ELSEVIER Journal of Computational and Applied Mathematics 72 (1996) 393–406

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Stiffness in numerical initial-value problems

M.N. Spijker*

Department of Mathematics and Computer Science, University of Leiden, P.O. Box 9512, 2300 RA leiden, The Netherlands

Received 2 October 1995; revised 8 December 1995

Abstract

This paper reviews various aspects of stiffness in the numerical solution of initial-value problems for systems of ordinary differential equations.

In the literature on numerical methods for solving initial value problems the term "stiff" has been used by various authors with quite different meanings, which often causes confusion. This paper attempts to clear up this confusion by reviewing some of these meanings and by giving a distinct difinition of a "stiff situation".

Further, the paper reviews classical as well as recent estimates, from the literature, of the Newton stopping error relevant to implicit step-by-step methods. These estimates illustrate the fact that the theoretical analysis of numerical procedures in the stiff situation generally requires more subtle arguments than in the nonstiff case. They also illustrate the interesting fact that classical error estimates (derived without taking stiffness into account) can be highly relevant in certain stiff situations while being deceptive in others.

The paper concludes by presenting various open problems, and putting forward a conjecture, pertinent to the theoretical analysis of step-by-step methods in the stiff situation.

Keywords: Ordinary differential equations; Stiff initial-value problems; Stiffness; Implicit step-by-step methods; Newton's method; Order reduction in the stiff situation; Numerical stability

AMS classification: 65L05, 65L20, 65L70, 65M12, 65M20

1. Introduction

This paper deals with initial-value problems for systems of ordinary differential equations written in the form

$$U'(t) = f(t, U(t)) \quad \text{for } 0 \leqslant t \leqslant T, \qquad U(0) = u_0. \tag{1.1}$$

* E-mail: spijker@wi.leidenuniv.nl.

0377-0427/96/$15.00 © 1996 Elsevier Science B.V. All rights reserved
PII S0377-0427(96)00009-X

Here u_0 is a given vector in the s-dimensional real vector space $\mathbb{R}^s$, and $U(t) \in \mathbb{R}^s$ is unknown. Further, f denotes a given mapping from $[0, T] \times D$ to $\mathbb{R}^s$, where D is some subset of $\mathbb{R}^s$.

In the literature on numerical methods for (1.1) some initial-value problems are referred to as *stiff* problems. One of the prominent features of these problems is that they are extremely hard to solve by standard explicit step-by-step methods.

Curtiss and Hirschfelder [8] were the first to use the term stiff in this context. They devised a numerical procedure which nowadays is called the backward differentiation formula (BDF). It is remarkable that this very early procedure is still a basic ingredient of some highly successful modern software packages for certain stiff problems, like LSODE [16, 17] and VODE [3].

Since 1952, numerical methods for stiff problems have been studied extensively, in particular during these last 20 years. Hundreds of papers appeared, both dealing with the construction of efficient procedures and with the theoretical analysis of such procedures. The question thus seems to be legitimate of whether there is still anything left in this area worth to be studied.

The present paper has been motivated partly by the above question. It reviews the concept of stiffness, and it intends to make clear that there are very interesting issues in this field which have been studied only recently; it also touches on some open problems.

In Section 2.1 we define what we call the *stiff situation*. Section 2.2 discusses various definitions of stiffness from the literature, and Section 2.3 reviews a few criteria for stiffness.

Section 3 addresses the interesting issue of estimating the Newton stopping error occurring in the implementation of implicit procedures. Theoretical estimates of this error given in [22, 12, 28] are reviewed. This chapter illustrates the fact that the theoretical analysis of numerical procedures for solving (1.1) usually requires much more subtle arguments in the stiff situation than in the classical nonstiff case.

Section 4 glances at some new lines of research that are relevant to the stiff situation.

2. What is stiffness?

2.1. The stiff situation

In the following we denote by $|x|$ an arbitrary *norm* for the vectors $x \in \mathbb{R}^s$. We define L to be the smallest value with

$$|f(t, y) - f(t, x)| \leqslant L|y - x| \quad \text{(for all } t \in [0, T] \text{ and } x, y \in D). \tag{2.1a}$$

In all of the following this value is assumed to be finite, and is called the *Lipschitz constant* of f. Further, we define M to be the smallest value with

$$|W(t) - V(t)| \leqslant e^{M(t-t')}|W(t') - V(t')| \quad \text{(for all } t, t' \in [0, T], \text{ with } t' < t, \text{ and all solutions } V, W \text{ to our differential equation on } [t', t]). \tag{2.1b}$$

We call the latter value the *logarithmic norm* of f. It is well known that $M \leqslant L$, and that M can be considerably smaller than L (see, e.g., [9, 10]).

In order to review the concept of stiffness we consider *step sizes* $h_n > 0$, and *grid points* t_n in $[0, T]$ defined by $t_0 = 0$, $t_n = t_{n-1} + h_n$ $(n \geqslant 1)$. We shall deal with numerical processes by means of which approximations $u_n \simeq U(t_n)$ are calculated successively for $n = 1, 2, 3, \ldots$. The following are two well known simple examples of such processes:

$$u_n = u_{n-1} + h_n f(t_{n-1}, u_{n-1}), \tag{2.2}$$

and

$$u_n = u_{n-1} + h_n f(t_n, u_n). \tag{2.3}$$

There exists a beautiful theory, for the case of *nonstiff* problems, which is relevant to many numerical methods. A basic contribution to this theory was made by Dahlquist [9]. Subsequent, important elaborations of Dahlquist's work were made among others by P. Henrici, J.C. Butcher and H.J. Stetter. Excellent reviews of this classical theory can be found, e.g., in [14, 6]. This theory covers three main questions: (1) *Convergence* (i.e., the behaviour of $u_n - U(t_n)$ as $h_n \to 0$); (2) *Stability* (i.e., the effect on u_n of arbitrary perturbations, like rounding errors, introduced at some stage of the computations); (3) *Implicitness* (existence and numerical approximation of u_n defined only implicitly, like in (2.3)).

The above, classical theory is quite general in that it applies to arbitrary, nonlinear problems (1.1). But, in substantial parts of this theory one starts from the assumption that the product $h_n L$ is (sufficiently) small. There exist important classes of initial value problems where it is *natural*, or *feasible*, to approximate $U(t)$ with step sizes h_n satisfying

$$h_n L \gg 1. \tag{2.4}$$

Clearly, in such cases the above theory breaks down. In the following we will refer to (2.4) as the *stiff situation*. We note that inequalities, closely related to (2.4), occur at various places in the literature, notably in [8].

As an illustration to the above, consider the test problem

$$U'(t) = -10^6 [U(t) - \tfrac{1}{3}] \quad \text{for } 0 \leqslant t \leqslant T = 1, \qquad U(0) = u_0. \tag{2.5}$$

In this case we have $L = 10^6$, $M = -10^6$, and the true solution can be written in the form

$$U(t) = (u_0 - \tfrac{1}{3})\exp(-10^6 t) + \tfrac{1}{3}.$$

The first term on the right-hand side of this equality is often called the *nonsmooth* or *transient component* of $U(t)$, whereas the second one is called the smooth or nontransient component. After a short *transient phase*, of length T_0, the first component will have died out and $U(t) \simeq \frac{1}{3}$. For instance, when $u_0 = 1$, $U(t)$ equals $\frac{1}{3}$ up to an error of less than one percent, as soon as $T_0 \leqslant t \leqslant T$, with $T_0 = 5 \cdot 10^{-6}$.

Consider the numerical solution of (2.5) by the methods (2.2), (2.3) using $h_n \equiv 0.1$. Since $h_n L = 10^5$, we have (2.4), and the classical theory cannot be applied so as to predict any differences in the behaviour of the two methods.

Still, the methods (2.2) and (2.3) behave quite differently, even in the absence of a transient component in the true solution. Consider the case where $u_0 = \frac{1}{3}$, so that $U(1) = \frac{1}{3}$. If no rounding errors are present, both methods yield the perfect approximation $u_n = \frac{1}{3}$ to $U(1)$. But, starting the computations with $\tilde{u}_0 = 0.333333$ instead of $\frac{1}{3}$, the approximation $\tilde{u}_n$ to $U(1) = \frac{1}{3}$, obtained by (2.3), equals $\tilde{u}_n = 0.333333$, whereas the one obtained from (2.2) equals $\tilde{u}_n = -0.333300 \cdot 10^{44}$ (both last values $\tilde{u}_n$ rounded to six decimal places, and obtained with $h_n \equiv 0.1$).

Remark 2.1. Relation (2.4), stating that $h_n L$ is very large, might be considered to be slightly vague. But, in the spirit of (2.4), one can define that a situation corresponding to h_n, L is *more stiff* than one corresponding to h'_n, L' if $h_n L > h'_n L'$; this definition is less vague. Moreover, also in the spirit of (2.4), one may define a class of situations to be a *stiff class* if there is no bound on $h_n L$ simultaneously valid for all situations in this class (cf. [24, p. 5]).

2.2. *The term "stiff" — a state of confusion*

In the literature the term "stiff" has been used with different meanings—this is confusing. In the following we attempt to clarify this situation by reviewing some of these meanings.

The most early, and probably *most straightforward*, definition of stiffness is as follows (cf. [8, 15]).

Definition 2.1. Initial value problems are *stiff* if they are (exceedingly) difficult to solve by ordinary, explicit step-by-step methods, whereas certain implicit methods perform quite well.

This definition is attractive in that it is simple and all other definitions of stiffness seem to fall under that head. The author is not aware of other definitions not covered essentially by the above. But, Definition 2.1 also suffers from two shortcomings: (1) It is rather nonspecific—one may feel that too many, and too diverse, difficult problems are termed stiff; (2) It does not hold out a prospect of a mathematical framework for the theoretical analysis of numerical processes.

The *most common*, more specific definition of stiffness encountered in the literature is as follows (cf., e.g., [13, 10, 1, 21, 25]).

Definition 2.2. *Stiffness* occurs if: for most explicit methods, the largest step size h_n^* guaranteeing numerical stability is much smaller than the largest step h_n for which the local discretization error is still sufficiently small (in norm), i.e., $h_n^* \ll h_n$.

In the above, the term "local discretization error" designates the difference between $U(t_n)$ and the approximation at t_n generated by one application of the numerical method under the localizing assumption $u_i = U(t_i)$ (for $i < n$). The step h_n in Definition 2.2 is sometimes called the *natural step size.* Note that its size depends on

(a) the accuracy requirements regarding the numerical approximation, and

(b) the smoothness of the solution U one wants to approximate.

One may also let its size depend on

(c) the stability properties of the differential equation.

Therefore, whether according to the above definition stiffness actually occurs, is related to the factors (a), (b) and (c).

As an illustration, consider the task of tracking the solution $U(t)$ to (2.5), with moderate accuracy ε, on whole of the interval $[0, T]$. The natural step size h_n then typically satisfies

$$h_n L \lesssim 1 \quad \text{on } [0, T_0], \qquad h_n L \gg 1 \text{ on } [T_0, T],$$

whereas $h_n^* L \simeq 1$ throughout $[0, T]$. According to the last, common definition, stiffness thus occurs on $[T_0, T]$, but not on the transient interval $[0, T_0]$. It occurs just when the natural step size satisfies (2.4).

A *less common* definition of stiffness encountered in the literature is as follows (cf., e.g., [8, 24]).

Definition 2.3. *Stiffness* occurs if: for most explicit methods, the (largest) step $\tilde{h}_n$ guaranteeing both numerical stability and a sufficiently small local discretization error, is excessively small (compared to T), i.e., $\tilde{h}_n \ll T$.

As an illustration, consider the task pertinent to (2.5) of approximating $U(t)$, with moderate accuracy ε, only at $t = T$. According to the last definition one may say that stiffness occurs on whole of $[0, T]$ since, throughout this interval, $\tilde{h}_n \leqslant h_n^* \simeq L^{-1} = 10^{-6} \ll 1 = T$. It is feasible to carry out the last mentioned task, e.g., using (2.3), with h_n satisfying (2.4) throughout $[0, T]$.

We note that according to Definition 2.2, taking into account the factors (a), (b) for determining the natural step size for the last mentioned task, stiffness would occur only on $[T_0, T]$.

In order to present a further illustration, we introduce, for a given constant $\rho > 0$, the linear test problem

$$\begin{aligned} &U_1'(t) = -10^6\, U_2(t), \qquad U_1(0) = \tfrac{1}{3} + \rho, \\ &U_2'(t) = 10^6\,[U_1(t) - \tfrac{1}{3}], \qquad U_2(0) = 0, \\ &\quad \text{with } 0 \leqslant t \leqslant T = 1. \end{aligned} \tag{2.6}$$

Using the Euclidean norm $|x|$ for $x \in \mathbb{R}^2$, the constants M, L (defined in Section 2.1) corresponding to this problem satisfy $L = 10^6$, $M = 0$. The true solution is highly oscillatory, and can be written in the form

$$U_1(t) = \rho \cos(10^6 t) + \tfrac{1}{3}, \qquad U_2(t) = \rho \sin(10^6 t).$$

Consider the task of approximating the true solution to (2.6) by vectors $u_n \in \mathbb{R}^2$ satisfying $|u_n - U(t_n)| < \varepsilon$ for $t_n \in [0, T]$, with moderate error tolerance $\varepsilon \geqslant 2\rho$. According to Definition 2.3, stiffness occurs on whole of $[0, T]$ since, for most explicit methods, the largest step h_n^* guaranteeing numerical stability satisfies $h_n^* L \lesssim 1$, so that $\tilde{h}_n \leqslant h_n^* \lesssim L^{-1} \ll T$. It is feasible to carry out the present task using (2.3), with steps h_n satisfying (2.4) throughout $[0, T]$.

We note that, according to Definition 2.2, there is *no* stiffness in the last example, when $\varepsilon \simeq 2\rho$, since for most explicit methods the natural step size h_n satisfies $h_n L \lesssim 1$ and is *not* much larger than h_n^*.

Apart from the task, pertinent to (1.1), of finding u_n with $|u_n - U(t_n)| < \varepsilon$ the following task is sometimes set: it is required to find, for $t_n \in [0, T]$ vectors u_n and points t_n' with $|u_n - U(t_n')| < \varepsilon$, $|t_n' - t_n| < \varepsilon$ (cf. [24, p. 111]). Further, in the spirit of a so-called backward error analysis, the task is considered by some authors of finding u_n with $u_n = V(t_n)$, where $|V'(t) - f(t, V(t))| < \varepsilon$ $(0 \leqslant t \leqslant T)$, $V(0) = u_0$ (cf. [25, p. 65]).

Since the size of the natural step size depends on the task being set, it is clear that also the question of whether, according to Definition 2.2 or 2.3, stiffness occurs, should be related to the actual task under consideration.

Remark 2.2. Evidently, all of the above three definitions are rather vague. The author feels that they are more vague than the concept of a stiff situation as defined in Section 2.1 by relation (2.4).

2.3. *Criteria for stiffness in terms of the differential equation only*

The above definitions of stiffness share the property of referring to numerical step sizes or to (ordinary, explicit) numerical methods. These definitions do not tell us directly which initial value problems are liable to give rise to stiffness. This circumstance led several authors to formulate criteria for such initial value problems.

Below we state some of these criteria from the literature. (By some authors these criteria were used instead for *defining* stiffness!) The criteria apply in fact only to linear equations $U'(t) = AU(t)$, but are often applied to more general equations as well. We denote by $J(t, x)$ the Jacobian matrix (with respect to x) of $f(t, x)$, and its eigenvalues by λ.

Criterion 1 (*Corresponding to Definition* 2.2). Solutions to the differential equation exist some components of which decay much more rapidly than others.

The following requirements (2.7), (2.8) aim at formalizing this criterion:

$$\max \operatorname{Re}(\lambda) < 0, \qquad \frac{\max|\operatorname{Re}\lambda|}{\min|\operatorname{Re}\lambda|} \gg 1 \tag{2.7}$$

(see [21, p. 217]), or

$$T(L - M) \gg 1 \tag{2.8}$$

(cf. [31, p. 202]).

Criterion 2 (*Corresponding to Definition* 2.3). There exists a solution to the differential equation a component of which has a variation which is large compared to $1/T$.

The following aims at formalizing the latter criterion:

$$T \max|\lambda| \gg 1 \tag{2.9}$$

(see [24, p. 20]).

Remark 2.3. None of the criteria or formulae (2.7)–(2.9) just presented is completely satisfactory (as often remarked already by the corresponding authors themselves). For instance, stiffness according to Definition 2.2 can already occur when there is just one differential equation ($s = 1$), in which case (2.7) fails. Moreover, whether stiffness actually manifests itself, depends not only on the differential equation, but also on the factors (a) and (b) of Section 2.2.

Remark 2.4. In view of the Remarks 2.2 and 2.3 the author would be in favor of reserving the term "stiff" to designate the situation defined by (2.4). Moreover, in Section 2.2 we have seen that (2.4) is sufficiently general to cope with cases that are stiff either in the sense of Definition 2.2 or Definition 2.3. In the rest of the paper the term "stiff" will exclusively be used to denote the situation (2.4).

3. The Newton stopping error in the stiff situation

3.1. Liniger's error estimate

In recent years much attention has been paid, in the stiff situation (2.4), to the three main questions already mentioned in Section 2.1, viz. convergence, stability and implicitness. In the stiff situation one often (but not always!) arrives at conclusions deviating substantially from those in the nonstiff case. This will be exemplified in the present Section 3.

We address the problem of implicitness. It occurs in implicit linear multistep methods, implicit Runge–Kutta methods and many other implicit methods for (1.1) (see, e.g., [15]). For the ease of exposition we confine ourselves in the following to method (2.3), with fixed step sizes $h_n = h$. Throughout this Section 3, the set $D \subset \mathbb{R}^s$ is assumed to be open and convex, and the partial derivatives of f, up to the second-order, are assumed to exist and to be continuous on $[0, T] \times D$.

The vectors u_n defined by (2.3) will typically be obtained as numerical approximations to the solution x^* of the equation

$$F(x) = 0, \tag{3.1}$$

where

$$F(x) = -x + hf(t_n, x) + u_{n-1} \quad (\text{for } x \in D).$$

In the stiff situation (2.4) the solution x^* to (3.1) is usually approximated by Newton's method or a variant thereof. We consider the so-called *modified Newton process*

$$F'(x_0)(x_j - x_{j-1}) = -F(x_{j-1}) \quad \text{for } j = 1, 2, 3, \ldots . \tag{3.2}$$

Here $F'(x)$ denotes the Jacobian matrix of F at x. Further, x_0 denotes an initial guess, and x_j ($j \geqslant 1$) denote (improved) approximations to x^*.

In the following we study, for a given $j \geqslant 1$, the order of the errors due to the stopping of the iteration (3.2) after j steps and replacing x^* simply by x_j. We shall analyse these errors in terms of the step size h, following closely Liniger [22], Dorsselaer and Spijker [12] and Spijker [28].

In assessing the norm $|x^* - x_j|$ of the *Newton stopping error* $x^* - x_j$ (for $j \geqslant 1$) we assume that the initial guess x_0 satisfies

$$|x^* - x_0| = \mathrm{O}(h^q) \tag{3.3}$$

(with O-constant of moderate size, and $q > 0$).

By using Taylor series expansions in a straightforward way, the corresponding errors $x^* - x_j$ can be estimated. In 1971, Liniger derived in this manner

$$|x^* - x_j| = \mathrm{O}(h^{R(j)}) \quad \text{with } R(j) = (j+1)q + j \quad \text{for } j \geqslant 1. \tag{3.4}$$

However, as was common practise at that time, Liniger replaced in his derivation certain quantities, among which the product hL, simply by $\mathrm{O}(h)$. Evidently, in the stiff situation (2.4) such a quantity $\mathrm{O}(h)$ cannot be interpreted, in the standard fashion, as the product of a moderate O-constant and a small step size h. Accordingly, one may expect that the O-constant in (3.4) is excessively large (or that (3.4) makes sense only for excessively small h).

Table 1
Ratios for the Newton stopping error in Example 1

h	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
$\lvert x^* - x_1 \rvert / h^3$	4.9	4.9	4.7	4.7	4.7

3.2. Numerical experiments

In order to check the relevance of (3.4) in the stiff situation we consider the following example.

Example 1.

$$U_1'(t) = -10^8[U_1(t) - (t-2)^3] + [U_2(t) - 2]^2 + 2(t-2)^2, \qquad U_1(0) = -8,$$
$$U_2'(t) = -10^8[U_1(t) - (t-2)^3] + [U_1(t) - (U_2(t) - 2)^3] + 1, \qquad U_2(0) = 0,$$

with $0 \leqslant t \leqslant \frac{1}{2}$.

The true solution is $U_1(t) = (t-2)^3$, $U_2(t) = t$. We use the notation

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},$$

and introduce

$$D = \{x\colon \ -\infty < \xi_1 < \infty, \ -\tfrac{1}{4} < \xi_2 < 1\}.$$

Further, we define T, u_0 and $f\colon [0, T] \times D \to \mathbb{R}^2$, in an obvious way, so that (1.1) becomes equivalent to Example 1. Using the norm

$$|x| = |x|_1 = |\xi_1| + |\xi_2|,$$

it turns out that

$$L \simeq 2 \cdot 10^8, \qquad M = -1.$$

Under moderate accuracy requirements, a natural step size h for Example 1 need not be very small — since the true solution is smooth, and $M = -1$. Accordingly, such a step size satisfies $hL \gg 1$, and we have, in the sense of Section 2.1, a stiff situation.

We consider the function F introduced in Section 3.1, corresponding to Example 1, with

$$t_n = 1/10, \qquad u_{n-1} = U(t_{n-1}),$$

and we choose the natural initial guess $x_0 = u_{n-1}$.

Using the material in [12] it can be shown that, with these definitions, the Eq. (3.1) has a unique solution $x^* \in D$. Further, it can be shown that the estimate (3.3) holds with an O-constant of moderate size and $q = 1$.

For $j = 1$ the error estimate (3.4) thus reduces to

$$|x^* - x_1| = O(h^3). \tag{3.4'}$$

In fact, we may *not* expect that this estimate is reliable in the present situation, since its derivation was based on the assumption of a moderately sized value for hL.

In order to check the estimate we have listed the (rounded) actual ratios $|x^* - x_1|/h^3$, for various choices of h, in Table 1.

From Table 1 it is evident that the estimate (3.4) is reliable in the present example. This is surprising at first sight, since it was derived under assumptions not being satisfied here.

We might be tempted to believe that, after all, Liniger's estimate is reliable in the general stiff situation (2.4). In order to check (3.4) further we consider a second, closely related example taken from [28].

Example 2.

$$U_1'(t) = -10^8[U_1(t) - (U_2(t) - 2)^3] + 3(U_2(t) - 2)^2, \qquad U_1(0) = -8,$$

$$U_2'(t) = 10^8[U_1(t) - (U_2(t) - 2)^3] + 1, \qquad U_2(0) = 0,$$

with $0 \leqslant t \leqslant \frac{1}{2}$.

The true solution is equal to the one in Example 1. We use the same set D and norm $|x|$ as in Example 1. It turns out that, for the function f corresponding to Example 2,

$$L \simeq 3 \cdot 10^9, \qquad M = 0.$$

Under moderate accuracy requirements, a natural step size h need not be very small. Therefore $hL \gg 1$, and we have again a stiff situation in the sense of Section 2.1.

We consider the function F, corresponding to Example 2, with the same t_n, u_{n-1}, x_0 as above, i.e.,

$$t_n = 1/10, \qquad u_{n-1} = U(t_{n-1}), \qquad x_0 = u_{n-1}.$$

Again it can be proved that (3.3) holds here with an O-constant of moderate size and $q = 1$. Liniger's estimate would thus again imply (3.4′).

Some (rounded) actual ratios $|x^* - x_1|/h^3$ have been listed in Table 2.

The entries in the Table 2 are quite different from those in Table 1. It is clear that (3.4) is *not* reliable in the present situation. It provides no insight regarding the actual error $x^* - x_1$ for Example 2.

3.3. *Theoretical explanations*

We note that Example 2 may be called a *strongly nonlinear problem* in that the large factor 10^8 affects both first- and second-order partial derivatives of the corresponding function f. For such problems, a general theoretical framework was devised [12, 28] in which reliable estimates of $|x^* - x_j|$ can be derived for arbitrary $L \geqslant 0$ and $M \leqslant 0$. These general estimates are as follows:

$$|x^* - x_j| = O(h^{Q(j)}) \quad \text{with } Q(j) = (j+1)q \quad \text{for } j \geqslant 1. \tag{3.5}$$

Here the O-constant, say K_j, is of moderate size, provided the O-constant K of (3.3) is moderate. In fact, K_j can be related to K via the formula

$$K_j = [(1+\sqrt{2})L_1 K]^j K. \tag{3.6}$$

Here L_1 denotes a so-called *"relative" Lipschitz constant* for the Jacobian matrix $J(t, x)$ (with respect to x) of $f(t, x)$; i.e., the constant L_1 is such that the following condition is fulfilled:

$$J(t, y) - J(t, x) = J(t, x)e(t, x, y) \quad \text{with an } s \times s \text{ matrix } e(t, x, y)$$
$$\text{satisfying } \|e(t, x, y)\| \leqslant L_1 |y - x| \text{ for all } t \in [0, T] \text{ and } x, y \in D. \tag{3.7}$$

In this condition, $\|\cdot\|$ denotes the matrix norm that is induced by an arbitrary, given vector norm $|\cdot|$ on $\mathbb{R}^s$, and L_1 is assumed to be of *moderate size*. Relation (3.6) was proved, essentially under the assumption that (cf.loc.cit.)

$$\text{(3.7) holds,} \quad \text{and} \quad M \leqslant 0.$$

In Example 2 the last assumption is satisfied, for the norm $|x| = |x|_1$ of Section 3.2 with $L_1 = 12$, so that (3.5) is in force with moderate O-constant K_j. For $q = 1, j = 1$ relation (3.5) reduces to

$$|x^* - x_1| = \mathrm{O}(h^2) \tag{3.5'}$$

This estimate is nicely in agreement with the numerical experiments displayed in Table 2.

Example 1 may be called a *weakly nonlinear problem* in that the large factor 10^8 affects only the first (and not the second) order partial derivatives of the corresponding function f. Also for such problems a general theoretical framework was devised [28] in which reliable estimates can be derived for arbitrary $L \geqslant 0$, $M \leqslant 0$. Within that framework Liniger's estimate was proved to be reliable.

In fact, (3.4) can be proved to hold with an O-constant.

$$K_j = \left[\frac{1+\sqrt{2}}{2} L_0 K\right]^j K. \tag{3.8}$$

Here L_0 denotes an *"absolute" Lipschitz constant* for the Jacobian $J(t, x)$; i.e., L_0 is such that the following condition is fulfilled:

$$\|J(t, y) - J(t, x)\| \leqslant L_0 |y - x| \quad \text{for all } t \in [0, T] \text{ and } x, y \in D. \tag{3.9}$$

In this condition L_0 is assumed to be of *moderate size*. Relation (3.8) was proved, essentially under the assumption that (cf.loc.cit.)

$$\text{(3.9) holds,} \quad \text{and} \quad M \leqslant 0.$$

In Example 1 the last assumption is satisfied, for the norm $|x| = |x|_1$ of Section 3.2 with $L_0 = 16$, so that (3.4′) is in force with moderate O-constant K_1. This nicely explains the moderate values for the ratios $|x^* - x_1|/h^3$ displayed in Table 1.

Table 2
Ratios for the Newton stopping error in Example 2

h	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
$\|x^* - x_1\|/h^3$	$0.10 \cdot 10^2$	$0.97 \cdot 10^2$	$0.96 \cdot 10^3$	$0.96 \cdot 10^4$	$0.96 \cdot 10^5$

It is clear that Example 2 does *not* fit within our last theoretical framework. It can be verified that the corresponding function f violates (3.9), with the norm of Section 3.2, for each $L_0 < 10^9$.

Remark 3.1. Similarly as the constant L in (2.1a), the value L_0 is *not* invariant under a rescaling of the independent variable t. Moreover, the values L_0, L_1 fail to be invariant under a rescaling of the dependent variable $U(t)$. On the other hand, the quantities

$$L' = hL, \qquad L_0' = hL_0 \max_t |U(t)|, \qquad L_1' = L_1 \max_t |U(t)|$$

are invariant under such rescalings. Therefore, from a formal point of view, it might have been more appropriate to require in connection with (3.9), (3.7) that L_0', L_1' (instead of L_0, L_1) are of moderate size. But, in the above discussion involving (3.9), (3.7) we assumed tacitly that $h \in (0, 1]$, whereas in the Examples 1 and 2 we have $\max_t |U(t)| = 8$. In this situation, moderate values of L_0, L_1 already imply moderate values of L_0', L_1'.

Remark 3.2. Orders $R(j)$, $Q(j)$, analogues to those in (3.4), (3.5), were derived for variants to the iteration (3.2) as well as for generalizations of the numerical process (2.3). For details we refer to the references already mentioned in Section 3.1.

4. About new lines of research relevant to the stiff situation

4.1. Nonlinear problems

In the second part of Section 3.3 we have seen that classical results (derived without taking stiffness into account) can still be relevant in certain nonlinear stiff situations. On the other hand, in the second part of Section 3.2, we have seen that the same results can also be misleading in other nonlinear stiff situations.

Apparently, this is related to the great diversity of nonlinear problems in which the stiff situation occurs. Due to this diversity, the theoretical analysis of methods in the nonlinear stiff situation has offered a great challenge to numerical analysis, and many questions are still open.

For instance, a rigorous theory is still missing which describes the error propagation (stability) satisfactorily of numerical procedures in the situation (3.7) (cf. [28]). Further, general estimates of the type (3.5) are still missing for certain important implicit Runge–Kutta methods (cf. [12, 27]). As a final example, we may mention that some interesting convergence results, established in the theoretical numerical analysis of dynamical systems, seem not yet to have been adapted so as to become relevant to (strongly nonlinear) stiff situations (cf. [20]).

4.2. Linear problems

Difficult questions and open problems not only arise due to the nonlinearity of (1.1). Even for the linear problem

$$U'(t) = AU(t) + r(t), \qquad U(0) = u_0, \tag{4.1}$$

with constant $s \times s$ matrix A, there are interesting open theoretical questions. We illustrate this point in the following.

Assume that the logarithmic norm M, corresponding to (4.1), satisfies $M \leqslant 0$. Consider the numerical solution of (4.1) by an arbitrary A-stable Runge–Kutta method (cf. [15]). Denoting the stability function of the Runge–Kutta method by φ, the corresponding numerical process, with step size h, can be written in the form

$$u_n = \varphi(hA)u_{n-1} + r_n.$$

We call this process *stable* if the norms $\|\varphi(hA)^n\|$ (induced by the given vector norm on $\mathbb{R}^s$, e.g., the maximum norm) are of moderate size.

Consider the problem of proving stability in the stiff situation, where $hL = h\|A\| \gg 1$.

It was proved by Lubich and Nevanlinna [23] that, under the above assumptions, there is a constant γ (only depending on the function φ, and not on A or hL) such that the following stability estimate holds:

$$\|\varphi(hA)^n\| \leqslant \gamma s \quad \text{(for all } s \geqslant 1,\ n \geqslant 1 \text{ and } h > 0).$$

We note that (4.1) can result from an application of the process of semi-discretization (method of lines) to a given initial-boundary value problem for a linear partial differential equation. In such a case the dimension s is related to the accuracy of the semi-discretization, and can attain (arbitrarily) large values. In view of such applications, it is an interesting question of whether the above stability estimate can be improved to

$$\|\varphi(hA)^n\| \leqslant \gamma s^\alpha \quad \text{(for all } s \geqslant 1,\ n \geqslant 1 \text{ and } h > 0), \tag{4.2}$$

for some fixed $\alpha < 1$. Up to now this question is open.

Conjecture. The estimate (4.2) holds with $\alpha = \frac{1}{2}$, and in general not with $\alpha < \frac{1}{2}$.

For evidence about this conjecture and for related material we refer to [11, 23, 26, 29].

4.3. Novel methods and variants to (1.1)

The open questions, related to stiffness, mentioned above concern classical numerical methods (Runge–Kutta methods and linear multistep methods) in the solution of problem (1.1). It is not surprising that many, equally interesting questions, related to stiffness, arise in the construction and analysis of novel numerical methods as well as in the numerical solution of problems that are variants to (1.1).

Recently, promising novel techniques have been developed, notably in the area of parallel numerical methods for problem (1.1). In this interesting field of research considerable attention has been given to the stiff situation. For recent work on parallel-methods, relevant to the stiff as well as the nonstiff situation, we refer to the review paper [4] and the monograph [5].

Situations that are very similar to (2.4) are also encountered in the numerical solution of initial-value problems that are not exactly of the form (1.1). Such initial-value problems arise notably in the context of delay-differential equations as well as certain partial differential equations. For some recent lines of research, related to stiffness in the numerical solution of these two sorts of differential equations, we refer to Bocharov and Romanyukha [2], Hout [19] and Zlatev [32], respectively.

Acknowledgements

The author is most thankful to Prof. Dr. S.P. Nørsett for inviting him to give a plenerary lecture on stiffness at the workshop ODE to NODE (Geiranger, June 1995). The present paper is an outgrowth of the lecture given at the workshop. The author also gratefully acknowledges discussions with Dr. K.J. in 't Hout and Dr. W.H. Hundsdorfer about the topic of this paper, as well as comments by two anonymous referees. The numerical results listed in the Tables 1 and 2 were obtained by J. Groeneweg.

References

[1] R.C. Aiken, *Stiff Computation* (Oxford University Press, New York 1985).
[2] G.A. Bocharov and A.A. Romanyukha, Numerical treatment of the parameter identification problem for delay-differential systems arising in immune response modelling, *Appl. Numer. Math.* **15** (1994) 307–326.
[3] P.N. Brown, G.D. Byrne and A.C. Hindmarsh, VODE: A variable-coefficient ODE solver, *SIAM J. Sci. Statist. Comput.* **10** (1989) 1038–1051.
[4] K. Burrage, Parallel methods for initial value problems, *Appl. Numer. Math.* **11** (1993) 5–25.
[5] K. Burrage, *Parallel and Sequential Methods for Ordinary Differential Equations* (Oxford University Press, Oxford, 1995).
[6] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations* (Wiley, Chichester, 1987).
[7] M. Crouzeix and A.L. Mignot, *Analyse Numérique des Équations Différentielles* (Masson, Paris, 1984).
[8] C.F. Curtiss and J.O. Hirschfelder, Integration of stiff equations, *Proc. Nat. Acad. Sci.* **38** (1952) 235–243.
[9] G. Dahlquist, *Stability and error bounds in the numerical integration of ordinary differential equations*, Trans. Roy. Inst. Techn. (Stockholm) **130** (1959).
[10] K. Dekker and J.G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations* (North-Holland, Amsterdam, 1984).
[11] J.L.M. van Dorsselaer, J.F.B.M. Kraaijevanger and M.N. Spijker, Linear stability analysis in the numerical solution of initial value problems, *Acta Numerica* (1993) 199–237.
[12] J.L.M. van Dorsselaer and M.N. Spijker, The error committed by stopping the Newton iteration in the numerical solution of stiff initial value problems, *IMA J. Numer. Anal.* **14** (1994) 183–209.
[13] C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations* (Prentice-Hall, Englewood Cliffs, 1971).

[14] E. Hairer, S.P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations*, Vol. I (Springer, Berlin, 1993).
[15] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations, Vol. II* (Springer, Berlin, 1991).
[16] A.C. Hindmarsh, LSODE and LSODI, Two new initial value ordinary differential equation solvers, *ACM-SIGNUM Newsletter* **15**(4) (1980) 10–11.
[17] A.C. Hindmarsh and K. Radhakrishnan, Description and use of LSODE, the Livermore Solver for Ordinary Differential Equations, NASA Ref. Publ. 1327, and LLNL Report UCRL-ID-113855, 1994.
[18] P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations* (Wiley, New York, 1962).
[19] K.J. in 't Hout, The stability of θ-methods for systems of delay differential equations, *Ann. Numer. Math.* **1** (1994) 323–334.
[20] A.R. Humphries and A.M. Stuart, Runge–Kutta methods for dissipative and gradient dynamical systems, *SIAM J. Numer. Anal.* **31** (1994) 1452–1485.
[21] J.D. Lambert, *Numerical Methods for Ordinary Differential Systems* (Wiley, Chichester, 1991).
[22] W. Liniger, A stopping criterion for the Newton–Raphson method in implicit multistep integration algorithms for nonlinear systems of ordinary differential equations, *Comm. ACM* **14** (1971) 600–601.
[23] C. Lubich and O. Nevanlinna, On resolvent conditions and stability estimates, *BIT* **31** (1991) 293–313.
[24] W.L. Miranker, *Numerical Methods for Stiff Equations* (Reidel Publ. Comp., Dordrecht, 1981).
[25] L.F. Shampine, *Numerical Solution of Ordinary Differential Equations* (Chapman & Hall, New York, 1994).
[26] M.N. Spijker, Numerical ranges and stability estimates, *Appl. Numer. Math.* **13** (1993) 241–249.
[27] M.N. Spijker, On the error committed by stopping the Newton iteration in implicit Runge–Kutta methods, *Ann. Numer. Math.* **1** (1994) 199–212.
[28] M.N. Spijker, The effect of the stopping of the Newton iteration in implict linear multistep methods, *Appl. Numer. Math.* **18** (1995) 367–386.
[29] M.N. Spijker and F.A.J. Straetemans, Error growth analysis, via stable regions, for discretizations of initial value problems, Report TW-96-01, Department of Mathematics, Leiden University (1996).
[30] H.J. Stetter, *Analysis of Discretization Methods for Ordinary Differential Equations* (Springer, Berlin, 1973).
[31] K. Strehmel and R. Weiner, *Numerik gewöhnlicher Differentialgleichungen* (Teubner, Stuttgart, 1995).
[32] Z. Zlatev, *Computer Treatment of Large Air Pollution Models* (Kluwer, Dordrecht, 1995).