Physica 28D (1987) 143–154
North-Holland, Amsterdam

NOISE AND STATISTICAL PERIODICITY

Andrzej LASOTA
Institute of Mathematics, Silesian University, 40-007 Katowice, Poland

Michael C. MACKEY
Department of Physiology, McGill University, Montreal, Quebec, Canada H3G 1Y6

Received 19 April 1986
Revised manuscript received 20 October 1986

In this paper we study the properties of a broad class of discrete time dynamical systems in the presence of added external noise. We prove three theorems showing that added noise will either induce a type of statistical periodicity or statistical stability in the asymptotic behaviour of these systems. This result is based on the fact that, in the presence of noise, the Markov operator describing the evolution of densities has smoothing properties which allow the application of a recently discovered asymptotic decomposition theorem. Using this result it is possible to evaluate the (limiting) period of the sequence of densities. This effect is numerically illustrated by the addition of noise to a discontinuous map studied by Keener.

1. Introduction

An immense effort has been invested in the study of stable, periodic, and chaotic dynamical systems over the past 15 years. It is generally agreed that all real systems are subjected to "noise" of one variety or another, and many discrete and continuous time systems with stochastic perturbations have been examined numerically. The discrete time systems that have been the most studied in the presence of noise are the quadratic map [3, 4, 8, 9, 18, 19, 22], the circle map [6], the standard mapping [11, 20], and the two-dimensional Kaplan–Yorke map [7, 10].

These studies have clearly highlighted an important and difficult question: What role does extraneous noise play in the development of asymptotic behaviour of the system? Or, put differently, how is the behaviour of the unperturbed dynamics altered by noise?

The answer to this question appears to depend on the properties of the unperturbed system. The results of Kifer [13, 14] indicate that addition of noise to dynamical systems with highly irregular trajectories (Axiom A systems) will not result in an alteration of the statistical behaviour of systems since the invariant measure changes continuously with the noise level. However, in other cases the addition of noise results in evident changes in dynamical behaviour of trajectories and may make the dynamics more regular by creating absolutely continuous invariant measures [1, 2]. A general answer to the question of how trajectories behave in the presence of perturbation appears to be quite difficult and involves the use of sophisticated topological methods in the definition of attracting sets [21].

This paper considers the effect of stochastic perturbation and, more generally, randomly applied stochastic perturbation on discrete time dynamical systems from a statistical point of view. Thus we consider the behaviour of sequences of densities, or ensembles of trajectories, corresponding to a given system. We show that for every system which concentrates trajectories in a bounded region of phase space [see condition (2)], the addition of noise always has the effect of making the sequence of densities asymptotically periodic.

In the next section we specify a class of stochastically perturbed systems to be considered. Section 3 contains a theorem relating the onset of asymptotic periodicity with the addition of noise to a deterministic system. Section 4 extends this with a second theorem on the existence of asymptotic stability, while section 5 points out three important special cases of the theorems of the preceding two sections. There we also illustrate the onset of asymptotic periodicity in a map that is normally capable of only very irregular behavior. The results of sections 3 and 4 are generalized in section 6 where we prove a theorem concerning the occurrence of asymptotic periodicity in a broad class of stochastically perturbed discrete time dynamical systems. The paper concludes with some brief comments in section 7.

2. Mathematical preliminaries

In the next three sections of this paper, we consider a stochastically perturbed d-dimensional discrete dynamical system

$$x_{n+1} = S(x_n) + \xi_n \quad \text{for } n = 0, 1, \ldots . \tag{1}$$

In eq. (1), S is a transformation that maps R^d into itself and the quantities $\xi_0, \xi_1, \ldots$ are independent d-dimensional random vectors. In considering the system (1), we assume the following:

(i) The transformation S: $R^d \to R^d$ is Borel measurable and such that

$$|S(x)| \le \alpha |x| + \beta \quad \text{for } x \in R^d, \tag{2}$$

where α and β are non-negative constants, $\alpha < 1$, and $|\cdot|$ is the norm in R^d.

(ii) The vectors $\xi_0, \xi_1, \ldots$ are independent and all have the same distribution with density g, i.e., for all n

$$\text{prob}(\xi_n \in B) = \int_B g(x)\, \mathrm{d}x$$

for $B \subset R^d$, B a Borel set.

(iii) The density g has a finite first moment, i.e.

$$m = \int_{R^d} |x| g(x)\, \mathrm{d}x < \infty . \tag{3}$$

Finally, in addition to assumptions (i)–(iii), we assume that the initial condition x_0 is independent of the sequence of perturbations $\{\xi_n\}$.

The set of all densities in R^d is denoted by D, and thus

$$D = \left\{ f \in L^1(R^d) \colon f \ge 0,\ \|f\|_{L^1} = 1 \right\}.$$

If we let the density of the distributions of the x_n be denoted by f_n, then it is straightforward [17] to show that

$$f_{n+1}(x) = \int_{R^d} f_n(y) g(x - S(y))\, \mathrm{d}y \qquad \text{for } n = 0, 1, \ldots . \tag{4}$$

Thus, given an arbitrary initial density f_0, the evolution of densities by the system (1) is described by the sequence of iterates $\{P^n f_0\}$, where

$$Pf(x) = \int_{R^d} f(y) g(x - S(y))\, \mathrm{d}y \tag{5}$$

is a linear (Markov) operator from L^1 into itself.

3. Asymptotic periodicity

Our first step in the study of the sequence $\{P^n f\}$ is to show that the operator P is weakly constrictive. By definition, an operator P is *weakly constrictive* if there exists a weakly precompact set $\mathscr{F} \subset L^1$ such that

$$\lim_{n \to \infty} \rho(P^n f, \mathscr{F}) = 0 \quad \text{for } f \in D, \tag{6}$$

where $\rho(f, \mathscr{F})$ denotes the distance, in L^1 norm, between the element f and the set $\mathscr{F}$. The importance of weak constrictiveness is a consequence of the following theorem of Komornik [15], also proved in a more restricted case in ref. 16:

Spectral decomposition theorem. Let P be a weakly constrictive Markov operator. Then there is an integer r, two sequences of non-negative functions $g_i \in D$ and $k_i \in L^\infty$, $i = 1,\dots,r$, and an operator Q: $L^1 \to L^1$ such that for all $f \in L^1$, Pf may be written in the form

$$Pf(x) = \sum_{i=1}^{r} \lambda_i(f) g_i(x) + Qf(x), \tag{7}$$

where

$$\lambda_i(f) = \int_{R^d} f(x) k_i(x)\,\mathrm{d}x.$$

The functions g_i and the operator Q have the following properties:

(1) $g_i(x) g_j(x) = 0$ for all $i \neq j$, so that the densities g_i have disjoint supports.

(2) For each integer i there exists a unique integer $\omega(i)$ such that $Pg_i = g_{\omega(i)}$. Further, $\omega(i) \neq \omega(j)$ for $i \neq j$ and thus the operator P just serves to permute the functions g_i.

(3) $\|P^n Qf\| \to 0$ as $n \to \infty$ for every $f \in L^1$.

From eq. (7) it is clear that $P^n f$ may be written as

$$P^n f = \sum_{i=1}^{r} \lambda_i(f) g_{\omega^n(i)} + Q_n(f), \tag{8}$$

where $Q_n = P^{n-1}Q$, and $\omega^n(i) = \omega(\omega^{n-1}(i)) = \dots$, and $\|Q_n f\| \to 0$ as $n \to \infty$. The terms in the summation in eq. (8) are permuted with each application of P. Since $\{\omega^n(1),\dots,\omega^n(r)\}$ is just a permutation of $\{1,\dots,r\}$, there is a unique i corresponding to each $\omega^n(i)$. Hence, the summation portion of (8) may be rewritten as

$$\sum_{i=1}^{r} \lambda_{\omega^{-n}(i)}(f) g_i(x),$$

where $\{\omega^{-n}(i)\}$ denotes the inverse permutation of $\{\omega^n(i)\}$.

Rewriting the summation portion of (8) in this way reveals precisely what is happening with every successive application of the operator P. As the densities $g_i(x)$ all have disjoint supports, each successive application of P leads to a new set of scaling coefficients $\lambda_{\omega^{-n}}(f)$ associated with each density $g_i(x)$. Because r is finite, the summation portion of (8) is periodic with period less than $r!$, and since $\|Q_n f\| \to 0$ as $n \to \infty$ we say that for a weakly constrictive Markov operator the sequence $\{P^n f\}$ is *asymptotically periodic.*

Having developed this background, we are ready to state our first main result:

Theorem 1. If the transformation S: $R^d \to R^d$ and the density of the distribution of the stochastic perturbation respectively satisfy inequalities (2) and (3), then the Markov operator defined by eq. (5) is weakly constrictive.

Proof. We define

$$E(f) = \int_{R^d} |x| g(x)\,\mathrm{d}x \tag{9}$$

and consider the sequence $\{E(P^n f)\}$ for an $f \in D$. From eq. (5) and inequalities (2) and (3), it follows immediately that

$$\begin{aligned} E(P^{n+1}f) &= \int_{R^d}\int_{R^d} |x| P^n f(y) g(x - S(y))\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{R^d}\int_{R^d} |z + S(y)| P^n f(y) g(z)\,\mathrm{d}z\,\mathrm{d}y \\ &\le \int_{R^d}\int_{R^d} |z| P^n f(y) g(z)\,\mathrm{d}z\,\mathrm{d}y \\ &\quad + \int_{R^d}\int_{R^d} (\alpha|y| + \beta) P^n f(y) g(z)\,\mathrm{d}z\,\mathrm{d}y \\ &= m + \beta + \alpha E(P^n f). \end{aligned}$$

As a consequence,

$$E(P^n f) \le \frac{m+\beta}{1-\alpha} + \alpha^n E(f).$$

Choose an arbitrary $M > (m+\beta)/(1-\alpha)$. If $E(f) < \infty$ then, since $0 \le \alpha < 1$, for sufficiently large n, say $n \ge n_0(f)$, we have

$$E(P^n f) \le M. \tag{10}$$

For $\varepsilon > 0$, denote by $\delta(\varepsilon)$ a positive number such that

$$\int_B g(x)\,\mathrm{d}x \le \varepsilon \quad \text{whenever } \mu(B) \le \delta(\varepsilon), \tag{11}$$

where μ denotes the standard Lebesgue measure on R. Let $\mathscr{F} \subset D$ be the set of all densities f satisfying the following two conditions:

$$\int_{|x| \ge r} f(x)\,\mathrm{d}x \le \frac{M}{r} \quad \text{for } r > 0 \tag{12}$$

and

$$\int_B f(x)\,\mathrm{d}x \le \varepsilon \quad \text{for every } \varepsilon > 0 \text{ if } \mu(B) \le \delta(\varepsilon). \tag{13}$$

From standard and well-known criteria for weak precompactness [5], it follows from (12) and (13) that the set $\mathscr{F}$ is weakly precompact.

In order to verify (6), and demonstrate that P is weakly constrictive, consider an $f_0 \in D$ such that $E(f_0) < \infty$. From inequality (10) and the Chebyshev inequality it follows that

$$\int_{|x| \ge r} P^n f_0(x)\,\mathrm{d}x \le \frac{M}{r} \quad \text{for } r > 0 \text{ and } n \ge n_0(f_0). \tag{14}$$

Further, let $B \subset R^d$ and $\varepsilon > 0$ be given, and let $\mu(B) \le \delta(\varepsilon)$. Then, from the definition of P in eq. (5) we have

$$\begin{aligned}
&\int_B P^n f_0(x)\,\mathrm{d}x \\
&= \int_B \left\{ \int_{R^d} P^{n-1} f_0(y) g(x - S(y))\,\mathrm{d}y \right\} \mathrm{d}x \\
&= \int_{R^d} \left\{ \int_B g(x - S(y))\,\mathrm{d}x \right\} P^{n-1} f_0(y)\,\mathrm{d}y \\
&= \int_{R^d} \left\{ \int_{B - S(y)} g(z)\,\mathrm{d}z \right\} P^{n-1} f_0(y)\,\mathrm{d}y.
\end{aligned}$$

When B satisfies $\mu(B) \le \delta(\varepsilon)$, then the set $B - S(y)$ has the same property and, as a consequence of (11),

$$\int_B P^n f_0(x)\,\mathrm{d}x \le \varepsilon \int_{R^d} P^{n-1} f_0(y)\,\mathrm{d}y = \varepsilon.$$

Thus for every fixed $n \ge n_0(f_0)$ the function $f = P^n f_0$ satisfies both (12) and (13), and as a consequence $P^n f_0 \in \mathscr{F}$ for $n \ge n_0(f_0) + 1$ whenever $E(f_0) < \infty$. Since the set of all f satisfying $E(f_0) < \infty$ is dense in D, this implies (6), and the proof is complete.

Remark. The constrictiveness of integral operators defined by a kernel has been shown previously [15–17], but the operators were not derived by considering the stochastic perturbation of systems.

As a consequence of Theorem 1 in conjunction with our comments and remarks of the previous section, we know that the addition of any stochastic perturbation with a continuous distribution to a deterministic transformation on R^d will make that transformation asymptotically periodic from a statistical point of view. The only requirements are that the density of the distribution of the stochastic perturbation must possess a finite first moment [inequality (3)], and the transformation S must satisfy the growth condition (2).

4. Asymptotic stability

In eq. (8) of the previous section for $\{P^n f\}$, if $r = 1$ so that the summation is reduced to a single term, then for every $f \in D$ the sequence $\{P^n f\}$ converges to the same limit as $n \to \infty$ independent of f. In such situations we say that the operator P is *asymptotically stable*. For applied situations, this is extremely interesting and useful because of the possibility of calculating statistical properties characterizing the dynamics.

Using Theorem 1 we may prove the following result concerning the appearance of asymptotic stability of the Markov operator P defined by eq. (5).

Theorem 2. Assume that the Borel measurable transformation S: $R^d \to R^d$ and the density g satisfy inequalities (2) and (3). Further assume that there exists a point $z_0 \in R^d$ and a number

$$r_0 > \frac{m\alpha + \beta}{1 - \alpha}$$

such that

$$g(x) > 0 \quad \text{a.e.} \quad \text{for } |x - z_0| < r_0. \tag{15}$$

Then the Markov operator defined by eq. (5) is asymptotically stable.

Proof. To prove this theorem, we employ the following.

Lemma. Let P be a weakly constrictive Markov operator. Assume there is a set $A \subset R^d$ of nonzero measure, $\mu(A) > 0$, with the property that for every $f \in D$ there is an integer $n_1(f)$ such that

$$P^n f(x) > 0 \tag{16}$$

for almost all $x \in A$ and all $n \geq n_1(f)$. Then $\{P^n f\}$ is asymptotically stable.

The proof of this lemma may be found in ref. 17.

Since, by Theorem 1, we know that P is weakly constrictive, we need only to demonstrate that P satisfies the rest of the assumptions of the Lemma. Let $f \in D$ be arbitrary. Since f is integrable there is a bounded subset $B \subset R^d$ such that

$$\int_B f(x)\,\mathrm{d}x = \tfrac{1}{2}.$$

Define $f_1(x) = 2f(x)1_B(x)$, where 1_B denotes the characteristic function of the set B. Clearly, $f_1 \in D$ and $E(f_1) < \infty$. Define

$$\sigma = r_0 - \frac{m\alpha + \beta}{1 - \alpha}, \quad M = \frac{1}{2}\sigma + \frac{m + \beta}{1 - \alpha},$$

and $$r = \sigma + \frac{m + \beta}{1 - \alpha}.$$

From inequality (14) it follows that

$$\int_{|x| \geq r} P^n f_1(x)\,\mathrm{d}x \leq \frac{M}{r} \quad \text{for } n \geq n_0(f_1).$$

Thus,

$$\begin{aligned} \int_{|x| \leq r} P^n f(x)\,\mathrm{d}x &\geq \tfrac{1}{2} \int_{|x| \leq r} P^n f_1(x)\,\mathrm{d}x \\ &= \tfrac{1}{2}\left\{1 - \int_{|x| \geq r} P^n f_1(x)\,\mathrm{d}x\right\} \\ &\geq \tfrac{1}{2}[1 - M/r] > 0 \quad \text{for } n \geq n_0(f_1). \end{aligned} \tag{17}$$

Now we may write

$$\begin{aligned} P^n f(x) &= \int_{R^d} P^{n-1} f(y) g(x - S(y))\,\mathrm{d}y \\ &\geq \int_{|y| \leq r} P^{n-1} f(y) g(x - S(y))\,\mathrm{d}y. \end{aligned} \tag{18}$$

Define $\varepsilon = (1 - \alpha)\sigma$. If $|x - z_0| < \varepsilon$ and $|y| \leq r$, then

$$\begin{aligned} |(x - S(y)) - z_0| &\leq |x - z_0| + |S(y)| \\ &\leq |x - z_0| + \alpha|y| + \beta \\ &< \varepsilon + \alpha r + \beta = r_0. \end{aligned}$$

Hence, according to inequality (15) we have

$$g(x - S(y)) > 0 \quad \text{for } |x - z_0| < \varepsilon \text{ and } |y| \le r. \tag{19}$$

From (17) and (19) it follows that for every x satisfying $|x - z_0| < \varepsilon$ the product

$$P^{n-1}f(y)g(x - S(y)) \quad \text{for } n \ge n_0(f_1) + 1,$$

as a function of y, does not vanish in the ball defined by $|y| \le$ r. As a consequence, inequality (18) implies (16) with

$$A = \{x\colon |x - z_0| < \varepsilon\} \quad \text{and} \quad n_1(f) = n_0(f_1) + 1.$$

Thus, the proof of the theorem is complete.

5. Some special cases

In Theorems 1 and 2 of the two previous sections, it was assumed that the transformation S is defined on the entire space R^d, which is quite unrestrictive in that further specifications of the nature of the perturbing vectors ξ_n are not required. However, if the transformation S is defined only on a subset $G \subset R^d$, then it is quite possible that, for some $x_{n_0} \in G$, the point

$$x_{n_0+1} = S(x_{n_0}) + \xi_{n_0}$$

may not belong to G and as a consequence $S(x_{n_0+1})$ may not be defined. Should this be the case, the sequence $\{x_n\}$ cannot be calculated for $n > n_0 + 1$.

However, in some special cases that may be important from an applications point of view these difficulties may be easily circumvented. We discuss three such special situations, which are not exhaustive but do serve as good illustrations of how one may proceed under such circumstances. In our discussion we will always assume that S is Borel measurable, and this assumption will *not* be repeated.

Case 1. Assume that S maps R^d_+ into itself and that $\xi_n \ge 0$, $n = 0, 1, 2, \ldots$, with probability one. Here, R_+ denotes $[0, \infty)$. In this case, for every initial $x_0 \ge 0$, the sequence $\{x_n\}$ is well defined with probability one. Thus we may proceed with all of our arguments and calculations precisely as in Theorems 1 and 2, noting simply that in the definition of the operator P given by eq. (5) the domain of integration must be altered so that

$$Pf(x) = \int_{A(x)} f(y)g(x - S(y))\,\mathrm{d}y, \tag{20}$$

where

$$A(x) = \{y \in R^d_+\colon x - S(y) \in R^d_+\}.$$

Therefore, having S: $R^d_+ \to R^d_+$ which satisfies inequality (2), and the sequence $\{\xi_n\}$ with $\xi_n \ge 0$ all having the same density g with a finite first moment (inequality (3)), Theorem 1 is immediately applicable. If, in addition, $g(x) > 0$ for a sufficiently large subset of R^d_+, e.g., if

$$g(x) > 0 \quad \text{for } x \in R^d_+,\ |x| \le r_0 \\ \text{with } r_0 > \sqrt{d}(m\alpha + \beta)/(1 - \alpha),$$

then Theorem 2 holds.

There is, however, another way to view this case. Thus, we could consider that S is indeed defined on the entire space R^d (by assuming, for example, that $S(x) = 0$ for $x \notin R^d_+$) and that $f_0(x) = 0$ for $x \notin R^d_+$. Then all of the successive f_n have the same property, and this situation is merely a special case of the more general situation with S defined on R^d.

Case 2. Now consider the situation where S maps an interval $[0, a]$ into itself, and S is such that

$$\sup S(x) = b < a.$$

Now we may consider the sequence $\{x_n\}$ defined by eq. (1), assuming that $x_0 \in [0, a]$ and $0 \le \xi_n \le$

$a - b$ with probability one. In this case, Theorem 1 holds since inequality (2) is automatically satisfied with $\alpha = 0$ and $\beta = b$, and inequality (3) holds due to the fact that the space is a finite interval. The operator P is again given by eq. (20) wherein the domain of integration $A(x)$ is given by

$$A(x) = \{ y \in [0, a]: x \geq S(y) \}.$$

For Theorem 2 to hold it is sufficient that $g(x) > 0$ on a sufficiently large subset of $[0, a]$, e.g., $g(x) > 0$ a.e. on $[0, a - b]$ with $b < \frac{1}{2}a$.

Case 3. Consider a transformation S that maps the d-dimensional torus T^d into itself. [Recall that T^d may be obtained from R^d (as a quotient space) if we identify all points x, $y \in R^d$ such that $x - y$ is a sequence of integers.] In this case, as in the previous one, inequalities (2) and (3) are trivially satisfied and the arguments in the proof of Theorem 1 proceed exactly as written. Thus, assuming that we have S: $T^d \to T^d$ and $\{\xi_n\}$ independent with values in T^d and all having the same density g: $T^d \to R$, we obtain Theorem 1. As before, to obtain Theorem 2 we need only assume that $g(x) > 0)$ on a sufficiently large subset of T^d, e.g., $g(x) > 0$ a.e. on T^d.

Theorem 1 implies that, for a very broad class of transformations, the addition of a stochastic perturbation will cause the limiting densities to become asymptotically periodic. For some transformations, this would not be at all surprising, e.g. the addition of a small stochastic perturbation to a transformation with an exponentially stable periodic orbit gives asymptotic periodicity. However, the surprising content of Theorem 1 is that even in a transformation S that has aperiodic limiting behaviour, the addition of noise will result in asymptotic periodicity.

This phenomenon is rather easy to illustrate numerically by considering

$$x_{n+1} = S(x_n) \pmod 1, \tag{21}$$

where $S(x) = \alpha x + \lambda$, $0 < \alpha < 1$ and $0 < \lambda < 1$.

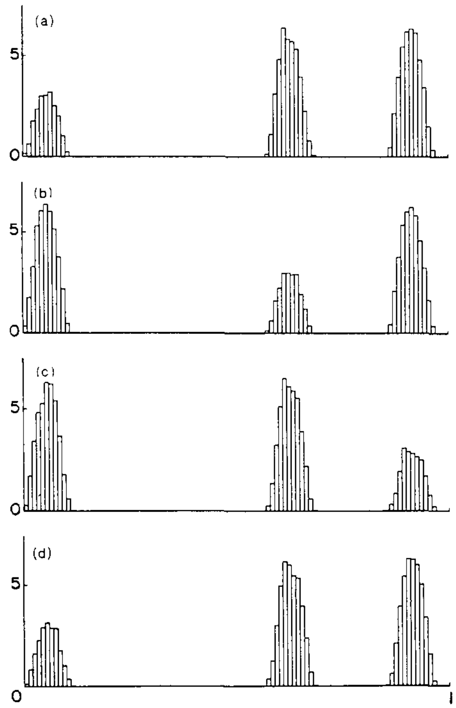

Fig. 1. Asymptotic periodicity illustrated. Here we show the histograms obtained after iterating 10^4 initial points uniformly distributed on $[0,1]$ with $\alpha = \frac{1}{2}$, $\lambda = \frac{17}{30}$, and $\theta = \frac{1}{5}$ in eq. (23). (a) $n = 10$; (b) $n = 11$; (c) $n = 12$; and (d) $n = 13$. The correspondence of the histograms for $n = 10$ and $n = 13$ indicate that, with these parameter values, numerically the sequence of densities has period 3.

This transformation is an example of a class of transformations considered by Keener [12]. From the results for general Keener transformations, for $\alpha \in (0,1)$ there exists an uncountable set Λ such that for each $\lambda \in \Lambda$ the rotation number corresponding to the transformation (21) is irrational. For each such λ the sequence $\{x_n\}$ is not periodic and the invariant limiting set

$$\bigcap_{k=0}^{\infty} S^k([0,1]) \tag{22}$$

is a Cantor set. The proof of Keener's general result offers a constructive technique for numerically determining values of λ that approximate elements of the set Λ.

From our remarks above (see Case 3), the transformation (21) clearly satisfies the conditions of Theorem 1 and is, therefore, an ideal candidate to illustrate the induction of asymptotic periodicity by noise in a transformation whose limiting behaviour is neither periodic nor asymptotically periodic in the absence of noise.

To be specific, we pick $\alpha = 1/2$ and use the results of Keener to show that $\lambda = 17/30$ is close to a value in the set Λ for which the invariant limiting set (22) should be a Cantor set. Asymptotic periodicity is illustrated by studying

$$x_{n+1} = (\alpha x_n + \lambda + \xi_n) \pmod 1 \tag{23}$$

where the ξ_n are random numbers uniformly distributed on $[0, \theta]$. Fig. 1 shows the numerically calculated effect of this stochastic perturbation for $\theta = 1/15$ and these values of α and λ. Figs. 1a through 1d, respectively, show the histograms obtained after 10, 11, 12, and 13 successive iterations of 10000 initial points uniformly distributed on [0, 1]. The 13th histogram is identical with the 10th, as is the 14th with the 11th, etc., thus indicating that numerically the sequence of densities is asymptotically periodic with period 3.

6. A generalization

To this point we have considered a relatively special class of situations in which perturbations were added to consecutive values of $S(x_n)$. This could be extended by considering 'multiplicative noise' with $x_{n+1} = S(x_n)\xi_n$ or, even more generally,

$$x_{n+1} = S(x_n, \xi_n). \tag{24}$$

Instead, however, we choose to describe a different process characterized by randomly applied perturbations. We assume that, in general, our system evolves according to a given transformation $S(x_n)$. The qualifying phrase 'in general' means that the transition $x_n \to S(x_n)$ occurs with probability $(1 - \varepsilon)$. In addition, with probability ε, the value of x_{n+1} is uncertain. If $x_n = y$ is given then, in this case, x_{n+1} may be considered as a random variable distributed with a density $K(x, y)$ which depends on y.

Our first goal in the description of this process is the derivation of an equation for the operator P which gives the prescription for passing from the density f_n of x_n to the density f_{n+1} of x_{n+1}. We assume that S maps a Borel measurable set $G \subset R^d$, $\mu(G) > 0$, into itself and that S is nonsingular. [Recall that S: $G \to G$ is nonsingular if S is measurable and such that

$$\mu(A) = 0 \quad \text{implies } \mu(S^{-1}(A)) = 0$$
$$\text{for all measurable } A \subset G.]$$

The requirement that S is nonsingular allows us to introduce the Frobenius-Perron operator P_S which describes the evolution of densities by the transformation S. The operator P_S is given implicitly by

$$\int_A P_S f(x)\,\mathrm{d}x = \int_{S^{-1}(A)} f(x)\,\mathrm{d}x \tag{25}$$

for every Borel subset A of G. The Radon-Nikodym theorem guarantees that for every $f \in L^1(G)$ there exists a unique $P_S f \in L^1(G)$ such that (25) holds.

Now assume that the density f_n of x_n is given and that a Borel set $A \subset G$ is given. We would like to calculate the probability that $x_{n+1} \in A$. As we outlined the randomly applied perturbation process, x_{n+1} may be reached in one of the two ways: deterministically with probability $(1 - \varepsilon)$ and stochastically with probability ε. Thus, in the deterministic case $x_{n+1} = S(x_n)$ and

$$\mathrm{Prob}_{\mathrm{I}}(x_{n+1} \in A) = \mathrm{Prob}_{\mathrm{I}}(S(x_n) \in A), \tag{26}$$

where the index I is used to denote the deterministic case. From the definition of the Frobenius–Perron operator, the density of $S(x_n)$ is $P_S f_n$ and, as a consequence,

$$\mathrm{Prob}_{\mathrm{I}}(S(x_n) \in A) = \int_A P_S f_n(x)\,\mathrm{d}x. \tag{27}$$

If the stochastic perturbation occurs and if $y = x_n$ then

$$\mathrm{Prob}_{\mathrm{II}}\,(x_{n+1} \in A | x_n = y) = \int_A K(x, y)\,\mathrm{d}x.$$

Since x_n is a random variable with density f_n, we also have

$$\mathrm{Prob}_{\mathrm{II}}\,(x_{n+1} \in A) = \int_G \mathrm{Prob}_{\mathrm{II}}\,(x_{n+1} \in A | x_n = y) f_n(y)\,\mathrm{d}y,$$

and combining this relation with the previous one we have

$$\mathrm{Prob}_{\mathrm{II}}\,(x_{n+1} \in A) = \int_A \left\{ \int_G K(x, y) f_n(y)\,\mathrm{d}y \right\} \mathrm{d}x. \tag{28}$$

From eqs. (26)–(28) we have

$$\begin{aligned}\mathrm{Prob}\,(x_{n+1} \in A) &= (1-\varepsilon)\,\mathrm{Prob}_{\mathrm{I}}\,(x_{n+1} \in A) + \varepsilon\,\mathrm{Prob}_{\mathrm{II}}\,(x_{n+1} \in A) \\ &= \int_A \left[(1-\varepsilon) P_S f_n(x) + \varepsilon \int_G K(x, y) f_n(y)\,\mathrm{d}y \right] \mathrm{d}x.\end{aligned}$$

Since A is an arbitrary Borel set this relation implies that the density f_{n+1} exists whenever f_n exists, and is given by

$$f_{n+1}(x) = (1-\varepsilon) P_S f_n(x) + \varepsilon \int_G K(x, y) f_n(y)\,\mathrm{d}y.$$

Thus, the expression for the operator P that describes the evolution of densities by our process is

$$Pf(x) = (1-\varepsilon) P_S f(x) + \varepsilon \int_G K(x, y) f(y)\,\mathrm{d}y. \tag{29}$$

Since $K(x, y)$ with fixed y is a density, it satisfies

$$K(x, y) \geq 0 \quad \text{and} \quad \int_G K(x, y)\,\mathrm{d}x = 1.$$

These two conditions in conjunction with the requirement that K: $G \times G \to R$ as a function of the two variables is measurable means that K is a stochastic kernel. We assume that $K(x, y)$ is uniformly integrable in x, i.e., for every $\eta > 0$ there is a $\delta > 0$ such that

$$\int_A K(x, y)\,\mathrm{d}x \leq \eta$$

for every $y \in G$ and A such that $\mu(A) \leq \delta$. Finally, we assume that

$$\int_G |x| K(x, y)\,\mathrm{d}x \leq \alpha|y| + \beta \quad \text{for } y \in G, \tag{30}$$

where α and β are non-negative constants and $\alpha < 1$. Note that condition (30) is automatically satisfied if G is bounded. However if G is unbounded, for example if G is the entire space R^d, then condition (30) is quite important since it prevents divergence of trajectories to infinity.

Then we have:

Theorem 3. If S: $G \to G$ is nonsingular and satisfies inequality (2) and K: $G \times G \to R$ is a uniformly integrable stochastic kernel satisfying (30), then with $0 < \varepsilon \leq 1$ the operator P given by (29) is weakly constrictive.

Proof. To simplify our notation we set $P_0 = P_S$ and

$$P_1 f(x) = \int_G K(x, y) f(y)\,\mathrm{d}y.$$

Furthermore, set $\varepsilon_0 = 1 - \varepsilon$ and $\varepsilon_1 = \varepsilon$. Then $P = \varepsilon_0 P_0 + \varepsilon_1 P_1$, and as a consequence

$$P^n f = \sum \varepsilon_{i_1} \varepsilon_{i_2} \cdots \varepsilon_{i_n} P_{i_1} P_{i_2} \cdots P_{i_n} f,$$

where the summation is taken over all possible sequences $(i_1 \cdots i_n)$ such that $i_k = 0$ or 1.

Next we define the set $\mathscr{F}_0$ to be all non-negative functions $f \in L^1$ such that

$$\|f\| \leq 1 \quad \text{and} \quad \int_G |x| f(x)\, \mathrm{d}x \leq \frac{M}{r} \quad \text{for } r > 0,$$

where $M > \beta/(1-\alpha)$ is a fixed constant. Using inequalities (2) and (30) it is easy to verify that $P_i(\mathscr{F}_0) \subset \mathscr{F}_0$, $i = 0, 1$, and as a consequence of the fact that $\mathscr{F}_0$ is convex, $P^n(\mathscr{F}_0) \subset \mathscr{F}_0$.

With $\mathscr{F}_0$ thus defined, for $f \in \mathscr{F}_0$ we set

$$\bar{P}^n f = \sum{}' \varepsilon_{i_1} \varepsilon_{i_2} \cdots \varepsilon_{i_n} P_{i_1} P_{i_2} \cdots P_{i_n} f, \tag{31}$$

where the prime (′) on the summation indicates that the term $\varepsilon_0^n P_0^n f$ is omitted. With these operators $\bar{P}^n$ defined from (31), we define a sequence of subsets in L^1 by

$$\mathscr{F}_n = \bar{P}^n(\mathscr{F}_0), \quad n = 1, 2, \ldots .$$

It is clear that $\mathscr{F}_n \subset \mathscr{F}_0$. From the sequence $\mathscr{F}_n$ we define $\mathscr{F}$ by

$$\mathscr{F} = \bigcup_{n=1}^{\infty} \mathscr{F}_n$$

taking care to note that $n = 0$ is not included. We are going to prove that the operator P is weakly constrictive, i.e. that condition (6) is fulfilled and that $\mathscr{F}$ is weakly precompact.

To demonstrate that the set $\mathscr{F}$ satisfies (6) is relatively easy. Since P is a Markov operator, it is sufficient to demonstrate that (6) holds for a dense subset $D_0 \subset D$. We take D_0 to be the set of all densities with a finite first moment, i.e.,

$$\int_G |x| f(x)\, \mathrm{d}x < \infty .$$

Proceeding as in the proof of Theorem 1, for every $f \in D_0$ we have $P^n f \in \mathscr{F}_0$ if n is sufficiently large. Therefore, it is sufficient to verify (6) for $f \in \mathscr{F}_0$. However, if $f \in \mathscr{F}_0$, then

$$P^n f - \varepsilon_0^n P_0^n f = \bar{P}^n f \in \mathscr{F}_n$$

and, as a consequence,

$$\rho(P^n f, \mathscr{F}) \leq \varepsilon_0^n,$$

so (6) is verified for $\mathscr{F}$.

To show that $\mathscr{F}$ is weakly precompact is much more difficult. However, note that since $\mathscr{F}_n \subset \mathscr{F}_0$, and $\mathscr{F} \subset \mathscr{F}_0$ as a consequence, to prove the weak precompactness of $\mathscr{F}$ it is sufficient to show that all functions $f \in \mathscr{F}$ are uniformly integrable [5]. Thus for every $\eta > 0$ pick $n_0 = n_0(\eta)$ such that

$$\varepsilon_0^{n_0} \leq \eta/2$$

and consider the set

$$\hat{\mathscr{F}}_{n_0} = \bigcup_{k=1}^{n_0} \left\{ \sum{}' \left(\varepsilon_{i_1} \cdots \varepsilon_{i_k} P_{i_1} \cdots P_{i_k} \mathscr{F}_0 \right) \right\}.$$

Note that the primed summation is a summation of sets of functions (not just functions), and that all of the terms in the primed summation contain the operator $P_1 = K$. From the uniform continuity of K it easily follows that the set $\hat{\mathscr{F}}_{n_0}$ is weakly precompact. Thus given $n_0(\eta)$ we may choose a $\delta(\eta) > 0$ such that for all $\hat{f} \in \hat{\mathscr{F}}_{n_0}$

$$\int_A \hat{f}(x)\, \mathrm{d}x \leq \eta/2 \tag{32}$$

for every $A \subset G$ satisfying $\mu(A) \leq \delta(\eta)$. Now fix $\eta > 0$ and pick an arbitrary $f \in \mathscr{F}$. Then there exists an integer $n \geq 1$ such that $f \in \mathscr{F}_n$. Consider two cases.

Case 1. $n \leq n_0$. In this case $\mathscr{F}_n \subset \hat{\mathscr{F}}_{n_0}$ and inequality (32), with $\hat{f} = f$, is satisfied.

Case 2. $n > n_0$. For this case set $f = \bar{P}^n g$ for some $g \in \mathscr{F}_0$. Then $f = \bar{P}^n g$ may be written in the form

$$f = \sum{}^1 \varepsilon_{i_1} \cdots \varepsilon_{i_n} P_{i_1} \cdots P_{i_n} g$$
$$+ \sum{}^2 \varepsilon_{i_1} \cdots \varepsilon_{i_n} P_{i_1} \cdots P_{i_n} g,$$

where the superscript 1 indicates that the summation is taken over all possible $(i_1 \cdots i_n)$ having indices $i_1 = \cdots = i_{n_0} = 0$, except $i_1 = \cdots = i_n = 0$, and the superscript 2 indicates that the summation contains all of the remaining terms. Since $g \in \mathscr{F}_0$ we have $P_{i_{n_0+1}} \cdots P_{i_n} g \in \mathscr{F}_0$ and it is easy to verify that Σ^2 is an element of the set $\hat{\mathscr{F}}_{n_0}$. Further, a simple calculation shows that

$$\left\| \sum{}^1 \varepsilon_{i_1} \cdots \varepsilon_{i_n} P_{i_1} \cdots P_{i_n} g \right\| \le \varepsilon_0^{n_0}.$$

Thus we may write $f = f_1 + f_2$ with $f_2 \in \hat{\mathscr{F}}_{n_0}$ and $\|f_1\| \le \eta/2$. From the foregoing and inequality (32) with $\hat{f} = f_2$ it is clear that

$$\int_A f(x)\,\mathrm{d}x \le \|f_1\| + \int_A f_2(x)\,\mathrm{d}x \le \eta.$$

Therefore, for arbitrary $f \in \mathscr{F}$ and $\eta > 0$ we have

$$\int_A f(x)\,\mathrm{d}x \le \eta$$

whenever $\mu(A) \le \delta(\eta)$ and, as a consequence, all $f \in \mathscr{F}$ are uniformly integrable. Hence $\mathscr{F}$ is weakly precompact and the proof is complete.

Now we are going to show that Theorem 3 is a generalization of Theorem 1, or more precisely that the process $\{x_n\}$ described by eq. (1) is a special case of the process with randomly applied perturbations. To achieve this we adopt what might appear to be a paradoxical approach, eliminating the deterministic portion of (29) and then reintroducing a deterministic process in the remaining term related to the random perturbation. The first essential point to note is that the value $\varepsilon_0 = 0$ in Theorem 3 is not excluded. Thus we may assume that $\varepsilon_0 = 0$ and $\varepsilon_1 = 1$ which means that the trajectory always evolves as a random walk. As in section 1, assume that we have a sequence $\{\xi_n\}$ of independent random variables with a common density g of the distribution and a transformation S: $R^d \to R^d$. Further assume that the random walk from $x_n = y$ to x_{n+1} is given by $x_{n+1} = S(y) + \xi_n$. In this case x_{n+1}, subject to the condition $x_n = y$, has the conditional density $g(x - S(y))$ and our eq. (29), with $K(x, y) = g(x - S(y))$, becomes identical with (5).

More generally if we have a given $\{\xi_n\}$ and a function S: $R^d \times R^d \to R$ of two variables, then we may assume that the system goes from $x_n = y$ to $x_{n+1} = S(y, \xi_n)$. Denote by $K(x, y)$ the density of $S(y, \xi_n)$. This density always exists if $S(y, z)$ as a function of z is nonsingular. If this is the case then eq. (29) with $\varepsilon_0 = 0$, $\varepsilon_1 = 1$ describes the evolution of densities corresponding to (24).

7. Discussion

Within the context of specific application, it would be of interest to know how the period of $P^n f$, as guaranteed by Theorems 1 and 3, depends on the perturbation applied to the system. A partial answer to this question may be obtained from eq. (8). For example, if $\{x_n\}$ is described by eq. (1) and the system is considered on the unit circle T^1, then we have

$$P^n f(x) = \int_{T^1} g(x - S(y)) P^{n-1} f(y)\,\mathrm{d}y$$
$$\le \gamma \int_{T^1} P^{n-1} f(y)\,\mathrm{d}y = \gamma,$$

where $\gamma = \sup g$ and this evaluation is valid for every density f. In particular, if $f = g_i$ then $P^n f = g_{\omega^n(i)}$ which implies that $g_i \le \gamma$. Since the g_i have disjoint supports, we also have

$$r = \int_{T^1} \sum_{i=0}^{r} g_i\,\mathrm{d}x \le \gamma, \tag{33}$$

and thus the period of the asymptotically periodic densities is equal to or less than $\gamma!$. If the noise amplitude is small, i.e., the ξ_n are small, then $\sup g$ is large and vice versa. Thus from (33) it follows that large noise amplitudes correspond to shorter periods of the asymptotically periodic sequence of densities. In particular, if $\gamma < 2$ then

$r=1$ (since r must be an integer) and the sequence $P^n f$ will, in this case, be asymptotically stable.

Though we discussed the unit circle for simplicity, in the general case the situation is much the same. However, the difference is that $\gamma = \sup g$ must be replaced by

$$\gamma = \int_G F(x)\,\mathrm{d}x,$$

where $F(x)$ is an asymptotic upper bound for the sequence of densities, i.e., F satisfies

$$P^n f \le F(x) + \varepsilon_n(x), \quad \lim_{n\to\infty} \|\varepsilon_n\| = 0.$$

It may be easily proved that for every constrictive operator P such a function exists, and F is strongly related to the properties of the set $\mathscr{F}$ in the spectral decomposition theorem.

Finally we would like to note that from eq. (8) it follows that any system in the presence of noise is quantized from a statistical point of view. Thus if n is large, which physically means that we have observed the system longer than the relaxation time, then

$$P^n f \approx \sum_{i=1}^{r} \lambda_i(y) g_{\omega^n(i)}.$$

Asymptotically $P^n f$ is either equal to one of the pure states g_i or to a mixture of these states, each having the weight λ_i.

References

[1] A. Boyarsky, Randomness implies order, J. Math. Anal. Applic. 76 (1980) 483–497.

[2] A. Boyarsky, On the significance of absolutely continuous invariant measures, Physica 11D (1984) 130–146.

[3] J.P. Crutchfield and B.A. Huberman, Fluctuations and the onset of chaos, Phys. Lett. A77 (1980) 407–410.

[4] J.P. Crutchfield, M. Nauenberg and J. Rudnick, Scaling for external noise at the onset of chaos, Phys. Rev. Lett. 46 (1981) 933–935.

[5] N. Dunford and J.T. Schwartz, Linear Operators, Part I: General Theory (Wiley, New York, 1957).

[6] L. Glass, C. Graves, G.A. Petrillo and M.C. Mackey, Unstable dynamics of a periodically driven oscillator in the presence of noise, J. Theor. Biol. 86 (1980) 455–475.

[7] R. Graham, Exact solution of some discrete stochastic models with chaos, Phys. Rev. A28 (1983) 1679–1691.

[8] J.E. Hirsch, B.A. Huberman and D.J. Scalapino, Theory of intermittency, Phys. Rev. A25 (1982) 519–532.

[9] J.E. Hirsch, M. Nauenberg and D.J. Scalapino, Intermittency in the presence of noise: A renormalization group formulation, Phys. Lett. A87 (1982) 391–393.

[10] J.L. Kaplan and J.A. Yorke, Numerical solution of a generalized eigenvalue problem for even mappings. Lect. Notes Math. 730 (1979) 228–237.

[11] C.F.F. Karney, A.B. Rechester and R.B. White, Effect of noise on the standard mapping, Physica 4D (1982) 425–438.

[12] J.P. Keener, Chaotic behavior in piecewise continuous difference equations, Trans. Amer. Math. Soc. 261 (1980) 589–604.

[13] Ju.I. Kifer, On the limiting behaviour of invariant measures of small random perturbations for some smooth dynamical systems, Dokl. Akad. Nauk. SSSR 216 (1974) 339–369.

[14] Ju.I. Kifer, On small perturbations of some smooth dynamical systems, Math. USSR-Izu. 8 (1974) 1083–1107.

[15] J. Komornik, Asymptotic periodicity of the iterates of weakly constrictive Markov operators, Tohoku Math. J. 38 (1986) 15–27.

[16] A. Lasota, T.Y. Li and J.A. Yorke, Asymptotic periodicity of the iterates of Markov operators, Trans. Amer. Math. Soc. 286 (1984) 751–764.

[17] A. Lasota and M.C. Mackey, Probabilistic Properties of Deterministic Systems (Cambridge Univ. Press, Cambridge, 1985).

[18] G. Mayer-Kress and H. Haken, Intermittent behavior of the logistic system, Phys. Lett. A82 (1981) 151–154.

[19] G. Mayer-Kress and H. Haken, Attractors of convex maps with positive Schwarzian derivative in the presence of noise, Physica 10D (1984) 329–339.

[20] A.B. Rechester, M.N. Rosenbluth, R.B. White and C.F.F. Karney, Statistical description of the Chirikov–Taylor model in the presence of noise, in: Long Time Prediction in Dynamics, C.W. Horton, L.E. Reichl and V.G. Szebehely, eds. (Wiley, New York, 1983), pp. 471–483.

[21] D. Ruelle, Small random perturbations of dynamical systems and the definition of attractors, Commun. Math. Phys. 82 (1981) 137–151.

[22] B. Shraiman, C.E. Wayne and P.C. Martin, Scaling theory for noisy period-doubling transitions to chaos. Phys. Rev. Lett. 46 (1981) 935–939.