Journal of Statistical Physics, Vol. 75, Nos. 1/2, 1994

Probabilistic Models of Multidimensional Piecewise Expanding Mappings

Jörn Loviscach[1]

Received August 10, 1993

Strongly chaotic systems (e.g., piecewise expanding mappings) exhibit diffusion-like behavior in the sense of central limit theorems. To find more precise statements about the similarity to probabilistic diffusion, we study how the evolution of probability densities under d-dimensional piecewise expanding mappings can be modeled by Markov processes with smooth transition probabilities (such as diffusion processes). Our results can be viewed as a special type of local limit theorem.

KEY WORDS: Multidimensional piecewise expanding mappings; deterministic chaos; diffusion; central limit theorem; local limit theorem; Markov processes.

1. INTRODUCTION

Diffusion generated by deterministic chaos has been studied both numerically and with the help of rigorous mathematics (cf. refs. 10 and 17 and the references cited therein). We investigate a class of discrete-time dynamical systems which are periodic extensions of piecewise expanding maps. These systems can be written

$$\left.\begin{aligned} X_{n+1} &= X_n + f(X_n, y_n) \\ y_{n+1} &= g(X_n, y_n) \end{aligned}\right\} \tag{1}$$

where $X_n \in \mathbb{R}^d$ and where y_n takes values from the e-dimensional torus $\mathscr{T}^e$ (which, as a set, will be identified with the cube $[0, 1)^e$ in the usual way). We assume the functions f and g to be integer- (i.e., $\mathbb{Z}^d$-) periodic with respect to X_n. (Further assumptions on f and g will be introduced in the following.)

---

[1] BiBoS/Fakultät für Physik, Universität Bielefeld, D-33615 Bielefeld, Germany.

0022-4715/94/0400-0189$07.00/0

X_n is regarded as an "observable"; its behavior is to be modeled by some diffusing *probabilistic* system. The value of y_n, the other variable, is not registered, however. One can think of a microscopic, internal, or "hidden" quantity, like the angle variables in ref. 1. The case $e = 0$, i.e., that none of the variables is hidden, is included.

We look for simple probabilistic systems which approximate the behavior of X_n as $n \to \infty$ in the following way: Let ρ_0 be a probability density[15] for X_0, and fix a probability distribution for y_0. Then, without further knowledge of the system, X_n becomes a random variable. Assume it possesses a probability density ρ_n^d. (We employ the superscript d as an abbreviation for "deterministic.") Now look at some stochastic process in $\mathbb{R}^d$. Starting from the same initial density ρ_0, it evolves in time n to a density ρ_n (assuming existence). Our aim is to find a stationary Markov process with the following property: Regardless of ρ_0 (up to smoothness assumptions) the natural "distance"

$$\|\rho_n^d - \rho_n\|_{L^1(\mathbb{R}^d)} := \int_{\mathbb{R}^d} d^dX\, |\rho_n^d(X) - \rho_n(X)|$$

of the densities decays like $n^{-\gamma}$ as $n \to \infty$, with γ as large as possible. This criterion means that the sought-for stochastic process matches the long-time behavior of X_n on both microscopic (~ 1) and macroscopic ($\sim \sqrt{n}$) scales. We will find that in the general case the decay of $\|\rho_n^d - \rho_n\|_{L^1(\mathbb{R}^d)}$ is faster than $n^{-\gamma}$ for all $\gamma < 1/2$ (for suitable random models). A decay faster than $n^{-1/2}$ is not possible in general (for Markov models with sufficiently smooth transition functions).

This probabilistic approximation is of global *and* local type. Most known results, however, are exclusively concerned with the behavior in the large. Especially the two well-known diffusionlike properties the central limit theorem and the functional central limit theorem, are statements about the scaled-down variable $X_{nt}/\sqrt{n}$, $n \to \infty$. Only refs. 11 and 18 also discuss local properties in this context, namely the local limit theorem and the renewal theorem.

To cast the system (1) into a more convenient form, split X_n into a macroscopic and a microscopic part: $X_n = \lfloor X_n \rfloor + \langle X_n \rangle$, where each entry of $\lfloor X_n \rfloor$ is the largest integer smaller than or equal to the corresponding entry of X_n. The second member, $\langle X_n \rangle$, is the fractional part in $[0, 1)^d \equiv \mathscr{T}^d$ of X_n. Let us denote $\langle X_n \rangle$ by x_n. There is a unique mapping T from (x_n, y_n) to (x_{n+1}, y_{n+1}). Therefore the integer vector $\lfloor X_{n+1} \rfloor - \lfloor X_n \rfloor$ can be written as a function of x_n and y_n only. Let us call this function $k(x_n, y_n)$, so that

$$X_N = x_N + \lfloor X_0 \rfloor + \sum_{n=0}^{N-1} k(x_n, y_n) \tag{2}$$

Thus Eq. (1) yields an "observable" X_n which is driven by the (x_n, y_n) motion. We can expect X_n to be diffusive if the system given by the mapping T on the phase space $\mathcal{T}^d \times \mathcal{T}^e$ is chaotic, because $k(x_n, y_n)$ can then be regarded as a "random" increment.

In this work we consider the case that T is a $(d+e)$-dimensional piecewise expanding map. (The case of T being an Anosov diffeomorphism will be treated in a later paper.)

As will be shown, on rather general conditions there exists a unique measure $d\mu$ which is invariant under T and absolutely continuous with respect to Lebesgue measure $d\lambda = d^d x\, d^e y$. If we want the system to exhibit diffusive motion, we have to ensure that the variable X_n typically leaves every bounded set. To this end, assume that for all $N = (N_1, \ldots, N_d) \in \mathbb{N}^d$ the restricted system

$$\left.\begin{aligned} X_{n+1} &= X_n + f(X_n, y_n) \quad (\mathrm{mod}\ N) \\ y_{n+1} &= g(X_n, y_n) \end{aligned}\right\} \tag{3}$$

is weakly mixing with respect to the invariant measure which can be constructed by periodic extension of the measure $d\mu$. This assumption will allow a reasoning as in refs. 7 and 14 to show that the effective diffusion tensor is nondegenerate.

As a toy example consider the system given by the mapping

$$X_{n+1} = X_n + \langle X_n \rangle \tag{4}$$

of $\mathbb{R}$ into itself. This can be treated within the above framework ($d = 1$, $e = 0$) by rewriting it as follows:

$$\left.\begin{aligned} x_{n+1} &= \langle 2x_n \rangle \\ X_N &= x_N + \lfloor X_0 \rfloor + \sum_{n=0}^{N-1} k(x_n) \end{aligned}\right\}$$

where $k(x) := \lfloor 2x \rfloor$. The natural choice for $d\mu$ is Lebesgue measure $d\lambda$ itself. For all $N \in \mathbb{N}$

$$X_{n+1} = X_n + \langle X_n \rangle \quad (\mathrm{mod}\ N)$$

is weakly mixing with respect to Lebesgue measure.[6] (Note that this reduced system, too, is a one-dimensional piecewise expanding map.)

Plan of This Paper. First, we give a definition of "functions of bounded variation" and recall the theorem of Ionescu-Tulcea and Marinescu. With the help of these tools we analyze multidimensional piecewize expanding mappings; as an extension to the work of M. L. Blank, we find both exponential decay of correlations and the central limit theorem for the large class of functions of bounded variation. The asymptotic behavior of probability densities under piecewise expanding mappings is estimated. A concluding analysis of the behavior of probability densities under Markov processes leads to the final result.

2. TOOLS

2.1. Generalized Variation

The function k which appears in the fundamental equation (2) is, by construction, $\mathbb{Z}^d$-valued and hence either trivial or discontinuous. Clearly, it must not be too singular. In the *one*-dimensional case where $d=1$ and $e=0$, it is possible[18] to assume that k is a function of bounded variation (in the well-known usual sense).

For treating the *multi*dimensional case the usual, one-dimensional notion of bounded variation has to be generalized; this can be done in a number of ways. The following definition is closely related to the ones given by Blank[3] and Keller[13]: Denote by d the distance on $\mathcal{T}^d \times \mathcal{T}^e$. For α, $\beta \in (0, 1)$ consider the set $BV_{\alpha,\beta}$ consisting of the complex-valued $L^1(d\lambda)$-functions[2] f with finite variation

$$\operatorname{var}_{\alpha,\beta}(f) := \inf_{\tilde{f}} \sup_{0<t<\beta} t^{-\alpha} \int d\lambda(z) \sup_{z_{1/2}:\, d(z,\, z_{1/2})<t} |\tilde{f}(z_1) - \tilde{f}(z_2)|$$

where the infimum runs over all functions $\tilde{f}$ which equal f almost everywhere with respect to $d\lambda$. (In the following this will be abbreviated "$d\lambda$-a.e.") Equip $BV_{\alpha,\beta}$ with the norm $\|f\|_{\alpha,\beta} := \|f\|_1 + \operatorname{var}_{\alpha,\beta}(f)$.

The following theorem gives a useful criterion for a function to be an element of $BV_{\alpha,\beta}$; besides, it shows that these functions can be quite irregular and that α plays the role of a Hölder exponent.

Theorem 1. For some bounded function f assume that there is a number $C<\infty$ and a subset D which cuts $\mathcal{T}^d \times \mathcal{T}^e$ into a countable union $\bigcup_i A_i = \mathcal{T}^d \times \mathcal{T}^e - D$ of disjoint open sets A_i such that the restriction of f to each of these sets A_i fulfills a Hölder condition

$$|f(z_1) - f(z_2)| \leqslant C d(z_1, z_2)^\alpha \qquad \text{for all} \quad z_1, z_2 \in A_i$$

[2] Function classes modulo equality $d\lambda$ almost everywhere, of course. In the following, the Lebesgue spaces L^p and their norms $\|\cdot\|_p$ refer to the measure $d\lambda$ if not indicated otherwise.

Assume furthermore that the upper capacity[3] $\bar{C}$ of D is smaller than $d+e-\alpha$. (This is the case, e.g., if D is the union of a finite number of smooth hypersurfaces.) Then $f \in BV_{\alpha,\beta}$ for β small enough.

We list several important properties of $BV_{\alpha,\beta}$:

1. For all $f \in BV_{\alpha,\beta}$ and all Lipshitz continuous functions ϕ on $\mathbb{C}$ also $\phi \circ f$ is an element of $BV_{\alpha,\beta}$ with $\mathrm{var}_{\alpha,\beta}(\phi \circ f) \leqslant \mathrm{Lip}(\phi)\, \mathrm{var}_{\alpha,\beta}(f)$.
2. There exists $C < \infty$ such that $\|f\|_\infty \leqslant C\, \|f\|_{\alpha,\beta}$ for all $f \in BV_{\alpha,\beta}$.
3. $\mathrm{var}_{\alpha,\beta}(fg) \leqslant \mathrm{var}_{\alpha,\beta}(f)\, \|g\|_\infty + \mathrm{var}_{\alpha,\beta}(g)\, \|f\|_\infty$ for all $f, g \in BV_{\alpha,\beta}$.

Of technical use is that $BV_{\alpha,\beta}$ with its above-defined norm is a Banach space. Its closed unit ball is compact as a subset of L^1.

2.2. The Theorem of Ionescu-Tulcea and Marinescu

Keystone to the study of chaotic systems by operator techniques is the following theorem,[16] which we present in a form specialized for our application:

Theorem 2 (Ionescu-Tulcea and Marinescu, special case). Fix α, $\beta \in (0, 1)$. Let P be a bounded operator in $BV_{\alpha,\beta}$ which can be extended to a bounded operator in L^1. Suppose that:

1. P is contracting with respect to L^1.
2. There exist $r_1 \in (0, 1)$, $r_2 \in \mathbb{R}_0^+$, and $m \in \mathbb{N}$ such that

$$\mathrm{var}_{\alpha,\beta}(P^m) \leqslant r_1\, \mathrm{var}_{\alpha,\beta}(f) + r_2\, \|f\|_1 \qquad \text{for all} \quad f \in \mathscr{L} \tag{5}$$

Then, P^n can for all $n \in \mathbb{N}$ be decomposed as $P^n = \sum_\gamma \gamma^n \Pi_\gamma + R^n$, where the sum runs over all eigenvalues γ of P which are of modulus 1 and belong to eigenvectors in $BV_{\alpha,\beta}$. There are only finitely many such eigenvalues, so that the sum is well defined. The operators Π_γ are $BV_{\alpha,\beta}$-projectors onto the corresponding eigenspaces. R maps $BV_{\alpha,\beta}$ into itself, and its $BV_{\alpha,\beta}$-spectral radius is strictly smaller than 1. Furthermore, $\Pi_\gamma \Pi_\delta = 0$ and $\Pi_\gamma R = 0 = R\Pi_\gamma$ for all $\gamma \neq \delta$ which occur as eigenvalues of modulus 1.

We need also a form of this theorem with weakened assumptions:

[3] Recall that the upper capacity $\bar{C}$ of a set D is given by $\limsup_{t \downarrow 0} (\log 1/t)^{-1} \log N(t)$, where $N(t)$ is the number of t-balls needed to cover D.

Theorem 3. In the situation of Theorem 2 suppose instead:

1. $\sup_{n \in \mathbb{N}_0} \|P^n\|_\infty < \infty$.
2. There exist $r_1 \in (0, 1)$ and $r_2 \in \mathbb{R}_0^+$ such that

$$\mathrm{var}_{\alpha,\beta}(Pf) \leqslant r_1 \, \mathrm{var}_{\alpha,\beta}(f) + r_2 \|f\|_\infty \qquad \text{for all} \quad f \in \mathscr{L} \tag{6}$$

Then the $BV_{\alpha,\beta}$-spectral radius of P is equal to or smaller than 1. It equals 1 if and only if there exists an eigenvector in $BV_{\alpha,\beta}$ with eigenvalue of modulus 1.

When applying these theorems, the following lemma is useful:

Lemma 4. Let $K \subset \mathbb{R}^d$ be a compact set. Assume $A: p \mapsto A_p$ maps K continuously (with respect to operator norm) to the bounded operators in some Banach space with norm $\|\cdot\|$. If the spectral radius of all A_p, $p \in K$, is strictly smaller than 1, then there exist $C < \infty$ and $\kappa \in (0, 1)$ such that $\|A_p^n\| \leqslant C\kappa^n$ for all $p \in K$ and all $n \in \mathbb{N}$.

3. PIECEWISE EXPANDING MAPPINGS

3.1. Basic Properties

For the notion "T is piecewise expanding" we employ a quite broad definition: There is a finite partition of $\mathscr{T}^{d+e}$ into Borel sets A_i such that:

1. On each A_i, the mapping T is the restriction $T_i|_{A_i}$ of some C^1-diffeomorphism T_i defined in an open neighborhood B_i of A_i; the Jacobi determinant of T_i is bounded from above by some $C < \infty$, and the inverse of the Jacobi determinant is Hölder continuous with growth constant c and exponent $\alpha \in (0, 1)$.
2. There exists a common $\gamma > 1$ such that every $T_i: B_i \mapsto T_i(B_i)$ expands distances by at least a factor γ.

The Perron–Frobenius operator P describes the evolution of probability densities under the action of T. It is defined by

$$\int d\lambda \, gP(f) \stackrel{!}{=} \int d\lambda (g \circ T) f \qquad \text{for all} \quad f \in L^1 \quad \text{and} \quad g \in L^\infty \tag{7}$$

It preserves the integral $\int d\lambda$ and is L^1-contracting. Note that

$$P(f)(z) = \sum_i 1_{TA_i}(z) \, J_i(z) \, f \circ T_i^{-1}(z) \tag{8}$$

where 1_{TA_i} is the indicator function of the set TA_i and $J_i(z) := |\det(DT_i^{-1})(z)|$.

In order to apply the theorem of Ionescu-Tulcea and Marinescu to P we use the following theorem. It is inspired by the work of Blank,[3] but covers a more general class of piecewise expanding maps, e.g., the boundaries of the sets A_i may be very singular in comparison.

Theorem 5. If β is small enough and if the upper capacities of the sets $\bigcup_i \partial A_i$ and $\bigcup_i \partial T A_i$ are strictly smaller than $d+e-\alpha$, then for all $f \in BV_{\alpha,\beta}$ we have

$$\operatorname{var}_{\alpha,\beta}(Pf) \leqslant \gamma^{-\alpha} A \operatorname{var}_{\alpha,\beta}(f) + B \|f\|_1 \tag{9}$$

where A, $B < \infty$ only depend on α, β, c, C, and the shape of the A_i.

The proof of this theorem is given in the appendix.

If the Jacobi determinants of each T_i do not oscillate too much, then c is small; if in addition the boundaries of the A_i and TA_i are piecewise smooth, then the expressions on the r.h.s. of Eqs. (26) and (27) (see appendix) can be made arbitrarily small by choosing β small. So it is not difficult to see that for a large class of piecewise expanding maps T one has indeed $\gamma^{-\alpha} A < 1$ in Eq. (9) for a certain choice of α and β. We are going to assume this property of T in the following.[4] Then Eq. (9) immediately yields condition 2 for Theorem 2.

Therefore, for all $n \in \mathbb{N}$ the nth power of the Perron–Frobenius operator in $BV_{\alpha,\beta}$ can be decomposed as

$$P^n := \sum_{\gamma} \gamma^n \Pi_\gamma + R^n \tag{10}$$

where the sum runs over all eigenvalues γ of modulus 1. Π_γ is a projector of $BV_{\alpha,\beta}$ onto the corresponding eigenspace, and the $BV_{\alpha,\beta}$-spectral radius of R is strictly smaller than 1.

From property 2 and Eq. (9) one can deduce by iteration that $\|P^n(1)\|_2$ is bounded by some $C < \infty$ for all $n \in \mathbb{N}_0$. From this and from Eq. (10) follows the existence of at least one $h \in BV_{\alpha,\beta}$ with $h > 0$, $\int d\lambda\, h = 1$, and $Ph = h$.

Now, $d\mu := h\, d\lambda$ defines an invariant measure for the mapping T. We

[4] If the expansion rate γ is too low to fulfill $\gamma^{-\alpha} A < 1$, one can consider some power T^m of the mapping (and hence P^m instead of P), so that γ is replaced by γ^m. However, A may increase, depending on the geometry of the A_i. (Ref. 5 gives some interesting examples of this problem.) In the *one*-dimensional case, however, the recipe of taking an appropriate power of T always leads to success.[18]

have assumed (cf. the introduction) that T is weakly mixing with respect to this measure. Therefore, h is the only eigenvector of P with an eigenvalue of modulus 1 in $BV_{\alpha,\beta}$ and even in the larger space L^∞. So Eq. (10) reduces to

$$P^n := \Pi + R^n \quad \text{for all} \quad n \in \mathbb{N} \tag{11}$$

where Π is the one-dimensional projector which maps every function f to $h \int d\lambda\, f$.

We assume in the following that each coordinate function of k lies in the space $BV_{\alpha,\beta}$ with α and β as above. [To achieve that, $\alpha \in (0, 1)$ and $\beta \in (0, 1)$ may be lowered at the very beginning, as long as $\gamma^{-\alpha} A < 1$.] From here on we will fix these α, β and suppress them mostly.

Exponential Decay of Correlations. Equation (11) immediately leads to an exponential decay of correlations for the large class of observables given by BV-functions: Take functions $f \in L^1$ and $g \in BV$ with $\int d\mu\, f = 0$. Then

$$\int d\mu\, (f \circ T^n)\, g = \int d\lambda\, fh \int d\lambda\, gh + \int d\lambda\, fR^n(gh)$$

The first term on the r.h.s. vanishes, because $\int d\lambda\, fh = \int d\mu\, f = 0$. The second term on the r.h.s. tends to 0 exponentially fast as $n \to \infty$, because $gh \in BV$ and because the BV-spectral radius of R is strictly smaller than 1.

3.2. The Characteristic Function of X_N

We want to examine the behavior of X_N given the initial probability density $\rho_0(X)\, v(\langle X \rangle, y)$ of (X_0, y_0). Here, ρ_0 is an arbitrary probability density of class C_0^1. The function v is held fixed; it is assumed to be of type C^1 and to be a probability density with respect to y for fixed x, i.e., $\int d^e y\, v(x, y) = 1$ for all $x \in \mathscr{T}^d$.

All information we need is encoded in the characteristic function of X_N for all $p \in \mathbb{R}^d$. It can be calculated with the help of Eq. (2):

$$\begin{aligned}
&E_{\rho_0}[\exp(ip \cdot X_N)] \\
&\quad = \sum_{l \in \mathbb{Z}^d} (\exp ip \cdot l) \int d\lambda(x, y) \\
&\qquad \times \exp\left[ip \cdot x(T^N(x, y)) + ip \cdot \sum_{n=0}^{N-1} k(T^n(x, y)) \right] \rho_0(x + l)\, v(x, y)
\end{aligned}$$

where, by abuse of notation, $x(\cdot)$ denotes the x coordinate $\in [0, 1)^d$ of a point in $\mathscr{T}^d \times \mathscr{T}^e$. Note that the sum over l is actually only finite. With the help of the Perron–Frobenius operator P define $P_p(f) := P(e^{ip \cdot k} f)$ for all $f \in L^1$. Because k is a $\mathbb{Z}^d$-valued function, P_p is $2\pi\mathbb{Z}^d$-periodic with respect to p. By N-fold application of Eq. (7) we find[18]

$$E_{\rho_0}[e^{ip \cdot X_N}] = \sum_{l \in \mathbb{Z}^d} e^{ip \cdot l} \int d\lambda\, e^{ip \cdot x} P_p^N(\rho_{0,l} v) \tag{12}$$

where $\rho_{0,l}(x) := \rho_0(x+l)$. To estimate the behavior of this expression as $N \to \infty$, we derive a decomposition of P_p similar to the decomposition (11) of the Perron–Frobenius operator. (The same idea is exploited in refs. 11 and 18.)

First, note that for $k \in BV$ the operator P_p in BV is bounded (properties 1 and 3 of BV). Even more, the spectral structure of P_p in the space BV is very similar to that of P itself: The spectral radius is equal to or less than 1; if it equals 1, then there exists an eigenfunction in BV with an eigenvalue of modulus 1. This is a consequence of Theorem 3 applied to P_p. [Note that Eq. (6) is a consequence of property 3 applied to the product $e^{ip \cdot k} f$.]

Let us examine for which p there exists a BV-eigenfunction with eigenvalue of modulus 1 (which by the preceding argument is equivalent to the statement that P_p has a BV-spectral radius of 1): Let P_p possess an eigenfunction f in L^∞ with eigenvalue of modulus 1, so that $P(e^{ip \cdot k} f) = e^{i\theta} f$ with some $\theta \in \mathbb{R}$. The Perron–Frobenius operator P has the property that $d\lambda$-a.e. $P(|f|) \geqslant |P(f)|$. Therefore $d\lambda$-a.e. $P(|f|) \geqslant |f|$ for the eigenfunction f. On the other hand, P is an L^1-contractive operator. So $d\lambda$-a.e. $P(|f|) = |f|$. But the 1-eigenspace of P in L^∞ is one-dimensional; therefore (perhaps after normalization) $d\lambda$-a.e. $f = e^{i\phi} h$ for some real-valued function ϕ. Combining the eigenvalue equation with the definition (7) of the Perron–Frobenius operator P we obtain

$$\int d\lambda\, g \circ T e^{ip \cdot k} e^{i\phi} = \int d\lambda\, g e^{i\theta} e^{i\phi} h \qquad \text{for all} \quad g \in L^\infty$$

We can choose $g := e^{-i\theta} e^{-i\phi}$, so that $\int d\mu \exp i(p \cdot k - \theta - \phi \circ T + \phi) = 1$, and hence

$$p \cdot k - \theta - \phi \circ T + \phi \quad d\mu\text{-a.e.} \quad \text{is an element of } 2\pi\mathbb{Z} \tag{13}$$

This result can be turned into an equivalence: If we are given p, θ, and ϕ fulfilling (13), then obviously $e^{i\phi} h$ is a L^∞-eigenfunction with eigenvalue $e^{i\theta}$. Then there also exists an eigenfunction of P_p with eigenvalue of modulus 1 which belongs not only to L^∞ but also to BV. (Use Theorem 3 and the denseness of BV in L^1.)

Call $\mathscr{S}$ the set of all $q \in \mathbb{R}^d$ for which the BV-spectral radius of $P_{2\pi q}$ is equal to 1. Obviously $\mathscr{S} \supset \mathbb{Z}^d$, and $\mathscr{S}$ is invariant under integer translations. Condition (13) shows that with every finite set $\{q_i\} \subset \mathscr{S}$ also $\{m + \sum_i n_i q_i : m, n_i \in \mathbb{Z}\}$ is a subset of $\mathscr{S}$. We proceed in two steps to prove that indeed $\mathscr{S} = \mathbb{Z}^d$:

Lemma 6. If $\mathscr{S}$ contains a vector with some irrational entry, then $\mathscr{S}$ has a limit point in $\mathbb{Q}^d - \mathbb{Z}^d$.

Proof. Let q be an element of $\mathscr{S}$. It is sufficient to show: If some entry of q is irrational, then the subset $\{\langle nq \rangle : n \in \mathbb{Z}\}$ of $\mathscr{T}^d$ has a nonzero rational limit point. This can be proved by considering the well-known dynamical system on the torus $\mathscr{T}^d$ which is defined by iterating the mapping $x \mapsto \langle x + q \rangle$. ■

Lemma 7. The BV-spectral radius of $P_{2\pi q}$ never equals 1 for $q \in \mathbb{Q}^d - \mathbb{Z}^d$.

Proof. Consider $0 \neq q = (m_1/N_1, ..., m_d/N_d)$, where $m_r \in \{0, ..., N_r - 1\}$, $N_r \in \mathbb{N}$, $1 \leqslant r \leqslant d$, and at least one of the m_r does not vanish. It is sufficient to show that for all such q there does not exist an eigenvector of $P_{2\pi q}$ with spectral radius of modulus 1. We have assumed that for all $N = (N_1, ..., N_r) \in \mathbb{N}^d$ the system (3) has the periodic extension $d\mu_N$ of $d\mu$ as a weakly mixing measure. A calculation of the Perron–Frobenius operator P_N of this enlarged system reveals: If $P_{2\pi q}$ *had* an eigenfunction f to an eigenvalue of modulus 1, then $(X, y) \mapsto \exp(-2\pi i q \cdot \lfloor X \rfloor) f(\langle X \rangle, y)$ would be an eigenfunction of P_N with the same eigenvalue. But due to the assumed weak mixing, this would imply

$$\exp\left(-2\pi i \sum_{r=1}^{d} \frac{m_r \lfloor X_r \rfloor}{N_r}\right) f(\langle X \rangle, y)$$
$$= \text{const} \cdot h(\langle X \rangle, y) \qquad \text{for} \quad d\mu_N\text{-almost all } (X, y)$$

This cannot be true, because the exponential factor is not invariant with respect to integer translations in X. ■

If $\mathscr{S}$ *was* different from $\mathbb{Z}^d$, then by Lemma 6 it would possess an element of $(\mathbb{Q}^d - \mathbb{Z}^d)$ either as an element *or* as a limit point. But such a limit point has to be a point of $\mathscr{S}$ itself, because $\mathscr{S}$ is a closed set: The mapping $p \mapsto P_p$ is BV-analytic and therefore the spectral radius of P_p is semicontinuous from above with respect to p (cf. Theorems IV.2.23a and IV.3.1 of ref. 12). So in any case $\mathscr{S}$ would contain an element of $(\mathbb{Q}^d - \mathbb{Z}^d)$, which is forbidden by Lemma 7. In conclusion, $\mathscr{S}$ cannot be different from $2\pi\mathbb{Z}^d$.

For small $|p|$, one can view P_p as an analytic perturbation of P: Along the lines of ref. 18 or, more generally, by ref. 12 one can show that there is an $\varepsilon > 0$ such that for all p, $|p| \leqslant \varepsilon$, we have a decomposition analogous to Eq. (11):

$$P^n_{p \cdot k} = K^n_p \Pi_p + R^n_p \qquad \text{for all} \quad n \in \mathbb{N} \tag{14}$$

where K_p is the eigenvalue with the largest modulus, Π_p a one-dimensional BV-projector, R_p a bounded operator, all three of them analytic with respect to p. The BV-spectral radius of all R_p is strictly smaller than 1, and K_p assumes the value 1 for $p = 0$; for all other p, $|p| \leqslant \varepsilon$, its modulus is strictly less than 1. For $p = 0$ the decomposition according to Eq. (14) corresponds to that of Eq. (11): $\Pi_0 = \Pi$ and $R_0 = R$.

Central Limit Theorem. With $B := \int d\mu\, k$, the random variable $(X_N - BN)/\sqrt{N}$ converges to a nondegenerate Gaussian distribution. This follow if we prove that for every $p \in \mathbb{R}^d$ the expectation $E_{\rho_0}[\exp(ip \cdot (X_N - BN)/\sqrt{N})]$ converges to $\exp(-p \cdot Dp/2)$, where D is a strictly positive real symmetric matrix. To see this, fix an arbitrary $p \in \mathbb{R}^d$ and choose M so large that for all $N \geqslant M$ Eq. (14) can be applied to $P^N_{p/\sqrt{N}}$. Then we have

$$\begin{aligned} &E_{\rho_0}[\exp(ip \cdot (X_N - BN)/\sqrt{N})] \\ &\quad = \int d\lambda \exp[ip \cdot (x/\sqrt{N} - B)\sqrt{N}] \\ &\qquad \times (K^N_{p/\sqrt{N}} \Pi_{p/\sqrt{N}} + R^N_{p/\sqrt{N}}) \left[\sum_{l \in \mathbb{Z}^d} \exp(ip \cdot l/\sqrt{N})\, \rho_{0,l} v \right] \end{aligned}$$

Because we have assumed $\rho \in C^1_0$, the expression $\sum_{l \in \mathbb{Z}^d} \exp(ip \cdot l/\sqrt{N})\rho_{0,l} v$ is an element of BV, whose BV-norm is bounded by some C_1 for all $N \geqslant M$. Now apply lemma 4 to $R_{p/\sqrt{N}}$, $N \geqslant M$. Thus, there exist $C_2 > 0$ and $\kappa \in (0, 1)$ such that $\|R^N_{p/\sqrt{N}}\|_{BV} \leqslant C_2 \kappa^N$. So it remains to show that for $N \to \infty$, $N \geqslant M$, we have

$$\exp(-p \cdot Dp/2) - K^N_{p/\sqrt{N}} \exp(-ip \cdot B\sqrt{N}) \to 0 \tag{15}$$

with B as above and some appropriate strictly positive real symmetric matrix D.

We know $K_0 = 1$, $|K_p| \leqslant 1$, and $\bar{K}_p = K_{-p}$, so that it is a simple matter to deduce $K_p = \exp[ip \cdot B - p \cdot Dp/2 + iE(p) + F(p)]$, where D is a nonnegative real symmetric matrix, and where E, F are real-valued C^{∞}-functions whose values and first two, respectively three derivatives vanish at $p=0$. Thus, D in Eq. (15) is known and the convergence is obvious.

The only point left to verify is that D is *strictly* positive. Assume the contrary, namely that there exists a nonzero vector $v \in \mathbb{R}^d$ such that $v \cdot Dv = 0$. Then the variance of the random variable $v \cdot \sum_{n=0}^{N-1} (k \circ T^n - B)/\sqrt{N}$ with respect to $d\mu$ would tend to 0 as $N \to \infty$. Now define $u := \sum_{n=1}^{\infty} P^n[v \cdot (k-B)]$, which converges exponentially in BV, because $\Pi(v \cdot (k-B)) = 0$. Clearly, $u - Pu = P[v \cdot (k-B)]$. Also the variance of

$$\sum_{n=0}^{N-1} [v \cdot (k-B) + u - u \circ T] \circ T^n / \sqrt{N}$$

will tend to 0 as $N \to \infty$. Since we have proved the (even exponential) decay of correlations, we can apply the Green–Kubo formula for the (vanishing) diffusion constant of this stochastic process:

$$0 = \int d\mu [v \cdot (k-B) + u - u \circ T]^2$$

$$+ 2 \sum_{n=1}^{\infty} \int d\mu [(v \cdot (k-B) + u - u \circ T) \circ T^n][v \cdot (k-B) + u - u \circ T]$$

But every member of the sum on the lower line vanishes, because

$$P^n[v \cdot (k-B) + u - u \circ T] = P^n[v \cdot (k-B)] - P^{n-1}(u - Pu) = 0$$

by the construction of u. Therefore, also the integral on the upper line vanishes, so that $d\mu$-a.e. $v \cdot k - v \cdot B + u - u \circ T = 0$. But then according to the eigenvalue criterion (13), P_v has an eigenvalue of modulus 1. Obviously this property remains true if we replace v by some multiple of itself. So we find that the set $\mathcal{S} := \{q \colon P_{2\pi q}$ has an eigenvalue of modulus $1\}$ contains the entire line $\mathbb{R}v$. This is a contradiction to the known identity $\mathcal{S} = \mathbb{Z}^d$.

3.3. The Probability Density of X_N

After n iterations of Eq. (1), the initial density ρ_0 is transformed into a density ρ_n^d of X_n. We want to model the evolution of this density by some other sequence of densities ρ_n [from $L^2(d^dX) \cap L^1(d^dX)$, to be precise]. Apply Eq. (12) to calculate the characteristic function of X_n; the Fourier–Pancherel theorem gives equality and existence for

$$(2\pi)^d \,\|\rho_n^d - \rho_n\|^2_{L^2(d^dX)}$$
$$:= (2\pi)^d \int_{\mathbb{R}^d} d^dX \,|\rho_n^d(X) - \rho_n(X)|^2$$
$$= \int_{\mathbb{R}^d} d^dp \left| E_{\rho_0}[e^{ip\cdot X_n}] - \int d^dX\, e^{ip\cdot X}\rho_n(X)] \right|^2$$
$$= \int_{[-\pi,\pi]^d} d^dp \int_{\mathscr{T}^d} d^dx \left| \left[\int_{\mathscr{T}^e} d^ey\, P_p^n \left(\sum_{l\in\mathbb{Z}^d} e^{ip\cdot l}\rho_{0,l}v \right) \right] - \sum_{l\in\mathbb{Z}^d} e^{ip\cdot l}\rho_n(x+l) \right|^2$$

Now choose $\varepsilon\in(0,\pi)$ so small that we can apply Eq. (14) for $|p|\leqslant\varepsilon$. One can split the integration domain of p into the two sets $\{p\colon |p|\leqslant\varepsilon\}$ and $\{p\in[-\pi,\pi]^d\colon |p|>\varepsilon\}$. The known statements about the spectral radii of R_p and P_p then allow us to apply Lemma 4 in order to find

$$(2\pi)^d \,\|\rho_n^d - \rho_n\|^2_{L^2(d^dX)}$$
$$\leqslant 2\int_{|p|\leqslant\varepsilon} d^dp \int_{\mathscr{T}^d} d^dx \left| \left[\int_{\mathscr{T}^e} d^ey\, K_p^n \Pi_p \left(\sum_{l\in\mathbb{Z}^d} e^{ip\cdot l}\rho_{0,l}v \right) \right] - \sum_{l\in\mathbb{Z}^d} e^{ip\cdot l}\rho_n(x+l) \right|^2$$
$$+ 2\int_{p\in[-\pi,+\pi]^d, |p|>\varepsilon} d^dp \int_{\mathscr{T}^d} d^dx \left| \sum_{l\in\mathbb{Z}^d} e^{ip\cdot l}\rho_n(x+l) \right|^2 + C\kappa^n \tag{16}$$

for all $n\in\mathbb{N}_0$, where $C<\infty$ and $\kappa\in(0,1)$ are constants which do not depend on n.

Assume we are given a Markov process (whose actual construction will be postponed until Section 4.2) with densities ρ_n such that

$$\sum_{l\in\mathbb{Z}^d} (\exp ip\cdot l)\,\rho_n(x+l)$$
$$= \exp\{n[ip\cdot B - p\cdot Dp/2 + iE'(p) + F'(p)]\}$$
$$\times \Pi'_p\left(\sum_{l\in\mathbb{Z}^d} (\exp ip\cdot l)\,\rho_{0,l} \right) + \Delta(n,p,x)$$

for $|p|\leqslant\varepsilon$, where

$$\sup_{|p|\leqslant\varepsilon} \int_{\mathscr{T}^d} d^dx\, |\Delta(n,p,x)|^2$$

decreases exponentially as $n \to \infty$, and where the operator-valued function Π' has the property that

$$\int_{\mathscr{T}^d} d^d x \left| \Pi'_p(f)(x) - \int d^e y\, \Pi_p(fv)(x, y) \right|^2 \leqslant |p|\, C_1\, \|fv\|_{\alpha,\beta}$$

with some $C_1 < \infty$ for $|p| \leqslant \varepsilon$ and $f(x, y)\, v(x, y) = f(x)\, v(x, y) \in BV$. The functions E', F' are real-valued, of type C^∞, and their value and first two, respectively three derivatives vanish at $p = 0$. Furthermore assume that

$$\sup_{p \in [-\pi, \pi]^d, |p| > \varepsilon} \left\| \sum_{l \in \mathbb{Z}^d} e^{ip \cdot l} \rho_n(\cdot + l) \right\|_{L^2(d^d x)}$$

decays exponentially as $n \to \infty$. Then we find from Eq. (16) that $\|\rho_n^d - \rho_n\|_{L^2(d^d X)}$ decays like $n^{-1/2 - d/4}$ or faster as $n \to \infty$.

This can be shown as follows: The r.h.s. of the preceding estimate (16) for $\|\rho_n^d - \rho_n\|^2_{L^2(d^d X)}$ can be estimated by the sum of an expression which decays exponentially as $n \to \infty$ plus the expression

$$\begin{aligned}
&4n^{-d/2} \int_{|q| \leqslant \varepsilon\sqrt{n}} d^d q \exp[-q \cdot Dq + 2nF'(q/\sqrt{n})] \int_{\mathscr{T}^d} d^d x \\
&\quad \times \left\{ \left| \left[\Pi'_{q/\sqrt{n}} \left(\sum_{l \in \mathbb{Z}^d} \exp(iq \cdot l/\sqrt{n})\, \rho_{0,l} \right) \right.\right.\right. \\
&\quad \left.\left. - \int_{\mathscr{T}^e} d^e y\, \Pi_{q/\sqrt{n}} \left(\sum_{l \in \mathbb{Z}^d} \exp(iq \cdot l/\sqrt{n})\, \rho_{0,l} v \right) \right] \right|^2 \\
&\quad + |1 - \exp\{n[iE(q/\sqrt{n}) + F(q/\sqrt{n}) - iE'(q/\sqrt{n}) - F'(q/\sqrt{n})]\}|^2 \\
&\quad \left. \times \left| \int_{\mathscr{T}^e} d^e y\, \Pi_{q/\sqrt{n}} \left(\sum_{l \in \mathbb{Z}^d} \exp(iq \cdot l/\sqrt{n})\, \rho_{0,l} v \right) \right|^2 \right\}
\end{aligned}$$

where a new integration variable $q := p\sqrt{n}$ has been introduced. Using that $|iE(p) + F(p) - iE'(p) - F'(p)|$ is of third order in p at $p \approx 0$, the desired estimate follows easily.

Now that a decay estimate for the L^2-difference $\|\rho_n^d - \rho_n\|_{L^2(d^d X)}$ has been derived, this has to be converted into the desired estimate for the decay of the L^1-difference $\|\rho_n^d - \rho_n\|_{L^1(d^d X)}$. To this end, we borrow some ideas of ref. 8:

For every $q \in \mathbb{N}$ there obviously exists a constant $C_q < \infty$ such that $E_{\rho_0}[|X_n - Bn|^{2q}] \leqslant C_q n^q$ for all $n \in \mathbb{N}$. Denote by $\mathscr{B}_M(X) \subset \mathbb{R}^d$ the ball with radius $M > 0$ and center X. Then for all $q \in \mathbb{N}$ and $n \in \mathbb{N}$

$$C_q n^q \geqslant E_{\rho_0}[|X_n - Bn|^{2q}] \geqslant M^{2q} \int_{\mathbb{R}^d - \mathscr{B}_M(Bn)} d^d X\, \rho_n^d(X)$$

Let vol_d be the volume of the unit ball in $\mathbb{R}^d$. Then

$$\begin{aligned}\|\rho_n^d-\rho_n\|_{L^1(d^dX)}&=2\int d^dX(\rho_n^d(X)-\rho_n(X))_+\\&\leqslant 2\int_{\mathscr{B}_M(Bn)} d^dX\,|\rho_n^d(X)-\rho_n(X)|+2\int_{\mathbb{R}^d-\mathscr{B}_M(Bn)} d^dX\,\rho_n^d(X)\\&\leqslant 2M^{d/2}(\mathrm{vol}_d)^{1/2}\|\,\rho_n^d-\rho_n\|_{L^2(d^dX)}+2C_q n^q M^{-2q}\end{aligned}$$

We know that $\|\rho_n^d-\rho_n\|_{L^2(d^dX)}\leqslant Cn^{-1/2-d/4}$ with some $C<\infty$ for all $n\in\mathbb{N}$. Hence, the choice $M:=n^{(q+1/2+d/4)/(d/2+2q)}$ results in

$$\|\rho_n^d-\rho_n\|_{L^1(d^dX)}\leqslant 2[(\mathrm{vol}_d)^{1/2}\,C+C_q]\,n^{-1/[2+d/(2q)]}$$

Thus, by choosing $\rho\in\mathbb{N}$ large enough, one can derive that $\|\rho_n^d-\rho_n\|_{L^1(d^dX)}$ decays faster than $n^{-\gamma}$ for all $\gamma<1/2$.

4. PROBABILISTIC MODELS

4.1. Markov Processes

Consider a Markov kernel Γ on $\mathbb{R}^d$ which describes the evolution of the density ρ_n of some probabilistic model:

$$\rho_{n+1}(X)=\int_{\mathbb{R}^d} d^dZ\,\Gamma(X,Z)\,\rho_n(Z) \tag{17}$$

Due to the periodicity of the deterministic system which we want to model, it is reasonable to restrict Γ to the periodic case:

$$\Gamma(X+l,Z)\overset{!}{=}\Gamma(X,Z-l)\qquad\text{for all}\quad X,Z\in\mathbb{R}^d\quad\text{and all}\quad l\in\mathbb{Z}^d$$

For our application, we have to examine the expression $\sum_{l\in\mathbb{Z}^d}e^{ip\cdot l}\rho_n(x+l)$. With the help of periodicity, this can be rewritten as follows:

$$\sum_{l\in\mathbb{Z}^d}e^{ip\cdot l}\rho_n(x+l)=P_p'^n\left(\sum_{l\in\mathbb{Z}^d}e^{ip\cdot l}\rho_0(\cdot+l)\right)$$

where the operator P_p' acts in $L^1(d^dx)$, i.e., on functions defined on $[0,1)^d$. It is given by

$$P_p'(f)(x):=\int_{\mathbb{R}^d} d^dZ\,\Gamma(x,Z)\,e^{-ip\cdot\lfloor Z\rfloor}f(\langle Z\rangle) \tag{18}$$

and thus is $L^1(d^dx)$-bounded by 1.

4.2. Diffusion Processes

Now, more specifically, consider an Itô diffusion in $\mathbb{R}^d$ with (strictly positive) diffusion tensor field $\mathscr{D}$ and drift field $\mathscr{B}$:

$$dX_t = \mathscr{B}(X_t)\,dt + \mathscr{D}^{1/2}(X_t)\,dW_t \tag{19}$$

where $\mathscr{D}^{1/2}$ denotes the positive square root of the positive symmetric matrix $\mathscr{D}$. This is a system with continuous time, but it can naturally be viewed as a system with discrete time, too, by considering the subset $X_0, X_1, X_2,...$ of a full trajectory $\{X_t\}_{t\geqslant 0}$.

In analogy to the deterministic problem which is to be modeled, we suppose that the spatial dependence of $\mathscr{B}$ and $\mathscr{D}$ is $\mathbb{Z}^d$-periodic. If, furthermore, $\mathscr{B}$ is of type C^1, if $\mathscr{D}$ is of type C^2, and if $\xi \cdot \mathscr{D}\xi$ is everywhere and for all $\xi \in \mathbb{R}^d$, $|\xi| = 1$, bounded from below by a positive constant, then we know[9] that there exists a strictly positive Markov kernel Γ of type C^1 for the evolution via Eq. (17) of the probability density of Eq. (19) in unit time. $\Gamma(X, Z)$ itself and also its first partial derivatives with respect to the components of X and Z are bounded in absolute value by $C_1 \exp(-C_2 |X-Z|^2)$ with appropriate $C_1, C_2 \in \mathbb{R}^+$. By the periodicity of $\mathscr{B}$ and $\mathscr{D}$ we find $\Gamma(X+l, Z) = \Gamma(X, Z-l)$ for all X, $Y \in \mathbb{R}^d$ and all $l \in \mathbb{Z}^d$. Note that P'_0 describes the evolution of probability densities under Eq. (19) as a diffusion process on the *torus* $\mathscr{T}^d$.

For P'_p the same spectral decomposition is valid which we have proved already for P_p, its deterministic counterpiece. To see this, note the following:

1. On the torus $\mathscr{T}^d$, the diffusion equation (19) possesses a unique invariant measure $d\mu'$ (Theorem 3.3.4 from ref. 2). Consequently, the eigenspace $\mathbb{C}\, d\mu'/d^dx$ of P'_0 to the eigenvalue 1 is one-dimensional. Define $h' := d\mu'/d^dx$. This function (i.e., the equilibrium density for the diffusion on the torus) is bounded from below by a positive constant d^dx-a.e. As a consequence of Theorem 3.3.2 from ref. 2, there exists no other eigenvalue of modulus 1.

2. For p not in $2\pi\mathbb{Z}^d$, there does not exist an $L^1(d^dx)$-eigenvector of P'_p belonging to an eigenvalue of modulus 1. To see this, assume that $P'_p(f) = e^{i\theta} f \neq 0$ for some $f \in L^1(d^dx)$ and some $\theta \in \mathbb{R}$. Then along the lines of Section 3.2 one can show that (after suitable normalization) d^dx-a.e. $f = e^{i\phi} h'$ with a real-valued function ϕ. From this follows

$$-p \cdot \lfloor Z \rfloor + \phi(\langle Z \rangle) - \theta - \phi(x) \in 2\pi\mathbb{Z}$$

for almost all $(x, Z) \in [0, 1)^d \times \mathbb{R}^d$. This can happen only if ϕ is constant d^dx-a.e., if $\theta = 0$ and

$$p \cdot \mathbb{Z}^d \overset{!}{\subset} 2\pi\mathbb{Z}$$

and hence if $p \in 2\pi\mathbb{Z}^d$.

3. P'_p depends analytically on p with respect to $\|\cdot\|_{\alpha,\beta}$ as a norm of the BV-functions defined on $\mathcal{T}^d$ instead of $\mathcal{T}^d \times \mathcal{T}^e$. To prove this, it is sufficient to prove that the sum in

$$P'_p(f)(x) = \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} p^{\otimes k} \cdot \int_{\mathbb{R}^d} d^dZ\, \Gamma(x, Z) \lfloor Z \rfloor^{\otimes k} f(\langle Z \rangle)$$

is absolutely convergent with respect to BV. This follows easily from the Gaussian decay of Γ and its partial derivatives. One even finds

$$\|P'_p(f)\|_{\alpha,\beta} \leqslant C\, \|f\|_1 \tag{20}$$

with a certain $C < \infty$ for all p with (for concreteness' sake) $|p| \leqslant 1$; so Theorem 3 can be applied to P'_p for all $p \in \mathbb{R}^d$.

These properties imply that the BV-spectral radius of P_p is strictly less than 1 for p not in $2\pi\mathbb{Z}^d$. Additionally, there is an $\varepsilon' > 0$ such that for $|p| \leqslant \varepsilon'$ we have a decomposition

$$P'^n_p = K'^n_p \Pi'_p + R'^n_p \tag{21}$$

for all $n \in \mathbb{N}$. Here, K'_p is the unique eigenvalue with largest modulus, Π'_p a one-dimensional BV-projector, and R'_p a bounded operator, all three of them analytic with respect to p. The BV-spectral radius of R'_p is strictly smaller than 1. The projector Π'_0 is given by $f \mapsto h' \int d^dx\, f(x)$. The modulus $|K'_p|$ is strictly less than 1 for $p \neq 0$. It can be written $K'_p = \exp[ip \cdot B' - p \cdot D'p/2 + iE'(p) + F(p)]$ with a $B' \in \mathbb{R}^d$, a strictly positive symmetric matrix $D' \in \mathbb{R}^{d \times d}$, and infinitely differentiable real-valued functions E' and F', which are of order 3 and 4, respectively, at $p = 0$.

Now apply these results—especially the decomposition (21)—to the expression

$$\sum_{l \in \mathbb{Z}^d} e^{ip \cdot l} \rho_n(x + l) = P'^n_p \left(\sum_{l \in \mathbb{Z}^d} e^{ip \cdot l} \rho_0(\cdot + l) \right)$$

The assumptions in the preceding sections which led to an approximation according to Section 3.3 can now obviously be fulfilled by the diffusion equation (19) at integer time steps if the following identities are valid:[5]

$$B \stackrel{!}{=} B', \qquad D \stackrel{!}{=} D' \tag{22}$$

$$\frac{d\mu}{d^d x} \stackrel{!}{=} \frac{d\mu'}{d^d x} \qquad d^d x\text{-a.e.} \tag{23}$$

The first two of these equations can readily be achieved by setting $\mathscr{B} := B$ and $\mathscr{D} := D$, i.e., by choosing *constant* coefficients in Eq. (19). But then, the equilibrium measure $d\mu'$ will not be correct in general. This problem can be overcome if the invariant density for the deterministic process of the torally restricted system is strictly positive and sufficiently smooth—namely, if there exists a strictly positive C^3-function h such that $d^d x$-a.e. $h = d\mu/d^d x$. Then one can proceed as follows: Consider the stochastic differential equation with constant coefficients

$$dY_t = B\,dt + D^{1/2}\,dW_t$$

and then take some appropriate C^3-diffeomorphism $X_t := \Phi(Y_t)$ of its random trajectories. By Itô's formula we know that X_t fulfills a stochastic differential equation in the form of Eq. (19) with appropriate $\mathscr{B}$ and $\mathscr{D}$. Choose for this transformation *such* a Φ of $\mathbb{R}^d$ that $\Phi(Y+l) = \Phi(Y) + l$ for all $Y \in \mathbb{R}^d$ and $l \in \mathbb{Z}^d$. Then the effective transport coefficients B' and D' [as in Eqs. (22) and (23)] will not be changed if we go from Y_t to X_t. The equilibrium measure of X_t as a diffusion on the torus is $|\det \partial\Psi/\partial x|$, $x \in [0, 1)^d \equiv \mathscr{T}^d$, where Ψ is the inverse of Φ. To fulfill Eq. (23) we need $|\det \partial\Psi/\partial X| = h(\langle X \rangle)$ for all $X \in \mathbb{R}^d$.

A simple way to construct such a C^3-invertible transformation $\Psi = (\Psi_1, \dots, \Psi_d)$ is the following: For $X = (X_1, \dots, X_d) \in \mathbb{R}^d$ define

$$\Psi_k(X) := \frac{\int_0^{X_k} dY_k \int_0^1 dy_{k+1} \cdots \int_0^1 dy_d\, h(\langle X_1 \rangle, \dots, \langle X_{k-1} \rangle, \langle Y_k \rangle, y_{k+1}, \dots, y_d)}{\int_0^1 dy_k \int_0^1 dy_{k+1} \cdots \int_0^1 dy_d\, h(\langle X_1 \rangle, \dots, \langle X_{k-1} \rangle, y_k, \dots, y_d)}$$

where $k = 1, \dots, d$. This expression is well defined, because from the assumptions it follows that h is bounded from below by a positive constant.

4.3. Beyond a Decay of Order −1/2?

Consider again the deterministic system described by Eq. (4): $X_{n+1} = X_n + \langle X_n \rangle$. The Perron–Frobenius operator P of the underlying mapping $T\colon x \mapsto \langle 2x \rangle$ acts on a function f by

$$P(f)(x) = \frac{1}{2}\left[f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right) \right]$$

[5] In the calculations of the preceding sections one has to replace ε by ε' if the latter is lower.

A straightforward calculation shows that the theorem of Ionescu-Tulcea and Marinescu[6] can for all α and $\beta \in (0, 1/2)$ be applied to P as an operator in $BV_{\alpha,\beta}$. The operator P_p, which generates the characteristic function of X_n, is given by

$$P_p(f)(x) = \frac{1}{2}\left[f\left(\frac{x}{2}\right) + e^{ip} f\left(\frac{x+1}{2}\right)\right]$$

Obviously, the constant functions are eigenfunctions of P_p with the eigenvalue $(1+e^{ip})/2$; this is the leading eigenvalue for small p. In the decomposition (14) we have therefore $K_p = (1+e^{ip})/2$, and Π_p is a projector onto the space of constant functions. By this we can immediately give a diffusion model for the behavior of this system:

$$dX_t = -i\frac{dK_p}{dp}\bigg|_{p=0} dt + \left(-\frac{dK_p}{dp}\bigg|_{p=0}^2 - \frac{d^2K_p}{dp^2}\bigg|_{p=0}\right)^{1/2} dW_t = \frac{1}{2}\,dt + \frac{1}{2}\,dW_t$$

According to the preceding subsection, this yields such an approximation that the $L^1(dX)$-distance between the density of the deterministic system and the density of this diffusion decays faster than $n^{\varepsilon - 1/2}$ as $n \to \infty$ for all $\varepsilon > 0$.

Can this decay velocity be enlarged? For that, it is necessary to approximate the projector Π_p up to *first* order in $p \approx 0$ by a projector Π'_p generated by a Markov process with smooth transition probabilities. But this is not possible: We are going to show that $d\Pi_p/dp\,|_{p=0}$ *cannot* be extended to a bounded operator in L^1, while $d\Pi'_p/dp\,|_{p=0}$ *can* (for Markov processes which are "smooth" in a way to be defined). Thus, these derivatives can never coincide.

First consider Π_p. Denote the function $x \mapsto \exp(2\pi i 2^n x)$ by f_n. For N, $n \in \mathbb{N}_0$, $N > n$, we have $P_p^N(f_n) = K_p^n P_p^{N-n}(f_1)$, so that, for $|p|$ small enough

$$\Pi_p(f_n) = \lim_{N\to\infty} K_p^{-N} P_p^N(f_n) = \Pi_p(f_1)$$

A short computation reveals that $\Pi_p(f_n)$ equals the constant function

$$\frac{1-e^{ip}}{1+e^{ip}} \prod_{m=2}^{\infty} (1 + e^{ip + 2\pi i 2^{-m}})/(1+e^{ip})$$

and therefore $d\Pi_p/dp\,|_{p=0}$ applied to f_1 is the constant function $-i/2 \prod_{m=2}^{\infty} [1 + \exp(2\pi i 2^{-m})]/2$. Now again consider the one-dimen-

[6] It is necessary to work with a suitable power P^m of P. Besides, note that the *measure-theoretic* properties of $x \mapsto \langle 2x \rangle$ can be deduced readily from the circumstance that this mapping is conjugated to a Bernoulli shift.

sional projector Π_p written as $f \mapsto 1\langle g_p, f\rangle$, where g_p is a continuous linear form on BV. Assume that the BV-continuous linear form $dg_p/dp|_{p=0}$ *could* be identified with a continuous linear form on L^1 and, hence, a function in L^∞. Hence, its Fourier coefficients are square-summable, which implies the finiteness of

$$\sum_{n \in \mathbb{Z}} |\langle dg_p/dp|_{p=0}, f_n\rangle|^2 \geqslant \sum_{n \in \mathbb{N}} \left| -i/2 \prod_{m=2}^{\infty} [1+\exp(2\pi i 2^{-m})]/2 \right|^2 = \infty$$

which is a contradiction. So we find that $d\Pi_p/dp|_{p=0}$ cannot be extended to a bounded operator in L^1.

Next we show that such a singular behavior cannot be reproduced by Markov processes whose transition kernel Γ is not too singular: Assume that Γ is of type C^1 and that $\Gamma(X, Z)$ and all its partial derivatives with respect to X are bounded in modulus by $C_1 \exp(-C_2 |X-Z|^2)$ for some $C_1, C_2 \in \mathbb{R}^+$. Then the analysis given before for an Itô diffusion can be carried over. We are exclusively interested in the case where there exists only a one-dimensional eigenspace of P'_0 with eigenvalue of modulus 1 (namely, 1 itself). The theorem of Ionescu-Tulcea and Marinescu (together with analyticity arguments) shows that under this assumption a decomposition

$$P_p'^n = K_p'^n \Pi'_p + R_p'^n \qquad \text{for all} \quad n \in \mathbb{N}$$

is valid for small $|p|$, where Π'_p is a one-dimensional BV-projector and where the BV-spectral radius of R'_p is strictly less than 1. Furthermore $\Pi'_p R'_p = 0 = R'_p \Pi'_p$. From this follows

$$\begin{aligned} \left.\frac{d\Pi'_p}{dp}\right|_{p=0} &= \left.\frac{dK_p'^{-1}\Pi'_p P'_p}{dp}\right|_{p=0} \\ &= \left.\frac{dK_p'^{-1}}{dp}\right|_{p=0} \Pi'_0 P'_0 + \left.\frac{d\Pi'_p}{dp}\right|_{p=0} P'_0 + \Pi'_0 \left.\frac{dP'_p}{dp}\right|_{p=0} \end{aligned}$$

The second term on the r.h.s. is a bounded operator in L^1, because P'_0 is, by Eq. (20), a bounded operator from L^1 into BV and because $d\Pi'_p/dp|_{p=0} P'_0$ is a bounded operator in BV. The third term on the r.h.s. is a bounded operator in L^1, because $dP'_p/dp|_{p=0}$ itself is bounded in L^1 [which can easily be proved with the help of the explicit representation (18) and the Gaussian decay of Γ].

5. CONCLUSIONS. OUTLOOK

We have shown the following: If Eq. (11) is valid and if there exists a C^3-function $h > 0$ such that $d^d x$-a.e. $d\mu/d^d x = h$, then one can model the deterministic system (1) by a diffusion process according to Eq. (19). In this case, $\|\rho_n^d - \rho_n\|_{L^1(d^dX)}$ decays faster than $n^{-\gamma}$ for all $\gamma < 1/2$. An example showed that this decay velocity can generally not be enlarged by considering (instead of diffusion) arbitrary Markov processes with transition probabilities which are "smooth" in a suitable way.

A future paper will be devoted to similar results for systems which are periodic extensions of Anosov diffeomorphisms instead of piecewise expanding mappings; an extension of the operator formalism to systems with continuous time is currently being investigated.

APPENDIX. PROOF OF THEOREM 5

Proof. Fix an $\varepsilon > 0$. Then for $f \in BV_{\alpha,\beta}$ there exists $\tilde{f}$ such that $d\lambda$-a.e. $\tilde{f} = f$, everywhere $|\tilde{f}| \leqslant \|f\|_\infty$, and

$$\int d\lambda(z) \sup_{z_{1,2}:\, d(z, z_{1,2}) < t} |\tilde{f}(z_1) - \tilde{f}(z_2)| \leqslant (1+\varepsilon)\, t^\alpha \operatorname{var}_{\alpha,\beta}(f) \quad \text{for all } t \in (0, \beta)$$

By Eq. (8) one finds for all z, z_1, and z_2

$$\begin{aligned}
&|P(\tilde{f})(z_1) - P(\tilde{f})(z_2)| \\
&\quad \leqslant \sum_i 1_{TA_i}(z)\, J_i(z)\, |1_{TA_i}(z_1)\, \tilde{f} \circ T_i^{-1}(z_1) - 1_{TA_i}(z_2)\, \tilde{f} \circ T_i^{-1}(z_2)| \\
&\qquad + \sum_i 1_{TA_i}(z_1)\, |\tilde{f} \circ T_i^{-1}(z_1)| \cdot |1_{TA_i}(z_1)\, J_i(z_1) - 1_{TA_i}(z)\, J_i(z)| \\
&\qquad + \sum_i 1_{TA_i}(z_2)\, |\tilde{f} \circ T_i^{-1}(z_2)| \cdot |1_{TA_i}(z_2)\, J_i(z_2) - 1_{TA_i}(z)\, J_i(z)|
\end{aligned}$$

Therefore, $\operatorname{var}_{\alpha,\beta}(Pf)$ is bounded by

$$\sup_{0 < t < \beta} t^{-\alpha} \sum_i \int_{TA_i} d\lambda(z)\, J_i(z)$$

$$\times \sup_{z_{1,2}:\, d(z, z_{1,2}) < t} |1_{TA_i}(z_1)\, \tilde{f} \circ T_i^{-1}(z_1) - 1_{TA_i}(z_2)\, \tilde{f} \circ T_i^{-1}(z_2)| \tag{A1}$$

$$+ 2 \sup_{0 < t < \beta} t^{-\alpha} \sum_i \int d\lambda(z) \sup_{z_1:\, d(z, z_1) < t} 1_{TA_i}(z_1)\, |\tilde{f} \circ T_i^{-1}(z_1)|$$

$$\times |1_{TA_i}(z_1)\, J_i(z_1) - 1_{TA_i}(z)\, J_i(z)| \tag{A2}$$

First, we estimate expression (A1). Let $\mathscr{B}_t$ denote the open ball with radius t around $0 \in \mathbb{R}^{d+e}$. We split TA_i into two subsets:

$$TA_i = (TA_i - (\partial TA_i + \mathscr{B}_t)) \cup (TA_i \cap (\partial TA_i + \mathscr{B}_t))$$

where the addition of $\mathscr{B}_t$ is to be understood mod $\mathbb{Z}^{d+e}$. Expression (A1) may now be written

$$\sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{TA_i-(\partial TA_i+\mathscr{B}_t)} d\lambda(z)\, J_i(z)$$
$$\times \sup_{z_{1/2}: d(z, z_{1/2})<t} |\tilde{f}\circ T_i^{-1}(z_1)-\tilde{f}\circ T_i^{-1}(z_2)|$$
$$+2\,\|f\|_\infty \sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{TA_i\cap(\partial TA_i+\mathscr{B}_t)} d\lambda(z)\, J_i(z)$$

In the first member on the r.h.s., the set in the supremum is enlarged (due to the expansion property) if we replace $\{z_{1/2}: d(z, z_{1/2})<t\}$ by $\{z_{1/2}: d(T_i^{-1}(z), T_i^{-1}(z_{1/2}))<t/\gamma\}$. We rename $T_i^{-1}(z_{1/2})$ as $z_{1/2}$, which yields by Eq. (8) that expression (A1) is bounded by

$$\sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{TA_i} d\lambda(z)\, J_i(z) \sup_{z_{1/2}: d(T_i^{-1}(z), z_{1/2})<t/\gamma} |\tilde{f}(z_1)-\tilde{f}(z_2)|$$
$$+2\,\|f\|_\infty \sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{TA_i} d\lambda(z)\, J_i(z)\, 1_{T_i^{-1}(TA_i\cap(\partial TA_i+\mathscr{B}_t))}(T_i^{-1}(z))$$
$$= \sup_{0<t<\beta} t^{-\alpha} \int d\lambda\, P(z' \mapsto \sup_{z_{1/2}: d(z', z_{1/2})<t/\gamma} |\tilde{f}(z_1)-\tilde{f}(z_2)|)$$
$$+2\,\|f\|_\infty \sup_{0<t<\beta} t^{-\alpha} \int d\lambda\, P(1_{\bigcup_i T_i^{-1}(TA_i\cap(\partial TA_i+\mathscr{B}_t))})$$
$$\leqslant \gamma^{-\alpha}(1+\varepsilon)\operatorname{var}_{\alpha,\beta}(f)+2\,\|f\|_\infty \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\bigcup_i A_i\cap(\partial A_i+\mathscr{B}_{t/\gamma})\right)$$

where we have used in the last step that $T_i^{-1}(TA_i\cap(\partial TA_i+\mathscr{B}_t))\subset A_i\cap(\partial A_i+\mathscr{B}_{t/\gamma})$. Note furthermore that

$$\sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\bigcup_i A_i\cap(\partial A_i+\beta_{t/\gamma})\right)$$
$$\leqslant \gamma^{-\alpha} \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\bigcup_i A_i\cap(\partial A_i+\mathscr{B}_t)\right)$$

Now we estimate expression (A2). Split the integration domain into the two subsets

$$(TA_i-(\partial T\, A_i+\mathscr{B}_t))\cup(TA_i-(\partial T\, A_i+\mathscr{B}_t))^c$$

Expression (A2) can thus be estimated from above by

$$2 \sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{TA_i-(\partial TA_i+\mathscr{B}_t)} d\lambda(z) \sup_{z_1:d(z,z_1)<t} |\tilde{f}\circ T_i^{-1}(z_1)| \cdot |J_i(z_1)-J_i(z_2)|$$

$$+2 \sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{(TA_i-(\partial TA_i+\mathscr{B}_t))^c} d\lambda(z) \sup_{z_1:d(z,z_1)<t} 1_{TA_i}(z_1)\,|\tilde{f}\circ T_i^{-1}(z_1)|$$

$$\times |1_{TA_i}(z_1)\,J_i(z_1)-1_{TA_i}(z)\,J_i(z)|$$

Note that the integration domain of the second member may be restricted to $\partial T A_i+\mathscr{B}_t$. In the first member we can apply the Hölder continuity of the Jacobi determinant to find that expression (A2) is bounded by

$$2 \sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{TA_i-(\partial TA_i+\mathscr{B}_t)} d\lambda(z) \sup_{z_1:d(z,z_1)<t} |\tilde{f}\circ T_i^{-1}(z_1)|\, c(2t)^\alpha$$

$$+2\,\|f\|_\infty \sup_{0<t<\beta} t^{-\alpha} \sum_i \int_{\partial TA_i+\mathscr{B}_t} d\lambda(z)$$

$$\times \sup_{z_1:d(z,z_1)<t} |1_{TA_i}(z_1)\,J_i(z_1)-1_{TA_i}(z)\,J_i(z_2)|$$

$$\leqslant 2^{1+\alpha}cC \sup_{0<t<\beta} \sum_i \int_{TA_i-(\partial TA_i+\mathscr{B}_t)} d\lambda(z)\,J_i(z) \sup_{z_1:d(z,z_1)<t} |\tilde{f}\circ T_i^{-1}(z_1)|$$

$$+4\,\|f\|_\infty\, \gamma^{-d-e} \sum_i \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}(\partial T A_i+\mathscr{B}_t)$$

where we have "artificially" introduced the Jacobi determinant $J_i(z)$. The upper line of the preceding expression is bounded by

$$2^{1+\alpha}cC \sup_{0<t<\beta} \sum_i \int_{TA_i-(\partial TA_i+\mathscr{B}_t)} d\lambda(z)\,J_i(z)\,|\tilde{f}\circ T_i^{-1}(z)|$$

$$+2^{1+\alpha}cC \sup_{0<t<\beta} \sum_i \int_{TA_i-(\partial TA_i+\mathscr{B}_t)} d\lambda(z)\,J_i(z)$$

$$\times \sup_{z_{1/2}:d(z,z_{1/2})<t} |\tilde{f}\circ T_i^{-1}(z_1)-\tilde{f}\circ T_i^{-1}(z_2)|$$

$$\leqslant 2^{1+\alpha}cC \int d\lambda(z)\,P(|\tilde{f}|)$$

$$+2^{1+\alpha}cC \sup_{0<t<\beta} \int d\lambda\, P(z'\mapsto \sup_{z_{1/2}:d(z',z_{1/2})<t/\gamma} |\tilde{f}(z_1)-\tilde{f}(z_2)|)$$

$$\leqslant 2^{1+\alpha}cC[\|f\|_1+\gamma^{-\alpha}(1+\varepsilon)\operatorname{var}_{\alpha,\beta}(f)]$$

Now collect all estimates for the expressions (A1) and (A2), take $\varepsilon \downarrow 0$, and apply property 2 of the space $BV_{\alpha,\beta}$—namely, that there exist $C_1, C_2 < \infty$ such that

$$\|f\|_\alpha \leqslant C_1 \|f\|_1 + C_2 \operatorname{var}_{\alpha,\beta}(f) \qquad \text{for all} \quad f \in BV_{\alpha,\beta}$$

Thus, for A and B in the theorem one can choose

$$A := 1 + 2C_2 \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\bigcup_i A_i \cap (\partial A_i + \mathscr{B}_t)\right)$$
$$+ 2^{1+\alpha} cC + 4C_2 \sum_i \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}(\partial T A_i + \mathscr{B}_t)$$

$$\mathrm{B} := 2C_1 \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\bigcup_i A_i \cap (\partial A_i + \mathscr{B}_t)\right)$$
$$+ 2^{1+\alpha} cC + 4C_1 \sum_i \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}(\partial T A_i + \mathscr{B}_t)$$

It remains to show that these A, B are *finite*.

To this end, note that

$$\sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\bigcup_i A_i \cap (\partial A_i + \mathscr{B}_t)\right) \leqslant \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\left(\bigcup_i \partial A_i\right) + \mathscr{B}_t\right) \tag{A3}$$

and

$$\sum_i \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}(\partial T A_i + \mathscr{B}_t) \leqslant \#\{A_i\} \sup_{0<t<\beta} t^{-\alpha} \operatorname{vol}\left(\left(\bigcup_i \partial T A_i\right) + \mathscr{B}_t\right) \tag{A4}$$

The statement that A and B are finite if the upper capacities of $\bigcup_i \partial A_i$ and $\bigcup_i \partial T A_i$ are strictly smaller than $d + e - \alpha$ and if β is small enough can now be proved as for Theorem 1. ■

Note that in the above proof the finiteness of the partition $\{A_i\}$ is only used to show that the l.h.s. of Eq. (A4) is finite. Therefore, also the case of a countably *infinite* partition can be covered, given that this purely geometric expression is finite for some other reason.

ACKNOWLEDGMENTS

I thank Prof. Dr. Philippe Blanchard for discussions. This work was supported by Studienstiftung des deutschen Volkes.

REFERENCES

1. H. D. I. Abarbanel, Non-linear diffusion in Hamiltonian systems exhibiting chaotic motion, *Physica D* **4**:89 (1981).
2. A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotical Analysis for Periodic Structures* (North-Holland, Amsterdam, 1978).
3. M. L. Blank, Stochastic properties of deterministic dynamical systems, *Sov. Sci. Rev. C* **6**:243 (1987).
4. M. L. Blank, Small perturbations of chaotic dynamical systems, *Russ. Math. Surv.* **44**:1 (1989).
5. M. L. Blank, Chaotic mappings and stochastic Markov chains, in *Mathematical Physics X*, K. Schmüdgen, ed. (Springer, Berlin, 1992), p. 341.
6. R. Bowen, Bernoulli maps of the interval, *Israel J. Math.* **28**:161 (1977).
7. L. A. Bunimovich and Ya, G. Sinai, Statistical properties of Lorentz gas with periodic configurations of scatterers, *Commun. Math. Phys.* **78**:479 (1981).
8. L. Devroye, *A Course in Density Estimation* (Birkhäuser, Boston, 1987).
9. A. Friedman, *Stochastic Differential Equations and Applications, Vol. I* (Academic Press, New York, 1975).
10. P. Gaspard, Diffusion, effusion, and chaotic scattering: An exactly solvable Liouvillian dynamics, *J. Stat. Phys.* **68**:673 (1992).
11. Y. Guivarc'h and J. Hardy, Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov, *Ann. Inst. H. Poincaré* **24**:73 (1988).
12. T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
13. G. Keller, Generalized bounded variation and application to piecewise monotonic transformations, *Z. Wahrsch. Verw. Geb.* **69**:461 (1985).
14. A. Knauf, Ergodic and topological properties of periodic Coulombic potentials, *Commun. Math. Phys.* **110**:89 (1987).
15. A. Lasota and M. C. Mackey, *Probabilistic Properties of Deterministic Systems* (Cambridge University Press, 1985).
16. F. Norman, *Markov Processes and Learning Models* (Academic Press, New York, 1972).
17. N. W. Murray, M. A. Lieberman, and A. J. Lichtenberg, Corrections to quasilinear diffusion in area-preserving maps, *Phys. Rev. A* **32**:2413 (1985).
18. J. Rousseau-Egele, Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux, *Ann. Prob.* **11**:772 (1983).