In Numerical Treatment of Differential Equations,
Proceedings of the Fourth Seminar "NUMDIFF - 4",
held in Halle, 1.987, Edited by Karl Strehmel,
Teubner-Texte zur Mathematik, Band 104,
Leipzig 1.988, pp.216-226.

# A DEFINITION OF STABILITY FOR NONLINEAR PROBLEMS

J. C. López-Marcos, J. M. Sanz-Serna
Valladolid, Spain

## 1. Introduction

We study the definition of stability in nonlinear situations. We show that the 'naive' concept obtained by extending the standard linear definition is too narrow to cater for truly nonlinear problems. The merit of two better definitions due to H.B. Keller and H. J. Stetter is then assessed. These improved definitions, while very satisfactory when dealing with ordinary differential equations, are not well suited for the partial differential equation case. An alternative concept is suggested which extends Keller's definition.

We begin by recalling in Section 2 the basic concepts of the linear case [5].

## 2. Preliminaries

Discretizations. We consider a fixed, given problem concerning a, not necessarily linear, differential or integral equation. Let u denote a solution of this problem. (Well-behaved nonlinear problems may of course possess more than one solution.) Often u cannot be readily expressed in terms of the data and one must obtain a numerical approximation $U_h$ to u. We have appended a subscript h in order to reflect that $U_h$ depends on a (small) parameter h such as a mesh-size, element diameter etc... In what follows we always assume that h takes values in a set of positive numbers H, with inf H = 0. The numerical approximation $U_h$ is reached by solving a discrete problem

(2.1) $$\Phi_h(U_h) = 0,$$

where, for each h in H, $\Phi_h$ is a fixed mapping with domain $D_h \subset X_h$ and taking values in $Y_h$. Here $X_h$ and $Y_h$ are vector spaces, either both real or both complex, with

(2.2) $$\dim X_h = \dim Y_h < \infty.$$

It is typical of nonlinear situations that $\Phi_h$ cannot be defined everywhere in $X_h$: it may involve logarithms, square roots etc... We assume, until further notice, that $D_h$ denotes the largest subset of $X_h$ where the analytic expression for $\Phi_h$ makes sense. As h ranges in H, the family of problems (2.1) is referred to as a discretization.

Global error, convergence. When a solution $U_h$ of (2.1) has been obtained, the question arises as to what extent does $U_h$ provide a good approximation to u. A first difficulty in answering this question stems from the fact that $U_h$ can be completely dissimilar to u. (Typically $U_h$ is a vector with, say, d entries, while u is a function of one or several continuous variables.) This difficulty is circumvented as follows. Since any solution $U_h$ of (2.1) is bound to be an element in $D_h$, we first make up our minds as to which element $u_h$ in $D_h$ should be regarded as the most desirable numerical result. (Typically $u_h$ contains d nodal values of u.) Once $u_h$ has been chosen, the vector $e_h = u_h - U_h$ is defined to be the global error in $U_h$. In order to measure the size of this error, we introduce, for each h in H, a norm $\|\cdot\|$ in $X_h$. (Norms in different spaces will simply be denoted by $\|\cdot\|$ without mention of the space.) We say that the discretization (2.1) is convergent if there exists $h_o > 0$ such that, for each h in H with $h \leq h_o$, (2.1) possesses a solution $U_h$ and as $h \to 0$, $\lim \|u_h - U_h\| = \lim \|e_h\| = 0$. If, furthermore, $\|e_h\| = O(h^p)$ as $h \to 0$, then the convergence is said to be of order p.

Local error, consistency. The local (discretization) error in $u_h$ is defined to be the element $l_h = \Phi_h(u_h) \in Y_h$. The measurement of the size of $l_h$ requires, therefore, the introduction of norms $\|\cdot\|$ in $Y_h$. When these norms have been chosen, (2.1) is said to be consistent (resp. consistent of order p) if, as $h \to 0$, $\|l_h\| \to 0$ (resp. $\|l_h\| = O(h^p)$).

Linear stability. So far no distinction has been made on the linear or nonlinear character of (2.1). In dealing with stability it is expedient to first present the linear case, where there is universal agreement in the literature. Let us then suppose that (2.1) takes the linear form

$$(2.3) \qquad \Phi_h(U_h) \equiv \Psi_h U_h - f_h = 0 \qquad \text{or} \qquad \Psi_h U_h = f_h,$$

where, for each h in H, $\Psi_h$ is a linear operator (matrix) mapping $X_h$ into $Y_h$ and $f_h$ is a fixed element in $Y_h$. If norms in $X_h$ and $Y_h$ have been introduced as above, then (2.3) is stable if there exist positive constants S (the stability constant) and $h_o$ such that for each h in H, $h \leq h_o$, and for each $V_h$ in $X_h$

$$(2.4) \qquad \|V_h\| \leq S \|\Psi_h V_h\| .$$

The bound (2.4) is employed to derive two kinds of properties:

(i) Assume that for h small (2.3) is uniquely solvable. (The solvability of the discrete equations is obvious in some applications, e.g. explicit methods in initial value problems, but cannot be taken for granted in many other settings.) Consideration of (2.4) for h varying leads to the easy proof of the implication "consistent (of order p) + stable → convergent (of order p)." In fact

$$(2.5) \qquad \|u_h - U_h\| \leq S \|\Psi_h(u_h - U_h)\| = \|\Phi_h(u_h) - \Phi_h(U_h)\|,$$

which leads to

$$(2.6) \qquad \|e_h\| = \|u_h - U_h\| \leq S \|\Phi_h(u_h)\| = S \|l_h\|$$

and the last term is o(1) (resp. $O(h^p)$) for consistent methods (resp. for methods consistent of order p).

Remark 2.1. If (2.3) represents an initial value problem, then, under the appropriate choice of norms, (2.4) is identical to the requirement of uniform boundedness of the powers of the evolution operator which maps a time-level into the next (Lax stability see [5]).

Remark 2.2. In practice the computed $U_h$ does not satisfy (2.1) but $\Phi_h(U_h) = \rho_h$, with $\rho_h$ accounting for round-off. There is no difficulty in incorporating this fact into the analysis; the bound (2.6) is obviously replaced by $\|e_h\| \leq S ( \|l_h\| + \|\rho_h\| )$.

(ii) If we now look at a fixed value of h, $h \leq h_o$, then (2.4) shows that $\Psi_h V_h = 0$ implies $V_h = 0$, or (see (2.2)) that $\Psi_h^{-1}$ exists. We emphasize that in the linear case the existence and uniqueness of solutions follows from the stability bound (2.4).

## 3. Nonlinear stability, first attemps

In the rest of the paper we return to the nonlinear discretization (2.1) and assume that we have made, for each h, definite choices of elements $u_h$ in $D_h$ and norms in $X_h$ and $Y_h$. Our task is to study how best to define the notion of stability under the restriction that the conclusions in points (i)-(ii) above are still valid.

On looking at (2.5)-(2.6), it is clear that if the existence of $U_h$ is guaranteed, then consistency implies convergence (point(ii)) provided that a bound

(*) $$\|V_h - W_h\| \le S\,\|\Phi_h(V_h) - \Phi_h(W_h)\|$$

holds for h in H, h sufficiently small, $V_h$ and $W_h$ arbitrary in $D_h$, S independent of h and $V_h$, $W_h$. This leads to the following temptative definition of stability.(N = naive).

Definition N. The discretization (2.1) is stable if there exist positive constants S (the stability constant) and $h_0$, such that for each $V_h$, $W_h$ in $D_h$, and each h in H, $h \le h_0$, the bound (*) holds.

The techniques presented in the next Section show that, under very mild hypotheses, a consistent discretization which is stable in the sense of definition N possesses unique solutions $U_h$ for small h (point (ii)). The conclusion appears to be that this definition is satisfactory and in fact this definition and minor variants of it have often been employed in the literature. However the Definition N is so restrictive that classifies as unstable many useful numerical methods.

Example. Let $\alpha$ be a fixed constant $0 \le \alpha < 1$. If N is a positive integer and $h = 1/N$, consider the discrete equations

(3.1) $$U_0 - \alpha = 0,$$
$$(U_{n+1} - U_n)/h - f(U_n) = 0, \quad n = 0, 1, \ldots, N-1;$$

where $f(v) = v^2$. This results from the application of Euler's rule to the problem $u(0) = \alpha$, $du/dt = u^2$, $0 \le t \le 1$, with solution $u(t) = \alpha/(1-\alpha t)$. Here $U_h = (U_0, U_1, \ldots, U_N)$ and $X_h = D_h$ and $Y_h$ are spaces of (N+1)-vectors. We choose $u_h$ equal to the grid restriction and norms

$$\|V_h\| = \max_n |V_n|, \text{ if } V_h = (V_0, V_1, \ldots, V_N) \in X_h,$$
$$\|F_h\| = |F_0| + \sum_{1 \le n \le N} h|F_n|, \text{ if } F_h = (F_0, F_1, \ldots, F_N) \in Y_h.$$

These are typical norms for initial value problems [5]. We now prove that (3.1) is not N-stable. It is enough to consider the (N+1)-vector $V_h$ in $X_h$ recursively defined

(3.2) $$V_0 - \alpha = 1 - \alpha,$$
$$(V_{n+1} - V_n)/h - f(V_n) = 0, \quad n = 0, 1, \ldots, N-1.$$

It follows from [6] that as $h \to 0$, $V_N \sim 1/(h \cdot \log|h|)$. Then, as $h \to 0$, $\|u_h - V_h\|$ increases without bound, while $\|\Phi_h(V_h)\| = 1 - \alpha$ and, by consistency, $\|\Phi_h(u_h)\| = O(h)$. Thus (*) cannot hold with S independent of h.

It is expedient to gain insight into this example. The equations (3.2), which so far have been viewed as the result of perturbing the right hand side of (3.1) are the Euler discretization of the problem $v(0) = 1$, $dv/dt = v^2$, whose solution $1/(1-t)$ is infinite at time $t = 1$. This explains why it is not reasonable to expect that $|u(1)-V_N|$ be boundable in terms of the size $|1-\alpha|$ of the perturbation as required by the definition N. It is clear that, for stability, we should not ask that (*) holds if the perturbation $\Phi_h(V_h)$ is too large or alternatively if $V_h$ is too far away from $u_h$. This introduces the important idea of stability threshold discussed next.

## 4. Stability, Stetter's definition

The following definition is used by Stetter [ 7]:

Definition S. The discretization (2.1) is stable if there exist positive constants S (the stability constant) and $h_o$, and a value r, $0 < r \leq \infty$ (the stability threshold) such that for each h in H, $h \leq h_o$ and for each $V_h$, $W_h$ in $D_h$ with

(4.1) $$\|\Phi_h(V_h) - \Phi_h(u_h)\| < r, \quad \|\Phi_h(W_h) - \Phi_h(u_h)\| < r,$$

the bound (*) holds.

It is easy to show that, when (2.1) takes the linear form (2.3), a discretization which is stable with threshold $r < \infty$ is also stable for the threshold $r = \infty$. For this choice (4.1) provides no restriction and we accordingly recover the usual linear definition. Note that the present definition is less demanding than the naive definition N, since (*) is only asked for vectors $V_h$, $W_h$ satisfying the threshold condition (4.1). In the context of the example above, it is not necessary now to cater for the unruly vector $V_h$ generated in (3.2): with $r < 1-\alpha$, $V_h$ would not pass the test (4.1) that vectors must undergo before we let then feature in (*). The main point is that the threshold condition (4.1) which leaves out unwelcome vectors is not in the way of the two elements $u_h$ and $U_h$ which we want to substitute in (*) to estimate the global error: with $V_h = u_h$, $W_h = U_h$ the condition (4.1) is satisfied for small h; this is obvious for $u_h$ and consistency takes care of $U_h$. Then the implication "stability + consistency → convergence" can be proved in the usual simple way, provided that the existence of solutions has been established.

Moreover the Definition S is also successful in proving existence of solutions (point ii). Here the crucial stepping-stone is the following Lemma due to Stetter

Lemma. Let X, Y be normed spaces with ths same finite dimension, $\Phi : D \subset X \to Y$ a mapping and u an element in D. Assume that: (L1) There exists $R \geq 0$ such that the open ball B(u,R) is contained in D. (L2) $\Phi$ is continuous in B(u,R). (L3) there exist $S > 0$ and r, $0 < r \leq \infty$ such that if v and w are in B(u,R) and

$$\|\Phi(v) - \Phi(w)\| < r, \quad \|\Phi(w) - \Phi(u)\| < r$$

then

$$\| v - w \| \leq S \| \Phi(v) - \Phi(w) \|.$$

Under the hypotheses (L1)-(L3), $\Phi$ possesses an inverse mapping $\Phi^{-1}$ defined in the open ball $B(\Phi(u), r_o)$ and taking values in B(u,R). Here $r_o = \min (R/S, r)$. Furthermore this inverse is Lipschitz continuous with constant S.

We now give the following important result (Stetter [7], Th. 1.2.3):

Theorem S. Let the discretization (2.1) be consistent and S stable. Assume that there exists a constant R (independent of h) such that, for each h in H, $\Phi_h$ is continuous in the open ball $B(u_h,R) \subset D_h$. Then:

(a) For h in H, h sufficiently small, the discrete equations (2.1) possess a solution $U_h$ in $B(u_h,R)$.

(b) This solution is unique in $D_h$.

(c) As $h \to 0$, the solutions converge and the order of convergence is not smaller than the order of consistency.

Proof. If S is the stability constant and r the stability threshold, application of the Lemma to $\Phi_h$ guarantees the existence of an inverse $\Phi_h^{-1}$ defined in $B(\Phi_h(u_h),r_o)$, $r_o = \min(R/S,r)$, h sufficiently small. Consistency implies that for h small $\|\Phi_h(u_h)\| < r_o$, i.e. $0 \in B(\Phi_h(u_h),r_o) = \mathrm{dom}(\Phi_h^{-1})$. Then $U_h = \Phi_h^{-1}(0)$ is in $B(u_h,R)$ and solves (2.1). This proves (a). The uniqueness follows without resorting to the Lemma. In fact for h small, any solution $Z_h$ of (2.1) satisfies (4.1). Therefore if $Z_{1h}$, $Z_{2h}$ are two solutions of (2.1)

$$\|Z_{1h}-Z_{2h}\| \leq S\,\|\Phi_h(Z_{1h})-\Phi_h(Z_{2h})\| = 0.$$

The easy proof of (c) has been given above.

## 5. Stability, H. B. Keller's definition

H.B. Keller [2] presented an alternative way of introducing thresholds in the naive definition.

Definition K. The discretization (2.1) is stable if there exist constants $S > 0$ (the stability constant), $h_o > 0$ and R, $0 < R \leq \infty$ (the stability threshold), such that for each h in H, $h \leq h_o$, the open ball $B(u_h,R)$ is contained in $D_h$ and for each $V_h$ and $W_h$ in this ball the bound (*) holds.

Again this definition reduces to the standard one in linear cases. Now the stability inequality is only asked for $V_h$ and $W_h$ near $u_h$. This involves a threshold condition which explicitly involves $V_h$ and $W_h$ in lieu of their images $\Phi_h(V_h)$, $\Phi_h(W_h)$ which feature in Definition S.

As distinct from the situation with Stetter's definition, it is far from obvious that the stability in the sense of Keller and the consistency of a discretization imply convergence, even if the existence of $U_h$ has been proved independently. In fact it seems that, even if the existence of $U_h$ is guaranteed, the stability estimate (*) cannot be applied to bound the global error without first proving the a priori bound $\|u_h-U_h\| < R$. This notion is wrong. There is no need for a priori bounds. Again the theory hinges on Stetter's lemma which can be readily applied to prove.

Theorem 1. Let the discretization (2.1) be consistent and stable in the sense of definition K. Assume that $\Phi_h$ is continuous in $B(u_h,R)$, where R is the threshold. Then the conclusions (a) (existence) and (c) (convergence) of the Theorem S hold. Furthermore:

(b') The solutions $U_h$ of (2.1) are unique in the ball $B(u_h,R)$.

It should be emphasized that the connection of Definition K with Stetter's lemma was first established in [3]. Keller [2] resorts to linearizations to prove convergence.

## 6. Practical implications

The definitions N, S, K will now be compared. It is useful to present first a simple example which gives much insight. We consider the problem

(6.1) $$du/dt = f(u), \quad 0 \leq t \leq 1, \quad u(0) = \alpha,$$

where $\alpha$ is a fixed real number and f a fixed $C_1$ real function defined for $-\infty < u < \infty$. It is assumed that (6.1) possesses a solution u(t) defined up to t = 1. If h = 1/N, N a positive integer, the problem (6.1) is discretized by Euler's rule as in (3.1). Here $U_h = (U_0, U_1, \dots, U_N)$ and $D_h = X_h, Y_h$ are spaces of (N+1)-vectors. The norms and the elements $u_h$ are chosen as in the example of Section 3. The discretization is consistent of the first order ( u is $C_2$ ) and the existence of a unique solution $U_h$ is evident.

(A) Following most textbooks, we first make the hypothesis that f is globally Lipschitz-continuous, i.e. that there exists a constant L such that for v, w, real

(6.2) $$|f(v)-f(w)| \leq L\ |v-w|.$$

This is a very restrictive hypothesis that not many smooth nonlinear functions f would satisfy.

In order to study stability, we take vectors $V_h$ and $W_h$, define $F_h = \Phi_h(V_h)$, $G_h = \Phi_h(W_h)$ i.e.

(6.3a) $$F_0 = V_0-\alpha, \quad F_{n+1} = (V_{n+1}-V_n)/h - f(V_n), \quad n = 0, \dots, N-1,$$

(6.3b) $$G_0 = W_0-\alpha, \quad G_{n+1} = (W_{n+1}-W_n)/h - f(W_n), \quad n = 0, \dots, N-1,$$

and try to bound $|V_n-W_n|$ in terms of $|F_n-G_n|$, uniformly in h. This is easily done. Subtraction in (6.3) leads to

(6.4) $$V_{n+1}-W_{n+1} = V_n-W_n + h\ (f(V_n)-f(W_n)) + h\ (F_{n+1}-G_{n+1}),$$

n = 0, 1, ... , N-1, and use of (6.2) yields

(6.5) $$|V_{n+1}-W_{n+1}| \leq (1+hL)\ |V_n-W_n| + h\ |F_{n+1}-G_{n+1}|.$$

A well-known induction argument, using induction in n, shows that

(6.6) $$\|V_h-W_h\| = \max_n |V_n-W_n| \leq e^L\ (|F_0-G_0| + \textstyle\sum_{1\leq n\leq N} h|F_n-G_n|)$$
$$= e^L\ \|\Phi_h(V_h)-\Phi_h(W_h)\|\ .$$

This proves stability in the sense of N, and a fortiori in the weaker senses S and K. (In S and K the thresholds can be respectively chosen $r = \infty$ or $R = \infty$, so that the threshold condition becomes void.)

(B) We now suppress the restrictive assumption (6.2) introduced in (A). In this general framework the Euler discretiztion (3.1) is not, in general, N-stable, as shown by the counterexample $f(u) = u^2$, studied above. (The argument leading to (6.6) breaks down, one cannot proceed from (6.4) to (6.5).) However it is still possible to show that (3.1) is K-stable and hence convergent. Let R be a positive number and consider the union M = M(R) of the intervals [u(t)-R, u(t)+R] for t varying, $0 \leq t \leq 1$. We can think of values not in M as being far from the values taken by the solution u and therefore irrelevant to the problem. The properties of f outside M should not affect the convergence or otherwise of the discretization. Now the smoothness of f implies that (6.2) must be valid, with a suitable constant $L = L(R)$, for v and w in M. To prove K-stability with threshold R, only vectors $V_h$, $W_h$ such that

$$\|V_h-u_h\| = \max_n|V_n-u(nh)| < R, \qquad \|W_h-u_h\| = \max_n|W_n-u(nh)| < R$$

must be considered. But this threshold condition implies that $V_n$ and $W_n$ lie in M, where (6.2) holds. Then it is again possible to proceed from (6.4) to (6.5) and (6.6) follows, proving K-stability and hence convergence via Theorem 1. Note that we have not needed any a priori bounds on $U_n$.

(C) The literature contains several ad hoc tricks to prove the convergence of (3.1) in cases where f is not globally Lipschitz-continuous and therefore N-stability does not hold. Some of these tricks have been reviewed in [3]. Note that our Theorem 1 easily bypasses the need for any ad hoc considerations and that proving stability in the sense of Keller for smooth f is not more difficult than proving N-stability for globally Lipschitz f.

(D) Until now the function f has been assumed to be defined for all real u. Very often (i.e. if f involves logarithms or square roots) f is only defined in a suitable neighbourhood $M = M(R)$ of the solution. In this case the existence of $U_h$ cannot be taken for granted: a value n can be reached in the time-stepping so that $U_{n+1}$ is outside M and then $U_{n+2}$ cannot be computed. The technique in (B) shows that for h small the discrete equations possess a solution.

(E) Stetter's definition. Let us now return to the situation in (B) (f smooth and defined for all real u, but not globally Lipschitz-continuous) and try to prove that (3.1) is stable in the sense of Stetter. We have now to show that a stability bound like (6.6) holds for $V_h$ and $W_h$ close to $u_h$, close meaning that (4.1) holds. As distinct from the situation in (B), where we had information relating directly to $V_h$,$W_h$ we now know that $F_n$, $G_n$ are small. Unfortunately it is not apparent how this information could be used to go from (6.4) to (6.5). Therefore Stetter's stability does not appear to be easily checked in practice. In fact we are not aware of many instances where S-stability has been proved from first principles. However it is fair to say that the more advanced linearization theory of Stetter [7] can be employed to prove the S-stability and convergence of (3.1).

Thus, while the conclusions in Theorem S and 1 are equal in strength, the practical application of the framework in Section 5 is easier than that of the material in Section 4. On these grounds we believe that Keller's definition should be favoured.

(F) Uniqueness. There is a final issue to be commented upon. Theorem S guarantees the uniqueness of $U_h$ in the whole domain $D_h$, whereas Theorem 1 only yields uniqueness in the ball $B(u_h,R)$. Many nonlinear problems, mainly in boundary value settings, posses several isolated solutions $u_i$ and correspondingly their consistent discretizations are expected to have, for each h, several solutions $U_{ih}$. Such discretizations cannot be S-stable: if they were they would contradict Theorem S, conclusion (b).

See [3], [4] for further analysis of the relation between definitions S and K.

## 7. Restricted stability

Just as the naive definition N was too restrictive to cater for most realistic ODE problems, the improved definitions S and K are too narrow to accommodate many PDE problems. This statement does not question the merits of the definitions by Stetter and Keller, as these authors were mainly concerned with ODEs.

Cases of convergent nonlinear PDE discretizations which are not stable in the senses S or K have been presented in [1], [8] and some insight into the situation will be obtained from the example in the next Section. For the time being, we anticipate that in order to cope successfully with nonlinear PDEs is essential to allow thresholds which depend on h. Several stability definitions using such thresholds have been suggested in the literature, see [3]. We introduce the following definition.

Definition. Suppose that, for each h in H, $R_h$ is a value $0 < R_h \leq \infty$. Then the discretization (2.1) is said to be stable (restricted to the thresholds $R_h$) if there exist positive constants $h_o$ and S (the stability constant) such that for h in H, $h \leq h_o$, the open ball $B(u_h, R_h)$ is contained in the domain $D_h$ and for any $V_h$ and $W_h$ in that ball the bound (*) holds.

Clearly Keller's definition is recovered in cases where the thresholds can be chosen independent of h. Again, in linear cases this definition reduces to the standard one. Furthermore with the present definition a smooth discretization is stable if and only if its linearization around $u_h$ is stable [4]. It is also possible to derive equivalence theorems.

The final theorem is provedas Theorem S.

Main Theorem. Assume that (2.1) is consistent and stable with thresholds $R_h$. If $\Phi_h$ is continuous in $B(u_h, R_h)$ and $\|l_h\| = o(R_h)$ as $h \to 0$, then:

(i) For h small the discrete equations (2.1) possess a solution in $B(u_h, R_h)$.

(ii) This solution is unique in the ball.

(iii) As $h \to 0$, the solutions in (i) converge. The order of convergence is not smaller than the order of consistency.

## 8. An application

A nontrivial example of the application of the theory just introduced is now presented which is useful in gaining insight into the need for h-dependent thresholds.

We consider the following 1-periodic initial value problem for a nonlinear Schroedinger equation [3]:

(8.1) $$iu_t + u_{xx} + f(u) = 0, \qquad -\infty < x < \infty, \qquad 0 \leq t \leq T < \infty,$$

(8.2) $$u(x,t) = u(x+1,t), \qquad -\infty < x < \infty, \qquad 0 \leq t \leq T,$$

(8.3) $$u(x,0) = u_o(x), \qquad -\infty < x < \infty,$$

where $i^2 = -1$ and u is complex. The initial datum $u_o$ is 1-periodic and the (nonlinear) function f is smooth and defined for all complex u, but not globally Lipschitz-continuous. (In applications of physical interest f is often a power.)

If J is a positive integer, we introduce the mesh size $\Delta x = 1/J$ and the grid $x_j$, j integer. With c a fixed positive constant, we define the time-step $h = c\Delta x$ and the time levels $t_n = nh$, $n = 0, 1, \dots, N$, $N = T/h$. Now (8.1)-(8.3) is discretized by means of standard spatial central differences together with a Crank-Nicolson time-stepping. Explicitly:

(8.4) $$U_j^0 - u_o(x_j) = 0, \quad j = 1, \dots, J;$$

(8.5) $$ih^{-1}(U_j^{n+1}-U_j^n) + \tfrac{1}{2}(\Delta x)^{-2}\delta^2(U_j^{n+1}+U_j^n) + \tfrac{1}{2}(f(U_j^{n+1})+f(U_j^n)) = 0,$$

$n = 0, 1, \dots, N-1$; $j = 1, \dots, J$. In (8.5) use must be made of the periodicity in the computation of $\delta^2 U_j^n$, $j = 1, J$.

It is convenient to collect the J complex values $U_j^n$, $j = 1, \dots, J$ in a vector $\underline{U}^n$ and analogously define vectors $\underline{u}_o$, $\underline{f}(\underline{U}^n)$. With this change in notation (8.4)-(8.5) become

(8.6) $$\underline{U}^0 - \underline{u}_o = \underline{0},$$

(8.7) $$h^{-1}A_h\underline{U}^{n+1} - h^{-1}B_h\underline{U}^n + \tfrac{1}{2}(\underline{f}(\underline{U}^{n+1})+\underline{f}(\underline{U}^n)) = \underline{0}, \quad 0 \le n \le N-1,$$

where $A_h$, $B_h$ are suitable complex JxJ matrices. This is a discretization of the form (2.1), where $D_h = X_h$ and $Y_h$ are spaces of block vectors with N+1 blocks, each block being in turn a vector with J complex entries. We employ the norms (cf. Section 3)

(8.8) $$\|V_h\| = \max_n \|\underline{V}^n\|_2, \quad \text{if} \quad V_h = (\underline{V}^0, \underline{V}^1, \dots, \underline{V}^N) \in X_h;$$

(8.9) $$\|F_h\| = \|\underline{F}^0\|_2 + \sum_{1\le n\le N} h\,\|\underline{F}^n\|_2, \text{ if } F_h = (\underline{F}^0, \underline{F}^1, \dots, \underline{F}^N) \in Y_h.$$

In (8.8)-(8.9) $\|\cdot\|_2$ denotes the usual discrete $L_2$-norm.

When the theoretical element $u_h$ is taken to be the obvious restriction of u to the grid, consistency of order $p = 2$ is easily proved via Taylor expansions.

The stability of (8.6)-(8.7) will be studied next. Let $V_h$ and $W_h$ denote two elements in $X_h$ and, as in Section 6, set $F_h = \Phi_h(V_h)$, $G_h = \Phi_h(W_h)$. Then, for $n = 0, \dots, N-1$ we can write

$$\underline{F}^{n+1} = h^{-1}A_h\underline{V}^{n+1} - h^{-1}B_h\underline{V}^n + \tfrac{1}{2}(\underline{f}(\underline{V}^{n+1})+\underline{f}(\underline{V}^n)),$$

$$\underline{G}^{n+1} = h^{-1}A_h\underline{W}^{n+1} - h^{-1}B_h\underline{W}^n + \tfrac{1}{2}(\underline{f}(\underline{W}^{n+1})+\underline{f}(\underline{W}^n)),$$

As in Section 3, these are subtracted to get

(8.10) $$A_h(\underline{V}^{n+1}-\underline{W}^{n+1}) = B_h(\underline{V}^n-\underline{W}^n) - (h/2)\,(\underline{f}(\underline{V}^{n+1})-\underline{f}(\underline{W}^{n+1})) - (h/2)\,(\underline{f}(\underline{V}^n)-\underline{f}(\underline{W}^n)) + h\,(\underline{F}^{n+1}-\underline{G}^{n+1}).$$

Now the relations

$$\|A_h^{-1}\|_2 \le 1, \qquad \|A_h^{-1}B_h\|_2 \le 1$$

are easily derived either by an energy estimate or by a von Neumann analysis. Thus:

$$(8.11)\qquad \|\underline{V}^{n+1}-\underline{W}^{n+1}\|_2 \le \|\underline{V}^n-\underline{W}^n\|_2 + (h/2)\,\|\underline{f}(\underline{V}^{n+1})-\underline{f}(\underline{W}^{n+1})\|_2 + (h/2)\,\|\underline{f}(\underline{V}^n)-\underline{f}(\underline{W}^n)\|_2 + h\,\|\underline{F}^{n+1}-\underline{G}^{n+1}\|_2.$$

In order to apply the standard induction argument, we would like to have bounds

$$(8.12)\qquad \|\underline{f}(\underline{V}^{n+1})-\underline{f}(\underline{W}^{n+1})\|_2 \le L\,\|\underline{V}^{n+1}-\underline{W}^{n+1}\|_2$$

$$\|\underline{f}(\underline{V}^n)-\underline{f}(\underline{W}^n)\|_2 \le L\,\|\underline{V}^n-\underline{W}^n\|_2.$$

A Keller threshold is of no use: even if the attention is restricted to elements $V_h$, $W_h$ with

$$\|V_h-u_h\| = \max_n\|\underline{V}^n-\underline{u}^n\|_2 < R, \quad \|W_h-u_h\| = \max_n\|\underline{W}^n-\underline{u}^n\|_2 < R,$$

the individual grid values $V_j^n$, $W_j^n$ may increase like $h^{-1/2}$ and it is not possible to bound $f(V_j^n)-f(W_j^n)$, as f is not globally Lipschitz-continuous. On the other hand, it is obvious that (8.12) holds if we introduce an h-dependent threshold condition as follows

$$\|V_h-u_h\| = \max_n\|\underline{V}^n-\underline{W}^n\|_2 < bh^{1/2}, \quad \|W_h-u_h\| = \max_n\|\underline{W}^n-\underline{u}^n\|_2 < bh^{1/2}.$$

When (8.12) holds, the standard induction argument leads to the stability bound (*). Thus we have proved stability restricted to the thresholds $bh^{1/2}$, b any positive constant. Since the order of consistency is 2, the condition $\|l_h\| = o(R_h)$ is fulfilled and our main Theorem shows that the discretization is convergent of the second order, i.e. (8.7) possesses a unique solution $\underline{U}^n$ near $\underline{u}^n$ (provided that h is small enough) and furthermore

$$\max_n \|U^n-u^n\|_2 = O(h^2).$$

Before we conclude the paper we would like to mention that even though our examples have dealt with finite differences, finite elements can be treated in an analogous way [4]. Also the restriction to nonlinearities of the form f(u) was dictated by simplicity in the exposition; more general nonlinear terms, involving derivatives of u, can also be studied in a similar manner.

## References


[1] de Frutos, J, Palencia, C. and Sanz-Serna, J.M., Stability and convergence of numerical schemes for nonlinear Dirac systems, submitted.

[2] Keller, H. B., Approximation methods for nonlinear problems with application to two-point boundary value problems, Math. Comput. 29, (1975), 464-474.

[3] López-Marcos, J. C. and Sanz-Serna, J. M., Stability and convergence in numerical analysis II: The definition of stability in nonlinear problems, unpublished manuscript.

[4] López-Marcos, J.C. and Sanz-Serna, J. M., Stability and convergence in numerical analysis III: Linear investigation of nonlinear stability, submitted.

[5] Sanz-Serna, J. M., Stability and convergence in numerical analysis I: Linear problems, a simple comprehensive account, in Nonlinear Differential Equations, (J. K. Hale and P. Martínez Amores, eds.), Pitman, Boston, 1985, 64-113.

[6] Sanz-Serna, J. M. and Verwer, J. G., A study of the recursion $y_{n+1} = y_n + y_m^m$, J. Math. Anal. Appl. 116, (1986), 456-464.

[7] Stetter, H. J., Analysis of Discretization Methods for Ordinary Differential Equations, Berlin, Springer 1973.

[8] Vadillo, F. and Sanz-Serna, J. M., Studies in numerical nonlinear instability II, a new look at $u_t + uu_x = 0$, J. Comput. Phys. 66, (1986), 225-238.