to apply Theorem 1 to the sequences $\mathbf{u}_n^h=\mathbf{Y}(nh)$ (the 'theoretical solution'), $\mathbf{v}_n^h=\mathbf{y}_n^h$ (the computed solution), with $h:=\|\mathbf{y}_1-\mathbf{y}_0\|$ (the n-independent Euclidean distance between consecutive computed points, cf. § 1, i)). Unfortunately, $\mathbf{u}_n^h=\mathbf{Y}(nh)$ does not satisfy the hypotheses in Theorem 1. On the one hand, a Taylor expansion reveals that the *truncation error*

$$\rho_{n+2}^h:=h^{-1}\boldsymbol{\Phi}(\mathbf{Y}((n+2)h),\ \mathbf{Y}((n+1)h),\ \mathbf{Y}(nh))$$

is only $O(h^2)$ and therefore hypothesis ii), requiring $O(h^3)$ is not fulfilled (i.e. the order of consistency is smaller than the index of generalized stability [3]). Also

$$\|\mathbf{Y}((n+1)h)-\mathbf{Y}(nh)\|=h+O(h^3)$$

so that iv) is vulnerated.

Strang [14] observed that difficulties of this sort could be circumvented by comparing the numerical solution not with the theoretical solution, but rather with a sequence which is close to the theoretical solution and satifies the discrete equations up to perturbation terms of order $O(h^j)$, with j large enough for those terms to come within the stability threshold. This technique yields an asymptotic expansion of the *global error* of the method.

We shall apply Theorem 1 to the sequence

$$\mathbf{u}_n^h := \mathbf{Y}(nh) + h^2 \mathbf{W}(nh), \tag{3.1}$$

where $\mathbf{W}(.) \in C^2([0,\gamma])$ must be chosen in such a way that $\mathbf{u}_n^h$ verifies the hypotheses of the theorem.

Lemma. *a) Let* $\mathbf{W}(.)$ *be an* $\mathbb{R}^d$*-valued function, twice continuously differentiable in* $[0,\gamma]$. *Define* $\mathbf{u}_n^h$ *according to* (3.1), $n=0,1,\ldots,[\gamma/h]$. *Then, there exists* $h_0>0$, *such that for* $0<h<h_0$, $\mathbf{u}_n^h \in \Omega$ *and v) in Theorem* 1 *holds.*

b) With the notations of a), $\mathbf{u}_n^h$, $0<h<h_0$, $n=0,1,\ldots,[\gamma/h]$, *satifies the requirement in ii) Theorem* 1, *if and only if* $\mathbf{W}(.)$ *satisfies*

$$\dot{\mathbf{W}} - (\dot{\mathbf{W}}^T \dot{\mathbf{Y}})\,\dot{\mathbf{Y}} = J(\mathbf{Y})\,\mathbf{W} - (1/6)\,\dddot{\mathbf{Y}} + (1/6)(\dddot{\mathbf{Y}}^T \dot{\mathbf{Y}})\,\dot{\mathbf{Y}}. \tag{3.2}$$

c) With the notations as above, the requirement in iv) Theorem 1 *is satisfied if and only if*

$$\dot{\mathbf{W}}^T \dot{\mathbf{Y}} = -(1/8)\,\|\ddot{\mathbf{Y}}\|^2 - (1/6)\,\dot{\mathbf{Y}}^T \dddot{\mathbf{Y}}. \tag{3.3}$$

d) The initial value problem

$$\begin{gathered} \mathbf{W}(0) = \mathbf{0}, \\ \dot{\mathbf{W}} = J(\mathbf{Y})\mathbf{W} - (1/6)\,\dddot{\mathbf{Y}} - (1/8)\,\|\ddot{\mathbf{Y}}\|^2\,\dot{\mathbf{Y}}, \qquad 0 \leqq s \leqq \gamma, \end{gathered} \tag{3.4}$$

has a unique solution. This solution is C^2 *and satisifies* (3.2) *and* (3.3).

Proof. a) is trivial and b), c) are obtained upon Taylor expanding $\boldsymbol{\Phi}(\mathbf{u}_{n+2}^h, \mathbf{u}_{n+1}^h, \mathbf{u}_n^h)$, $\|\mathbf{u}_{n+1}^h - \mathbf{u}_n^h\|^2$, respectively. (Properties (P1)-(P3) in the previous section are essential when carrying out the expansions.) To prove d), we note that (3.4) is a linear IVP whose matrix $J(\mathbf{Y}(s))$ and forcing term are continuously differentiable, after (P1), (P2). Therefore (3.4) has a unique solution with C^2 continuity. In order to see that (3.4) implies (3.3), it is enough to take the inner product of (3.4) with $\dot{\mathbf{Y}}$ and use (2.4). Finally, substraction from (3.4) of the product of (3.3) with $\dot{\mathbf{Y}}$ yields the equality (3.2).

We are now ready to prove the convergence of (1.2). In order to cater for the presence of *round-off errors*, we assume that the computed solution is obtained as follows

$$\begin{gathered} \mathbf{y}_0^h, \mathbf{y}_1^h \quad \text{given with } \|\mathbf{y}_0^h - \mathbf{y}_1^h\| = h \\ \boldsymbol{\varepsilon}_{n+2}^h = \boldsymbol{\Phi}(\mathbf{y}_{n+2}^h, \mathbf{y}_{n+1}^h, \mathbf{y}_n^h), \end{gathered} \tag{3.5}$$

where $\boldsymbol{\varepsilon}_{n+2}^h$ is the round-off error perpetrated in the computation of $\mathbf{y}_{n+2}^h$. It is assumed that

(H4) $\max_n \|\varepsilon_{n+2}^h\| = O(h^4)$, $n=0,1,\dots,[\gamma/h]-2$, and $\|\mathbf{y}_{n+1}^h - \mathbf{y}_n^h\| = h + O(h^4)$, $n=0,1,\dots,[\gamma/h]-1$.

(Recall from the introduction that in the absence of round-off $\|\mathbf{y}_{n+1}^h - \mathbf{y}_n^h\| = \|\mathbf{y}_1^h - \mathbf{y}_0^h\|$. It is easy to prove that (H4) is in particular verified if $\max_n \|\varepsilon_{n+2}^h\| = O(h^5)$. Of course a hypothesis like (H4) cannot hold in practice if a *given* computing tool is used for all values of h: it demands a sequence of incresingly accurate computing machines.)

Theorem 2. *Let* (H1)–(H4) *hold and assume that the points* $\mathbf{y}_n^h$ *are computed according to* (3.5) *with* $\|\mathbf{y}_0^h - \mathbf{Y}(0)\| = O(h^3)$, $\|\mathbf{y}_1^h - \mathbf{Y}(h)\| = O(h^3)$, *then for* h *small* $\mathbf{y}_n^h$ *is well defined* (*i.e.* $\mathbf{y}_n^h \in \Omega = \mathrm{dom}(\mathbf{F})$) *for* $n=0,1,\dots,[\gamma/h]$ *and*

$$\max_n \|\mathbf{y}_n^h - \mathbf{Y}(nh) - h^2\mathbf{W}(nh)\| = O(h^3), \tag{3.6}$$

where $\mathbf{W}(.)$ *is the solution of* (3.4).

Proof. If for h small, $\mathbf{y}_n^h \in \Omega$ for $n \leqq [\gamma/h]$, then, setting $\mathbf{v}_n^h = \mathbf{y}_n^h$, and taking $\mathbf{u}_n^h$ from (3.1), all the hypotheses in Theorem 1 are fulfilled and therefore

$$\|\mathbf{y}_n^h - \mathbf{Y}(nh) - h^2\mathbf{W}(nh)\| \leqq Ch^3$$

for $0 < h < h_0$, $n=0,1,\dots,[\gamma/h]$, with h_0, C suitable positive constants. Now reduce h if necessary to ensure that $\mathbf{y}_0^h, \mathbf{y}_1^h$ are both in Ω and that Ch^3 is less than the distance from the complement of Ω to the compact set $\{\mathbf{Y}(s) - h^2\mathbf{W}(s): 0 \leqq s \leqq \gamma,\ 0 \leqq h \leqq h_0/2\}$. For such values of h, induction with respect to n shows, in a standard way, that $\mathbf{y}_n^h \in \Omega$. The proof is now complete.

Remark. It was noted in the introduction that for one-dimensional problems ($d=1$) the scheme (1.2) is identical with the unstable recursion $y_{n+2} = 2y_{n+1} - y_n$. This shows that (H4) cannot be weakened if (3.6) is to hold. It could be concluded at first that the Lambert-McLeod method behaves badly with regard to round-off. However the following consideration should be taken into account. Let $\mathbf{y}_n, \mathbf{y}_{n+1}, \mathbf{y}_{n+2}$ satisfy (1.2) and suppose that $\mathbf{y}_n, \mathbf{y}_{n+1}$ are perturbed and become $\mathbf{y}_n + \varepsilon_n$, $\mathbf{y}_{n+1} + \varepsilon_{n+1}$ with $\varepsilon_n, \varepsilon_{n+1}$ orthogonal to $\mathbf{f}(\mathbf{y}_{n+1})$, i.e. approximately *orthogonal to the trajectory*; then $\mathbf{y}_{n+2}$ becomes, according to (1.2) $\mathbf{y}_{n+2} + \varepsilon_n$. Thus errors orthogonal to the trajectory propagate in a stable manner, and, by definition, it is only those errors we are interested in when integrating a trajectory problem. Round-off may cause the computed points to glide parallel to the trajectory, without affecting seriously the distance of those points to the trajectory.

A procedure to generate the missing starting point $\mathbf{y}_1$ is now presented. We consider a standard, second order one-step method with step size k

$$\mathbf{z}_{n+1} = \mathbf{z}_n + k\boldsymbol{\Psi}(\mathbf{z}_n, \mathbf{z}_{n+1}, k; \mathbf{g}) \tag{3.7}$$

for the discretization of systems $\mathbf{z}' = \mathbf{g}(\mathbf{z})$. Then, ignoring for simplicity round-off errors, we have

Theorem 3. *Assume that* (H1)-(H3) *hold,* $\mathbf{y}_0=\boldsymbol{\eta}$, $\mathbf{y}_1=\boldsymbol{\eta}+k\boldsymbol{\Psi}(\boldsymbol{\eta},\mathbf{y}_1,k;\mathbf{F})$, *with* $\boldsymbol{\Psi}$ *as above. Then for k small enough the recursion* (1.2) *is well defined* (*i.e.* $\mathbf{y}_n\in\Omega=\mathrm{dom}(\mathbf{F})$) *for* $n=0,1.\ldots,[\gamma/k]$ *and*

$$\max_n \|\mathbf{y}_n-\mathbf{Y}(nk)\|=O(k^2).$$

Proof. It is enough to observe that, upon setting $h:=\|\mathbf{y}_1-\mathbf{y}_0\|$, we can write

$$h=\|\mathbf{y}_1-\mathbf{Y}(k)\|+\|\mathbf{Y}(k)-\mathbf{Y}(0)\|+\|\mathbf{y}_0-\mathbf{Y}(0)\|=k+O(k^3)$$

and apply Theorem 2. Note that (3.7) must be of the second order in order to satisfy the condition $\|\mathbf{y}_1^h-\mathbf{Y}(h)\|=O(h^3)$ in the hypotheses of Theorem 2.

Remark. The *trapezoidal rule* is a good canditate for starting procedure, as it preserves the *circular exactness* [8]. Techniques similar to the ones in this paragraph can be employed to demonstrate that the starting method can be taken to be *Euler's rule* without damaging the order of convergence.

4. The CELF Method

A convergence result for the CELF scheme will now be presented. Round-off errors are not considered here, as their propagation has been discussed and tested in [10].

Theorem 4. *Assume that* (*H*1)-(*H*3) *hold. If* $\mathbf{y}_0$, $\mathbf{y}_1$ *are as in Theorem* 3, $t_0=0$, $t_1=k\|\mathbf{f}(\boldsymbol{\eta})\|^{-1}$, *then for k small enough the recurrence* (1.3) *is well defined* (*i.e.* $\mathbf{y}_n\in\Omega=\mathrm{dom}(\mathbf{F})$) *for* $n=0,1,\ldots,[\gamma/k]$ *and*

$$\max_n \|\mathbf{y}_n-\mathbf{y}(t_n)\|=O(k^2). \tag{4.1}$$

Proof. From Theorem 3 we know that $\|\mathbf{y}_n-\mathbf{Y}(nk)\|=O(k^2)$. Therefore we must establish an estimate $\|\mathbf{Y}(nk)-\mathbf{y}(t_n)\|=O(k^2)$. With the notation of Section 2, one has $\mathbf{Y}(nk)=\mathbf{y}(t(nk))$ and thus it is enough to prove $|t_n-t(nk)|=O(k^2)$. When $n=0,1$ we can write

$$t_0-t(0)=0$$
$$t_1-t(k)=k\|\mathbf{f}(\boldsymbol{\eta})\|^{-1}-(k\|\mathbf{f}(\boldsymbol{\eta})\|^{-1}+O(k^2))=O(k^2),$$

because $t(.)$ satisfies

$$t(0)=0, \qquad dt/ds=\|\mathbf{f}(\mathbf{Y}(s))\|^{-1}. \tag{4.2}$$

Set $h:=\|\mathbf{y}_0-\mathbf{y}_1\|$. As in the proof of Theorem 3, $h=k+O(k^3)$, while according to Theorem 2, $\mathbf{y}_n=\mathbf{Y}(nh)+h^2\mathbf{W}(nh)+O(h^3)$. This implies, after Taylor expansion, that if τ_{n+1} is given by (1.3b), then $\tau_{n+1}=h\|\mathbf{f}(\mathbf{Y}(nh))\|^{-1}+O(h^3)$. Therefore (1.3c) reads

$$t_{n+2}-t_n=2h\|\mathbf{f}(\mathbf{Y}(nh)\|^{-1}+O(h^3),$$

i.e. the values t_n arise from the midpoint rule applied to the quadrature problem (4.2), except for $O(h^2)$ perturbations in the integrand. Accordingly $t_n = t(nh) + O(h^2) = t(nk) + O(k^2)$, as required.

Remark. Note that the CELF method produces an equispacing $\|\mathbf{y}_{n+1} - \mathbf{y}_n\| = \text{constant}$, $n = 0, 1, \ldots$ of the depending variable and adjusts the increments $t_{n+1} - t_n$ of the independent variable. This behaviour should be compared with that of standard ODE solvers, which in fixed step implementations entail constant increments $t_{n+1} - t_n$ and variable distances $\|\mathbf{y}_{n+1} - \mathbf{y}_n\|$.

Again, it can be shown that *Euler's rule* can be used to generate $\mathbf{y}_1$ without impairing the order of convergence.

Final remark. Prof. M.N. Spijker has recently let us know that he had developed considerably [16] the idea of Strang used in this paper.

References

1. Calvo, M., Quemada, M.M.: On the stability of rational Runge-Kutta methods. J. Comp. Appl. Math. **8**, 289–292 (1982)
2. Dekker, K., Verwer, J.G.: Stability in the method of lines (To appear)
3. Kuo Pen-Yu: On stability of discretization. Sci. Sinica (Series A) **XXV**, 702–715 (1982)
4. Kuo Pen-Yu, Sanz-Serna, J.M.: Convergence of methods for the numerical solution of the Korteweg-de Vries equation. IMA J. Num. Anal. **1**, 215–221 (1981)
5. Kuo Pen-Yu, Wu Hua-Mo: Numerical solution of KDV equation. J. Math. Anal. Appl. **82**, 334–345 (1981)
6. Lambert, J.D., McLeod, R.J.Y.: Numerical methods for phase plane problems in ordinary differential equations. In: Numerical Analysis Proceedings Dundee 1979. A.G. Watson (ed.), pp. 83–97. Berlin, Heidelberg, New York: Springer 1980
7. Laurie, D.P.: Equispacing numerical methods for trajectory problems. In: Proceedings of the Sixth South Africal Symposium in Numerical Analysis. G.R. Joubert (ed.), pp. 39–45. Durban: Computer Science Department University of Natal 1980
8. McLeod, R.J.Y., Sanz-Serna, J.M.: Geometrically derived difference formulae for the numerical integration of trajectory problems. IMA J. Numer. Anal. **2**, 357–370 (1982)
9. Morton, K.W.: Initial value problems by finite difference and other methods. In: The state of the art in numerical analysis. D.A.H. Jacobs (ed.), pp. 699–756. London: Academic Press 1977
10. Sanz-Serna, J.M.: An explicit finite-difference scheme with exact conservation properties. J. Comput. Phys. **47**, 199–210 (1982)
11. Sanz-Serna, J.M., Manoranjan, V.S.: A method for the integration in time of certain partial differential equations. J. Comput. Phys. **52**, 273–289 (1983)
12. Stetter, H.J.: Stability of nonlinear discretization algorithms. In: Numerical solution of partial differential equations. J. Bramble (ed.), pp. 111–123. New York: Academic Press 1966
13. Stetter, H.J.: Analysis of discretization methods for ordinary differential equations. Berlin, Heidelberg, New York: Springer 1973
14. Strang, G.: Accurate partial difference methods. II. Nonlinear problems. Numer. Math. **6**, 37–46 (1964)
15. Wambeck, A.: Rational Runge-Kutta methods for solving systems of ordinary differential equations. Computing **20**, 333–342 (1978)
16. Spijker, M.N.: Equivalence theorems for nonlinear finite-difference methods. In: Numerische Behandlung nichtlinearer Integrodifferential und Differentialgleichungen. R. Ansorge, W. Törning (eds.), pp. 109–122. Lecture Notes in Mathematics 395. Berlin, Heidelberg, New York. Springer 1974