NON-LINEAR INSTABILITY, THE DYNAMIC APPROACH

J. M. Sanz - Serna and F. Vadillo

IN
NUMERICAL ANALYSIS
D. F. GRIFFITHS & G. A. WATSON EDS.
RESEARCH NOTES IN MATEHEMATICS 140
LONGMAN SCIENTIFIC & TECHNICAL 1986
pp. 187-199

1. INTRODUCTION

Consider an evolutionary differential equation (ordinary or partial) $u' = A(u)$, where a prime represents differentiation with respect to the time t. In the ODE case, and for each fixed t, $u(t)$ is a d-dimensional vector and A a vector valued function of u. In the PDE case, for each fixed t, $u(t)$ is a scalar or vector valued function of the 'space variables' $x, y, \ldots$ and A an operator involving differentiation w.r. to the space variables.

Numerical methods generate approximations U_n to the true value $u(t_n)$, $t_n = nk$, which depend on the time-step parameter k. Usually the first stage in the analysis of a numerical method consists of the investigation of its convergence: does U_n tend to $u(t_n)$ as k tends to 0, n tends to ∞, $nk = t_n$ constant? It is in this analysis that the concepts of Lax-stability, Dahlquist-stability, etc... are of the outmost importance (see [11] for a survey).

Assume that the convergence of a numerical method has been established; it is still possible that for a given choice of k, or even for any such a choice, the qualitative behaviour of the numerical sequence $U_0, U_1, \ldots, U_n, \ldots$ be completely different from that of the theoretical sequence $u(t_0), u(t_1), \ldots, u(t_n), \ldots$ This discrepancy which refers to n tending to ∞, k fixed cannot be ruled out by the convergence requirement, as this involves a different limit process (namely k tending to 0).

In linear, constant coefficient problems it is often possible to derive closed expressions for U_n. These can be used to derive 'stability' conditions on k (or on k and the space grid-sizes in the PDE case) which guarantee that U_n possesses the right qualitative behaviour as n increases.

In nonlinear problems it is customary to linearize the discrete equations, freeze the resulting coefficients and then demand that k satisfies all the 'stability' conditions of the resulting linear, constant coefficient problems. It has been known for a long time that this approach may fail: in 1959, Phillips [10] showed an example where U_n grew with n in spite of the fact that no growth was predicted from the linearizations. This behaviour is often referred to as nonlinear instability.

The fact that analyses based on linearization cannot accurately predict the qualitative behaviour of U_n for fixed k should not be surprising: there is a host of nonlinear phenomena (chaos, bifurcations, limit cycles ...) which cannot possibly be mimicked by a linear model.

The nonlinear instability phenomenon may be considered from a number of different points of view. A thorough survey of the literature is out of our scope here, but at least the following ideas should be briefly mentioned: (i) The role of spatial discretizations which conserve positive definite quantities in order to prevent blow ups (Arakawa [1]). (ii) The link between such discretizations and Galerkin methods (see the survey [8] by Morton). (iii) The study of time integrators which behave dissipatively in dissipative nonlinear situations (contractivity, B-stability, Dahlquist, Butcher, Spijker; see the recent book by Dekker and Verwer [5]). (iv) Fourier analysis; role of aliasing errors. (v) The important and widely quoted paper by Fornberg [6]. (vi) The connection with stability ideas in fluid mechanics investigated by Newell and his co-workers [4].

Recently there has been a growing interest in studying the fixed k behaviour from the point of view of the theory of dynamical systems [2], [7], [14]. In this paper we try to convey to numerical analysts the flavour of the powerful dynamic approach. To this end an example will be presented which illustrates several important features. The treatment of this simple case (the pendulum equation with central differences) is based on our earlier articles [12], [13], [16] and on the thesis [15]. The reader is referred to these references for applications to PDEs and for several extensions of the results discussed here.

2. THE PENDULUM EQUATION

We consider the well-known ODE system

$$u' = -\sin v, \quad v' = u, \tag{2.1}$$

describing the evolution in time of the angle v and angular velocity u of a mathematical pendulum. The phase-plane of (2.1) is displayed in many textbooks. It consists of (i) the stable equilibria $v = 2m\pi$, $u = 0$, m integer (the pendulum remains at its lowest position), (ii) the unstable equilibria $v = (2m+1)\pi$, $u = 0$, m integer (the pendulum stays at its highest position), (iii) libration orbits (the pendulum oscillates around a stable equilibrium), (iv) rotation orbits (v increases or decreases monotonically) and (v) separatrices between the libration and rotation orbits.

We also recall that in any libration domain of the (u,v)-plane (i.e. near a stable equilibrium) it is possible to change the dependent variables from (u,v) into the so-called action/angle variables (I,ϕ) [3, chap. 10]. (The abstract angle ϕ is not to be confused with the physical angle v.) Among the properties of (I,ϕ) we need the following three: (i) $u = u(I,\phi)$, $v = v(I,\phi)$ are 2π-periodic with respect to ϕ, i.e. ϕ behaves like a genuine angle. (ii) I takes the value 0 at the stable equilibrium and increases away from it. (iii) In the new variables (2.1) takes the simple form

$$I' = 0, \quad \phi' = \omega(I), \tag{2.2}$$

where ω is a known function of I. It is possible to give explicit, closed expressions of the transformation $I(u,v)$, $\phi(u,v)$ and of the function $\omega(I)$ in terms of elliptic integrals, but they are not required here.

The main advantage of the new variables is that now (2.2) is readily integrated to yield

$$I(t) = I(0), \quad \phi(t) = \omega(I(0))t + \phi(0). \tag{2.3}$$

In particular it is clear that $I(u,v)$ = constant provides the equations of the orbits in the (u,v) variables.

If Δt is a given, fixed time increment, it is convenient to introduce the Δt-flow of (2.1)/(2.2). By definition this is the mapping which sends the generic point (u_0,v_0) of the phase plane into the point $(\underline{u}_0,\underline{v}_0) = (u(\Delta t), v(\Delta t))$, where $(u(t),v(t))$ is the solution of (2.1) which satisfies the initial conditions $u(0) = u_0$, $v(0) = v_0$. It follows trivially from (2.3) that in action/angle variables, the t-flow is given by the transformation $(I_0,\phi_0) \to (\underline{I}_0,\underline{\phi}_0)$, with

$$\underline{I}_0 = I_0, \quad \underline{\phi}_0 = \phi_0 + \omega(I_0)\,\Delta t. \tag{2.4}$$

When (I,ϕ) are interpreted as the radius and the angle respectively in a system of polar coordinates, (2.4) is easily described: each point rotates around the origin by an angle $\omega(I_0)\Delta t$ which depends on the radius I_0. In other words, the flow leaves invariant each circle with centre at the origin. Within a circle the transformation is a rotation, but the overall effect is not that of a rigid rotation, as the angle being rotated varies with the particular circle. Transformations of this kind are called twist mappings [9].

It is useful to realize that if $\omega(I_o)\,\Delta t/2$ is a rational number p/q, then q applications of the twist (2.4) send the point back to its initial position after having completed p revolutions around the origin. On the other hand, irrational values imply that the corresponding points never return to their initial positions in the repeated iteration of the twist. In fact in this case the iterates fill densely the corresponding invariant circle and even possess a property of ergodicity: the number of iterants on any arc of the circle is proportional to the size of the arc [3].

3. LEAP FROG DISCRETIZATIONS

The system (2.1) is discretized by means of the leap frog (explicit midpoint) technique to yield, $n = 1,2, \ldots$

$$\begin{aligned} U_{n+1} &= U_{n-1} - 2k \sin V_n, \\ V_{n+1} &= V_{n-1} + 2k\, U_n. \end{aligned} \tag{5.1}$$

Equivalently, (5.1) can be rewritten in the form

$$U_{2n} = U_{2n-2} - 2k \sin V_{2n-1}, \tag{5.2a}$$

$$V_{2n} = V_{2n-2} + 2k\, U_{2n-1}, \tag{5.2b}$$

$$U_{2n+1} = U_{2n-1} - 2k \sin V_{2n}, \tag{5.2c}$$

$$V_{2n+1} = V_{2n-1} + 2k\, U_{2n}, \tag{5.2d}$$

$n = 1,2, \ldots$, where we have simply displayed two consecutive steps of (5.1) (unrolled the loop in computer science jargon). Now (5.2) presents two remarkable features:

(i) Formulae (5.2a)/(5.2d), which compute U at an even numbered grid point and V at an odd numbered grid point, are uncoupled from formulae (5.2b)/(5.2c), which compute odd numbered Us and even numbered Vs. We refer to (5.2a)/(5.2d) as the even/odd iteration and to (5.2b)/(5.2c) as the odd/even iteration. This splitting is a result of the simple structure of (2.1) and would not carry over to more complex systems.

(ii) If U_{2n}, V_{2n+1}, U_{2n+1}, V_{2n} are regarded as approximations to $p(2nk)$, $s(2nk)$, $r(2nk)$, $q(2nk)$ respectively, where p, s, r, q are functions of t which satisfy

$$p' = -\sin s, \tag{5.3a}$$

$$q' = r, \tag{5.3b}$$

$$r' = -\sin q, \tag{5.3c}$$

$$s' = p, \tag{5.3d}$$

then (5.2) provides a consistent discretization of the system (5.3). The fact that leap frog points originating from a d-dimensional system (d = 2 in (2.1)) can also be regarded as consistent approximations of a 2d-dimensional system (given here by (5.3)) is universal and plays a major role in understanding leap frog discretizations. See [12] for a thorough discussion.

The system (5.3) also splits: (5.3a)/(5.3d) are not coupled to (5.3b)/(5.3c). The situation is then as follows. The values U_{2n}, V_{2n+1} originating from the even/odd iteration approximate the velocity p and position s of the pendulum (5.3a)/(5.3d), which we call the even/odd pendulum. The values U_{2n+1}, V_{2n} of the odd/even iteration approximate the velocity r and position q of a different pendulum (5.3b)/(5.3c), the odd/even pendulum. We emphasize that there is no coupling between both pendulums.

A final property of the leap frog recursion is given next.

(iii) The transformations $T_{eo}: (U_{2n-2}, V_{2n-1}) \to (U_{2n}, V_{2n+1})$, $T_{oe}: (U_{2n-1}, V_{2n-2}) \to (U_{2n+1}, V_{2n})$ preserve the area in the (u,v)-plane, i.e. whenever D is a plane domain, D, $T_{oe}(D)$, $T_{eo}(D)$ possess the same area. As in (ii), this is a general property of leap frog discretizations [12].

4. LINEARIZATIONS

Linearization of the equations of the even/odd iteration near the equilibrium U = V = 0 results in a recursion

$$U_{2n} = U_{2n-2} - 2k\, V_{2n-1}, \quad V_{2n+1} = V_{2n-1} + 2k\, U_{2n},$$

whose solutions can be written in closed form in terms of the corresponding eigenvalues/vectors. It can be shown that if $k \geq 1$ then the solutions grow with n. For $k < 1$ the successive U, V values remain on an ellipse of eccentricity $(2k/(1+k))^{1/2}$ and major axis along the bisectrix of the first quadrant. Note that as k tends to 0 the orbits tend to be circles in the (u,v)-plane, thus mimicking a property of the linearized pendulum system $u' = -v$, $v' = u$. On the other hand values of k just below 1 will result in very elongated ellipses.

For the linearization of the odd/even iteration the same results hold, except that now the major axis is directed along the bisectrix of the second and fourth quadrants.

5. THE TWIST THEOREM

We have just seen that, provided that $k < 1$, the linearized theory predicts that the even/odd points will remain on an ellipse in the (u,v)-plane. The numerical experiments in the next section show that the real behaviour is far more complicated. A rigorous nonlinear analysis will now be presented. Recall that the even/odd transformation T_{eo} is an area preserving mapping which approximates consistently, i.e. up to $O(k^2)$, the $2k$-flow of the even/odd pendulum. This flow is in turn a twist mapping when written in action/angle variables. The behaviour of area preserving perturbations of twist mappings is well understood. The main result is the so-called twist theorem due to Moser, see e.g. [9, p.51].

THEOREM. Assume that in (2.4) $\omega(I_o)$ is any smooth function with nonvanishing derivative and defined for $0 < a \leq I_o \leq b$. Let ε denote a positive number. Then there exists δ, depending on ω, a, b and ε but not on Δt such that any smooth area preserving mapping

$$\underline{I}_o = f(I_o,\phi_o), \quad \underline{\phi}_o = \phi_o + g(I_o,\phi_o), \quad a \leq I_o \leq b, \tag{5.1}$$

where f, g are 2π-periodic in ϕ_o, that is close to (2.4) in the sense that

$$|f(I_o,\phi_o) - I_o| + |g(I_o,\phi_o) - \omega(I_o)\,\Delta t| < \delta\,\Delta t, \tag{5.2}$$

possesses an invariant curve contained in $a \leq I_o \leq b$ with parametric equations

$$I_o = c + z(\xi), \quad \phi_o = \xi + y(\xi), \tag{5.3}$$

where c is a constant and z and y are smooth 2π-periodic functions of ξ and satisfy $|z| + |y| < \varepsilon$. (Here $|\cdot|$ denotes a suitable norm, see [9].) Furthermore, on the invariant curve (5.3) the transformation (5.1) is simply given by $\xi \to \xi+\alpha$, where $\alpha/2\pi$ is 'very irrational' (again see [9] for a precise definition). In fact each value α in the range of $\omega(I_o)\,\Delta t$, with $\alpha/2\pi$ very irrational originates such an invariant curve.

In our context $t = 2k$, (2.4) is the $2k$-flow of the even/odd pendulum and (5.1) the transformation T_{eo}. The condition (5.2) is satisfied for k small enough because, by consistency, the right hand-side is $O(k^2)$. The theorem predicts that, for k small (how small depends on the distance to the equilibrium), T_{eo} possesses invariant curves which are close to the orbits of the pendulum system. On these invariants curves the transformation just acts as a rotation of irrational angle. There are also initial data which do not lie on an invariant curve. For these, the behaviour of the succesive points U_{2n}, V_{2n-1}, $n = 1, 2, \ldots$ is very involved. In any case, for k small those points must be surrounded by an invariant curve and therefore cannot scape away from the equilibrium $U = V = 0$, thus rigorously guaranteeing stability.

It is not necessary to mention that the results above also hold for the odd/even iteration.

The multidimensional analogue of the Moser twist mapping is given by the KAM (Kolmogorov-Arnold-Moser) theorem. For applications of this theorem to the analysis of numerical methods see [16].

6. NUMERICAL EXPERIMENTS

The leap frog iteration (5.1) was implemented on a computer with $U_o = 0$, V_o a parameter and U_1, V_1 taken from the application of Euler's rule. Some results will now be discussed.

In figure 1, $k = .1$ and each plotted point has coordinates U_{2n+1}, V_{2n}, so that the odd/even iteration is displayed. There are three initial conditions and for each 2,500 points were computed. For $V_o = 1, 2$ we find that the points place themselves on curves similar to the true pendulum orbits. We are thus facing the invariant curves predicted by the twist theorem. However for $V_o = 3.14$ it is quite clear that the points do not lie on a curve, but rather fill a region of finite width. Note that all three plots are elongated along the second and fourth quadrants as predicted by the linearized analysis of section 4. (A similar remark applies in subsequent experiments.)

In figure 2, k has been increased up to .625, while still displaying the odd/even iteration. There are two invariant curves corresponding to $V_o =$ 1, 1.5. The value $V_o = 2$, which lead before to an invariant curve, originates now a remarkable phenomenon. The orbit consists of six suborbits so that after six iterations the computed point returns to its original suborbit. Thus, if only every sixth iterant were plotted or in other words the power T_{oe}^6 were considered rather than T_{oe}, then only one suborbit would be seen. In fact the suborbits are twist theorem invariant curves of T_{oe}^6 around a fixed point of T_{oe}^6 (or equivalently a 6-periodic point of T_{oe}).

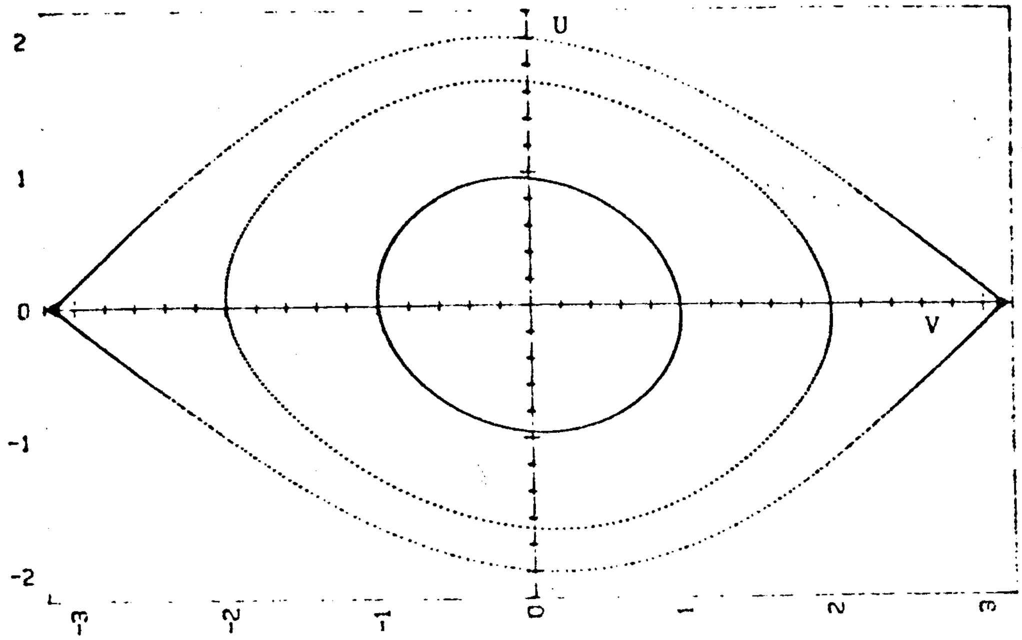

Figure 1. $k = .1$, $V_o = 1., 2., 3.14$.

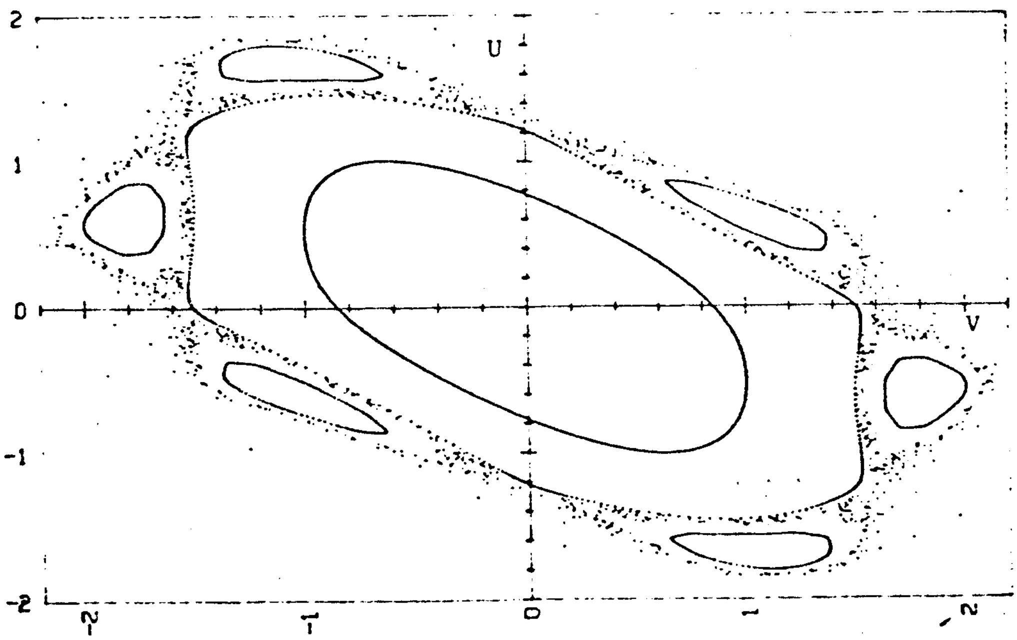

Figure 2. $k = .625$, $V_o = 1., 1.5$ (libration-like orbits), 2. (six suborbits), 2.1 (scattered points with eventual escape).

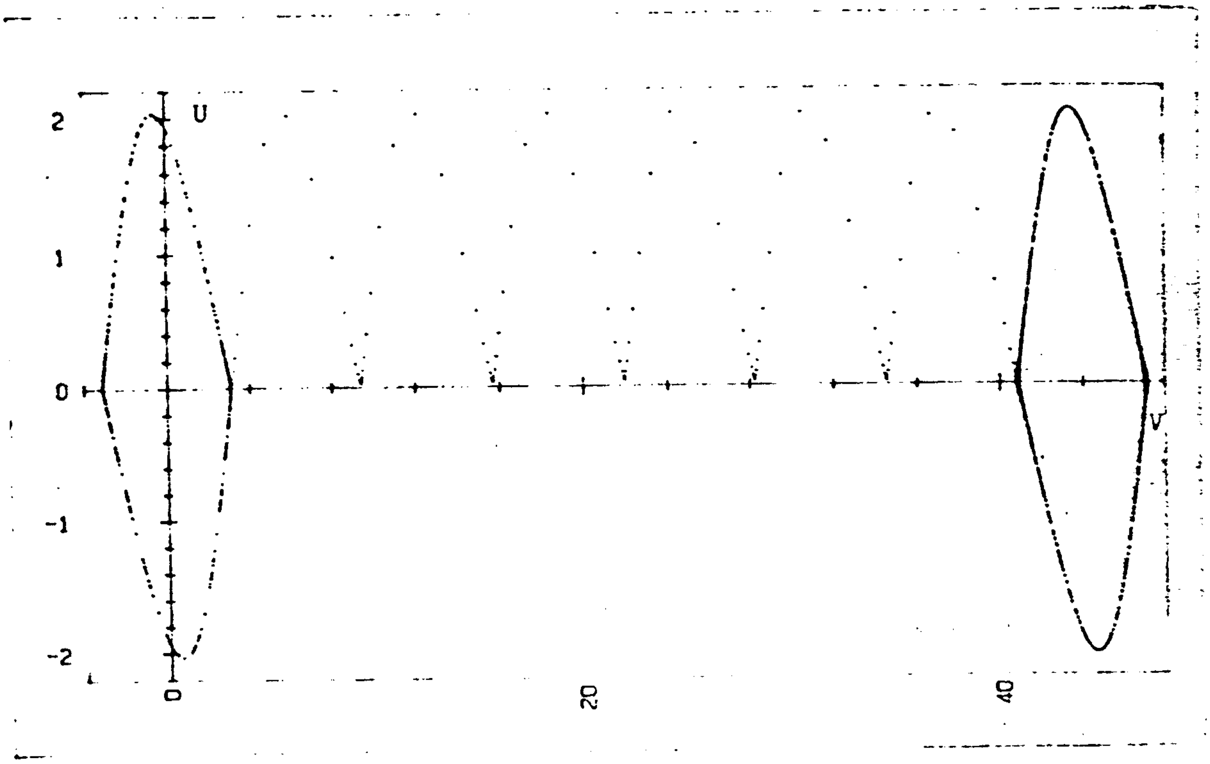

Figure 3. $k = .3$, $V_o = 3.1$. Spurious switch from libration to rotation near saddles.

Such a periodic point of T_{oe} cannot give rise to a twist invariant curve of T_{oe}, since these curves only originate from points which rotate by a very irrational angle and, of course a 6-periodic point rotates by $2p\,\pi/6$ radians.

Also displayed in figure 2 is the initial condition $V_o = 2.1$. At first, the points are scattered outside the islands corresponding to $V_o = 2.$, but eventually they escape from the plotted area. This behaviour is now possible because for the large value of k being used, the point $V_o = 2.1$, $U_o = 0$ is not surrounded by any invariant curve.

Another instance of escape is given in figure 3, where $k = .3$ and $V_o = 3.1$. There are 4,000 computed points. These first describe the expected libration orbit around $V = 0$, but later they switch to a rotation orbit and still later they settle in a libration orbit around $V = 14$. The switches occur near the unstable equilibria (saddles) $V = \pi$, $V = 13\pi$. Ushiki [14] has analyzed in detail this sort of behaviour and the reader is referred to his paper for further details. We just point out that such switches near saddles take place in an essentially un predictable way and that they are not caused by round off errors.

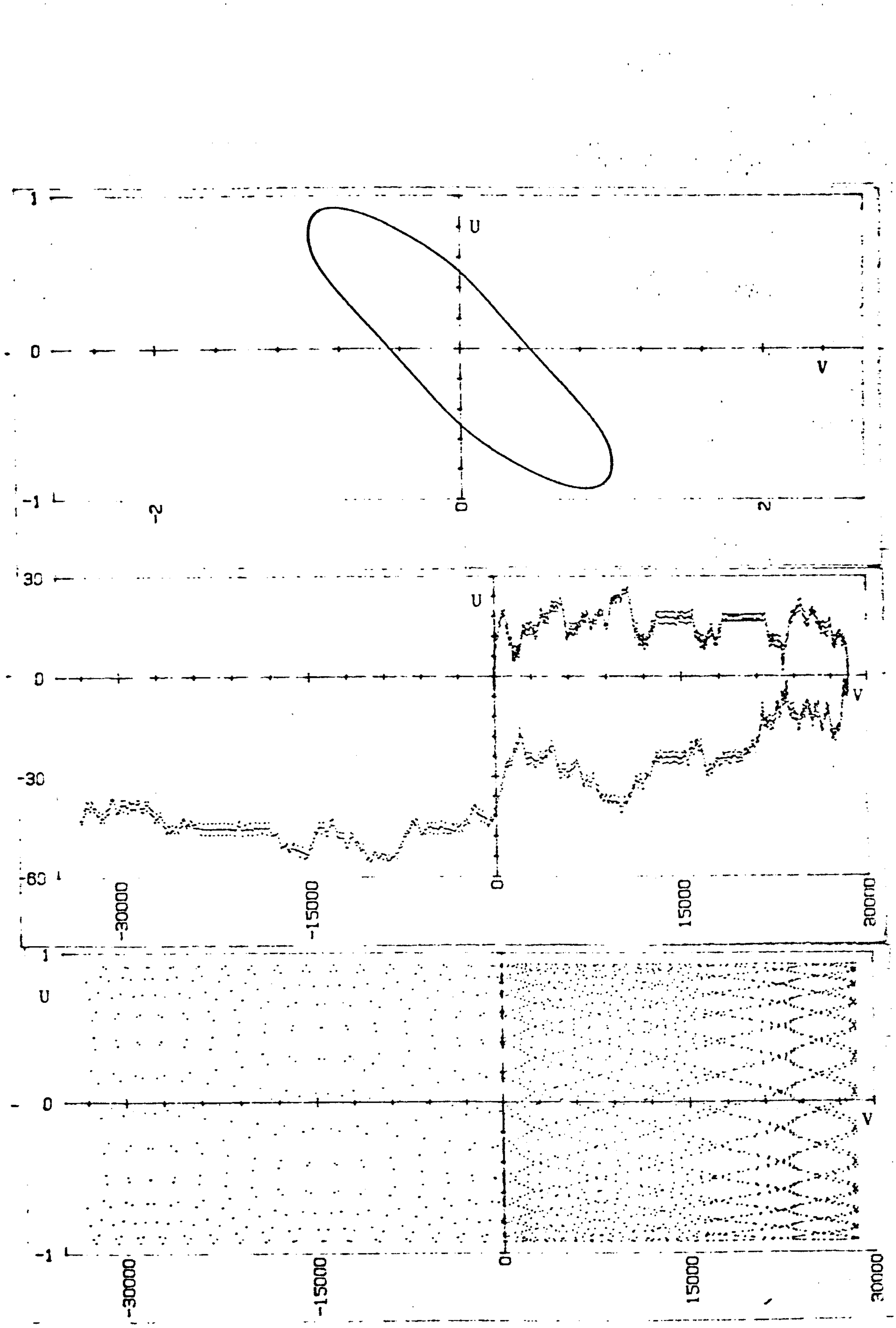

Figure 4. $k = .9$, $V_o = 1$. One numerical run mimics two pendulums. Top: odd/even points. Centre: even/odd points. Bottom: even/even points on a Lissajous-like curve.

The final experiment, in figure 4, has k= .9, near the maximum allowed by the linear condition, and $V_o = 1$. We first display the odd/even iteration which shows a familiar libration orbit with the elongation predicted by the linear theory. The even/odd iteration showed next is more interesting. (Note the change in scale of the plot.) Here we find a escape to a rotation orbit. When 2370 points have been represented, the orbit approaches a saddle, where it switches to a rotation in the opposite direction.

Comparison of the even/odd and odd/even points clearly illustrates that the leap frog points describe the motion of two uncoupled pendulums (5.3). There is no reasonable way of accounting for the behaviour of (U_n, V_n) when these are regarded, as they would normally be, as approximations to $(u(t_n), v(t_n))$, with u, v satisfying the single pendulum system (2.1). This is so even if we separate the behaviour of the odd (U_{2n+1}, V_{2n+1}) and even (U_{2n}, V_{2n}) points, a separation which has been suggested as a means for understanding the behaviour of leap frog schemes. The bottom part of figure 4 represents the even points (U_{2n}, V_{2n}). It is obvious that the dynamics displayed is unaccountable in terms of the pendulum system (2.1). Perhaps it is not without interest to point out that the plot has the appearance of a Lissajous curve. This is no surprise if we think once more of two pendulums with different frequencies.

7. CONCLUSIONS

In general, the investigation of the convergence of a numerical method provides little or no information on the qualitative behaviour of the sequence $U_0, U_1, \ldots, U_n, \ldots$ It is a (methodologically unfortunate) coincidence that in some simple, linear cases the same relations that must be imposed to achieve convergence guarantee the right qualitative behaviour. An instance is given by the familiar explicit scheme for the heat equation: the condition $r \leq \frac{1}{2}$ gives convergence and at the same time rules out the growth in time of the solution on a given grid and ensures the validity of the discrete maximum principle. (A more detailed discussion of this point has been provided by J.M. Sanz-Serna in a set of unpublished notes, available on request.)

In nonlinear situations the behaviour of $U_0, U_1, \ldots, U_n, \ldots$, when n is large, may be extraordinarily involved, as illustrated by the example considered in this paper. There is no hope of fully accounting for the phenomena involved by means of analyses based on linearization.

Fortunately, for a given, finite length of time $0 \le t \le T$, the use of high accuracy methods, together with small mesh-sizes guarantees that U_n is close to $u(t_n)$, thus ruling out pathological behaviours. Of course, a larger value of T will require smaller mesh-sizes.

We feel that case studies like the one presented in this paper together with estimates of the time needed for the pathologies to show up (see [16]) will play an essential role in the construction of a (Lax-like) definition of the concept of stability in nonlinear situations, a definition, in our opinion, missing at the moment. Steps in that direction have been taken within our research group and will be reported elsewhere.

ACKNOWLEDGEMENT

One of the authors (J.M.S.) has received financial support from the British Council in Spain.

REFERENCES

1. ARAKAWA, A. Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompresible flows. Part I, J. Comput. Phys. 1 (1966) 119-143.

2. AREF, H. and DARIPA, P. K. Note on finite difference approximations to Burgers' equation, SIAM J. Sci. Stat. Comput. 5 (1984) 856-864.

3. ARNOLD, V. I. Mathematical Methods of Classical Mechanics, Springer, New York 1978.

4. BRIGGS, W. L., NEWELL, A. C. and SARIE, T. Focusing: A mechanism for instability of nonlinear difference equations, J. Comput. Phys. 51 (1983) 83-106.

5. DEKKER, K. and VERWER, J. G. Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations, North Holland, Amsterdam 1984.

6. FORNBERG, B. On the instability of leap-frog and Crank-Nicolson approximations of a nonlinear partial differential equation, Math. Comp. 27 (1973) 45-57.

7. LAUWERIER, H. A. Dynamical Systems and Numercial Integration, Report AM-R8413, Centrum voor Wiskunde en Informatica, (Stichting Mathematical Centre), Amsterdam 1984.

8. 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. Academic Press, London 1977.

9. MOSER, I. Stable and Random Motions in Dynamical Systems, Princeton University Press, Princeton N.J. 1973.

10. PHILLIPS, N. A. An example of nonlinear computational instability, in "The Atmosphere and the Sea in Motion", B. Bolin ed., Rockefeller Institute, New York, 1959.

11. SANZ-SERNA, J. M. Stability and convergence in Numerical Analysis I: Linear problems, a simple, comprehensive account, in Nonlinear Differential Equations and Applications, J. K. Hale and P. Martínez-Amores eds., Pitman (to appear).

12. SANZ-SERNA, J. M. Studies in numerical nonlinear instability I: Why do leapfrog schemes go unstable?, SIAM J. Sci. Stat. Comput. 6 (1985) in the press.

13. SANZ-SERNA, J. M. and VADILLO, F. Studies in numerical nonlinear instability III: Augmented Hamiltonian Systems, submitted 1985.

14. USHIKI, S. Central difference scheme and chaos, Physica 4D (1982) 407-424.

15. VADILLO, F. Inestabilidad no Lineal en Análisis Numérico: Un Estudio Dinámico, Tesis, Universidad de Valladolid, mayo 1985.

16. VADILLO, F. and SANZ-SERNA, J. M. Studies in numerical nonlinear instability II: A new look at $u_t+uu_x = 0$, submitted 1985.

J.M. Sanz-Serna
Departamento de Ecuaciones Funcionales
Facultad de Ciencias
Universidad de Valladolid
Valladolid, Spain

F. Vadillo
Departamento de Matemática Aplicada
Facultad de Ciencias
Universidad del País Vasco,
Bilbao, Spain