

Ecole polytechnique fédérale de Zurich Politecnico federale di Zurigo Swiss Federal Institute of Technology Zurich

On the Definition of Nonlinear Stability for Numerical Methods

Magnus Pirovino

Research Report No. 91-07 October 1991

Seminar für Angewandte Mathematik Eidgenössische Technische Hochschule CH-8092 Zürich Switzerland

On the Definition of Nonlinear Stability for Numerical Methods

Magnus Pirovino Seminar für Angewandte Mathematik Eidgenössische Technische Hochschule CH-8092 Zürich Switzerland

Research Report No. 91-07 October 1991

Abstract

In this paper a new definition of nonlinear stability for the general nonlinear problem F(u) = 0 and the corresponding family of discretized problems $F_h(u_h) = 0$ is given. The notion of nonlinear stability introduced by Keller [3] and later by Lopéz-Marcos and Sanz-Serna [4] have the disadvantage that the Lipschitz constant of the derivative of $F_h(u_h)$ has to be known which, in many applications, is not practicable. The modification proposed here allows us to use linearized stability in a ball containing the solution u_h to get nonlinear stability. The usual result remains true: nonlinear stability together with consistency implies convergence.

Keywords: nonlinear stability, linearized stability Mathematics Subject Classification: 65H10, 65L20, 65M12, 65N12 **1. Introduction and preliminaries.** Let us consider the general nonlinear problem

$$F(u) = 0 \tag{1}$$

where $F: D_F \subset X \to Y$ is a differentiable mapping between the Banach spaces X, Y with domain D_F . We only consider isolated solutions u^* of (1), i.e. the Fréchet derivative $d_{u^*}F: X \to Y$ shall be boundedly invertible.

A numerical method may be applied in order to solve equation (1). This leads in general to a family of equations

$$F_h(u_h) = 0 \tag{2}$$

where the $F_h: D_{F_h} \subset X^h \to Y^h$ are mappings between finite dimensional Banach spaces. Here the domain D_{F_h} is an open set and $F_h(\cdot)$ is continuous on D_{F_h} . The subscript h indicates the dependence of the discretization on a small parameter such as the mesh size. Let us assume that h takes values in a set H of positive parameters with H = 0 and $\sup H = h_0 < \infty$. In order to define convergence of the approximating solution u_h^* of (2) to u^* , we require that there exist linear mappings

$$R_X^h: D_{R_Y^h} \subset X \to X^h$$
 and $R_Y^h: D_{R_Y^h} \subset Y \to Y^h$

obeying the condition

$$\lim_{h \to 0} \|R_{\nu}^{h}(x)\| = \|x\|, \quad x \in D_{R_{\nu}^{h}}, \, \nu \in \{X, Y\}^{2}.$$

Here the $D_{R_{\nu}^{h}}$ are required to be dense in ν . For example, it may be reasonable to work with a function space X, say $L^{2}(\Omega)$, and with a restriction represented by a grid function U(i, j) = u(ih, jh), $(i, j) \in \Delta \subset \mathbb{Z}^{2}$. The domain of this restriction is a set containing all continuous functions $C(\Omega)$. But it is impossible to define such a restriction for all $u \in L^{2}(\Omega)$. Let us use henceforth the notation

$$[z]_h \equiv R^h_\nu(z) \,, \quad \nu \in \{X, Y\} \,.$$

Convergence of the approximating solution u_h^* to u^* is now meant in the sense that

$$\lim_{h \to 0} \|u_h^* - [u^*]_h\| = 0$$

²We suppress indices of norms: $||z||_{\nu} \equiv ||z||, z \in \nu, \nu \in \{X, Y, X^h, Y^h\}.$

Also the definition of consistency makes no difficulties. We introduce the local discretization error at a point $u \in D_{R_{r}^{h}}$:

$$\tau_h(u) \equiv F_h([u]_h) - [F(u)]_h$$

Definition 0.1. The family $\{F_h(\cdot)\}$ is said to be consistent of order p with $F(\cdot)$ at u iff for some constant M, independent of h,

$$\|\tau_h(u)\| \le Mh^p \,. \tag{3}$$

A further definition will be useful because our original problem (1) as well as the discretized problems (2) might have more than one isolated solution. Hence, the original solutions have to be related to the right discretized ones.

Definition 0.2. We call the family $\{F_h(\cdot)\}$ a proper discretization of $F(\cdot)$ for the solution u^* iff there exists a radius $\rho > 0$, independent of h, such that the following two conditions are satisfied for all relevant values of h:

- (i) There are solutions u_h^* of (2) such that $B_{\rho}(u_h^*) \subset D_{F_h}$ and $F_h(\cdot)$ restricted on $B_{\rho}(u_h^*)$ is one-to-one.
- (*ii*) $[u^*]_h \in B_\rho(u_h^*)$

2. Classical definitions of stability. There are several attempts to define nonlinear stability in the literature, see [8], [3], [4] or [5]. For a summary of these different notions see [4]. They all require the existence of a stability constant S, independent of h, such that a stability inequality

$$||u_h - v_h|| \le S ||F_h(u_h) - F_h(v_h)||$$
(4)

holds for u_h, v_h in some subset $D_h \subset D_{F_h}$. Let us call D_h the domain of stability. It is exactly this domain of stability which distinguishes all the notions of nonlinear stability mentioned above. Clearly, if we choose D_h too large it may be difficult to find a suitable stability constant S satisfying (4). On the other hand, if we choose D_h too small, e.g. $\operatorname{diam}(D_h) = O(h^q)$, it may occur that we must impose a condition $||[u^*]_h - u_h^*|| = O(h^q)$, which itself guarantees convergence and our stability analysis becomes meaningless.

Let us select H. B. Keller's definition of stability. He chooses $D_h = B_{\rho}(u_h^*)^2$, the open ball of radius ρ about u_h^* . Here the stability threshold

²In fact he chooses $D_h = B_{\rho}([u^*]_h)$, but this difference is only of technical nature.

 ρ must be positive and independent of h. The problem now is how to determine ρ and S. Let us assume that the linearized problem at u_h^* is stable with stability constant S_0 independent of h, which is equivalent to the condition

$$\|[d_{u_h^*}F_h]^{-1}\| \le S_0$$

If, in addition, the Fréchet derivative $d_{u_h^*}F_h$ has a Lipschitz constant K_{ρ_0} independent of h in some open ball $B_{\rho_0}(u_h^*)$, i.e.

$$\|d_{u_h}F_h - d_{v_h}F_h\| \le K_{\rho_0}\|u_h - v_h\|, \quad u_h, v_h \in B_{\rho_0}(u_h^*),$$
(5)

then it can be shown (see [5]) that we may choose an $S > S_0$ arbitrarily to get the condition

$$\rho = \min\{\rho_0, (S_0^{-1} - S^{-1})/K_{\rho_0}\}.$$

We have transformed the problem of finding ρ and S into another problem, namely finding S_0 and K_{ρ_0} . It may be easy to handle the linearized problem, but to determine K_{ρ_0} means in practice that we must estimate the norm of the second derivative $d_{u_h}^2 F_h$, which for a wide series of applications does not even exist.

Nevertheless, this definition of stability might be useful. An important lemma of Stetter gives us the connection to convergence.

Lemma 0.3 ([8]). Let X^h , Y^h be finite dimensional Banach spaces and let $F_h : B_\rho(u_h^*) \subset X^h \to Y^h$ be a continuous mapping with $F_h(u_h^*) = 0$. If the bound (4) holds on $D_h = B_\rho(u_h^*)$ for some constant S, independent of h, then the inverse $F_h^{-1} : F_h(B_\rho(u_h^*)) \to B_\rho(u_h^*)$ is Lipschitz continuous with Lipschitz constant S and

$$B_{\rho/S}(0) \subset F_h(B_\rho(u_h^*)) \,.$$

Now we may state a convergence result.

Theorem 0.4. Let the family $\{F_h(\cdot)\}$ be consistent of order p > 0 with $F(\cdot)$ at u^* and stable on the stability domain $B_\rho(u_h^*)$ with stability constant S. If, in addition, it is a proper discretization for $F(\cdot)$ at u^* (i.e. $[u^*]_h \in B_\rho(u_h^*)$), then

$$||u_h^* - [u^*]_h|| \le MSh^p, \quad h \le \left(\frac{\rho}{MS}\right)^{1/p}.$$

The order of convergence is not smaller than the order of consistency.

Proof. Since by (3)

$$\|\tau_h(u^*)\| = \|F_h([u^*]_h) - [F(u^*)]_h\| = \|F_h([u^*]_h)\| \le Mh^p,$$

we see that $h \leq (\frac{\rho}{MS})^{1/p}$ implies $F_h([u^*]_h) \in B_{\rho/S}(0)$. Thus, Lemma 0.3 applies and the result follows by the stability inequality (4).

From the proof of Theorem 0.4 we learn the following: the inverse F_h^{-1} is not required to be Lipschitz continuous on the whole range $F_h(B_\rho(u_h^*))$ to imply convergence for the numerical method. It would suffice for F_h^{-1} to be Lipschitz continuous on $B_{\rho/S}(0)$ only. This is exactly the point where we want to modify the definition of nonlinear stability.

3. Stability by range. Since we only consider isolated solutions u^* of (1), the Fréchet derivative $d_{u^*}F: X \to Y$ is boundedly invertible. If $F(\cdot)$ is continuously differentiable², then, by the inverse mapping theorem (see e.g. [7]), there exists a ball $B_{\rho}(u^*)$ such that $F: B_{\rho}(u^*) \to F(B_{\rho}(u^*))$ is an isomorphism. This allows us to define an inverse mapping F^{-1} on $F(B_{\rho}(u^*))$. It is the inverse mapping we want to compute, therefore we think stability in some sense has to be defined for this inverse mapping.

Definition 0.5. Let $F_h(u_h^*) = 0$ and let $R_h \subset Y^h$ be an open set with $0 \in R_h$. The family $\{F_h(\cdot)\}$ is said to be stable on the stability range R_h with stability constant S, independent of h, iff for all relevant h > 0

- (i) $F_h^{-1}: R_h \to X^h$ can be uniquely defined with $u_h^* \in F_h^{-1}(R_h)$,
- (ii) $F_h^{-1}(\cdot)$ is Lipschitz continuous on R_h with Lipschitz constant S.

Again, we get the convergence result.

Theorem 0.6. Let the family $\{F_h(\cdot)\}$ be consistent of order p > 0 with $F(\cdot)$ at u^* and stable on the stability range $B_r(0)$ with stability constant S. If, in addition, it is a proper discretization for $F(\cdot)$ at u^* with $\rho \geq Sr$, then

$$||u_h^* - [u^*]_h|| \le MSh^p, \quad h \le \left(\frac{r}{M}\right)^{1/p}$$

²Recall that in general $F(\cdot)$ need not be continuously differentiable since it is only defined on a subset D_F , but we will see that this requirement (at least the continuity) is essential for the family $\{F_h(\cdot)\}$.

Proof. As in the proof of Theorem 0.4 inequality (3) yields

$$\|\tau_h(u^*)\| = \|F_h([u^*]_h)\| \le Mh^p$$
.

Thus, $\tau_h(u^*)$ is in $B_r(0)$ if only $h \leq (r/M)^{1/p}$. $F_h^{-1}(\cdot)$ is Lipschitz continuous on $B_r(0)$ with Lipschitz constant S. This implies $F_h^{-1}(B_r(0)) \subset B_{Sr}(u_h^*) \subset B_\rho(u_h^*)$. Since the discretization is proper we conclude $[u^*]_h \in F_h^{-1}(B_r(0))$ and $[u^*]_h = F_h^{-1}(\tau_h(u^*))$. The result follows now by the Lipschitz continuity of $F_h^{-1}(\cdot)$.

4. Nonlinear stability by linearized stability. We want to have a practicable tool that allows us to verify the assumptions of Theorem 0.6. It would be helpful if we could work with the linearized problems

$$L_{h}(u_{n}) \equiv F_{h}(v_{h}) + d_{v_{h}}F_{h}(u_{h} - v_{h}) = 0$$
(6)

for suitable $v_h \in X^h$. Clearly, the linear family $\{L_h(\cdot)\}$ is stable in all above senses iff there is a constant S, independent of h, such that

$$\|[d_{v_h}F_h]^{-1}\| \le S.$$
(7)

We will now state our main theorem.

Theorem 0.7. Let the family $\{F_h(\cdot)\}$ satisfy the following conditions for all relevant values of h:

- (a) $\{F_h(\cdot)\}$ is a proper discretization for $F(\cdot)$ at u^* on $B_\rho(u_h^*)$.
- (b) $F_h(\cdot)$ is continuously differentiable on $B_{\rho}(u_h^*)$.
- (c) The linearized families $\{L_h(\cdot)\}$ of (6) are stable with stability constant S uniformly for $v_h \in B_{\rho}(u^*)$.
- (d) The constants ρ and S are independent of h.

Then the family $\{F_h(\cdot)\}$ is stable on the stability range $B_{\rho/S}(0)$ with stability constant S.

Corrolary 0.8. Let the assumptions of Theorem 0.7 hold. If, in addition, $\{F_h(\cdot)\}$ is consistent of order p with $F(\cdot)$ at u^* then

$$||u_h^* - [u^*]_h|| = O(h^p).$$

The order of convergence is not smaller than the order of consistency.

To proof our main theorem, it suffices to proof the following, more general lemma.

Lemma 0.9. Let X, Y be arbitrary Banach spaces and let the mapping $f: B_{\rho}(u^*) \subset X \to Y$ satisfy the conditions

- (a) $f(u^*) = 0$
- (b) $f(\cdot)$ is continuously differentiable and one-to-one on $B_{\rho}(u^*)$.
- (c) $d_u f: X \to Y$ is a linear isomorphism and $||[d_u f]^{-1}|| \leq S$, uniformly for $u \in B_{\rho}(u^*)$.

Then the following holds.

- (i) $f : B_{\rho}(u^*) \to V \equiv f(B_{\rho}(u^*))$ is a diffeomorphism. In particular $f^{-1}(\cdot)$ can be uniquely defined on V.
- (*ii*) $B_{\rho/S}(0) \subset V$
- (iii) f^{-1} is Lipschitz continuous with Lipschitz constant S on $B_{\rho/S}(0)$.

Proof. Because $f(\cdot)$ on $B_{\rho}(u^*)$ is one-to-one, we get (i) from the inverse mapping theorem. For the rest we may assume that $Y \setminus V$ is nonempty, for otherwise (ii) is always satisfied and (iii) follows from Taylor's formula

$$f^{-1}(y) - f^{-1}(y') = \int_0^1 d_{ty+(1-t)y'} f^{-1}(y-y') dt$$
(8)

(see e.g. [2, Theorem 8.14.3]) together with hypothesis (c). Thus, let $y \in Y \setminus V$ and $r_1 = ||y|| + 1$, i.e. $B_{r_1}(0) \setminus V \neq \emptyset$. By the inverse mapping theorem we know that there is a $0 < r_0 < r_1$ with $B_{r_0}(0) \subset V$. We define

$$r = \sup\{s \in [r_0, r_1] : B_s(0) \subset V\}.$$
(9)

Thus, $B_r(0)$ is the biggest ball around zero, which lies entirely in V. We may now choose $\epsilon > 0$ arbitrarily. $B_{r+\epsilon}(0)$ is no more a subset of V. We may therefore choose $y \in (Y \setminus V) \cap B_{r+\epsilon}(0)$. Let

$$T_y = \{ty : 0 \le t < 1\}$$
 and $t^* = \sup\{\tau : T_{\tau y} \subset V\}.$

If we could walk on T_y beginning at zero, we would leave V for the first time at the point $y^* = t^*y$. Clearly, $y^* \notin V$ by the definition of t^* , but $y^* \in V = \overline{V} \setminus V$ and by (9) we have

$$r \le \|y^*\| \le r + \epsilon \,. \tag{10}$$

If we choose a sequence $\{y_k\} \subset T_{y^*} \subset V$ with limit y^* , then the sequence $\{v_k = f^{-1}(y_k)\}$ is well defined in $B_{\rho}(u^*)$. Since T_{y^*} is convex, we may apply Taylor's formula to get

$$v_k - v_m = \int_0^1 d_{ty_k + (1-t)y_m} f^{-1}(y_k - y_m) dt, \quad k, m \ge 1.$$

and together with hypothesis (c)

$$||v_k - v_m|| \le S ||y_k - y_m||, \quad k, m \ge 1.$$
 (11)

By this, both $\{y_k\}$ and $\{v_k\}$ are Cauchy sequences. Let $v^* = \lim_k v_k$. It is impossible for v^* to be an element of $B_{\rho}(u^*)$, for else, by the continuity of $f(\cdot)$, $y^* = f(x^*) \notin V$, which is a contradiction. Hence, $v^* \in \underline{B}_{\rho}(u^*)$, i.e. $\|v^* - u^*\| = \rho$. We may choose $y_m = 0$ in (11) and pass to the limit; then we have , recalling (10),

$$\rho = \|v^* - u^*\| \le S \|y^*\| \le S(r + \epsilon) \,.$$

Since ϵ can be chosen arbitrarily, (*ii*) is proven. $B_r(0)$ is convex; thus again, we can apply Taylor's formula and verify (*iii*) by hypothesis (c). This completes our proof.

5. Numerical examples. Let us look at the nonlinear problem

$$\begin{aligned} -u_{xx} + g(u) - f &= 0, \quad x \in [0, \pi] \\ u(0) &= u(\pi) &= 0, \end{aligned}$$
 (12)

where $g: \mathbb{R} \to \mathbb{R}$ is a contraction (i.e. $g(\cdot)$ is Lipschitz continuous with Lipschitz constant L < 1) and $f \in L^2(0, \pi)$. Here F(u) = Au + g(u) - f, where A is an unbounded self-adjoint linear operator with dense domain D_F in $L^2(0,\pi)$ and spectrum $\sigma(A) = \{k^2 : k \ge 1\}$. Problem (12) has a unique solution $u^* \in L^2(0,\pi)$ (see [1]). If f is continuous, i.e. $f \in C^0(0,\pi)$, then $u^* \in C^2(0,\pi)$. In the following we will always assume that f is continuous. Let us discretize (12) with central differences on the grid $\{ih : i = 0, \ldots, n+1, h = \pi/(n+1)\}$. We get a system

$$F_h(u_h) = A_h u_h + g(u_h) - f_h = 0$$
(13)

with $u_h, f_h \in \mathbb{R}^n$ and

$$A_{h} = \frac{1}{h^{2}} \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Let us use the notation $U \equiv u_h$. Then by g(U) we denote the vector $[g(U_1), \dots, g(U_n)]^T \in \mathbb{R}^n$. The boundary conditions $U_0 = U_{n+1} = 0$ imply that we can work with the reduced grid $\{ih : i = 1, \dots, n\}$. The restriction from $L^2(0, \pi)$ on this reduced grid is surely defined on the dense set $C^0(0, \pi)$. We write $f_h = [f]_h \equiv [f(h), f(2h), \dots, f(nh)]^T$. Let us use the discrete L^2 -norm on grid functions

$$||U|| = (\sum_{i=1}^{n} U_i^2 h)^{1/2}$$

We first show that the family $\{F_h(\cdot)\}$ is a proper discretization for the solution u^* of (12) with $\rho = \infty$. As one can easily check, this can be done by showing that the system (13) has a unique solution $U^* \in \mathbb{R}^n$. It is well-known that A_h has eigenvalues $\lambda_k = 4\sin^2(kh/2)/h^2$, $k = 1, \ldots, n$. This implies

$$||A_h^{-1}|| = \frac{h^2}{4\sin^2(h/2)}.$$

From now on we only admit values of h so small that a condition

$$L\frac{h^2}{4\sin^2(h/2)} \le L' < 1$$

can be satisfied. System (13) has a unique solution iff the function G_h : $\mathbb{R}^n \to \mathbb{R}^n$ defined by

$$G_h(U) = -A_h^{-1}(g(U) - f_h)$$

has a single fixed-point. Since $g(\cdot)$ is a contraction, we have

$$||G_h(U) - G_h(V)|| \le ||A_h^{-1}|| ||g(U) - g(V)|| \le L ||A_h^{-1}|| ||U - V|| \le L' ||U - V||$$

hence $G_h(\cdot)$ is a contraction, too. Now the result follows by the contraction theorem.

Next we verify that the linearizations of (13) are uniformly stable with stability constant

$$S = \frac{L'}{L(1-L')}$$

provided that $g(\cdot)$ is continuously differentiable. From (13) we derive

$$d_U F_h = A_h + \operatorname{diag}(g'(U)) = A_h [Id + A_h^{-1} \operatorname{diag}(g'(U))].$$

Thus, by the Banach lemma [6, p. 333]

$$\|d_U F_h^{-1}\| \le \frac{\|A_h^{-1}\|}{1 - \|g'(U)\|_{\infty} \|A_h^{-1}\|} \le \frac{L'}{L(1 - L')}.$$

Now all assumptions of Theorem 0.7 hold for $\rho = \infty$. We conclude that the family $\{F_h(\cdot)\}$ is stable on the stability range \mathbb{R}^n with stability constant S = L'/(L(1-L')).

It is a well-known property of the central difference discretization that it is consistent of order p = 1 provided the solutions u^* lie in $C^3(0, \pi)$, which is true if both $f(\cdot)$ and $g(\cdot)$ are continuously differentiable². The solutions U^* of (13) therefore converge to u^* with order not smaller than p = 1.

We mention that we have never imposed a condition on the Lipschitz continuity of the derivative $d_U F_h$ as in (5). This is the main advantage of our analysis.

We conclude this paper with a remark on partial differential equations. Exactly the same kind of analysis as we have made for equation (12) can be done on related two-dimensional problems such as

$$u_{tt} - u_{xx} + g(u) - f = 0$$

 $u(0,t) = u(\pi,t) = 0$
 $u \ 2\pi$ -periodic.

For more details on these problems see [1].

²If $f(\cdot)$ and $g(\cdot)$ are twice continuously differentiable then $u^* \in C^4(0,\pi)$ which implies consistency of order p = 2.

References

- [1] R.A. Brawer. Numerik und Existenzsätze schwacher Lösungen nichtlinearer partieller Differentialgleichungen. PhD thesis, ETH Zürich, 1991.
- [2] J. Dieudonné. Foundations of Modern Analysis. Academic Press, New York, 1969.
- H.B. Keller. Approximation methods for nonlinear problems with application to two-point boundary value problems. *Math. Comp.*, 29:464–474, 1975.
- [4] J.C. López-Marcos and J.M. Sanz-Serna. Definition of stability for nonlinear problems. In K. Strehmel, editor, *Numerical Treatment of Differential Equations*, Leibzig, 1988. Teubner.
- [5] J.C. López-Marcos and J.M. Sanz-Serna. Stability and convergence in numerical analysis III: Linear investigation of nonlinear stability. *IMA J. Numer. Anal.*, 8:71–84, 1988.
- [6] M.Z. Nashed. Perturbations and approximations for generalized inverses and linear operator equations. In M.Z. Nashed, editor, *Generalized In*verses and Applications, New York, 1976. Academic Press.
- [7] W. Rudin. *Functional Analysis*. Tata McGraw-Hill Company Ltd., New Dehli, 1973.
- [8] H.J. Stetter. Analysis of Discretization Methods for Ordinary Differential Equations. Springer, Berlin, 1973.

Research Reports

Report No.	Authors	Title
91-07	M. Pirovino	On the Definition of Nonlinear Stability
		for Numerical Methods
91-06	Ch. Lubich,	Runge-Kutta Methods for Parabolic
	A. Ostermann	Equations and Convolution
		Quadrature
91-05	C. W. Schulz-Rinne	Classification of the Riemann Problem
		for Two-Dimensional Gas Dynamics
91-04	R. Jeltsch, J. H. Smit	Accuracy Barriers of Three Time
		Level Difference Schemes for Hyper-
		bolic Equations
91-03	I. Vecchi	Concentration-cancellation and Hardy
		spaces
91-02	R. Jeltsch, B. Pohl	Waveform Relaxation with Overlap-
		ping Splittings
91-01	R. Brawer, M. Pirovino	Das Bernoullische Potenzsummen-
	·	Problem und die Stirling-Zahlen 2. Art