*Journal of Statistical Physics, Vol. 69, Nos. 3/4, 1992*

# Asymptotic Theory of Multidimensional Chaos

**Sergey V. Ershov**[1]

*Received July 19, 1991; final May 15, 1992*

A delay-differential equation $\varepsilon\dot{u}(t)+u(t)=f(u(t-1))$, $0\leqslant t<\infty$, and its generalization are investigated in the limit $\varepsilon\to 0$, when the attractor's dimension increases infinitely. It is shown that a number of statistical characteristics are asymptotically independent of $\varepsilon$. As for the attractor, it can be regarded as a direct product of $O(1/\varepsilon)$ equivalent "subattractors," their statistical characteristics being asymptotically independent of $\varepsilon$. The results enable one to predict some characteristics of the attractor with fractal dimension $D\gg 1$ for the case $\varepsilon\ll 1$, when they are inaccessible numerically. The approach developed seems to be applicable for a wide class of spatiotemporal systems.

**KEY WORDS:** Chaos; delay-differential equation; domain structure; invariant distribution.

## 1. INTRODUCTION

In past decade a good deal of success has been achieved in the study of the attractors of low-dimensional dynamical systems. However, most such systems were derived as approximations to original infinite-dimensional models, so a number of questions arise, e.g.: Is it common that their properties can be satisfactorily described by a low-dimensional approximation, or do there exist models allowing no such approximation? How does the chaotic behavior become complicated as the dimension increases?

To study these questions a model is needed whose attractor's dimension increases infinitely as some control parameter is varied. The technique developed for low-dimensional dynamical systems analysis would be inapplicable to this model, and one could expect some "infinite-dimensional" features to reveal themself.

As such a model a differential-delay equation of the form

$$\varepsilon\dot{u}(t)+u(t)=f(u(t-1)), \qquad 0\leqslant t<\infty \tag{1.1a}$$

[1] Keldysh Institute for Applied Mathematics, Moscow 125047, Russia.



0022-4715/92/1100-0781$06.50/0 

can be used. The model arises in the field of nonlinear optics, medicine, control theory, etc.,[1–6] and is rather convenient for numerical analysis. Indeed, if $u_n(t)$ $(0 \leqslant t \leqslant 1)$ denotes the solution on the interval $[n, n+1]$, Eq. (1.1a) can be rewritten as

$$\varepsilon \dot{u}_{n+1}(t) + u_{n+1}(t) = f(u_n(t)), \qquad 0 \leqslant t \leqslant 1, \quad u_{n+1}(0) = u_n(1) \tag{1.1b}$$

Now, given the initial data, i.e., $u_0(t)$, one solves the linear equation (1.1b) and obtains $u_1(t)$, then substitutes it in the right-hand side of (1.1b) and obtains $u_2(t)$, and so on.

Usually the variable $t$ is regarded as a time. But the delay-differential equation, if written in the form (1.1b), allows another interpretation. As a matter of fact, this equation is a map $u_{n+1} = \hat{S}_\varepsilon[u_n]$ in the functional space, so it is convenient to regard $n$ as "an actual time," with $t$ as "a spatial coordinate" describing the profile of $u_n$ in a given "time" $n$. In other words, it is convenient to regard Eq. (1.1) as an evolution of functions $\{u_n(t)\}$ defined on the "space" sample $0 \leqslant t \leqslant 1$.

Many numerical experiments[1–6] indicate that the model exhibits chaotic behavior with its attractor's dimension growing as $O(1/\varepsilon)$. Nevertheless a number of statistical characteristics of the model, e.g., the scaled correlation function $C(\tau/\varepsilon)$, stationary distribution $p(u(t))$, Lyapunov spectrum, etc., are asymptotically independent of $\varepsilon$ (as $\varepsilon \to 0$); thus in this very complex behavior some simple laws are expected to be found. This paper presents such an asymptotic theory.

Along with the original model (1.1) we will consider its modification

$$\varepsilon \dot{u}_{n+1}(t) + u_{n+1}(t) = f(u_n(t)), \qquad 0 \leqslant t \leqslant 1, \quad u_{n+1}(0) = z_n \tag{1.2}$$

where $\{z_n\}$ is some random sequence, "noise." It is dubbed "a model with an external excitation," for in the absence of such an excitation (i.e., when $z_n = \text{const}$) any oscillations damp, the solution converging to a steady state $u^*(t)$:

$$u_n(t) \xrightarrow{n \to \infty} u^*(t)$$

So only due to this "external force" is a complex behavior (which does exist provided $\delta z \equiv \max z_n - \min z_n \neq 0$) possible. A similar "noise-sustained" chaos was observed by Deissler in a PDE.[18,19] A marvelous feature of the system is that the external excitation that gives rise to chaotic oscillations exerts almost no influence on their *characteristics*!

The model (1.2) thus can be said to be "an amplifier of chaos" as the "output amplitude" (at the point $t = 1$) appears to be considerably greater than the "input" one (at the point $t = 0$), that is, the amplitude

of the external excitation (see Figs. 7 and 8). But, in the case $\varepsilon \to 0$ this "amplifier" is also a "transformer" and a "restrictor," as the characteristics of its output $\{u_n(t)\}$ are almost independent of those of its input $\{z_n\}$. And, being almost independent of $\{z_n\}$, these characteristics converge to those of the basic model (1.1) as $\varepsilon$ goes to 0!

The theory advanced can be applied not only to the delay-differential equation, but to a wide class of models of spatiotemporal chaos, e.g., to coupled map lattices and reaction-diffusion systems, in which it predicts statistical characteristics with asymptotic independence of the area size. Also predicted is "forgetting boundary conditions": the influence of the latter exponentially damps and almost vanishes at internal points. As for the model with an external excitation, its analysis seems to be useful in studying noise-sustained chaos in open flow systems.

## 2. BASIC PROPERTIES OF THE MODEL

As said above, Eq. (1.1b) can be regarded as a map (in functional space) if one derives $u_{n+1}(t)$ from $u_n(t)$:

$$u_{n+1}(t) = (\hat{S}_\varepsilon u_n)(t) \equiv u_n(1)\, e^{-t/\varepsilon} + \int_0^t e^{(-t+\tau)/\varepsilon} f(u_n(\tau))\, d\tau/\varepsilon \qquad (2.1)$$

In the same way the model with external excitation leads to the "noisy" map

$$u_{n+1}(t) = z_n e^{-t/\varepsilon} + \int_0^t e^{(-t+\tau)/\varepsilon} f(u_n(\tau))\, d\tau/\varepsilon \qquad (2.2)$$

Hereafter we suppose that there exists an interval $[-C, C]$ mapped by $f$ into itself: $f([-C, C]) \subseteq [-C, C]$.

### 2.1. Boundedness

Let $|u_n(t)| \leqslant C$; then $|f(u_n(t))| \leqslant C$, from which it follows that [see (2.1)]

$$|u_{n+1}(t)| \leqslant Ce^{-t/\varepsilon} + C(1 - e^{-t/\varepsilon}) = C; \qquad |\dot{u}_{n+1}(t)| \leqslant 2C/\varepsilon$$

Hence, there exists a bounded set in $\mathbb{C}^{(1)}$ space which is mapped into itself by the operator $\hat{S}_\varepsilon$. In other words, the infinite-dimensional map (2.1) possesses a compact attractor.

The same conclusions hold for map (2.2), too, provided $|z_n| \leqslant C$.

Thus, the solutions of both models along with their derivatives are bounded, and $|\dot{u}_n| \leqslant O(1/\varepsilon)$. But perhaps the latter estimate is too excessive and actually $\lim_{\varepsilon \to 0} \varepsilon\, |\dot{u}_n| = 0$? Were this so, it would be possible to obtain an approximate model for the case $\varepsilon \to 0$ by discarding the term $\varepsilon \dot{u}$, thus

reducting the original system to a lattice of uncoupled 1D maps: $u_{n+1}(t) = f(u_n(t))$, $0 \leqslant t \leqslant 1$.

The possibility of doing so depends on the features of the map $x_{n+1} = f(x_n)$. Namely, experiments show that if this map possesses a stable limit cycle, the solution $\{u_n(t)\}$ proves to be meanderlike [i.e., approximately constant except for some tiny intervals of length $O(\varepsilon)$, where $|\dot{u}_n|$ is of order $1/\varepsilon$]. That in most of the interval $0 \leqslant t \leqslant 1$ the term $\varepsilon \dot{u}_n(t)$ is $\ll 1$ implies that the approximation $u_{n+1}(t) = f(u_n(t))$ proves to be valuable, its use leading to interesting results.[2,3] Provided some conditions are satisfied, it is possible to prove that the existence of a stable limit cycle in the sequence $x_{n+1} = f(x_n)$ results in the existence of a stable periodic solution in Eq. (1.1).

The situation is absolutely different if the map $x_{n+1} = f(x_n)$ is chaotic, i.e., its Lyapunov exponent $\lambda > 0$ (this is the case when the attractor dimension increases infinitely as $\varepsilon$ goes to 0). Really, were $|\dot{u}_n|$ bounded (in the limit $\varepsilon \to 0$), the term $\varepsilon \dot{u}_n(t)$ would be negligible, thus the equation would reduce to

$$u_n(t) = f(u_{n-1}(t)) = \cdots = f^n(u_0(t))$$

$$|\dot{u}_n(t)| = |(f^n)'(u_0(t))| \cdot |\dot{u}_0(t)|$$

From the definition of the Lyapunov exponent it follows that $|(f^n)'(x)| = O(e^{\lambda n})$ for $n$ large enough and most points $x$, which leads to an estimate $|\dot{u}_n(t)| = O(e^{\lambda n})\,|\dot{u}_0(t)|$ holding for $n \gg 1$. Hence, the derivative $|\dot{u}_n(t)|$ grows as $O(e^{\lambda n})$ (at least at most points of the interval $0 \leqslant t \leqslant 1$), which contradicts boundedness of the solution derivative. So the supposition that in the limit $\varepsilon \to 0$ the term $\varepsilon \dot{u}_n(t)$ is negligible is wrong. In other words, it means that $|\dot{u}_n|$ cannot grow more slowly than $1/\varepsilon$ and, because it also cannot grow faster, it has to grow precisely as $O(1/\varepsilon)$.

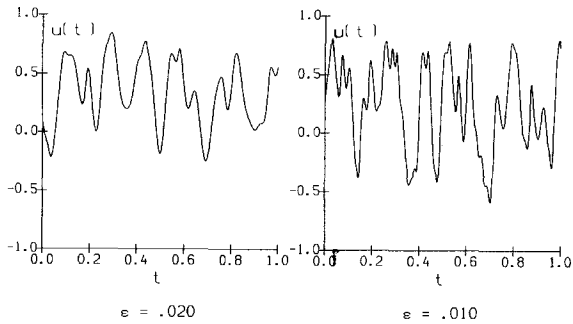



Fig. 1. $u_n(t)$ in the model (1.1) for $\varepsilon = 0.02$ (left) and $\varepsilon = 0.01$ (right).

In this work $f(x)$ of the form $f(x)=1-a|x|$ was chosen, which, for $a>1$ ($a$ is $a=1.9$ throughout the paper) is chaotic because $|f'|>1$. Therefore $|\dot{u}_n|=O(1/\varepsilon)$, and the smaller the $\varepsilon$, the more ragged the solution (see Fig. 1); the same holds for the model with an external excitation (Fig. 2). This implies that the integration step should be very small, and a simple reasoning indicates that "small" means "not greater than $\varepsilon$." To obtain this estimate, let us approximate the integral in (2.2) by a piecewise constant formula, which is equivalent to integrating Eq. (1.2) by the Euler method:

$$u_{n+1}(t_i)=\alpha^i z_n+\frac{h}{\varepsilon}\sum_{j=1}^{i}\alpha^{i-j}f(u_n(t_j)),\quad \alpha=e^{-h/\varepsilon},\ t_i=(i-1)h \tag{2.3}$$

If $z_n=\text{const}$, the original equation (1.2) has a stable fixed point $u^*(t)$, i.e., in this case

$$u_n(t)\xrightarrow{n\to\infty}u^*(t)$$

which will be proved in Section 4. Having linearized the transformation (2.3) in the vicinity of the fixed point, we obtain the Jacobi matrix

$$\frac{\partial(u_{n+1}(t_1),\ldots,u_{n+1}(t_N))}{\partial(u_n(t_1),\ldots,u_n(t_N))}=\alpha\frac{h}{\varepsilon}\begin{pmatrix} f_1' & 0 & 0 & 0 & 0 & 0\\ \alpha f_1' & f_2' & 0 & 0 & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ \alpha^{N-1}f_1' & \cdots & \cdots & \cdots & \cdots & f_N'\end{pmatrix}$$

It can be easily seen that its eigenvalues are $\alpha(h/\varepsilon)f_i'$, and so, if $\alpha(h/\varepsilon)\max|f'|<1$, all of them are less than 1. Therefore, if $h\leqslant\varepsilon/\max|f'|$, the discrete equation (2.3) also may have a stable fixed point.

On the contrary, if one uses the integration step $h>\varepsilon/\max|f'|$, then all eigenvalues are $>1$, and thus the fixed point (if exists) is at any rate

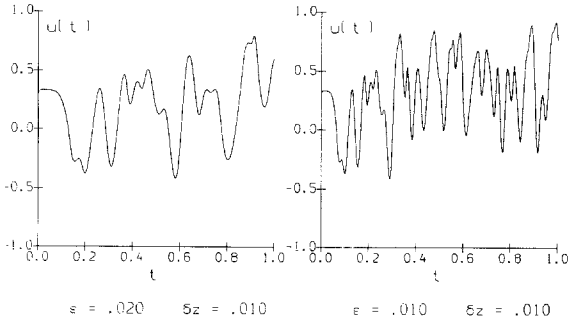



Fig. 2. $u_n(t)$ in the model (1.2) for $\varepsilon=0.02$ (left) and $\varepsilon=0.01$ (right). In both cases the excitation is uniformly distributed noise with amplitude $\delta z=0.01$.

unstable. So the discrete transformation (2.3) is unable to mimic even the qualitative properties of the original equation.

This means that for Eq. (2.2) to be integrated correctly (by the Euler method) the step should be less than $\varepsilon/\max|f'|$. If one uses more sophisticated methods (that of Adams, etc.) an insignificant factor of about $O(1)$ will precede the estimate.

Since the models (1.2) and (1.1) have many common properties, one may expect the condition $h \leqslant \varepsilon O(1)$ to hold for the latter, too. Numerical experiments have confirmed this and showed that $h = (0.1\text{–}0.2)\varepsilon$ is enough for integration, further decreasing it only wastes CPU time, the computational accuracy increasing insignificantly.

## 2.2. Decay of Correlations

As the integrals in both (2.1) and (2.2) contain the rapidly damping factor $\exp[(-t+\tau)/\varepsilon]$, which may be neglected for $\tau < t - r\varepsilon$ [within the accuracy $O(e^{-r})$], one obtains for $t > r\varepsilon$:

$$u_{n+1}(t) = \int_0^r e^{-\xi} f(u_n(t-\varepsilon\xi))\, d\xi + O(e^{-r}) \tag{2.4}$$

for both models (1.1) and (1.2). Hereafter $r = 5$ will be used for convenience in the evaluations.

To put it in another way, Eq. (2.4) means that $u_{n+1}(t)$ is determined by $u_n(\tau)$ on the interval $t - 5\varepsilon \leqslant \tau \leqslant t$ only, the latter in turn being determined by $u_{n-1}(\tau')$ on the interval $t - 10\varepsilon \leqslant \tau' \leqslant t$, and so on. Therefore the autocorrelation function

$$C(\tau) \equiv \frac{\langle u(t)\, u(t+\tau)\rangle - \langle u(t)\rangle^2}{\langle u^2(t)\rangle - \langle u(t)\rangle^2} \tag{2.5}$$

should be sufficiently nonzero only for $\tau \cong 0, 1, 2, \ldots$ (Fig. 3). Note that here it is more convenient to define the autocorrrelation function of the "continuous" solution $u(t)$, $0 \leqslant t < \infty$, rather than of the sequence $\{u_n(t)\}$, $0 \leqslant t \leqslant 1$. One can easily find that $C(\tau)$ for $0 \leqslant \tau \leqslant 1$ corresponds to the correlation between $u_n(t)$ and $u_n(t+\tau)$, while for $1 < \tau \leqslant 2$ it corresponds to that between $u_n(t)$ and $u_{n+1}(t + [\tau - 1])$, and so on. Such a behavior of the autocorrelation function is a first confirmation of the idea that the interval $0 \leqslant t \leqslant 1$ partitions into $O(1/\varepsilon)$ domains of length $O(\varepsilon)$, the dynamics on different domains being almost independent.

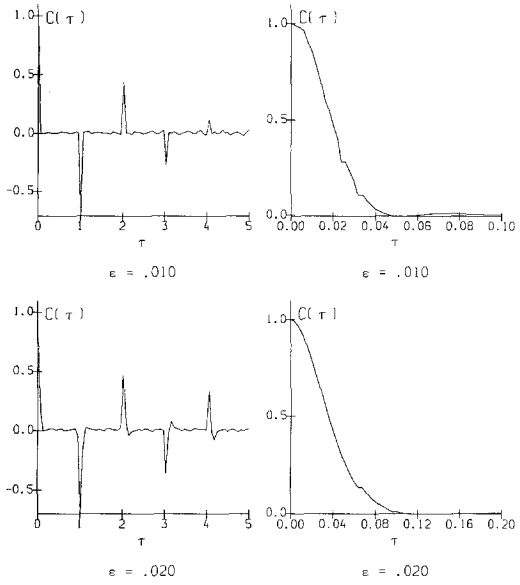



Fig. 3. Correlation function $C_\varepsilon(\tau)$ in the model (1.1) for $\varepsilon = 0.01$ (top) and $\varepsilon = 0.02$ (bottom). Right column shows the scaled initial part of $C_\varepsilon(\tau)$.

## 2.3. The Model's Statistical Properties: Preliminary Formulas

For our further deductions to be more convenient, let us make some agreement on the notations. Namely, the solution on the $n$th iteration, being "the argument" of the transformation (2.1)–(2.2), will be denoted as $v(t)$, while that on $(n+1)$th iteration will be denoted as $u(t)$ [just as it is sometimes more convenient to write $y = f(x)$ instead of $x_{n+1} = f(x_n)$].

Then, the sets of the associated lattice function values will be denoted as $\mathbf{u} = \{u(t_1),..., u(t_N)\}$ and $\mathbf{v} = \{v(t_1),..., v(t_N)\}$, and, finally, let $u_i = u(t_i)$ and $v_i = v(t_i)$ (without an argument and vector sign) denote the values of the corresponding lattice functions at the point $t_i = (i-1)h = (i-1)/(N-1)$.

Hereafter we will consider finite-dimensional truncation of the original map, i.e., the $N$-dimensional transformation $\{v_1,..., v_N\} \mapsto \{u_1,..., u_N\}$, $u_i \equiv u(t_i)$, $i = 1,..., N$. Note that the particular form of this map depends on the actual method of truncation and is of no significance for further deductions.

Since in the original map $u(t)$ for $t > 5\varepsilon$ is determined by $v(t')$ on the interval $t - 5\varepsilon \leqslant t' \leqslant t$ [see Eq. (2.4)] accurately enough, so it will be in the difference equation: in the latter $u(t_i)$ is determined by $\{v(t_{i-L}),..., v(t_i)\}$,

$L = 5\varepsilon/h$. We will choose (for convenience to be revealed below) $h = \varepsilon/2$; therefore $L = 10$, and, within the accuracy $O(\alpha^{10}) = O(e^{-5})$ we have

$$\begin{aligned} u_1 &= v_N \\ u_2 &= F_2(v_1, v_2) + \alpha v_N, \qquad \alpha = e^{-h/\varepsilon} \\ &\vdots \\ u_i &= F(v_{i-10},..., v_i), \qquad i = 11,..., N \end{aligned} \tag{2.6}$$

or, in a brief form, $\mathbf{u} = F(\mathbf{v})$.

As was already mentioned, the form of the functions $F$, $F_i$ depends on the difference scheme used, and is of no importance for us. But that for a given scheme the functions $F_i$ (for $i > 10$) are *the same* [within the accuracy $O(e^{-5})$] is a very significant fact resulting from the independence of the kernel in the integral (2.5) of $t$. Moreover, if one keeps the relation $h/\varepsilon$ fixed, the functions $F_1,..., F_{10}$ and $F$ are independent of $\varepsilon$ (i.e., of $N$); it is only the number of variables that varies while the connections between them do not: $u_i = F(v_{i-10},..., v_i)$, for $i \geqslant 11$.

The map $\mathbf{u}_{n+1} = F(\mathbf{u}_n)$ gives rise to a transformation of probability distributions $p_N^{(n+1)}(\mathbf{u}) = (\hat{\mathscr{L}}_N p_N^{(n)})(\mathbf{u})$, where $p_N^{(n)}(\mathbf{u})$ denotes the distribution of $\{\mathbf{u}_n\}$ on $n$th iteration, and $\hat{\mathscr{L}}_N$ is the Frobenius–Perron operator (FPO),[9–12]

$$\begin{aligned} p_N^{(n+1)}(\mathbf{u}) &= (\hat{\mathscr{L}}_N p_N^{(n)})(\mathbf{u}) \\ &\equiv \int \delta(\mathbf{u} - \mathbf{F}(\mathbf{v}))\, p_N^{(n)}(\mathbf{v})\, d\mathbf{v} \\ &= \int \delta(u_1 - v_N) \cdot \cdots \cdot \delta(u_N - F(v_{N-10},..., v_N)) \\ &\quad \times p_N^{(n)}(\mathbf{u})(v_1,..., v_N)\, dv_1 \cdots dv_N \end{aligned} \tag{2.7}$$

Sometimes the operator $\hat{\mathscr{L}}_N$ is written in the form

$$(\hat{\mathscr{L}}_N p_N)(\mathbf{u}) = \sum_{\forall \mathbf{v}:\ \mathbf{F}(\mathbf{v}) = \mathbf{u}} p_N(\mathbf{v}) / |\mathrm{Det}(D\mathbf{F}(\mathbf{v}))|$$

The Krylov–Bogolubov theorem states that if the iterated map $\mathbf{u}_{n+1} = F(\mathbf{u}_n)$ possesses a compact attractor (and it does possess one, as we have seen above), it possesses an invariant probability density $p_N(\mathbf{u})$. It is the density which does not change during the iterations, that is, it obeys the equation:

$$p_N(\mathbf{u}_N) = (\hat{\mathscr{L}}_N p_N)(\mathbf{u})$$

called the Frobenius–Perron equation, which is widely used in ergodic theory.[9,12]

The particular form of the map (2.6) allows us to reduce the "full" Frobenius–Perron operator (2.7) so as to derive the distribution $p_k^{(n+1)}(u_{N-k},..., u_N)$ of $k+1$ variables $\{u_{N-k},..., u_N\}$ only from $p_{k+10}^{(n)}(u_{N-k-10},..., u_N)$. The partial distribution $p_k^{(n+1)}$ is

$$p_k^{(n+1)}(u_{N-k},..., u_N) \equiv \int p_N^{(n+1)}(\mathbf{u})\, du_1 \cdots du_{N-k-1}$$

and substituting in this integral $p_N^{(n+1)} = \hat{\mathscr{L}}_N p_N^{(n)}$, one obtains

$$\begin{aligned}
&p_k^{(n+1)}(u_{N-k},..., u_N) \\
&\quad \equiv \int p_N^{(n+1)}(\mathbf{u})\, du_1 \cdots du_{N-k-1} \\
&\quad = \int (\hat{\mathscr{L}}_N p_N^{(n)})(\mathbf{u})\, du_1 \cdots du_{N-k-1} \\
&\quad = \int \left( \int \delta(u_1 - v_N) \cdot \cdots \cdot \delta(u_N - F(v_{N-10},..., v_N)) \cdot p_N^{(n)}(\mathbf{v})\, d\mathbf{v} \right) \\
&\qquad \times du_1 \cdots du_{N-k-1} \\
&\quad = \int \delta(u_{N-k} - F(v_{N-k-10},..., v_{N-k})) \cdot \cdots \cdot \delta(u_N - F(v_{N-10},..., v_N)) \\
&\qquad \times \left( \int p_N^{(n)}(v_1,..., v_N)\, dv_1 \cdots dv_{N-k-11} \right) dv_{N-k-10} \cdots dv_N \\
&\quad \equiv \int \delta(u_{N-k} - F(v_{N-k-10},..., v_{N-k})) \cdot \cdots \cdot \delta(u_N - F(v_{N-10},..., v_N)) \\
&\qquad \times p_{k+10}^{(n)}(v_{N-k-10},..., v_N)\, dv_{N-k-10} \cdots dv_N
\end{aligned} \tag{2.7'}$$

which is equivalent to

$$\begin{aligned}
&p_k^{(n+1)}(u_{N-k},..., u_N) \\
&\quad = \int du_{N-k-10} \cdots du_{N-k-1} \int \delta(u_{N-k-10} - v_N) \times \cdots \\
&\qquad \times \delta(u_{N-k} - F(v_{N-k-10},..., v_{N-k})) \times \cdots \times \delta(u_N - (v_{N-10},..., v_N)) \\
&\qquad \times p_{k+1}^{(n)}(v_{N-k-10},..., v_N)\, dv_{N-k-10} \cdots dv_N
\end{aligned}$$

The second integral in this expression is easily recognized as nothing but the Frobenius–Perron operator $\hat{\mathscr{L}}_{k+10}$ of (2.7), so

$$p_k^{(n+1)}(u_{N-k},\ldots,u_N)=\int du_{N-k-10}\cdots du_{N-k-1}(\hat{\mathscr{L}}_{k+10}p_{k+10}^{(n)})$$
$$\times(u_{N-k-10},\ldots,u_N) \qquad \text{if} \quad N>k-10 \tag{2.8}$$

In absolutely the same way a second iteration of this formula can be obtained which reads

$$p_k^{(n+2)}(u_{N-k},\ldots,u_N)$$
$$=\int(\hat{\mathscr{L}}_{k+20}^2 p_{k+20}^{(n)})(u_{N-k-20},\ldots,u_N)\,du_{N-k-20}\cdots du_{N-k-1}$$

[having done the integration, one easily finds that it is nothing but the transformation (2.7′) applied twice]. Further generalization leads to

$$p_k^{(n+m)}(u_{N-k},\ldots,u_N)$$
$$=\int(\hat{\mathscr{L}}_{k+10m}^m p_{k+10m}^{(n)})(u_{N-k-10m},\ldots,u_N)\,du_{N-k-10m}\cdots du_{N-k-1} \tag{2.9}$$

It is remarkable that for $k<N-11$ this operator $\hat{\mathscr{L}}_k$ depends on $k$ only, thus being independent of $\varepsilon$ (or $N$) provided $h/\varepsilon$ is kept fixed (i.e., if $N\propto 1/\varepsilon$)!

When studying one-dimensional maps $x_{n+1}=f(x_n)$,[12,25] it was found that if the invariant density is absolutely continuous, the iterations of any smooth initial distribution $p^{(0)}$ exponentially converge to the invariant one:

$$\|\hat{\mathscr{L}}^m p^{(0)}-p\|=O(\kappa^m), \qquad \kappa=\text{const}<1$$

where $\hat{\mathscr{L}}$ is the FPO associated with the 1D map $x_{n+1}=f(x_n)$ and here and henceforth $\|\cdot\|_{L_1}$ denotes the $L_1$ norm. This implies that the iterates of two initial distribution quickly converge:

$$\|\hat{\mathscr{L}}^m(\tilde{p}^{(0)}-p^{(0)})\|=O(\kappa^m)$$

The same may be expected[2] to hold for the multidimensional map (2.6):

$$\|\hat{\mathscr{L}}_N^m(\tilde{p}_N^{(0)}-p_N^{(0)})\|=O(\kappa^m[N]) \tag{2.10}$$

[2] In the presence of noise, however small, the estimate can be proved rigorously (see, e.g., ref. 14).

Notice that these relations at least require that the corresponding measures have densities (we might work with generalized densities, regarding, e.g., the Dirac measure as possessing a density; but there is no appropriate norm in this case). Fortunately, we do not need that the *full* measure itself have a density—because throughout the work we actually only deal with its *finite-dimensional projections.* And if the spatial correlations decay, then these projections are absolutely continuous; see ref. 26. This is also justified by experiments: one- and two-point distributions are obviously continuous (see Figs. 4 and 5).

If these suppositions are satisfied, then (2.9) results in

$$\|p_k^{(m)} - p_k\| = \|\hat{\mathscr{L}}^m_{k+10m}(p_{k+10m} - p^{(0)}_{k+10m})\| = O(\kappa[k+10m]^m) \quad (2.11)$$

the estimate not including the total number of variables $N$ (or $\varepsilon$), as it describes convergence to the invariant distribution of a given part of variables.

To put it loosely, it indicates that the equilibration at distant points goes on independently.

Now a remark should be made on the dependence of "the convergence exponent" $\kappa$ on $N$ (or $\varepsilon$). As during one iteration of (2.1) [or (2.2)] the influence area of the "boundary condition" (imposed at the point $t=0$) moves right of about $5\varepsilon$, it would take, generally, $O(1/\varepsilon)$ iterations for an equilibrium state to settle throughout the interval $0 \leqslant t \leqslant 1$. Such a slow (in the limit $\varepsilon \to 0$) convergence would mean that $\lim_{\varepsilon \to 0} \kappa = 1$. But, in Eq. (1.1a) [and, therefore, in (1.1b)[3] and all its discrete derivatives] $u(t_i)$ behaves equally. So, if the initial distribution $p_{\text{init}}(u_1, ..., u_N)$ is such that all its arguments $u(t_i)$ are equal [that is, the function $p_{\text{init}}(u_1, ..., u_N)$ equally depends on its arguments], all parts of the interval $0 \leqslant t \leqslant 1$ will contain the same information; thus one need not wait till the influence of its left side $t=0$ reaches the domain $t=1$, which process would take $O(1/\varepsilon)$ iterations. Roughly speaking, at first a stationary distribution in each time domain of the length $\Delta t$ ($\varepsilon \ll \Delta t \ll 1$) will settle, and then the correlations between close domains will settle (the distant domain intercorrelations need not settle, due to their independence, as the autocorrelation function indicates)—and so the invariant distribution will settle simultaneously throughout the interval. Both processes take place on the scale $O(\varepsilon)$, so their "rate" is independent of $\varepsilon$ [that is clear if one substitutes a "global" time $t$ in (2.4) for a "local" one $t' = (t - t^*)/\varepsilon$, which results in $\varepsilon$ being excluded from the equation].

[3] The point $t=0$ [or $t=n$, $n=0, 1, ...$, in Eq. (1.1a)] appears to be artificially distinguished, because while dividing the time domain into the (fundamental) intervals $[n, n+1]$ it falls on their left end. As for the original Eq. (1.1a), no point is distinguished at all.

The physical explanation allows a mathematical formulation of which we will produce only the gist, omitting the details.

If spatial correlations between distant points decay quickly enough, then in the conditional distribution $P(u_i \mid u_{i+1},\ldots)$ all $u_j$ can be discarded for which $|i-j| \gg 1$. (For the sake of clarity only $u_{i+1}$ will be retained). It can be shown that the distribution $\rho_N(u_1,\ldots, u_N)$ (which should not be confused with the *invariant distribution* $p_N$) allows the following representation:

$$\rho_N(u_1,\ldots, u_N) = \int (P(u_1 \mid u_2) \cdot P(u_2 \mid u_3) \cdot \cdots \cdot P(u_N \mid u_{N+1}) \cdot \cdots)\, du_{N+1} \cdots$$

with the accuracy increasing as more factors follow $P(u_N \mid u_{N+1})$. Let us retain ten of them. Now let us consider two close distributions, $\rho_N(u_1,\ldots, u_N)$ and $\tilde{\rho}_N(u_1,\ldots, u_N)$ and denote $\tilde{P}(u_i \mid u_{i+1}) - P(u_i \mid u_{i+1}) \equiv \Delta P(u_i \mid u_{i+1}) \ll 1$. Then the difference $\|\tilde{\rho}_N - \rho_N\|_{L_1}$ can be evaluated as

$$\begin{aligned}
&\|\tilde{\rho}_N - \rho_N\|_{L_1} \\
&\quad = \int P(u_1 \mid u_2) \cdot \cdots \cdot P(u_{N+9} \mid u_{N+10}) \\
&\qquad \times \left| \left( \frac{\Delta P(u_1 \mid u_2)}{P(u_1 \mid u_2)} + \cdots + \frac{\Delta P(u_{N+9} \mid u_{N+10})}{P(u_{N+9} \mid u_{N+10})} \right) \right| du_1 \cdots du_{N+10} \\
&\quad \equiv (N+10)^{1/2} \int \rho_{N+10}(u_1,\ldots, u_{N+10}) \\
&\qquad \times \frac{\varphi(u_1, u_2) + \cdots + \varphi(u_{N+9}, u_{N+10})}{(N+10)^{1/2}}\, du_1 \cdots du_{N+10} \\
&\quad \equiv (N+10)^{1/2} \int \rho_{N+10}(u_1,\ldots, u_{N+10}) \\
&\qquad \times \delta\left( z - \frac{\varphi(u_1, u_2) + \cdots}{(N+10)^{1/2}} \right) |z|\, du_1 \cdots du_{N+10} \\
&\quad \equiv (N+10)^{1/2} \int |z|\, \Phi(z)\, dz
\end{aligned}$$

Applying to this expression the method of Fourier transformation used in the proof of the central limit theorem, we get an asymptotic (for $N \gg 1$) expression for the function $\Phi$:

$$\Phi(z) = \frac{1}{R} e^{-(z/2R)^2}, \qquad R \equiv \sqrt{\int \left( \frac{\Delta P(u_1 \mid u_2)}{P(u_1 \mid u_2)} \right)^2 \rho_2(u_1, u_2)\, du_1\, du_2}$$

Thus

$$\|\tilde{\rho}_N - \rho_N\|_{L_1} = C(N+10)^{1/2} \sqrt{\int \left(\frac{\Delta P}{P}\right)^2 \rho_2(u_1, u_2)\, du_1\, du_2}$$

so for two different $N$'s (both $\gg 1$) we get

$$\|\tilde{\rho}_N - \rho_N\|_{L_1} = (N/N')^{1/2}\, \|\tilde{\rho}_{N'} - \rho_{N'}\|_{L_1}$$

It obviously holds for $\hat{\mathscr{L}}^m_N \rho_N$, $\hat{\mathscr{L}}^m_N \tilde{\rho}_N$, $\hat{\mathscr{L}}^m_{N'} \rho_{N'}$, and $\hat{\mathscr{L}}^m_{N'} \tilde{\rho}_{N'}$, which are distributions of the same type as $\rho_N$, $\tilde{\rho}_N$, $\rho_{N'}$, and $\tilde{\rho}_{N'}$, hence

$$\|\hat{\mathscr{L}}^m_N \tilde{\rho}_N - \hat{\mathscr{L}}^m_N \rho_N\|_{L_1} / \|\tilde{\rho}_N - \rho_N\|_{L_1} = \|\hat{\mathscr{L}}^m_{N'} \tilde{\rho}_{N'} - \hat{\mathscr{L}}^m_{N'} \rho_{N'}\|_{L_1} / \|\tilde{\rho}_{N'} - \hat{\rho}_{N'}\|$$

which means that the convergence exponent $\kappa$ [see (2.10)] is asymptotically independent of $N$.

All further considerations will concern only distributions whose arguments $u(t_i)$ act equally, the estimate (2.9) containing a convergence exponent $\kappa$ independent of $\varepsilon$ for such distributions.

## 3. STATISTICAL PROPERTIES OF THE SOLUTIONS OF (1.1) AND THEIR DOMAIN STRUCTURE IN THE CASE $\epsilon \to 0$

Let us denote a *stationary* distribution of the variables $\{u_{N-k}, \ldots, u_N\}$ [here $u_s = u(t_s) = u(s \cdot h)$] as $p_{\varepsilon,k}(u_{N-k}, \ldots, u_N)$, and compare it with $p_{\varepsilon',k}(w_{N'-k}, \ldots, w_{N'})$, where $\{w_i\}$ denote the variables in the map for $\varepsilon'$ ($\varepsilon' < \varepsilon$). Applying (2.9) to $p_{\varepsilon,k}$, which does not change during iterations, we get

$$p_{\varepsilon,k}(u_{N-k}, \ldots, u_N) = \int (\hat{\mathscr{L}}^m_{k+10m}\, p_{\varepsilon,k+10m})(u_{N-k-10m}, \ldots, u_N)\, du_{N-k-10m} \cdots du_{N-k-1}$$

the operator being independent of $\varepsilon$ and $N$ for $k < N - 10m$. Therefore $p_{\varepsilon',k}$ should obey the same equation:

$$p_{\varepsilon',k}(u_{N'-k}, \ldots, u_{N'}) = \int (\hat{\mathscr{L}}^m_{k+10m}\, p_{\varepsilon',k+10m})(u_{N'-k-10m}, \ldots, u_{N'})\, du_{N'-k-10m} \cdots du_{N'-k-1}$$

and using (2.10), one obtains

$$\|p_{\varepsilon,k} - p_{\varepsilon',k}\| = \|\hat{\mathscr{L}}^m_{k+10m}(p_{\varepsilon,k+10m} - p_{\varepsilon',k+10m})\| = O(\kappa^m)$$

The estimate holds if $k+10m<N-10$, i.e., $m\leqslant M=(N-k-11)/10$; thus

$$\|p_{\varepsilon,k}-p_{\varepsilon',k}\|=O(\kappa^{(N-k-11)/10})=O(\kappa^{O(N)})=O(\kappa^{1/\varepsilon}),\qquad \varepsilon'\leqslant\varepsilon \tag{3.1}$$

Certainly, this estimate holds for any $k+1$ variables $\{u_{j-k},\ldots,u_j\}$, as all $u_j$ are equal. We used $j=N$ only because $t_N=1$ in Eq. (1.1b) [or points $t=1, 2, 3,\ldots$ in (1.1a)] are artificially distinguished as end(s) of the fundamental intervals $[n, n+1]$. Obviously by adjusting this interval (e.g., $[\tau+n, \tau+n+1]$) one can make any $t_j$ to be such an "endpoint" instead of $t_N$.

It is worth mentioning that the exponent $\kappa$ in (3.2) is independent of $k$; see end of Section 2.

The computations confirm the conclusion that the invariant distribution is asymptotically independent of $\varepsilon$; see Figs. 4 and 5, where the one-point distribution $p(u(t))$ and contours of the two-point distribution $p(u(t), u(t+\tau))$ (for $\tau/\varepsilon=\text{const}$) are plotted.

In the very beginning a supposition was advanced about the attractor's domain structure for $\varepsilon\to 0$, that is, about a mutual independence (in a sense) of the solution $u_n(t)$ on distant intervals. Now we should define this mathematically. To do this, let us look at two groups (I and II) each

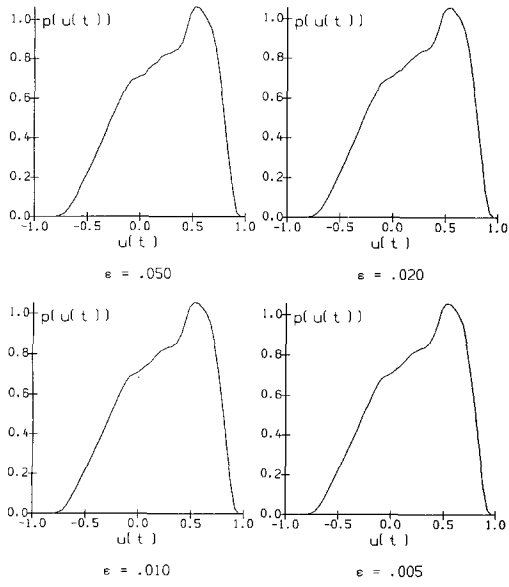



Fig. 4. Stationary one-point distribution $p_\varepsilon(u(t))$ in the model (1.1) for $\varepsilon=0.05$, 0.02, 0.01, and 0.005.

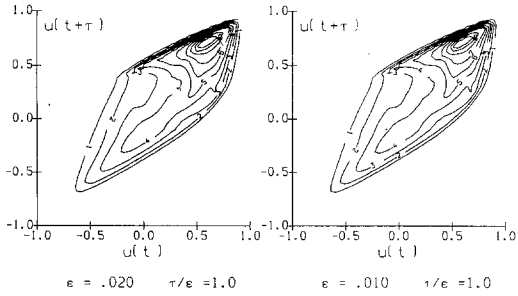
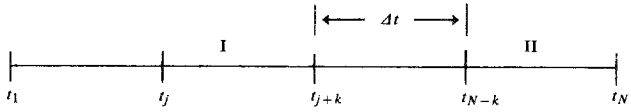



Fig. 5. Stationary two-point distribution $p_\varepsilon(u(t), u(t+\tau))$ in the model (1.1) for $\varepsilon = 0.02$ (left) and $\varepsilon = 0.01$ (right). In both cases $\tau = \varepsilon$.

of $k+1$ variables, $\{u_j, \ldots, u_{j+k}\}$ and $\{u_{N-k}, \ldots, u_N\}$, the associated intervals $[t_j, t_{j+k}]$ and $[t_{N-k}, t_N]$ being disjoint (i.e., $j+k<N-k$). It is these intervals that are dubbed domains; here domains of equal length ($=kh$) are chosen for simplicity.




Let

$$p^{(n)}_{k,j}(u_j, \ldots, u_{j+k}; u_{N-k}, \ldots, u_N) \equiv \int p^{(n)}_N(\mathbf{u})\, du_1 \cdots du_{j-1}\, du_{j+k+1} \cdots du_{N-k-1}$$

be invariant distributions of these variables on the $n$th iteration. Substituting in this integral $\hat{\mathscr{L}}_N p^{(n)}_N = p^{(n+1)}_N$ and transforming it in the way used in the derivation of (2.8), one gets for $p^{(n+1)}_{k,j}$

$$\begin{aligned} p^{(n+1)}_{k,j}(u_j, \ldots, u_{j+k}; u_{N-k}, \ldots, u_N) &= \int \delta(u_j - F(v_{j-10}, \ldots, v_j)) \cdot \cdots \cdot \delta(u_{j+k} - F(v_{j+k-10}, \ldots, v_{j+k})) \\ &\times \delta(u_{N-k} - F(v_{N-k-10}, \ldots, v_{N-k})) \cdot \cdots \cdot \delta(u_N - F(v_{N-10}, \ldots, v_N)) \\ &\times p^{(n)}_{k+10,\, j-10}(v_{j-10}, \ldots, v_{j+k}; v_{N-k-10}, \ldots, v_N) \\ &\times dv_{j-10} \cdots dv_{j+k}\, dv_{N-k-10} \cdots dv_N \end{aligned} \tag{3.2}$$

which is equivalent to

$$
\begin{aligned}
&p_{k,j}^{(n+1)}(u_j,\ldots,u_{j+k};u_{N-k},\ldots,u_N)\\
&\quad=\int du_{j-10}\cdots du_{j-1}\,du_{N-k-10}\cdots du_{N-k-1}\\
&\qquad\times\int\delta(u_{j-10}-v_N)\cdot\cdots\cdot\delta(u_{j-1}-F_{10}(v_{j-10},\ldots,v_{j-1})-\alpha^9 v_N)\\
&\qquad\times\delta(u_j-F(v_{j-10},\ldots,v_j))\cdot\cdots\cdot\delta(u_{j+k}-F(v_{j+k-10},\ldots,v_{j+k}))\\
&\qquad\times\delta(u_{N-k-10}-F(v_{j+k-9},\ldots,v_{j+k},v_{N-k-10}))\times\cdots\\
&\qquad\times\delta(u_{N-k-1}-F(v_{j+k},v_{N-k-10},\ldots,v_{N-k-1}))\\
&\qquad\times\delta(u_{N-k}-F(v_{N-k-10},\ldots,v_{N-k}))\cdot\cdots\cdot\delta(u_N-F(v_{N-10},\ldots,v_N))\\
&\qquad\times p_{k+10,j-10}^{(n)}(v_{j-10},\ldots,v_{j+k};v_{N-k-10},\ldots,v_N)\\
&\qquad\times dv_{j-10}\cdots dv_{j+k}\,dv_{N-k-10}\cdots dv_N
\end{aligned}
$$

The internal integral is recognized as the FPO $\hat{\mathscr{L}}_{2k+20}$ in which the variables are arranged as follows: $\{v_{j-10},\ldots,v_{j+k},v_{N-k-10},\ldots,v_N\}$:

$$
\begin{aligned}
&p_{k,j}^{(n+1)}(u_j,\ldots,u_{j+k};u_{N-k},\ldots,u_N)\\
&\quad=\int du_{j-10}\cdots du_{j-1}\,du_{N-k-10}\cdots du_{N-k-1}\\
&\qquad\times(\hat{\mathscr{L}}_{2k+20}\,p_{k+10,j-10}^{(n)})(u_{j-10},\ldots,u_{j+k};u_{N-k-10},\ldots,u_N)
\end{aligned}
\tag{3.3}
$$

In absolutely the same way we can derive the $m$th iteration of this formula:

$$
\begin{aligned}
p_{k,j}^{(n+m)}=\int&(\hat{\mathscr{L}}_{2k+10m}^{(m)}\,p_{k+10m,j-10m}^{(n)})\\
&\times du_{j-10m}\cdots du_{j-1}\,du_{N-k-10m}\cdots du_{N-k-1}
\end{aligned}
\tag{3.4}
$$

[having done the integration one finds that it is nothing but (3.2) applied $m$ times]. From (3.4) an estimate similar to (2.11) follows:

$$
\|p_{k,j}^{(m)}-p_{k,j}\|_{L_1}=O(\kappa^m)
\tag{3.5}
$$

Now let us take the initial distribution of the form

$$
\begin{aligned}
&p_{k',j'}^{(0)}(u_{j'},\ldots,u_{j'+k';}u_{N-k'},\ldots,u_N)\\
&\quad=p_{k'}(u_{j'},\ldots,u_{j'+k'})\cdot p_{k'}(u_{N-k'},\ldots,u_N)
\end{aligned}
$$

where $p_{k'}$ is the *invariant* distribution of $k+1$ variables, which due to the equality of different $t$'s is the same for $\{u_{j'},..., u_{j'+k'}\}$ and $\{u_{N-k'},..., u_{N'}\}$; and let $k'=k+10m$ and $j'=j-10m$ for some $m$ to be specified below. Then from (3.2) [which equivalent to (3.3)] it follows that

$$p^{(1)}_{k'-10,j'+10}(u_{j'+10},..., u_{j'+k'}; u_{N-k'+10},..., u_N)$$
$$= p_{k'-10}(u_{j'+10},..., u_{j'+k'}) \cdot p_{k'-10}(u_{N-k'+10},..., u_N)$$

Obviously the distribution on the $m$th iteration will also be "factorizable":

$$p^{(m)}_{k,j}(u_j,..., u_{j+k}; u_{N-k},..., u_N) = p_k(u_j,..., u_{j+k})\, p_k(u_{N-k},..., u_N)$$

(recall that $k=k'-10m$, $j=j'+10m$), and at the same time (3.5) results in $\|p^{(m)}_{k,j}-p_{k,j}\|_{L_1}=O(\kappa^m)$, thus $\|p_{k,j}-p_k\cdot p_k\|=O(\kappa^m)$. The choice of initial distribution allows any $m$ such that $N-k'>j'+k'$ (so that $p^{(0)}_{k',j'}$ corresponds to disjoint[4] domains). Recalling that $k'=k+10m$ and $j'=j-10m$, we obtain the restriction on $m$: $m\leqslant M=[(N-k)-(j+k)]/10$. The difference $\|p_{k,j}-p_k\cdot p_k\|$ is obviously independent of our choice of $m$, so we can take $m=M$ to get the best estimate:

$$\|p_{k,j}-p_k\cdot p_k\| = O(\kappa^m) = O\left(\exp\left(\frac{t_{N-k}-t_{j+k}}{10\varepsilon}\frac{\varepsilon}{h}\ln\kappa\right)\right)$$
$$= O\left(\exp\frac{-C\,\Delta t}{\varepsilon}\right) \tag{3.6}$$

Were the behavior on the domains mutually independent, the joint invariant distribution $p_{k,j}(u_j,..., u_{j+k}; u_{N-k},..., u_N)$ would be a product of partial ones: $p_{k,j}(\mathbf{u}_{\mathrm{I}}, \mathbf{u}_{\mathrm{II}})=p_k(\mathbf{u}_{\mathrm{I}})\cdot p_k(\mathbf{u}_{\mathrm{II}})$. But since this independence is not absolute, a discrepancy $\|p_{k,j}-p_k\cdot p_k\|$ arises that measures the degree of independence. The estimate (3.6) indicates that the spatial correlation decays exponentially.

So, if the interval $0\leqslant t\leqslant 1$ is considered as partitioned into $L$ domains, the, say, odd domains (separated from each other by an interval $\Delta t=1/L$) will be mutually independent within the accuracy $O(e^{-O(1/L\varepsilon)})$. The same holds for the even domains; and, though they will be correlated with odd ones (as being neighbors), the total number of "independent pieces" in the interval $0\leqslant t\leqslant 1$ is $O(L)$, the degree of independence being $O(e^{-O(1/L\varepsilon)})$. Note that if $L=O(1/\varepsilon)$, the domain length is $O(\varepsilon)$, so for $h\sim\varepsilon$ they contain the same number of variables, their statistical characteristics, according to (3.3), being almost independent of $\varepsilon$.

[4] We suppose that the domains are much closer to each other than to the point $t=0$, so that $j=j'-10m>10$ for such $m$.

Now, the main results of this section may be summarized as follows:

For a given accuracy $\delta > 0$ the interval $0 \leqslant t \leqslant 1$ can be regarded as consisting of $L = \text{const}/(\varepsilon \log \delta)$ domains, the dynamics on which are mutually independent within this accuracy [see (3.8)]. In other words, the original dynamical system looks like a product of $L$ equal, independent within this accuracy, dynamical systems, each located on its "own" domain. Moreover, the statistical characteristics of these systems are almost independent of $\varepsilon$ (within the accuracy $\delta \sim e^{-O(1/L\varepsilon)}$). As for the attractor, it looks (within the same accuracy!) like a direct product of $L$ subattractors associated with these "elementary" dynamical systems.

The higher is $L\varepsilon$ ($L$ is proportional to $1/\varepsilon$: $L = C/\varepsilon$, and may vary only through the prefactor $C$), the poorer is the accuracy $-e^{-O(1/L\varepsilon)}$, though the simpler are these "elementary" dynamical systems (their dimension is $\sim 1/\varepsilon L$). So, one can regard the original dynamical system (1.1) as a very simple object—a product of, say, $1000/\varepsilon$ elementary dynamical systems each of dimension of, say, $\sim 3$—but within a poor accuracy. If one wants to increase the accuracy, one should consider it as a more complex object—a product of, say, $100/\varepsilon$ elementary dynamical systems of a rather complex nature, their dimension being $\sim 30$, and so on.

As the properties of these elementary dynamical systems are independent of $\varepsilon$ (if $\varepsilon L$ is kept fixed), one can predict within some good accuracy the attractor's statistical characteristics in the case $\varepsilon \to 0$, when numerical investigation is impossible.

In addition, such a "domain structure" provides the dimension estimate $D = O(1/\varepsilon)$ consistent with Farmer's experiments.[1]

## 4. THE MODEL WITH AN EXTERNAL EXCITATION

This model is described by the equation

$$\varepsilon \dot{u}_{n+1}(t) + u_{n+1}(t) = f(u_n(t)), \qquad 0 \leqslant t \leqslant 1, \quad u_{n+1}(0) = z_n$$

or

$$\dot{u}_{n+1}(t) = z_n e^{-t/\varepsilon} + \int_0^t e^{-(t-\tau)/\varepsilon} f(u_n(\tau))\, d\tau/\varepsilon \tag{4.1}$$

and has chaotic solutions only due to the external excitation $z_n$, i.e., when its amplitude $\delta z \equiv \max z_n - \min z_n \neq 0$. At the same time, the model exhibits a marvelous feature: the statistical characteristics of $\{u_n(t)\}$ for $\varepsilon \ll 1$ and $t \gg \varepsilon$ are almost independent of $\varepsilon$, $t$, and the excitation characteristics.

## 4.0. Time Scaling

Let us denote $w_n(t) = u_n(\gamma t)$, $0 < \gamma < 1$. Then $w_n$ obeys a (4.1)-type equation

$$\mu \dot{w}_{n+1}(\tau) + w_{n+1}(\tau) = f(w_n(\tau)), \qquad 0 \leqslant \tau \leqslant \gamma, \quad w_n(0) = z_n$$

where $\mu = \varepsilon/\gamma > \varepsilon$. That is, the solution in "the initial interval" $0 \leqslant t \leqslant \gamma$ can be described by the same model, but with greater $\tilde{\varepsilon} = \varepsilon/\gamma$. To put it in another way, the initial part of the solution is only "compressed" by $\varepsilon$ decreasing (see Fig. 2).

## 4.1. Convergence to a Steady State in the Case $\delta z = 0$

First, let us prove this steady state does exist for the case $z_n = z_0 = \text{const}$. Obviously the stationary solution obeys the equation

$$u^*(t) = z_0 e^{-t/\varepsilon} + \int_0^t e^{-(t-\tau)/\varepsilon} f(u^*(\tau))\, d\tau/\varepsilon \equiv \hat{S}_{\varepsilon, z_0}[u^*]$$

that is, it is a fixed point of the operator $\hat{S}_{\varepsilon, z_0}$. It is easy to prove this operator maps the set

$$\mathbb{U} = \{u \mid \|u\|_{\mathbb{C}} \leqslant C, \ \|\dot{u}\|_{\mathbb{C}} \leqslant C/2\varepsilon, \ u(0) = z_0\}$$

into itself: $\hat{S}_{\varepsilon, z_0}[\mathbb{U}] \subseteq \mathbb{U}$; $C$ is a constant such that $f[-C, C] \subseteq [-C, C]$ (Section 2.1). The set $\mathbb{U}$ is convex and closed, and, in addition, compact in $\mathbb{C}$ due to the Arzela theorem, so Shauder's theorem yields the existence of a fixed point $u^* \in \mathbb{U}$, which is the sought for steady state: $u^* = \hat{S}_{\varepsilon, z_0}[u^*]$.

Now we consider its stability; due to the condition $u_n(0) = z_0$, the perturbations should obey $\delta u_n(0) = 0$. To do this, let us first consider the behavior of $u_n(t)$ on the interval $0 \leqslant t \leqslant \gamma$ for some small $\gamma$. It is described by the equation (see Section 4.0)

$$\mu \dot{w}_{n+1}(\tau) + w_{n+1}(\tau) = f(w_n(\tau)), \qquad 0 \leqslant \tau \leqslant \gamma, \quad w_n(0) = z_0$$

where $\mu = \varepsilon/\gamma$, $\tau = t/\gamma$, and $w_n(\tau) = u_n(\gamma\tau)$, the linearized equation governing perturbations being

$$\mu \delta \dot{w}_{n+1}(\tau) + \delta w_{n+1}(\tau) = f'(w^*(\tau))\, \delta w_n(\tau), \qquad 0 \leqslant \tau \leqslant \gamma, \quad \delta w_n(0) = 0$$

[since due to the conditon $u_n(0) = z_0$ the perturbations obey $\delta u_n(0) = 0$]; thus

$$\delta w_{n+1}(\tau) = \int_0^t e^{-(\tau - \tau')/\mu} f'(w^*(\tau'))\, \delta w_n(\tau')\, d\tau'/\mu$$

This obviously leads to the estimate

$$\|\delta w_{n+1}\|_C \leqslant \|\delta w_n\|_C \frac{\gamma \max |f'|}{\mu} \leqslant \|\delta w_n\|_C \frac{\gamma \max |f'|}{\varepsilon}$$

Hence for $\gamma \leqslant (\varepsilon/\max |f'|)^{1/2}$ the perturbations on the interval $0 \leqslant t \leqslant \gamma$ do damp, and

$$\lim_{n \to \infty} u_n(t) = u^*(t) \qquad \forall t \in [0, \gamma]$$

After the steady state $u^*$ has been settled on $[0, \gamma]$, one can consider the behavior on $[\gamma, 2\gamma]$. It is described by an equation of the same type as on $[0, \gamma]$:

$$\mu \dot{\tilde{w}}_{n+1}(\tau) + \tilde{w}_{n+1}(\tau) = f(\tilde{w}_n(\tau)), \qquad 0 \leqslant \tau \leqslant \gamma, \quad \tilde{w}_n(0) = u_n(\gamma) = u^*(\gamma)$$

[here $\tilde{w}_n(\tau) = u_n(\gamma + \gamma \cdot \tau)$]; the perturbations will obviously damp on this interval as well. So, after the solution has converged to the steady state on $[0, \gamma]$, the convergence on $[\gamma, 2\gamma]$ begins, and

$$\lim_{n \to \infty} u_n(t) = u^*(t) \qquad \forall t \in [0, 2\gamma]$$

and so on. The convergence is of a "propagating type," that is, the transient oscillations on the right end ($t \sim 1$) are still in effect while on the left end ($t \sim 0$) they already have damped.

## 4.2. "Spatial" Growth of the Oscillations Amplitude

Now let $\delta z \neq 0$, but $\ll 1$, so that the amplitude of the chaotic oscillations for $t \approx 0$ is very small. As said above, we can divide the whole interval $0 \leqslant t \leqslant 1$ into two subintervals, $0 \leqslant t \leqslant 1/2$ with a scaled solution on it denoted as $w_n(t') \equiv u_n(t'/2)$, $0 \leqslant t \leqslant 1$, and $1/2 \leqslant t \leqslant 1$ with one denoted as $\tilde{w}_n(t) \equiv u_n(1/2 + t/2)$:

$$2\varepsilon \dot{w}_{n+1}(t) + w_{n+1}(t) = f(w_n(t)), \qquad 0 \leqslant t \leqslant 1, \quad w_{n+1}(0) = z_n$$

$$2\varepsilon \dot{\tilde{w}}_{n+1}(t) + \tilde{w}_{n+1}(t) = f(\tilde{w}_n(t)), \qquad 0 \leqslant t \leqslant 1, \quad \tilde{w}_n(0) = u_n(\tfrac{1}{2}) = w_n(1)$$

So, the dynamics can be completely described by two coupled models with $\varepsilon' = 2\varepsilon$ with the first one's "output" [i.e., $w_n(1)$] being an excitation for the second one.

Numerical experiments show that for "small oscillations" the "output" amplitude is always greater than the "input" one. This can be written as

follows: if $Du(1) \ll 1$, then $Du(1) \geqslant \vartheta(\varepsilon)\, Du(0) = \vartheta(\varepsilon)\, Dz$, $\vartheta(\varepsilon) > 1$ $\{\sqrt{Du(t)}$ is an "average amplitude" at the point $t$: $Du(t) \equiv \langle [u_n(t) - \langle u_n(t) \rangle^2] \rangle^{1/2}$, where $\langle \cdot \rangle$ denotes an averaging over $n\}$. Obviously, then, $Dw(1) \geqslant \vartheta(2\varepsilon)\, Dw(0) = \vartheta(2\varepsilon)\, Du(0)$; $D\tilde{w}(1) \geqslant \vartheta(2\varepsilon)\, D\tilde{w}(0) = \vartheta(2\varepsilon)\, Dw(1)$ (see Fig. 6); thus, $Du(1) = D\tilde{w}(1) \geqslant [\vartheta(2\varepsilon)]^2\, Du(0)$. This means that $\vartheta(\varepsilon) = [\vartheta(2\varepsilon)]^2$ and in the same way it can be shown that $Du(1) \geqslant [\vartheta(k\varepsilon)]^k\, Du(0)$, which can be generalized:

$$Du(1) \geqslant [\vartheta(\varepsilon_0)]^{\varepsilon_0/\varepsilon}\, Du(0) \qquad \text{provided} \quad Du(1) \ll 1 \tag{4.2}$$

Hence $\vartheta(\varepsilon) = O(\beta^{1/\varepsilon}) \to \infty$ as $\varepsilon \to 0$.

Now note that $u_\varepsilon(t_0) = u_{\varepsilon/t_0}(1)$ (see Section 4.2). This leads to $Du_\varepsilon(t_0) = Du_{\varepsilon/t_0}(1)$, or

$$Du(t) \geqslant [\vartheta(\varepsilon_0)]^{t\varepsilon_0/\varepsilon}\, Du(0) = O(\beta^{t/\varepsilon})\, Dz \qquad \text{provided} \quad Du(t) \ll 1 \tag{4.3}$$

For large amplitudes this estimate ceases to hold, since, due to the boundedness of the solution, the dispersion $Du(t)$ cannot grow to infinity, so for $t/\varepsilon \gg 1$ a saturation occurs—i.e., $Du(t)$ ceases to depend on $t$, $\varepsilon$, or $\delta z$ (see Fig. 7 and 8). Moreover, it appears that the statistical characteristics of $\{u_n(t)\}$ for $t \simeq 1$, such as the probability distribution $p_{\varepsilon,k}$, are asymptotically independent of $\varepsilon$ and $\delta z$; see Figs. 7 and 8 with $p(u(1))$ $(\equiv p_{\varepsilon,0})$ plotted for different values of the parameters.

So the model acts as a "restricting amplifier" of a chaotic signal.

Let us note that the original model (1.1b) can be regarded as a model with external excitation, though of a special type: $z_n = u_n(1)$. Now the "output" is joined to the "input," the operation transforming the "amplifier" of chaos into a "generator" of it. And as the excitation properties do not influence substantially the solution characteristics (as experiments and some theoretical considerations below indicate), they should be almost the same in both models (1.1) and (1.2)—compare Fig. 4 with Figs. 7 and 8.

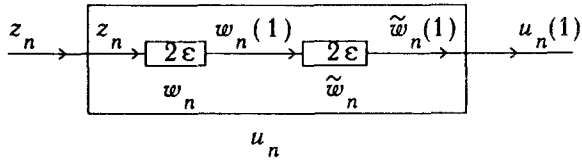



Fig. 6.

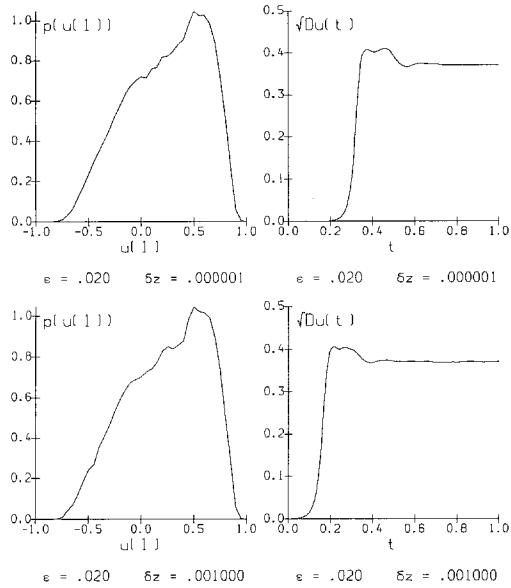



Fig. 7. Stationary one-point distribution $p_\varepsilon(u(1))$ in the model (1.2) (left column) and average amplitude of oscillations $\sqrt{Du(t)}$ vs. $t$ (right column). The excitation is uniformly distributed noise with amplitude $\delta z = 10^{-6}$ (top) and $\delta z = 10^{-3}$ (bottom). In both cases $\varepsilon = 0.02$.

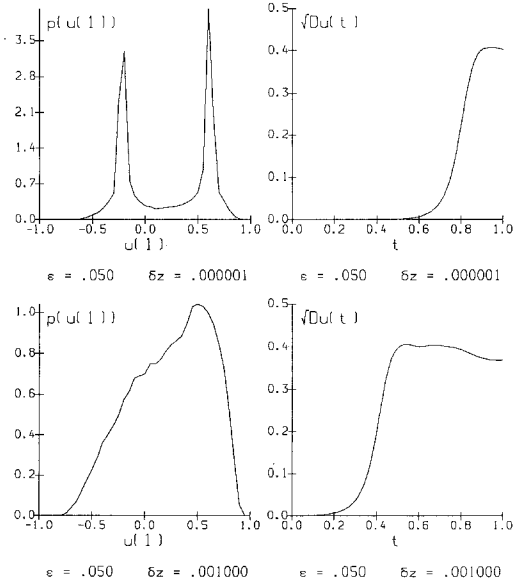



Fig. 8. Same as Fig. 7, but for greater $\varepsilon$: $\varepsilon = 0.05$. Left top panel shows that for very small $\delta z$ and $\varepsilon$ large enough the distribution may substantially differ from that in the original model; see Fig. 4.

## 4.3. Asymptotics of $p_{\epsilon,k}$ for $\epsilon \to 0$ in the Model with External Excitation

To study this, let us discretisize Eq. (4.1) as in Section 3, discarding terms smaller than $O(\alpha^{10}) = O(e^{-10h/\varepsilon}) = O(e^{-5})$:

$$
\begin{aligned}
u_{n+1}(t_1) &= z_n \\
u_{n+1}(t_2) &= F_2(u_n(t_1), u_n(t_2)) + \alpha z_n \\
&\vdots \\
u_{n+1}(t_{10}) &= F_{10}(u_n(t_1),..., u_n(t_{10})) + \alpha^9 z_n \\
u_{n+1}(t_i) &= F(u_n(t_{i-10}),..., u_n(t_i)), \qquad i = 11,..., N
\end{aligned}
$$

the only difference from (2.6) being the "boundary conditions": $u_n(t_N)$ is replaced by $z_n$, while the functions $F_i$, $F$ are the same as in (2.6). Unless the excitation is a pure noise, the dynamical equation for $z_n$ should be added. Here $z_{n+1}$ is supposed to depend on $z_n$ and $\mathbf{u}_n$ only with some "pure" noise added: $z_{n+1} = \Phi(z_n, \mathbf{u}_n) + \xi_n$. This allows one to describe a sheer random excitation if $z_n = \xi_n$ as well as a purely deterministic one if $z_{n+1} = \Phi(z_n)$. Finally, for $z_{n+1} = u_{n+1}(t_N) = F(u_n(t_{N-10}),..., u_n(t_N))$ even the original model [in the form (2.6)] can be described by these equations.

Thus, the equations governing the model are

$$
\begin{aligned}
u_{n+1}(t_1) &= z_n \\
u_{n+1}(t_2) &= F_2(u_n(t_1), u_n(t_2)) + \alpha z_n \\
&\vdots \\
u_{n+1}(t_{10}) &= F_{10}(u_n(t_1),..., u_n(t_{10})) + \alpha^9 z_n \\
u_{n+1}(t_i) &= F(u_n(t_{i-10}),..., u_n(t_i)), \qquad i = 11,..., N \\
z_{n+1} &= \Phi(z_n, u_n(t_1),..., u_n(t_N)) + \xi_n
\end{aligned}
\tag{4.4}
$$

Were $\xi$ a constant, the invariant distribution of the map (4.4) would obey the Perron–Frobenius equation:

$$
\begin{aligned}
p_\varepsilon(\mathbf{u}, z \mid \xi) = \int & \delta(u_1 - \varkappa) \cdot \delta(u_2 - F_2(v_1, v_2) - \alpha\varkappa) \\
& \times \delta(u_3 - F_3(v_1, v_2, v_3) - \alpha^2\varkappa) \\
& \times \delta(u_i - F(v_{i-10},..., v_i)) \cdot \cdots \cdot \delta(u_N - F(v_{N-10},..., v_N)) \\
& \times \delta(z - \Phi(\varkappa, v_1,..., v_N) - \xi) \\
& \times p_\varepsilon(\mathbf{v}, \varkappa \mid \xi)\, dv_1 \cdots dv_N\, d\varkappa
\end{aligned}
$$

When going on to the independent random $\xi$'s distributed with the density $H(\xi)$, one should simply average this equation over all $\xi$'s,[13]

$$p_\varepsilon(\mathbf{u}, z) = \int p_\varepsilon(\mathbf{u}, z \mid \xi)\, H(\xi)\, d\xi$$

So, for

$$p_{\varepsilon,k}(u_{N-k},\ldots, u_N) \equiv \int p_\varepsilon(\mathbf{u}, z)\, dz\, du_1 \cdots du_{N-k-1}$$

one obtains

$$\begin{aligned}
&p_{\varepsilon,k}(u_{N-k},\ldots, u_N) \\
&\quad \equiv \int p_\varepsilon(\mathbf{u}, z)\, dz\, du_1 \cdots du_{N-k-1} \\
&\quad = \int \delta(u_1 - \varkappa)\cdot\delta(u_2 - F_2(v_1, v_2) - \alpha\varkappa)\cdot\delta(u_3 - F_3(v_1, v_2, v_3) - \alpha^2\varkappa) \\
&\qquad \times \delta(u_i - F(v_{i-10},\ldots, v_i))\cdot\cdots\cdot\delta(u_N - F(v_{N-10},\ldots, v_N)) \\
&\qquad \times \delta(z - \Phi(\varkappa, v_1,\ldots, v_N) - \xi) \\
&\qquad \times p_\varepsilon(\mathbf{v}, \varkappa \mid \xi)\, H(\xi)\, dv_1 \cdots dv_N\, d\varkappa\, d\xi\, dz\, du_1 \cdots du_{N-k-1} \\
&\quad = \cdots \\
&\quad = \int \delta(u_{N-k} - F(v_{N-k-10},\ldots, v_{N-k}))\cdot\cdots\cdot\delta(u_N - F(v_{N-10},\ldots, v_N)) \\
&\qquad \times (p_\varepsilon(\mathbf{v}, \varkappa)\, d\varkappa\, dv_1 \cdots dv_{N-k-11})\, dv_{N-k-10} \cdots dv_N
\end{aligned}$$

which coincides with (2.7) and so is equivalent to

$$\begin{aligned}
p_{\varepsilon,k}(u_{N-k},\ldots, u_N) = \int (\hat{\mathscr{L}}^m_{k+10m}\, p_{\varepsilon,k+10m})(u_{N-k-10m},\ldots, u_N) \\
\times\, du_{N-k-10m} \cdots du_{N-k-1}
\end{aligned} \tag{4.5}$$

as when deriving (2.9).

Let $\mathbb{E}$ denote the set of excitation parameters ($\Phi$ and $H$ in the case at issue). If this excitation gives rise to a *chaotic* regime, then (2.11) is satisfied, and in quite the same way as the estimate (3.1) was derived, one gets

$$\| p_{\varepsilon,k,\mathbb{E}} - p_{\varepsilon',k,\mathbb{E}} \| = O(\kappa^{1/\max(\varepsilon,\varepsilon')}) \qquad \text{for} \quad \varepsilon, \varepsilon' \to 0 \tag{4.6}$$

As $\hat{\mathscr{L}}_k$ is independent of excitation, the same holds if it is the latter that is varied (so that both $\mathbb{E}$ and $\mathbb{E}'$ give rise to chaos):

$$\| p_{\varepsilon,k,\mathbb{E}} - p_{\varepsilon,k,\mathbb{E}'} \| = O(\kappa^{1/\varepsilon}) \qquad \text{for} \quad \varepsilon \to 0 \tag{4.7}$$

Moreover, as the operator $\hat{\mathscr{L}}_k$ in (2.8) and (4.5) is the same, we can compare $p_{\varepsilon,k,\mathbb{E}}$ with the corresponding distribution *for the original model* (1.1) denoted as $p_{\varepsilon,k}$:

$$\| p_{\varepsilon,k,\mathbb{E}} - p_{\varepsilon,k} \| = O(\kappa^{1/\varepsilon}) \qquad \text{for} \quad \varepsilon \to 0 \tag{4.8}$$

In contrast with the original equation (1.1), the model with an external excitation possesses a distinguished boundary $t = 0$, so different $t$'s are not equal and the statistical characteristics depend on $t$. To derive this dependence, let us consider a one-point distribution *at the point* $t$: $p_{\varepsilon,\mathbb{E}}(u(t)) = p_{\varepsilon,\mathbb{E}}(u \mid t)$. In absolutely the same way as was done in Section 4.2 for $Du_\varepsilon(t)$, one obtains that $p_{\varepsilon,\mathbb{E}}(u \mid t) = p_{\varepsilon/t,\mathbb{E}}(u \mid 1)$, so (4.8) reads

$$\| p_{\varepsilon,\mathbb{E}}(\cdot \mid t) - p_* \| = O(e^{Ct/\varepsilon}) \qquad \text{for} \quad \varepsilon/t \to 0 \tag{4.9}$$

where $p_*(u) = \lim_{\varepsilon \to 0} p_\varepsilon(u)$ is the limit distribution for the "boundless" model (1.1). A similar estimate holds for the two-point distribution $p_\varepsilon(u(t), u(t+\tau))$, and so on. In other words, the influence of the boundary on local (at a given point) statistical characteristics exponentially damps and almost vanishes at internal points.

If, on the contrary, even one of these excitations—say, $\mathbb{E}$—relates to a *periodic* regime, then the associated distribution will be $\delta$-functions [e.g., if $\delta z = 0$, then

$$p_{\varepsilon,k,\mathbb{E}}(u_{N-k}, \ldots, u_N) = \delta(u_{N-k} - u^*_{N-k}) \times \cdots \times \delta(u_N - u^*_N)$$

see Section 4.1] which do not converge under the FPO iterations. (Indeed, the $L_1$ distance between two almost all $\delta$-functions is 2.) Thus, in this case the estimate (2.9) ceases to hold, and so do (4.7)–(4.9).

However, in the presence of noise *any* initial distribution (even a $\delta$-like one) converges to the invariant distribution,[14] and hence (4.8) holds if both excitations include some amount of noise.

So: If $\varepsilon \to 0$, while one fixes both excitations, with both of them providing a chaotic behavior of $\{\mathbf{u}_n\}$ (for which any small noise is sufficient), then (2.10) holds (i.e., iterations converge), which implies (4.7)–(4.9). On the contrary, if $\varepsilon$ is fixed, and, say, one excitation is varied

in such a way that its amplitude $\delta z \to 0$, then, due to the continuous dependence of the distribution on noise,[15]

$$p_{\varepsilon,k,\mathbb{E}}(u_{N-k},..., u_N) \to \delta(u_{N-k} - u^*_{N-k}) \times \cdots \times \delta(u_N - u^*_N)$$

as $\delta z \to 0$ and (4.7)–(4.9) cannot hold.

The conclusions were confirmed by numerical results, where uniformly distributed values $z_n$, their mean square deviation being $\delta z$, were used as an excitation. The distribution $p(u(1))$ $(\equiv p_{\varepsilon,0})$ for different $\varepsilon$'s and $\delta z_n$ is plotted in Fig. 7 and 8. One can see, that, if $\delta z > 0$ is fixed while $\varepsilon \to 0$, it converges to the same limit distribution as that of the model (1.1); compare Fig. 4 with Figs. 7 and 8. But if $\varepsilon$ is fixed while $\delta z \to 0$, the distribution converges to another limit, as Fig. 8a indicates.

Unfortunately, computation of even the two-point distribution (to say nothing about the three-point one!) requires too large a sequence $\{u_n(1), u_n(1-\tau)\}$ for the statistics to be obtained. As for the original model (1.1), its uniformity in time enables one to reduce the calculations. Namely, since for any $t$, $\{u_n(t)\}$ are equally distributed, one can use all of them to obtain the statistics—i.e., to plot a histogram by

$$\{..., u_n(t_1),..., u_n(t_N);..., u_{n+1}(t_1),..., u_{n+1}(t_N);...\}$$

rather than by

$$\{...; u_n(t_N),...; u_{n+1}(t_N),...\}$$

This "small ruse" reduces the amount of (1.1b) iterations to be done approximately $N$ times,[5] thus rendering the computation of the two-point distributions in Fig. 5 rather a simple problem.

From the above consideration, an interesting conclusion on the original model's properties can be made. Consider a domain $[0, t_b]$ of the length $\Delta t = t_b \equiv \varepsilon/\mu$. The dynamics on it is governed by the following equation (see Section 4.0):

$$\mu \dot{w}_{n+1}(\tau) + w_{n+1}(\tau) = f(w_n(\tau)), \qquad w_{n+1}(0) = u_n(1), \quad 0 \leqslant \tau \leqslant 1 \quad (4.10)$$

where $\mu = \varepsilon/\Delta t$, $w_n(\tau) \equiv u_n(\tau\, \Delta t)$. In this equation $\mathbf{w}_{n+1}$ depends not only on $\mathbf{w}_n$, but on some *external* (with respect to $\mathbf{w}_n$) "boundary value" $u_n(1)$. So, (4.10) is a sort of model with external excitation, the latter being $\{u_n(1)\}$.

Now let us take two different values $\varepsilon$ and $\varepsilon' < \varepsilon$ and the domain length proportional to $\varepsilon$, so that $\Delta t/\varepsilon = \Delta t'/\varepsilon'$ and the parameter $\mu$ is the same in both cases. So, the only thing changing in (4.10) while replacing

[5] As each iteration adds $N$ values to the histogram at once.

$\varepsilon$ for $\varepsilon'$ and $\Delta t$ for $\Delta t'$ is the characteristics of the excitation $\{u_n(0)\}$. In this case with only excitation varying one can apply the estimate (4.9), which now reads

$$\|\rho_{\mu,\varepsilon}(\cdot \mid \tau) - \rho_{\mu,\varepsilon'}(\cdot \mid \tau)\| = O(e^{-C\tau/\mu}) \tag{4.11}$$

where $\rho_{\mu,\varepsilon}(w \mid \tau)$ is the distribution of $\{w_n(\tau)\}$, $\varepsilon$ playing the role of the excitation characteristics. That is, the greater the time $\tau$, the smaller the difference of distributions! In other words, for some $\tau \in [0, 1]$ one finds

$$\|\rho_{\mu,\varepsilon}(\cdot \mid \tau) - \rho_{\mu,\varepsilon'}(\cdot \mid \tau)\| < \|\rho_{\mu,\varepsilon}(\cdot \mid 0) - \rho_{\mu,\varepsilon'}(\cdot \mid 0)\| \tag{4.12}$$

Now, as $w_n(\tau/\Delta t) = u_n(\tau)$, their distributions are the same: $\rho_{\mu,\varepsilon}(w \mid t/\Delta t) = p_\varepsilon(w \mid t)$; thus, from (4.12) it follows that

$$\|p_\varepsilon(\cdot \mid t) - p_{\varepsilon'}(\cdot \mid t')\| < \|p_\varepsilon(\cdot \mid 0) - p_{\varepsilon'}(\cdot \mid 0)\| \tag{4.13}$$

for some $t$, $t'$. But in the original model the distribution of $\{u_n(t)\}$ is independent of $t$; hence $p_\varepsilon(\cdot \mid t) = p_\varepsilon(u)$ and (4.13) means that $\|p_\varepsilon - p_{\varepsilon'}\| \leqslant \|p_\varepsilon - p_{\varepsilon'}\|$, which is possible only if $p_\varepsilon = p_{\varepsilon'}$!

This marvelous complete independence of $\varepsilon$ is confirmed by numerical experiments: the distributions $p_\varepsilon(u(t))$ plotted in Fig. 4 are independent of $\varepsilon$ within at least 0.1% accuracy for a wide range of $\varepsilon$. The same independence of $\varepsilon$ seems to hold for the two-point distributions $p_\varepsilon(u(t), u(t+\tau))$ (if $\tau/\varepsilon = \text{const}$), and so on.

Unfortunately, the consideration is not rigorous, as (4.9) is not proved to hold for any type of excitation, while the excitation in (4.10) is not of the type $z_{n+1} = \Phi(z_n, \mathbf{w}_n) + \xi_n$ used in the evaluation of (4.9). Moreover, the very proof of (4.9) is based on the "basic" estimate (2.9) with $\kappa(\varepsilon)$ being independent of $\varepsilon$, which is not yet proved.

## 5. DISCUSSIONS

1. For the continuous-time process $u(t)$ the estimate (3.3) (and others of this sort) takes the form

$$\|p_\varepsilon - p_*\| = O(\kappa^{1/\varepsilon}) \tag{5.1}$$

where $p_\varepsilon$ is the distribution of $\{u(t), u(t+\vartheta_1\varepsilon), \ldots, u(t+\vartheta_k\varepsilon)\}$, $\vartheta_i = \text{const}$ ($\Delta t_i/\varepsilon = \text{const}$), and $p_*$ is the "universal" (limit) distribution. The same holds for any statistical characteristics derived from the invariant distribution, say, for the autocorrelation function $C_\varepsilon(\tau)$:

$$|C_\varepsilon(\tau) - C_*(\tau/\varepsilon)| = O(\kappa^{1/\varepsilon}) \qquad \text{if} \quad \tau < 1 \tag{5.2}$$

the universal function $C_*(\eta)$ being almost zero for $\eta \geqslant 10$.

2. The behavior of the Lyapunov exponents (LEs) is somewhat different. Though they do converge to the limit value $\lambda_* \simeq 0.3$ (the same for all of them) as $\varepsilon \to 0$, $\lim_{\varepsilon \to 0} \lambda_{\varepsilon,k} = \lambda_*$, this occurs much more slowly, as $|\lambda_{\varepsilon,k} - \lambda_*| = O(k\varepsilon)$; see Fig. 9. Comparing this with the estimate (5.1) for probability distributions, one gets

$$|\lambda_{\varepsilon,k} - \lambda_*| = O(1/\log \|p_\varepsilon - p_*\|)$$

A similar behavior was found in two weakly coupled identical 1D maps[20]:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = F_d \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} rx_n(1-x_n) - d(x_n - y_n) \\ ry_n(1-y_n) + d(x_n - y_n) \end{pmatrix}$$

where LEs appeared to be very sensitive to the coupling strength $d$: $\lambda_{d,m} = \lambda_{0,m} + O(1/\log d)$ (even for $d \sim 10^{-10}$, $\lambda_{d,m} - \lambda_{0,m} \sim 0.1$), while the distribution was not: $\|p_d - p_0\| = O(d)$; thus we also have

$$|\lambda_{d,k} - \lambda_{0,k}| = O(1/\log \|p_d - p_0\|)$$

The similarity seems to result from the fact that for $\varepsilon \ll 1$ the model (1.1) looks as if composed of $O(1/\varepsilon)$ weakly coupled identical "subsystems" located on associated domains (see Section 3).

3. It should be noted that the particular form of Eq. (1.1) is not too important, as only the "local spatial dependence" (LSD) was used. The term LSD means that (sufficiently far from the boundary) $u_{n+1}(t)$ depends on $u_n(t')$ for $t - 5\varepsilon \leqslant t' \leqslant t$ only (recall that $t$ in this model is rather "a spatial" coordinate; Section 1).

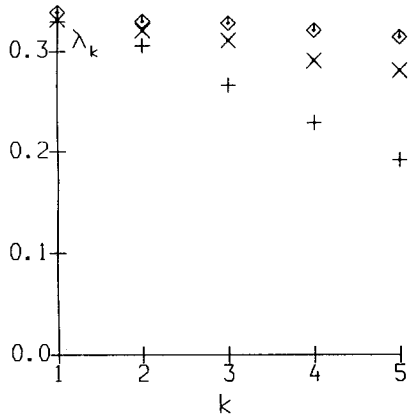



Fig. 9. Initial part of the Lyapynov spectrum $\{\lambda_k\}$. First five LEs are plotted vs. $k$ for $\varepsilon = 0.05$ (+), 0.02 (×), and 0.01 (◇).

But the LSD property is rather common: for example, the reaction-diffusion system

$$\mathbf{u}_t = D\mathbf{u}_{xx} + \mathbf{f}(\mathbf{u}), \qquad \mathbf{u} = \mathbf{u}(x, t)$$

determines a mapping in the functional space: $\mathbf{u}_{n+1} = \hat{T}[\mathbf{u}_n]$ [where $\mathbf{u}_n(x) \equiv \mathbf{u}(t_n, x) = \mathbf{u}(n\tau, x)$] with $u_{n+1}(x)$ effectively depending on $u_n(x')$ for $x - O(\sqrt{D}) \leqslant x' \leqslant x + O(\sqrt{D})$ only. So this model's statistical characteristics are expected to be asymptotically independent of the diffusion coefficient $D$ for $D \to 0$.

However, accurate numerical integration for this model is rather difficult, and it is natural to put forward a more convenient system with similar properties. As the reaction-diffusion system arises from, to put it loosely, adding a diffusion term to ODEs, its straightforward simplification is to replace ODEs by maps:

$$u_{n+1}(x) = \mu^2(\partial^2 u_{n+1}/\partial x^2)(x) + f(u_n(x)), \qquad -L \leqslant x \leqslant L \tag{5.3}$$

where $x$ is a spatial coordinate and $n$ is a discrete time. The equation allows the following interpretation: there is a family of (identical) maps $u_{n+1} = f(u_n)$ located at each point $x$, which are then diffusively coupled. Indeed, Eq. (5.3) can be transformed into (this representation was used in ref. 17):

$$u_{n+1}(x) = \int_{-L}^{L} G(x-y)\, f(u_n(y))\, dy \tag{5.3'}$$

where $G$ is the Green's function: $G(x) = \mu^2 G''(x) + \delta(x)$.

Due to the similarity of (5.3') and (2.1), this model enables almost straightforward application of the theory developed for the delay-differential equation. Experiments have confirmed that for $\mu \ll 1$ one- and two-point distributions $p_\mu(u(x))$ and $p_\mu(u(x), u(x+\Delta x))$ as well as the scaled correlation function $C_\mu(x/\mu)$ are almost independent of $\mu$; see Figs. 10–12. Moreover, an estimate similar to (4.9)

$$\| p_\mu(\cdot \mid x) - p_* \| = O(\exp(-C\,|x - x_b|/\mu) \tag{5.4}$$

indicates that the influence of the boundary conditions and the position $x$ exponentially damps as the distance $|x - x_b|$ grows between $x$ and the nearest boundary. Hence the local (=in the vicinity of a given point $x$) statistical characteristics [e.g., $p_\mu(u(x) \mid x)$] at internal points are almost independent of $\mu$, $x$, and the boundary conditions. Such a property seems to be somewhat similar to the forgetting *initial* conditions in ODEs.

Further simplification of the model is achieved by replacing the

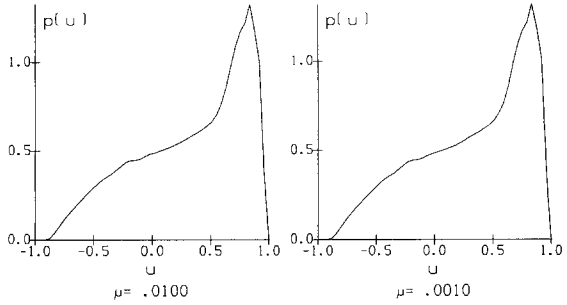



Fig. 10. One-point distribution $p_\mu(u(x))$ for the "diffusively coupled maps" model (5.3) for $\mu = 0.01$ (left) and $\mu = 0.001$ (right). The area length $L = 0.5$, the b.c. being periodic.

continuous spatial coordinate $x$ by a discrete one $i$, i.e., by replacing the integral in (5.3′) by a sum, e.g.,

$$u_{n+1}(i) = \tfrac{1}{3}\{f(u_n(i-1)) + f(u_n(i)) + f(u_n(i+1))\}, \qquad -N \leqslant i \leqslant N \tag{5.5}$$

Such models, called CMLs (coupled map lattices), are very popular in studying spatiotemporal chaos.[16] The map (2.6) is of this type, so the theory developed for it is obviously applicable to CMLs as confirmed experimentally.

4. The asymptotics obtained appeared very useful in investigating the "nonchaotic (transient) turbulence" discovered by Crutchfield and Kaneko[21] in CMLs with a *nonchaotic*, i.e., nonexpanding local map $f$ of the form $f(x) = \omega x + s \pmod 1$, $\omega \cong 0.9$, $s \cong 0.1$. For this map $|f'| < 1$ this implies that a "real" chaos in the model with such an $f$ is impossible; and the attractor(s) should be "*laminar*." Nevertheless, during an enormous time a "turbulent" (i.e., irregular) behavior was recorded and only later does the trajectory reach the "actual" laminar attractor. The duration of the transients was found to grow as $T_N \sim \exp(cN^v)$, $2 \leqslant v \leqslant 3$.[21]

Our investigations[22] have revealed that the phenomenon is related to

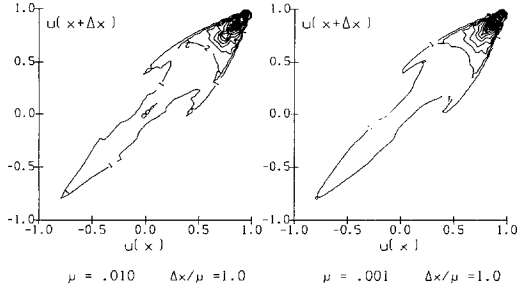



Fig. 11. Same as Fig. 10, but for two-point distribution $p_\mu(u(x), u(x+\mu))$.

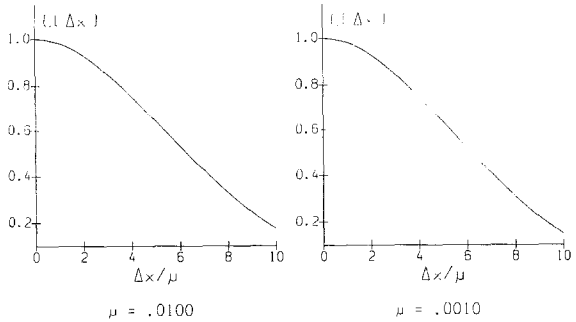



Fig. 12. Same as Fig. 10, but for correlation function $C_\mu(x)$.

the attractor crisis: with slightly varied $f$ one can transform the transient turbulence into an "eternal" chaos. In the original model the chaotic attractor disappears: a tiny "hole" arises in it through which all trajectories sooner or later (rather later, as it were) escape the "turbulent area" and reach the laminar attractor. Until then, i.e., during the "turbulent" period, spatial correlations decay; thus, the "hole" volume scales as $V_N \sim e^{-\gamma N}$,[22] and the duration of the transients can be estimated in the usual way in dealing with this crisis[23]: $T_N \sim 1/P_N$, where $P_N = V_N \rho_N(\mathbf{u}_h)$ is the probability of "falling" into the hole located near the point $\mathbf{u} = \mathbf{u}_h$, $\rho_N$ is the density of the quasistationary distribution. It appeared that the quasistationary distribution enables the same asymptotic estimates as the invariant one, that is, $\delta_N \equiv |\rho_N(\mathbf{u}_h) - \rho_*(\mathbf{u}_h)| \leqslant O(e^{-O(N)})$, from which it follows that

$$T_N \sim e^{O(N)}/[\rho_*(\mathbf{u}_h) + \delta_N] \tag{5.6}$$

Generally, $\rho_*(\mathbf{u}_h) \neq 0$, and thus $T_N \sim e^{cN}$,[22] which is the case for another critical CML[24]; but in the Crutchfield–Kaneko model it is zero. Hence the "corrections" $\delta_N$ prevail and $T_N \geqslant \exp(cN^2)$. As the convergence rate $|\rho_N(\mathbf{u}_h) - \rho_*(\mathbf{u}_h)| = O(e^{-O(N)})$ is only the *slowest* possible, $T_N$ may grow even faster.

Therefore it is the exponential convergence of the statistical characteristics that may cause a hyperexponential growth of the duration of transients in distributed systems. And there are reasons to believe[21] that such phenomena may occur in hydrodynamics.

## ACKNOWLEDGMENTS

I am grateful to A. Potapov, Ya. Sinai, and G. Malinetsky for fruitful discussions, and to my mother, V. Toroptseva, for reading the first version of the manuscript and improving its style.

## REFERENCES


1. J. D. Farmer, Chaotic attractors of infinite-dimensional systems, *Physica* **4D**:366–392 (1982).
2. K. Ikeda and K. Matsumoto, High-dimensional chaotic behavior in systems with delayed feedback, *Physica* **29D**:223–236 (1987).
3. K. Ikeda, K. Kondo, and O. Akimoto, Successive higher-harmonic bifurcations in systems with delayed feedback, *Phys. Rev. Lett.* **49**:1467–1470 (1982).
4. K. Ikeda and K. Matsumoto, Study of a high dimensional chaotic attractor, *J. Stat. Phys.* **44**:955–982 (1986).
5. F. A. Hopf, D. L. Kaplan, *et al.*, Bifurcations to chaos in optical bistability, *Phys. Rev.* **25a**:2172–2183 (1982).
6. B. Dorrizi, B. Grammaticos, M. Berre, Y. Pomeau, E. Ressayre, and A. Tallet, Statistics and dimension of chaos in differential delay systems, *Phys. Rev.* **35a**:328–339 (1987).
7. A. N. Sharkovsky, Yu. L. Maistrenko, and E. Yu. Romanenko, *Difference Equations and Their Applications* (1986) [in Russian].
8. J. Mallet-Parret and R. Nussbaum, A bifurcation gap for a singularly perturbed delay equation, in *Chaotic Dynamics and Fractals* (1986).
9. J.-P. Eckmann and D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.* **57**:617–656 (1985).
10. R. Shaw, Strange attractors, chaotic behavior and information flow, *Z. Naturforsch.* **36a**:80–112 (1981).
11. H.-G. Schuster, *Deterministic Chaos* (Physik Verlag, Weiheim, 1984).
12. S. V. Ershov and G. G. Malinetsky, On the solution of the direct and inverse problem for the Frobenius–Perron equation, IAM preprint #166 (1988) [in Russian].
13. H. Haken and G. Mayer-Kress, Influence of noise on the logistic model, *J. Stat. Phys.* **26**:149–171 (1981).
14. J. L. Doob, *Stochastic Processes* (1953).
15. J. I. Kifer, On small random perturbations of some smooth dynamical systems, *Math. USSR Izv.* **8**:1083 (1974).
16. K. Kaneko, Pattern dynamics in spatio-temporal chaos, *Physica* **34D**:1–41 (1989).
17. A. S. Pikovsky, Spatial development of chaos in nonlinear media, *Phys. Lett. A* **137**:121–127 (1989).
18. R. J. Deissler, Noise-sustained structure, intermittency and the Ginzburg–Landau equation, *J. Stat. Phys.* **40**:371–395 (1988).
19. R. J. Deissler, External noise and the origin and dynamics of structure in convectively unstable systems, *J. Stat. Phys.* **54**:1458–1488 (1989).
20. H. Daido, Coupling sensitivity of chaos: Theory and further numerical evidence, *Phys. Lett. A* **121A**:60–66 (1988).
21. J. P. Crutchfield and K. Kaneko, Are attractors relevant to turbulence? *Phys. Rev. Lett.* **60**:2715–2718 (1988).
22. S. V. Ershov and A. B. Potapov, On the nature of nonchaotic turbulence, IAM preprint #7 (1991) [in Russian]; *Phys. Lett. A*, submitted.
23. C. Grebogi, E. Ott, and J. A. Yorke, Crises, sudden changes in chaotic attractors and transient chaos, *Physica* **7D**:181–200 (1983).
24. T. Schreiber, Spatio-temporal structure in coupled map lattices: Two-point correlations versus mutual information, *J. Phys. A* **23A**:L393–398 (1990).
25. F. Hofbauer and G. Keller, Ergodic properties of invariant measures of piecewise monotonic transformations, *Math. Z.* **180**:119–140 (1980).
26. V. L. Volevich, Kinetics of coupled map lattices, *Nonlinearity* **4**:37–48 (1991).