THE USE OF THE METHOD OF AVERAGING IN PREDICTING CHAOTIC MOTION

A.C. FOWLER
Dept. of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 26 July 1983

The method of averaging is discussed in terms of its possible application to the prediction of chaotic orbits in differential equations.

1. In the last few years, approximate methods for the "construction" of chaotic orbits of low order differential equations have begun to appear, with varying success. Standard amplitude expansions, as applied to (Hopf-type) bifurcating systems, are inappropriate, and other, fully non-linear, approaches are required. Amongst these, the method of averaging and the related method of Kuzmak are prominent, and recently, Shimizu and Ichimura [1] have attempted to explain chaotic motion in a weakly dissipative three-dimensional oscillator by using Kuzmak's method. For a system of third order which has two constants of the motion E, Q, and corresponding oscillatory solutions, when a parameter $E = 0$, the method of averaging, or the usually more elegant method of Kuzmak [2], see also refs. [3–5], leads one to a pair of (autonomous) ordinary differential equations for E and Q as functions of $\tau = \epsilon t$, when $\epsilon \neq 0$ but is small. The simplification from three to two dimensions is clear, but since the worst an autonomous second order equation will usually do is have periodic solutions, one would not expect averaging of third order equations to be able to predict anything other than quasiperiodic solutions on tori. Shimizu and Ichimura's idea was that the underlying (fast) oscillatory solutions were of two types, depending on the region of the (E, Q) phase plane one was in, and that the switching between the two types would be effectively a chaotic process. The present note was motivated by these latter observations, and its purpose is to elaborate the explanation somewhat. It seems that all the "ingredients" for chaos have been identified, but exactly how they conspire to produce chaos is not clear. In particular, we wish to point out the dominant importance of the homoclinic orbit which exists in Shimizu and Ichimura's system when $\epsilon = 0$: this is in line with other recent studies, in particular ref. [6].

2. There have been a number of attemps to use the simplifications inherent in averaging in trying to deduce something of chaotic type [7–14]. One should also mention related ideas which use local approximate Poincaré maps without the full use of averaging, in order to prove the existence of strange invariant sets. In particular, Holmes [8] (see also refs. [15,16] applied Melnikov's method to systems of averaging type, to prove the existence of a strange invariant set of trajectories in the vicinity of homoclinic or heteroclinic orbits. This provides an explicit criterion for the existence of such a set, but stops short of analysing the global dynamics of weakly non-conservative systems, and is thus unable to distinguish a strange attractor. It is probably reasonable to say that the use of averaging has not been wholly successful at either predicting or explaining chaos, and the paper of Shimizu and Ichimura [1] is a notable step forward in this direction, as is also the lucid discussion of the Lorenz equations in Sparrow's book [6]. The aim of the present note is therefore to indicate how averaging can be formally completed to accomodate the switching process [1], and thus how it can (in principle) be used as a predictor for chaotic solutions. Our discussion will be fairly general, but also fairly qualitative.

$$\mathrm{d}E/\mathrm{d}t = \epsilon f(x, \dot{x}, Q), \quad \mathrm{d}Q/\mathrm{d}t = \epsilon g(x, \dot{x}, Q), \tag{1}$$

where we have written

$$Q = 1 - m, \quad E = \tfrac{1}{2}\dot{x}^2 + b[\tfrac{1}{4}x^4 - \tfrac{1}{2}Qx^2]. \tag{2}$$

E and Q are slowly varying functions of $\tau = \epsilon t$, and the second equation in (2) is then essentially a non-linear oscillator, whose solutions $x(t; E, Q)$ can be written in terms of elliptic functions [17]. In particular, x has a "period" $P = 4K$, where $K(k)$ is the complete elliptic integral of the first kind. By averaging (1) over the period P, one derives the averaged equations

$$\mathrm{d}E/\mathrm{d}\tau \approx F(E, Q), \quad \mathrm{d}Q/\mathrm{d}\tau \approx G(E, Q), \tag{3}$$

which govern the long-time evolution of E and Q: here $\tau = \epsilon t$.

A comment is perhaps in order. In order to obtain higher order terms, Kuzmak's method, as elaborated in refs. [4,5], for example, is preferable, since systems of the type under consideration are rarely given in the simple form to which the averaging theorem applies. However, at leading order, the two methods are equivalent, the only difference being that the fast time scale also evolves (slowly) in Kuzmak's method so that the period remains constant.

Now, it is evident from (2) that the fast oscillation in x has a homoclinic orbit when $E = 0$. Consequently, the underlying period $P(E, Q)$ tends to infinity as $E \to 0$. In fact, the modulus k is related to E by $E \propto 1 - k^2$, and thus k goes through 1 when E goes through zero; also $P \sim \ln(1/|E|)$ as $E \to 0$. It is the line $E = 0$ which separates the two types of oscillatory solution for $x(t; E, Q)$, and it is the crossing of this line in (E, Q) space which heralds chaotic behavior.

However, all is not well. Both the method of averaging and the method of Kuzmak becomes invalid when the period of the underlying oscillation P becomes too large [of $O(1/\epsilon)$]. This is exactly what happens when $E \to 0$. Consequently, *averaging does not straightforwardly apply in the vicinity of $E = 0$*, and we need to extend the analysis to cope with this difficulty. It is interesting to examine the solutions of the averaged equations (3) as $E \to 0$. One finds that E reaches zero in finite time [essentially $\mathrm{d}E/\mathrm{d}\tau \sim 1/\ln(1/|E|)$], but also $\mathrm{d}E/\mathrm{d}\tau = 0$ when $E = 0$. According to the averaged equations, $E = 0$ is strictly an invariant line, and the averaged variables behave non-uniquely, since they can pass an arbitrary distance along $E = 0$ and then go into $E \neq 0$. Substantial discussion of this phenomenon in the Lorenz equations is given by Sparrow [6], and it is apparent that the possibility exists of "anomalous" orbits occurring, which spend some time "on" $E = 0$; a similar possibility exists here, too, and may be more germane to the existence of chaos than any switching which occurs as E crosses zero.

3. The analysis of what exactly happens when $E \to 0$ is clearly necessary before one can describe, even qualitatively, the relation between the averaged system and the full system. Some discussion of this difficulty has been given by Sparrow [6] and Swinnerton-Dyer [18], and we essentially follow Sparrow's prescription for an appropriate technique.

The key is the realization that the averaged (differential) equations are obtained by approximating the first Euler difference ($[E(t + P) - E(t)]/P$, say) by the limiting differential, $\mathrm{d}E/\mathrm{d}t$, but that though *this* approximation breaks down when $P \to \infty$ [$P \sim O(1/\epsilon)$], nevertheless, the construction of the Euler difference is still feasible. To be specific, let us consider a system of the form

$$\mathrm{d}E/\mathrm{d}t = \epsilon f(\boldsymbol{x}; E, Q), \quad \mathrm{d}Q/\mathrm{d}t = \epsilon g(\boldsymbol{x}; E, Q), \tag{4a}$$

$$\mathrm{d}\boldsymbol{x}/\mathrm{d}t = \boldsymbol{h}(\boldsymbol{x}; E, Q). \tag{4b}$$

The systems in refs. [1,6], for example, can be written in this form. We suppose that (4b) has a periodic solution

$$\boldsymbol{x} = \boldsymbol{u}(t) \tag{5}$$

of period

$$P = P(E, Q). \tag{6}$$

$\boldsymbol{u}$ will additionally have various constants of integration (phases), whose presence, however, does not affect the slow evolution of E and Q [4,5]. We assume that $P \to \infty$ as $E \to 0$. For example, elliptic functions have $P \sim \ln(1/|E|)$ as $E \to 0$, and this will be generally true, since as $E \to 0$, $\boldsymbol{u} \to \boldsymbol{u}_0$ which is (assumed to be) a homoclinic orbit in the neighborhood of which the period is controlled by the linear behavior of $\boldsymbol{x}$ near the fixed point on $\boldsymbol{u}_0$ (which we take to be the origin) [6].

We average by integrating (4) over a period of $\boldsymbol{u}$:

$$E(t+P)-E(t)\approx P\epsilon F(E(t),Q(t))\,,$$

$$Q(t+P)-Q(t)=P\epsilon G(E(t),Q(t))\,, \tag{7}$$

where F and G are the averages of f and g. Normally, one considers (7) as approximate differential equations, but equally well, they may be considered (and in fact *are*) approximate Poincaré maps which relate successive values of E and Q [actually (7) is a projection of the full map, which is sufficient, since the phase constants do not enter the relation]. Thus, a continuous orbit of the averaged equations corresponds to a sequence of iterates of E, Q, say $\{E_n\}$, $\{Q_n\}$, spaced $\mathrm{O}(\epsilon)$ apart. Although (if we assume PF is bounded and non-zero as $E\to 0$ (as is the case here), the averaged variable $E(\tau)$ reaches zero in finite τ, and is then indeterminate, it is apparent from the Poincaré map that nevertheless $E_{n+1}-E_n\sim\epsilon$: consequently, the iterates of E_n will simply continue across $E=0$ in a perfectly determinate manner; for the same reason, numerically computed trajectories will generally do so as well.

Shimizu and Ichimura [1] ascribe chaos to the indeterminancy of picking the sign of the solution (x) when $E<0$ (when symmetric positive and negative fast oscillations exist). Such "phase-mixing" chaos is at least very "weak", since x^2 is doubly periodic, for example. But there is another, more dramatic, mechanism of chaos present. That is, suppose that some iterate E_n in the sequence comes so close to $E=0$ that $P\sim 1/\epsilon$. The derivation of (7) is then suspect, since if $P\sim 1/\epsilon$, E_n is so close to the stable manifold of the origin that the x trajectory will spend a time $t\sim 1/\epsilon$ near the origin, and consequently, a different approach to the problem is required. It is still possible to construct the Poincaré map, but the argument is now similar to that used in constructing a map in the neighborhood of a homoclinic bifurcation [6].

We suppose

$$E_n=\exp[-R/\epsilon]\,,\quad R\sim\mathrm{O}(1)\,, \tag{8}$$

and $Q=Q_n, x=x_n$, let us say at time $t=0$. We integrate forward in time to find x. Since x is close to the homoclinic orbit, it will closely approach the origin for large t. At such t, x will be approximately

$$x\approx a\mathbf{v}_{\mathrm{s}}\mathrm{e}^{-\lambda t}+bE_n\mathbf{v}_{\mathrm{u}}\mathrm{e}^{\mu t}\,, \tag{9}$$

where a and b are [O(1)] amplitudes which will depend on x_n and Q_n, $\mathbf{v}_{\mathrm{s}}$ is the eigenvector corresponding to the *least* stable (smallest λ) eigenvalue $-\lambda<0$ of the linearized system for x at the origin, and $\mathbf{v}_{\mathrm{u}}$ is the eigenvector corresponding to the *most* unstable eigenvalue $\mu>0$. We suppose λ and μ are real. We can neglect all the other eigenvectors, since the corresponding eigenmodes will be much smaller. With E_n given by (8), x is going to become *transcendentally* small ($x\sim\exp[-\mathrm{O}(1/\epsilon)]$), and its evolution is governed by [from (4)]

$$\dot{x}=[Dh(0;\ 0,Q)]\,x=L(Q)x\,. \tag{10}$$

From (9), x is going to be near the origin for a time $t\sim 1/\epsilon$, or $\tau\sim 1$; during this time, Q is governed by, from (4),

$$\mathrm{d}Q/\mathrm{d}\tau=g(0;0,Q)=g_0(Q)\,, \tag{11}$$

whence we determine $Q=Q(\tau)$, with $Q(0)=Q_n$. Asymptotic solutions of the linear equation (10) are [to $\mathrm{O}(\epsilon)$],

$$x\sim\mathbf{v}_\rho(\tau)\exp[\Lambda(\tau)/\epsilon]\,,\quad \mathrm{d}\Lambda/\mathrm{d}\tau\sim\rho\,, \tag{12}$$

and ρ and $\mathbf{v}_\rho$ are an eigenpair for $L(\tau)$:

$$L\mathbf{v}_\rho=\rho\mathbf{v}_\rho\,. \tag{13}$$

We assume that the eigenvalue μ remains maximal as Q varies. For a third order system, or where x is two dimensional, this is necessarily the case. Thus the solution for x which matches to (9) is

$$x\sim aV_{\mathrm{s}}(\tau)\exp[-\Lambda_{\mathrm{s}}(\tau)/\epsilon]$$
$$\quad+bV_{\mathrm{u}}(\tau)\exp\{[-R+\Lambda_{\mathrm{u}}(\tau)]/\epsilon\}\,, \tag{14}$$

where

$$\Lambda_{\mathrm{s}}(0)=0\,,\quad \Lambda_{\mathrm{u}}(0)=0\,,$$

$$V_{\mathrm{s}}(0)=\mathbf{v}_{\mathrm{s}}\,,\quad V_{\mathrm{u}}(0)=\mathbf{v}_{\mathrm{u}}\,, \tag{15}$$

and

$$\mathrm{d}\Lambda_{\mathrm{s}}/\mathrm{d}\tau=\lambda(\tau)\,,\quad \mathrm{d}\Lambda_{\mathrm{u}}/\mathrm{d}\tau=\mu(\tau)\,, \tag{16}$$

where $-\lambda$ is the maximal negative, μ is the maximal positive eigenvalue as Q varies. Of interest is when x increases to O(1) again, for then x leaves the origin, essentially along the unstable manifold of another homoclinic orbit. However, as t increases, both Q and E will increase by $\mathrm{O}(\epsilon)$ again when the trejectory hits the return plane. What this means is that, in general E_{n+1} will *not* be transcendentally small (one leaves the invariant line), and that Q_{n+1} will essentially be given by its value as x leaves the neighborhood of the origin. From (14), this is approximately when $\tau = \tau^*$, where $\Lambda_u(\tau^*) = R$. Consequently, we derive an approximate difference equation for Q_n,

$$Q_{n+1} = Q(\tau^*), \tag{17}$$

where Q is the solution of (11), $Q(0) = Q_n$, and τ^* is given by

$$R = \epsilon \ln(1/|E_n|) = \int_0^{\tau^*} \mu(\tau)\,\mathrm{d}\tau\,, \tag{18}$$

μ being the largest (assumed real) eigenvalue of $L(Q)$. A similar argument would seem to indicate $E_{n+1} \sim \mathrm{O}(1)$, but in the Lorenz equations [6], and Shimizu and Ichimura [1], one finds that PF in (7) is finite as $E \to 0$: equivalently

$$\int_{-\infty}^{\infty} f(u_0, 0, Q)\,\mathrm{d}\tau$$

is finite. In ref. [1], this is because $f = \alpha_1 u_0^2 + \alpha_2 u_0^4$, and u_0 is essentially sech(t). This is liable to be quite general, e.g., for almost conservative systems of the form $\ddot{x} + \nabla V(x) = \epsilon f(x, \dot{x}, t)$, if f is well behaved. In that case E remains small, and $E_{n+1} \sim \mathrm{O}(\epsilon)$ if $E_n \approx 0$.

4. There is little point elaborating the quantitative aspects of the discussion, since one needs the numerical solution of the averaged equation away from $E = 0$, anyway. We therefore proceed qualitatively. Behavior in any particular case will depend on the phase portrait of the averaged equations. We concentrate on the situation in refs. [1,12], where one can have a stable "limit cycle" of the averaged equations which crosses $E = 0$. By crossing, we mean that E_n is bounded away from zero as the trajectory crosses (i.e., does not get exponentially close). We then have a phase portrait as shown in fig. 1, where now we take $Q = Q$ as scalar. For any trajectory through a point $(Q, 0)$ on $E = 0$, we can define the "number" of passages through the Poincaré section, N as

$$N = \epsilon^{-1} \int_{Q'}^{Q''} \frac{\mathrm{d}\tau}{P(E, Q)}\,, \tag{19}$$

(i.e., number of periods per unit time = $1/P$, where Q' and Q'' are successive values of Q on $E = 0$. We would like to construct a mapping relating successive Q values on $E = 0$, using the notions expressed here. This is not strictly simple in general (one still needs a 2-D map), but it *is* possible to construct a 1-D map which, in some sense, describes the behavior of the system. (See also ref. [12].)

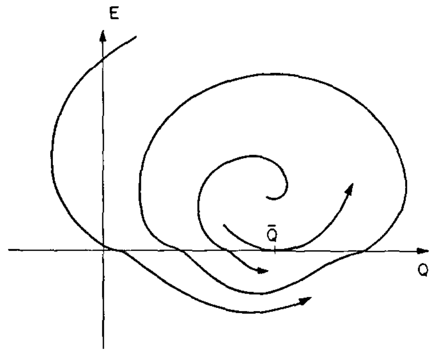

Fig. 1. Schematic phase portrait in the averaged (E, Q) phase plane, showing a limit cycle.

Suppose that $Q = \bar{Q}$ is the point of tangency of trajectories to $E = 0$. We can then pick a line C: $E_0 = \epsilon\phi(Q)$, $\phi(\bar{Q}) = 0$, $(Q - \bar{Q})\phi(Q) > 0$, and ask where trajectories leaving C first return there. This will define (since C is effectively parameterized by Q) a difference map $Q_n \to Q_{n+1}$. If the complications associated with $E = 0$ did not exist, this map, f, say, would be as shown in fig. 2, and its second iterate f^2 would be monotone increasing with the larger stable fixed point at Q^*, say. The second iterate is sketched in fig. 2. However, there will evidently be a sequence of values of Q, at intervals $\mathrm{O}(1/\epsilon)$, when N in (19) is such that the Nth iterate of E_0, $E_N = 0$ (see also ref. [6]). Denoting these values by $q_1, q_2, \ldots$, it is clear that if $Q_n = q_r + \exp[-\mathrm{O}(1/\epsilon)]$, then $E_n = \exp[-\mathrm{O}(1/\epsilon)]$, and consequently an excursion up the Q axis can occur. Thus the difference map will be as shown in fig. 2, but with some anomalous behavior near q_r. To specify what this is, take

$$Q_n = q_r \pm \exp[-R/\epsilon]\,, \tag{20}$$

then

$$E_N = \pm\,\alpha(q_r) \exp[-R/\epsilon]\,, \tag{21}$$

(by Taylor expansion about q_r), and to leading order, (17) and (18) are

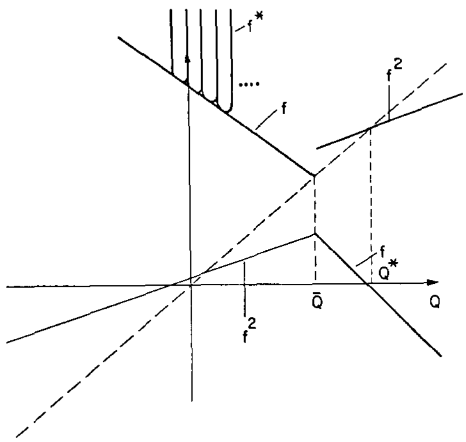

Fig. 2. Schematic form of the maps f, f^2, f^* (see text). f consists of two discontinuous branches, as does f^2.

$$Q_{n+1} = Q(\tau^*), \quad R = \int_0^{\tau^*} \mu(\tau)\,\mathrm{d}\tau . \tag{22}$$

As an example, suppose Q migrates on $E \mp 0$ towards $Q = 1$, e.g.,

$$\dot{Q} = c(1 - Q) \tag{23}$$

(this is actually the case in ref. [1], with $Q = 1 - m_0$). In ref. [1], we also have $\mu = (bQ)^{1/2}$, but for simplicity, we shall take μ constant. Then (22) and (23) give, for Q_n near q_r

$$Q_{n+1} \approx 1 - [1 - f(Q_n)]\,|Q_n - q_r|^{\epsilon c/\mu} \tag{24}$$

and the complete map, f^*, is like f, but with superimposed cusps, as partially indicated in fig. 2. These cusps are $\mathrm{O}(\epsilon)$ apart, and of thickness $\exp[-\mathrm{O}(1/\epsilon)]$, i.e., *very* narrow. One should compare this result with that of Fowler and McGuinness [19].

Of course, this is not entirely satisfactory, since after one iterate, say $Q_1 \to Q_2$, the point labeled by Q_2 will be out of phase, and the set of anomalous q values will be different. What one must really do, is to construct a two dimensional map relating successive values of Q "at" $E = 0$, and including the *phase* of the E-value there. However, (24) is of some interest, since one can expect (in a numerical integration) that most iterates of Q will follow the envelope curve towards the fixed point, because on any particular circuit, the range of Q-values which can give anomalous behaviour is very small ($\exp[-\mathrm{O}(1/\epsilon)]$). Also, numerical integrations of Marzec and Spiegel [12], show that on *one* branch of an attractor in the (E, Q) plane, a map such as f is actually numerically observed, which suggests that the mechanism behind the derivation of (24) is of some relevance.

On the limit cycle trajectory one can define a map relating successive values of E in a range $(0, \epsilon\phi^*)$ where $\phi^* = \lim_{E\to 0} PF$ [in (7)] evaluated at $Q = Q^*$ (the positive fixed point of f^2). One can think of this as essentially a rotation on the circle $\theta \in [0, 2\pi)$, where $E = \epsilon\phi^*\theta/2\pi$. Thus, provided motion on the (E, Q) limit cycle has a period $\widetilde{P}$ which is such that N in (19) is irrational, every so often (once every $\epsilon \exp[\mathrm{O}(1/\epsilon)]$ circuits) E becomes $\exp[-\mathrm{O}(1/\epsilon)]$, and an excursion along $E = 0$ occurs. This heralds a chaotic attractor of intermittent type, with excursions separated by t-intervals of $\exp[\mathrm{O}(1/\epsilon)]$. With $\epsilon = 10^{-3}$, as in ref. [1], this is upwards of 10^{400}!

5. The purpose of this note is to show how the method of averaging can be supplemented by a local multiple scale analysis in the vicinity of homoclinic orbits, to produce (in principle) a globally applicable Poincaré map. For a system with an attracting torus (i.e., a limit cycle in the averaged phase plane) which intersects an invariant line of homoclinic orbits, we predict an attracting chaotic set of intermittent type with laminar bursts of duration $\epsilon \exp[\mathrm{O}(1/\epsilon)]$ as $\epsilon \to 0$ (cf. ref. [20]). This is in addition to a weaker "chaos" associated with phase-switching across the invariant line. The bifurcation describing the transition from torus to chaotic attractor is associated with the point in parameter space where the averaged limit cycle is tangent to the invariant line. In terms of a Poincaré map, the attracting set for points $E \neq 0$ intersects $E = 0$, which is attracted to the (suitably defined) origin; this is a *crisis* for the attractor [21]. In terms of the flow, the torus become homoclinic to the origin ($Q = 1, E = 0$ in our example); that is, if there is a bifurcation parameter r [1] which describes the evolution of the averaged limit cycle, and tangency to $E = 0$ occurs at $r = r_\mathrm{h}$, then the interpretation of the results here is that for small ϵ, there is a homoclinic torus in the full system for r close to r_h. In analogy with ref. [6], we might expect an infinite number of tori to be produced at $r = r_\mathrm{h}$, and an associated infinite set of windows of period-doubling tori. Thus, the chaotic attractor occurs as a result of a homoclinic bifurcation.

References

[1] T. Shimizu and A. Ichimura, Phys. Lett. 91A (1982) 52.
[2] G.E. Kuzmak, J. Appl. Math. Mech. 23 (1959) 730.
[3] J.C. Luke, Proc. R. Soc. A292 (1966) 403.
[4] S. Kogelman and J.B. Keller, SIAM J. Appl. Math. 24 (1973) 352.
[5] J. Kevorkian and J.D. Cole, Perturbation methods in applied mathematics (Springer, Berlin, 1981).
[6] C. Sparrow, The Lorenz equations: bifurcations, chaos, and strange attractors (Springer, Berlin, 1982).
[7] P. Holmes, Philos. Trans. R. Soc. A292 (1979) 419.
[8] P.J. Holmes, SIAM J. Appl. Math. 38 (1980) 65.
[9] M. Nauenberg, Perturbative solution of Rössler's equation for chaotic motion, preprint (1981).
[10] J.R. Buchler, in: Nonlinear equations in physics and mathematics, ed. A.O. Barut (Reidel, Dordrecht, 1978) p. 85.
[11] N.H. Baker, D.W. Moore and E.A. Spiegel, Quart. J. Mech. Appl. Math. 24 (1971) 391.
[12] C.J. Marzec and E.A. Spiegel, SIAM J. Appl. Math. 38 (1980) 403.
[13] E.A. Spiegel, Ann. N.Y. Acad. Sci. 357 (1980) 305.
[14] J.P. Poyet, Thesis, Columbia University (1980).
[15] J.P. Keener, SIAM J. Appl. Math. 41 (1981) 127.
[16] J.P. Keener, Stud. Appl. Math. 67 (1982) 25.
[17] S.-N. Chow, J.K. Hale and J. Mallet-Paret, J. Diff. Eq. 37 (1980) 351.
[18] H.P.F. Swinnerton-Dyer, J. Lond. Math. Soc. 22 (1980) 534.
[19] A.C. Fowler and M.J. McGuinness, Physica 5D (1982) 149.
[20] P. Manneville and Y. Pomeau, Physica 1D (1980) 219.
[21] C. Grebogi, E. Ott and J.A. Yorke, Phys. Rev. Lett. 50 (1983) 935.