ELSEVIER

Physica D 104 (1997) 32–60

PHYSICA D

Diffusion approximation and hyperbolic automorphisms of the torus

Claude Bardos [a], François Golse [b,*], Jean-François Colonna [c]

[a] Université Paris VII and Ecole Normale Supérieure de Cachan, CMLA, 61 av. Président Wilson, 94235 Cachan, France
[b] Université Paris VII and Ecole Normale Supérieure, DMI, 45 rue d'Ulm, 75005 Paris, France
[c] LACTAMME, CNET and Ecole Polytechnique, 91128 Palaiseau, France

Received 19 December 1994; revised 15 August 1995; accepted 20 September 1995
Communicated by H. Flaschka

Abstract

In this article a diffusion equation is obtained as a limit of a reversible kinetic equation scaled appropriately. This limiting diffusion is produced by the collisions of the particles with the boundary. Indeed, these particles follow a reversible reflection law having convenient mixing properties. This model, based on "Arnold's cat map", can be handled with Fourier series instead of the symbolic dynamics associated to a Markov partition. As a consequence, optimal convergence results can be obtained by elementary means and illustrate the apparition of irreversibility in macroscopic limits.

0. Introduction

"Irreversibility" is a fundamental notion in Statistical Mechanics. It is also a misleading one, and we feel useful to recall some basic facts. In doing so, we follow suggestions from the referee, whom we thank for his help in clarifying the presentation of this paper.

Consider a mechanical system made of a (large) number N of identical point particles; denote by $x_j(t)$ and $v_j(t)$, respectively, the position of the jth particle at time t and its velocity. Starting at time $t = 0$ with an initial state of the system denoted by $(x_j(0), v_j(0))_{1\le j\le N}$, one lets the system evolve until time $t > 0$ into $(x_j(t), v_j(t))_{1\le j\le N}$. One then changes instantaneously the velocity of each particle into its opposite, which leads to the new state $(x_j(t), -v_j(t))_{1\le j\le N}$, and lets the system evolve further until time $2t$. If the final state thus reached is identical to $(x_j(0), v_j(0))_{1\le j\le N}$, the dynamics considered will be called "reversible"; otherwise it will be called irreversible.

One of the central issues in nonequilibrium statistical mechanics is the derivation of macroscopic dynamics such as hydrodynamics from first principle equations. For example, the Euler or Navier–Stokes systems for a perfect ideal gas can be derived from the Boltzmann equation for, say, a gas of hard spheres, which in turn can be derived from the Liouville equation (or equivalently the system of Hamilton's equations) for the N-body classical hamiltonian with hard-sphere interactions. These derivations are mostly "formal", except in particular regimes where mathematical

* Corresponding author. E-mail: francois.golse@ens.fr.

PII S0167-2789(96)00128-5

proofs have been given: see for example [1–3,9,10,13,17,19]. There are structural obstacles to a global understanding of these questions. One difficulty is that the partial differential equations (PDEs) involved are mostly nonlinear. Another conceptual difficulty is related to the notion of irreversibility. Indeed, whereas the Liouville equation is reversible in the sense described above, neither the Boltzmann equation nor the Euler and Navier–Stokes systems are. Indeed, these models contain in some sense Carnot's "Second Principle of Thermodynamics", and this can be seen most easily by noticing that these models have a natural notion of entropy which increases until the system reaches an equilibrium.

It appears therefore desirable to isolate the difficulties described above, and for one thing, to discuss the apparition of irreversibility on macroscopic dynamics on examples which can be analyzed as explicitly as possible (which, by the way, almost rules out dynamics described by nonlinear PDEs). Such an example has been studied in a remarkable series of papers by Bunimovich and Sinai [5,6] and more recently by Bunimovich et al. [7,8]. They have considered the periodic Lorentz gas, i.e. the dynamics of a population of point particles interacting with a periodic array of infinitely heavy circular obstacles through elastic collisions. They showed that, under the assumption that the billiard has the property of "finite horizon" (see [6] for more details) the long time, large scale limiting dynamics of these particles is described by a linear diffusion equation with diffusion constant given by the Green–Kubo formula. However, their analysis relies on sophisticated tools from ergodic theory, namely the construction of a "Markov sieve" for the billiard dynamics thus generated.

The goal of the present paper is to give another example of a "reversible" dynamics, which, after some appropriate scaling, leads to an "irreversible" limiting dynamics. In particular, we believe that it might be of some interest to provide such an example that could be completely analyzed by elementary and explicit techniques and without appealing to general and abstract constructions from ergodic theory, as in the work of Bunimovich et al.

Before describing the particular (class of) models described in this paper, we want to make as clear as possible some very simple remarks concerning irreversibility. We refer in particular to [16,20] for a precise and yet elementary presentation of these notions.

First, we insist that, although the model treated by Bunimovich et al. or ours give some insight on the apparition of irreversibility when macroscopic limits are taken, it is likely that the apparition of irreversibility when the Boltzmann equation is derived from Classical Mechanics is governed by a different mechanism. Indeed, in this latter case, there are two different notions of entropy involved:

(a) *Gibbs entropy*

$$S_G(t) = -\int f_t^N(x_1, v_1, \ldots, x_N, v_N) \log f(x_1, v_1, \ldots, x_N, v_N)\, \mathrm{d}x_1\, \mathrm{d}v_1 \cdots \mathrm{d}x_n\, \mathrm{d}v_N,$$

where f_N is the solution of the N-body Liouville equation

(b) *Boltzmann H function*

$$H(t) = -\int F_t(x_1, v_1) \log F(x_1, v_1)\, \mathrm{d}x_1\, \mathrm{d}v_1,$$

where F is the number density which is the solution of the Boltzmann equation.

It is believed (and in some cases proved) that, under the assumption that the number of particles tends to infinity as the radius of each particle tends to zero according to the very particular Boltzmann–Grad limit (see [10]), and if the initial state is factorized, i.e.

$$f_0^N(x_1, v_1, \ldots, x_N, v_N) = F_0(x_1, v_1) \cdots F_0(x_N, v_N),$$

then at further times the solution of the Liouville equation will remain approximately so with F being a solution of the Boltzmann equation. In other words, the first marginal of the solution of the N-body Liouville equation follows approximately the dynamics prescribed by the Boltzmann equation. However

- like any integral of any function of f^N, $S_G(t) = S_G(0)$ (a consequence of Liouville's theorem for hamiltonian dynamics);
- by Boltzmann's H theorem, $H(t)$ increases until F_t reaches an equilibrium state.

Observe however that for a factorized state

$$S_G(t) = -\int F_t(x_1, v_1)\cdots F_t(x_N, v_N)\log\left(F_t(x_1, v_1)\cdots F_t(x_N, v_N)\right)\mathrm{d}x_1\,\mathrm{d}v_1\cdots\mathrm{d}x_n\,\mathrm{d}v_N$$
$$= NH(t)\left(\int F_t(x, v)\,\mathrm{d}x\,\mathrm{d}v\right)^{N-1}$$

so that, in the limit as $N \to +\infty$, one cannot infer the behavior of the Boltzmann entropy from that of the Gibbs entropy. Also, as explained in [20], irreversibility is described by the Boltzmann entropy and not by the Gibbs entropy. Notice however, that one can hardly speak of apparition of irreversibility in this case, since irreversibility is introduced from the beginning by considering the marginals of the solution of the Liouville equation.

After this lengthy introduction on irreversibility, let us explain in which sense the model that we analyze below (as well as the one studied by Bunimovich et al.) might provide some understanding of the apparition of irreversibility as macroscopic limits are taken.

Our model will be reversible in the following sense: it is described by a continuous unitary group of $L^2(X, \mu)$ where X is a phase space of the form $\mathbb{R}^d \times Y$ and μ a measure on X whose integral along the fibers Y gives the Lebesgue measure on $\mathbb{R}^d$. When appropriately scaled, its long time, large scale limiting evolution is governed by the heat equation on $\mathbb{R}^d$ which is the archetype of irreversible systems and generates a continuous contraction semigroup on $L^2(\mathbb{R}^d)$. Moreover, as is well known, the L^2 norm of the solution of the heat equation decreases unless the solution is in equilibrium (the only equilibrium solution in $L^2(\mathbb{R}^d)$ being 0). What happens can be explained very easily in terms of weak topologies: denote by G_t^ϵ the unitary group corresponding to the scaled reversible evolution and by $\exp(tD\Delta)$ the diffusion semigroup with diffusion constant D (here $\epsilon > 0$ denotes a small scaling parameter, and the macroscopic limit corresponds to letting $\epsilon \to 0$). For all $t > 0$ and all smooth initial data $\phi \in C_0^\infty(\mathbb{R}^d)$, the scaled family $G_t^\epsilon \cdot \phi$ converges weakly to $\exp(tD\Delta)\cdot\phi$, but not in the norm topology. As a consequence, nonlinear functionals of $G_t^\epsilon \cdot \phi$ do not converge to the corresponding expressions with $\exp(tD\Delta)\cdot\phi$ in place of $G_t^\epsilon\cdot\phi$. This applies in particular to the L^2 norm, and this explains how the L^2 norm of the limiting solution decreases while that of the scaled microscopic dynamics (the one corresponding to the unitary group G_t^ϵ) remains constant under the evolution.

Before describing in detail the mechanical system that we shall analyze in this paper, we conclude this impressionistic survey of the notion of irreversibility as related to our model by recalling for readers who might have more background in Physics than in Mathematics the notion of weak convergence which is central to our discussion, as can be understood from the explanation above.

We believe that the most intuitive way of understanding weak topologies in Hilbert spaces consists in looking at the particular model $H = L^2(S^1)$, the space of square-integrable functions on the circle (or equivalently defined on the real line $\mathbb{R}$ and periodic of period, say, one). Then, functions can be thought as superpositions of elementary waves via the theory of Fourier series. For a function $u \in H$, we denote by $\hat{u}(k)$ its Fourier coefficient of order $k \in \mathbb{Z}$. Then a family of functions (u_ϵ) indexed by $\epsilon \in]0, 1[$ is said to:

(i) converge to u in the weak topology of H if and only if, for all $k \in \mathbb{Z}$, the family of Fourier coefficients of order k, $\hat{u}_\epsilon(k)$ converges to $\hat{u}(k)$ as $\epsilon \to 0$;

(ii) converge to u in the strong (or norm) topology of H if and only if, as $\epsilon \to 0$,

$$\int_{S^1} |u_\epsilon(x) - u(x)|^2\,\mathrm{d}x = \sum_{k\in\mathbb{Z}} |\hat{u}_\epsilon(k) - \hat{u}(k)|^2 \to 0.$$

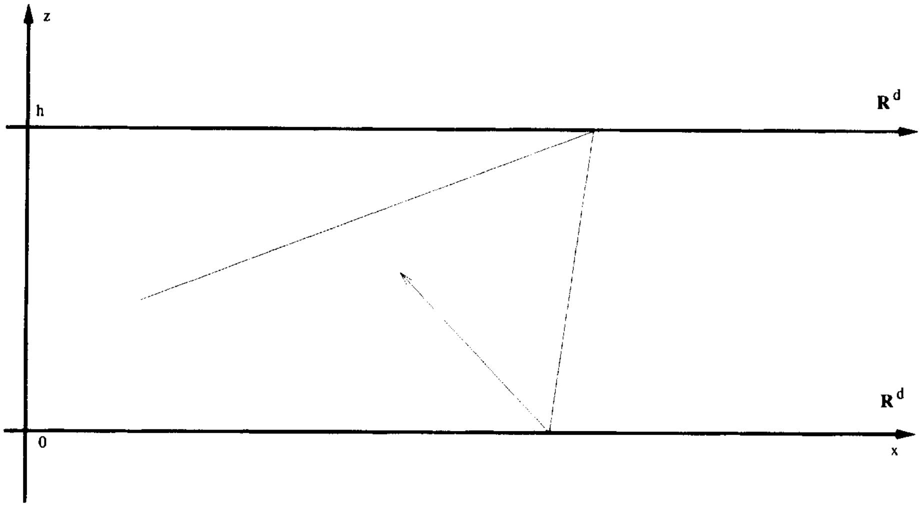

Fig. 1.

Another formulation of (i) is as follows:
(i′) the family (u_ϵ) converges to u in the weak topology of H as $\epsilon \to 0$ if and only if, for any interval I of S^1 one has

$$\int_I u_\epsilon(x)\,\mathrm{d}x \to \int_I u(x)\,\mathrm{d}x.$$

This makes more intuitive the relation between weak convergence and coarse graining.

It is clear from this presentation that strong convergence implies weak convergence, whereas the converse is wrong, unless additional assumptions are made on the decay of the Fourier coefficients at high wave numbers (uniform – in ϵ – smallness as the wave number tends to infinity). Since nonlinear functionals of u_ϵ would involve a nonlocal interaction of the Fourier coefficients of u_ϵ, it is also clear that such nonlinear functionals of u_ϵ, as for example, the quadratic mean, are not continuous with respect to the weak topology. We refer to [15] for a comprehensive introduction to the notion of weak convergence in L^p spaces.

To conclude this list of remarks, we hope to have made clear what is meant here by reversibility and irreversibility, and how the way these notions are used in the present work differs from the more classical situation considered in statistical mechanics, i.e. the derivation of kinetic theory from hamiltonian dynamics.

1. Presentation of the model

Between two horizontal plates, a family of particles evolve as a Knudsen gas, i.e. a gas with no interparticle collisions. The vertical components of the velocities of the particles are all assumed to have modulus $c > 0$. The horizontal components of their velocities $ca(\omega)$ are parametrized by $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$. Whenever the particles hit the top or bottom plate, their vertical velocities are changed into their opposite while their horizontal velocities are modified by the right action of a hyperbolic automorphism of $\mathbb{T}^2$ (cf. Fig. 1).

We shall use the following notations: the position of the particles is denoted by $(x, z) \in \mathbb{R}^d \times (0, h)$ and the vertical component of the velocity of the particles by $\pm c$. The horizontal component of this velocity is given by $ca(\omega)$, $\omega \in \mathbb{T}^2$, where $a : \mathbb{T}^2 \to \mathbb{R}^d$ denotes a smooth zero mean vector field. The nonnegative functions $f_+(t, x, z, \omega)$

(resp. $f_-(t, x, z, \omega)$) represent the number density of particles which at time t occupy the position (x, z) and move with velocity $(ca(\omega), +c)$ (resp. $(ca(\omega), -c)$).

The following hyperbolic automorphism T of the torus (Arnold's cat map) defined by

$$T\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix} = \begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix} \pmod{2\pi} \tag{1}$$

will be the only case treated here for computational simplicity. However, it will be clear throughout the paper that our method would apply to any hyperbolic automorphism of $\mathbb{T}^n$. The map $T : \mathbb{T}^2 \to \mathbb{T}^2$ is one-to-one and C^∞; it preserves the measure $d\omega_1\, d\omega_2/4\pi^2$ and its inverse (which also is a C^∞ map) is given by

$$T^{-1}\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix} = \begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix} \pmod{2\pi}. \tag{2}$$

The densities $f^\pm$ satisfy the Liouville system of equations

$$\partial_t f^\pm + ca(\omega)\cdot\partial_x f^\pm \pm c\partial_z f^\pm = 0, \quad x \in \mathbb{R}^d, \quad 0 < z < h, \quad \omega \in \mathbb{T}^2 \tag{3}$$

with the following boundary conditions on the plates:

$$f^+(t, x, 0, \omega) = f^-(t, x, 0, T\omega), \quad x \in \mathbb{R}^d, \quad \omega \in \mathbb{T}^2, \tag{4a}$$

$$f^-(t, x, h, \omega) = f^+(t, x, h, T\omega), \quad x \in \mathbb{R}^d, \quad \omega \in \mathbb{T}^2. \tag{4b}$$

Their value at $t = 0$ is given by the following initial condition:

$$f^\pm(0, x, z, \omega) = \phi(x), \quad x \in \mathbb{R}^d, \quad 0 < z < h, \quad \omega \in \mathbb{T}^2, \tag{5}$$

which is compatible with an approximation, as $h \to 0$, by a horizontal diffusion, and avoids initial layers in the limiting process.

Since the densities $f^\pm$ satisfy Eq. (3), they are constant along the characteristic lines of the system; hence

$$f^\pm(t, x, z, \omega) = \phi\left(x - h\sum_{k=0}^{[ct/h]} a(T^k\omega)\right) + O(h). \tag{6}$$

The asymptotic limit leading to a horizontal diffusion will be obtained by letting h go to zero and observing the system for large positive times. A small parameter ϵ being introduced, h is changed into ϵh (thus letting the collision frequency go to ∞) and t into t/ϵ. The problem of interest (3)–(5) becomes

$$\epsilon^2\partial_t f_\epsilon^\pm + c\epsilon a(\omega)\cdot\partial_x f_\epsilon^\pm \pm c\partial_z f_\epsilon^\pm = 0, \quad x \in \mathbb{R}^d, \quad 0 < z < h, \quad \omega \in \mathbb{T}^2. \tag{7}$$

$$f_\epsilon^+(t, x, 0, \omega) = f_\epsilon^-(t, x, 0, T\omega), \quad x \in \mathbb{R}^d, \quad \omega \in \mathbb{T}^2, \tag{8a}$$

$$f_\epsilon^-(t, x, h, \omega) = f_\epsilon^+(t, x, h, T\omega), \quad x \in \mathbb{R}^d, \quad \omega \in \mathbb{T}^2, \tag{8b}$$

$$f_\epsilon^\pm(0, x, z, \omega) = \phi(x), \quad x \in \mathbb{R}^d, \quad 0 < z < h, \quad \omega \in \mathbb{T}^2. \tag{9}$$

Its solution is given by (see (6))

$$f_\epsilon^\pm(t, x, z, \omega) = \phi\left(x - \epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right) + O(\epsilon), \tag{10}$$

therefore most of the analysis is reduced to studying the limit, as $\epsilon \to 0$, of the expression

$$\psi_\epsilon(t,x,\omega) = \phi\left(x - \epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right). \tag{11}$$

2. Notations and main results

The following notation will be systematically used below:

$$\langle F\rangle = \frac{1}{4\pi^2}\int_{\mathbb{T}^2} F(\omega)\,\mathrm{d}\omega.$$

The formula $U_T f = f \circ T$ with the mapping T given by (1) defines an operator U_T in the space $L^2(\mathbb{T}^2)$; this operator is unitary and therefore its adjoint is given by $U_T^* f = U_T^{-1} f = f \circ T^{-1}$.

Definition 1. A coboundary is an element of $\mathrm{Im}(I - U_T)$. Two functions f and g belonging to $L^2(\mathbb{T}^2)$ are said to be cohomologous if and only if $f - g$ is a coboundary and this equivalence relation will be denoted $f \sim g$.

The next proposition describes the elementary properties of what will be, in the limit $\epsilon \to 0$, the diffusion coefficient. It is essentially based on the ergodic and mixing properties of the mapping T.

Proposition 2.

(1) Any function $a \in L^2(\mathbb{T}^2)$ which satisfies the relation $a = a \circ T$ is constant and the subspace $\mathrm{Im}(I - U_T)$ is dense in the space of functions $a \in L^2(\mathbb{T}^2)$ such that $< a >= 0$ (notice that this space is invariant under T).

Let $s > 0$ and $a : \mathbb{T}^2 \to \mathbb{R}^d$ belong to the Sobolev class $H^s(\mathbb{T}^2)$, with mean value $\langle a\rangle = 0$.

(2) One has

$$D(a) = \tfrac{1}{2}\langle a^2\rangle + \sum_{k\geq 1}\langle a\circ T^k \otimes a\rangle = \frac{1}{2}\lim_{N\to\infty}\left\langle\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} a\circ T^k\right)^{\otimes 2}\right\rangle \geq 0, \tag{12}$$

where the series

$$\sum_{k\geq 1}\|\langle a\circ T^k\otimes a\rangle\| < +\infty$$

for any norm $\|\cdot\|$ on $M_d(\mathbb{R})$.

(3) Let $\xi \in \mathbb{R}^d$ and $b \in H^s(\mathbb{T}^2;\mathbb{R}^d)$ such that $\langle b\rangle = 0$. If $a\cdot\xi \sim b\cdot\xi$ then $\xi\cdot D(a)\xi = \xi\cdot D(b)\xi$.

(4) For any $\xi \in \mathbb{R}^d$, the following properties are equivalent:

(i) $D(a)\xi = 0$;

(ii) $\xi\cdot D(a)\xi = 0$;

(iii) the sequence of functions $f_N\cdot\xi$ of $L^2(\mathbb{T}^2)$

$$f_N\cdot\xi = \sum_{k=1}^{N}(a\circ T^k)\cdot\xi$$

is (uniformly with respect to N) bounded in $L^2(\mathbb{T}^2)$;

(iv) the function $a\cdot\xi$ is a coboundary.

The main result of this paper is the following.

Theorem 3. Let $a : \mathbb{T}^2 \to \mathbb{R}^d$ be in the class $C^3(\mathbb{T}^2)$ with mean value $\langle a \rangle = 0$ and let $\phi \in C_0^\infty(\mathbb{R}^d)$ be an initial data. Denote by $u(t,x)$ the solution of

$$\partial_t u = \tfrac{1}{2} hc \nabla_x \cdot (D(a) \nabla_x u), \quad u(0,x) = \phi(x). \tag{13}$$

Then the family of functions $f_\epsilon^\pm$ defined by (7)–(9) converges to $u(x,t)$ as $\epsilon \to 0$ in the following sense: for any $\tau > 0$ and any compact $K \subset \mathbb{R}_x^d$

$$\langle f_\epsilon^\pm(t,x,z,\omega) \rangle \to u(t,x), \quad C^0([0,\tau] \times K \times \mathbb{T}^2 \times]0,h[), \tag{14a}$$

$$f_\epsilon^\pm(t,x,z,\omega) \to u(t,x), \quad C^0([0,\tau], w^* - L^\infty(\mathbb{R}^d \times \mathbb{T}^2)). \tag{14b}$$

Furthermore with $\psi_\epsilon(x,t)$ defined by formula (11) one has

$$\| f_\epsilon^\pm - \psi_\epsilon \|_{L^\infty(\mathbb{R} \times \mathbb{R}^d \times (0,h) \times \mathbb{T}^2)} = O(\epsilon). \tag{15}$$

The proof of this theorem is tailored on the proof of the Ito formula (cf. for instance [12]); in particular it will be shown that the average of the different products appearing in a Taylor expansion are, in the limit, completely decorrelated and therefore converge to the product of the corresponding limiting averages. In the original Ito formula, this point is straightforward since (by construction) the Brownian motion has independent increments. At variance in the present paper, the independence is obtained only in the limit $\epsilon \to 0$ and, as will be shown below, is a consequence of the different mixing properties of the map T.

An analogous result has been proven by Denker and Philipp [14] for suspensions of finite type subshifts under Hölder continuous maps. It would be theoretically possible to reduce the present analysis to this situation by coding the mapping T with a Markov partition. However, we recall that our goal in the present paper is to give an example of diffusion approximation of a reversible system which is as explicit as possible and uses only elementary techniques. In the present case, Fourier series expansions are used instead of a Markov partition as a coding of the system. [1]

The outline of the paper is as follows. In Section 3, the basic mixing properties of the map T are given (property (H1) of Proposition 5); they are based on the fact that the eigenvalues of the matrix defining the cat map are algebraic of degree 2 and therefore satisfy a diophantine condition. The mixing property (H1) is used to prove Proposition 2. Then the decorrelation properties of the process (property (H2)) are first proven in Proposition 6. This implies property (H3) of Corollary 7, used to estimate the remainder of the Taylor expansion formula. Then Proposition 8 and Corollary 9 provide all the material for the proof of Theorem 3 which belongs to Section 4. Section 5 contains final remarks and comments on the numerical simulations.

3. Proofs of the mixing properties and of Proposition 2

As announced above, the proofs are based on Fourier series techniques. Indeed the properties of the dynamical system induced by the map T on the torus $\mathbb{T}^2$ are most conveniently translated into those of the dynamical system defined in $\mathbb{Z}^2$ (the dual group of $\mathbb{T}^2$) by the iteration of the matrix

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

[1] While going through the revision of this manuscript, we became aware of an independent contribution by Babovski [22] on the same model. He uses the coding based on a Markov partition instead of the explicit method of the present article, and hence his analysis is a particular case of [14]. However, he provides a phenomenological derivation of models similar to (3)–(5) as a linearization of the law of specular reflection at a rough surface. This might add some practical interest for this class of examples.

The matrix M is strictly hyperbolic, i.e. it has two (distinct) real eigenvalues given by

$$\lambda_+ = 1+\theta, \quad \lambda_- = \lambda_+^{-1}, \qquad \text{with } \theta = \tfrac{1}{2}(1+\sqrt{5}) \tag{16}$$

and the corresponding eigenvectors (generating the unstable and stable manifolds) are

$$e_+ = \mu\begin{pmatrix}\theta\\ 1\end{pmatrix}, \quad e_- = \mu\begin{pmatrix}1\\ -\theta\end{pmatrix} \tag{17}$$

with $\mu = (1+\theta^2)^{-1/2}$. The vectors (e_+, e_-) define an orthonormal basis. In this basis M reduces to the diagonal form

$$M \sim \begin{pmatrix}\lambda_+ & 0\\ 0 & \lambda_-\end{pmatrix}.$$

Most of the results that we shall obtain are based on the fact that the eigenvalues of M are algebraic numbers and therefore satisfy a diophantine estimate. We recall this simple fact (originally due to Kronecker) in the following lemma.

Lemma 4. Let $\theta = \frac{1}{2}(1+\sqrt{5})$. Then:
(i) *best rational approximation for* θ

$$\inf_{(p,q)=1} |\theta - p/q| \geq \frac{1}{(1+\sqrt{5})q^2}, \tag{18}$$

(ii) *diophantine estimate for* θ

$$\inf_{(p,q)\in\mathbb{Z}^2\setminus\{0\}} |q\theta - p| \geq \frac{1}{(1+\sqrt{5})\sup(|p|,|q|)}. \tag{19}$$

Proof. The minimal polynomial of θ over $\mathbb{Q}$ is $P(X) = X^2 - X - 1 = (X-\theta)(X+\theta^{-1})$. Let p and q be two integers $(p,q) = 1$. First $P(p/q) \neq 0$ since P has no rational root, and therefore

$$|P(p/q)| = \frac{|p^2 - qp - q^2|}{q^2} \geq \frac{1}{q^2}. \tag{20}$$

On the other hand

$$|p/q - \theta| = \frac{|P(p/q)|}{|p/q - \theta + \theta + \theta^{-1}|} = \frac{|P(p/q)|}{|p/q - \theta + \sqrt{5}|} \geq \frac{1}{q^2(|p/q-\theta| + \sqrt{5})}.$$

Considering both cases $|p/q - \theta| > 1$ and $|p/q - \theta| < 1$, one obtains for any pair of integers (p, q) with $q \neq 0$

$$|p/q - \theta| \geq \inf\left(1, \frac{1}{(1+\sqrt{5})q^2}\right), \tag{21}$$

which proves (i).
(ii) follows directly from (i). □

Proposition 5. Let $0 \leq \chi(R)$ be a nonincreasing positive function such that

$$\lim_{R\to\infty} \chi(R) = 0.$$

Consider the class of functions

$$H_\chi = \left\{ f \in L^2(\mathbb{T}^2) \text{ s.t. } \sum_{|k_1|,|k_2|>R} |\hat{f}(k)|^2 \le \chi(R)^2 \|f\|_2^2 \right\}.$$

Then one has the following:

(i) *Rate of Mixing.* For all $(f, g) \in H_\chi$ such that $\langle f \rangle = \langle g \rangle = 0$,

$$|\langle f \circ T^n \cdot g \rangle| \le \tfrac{1}{2\pi^2} \|f\|_2 \|g\|_2 \chi \left(C_0^{-1/2} \left(\tfrac{3+\sqrt{5}}{2} \right)^{n/2} \right) \quad \text{with } C_0 = \frac{1+\sqrt{5}}{\sqrt{2}}. \tag{22}$$

(ii) *Property* (H1): *Exponential mixing.* For all $s > 0$ and all $f \in H^s(\mathbb{T}^2)$ such that $\langle f \rangle = 0$, the self-correlation coefficient defined as

$$C_f(n) = \langle f \circ T^n \cdot f \rangle$$

satifies the decay estimate:

$$|C_f(n)| \le C \|f\|_s^2 \mathrm{e}^{-sn\alpha} \quad \text{with } \alpha = \log\left(\tfrac{1}{2}(3+\sqrt{5})\right). \tag{23}$$

Proof. With the Plancherel formula one has, for any pair $(f, g) \in L^2(\mathbb{T}^2)$ with mean value $\langle f \rangle = \langle g \rangle = 0$, the formula:

$$\langle f \circ T^n \cdot g \rangle = \frac{1}{4\pi^2} \sum_{k \ne 0} \hat{f}(M^{-n}k)\hat{g}(-k). \tag{24}$$

For any $R > 0$, decompose the above sum into two parts corresponding to K_R and K_R^c with K_R given by

$$K_R = \{k \in \mathbb{Z}^2 \text{ s.t. } \sup(|k_1|, |k_2|) \le R\}.$$

Since g belongs to the class H_χ, the Cauchy–Schwarz inequality yields the estimate:

$$\left| \sum_{k \in K_R^c} \hat{f}(M^{-n}k)\hat{g}(-k) \right| \le \|f\|_2 \|g\|_2 \chi(R). \tag{25}$$

For $k \in K_R$, one introduces the decomposition $k = (k \cdot e_+)e_+ + (k \cdot e_-)e_-$; Kronecker's estimate (19) shows that

$$|k \cdot e_-| \ge \theta^{-1}|k|^{-1} \ge (\sqrt{2}R\theta)^{-1},$$

whence

$$|M^{-n}k| \ge \frac{\lambda_+^n}{\sqrt{2}R\theta}.$$

Using that $f \in H_\chi$ and that χ is nonincreasing, this implies the estimate:

$$\left| \sum_{k \in K_R - \{0\}} \hat{f}(M^{-n}k)\hat{g}(-k) \right| \le \|f\|_2 \|g\|_2 \chi \left(\frac{\lambda_+^n}{\sqrt{2}R\theta} \right). \tag{26}$$

Relation (22) is obtained by choosing $R = C_0^{-1/2}\lambda_+^{n/2}$ in (25) and (26). To obtain the exponential rate of mixing (23), one specializes (22) to the case $f = g$ and uses the expression $\chi(R) = R^{-s}(\|f\|_s \|f\|_2^{-1})^{1/2}$. □

Proof of Proposition 2. Let $k \in \mathbb{Z}^2 \setminus 0$. Since $\ker(M - \lambda_- I)$ is a line with irrational slope, the orthogonal projection of k on $\ker(M - \lambda_- I)$ is not 0 and therefore $|M^{-n}k| \to +\infty$ as $n \to +\infty$.

Let $a \in L^2(\mathbb{T}^2)$ and consider its Fourier series:

$$a(\omega) = \sum_{k \in \mathbb{Z}^2} \hat{a}(k) \mathrm{e}^{\mathrm{i}k\omega}.$$

The relation $a = a \circ T$ shows that for any $k \in \mathbb{Z}^2$ and any n one has

$$\hat{a}(k) = \hat{a}(M^{-n}k).$$

But $a \in l^2(\mathbb{Z}^2)$ since $a \in L^2(\mathbb{T}^2)$: hence for all $k \neq 0$ in $\mathbb{Z}^2$ $\hat{a}(k) = \hat{a}(M^{-n}k) \to 0$ as $n \to +\infty$. Therefore $\hat{a} \equiv 0$ on $\mathbb{Z}^2 \setminus \{0\}$, which means that a is a constant. This proves the ergodic property $\ker(I - U_T) = \mathbb{R}$. The same remark applies to the kernel of the adjoint $(I - U_T)^* = U_T^{-1}(U_T - I)$ in $L^2(\mathbb{T}^2)$ which also is reduced to the constants. Therefore the set of coboundaries is dense in the space of functions $a \in L^2(\mathbb{T}^2)$ such that $\langle a \rangle = 0$.

For condition (2), start with the formula

$$\langle a \circ T^k \otimes a \circ T^l \rangle = \langle a \circ T^{k-l} \otimes a \rangle, \tag{27}$$

which follows from the invariance of the measure $\mathrm{d}\omega_1\, \mathrm{d}\omega_2$ under T. Summing with respect to $m = k - l$ yields

$$\left\langle \left(\sum_{k=0}^{N-1} a \circ T^k \right)^{\otimes 2} \right\rangle = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \langle a \circ T^{k-l} \otimes a \rangle = N \langle a^{\otimes 2} \rangle + 2 \sum_{m=1}^{N-1} (N - m) \langle a \circ T^m \otimes a \rangle \tag{28}$$

or in other words

$$\langle a^{\otimes 2} \rangle + 2 \sum_{m=1}^{N-1} \langle a \circ T^m \otimes a \rangle = \left\langle \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a \circ T^k \right)^{\otimes 2} \right\rangle + \frac{2}{N} \sum_{m=1}^{N-1} m \langle a \circ T^m \otimes a \rangle. \tag{29}$$

Since the vector field a has mean zero and is of Sobolev class H^s with $s > 0$, it follows from Proposition 5 that the exponential estimate (23) holds; it implies the absolute convergence of the series

$$\sum_{m \geq 1} \langle a \circ T^m \otimes a \rangle,$$

which appears in the left-hand side of (12) (Proposition 2 (2)) and of

$$\sum_{m \geq 1} m \langle a \circ T^m \otimes a \rangle .$$

Therefore, the last term of the right-hand side of (29) goes to zero, leading to the relation

$$D(a) = \frac{1}{2} \lim_{N \to \infty} \left\langle \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a \circ T^k \right)^{\otimes 2} \right\rangle \tag{30}$$

and completing the proof of (2).

To prove condition (3), write $a \sim b$ as $a - b = \phi - \phi \circ T$ and use the relation:

$$\xi \cdot D(b)\xi = \lim_{N \to \infty} \left\| \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} (\xi \cdot b) \circ T^k \right\|_2^2$$

$$= \lim_{N\to\infty} \left\| \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} (\xi \cdot a) \circ T^k + (\xi \cdot \phi) \circ T^k - (\xi \cdot \phi) \circ T^{k+1} \right\|^2$$

$$= \lim_{N\to\infty} \left\| \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} (a \circ T^k \cdot \xi) \right\|_2^2 = \xi \cdot D(a)\xi. \tag{31}$$

Since the matrix $D(a)$ is symmetric and nonnegative, conditions (4) (i) and (ii) are equivalent. To prove that (4) (ii) implies 4 (iii), write (29) in the form

$$\left\| \sum_{k=0}^{N-1} (\xi \cdot a) \circ T^k \right\|_2^2 = N\xi \cdot D(a)\xi - 2N \sum_{m\geq N} \langle (\xi \cdot a) \circ T^m (\xi \cdot a) \rangle$$
$$- 2 \sum_{m=1}^{N-1} m \langle (\xi \cdot a) \circ T^m (\xi \cdot a) \rangle. \tag{32}$$

We recall that, thanks to the exponential decay estimate (23), the series

$$\sum_{m\geq 1} |m \langle (\xi \cdot a) \circ T^m (\xi \cdot a) \rangle| < +\infty \quad \text{and} \quad \sum_{m\geq N} |\langle (\xi \cdot a) \circ T^m (\xi \cdot a) \rangle| = \mathrm{o}\left(\frac{1}{N}\right). \tag{33}$$

Hence, if one assumes that $\xi \cdot D(a)\xi = 0$, one has

$$\left\| \sum_{k=0}^{N-1} (\xi \cdot a) \circ T^k \right\|_2^2 \leq 2 \sup_{N>0} \left(N \sum_{m\geq N} |\langle (\xi \cdot a) \circ T^m (\xi \cdot a) \rangle| \right)$$
$$+ 2 \sum_{m\geq 1} |m \langle (\xi \cdot a) \circ T^m (\xi \cdot a) \rangle| < +\infty. \tag{34}$$

Finally, to prove that (iii) implies (iv) observe that the sequence $f_N \cdot \xi$ is, in any case, the "formal inverse" for the equation

$$a \cdot \xi = g^\xi - g^\xi \circ T. \tag{35}$$

Since the sequence

$$f_N \cdot \xi = \sum_{k=1}^{N} a \circ T^k \cdot \xi$$

is bounded in $L^2(\mathbb{T}^2)$, one can consider f, one of its weak $L^2(\mathbb{T}^2)$ limit points and, using Lemma 4, one has, for any function $\phi \in C^\infty(\mathbb{T}^2)$,

$$\langle (f_N - f_N \circ T)\phi \rangle - \langle \xi \cdot a\phi \rangle = -\langle \xi \cdot a \circ T^{N+1} \phi \rangle \to 0, \quad \text{for } N \to \infty, \tag{36}$$

which shows that $f - f \circ T = a \cdot \xi$. □

As we said above, the proof of Theorem 2 follows in many respects the proof of the Ito formula. Therefore it will be important to study the decorrelation of events occurring in two separate intervals of time, uniformly with respect to the size of these intervals, and under the only assumption that their mutual distance is large enough. This is dealt with in the following proposition. What we prove is a property similar to the "Weak Bernouilli property" as introduced by Ornstein and Weiss (cf. [4]). The main difference with the classical formulation as in [4] is the use of trigonometric polynomials instead of indicator functions of the elements of a partition.

Proposition 6. The transformation T has the following property:

Property (H2). There exist two constants $\beta_0 > 0$ and β_1 such that for all $l, m \in \mathbb{N}$, $U \subset \{n, \dots, n+l\}$, $V \subset \{n, \dots, n+m\}$ and for all pair of trigonometric polynomials P, Q, of degree less than R, one has, for all $n \geq \beta_0 \log R + \beta_1$,

$$\left\langle \prod_{k\in U} P \circ T^{-k} \prod_{k \in V} Q \circ T^k \right\rangle - \left\langle \prod_{k\in U} P \circ T^{-k} \right\rangle \left\langle \prod_{k\in V} Q \circ T^k \right\rangle = 0. \tag{37}$$

Proof. This proof follows that of Proposition 5. The method is similar to the one used by Katznelson [18], but here a more precise result is needed and proven. Observe that it is enough to study expressions of the following type:

$$\begin{aligned} &\frac{1}{4\pi^2} \int_{\mathbb{T}^2} \exp\left(\mathrm{i}\omega \cdot \sum_{k\in U} M^{-k}\xi_k\right) \exp\left(\mathrm{i}\omega \cdot \sum_{k\in V} M^k \eta_k\right) \mathrm{d}\omega \\ &- \left(\frac{1}{4\pi^2} \int_{\mathbb{T}^2} \exp\left(\mathrm{i}\omega \cdot \sum_{k\in U} M^{-k}\xi_k\right) \mathrm{d}\omega\right) \left(\frac{1}{4\pi^2} \int_{\mathbb{T}^2} \exp\left(\mathrm{i}\omega \cdot \sum_{k\in V} M^k \eta_k\right) \mathrm{d}\omega\right) \end{aligned} \tag{38}$$

with $\xi_k \in K_R$, for all k such that $n \leq k \leq n+l$ and $\eta_k \in K_R$, for all k such that $n \leq k \leq n+m$ (where, as in the proof of Proposition 5, $K_R = \{k \in \mathbb{Z}^2 \text{ s.t. } \sup(|k_1|, |k_2|) \leq R\}$).

With the notations

$$X_U^- = \sum_{k\in U} M^{-k}\xi_k, \qquad X_V^+ = \sum_{k\in V} M^k \eta_k,$$

relation (37) is equivalent to the following assertion:

There exist $\beta_0 > 0$ and β_1 such that

$$X_U^- + X_V^+ = 0 \ \Rightarrow \ X_U^- = 0 \text{ and } X_V^+ = 0 \quad \forall n > \beta_0 \log R + \beta_1, \ \forall l, m \in \mathbb{N}. \tag{39}$$

Denote by $I = \mathbb{R}e_+$ and $S = \mathbb{R}e_-$ the unstable and stable manifolds of M acting on $\mathbb{R}^2$. Let $\xi \in K_R$; one has

$$|X_U^- \cdot e_+| \leq \sum_{k\in U} |M^{-n}\xi_k \cdot e_+| = \sum_{k\in U} |\xi_k \cdot M^{-n} e_+| = \sum_{k\in U} \lambda_-^k |\xi_k \cdot e_+|.$$

Proceeding likewise with X_V^+, for all ξ and η in K_R one has

$$|X_U^- \cdot e_+| \leq \sqrt{2}R \frac{\lambda_-^n}{1-\lambda_-}, \qquad |X_V^+ \cdot e_-| \leq \sqrt{2}R \frac{\lambda_-^n}{1-\lambda_-}. \tag{40}$$

This implies that X_U^- belongs to a neighborhood S_R^n of S while X_V^+ belongs to a neighborhood I_R^n of I given by the formulas:

$$\begin{aligned} I_R^n &= \left\{X \in \mathbb{R}^2 \text{ s.t. } |X \cdot e_-| \leq \sqrt{2}R \frac{\lambda_-^n}{1-\lambda_-}\right\}, \\ S_R^n &= \left\{X \in \mathbb{R}^2 \text{ s.t. } |X \cdot e_+| \leq \sqrt{2}R \frac{\lambda_-^n}{1-\lambda_-}\right\}, \end{aligned} \tag{41}$$

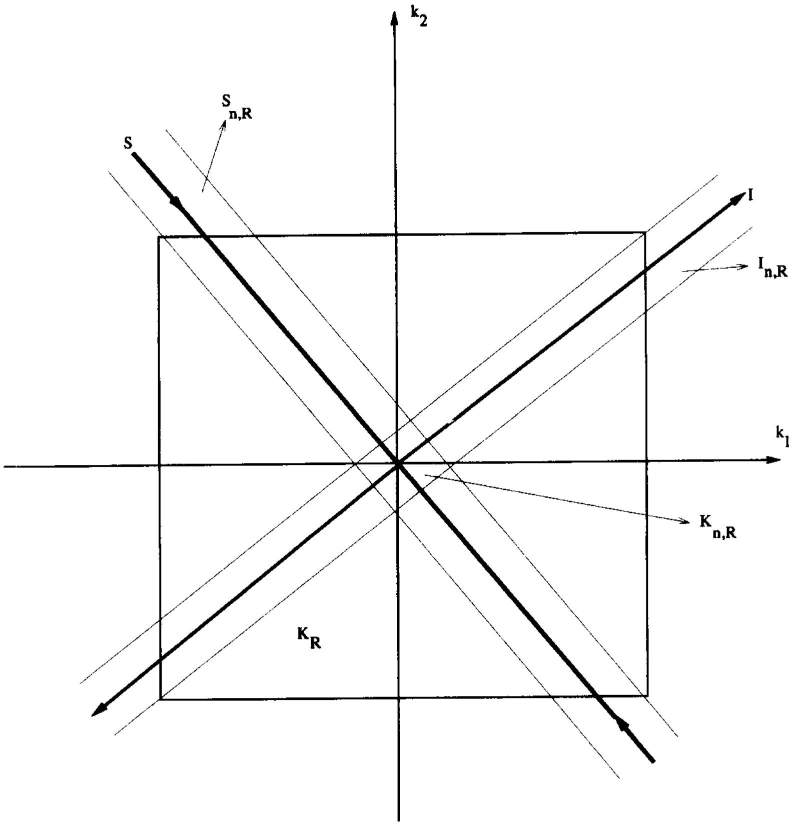

Fig. 2.

Since $X_U^- + X_V^+ = 0$, both X_U^- and X_V^+ belong to $K_{n,R} = S_R^n \cap I_R^n$ (cf. Fig. 2) which for an n greater than a given value N_0 is contained in K_R. Whenever they are not both equal to zero, the diophantine estimate (19) of Lemma 4 can be used to give.

$$|X_U^- \cdot e_+| \geq \frac{1}{(1+\sqrt{5})R} \quad \text{or} \quad |X_V^+ \cdot e_-| \geq \frac{1}{(1+\sqrt{5})R}. \tag{42}$$

For $n > (2\log R + \log(\sqrt{2}(1+\sqrt{5}))/\log\lambda_+$, one has

$$\frac{1}{(1+\sqrt{5})R} > \sqrt{2}R\frac{\lambda_-^n}{1-\lambda_-}. \tag{43}$$

Therefore, if n is greater than

$$\sup\left(n_0, \frac{2\log R + \log(\sqrt{2}(1+\sqrt{5}))}{\log\lambda_+}\right),$$

X_U^- and X_V^+ are both equal to zero. □

Corollary 7. Let $0 \leq \chi(R)$ be a nonincreasing positive function going to 0 as R tends to infinity, such that

$$\chi(R) = \mathrm{O}((1/\log R)^6) \quad \text{for} \quad R \to +\infty. \tag{44}$$

Consider the class of functions defined by

$$W_\chi = H_\chi \cap \left\{ f \in L^\infty(\mathbb{T}^2) \text{ s.t. } \frac{1}{4\pi^2} \sum_{\sup(|k_1|,|k_2|)>R} |\hat{f}(k)| \le \|f\|_\infty \chi(R)\ \forall R > 0 \right\}. \tag{45}$$

Then one has:

Property (H3). For all $f \in W_\chi$ such that $\langle f \rangle = 0$

$$\frac{1}{\sqrt{N}} \sum_{k=0}^{N} f \circ T^k$$

is uniformly bounded (with respect to N) in $L^4(\mathbb{T}^2)$.

Proof. The proof follows the same line as the one indicated by Ratner [21] when the mapping T is replaced by any K system. For the sake of being complete, we give a proof based only on property (H2) (which is more restrictive than assuming the K property). First observe that

$$\begin{aligned}\left\langle \left| \sum_{k=0}^{N} f \circ T^k \right|^4 \right\rangle &= \sum_{0\le k_1,k_2,k_3,k_4 \le N} \langle f \circ T^{k_1} f \circ T^{k_2} f \circ T^{k_3} f \circ T^{k_4} \rangle \\ &= 4! \sum_{0\le k_1 \le k_2 \le k_3 \le k_4 \le N} \langle f \circ T^{k_1} f \circ T^{k_2} f \circ T^{k_3} f \circ T^{k_4} \rangle,\end{aligned} \tag{46}$$

and introduce the following sets of indices:

$$\begin{aligned}A &= \left\{ (k_1,k_2,k_3,k_4) \text{ s.t. } 0 \le k_1 \le k_2 \le k_3 \le k_4 \le N,\quad \sup_{2\le i\le 4} |k_i - k_{i-1}| \le N^{1/3} \right\},\\ B &= \{(k_1,k_2,k_3,k_4) \text{ s.t. } 0 \le k_1 \le k_2 \le k_3 \le k_4 \le N,\ k_2 - k_1 > N^{1/3}\},\\ C &= \{(k_1,k_2,k_3,k_4) \text{ s.t. } 0 \le k_1 \le k_2 \le k_3 \le k_4 \le N,\ k_3 - k_2 > N^{1/3}\},\\ D &= \{(k_1,k_2,k_3,k_4) \text{ s.t. } 0 \le k_1 \le k_2 \le k_3 \le k_4 \le N,\ k_4 - k_3 > N^{1/3}\}.\end{aligned} \tag{47}$$

One has

$$A \cap B = A \cap C = A \cap D = \emptyset \quad \text{and} \quad 0, 1, \ldots, N^4 = A \cup B \cup C \cup D \tag{48}$$

with the subsets B, C and D having nonempty intersection. From relation (48) one deduce the estimate:

$$\begin{aligned}\left\langle \left| \sum_{k=0}^{N} f \circ T^k \right|^4 \right\rangle \le\ & 4! \sum_{(k_1,k_2,k_3,k_4)\in A} |\langle f \circ T^{k_1} f \circ T^{k_2} f \circ T^{k_3} f \circ T^{k_4} \rangle| \\ &+ 4! \sum_{(k_1,k_2,k_3,k_4)\in B} |\langle f \circ T^{k_1} f \circ T^{k_2} f \circ T^{k_3} f \circ T^{k_4} \rangle| \\ &+ 4! \sum_{(k_1,k_2,k_3,k_4)\in C} |\langle f \circ T^{k_1} f \circ T^{k_2} f \circ T^{k_3} f \circ T^{k_4} \rangle| \\ &+ 4! \sum_{(k_1,k_2,k_3,k_4)\in D} |\langle f \circ T^{k_1} f \circ T^{k_2} f \circ T^{k_3} f \circ T^{k_4} \rangle|.\end{aligned} \tag{49}$$

To prove the corollary we show that all the terms which appears in (49) are uniformly bounded with respect to N^2. This will be done in three steps: step 1: terms of (49) with support in A; step 2: terms with support in B or D (the proof being similar in these two cases); and step 3: terms with support in C. For steps 2 and 3 the Fourier expansion of f truncated at degree R will be used; it is denoted by $P_R(f)$ and since $f \in W_\chi$ one has

$$|f(\omega) - P_R(f)(\omega)| \le \chi(R)\|f\|_\infty. \tag{50}$$

Step 1: Summation with support in A. Observe that $\sharp A \le 3!N(N^{1/3}+1)^3$; this implies the estimate:

$$\sum_{(k_1,k_2,k_3,k_4)\in A} |\langle f\circ T^{k_1}\cdot f\circ T^{k_2}\cdot f\circ T^{k_3}\cdot f\circ T^{k_4}\rangle| \le 3!N(N^{1/3}+1)^3\|f\|_\infty^4. \tag{51}$$

Step 2: Summation with support in B and D. As already noticed, the proof is similar for these two terms and therefore only the sum with support in B is considered.

Since f belongs to W_χ, estimate (50) shows that

$$\begin{aligned}\sum_{(k_1,k_2,k_3,k_4)\in B} &|\langle f\circ T^{k_1}\cdot f\circ T^{k_2}\cdot f\circ T^{k_3}\cdot f\circ T^{k_4}\rangle| \\ &\le \sum_{(k_1,k_2,k_3,k_4)\in B} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\cdot P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle| \\ &\quad + 4N^4\|f\|_\infty^4\chi(R)(\sup(\chi(R),1))^3.\end{aligned} \tag{52}$$

To use property (H2), write

$$\begin{aligned}&\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\cdot P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle \\ &\quad = \langle P_R(f)\circ T^{k_1-\kappa}\cdot P_R(f)\circ T^{k_2-\kappa}\cdot P_R(f)\circ T^{k_3-\kappa}\cdot P_R(f)\circ T^{k_4-\kappa}\rangle.\end{aligned} \tag{53}$$

In (53) choose $\kappa = [\frac{1}{2}(k_1+k_2)]$, notice that $P_R(f)$ is of mean value zero (because f is of mean value zero) and apply Proposition 6 to the sets

$$U = \{\kappa - k_1\}, \qquad V = \{k_2-\kappa, k_3-\kappa, k_4-\kappa\}.$$

It shows that, for

$$R = \exp\left(\frac{N^{1/3}-2\beta_1}{2\beta_0}\right)$$

and for any $(k_1,k_2,k_3,k_4)\in B$, one has

$$\langle P_R(f)\circ T^{k_1}P_R(f)\circ T^{k_2}P_R(f)\circ T^{k_3}P_R(f)\circ T^{k_4}\rangle = 0. \tag{54}$$

Therefore with $\chi(R) = \mathrm{O}((1/\log R)^6)$ for $R\to+\infty$, the estimate

$$\sum_{(k_1,k_2,k_3,k_4)\in B} |\langle f\circ T^{k_1}\cdot f\circ T^{k_2}\cdot f\circ T^{k_3}\cdot f\circ T^{k_4}\rangle| \le 4N^4\|f\|_\infty^4\chi\left(\exp\left(\frac{N^{1/3}-2\beta_1}{2\beta_0}\right)\right) = \mathrm{O}(N^2) \tag{55}$$

is deduced from (52).

Step 3: Summation with support in C. As in step 2 write

$$\sum_{(k_1,k_2,k_3,k_4)\in C} |\langle f\circ T^{k_1}\cdot f\circ T^{k_2}\cdot f\circ T^{k_3}\cdot f\circ T^{k_4}\rangle|$$

$$\leq \sum_{(k_1,k_2,k_3,k_4)\in C} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\cdot P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle|$$

$$+\,4N^4\|f\|_\infty^4\chi(R)(\sup(\chi(R),1))^3. \tag{56}$$

Then for any $(k_1,k_2,k_3,k_4)\in C$, apply Proposition 6 to the sets

$$U=\{\kappa-k_1,\kappa-k_2\}\quad\text{and}\quad V=\{k_3-\kappa,k_4-\kappa\}$$

with $\kappa=[\frac{1}{2}(k_2+k_3)]$. It shows that, for

$$R=\exp\left(\frac{N^{1/3}-2\beta_1}{2\beta_0}\right),$$

one has

$$\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\cdot P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle$$

$$=\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle\langle P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle. \tag{57}$$

Therefore

$$\sum_{(k_1,k_2,k_3,k_4)\in C} |\langle f\circ T^{k_1}\cdot f\circ T^{k_2}\cdot f\circ T^{k_3}\cdot f\circ T^{k_4}\rangle|$$

$$\leq \sum_{(k_1,k_2,k_3,k_4)\in C} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle\langle P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle|$$

$$+\,4N^4\|f\|_\infty^4\chi(R). \tag{58}$$

On the other hand

$$\sum_{(k_1,k_2,k_3,k_4)\in C} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle\langle P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle|$$

$$\leq \sum_{0\leq k_1,k_2,k_3,k_4\leq N} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle||\langle P_R(f)\circ T^{k_3}\cdot P_R(f)\circ T^{k_4}\rangle|$$

$$=\left(\sum_{0\leq k_1,k_2\leq N} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle|\right)^2. \tag{59}$$

Since f belongs to the class $W_\chi\subset H_\chi$, $P_R(f)$ also belongs to the class H_χ and the estimate (22) of Proposition 5 can be used to give

$$|\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle|\leq \tfrac{1}{2\pi^2}\|f\|_2^2\chi(C_0\lambda_+^{(k_2-k_1)/2}). \tag{60}$$

The sum $\sum_{n\geq 0}\chi(C_0\lambda_+^{n/2})$ converges (by assumption: see (44)) and therefore

$$\left(\sum_{0\leq k_1,k_2\leq N} |\langle P_R(f)\circ T^{k_1}\cdot P_R(f)\circ T^{k_2}\rangle|\right)^2 \leq \left(\tfrac{1}{2\pi^2}\|f\|_2^2\right)^2 \sum_{0\leq k_1,k_2\leq N}\chi(C_0\lambda_+^{(k_2-k_1)/2}) = \mathrm{O}(N^2). \tag{61}$$

With (56)–(58) and (61) one obtains, for

$$R = \exp\left(\frac{N^{1/3}-2\beta_1}{2\beta_0}\right),$$

the formula

$$\sum_{(k_1,k_2,k_3,k_4)\in C} |\langle f\circ T^{k_1}\cdot f\circ T^{k_2}\cdot f\circ T^{k_3}\cdot f\circ T^{k_4}\rangle|$$

$$\leq 4N^4\|f\|_\infty^4\chi\left(\exp\left(\frac{N^{1/3}-2\beta_1}{2\beta_0}\right)\right) + \mathrm{O}(N^2) \leq CN^2,$$

which concludes the proof of step 3 and of Corollary 7. □

To further extend the decorrelation properties, this section is concluded with Proposition 8 involving functions $f\in H^s(\mathbb{R}^d)$ and Corollary 9 which allows to consider smooth functions with subquadratic growth at infinity.

Proposition 8. Assume that the vector field a is smooth enough (say $a\in C^3(\mathbb{T}^2,\mathbb{R}^2)$) and satisfies $<a>=0$. Then for any pair of functions f and $g\in H^s(\mathbb{R}^d)$ with $\frac{1}{2}d<s$ one has

$$\left|\left\langle f\left(\epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\right) g\left(\epsilon h\sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\right)\right\rangle\right.$$

$$\left. - \left\langle f\left(\epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\right)\right\rangle\left\langle g\left(\epsilon h\sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\right)\right\rangle\right|$$

$$\leq \|f\|_s\|g\|_s\left(C\frac{\epsilon^{s-d/2}}{s-d/2} + \epsilon^{-d}H(K,\delta,\epsilon)\right) \tag{62}$$

with $H(K,\delta,\epsilon)$ given by

$$H(K,\delta,\epsilon) = \frac{C''}{\epsilon^2}\exp\left(-\frac{C'\delta}{\epsilon^2}\right)\exp\left(\frac{C''}{\epsilon^2}\exp\left(-\frac{C'\delta}{\epsilon^2}\right)\right). \tag{63}$$

In (62) and (63) C, C' and C'' denote some constants independent of ϵ, δ and t, $t+\tau$ in the bounded interval $[0, K]$.

Proof. First represent f and g in terms of their Fourier transforms:

$$f(x) = \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\xi\cdot x}\hat{f}(\xi)\frac{\mathrm{d}\xi}{(2\pi)^d}, \qquad g(x) = \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\eta\cdot x}\hat{g}(\eta)\frac{\mathrm{d}\eta}{(2\pi)^d} \tag{64}$$

and observe that since f and g belong to $H^s(\mathbb{R}^d)$ the above integrals can, in the sequel, be replaced by

$$f_\epsilon(x) = \int_{\mathbb{R}^d\cap\{|\epsilon\xi|\leq 1\}} \mathrm{e}^{\mathrm{i}\xi\cdot x}\hat{f}(\xi)\,\mathrm{d}\xi/(2\pi)^d, \qquad g_\epsilon(x) = \int_{\mathbb{R}^d\cap\{|\epsilon\xi|\leq 1\}} \mathrm{e}^{\mathrm{i}\eta\cdot x}\hat{g}(\eta)\,\mathrm{d}\eta/(2\pi)^d. \tag{65}$$

This truncation introduces in the proof an error of the order of $\|f\|_s\|g\|_s(\epsilon^{s-d/2})/(s-d/2)$ which correspond to the first term in the right-hand side of (62) and reduces the proof of the Proposition 8 to estimating the expression:

$$\epsilon^{-d}\sup_{|\xi|\le 1,|\eta|\le 1}\left|\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}\exp(\mathrm{i}\xi\cdot\epsilon ha\circ T^k)\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}\exp(\mathrm{i}\eta\cdot\epsilon ha\circ T^l)\right\rangle - \left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}\exp(\mathrm{i}\xi\cdot\epsilon ha\circ T^k)\right\rangle\left\langle\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}\exp(\mathrm{i}\eta\cdot\epsilon ha\circ T^l)\right\rangle\right|. \tag{66}$$

However Proposition 6 (property (H2) or relation (37)) cannot be directly applied to the above formula because the functions $\exp(\mathrm{i}\xi\cdot\epsilon ha)$ and $\exp(\mathrm{i}\eta\cdot\epsilon ha)$ are not trigonometric polynomials in the variable ω. Therefore these functions have to be approximated by their truncated Fourier series. In order to do so it is convenient to introduce the following notations:

$$\begin{aligned}\theta_\xi(\omega) &= \exp(i\xi\cdot ha(\omega)),\\ P_R(\theta_\xi) &= \sum_{\sup(|k_1|,|k_2|<R)}\hat\theta_\xi(k)\exp(\mathrm{i}k\cdot\omega),\\ A_{\xi,R} &= \sum_{\sup(|k_1|,|k_2|\ge R)}|\hat\theta_\xi(k)|.\end{aligned} \tag{67}$$

Assuming that $a\in C^3(\mathbb{T}^2,\mathbb{R}^2)$, the quantities $A_{\xi,R}$ are finite and one has

$$\begin{aligned}\|P_R(\theta_\xi)-\theta_\xi\|_\infty &\le \sum_{\sup(|k_1|,|k_2|\ge R}|\hat\theta_\xi(k)| = A_{\xi,R},\\ \||P_R(\theta_\xi)|-1\|_\infty &\le \|P_R(\theta_\xi)-\theta_\xi\|_\infty\le A_{\xi,R},\\ \|P_R(\theta_\xi)\|_\infty &\le (1+A_{\xi,R}),\end{aligned} \tag{68}$$

and

$$\forall\sigma,\ 0<\sigma<1,\quad A_R=\sup_{|\xi|\le 1}A_{\xi,R}\le\frac{C}{(1-\sigma)R^{1-\sigma}}. \tag{69}$$

Going back to (66), one has the inequality

$$\begin{aligned}&\left|\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}\theta_{\epsilon\xi}\circ T^k\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}\theta_{\epsilon\eta}\circ T^k\right\rangle-\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}\theta_{\epsilon\xi}\circ T^k\right\rangle\left\langle\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}\theta_{\epsilon\eta}\circ T^k\right\rangle\right|\\
&\le\left|\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}P_R(\theta_{\epsilon\xi})\circ T^k\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}P_R(\theta_{\epsilon\eta})\circ T^k\right\rangle\right.\\
&\quad\left.-\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}P_R(\theta_{\epsilon\xi})\circ T^k\right\rangle\left\langle\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}P_R(\theta_{\epsilon\eta})\circ T^k\right\rangle\right|\\
&\quad+\left|\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}\theta_{\epsilon\xi}\circ T^k\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}\theta_{\epsilon\eta}\circ T^k\right\rangle\right.\\
&\quad\left.-\left\langle\prod_{k=0}^{[ct/\epsilon^2 h]}P_R(\theta_{\epsilon\xi})\circ T^k\prod_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]}P_R(\theta_{\epsilon\eta})\circ T^k\right\rangle\right|\end{aligned}$$

$$+\left|\left\langle\prod_{k=0}^{[ct/\epsilon^2h]}\theta_{\epsilon\xi}\circ T^k\right\rangle\left\langle\prod_{l=[c(t+\delta)/\epsilon^2h]}^{[c(t+\tau)/\epsilon^2h]}\theta_{\epsilon\eta}\circ T^k\right\rangle\right.$$
$$\left.-\left\langle\prod_{k=0}^{[ct/\epsilon^2h]}P_R(\theta_{\epsilon\xi})\circ T^k\right\rangle\left\langle\prod_{l=[c(t+\delta)/\epsilon^2h]}^{[c(t+\tau)/\epsilon^2h]}P_R(\theta_{\epsilon\eta})\circ T^k\right\rangle\right|. \tag{70}$$

The property (H3) can be used for the first term of the right-hand side of (70) which turns out to be zero for

$$R=\exp\left(\frac{1}{\beta_0}([c(t+\delta)/\epsilon^2h]-[ct/\epsilon^2h]-\beta_1)\right). \tag{71}$$

The estimates (68) and (69) can be used to control the second and third term of (70), with $|\epsilon\xi|\le 1$ and $|\epsilon\eta|\le 1$ their sum is bounded by

$$2A_R\left[\frac{c(t+\tau)}{\epsilon^2h}\right](1+A_R)^{[c(t+\tau)/\epsilon^2h]}\le 2A_R\left[\frac{c(t+\tau)}{\epsilon^2h}\right]\exp\left(A_R\left[\frac{c(t+\tau)}{\epsilon^2h}\right]\right). \tag{72}$$

With (69) and the choice of R given by (71), the right-hand side of (72) satisfies the following estimate:

$$A_R\left[\frac{c(t+\tau)}{\epsilon^2h}\right]\exp\left(A_R\left[\frac{c(t+\tau)}{\epsilon^2h}\right]\right)\le\frac{C"}{\epsilon^2}\exp\left(-\frac{C'\delta}{\epsilon^2}\right)\exp\left(\frac{C"}{\epsilon^2}\exp\left(-\frac{C'\delta}{\epsilon^2}\right)\right), \tag{73}$$

which leads to the function $H(K,\delta,\epsilon)$ and completes the proof. □

Corollary 9. Assume that the vector field a satisfies the assumptions of Proposition 8. Let $f\in C_0^\infty(\mathbb{R}^d)$ and $g\in C^\infty(\mathbb{R}^d)$ with subquadratic growth at infinity as follows:

$$|g(x)|+\sum_{1\le l\le[d/2+1]}|\nabla_x^l g(x)|\le C_g(1+|x|^2). \tag{74}$$

Then for any positive constant M

$$\left|\left\langle f\left(\epsilon h\sum_{k=0}^{[ct/\epsilon^2h]}a\circ T^k\right)g\left(\epsilon h\sum_{l=[c(t+\delta)/\epsilon^2h]}^{[c(t+\tau)/\epsilon^2h]}a\circ T^l\right)\right\rangle\right.$$
$$\left.-\left\langle f\left(\epsilon h\sum_{k=0}^{[ct/\epsilon^2h]}a\circ T^k\right)\right\rangle\left\langle g\left(\epsilon h\sum_{l=[c(t+\delta)/\epsilon^2h]}^{[c(t+\tau)/\epsilon^2h]}a\circ T^l\right)\right\rangle\right|$$
$$\le 2\|f\|_\infty C_g\sup_{N>0}\left\|\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N}a\circ T^k\right)\right\|_4^4\left(\frac{\sqrt{\tau-\delta+2\epsilon^2}}{M}\right)^2$$
$$+C_g\|f\|_sM^{(d+2)/2}\left(C_1\frac{\epsilon^{s-d/2}}{s-d/2}+\epsilon^{-d}H(K,\delta,\epsilon)\right), \tag{75}$$

where s is chosen equal to $[d/2+1]$, $\|f\|_s$ represent the H^s Sobolev norm of f and $H(K,\delta,\epsilon)$ is the function defined by (63). As before, the constants are independent of t, τ, ϵ and δ.

Proof. Introduce a cutoff function $\chi_M\in C^\infty(\mathbb{R}^2)$ with the following properties:

$$\forall x,\quad 0\le\chi_M\le 1,\qquad \chi_M(x)=1\ \text{if}\ |x|\le M,\quad \chi_M(x)=0\ \text{for}\ |x|\ge 2M \tag{76}$$

and denote by g_M the function $g\chi_M$. From the formula $|g-g_M| \le |g|\mathbf{1}_{|x|\ge M}$ one deduces the estimate

$$\begin{aligned}
&\Bigg|\Bigg\langle f\Bigg(\epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\Bigg) g\Bigg(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\Bigg)\Bigg\rangle \\
&\quad - \Bigg\langle f\Bigg(\epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\Bigg)\Bigg\rangle \Bigg\langle g\Bigg(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\Bigg)\Bigg\rangle\Bigg| \\
&\le \Bigg|\Bigg\langle f\Bigg(\epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\Bigg) g_M\Bigg(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\Bigg)\Bigg\rangle \\
&\quad - \Bigg\langle f\Bigg(\epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\Bigg)\Bigg\rangle \Bigg\langle g_M\Bigg(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\Bigg)\Bigg\rangle\Bigg| \\
&\quad + 2\|f\|_\infty \int\limits_{|\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l(\omega)|\ge M} \Bigg| g\Bigg(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l(\omega)\Bigg)\Bigg| \,\mathrm{d}\omega.
\end{aligned} \tag{77}$$

Corollary 7 shows that

$$\begin{aligned}
&\mathrm{mes}\left\{\omega \text{ s.t. } \Bigg|\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l(\omega)\Bigg| \ge M\right\} \\
&\le \sup_{N>0} \left\|\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N} a\circ T^k\right)\right\|_4^4 \left(\frac{\sqrt{\tau-\delta+2\epsilon^2}}{M}\right)^4.
\end{aligned} \tag{78}$$

Using the Cauchy–Schwarz inequality and the subquadratic growth of g (74), this implies

$$\begin{aligned}
&\int\limits_{\left|\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l(\omega)\right|\ge M} \Bigg| g(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l(\omega))\Bigg| \,d\omega \\
&\le C_g \sup_{N>0} \left\|\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N} a\circ T^k\right)\right\|_4^4 \left(\frac{\sqrt{\tau-\delta+2\epsilon^2}}{M}\right)^2.
\end{aligned} \tag{79}$$

Therefore, with estimate (62), one concludes that

$$\begin{aligned}
&\Bigg|\Bigg\langle f\Bigg(\epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\Bigg) g\Bigg(\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\Bigg)\Bigg\rangle \\
&\quad - \Bigg\langle f\Bigg(\epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a\circ T^k\Bigg)\Bigg\rangle \Bigg\langle g\epsilon h \sum_{l=[c(t+\delta)/\epsilon^2 h]}^{[c(t+\tau)/\epsilon^2 h]} a\circ T^l\Bigg)\Bigg\rangle\Bigg|
\end{aligned}$$

$$\le 2\|f\|_\infty C_g \sup_{N>0}\left\|\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N} a\circ T^k\right)\right\|_4^4 \left(\frac{\sqrt{\tau-\delta+2\epsilon^2}}{M}\right)^2$$
$$+\|f\|_s\|g_M\|_s\left(C\frac{\epsilon^{s-d/2}}{s-d/2}+\epsilon^{-d}H(K,\delta,\epsilon)\right) \tag{80}$$

Since (74) shows that $\|g_M\|_s$ is bounded by $C_g M^{(d+2)/2}$, (80) completes the proof of Corollary 9. □

4. Proof of Theorem 2

As it was said in the introduction the proof of Theorem 2 is inspired by the proof the Ito formula for the Brownian motion. Therefore the starting point is the Taylor formula at order three for the increment.

$$\langle\psi_\epsilon(t+\tau,x,\cdot,\cdot)\rangle-\langle\psi_\epsilon(t,x,\cdot,\cdot)\rangle$$
$$=\left\langle\nabla\phi\left(x-\epsilon h\sum_{k=0}^{[ct/\epsilon^2h]}a(T^k\omega)\right)\cdot\epsilon h\sum_{k=[ct/\epsilon^2h]+1}^{[c(t+\tau)/\epsilon^2h]}a(T^k\omega)\right\rangle$$
$$+\frac{1}{2}\left\langle\nabla^2\phi\left(x-\epsilon h\sum_{k=0}^{[ct/\epsilon^2h]}a(T^k\omega)\right):\left(\epsilon h\sum_{k=[ct/\epsilon^2h]+1}^{[c(t+\tau)/\epsilon^2h]}a(T^k\omega)\right)^{\otimes 2}\right\rangle$$
$$+\mathrm{O}\left(\left\langle\left|\left(\epsilon h\sum_{k=[ct/\epsilon^2h]+1}^{[c(t+\tau)/\epsilon^2h]}a(T^k\omega)\right)\right|^3\right\rangle\right). \tag{81}$$

The analysis of the limit for $\epsilon\to 0$ in the above expression will be done in three steps and C will be used to denote various constants independent of ϵ and τ.

Step 1: Estimate of the remainder. One has

$$\left\langle\left|\left(\epsilon h\sum_{k=[ct/\epsilon^2h]+1}^{[c(t+\tau)/\epsilon^2h]}a(T^k\omega)\right)\right|^3\right\rangle\le\left\langle\left|\left(\epsilon h\sum_{k=0}^{[c\tau/\epsilon^2h]}a(T^k\omega)\right)\right|^3\right\rangle+\mathrm{O}\left((\epsilon h)^3\right) \tag{82}$$

or, using Hölder's inequality with $N=[c\tau/\epsilon^2h]$

$$\left\langle\left|\left(\epsilon h\sum_{k=[ct/\epsilon^2h]+1}^{[c(t+\tau)/\epsilon^2h]}a(T^k\omega)\right)\right|^3\right\rangle\le\sqrt{2\pi}(\epsilon h\sqrt{N})^3\left(\left\langle\left|\frac{1}{\sqrt{N}}\sum_{k=0}^{N}a(T^k\omega))\right|^4\right\rangle\right)^{3/4}+\mathrm{O}\left((\epsilon h)^3\right). \tag{83}$$

Using Corollary 7, the following bound is deduced from (83)

$$\limsup_{\epsilon\to 0}\left\langle\left|\left(\epsilon h\sum_{k=[ct/\epsilon^2h]+1}^{[c(t+\tau)/\epsilon^2h]}a(T^k\omega)\right)\right|^3\right\rangle\le C(\sqrt{ch})^{3/2}\tau^{3/2}. \tag{84}$$

Step 2: Decorrelation in (81). Since the treatment of the linear term is simpler than the treatment of the quadratic term (but follows the same lines), only the latter will be considered in detail. A "small" positive time δ is introduced and one has

$$\begin{aligned}
&\left|\left\langle \nabla_x^2\phi\left(x-\epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right):\left(\epsilon h\sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega)\right)^{\otimes 2}\right\rangle\right.\\
&\left.-\left\langle \nabla_x^2\phi\left(x-\epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right):\left(\epsilon h\sum_{k=[c(t+\delta)/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega)\right)^{\otimes 2}\right\rangle\right|\\
&\le\left\langle \left|\nabla_x^2\phi\left(x-\epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right)\right|\left|\left(\epsilon h\sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\delta)/\epsilon^2 h]} a(T^k\omega)\right)\right|^2\right\rangle\\
&\quad+2\left\langle \left|\nabla_x^2\phi\left(x-\epsilon h\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right)\right|\left|\left(\epsilon h\sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\delta)/\epsilon^2 h]} a(T^k\omega)\right)\right|\right.\\
&\quad\left.\times\left|\left(\epsilon h\sum_{k=[c(t+\delta)/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega)\right)\right|\right\rangle. \qquad (85)
\end{aligned}$$

The first term of the right-hand side of (85) is bounded by

$$\begin{aligned}
&\|\nabla_x^2\phi\|_\infty\left\|\left(\epsilon h\sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\delta)/\epsilon^2 h]} a\circ T^k\right)\right\|_2^2\\
&\le\|\nabla_x^2\phi\|_\infty\sup_{N>0}\left\|\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N} a\circ T^k\right)\right\|_4^2(c\delta+2\epsilon^2 h)h. \qquad (86)
\end{aligned}$$

Similarly the second term is bounded by

$$\|\nabla_x^2\phi\|_\infty\sup_{N>0}\left\|\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N} a\circ T^k\right)\right\|_4^2\sqrt{(c\tau+2\epsilon^2 h)(c\delta+2\epsilon^2)}h. \qquad (87)$$

The only remaining term is

$$\left\langle \nabla_x^2\phi\left(x-\frac{\epsilon h}{c}\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right):\left(\frac{\epsilon h}{c}\sum_{k=[c(t+\delta)/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega)\right)^{\otimes 2}\right\rangle.$$

Here Corollary 9 is used with $g(x)=x^{\otimes 2}$ leading to the following estimate:

$$\begin{aligned}
&\left|\left\langle \nabla_x^2\phi\left(x-\frac{\epsilon h}{c}\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right):\left(\frac{\epsilon h}{c}\sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega)\right)^{\otimes 2}\right\rangle\right.\\
&\left.-\left\langle \nabla_x^2\phi\left(x-\frac{\epsilon h}{c}\sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right)\right\rangle:\left\langle\left(\frac{\epsilon h}{c}\sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega)\right)^{\otimes 2}\right\rangle\right|
\end{aligned}$$

$$\begin{aligned}
&\le \|\nabla_x^2\phi\|_\infty \sup_{N>0} \left\| \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N} a \circ T^k \right) \right\|_4^2 (c\delta + 2\epsilon^2 h)h \\
&+ \|\nabla_x^2\phi\|_\infty \sup_{N>0} \left\| \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N} a \circ T^k \right) \right\|_4^2 \sqrt{(c\tau + 2\epsilon^2 h)(c\delta + 2\epsilon^2 h)h} \\
&+ \|\nabla_x^2\phi\|_\infty \sup_{N>0} \left\| \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N} a \circ T^k \right) \right\|_4^2 (c\delta + 2\epsilon^2 h)h \\
&+ 2\|\nabla_x^2\phi\|_\infty \sup_{N>0} \left\| \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N} a \circ T^k \right) \right\|_4^4 \left(\frac{\sqrt{\tau - \delta + 2\epsilon^2}}{M} \right)^2 \\
&+ C\|\nabla_x^2\phi\|_s M^{(d+2)/2} \left(C_1 \frac{\epsilon^{s-d/2}}{s - d/2} + \epsilon^{-d} H(K, \delta, \epsilon) \right).
\end{aligned} \tag{88}$$

For a given $\tau \in (0, 1)$, and $0 < \epsilon < \tau$ one uses the special form of the function $H(K, \delta, \epsilon)$; it shows that, by choosing

$$\delta = \epsilon \quad \text{and} \quad M = \epsilon^{-(2s-d)/4(d+2)}$$

one has

$$\begin{aligned}
\limsup_{\epsilon\to 0} \Bigg\{ \Bigg| &\left\langle \nabla_x^2\phi\left(x - \frac{\epsilon h}{c} \sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega) \right) : \left(\frac{\epsilon h}{c} \sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega) \right)^{\otimes 2} \right\rangle \\
&- \left\langle \nabla_x^2\phi\left(x - \frac{\epsilon h}{c} \sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega) \right) \right\rangle : \left\langle \left(\frac{\epsilon h}{c} \sum_{k=[ct/\epsilon^2 h]+1}^{[c(t+\tau)/\epsilon^2 h]} a(T^k\omega) \right)^{\otimes 2} \right\rangle \Bigg| \Bigg\} = 0.
\end{aligned} \tag{89}$$

Step 3: Weak and strong limits. Denote by $u_\epsilon(t, x)$ the family of functions $\langle \psi_\epsilon(t, x, \cdot) \rangle$. Starting with (81), using the estimate for the remainder (84), the L^∞ bounds on $\nabla\phi$ and $\nabla^2\phi$, the inequality

$$\#\{k \text{ s.t. } [ct/\epsilon^2 h] + 1 \le k \le [c(t + \tau)/\epsilon^2 h]\} \le [c\tau/\epsilon^2 h] + 1, \tag{90}$$

and property (H3) from Corollary 7, we arrive at the (uniform in $\tau \in [0, 1]$ and $\epsilon \in [0, 1]$) bound:

$$|u_\epsilon(t + \tau, x) - u_\epsilon(t, x, \cdot)| \le C\sqrt{\tau + \epsilon^2}. \tag{91}$$

It follows from Ascoli's theorem that the family u_ϵ is relatively compact in $C^0([0, \tau]; w^* - L^\infty(\mathbb{R}^d))$. Let u be a limit point of this family and rename as usual, $f_\epsilon^\pm$, ψ_ϵ and u_ϵ the corresponding subfamilies, with u_ϵ converging to u. Starting from the "Ito" formula (82), using (84) and (89) one obtains

$$\begin{aligned}
&u(t + \tau, x) - u(t, x) \\
&\quad = \lim_{\epsilon\to 0} \{u_\epsilon(t + \tau, x) - u_\epsilon(t, x)\} \\
&\quad = \tau c h \lim_{\epsilon\to 0} \left\{ \frac{1}{2} \nabla_x^2 u_\epsilon(t, x) : \lim_{\epsilon\to 0} \left\langle \left(\frac{\epsilon\sqrt{h}}{\sqrt{c\tau}} \sum_{k=0}^{[c(t+\tau)/\epsilon^2 h]-[ct/\epsilon^2 h]-1} a(T^k\omega) \right)^{\otimes 2} \right\rangle \right\} + O(\tau)^{3/2}
\end{aligned} \tag{92}$$

or with (12) (cf. Proposition 2(2).)

$$u(t+\tau, x) - u(t, x) = \tfrac{1}{2}hc\tau D(a) : \nabla_x^2 u(t, x) + \mathrm{O}(\tau)^{3/2}. \tag{93}$$

Dividing (93) by τ and letting τ go to zero, one sees that u solves the initial value problem for the diffusion equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} - \tfrac{1}{2}hcD(a) : \nabla_x^2 u, \quad u(x, 0) = \phi(x). \tag{94}$$

The solution of (94) being unique, it follows that the whole families

$$\langle f_\epsilon^{\pm}(t, x, \omega)\rangle \quad \text{or} \quad u_\epsilon(x, t) = \langle \psi_\epsilon(t, x, \omega)\rangle = \left\langle \phi\left(x - \epsilon h \sum_{k=0}^{[ct/\epsilon^2 h]} a(T^k\omega)\right)\right\rangle$$

converge in $C^0([0, \tau]; w^* - L^\infty(\mathbb{R}^d))$ to $u(t, x)$.

Observing that the problem (7)–(10) is translation invariant in the variable $x \in \mathbb{R}^d$ and using the regularity of $\phi \in C^2(\mathbb{R}^d)$, one can see that the family of functions $u_\epsilon(t, x)$ satisfies the (uniform in ϵ) estimate

$$\|u_\epsilon(t, x)\|_{L^\infty(\mathbb{R}_t^+, C^2(\mathbb{R}_x^d))} \leq C. \tag{95}$$

Acoli's theorem, (91) and (95) show that for any $\tau > 0$ and any compact $K \subset \mathbb{R}^d$, the sequence $u_\epsilon(t, x)$ converges to $u(t, x)$ in $C([0, \tau] \times K)$. This argument completes the proof of the strong convergence of the averages (14a).

Finally let f be in the w^* closure in $L^\infty(\mathbb{R}_t^+ \times \mathbb{R}_x^d \times \mathbb{T}^2)$ of the family $\psi_\epsilon(t, x, \omega)$. One deduces from (11) that f is invariant under the action of T

$$f(t, x, \omega) = f(t, x, T\omega). \tag{96}$$

The ergodicity property (Proposition 2 (1)) implies that f is independent of ω and therefore coincide with the function $u(t, x)$: this demonstrates the convergence (14b) and concludes the proof of Theorem 3. □

5. Final remarks and numerical experiments

The model (3)–(5) defines a global broken hamiltonian flow which induces an isometry on $L^p(\mathbb{R}^d \times \mathbb{T}^2)$ for all $1 \leq p \leq \infty$. In particular the quantity $\|f_\epsilon^+(t, \cdot, \cdot)\|_2 + \|f_\epsilon^-(t, \cdot, \cdot)\|_2$ is conserved for any $\epsilon > 0$ at variance with the quantity $\|u(t, \cdot)\|_2$. Therefore the type of convergence which is given in Theorem 3 (strong for the average and weak for the solution itself) is optimal with the exception of the trivial case where $D(a) = 0$ which correspond to no diffusion. In this situation the solution exhibits an initial layer near $t = 0$ which is taken care of by the time scaling and the solution $f_\epsilon^{\pm}(x, t, z\omega)$ converges strongly to its initial value $\phi(x)$

The paradox of deriving a well-posed irreversible problem for $t > 0$ from a reversible problem can be explained from the following facts.

(1) The scaling has been done with the a priori choice of considering the solution for large positive times.

(2) It gives the correct approximation at the order ϵ of local averages of the solution in terms of the local averages of its initial data; the dependence on ω is lost in the approximation. In some sense, this decay of information can be estimated by the decay of the L^2 norm of u which, for the diffusion equation, is the linearized version of the classical entropy.

The relevance of the above remarks depends of course on the analysis of the strict positivity of the diffusion matrix $D(a)$. In fact it results from proposition 2 that $D(a) = 0$ for a varying in a dense subset of $L^2(\mathbb{T}^2)/\mathbb{R}$ (isomorphic to

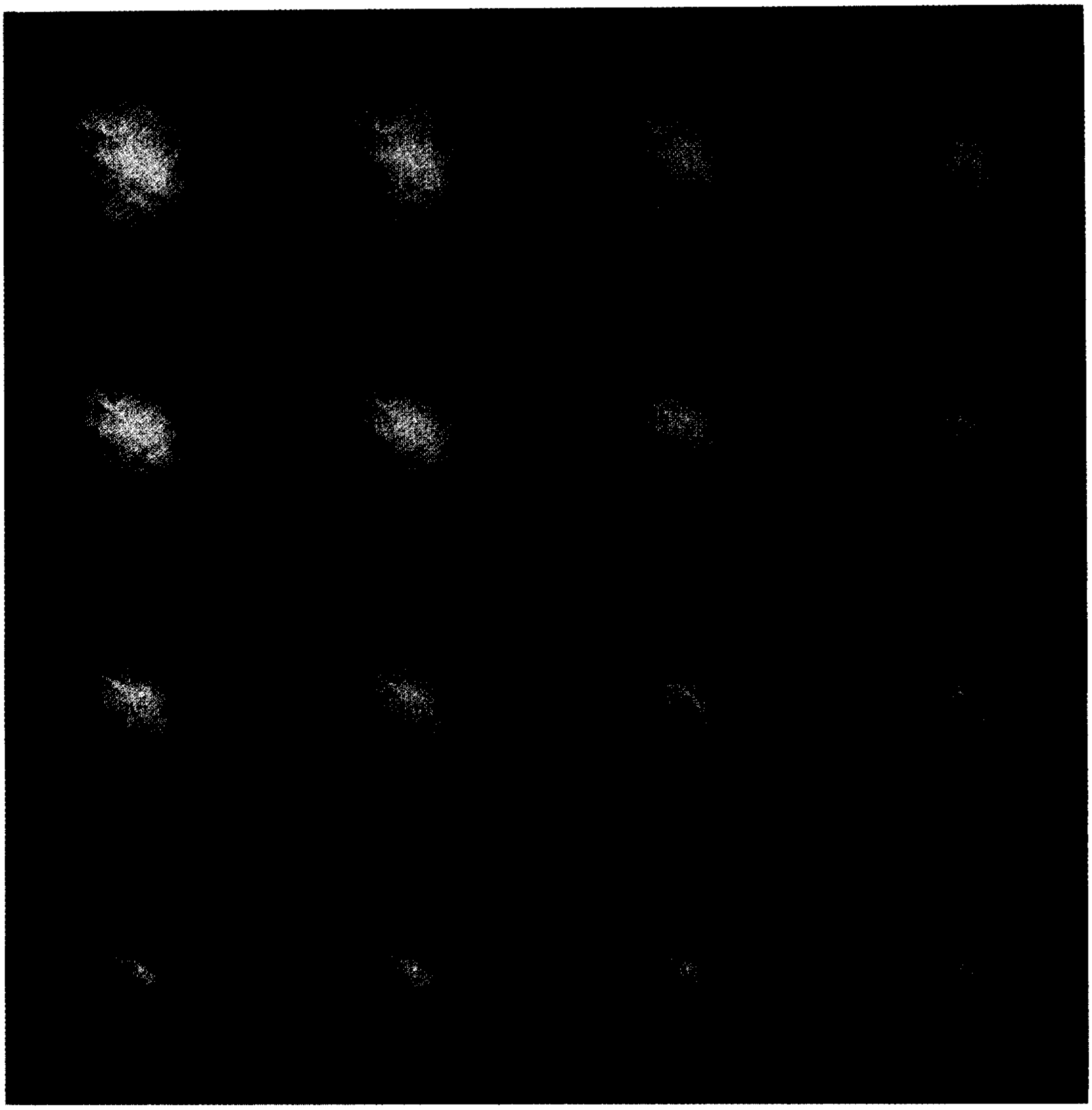

Fig. 3. The time evolution of the particle density when a is not a coboundary.

the space of mean zero functions in $L^2(\mathbb{T}^2)$). Furthermore the diffusion is degenerate in the directions ξ for which the "excursion length"

$$f_N \cdot \xi = \sum_{k=1}^{N} (a \circ T^k) \cdot \xi$$

is (uniformly with respect to N) bounded in $L^2(\mathbb{T}^2)$, an observation which turns out to be in agreement with the intuition.

Although the space of coboundaries is dense in the subspace of $L^2(\mathbb{T}^2)$ consisting of mean zero functions, it is not L^2 closed. In other words, it is possible to find smooth funtions of $L^2(\mathbb{T}^2)$ not being coboundaries. For example, observe that for $d = 1$ and $a(\omega_1, \omega_2) = \cos \omega_1$ one has $D(a) = \frac{1}{4}$.

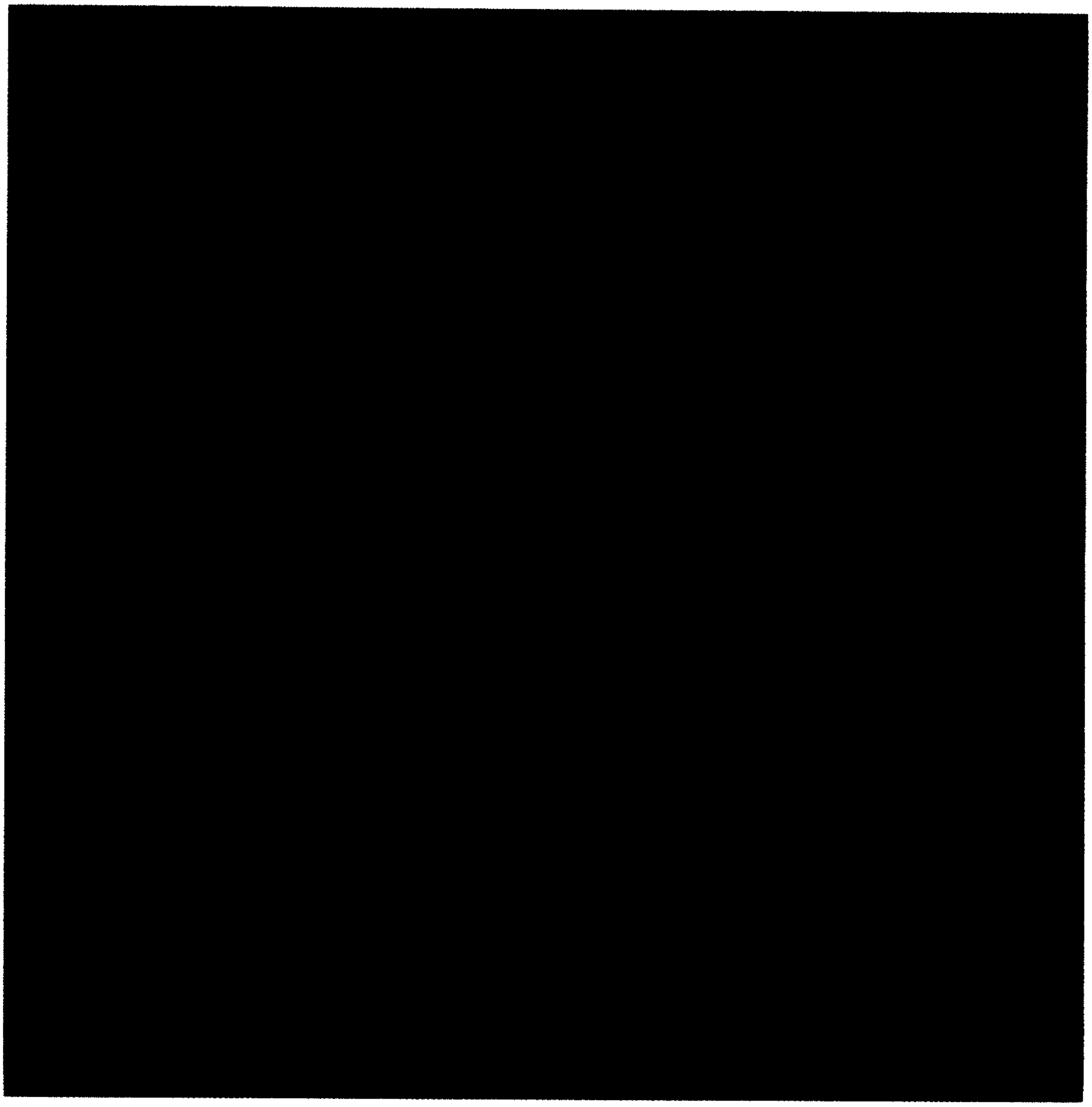

Fig. 4. Some individual trajectories.

More generally it would be extremely useful to have an explicit expression of $D(a)$. For $d = 1$ it can be easily obtained in terms of the Fibonacci sequence F_n

$$F_0 = 0, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2}. \tag{97}$$

Observe that the minimal polynomial of the matrix M is $X^2 - 3X + 1$ and introduce the matrix $P = M - I$ which has minimal polynomial $X^2 - X - 1$, then on one hand one has

$$M = P + 1 = P^2 \tag{98}$$

and on the other hand

$$P^n = P^{n-1} + P^{n-2} \tag{99}$$

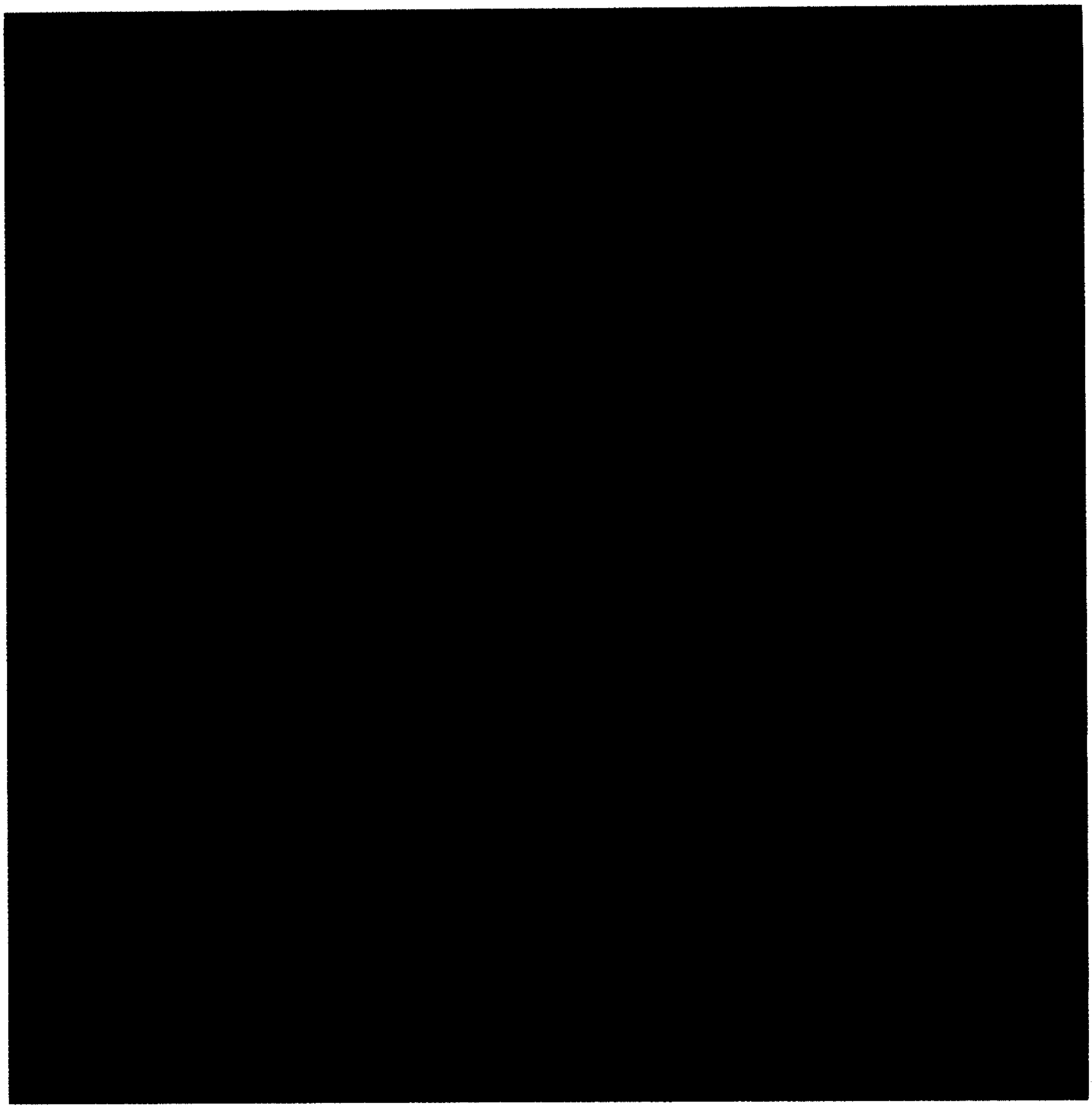

Fig. 5. The time evolution of the particle density when a is a coboundary.

and this implies the formula

$$M^n = \begin{pmatrix} F_{2n+1} & F_{2n} \\ F_{2n} & F_{2n-1} \end{pmatrix}.$$

Therefore, with the diffusion matrix given by (12) one has, for any function $f : \mathbb{T}^2 \to \mathbb{R}$ in H^s with $s > 0$:

$$D(f) - \frac{1}{2}\langle f^2 \rangle$$

$$= \sum_{n \geq 1} \langle f \circ T^n \cdot f \rangle = \frac{1}{4\pi^2} \sum_{n \geq 1} \sum_{k \in \mathbb{Z}^2 - \{0\}} \hat{f}(M^n k) \hat{f}(-k)$$

$$= \frac{1}{4\pi^2} \sum_{k\in\mathbb{Z}^2-\{0\}} \sum_{n\geq 1} \hat{f}(F_{2n+1}k_1 + F_{2n}k_2, F_{2n}k_1 + F_{2n-1}k_2)\hat{f}(-k_1, -k_2). \tag{100}$$

In particular, since the Fibonacci sequence is rapidly increasing this will provide an exact formula (involving a small number of nonzero terms) for $D(f)$ whenever f is a trigonometric polynomial.

The numerical experiments were done by the third author. They intend to illustrate the difference between the diffusive and the nondiffusive case. In two space variables the trajectories of 128 particles over 1000 interactions with the upper and lower boundary have been obtained. The diffusive case (see Fig. 3) corresponds to horizontal velocity field:

$$a(\omega) = (a_1(\omega),\ a_2(\omega)) = (\cos\omega_1,\ \cos\omega_2). \tag{101}$$

As shown on Fig. 4 the trajectory of a single particle is in general ergodic, however notice that even in this case some exceptional trajectories are not ergodic. This will be the case for any particle starting with a velocity $a(\omega_0)$ with ω_0 any periodic point for the mapping T. For instance in Fig. 4 is plotted the path of a single particle driven by the flow given by (101) with initial velocity:

$$a(\omega_0),\quad \omega_0 = \left(0,\ \tfrac{1}{2}\pi\right). \tag{102}$$

Observe the relation $\omega_0 = T^3\omega_0$ which correspond to the behavior of the particle.

Fig. 5 is devoted to the simulation (128 particles and 1000 collisions) of the nondiffusive case with a vector field given by the formula:

$$a'(\omega) = a(T\omega) - a(T)$$

First there is an initial layer then the process stabilizes to a stationary state. The various colors in Figs. 4 and 5 are coding the different direction of the velocity vectors in the 3-dimensional space.

Acknowledgements

We are grateful to Prof. J.-M. Strelcyn who indicated to us Ref. [18] and to Prof. M. Farge, who suggested the numerical experiments illustrating this work. C. Bardos was supported by the Humboldt Foundation during the Academic Year 1994–1995 and thanks the Arbeitsgruppe Technomathematik in the University of Kaiserlautern where part of this work has been done.

References

[1] C. Bardos, F. Golse and D. Levermore, J. Stat. Phys. 63 (1991) 323–344.
[2] C. Bardos, F. Golse and D. Levermore, Commun. Pure Appl. Math. 46 (1993) 667–753.
[3] C. Bardos and S. Ukai, Math. Methods and Models in the Appl. Sci. 1 (1991) 235–257.
[4] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, Lectures Notes in Mathematics, Vol. 470 (Springer, Berlin, 1975).
[5] L. Bunimovich and Ya. Sinai, Commun. Math. Phys. 78 (1980) 247–280.
[6] L. Bunimovich and Ya. Sinai, Commun. Math. Phys. 78 (1980) 479–497.
[7] L. Bunimovich, Ya. Sinai and N. Chernov, Russian Math. Surveys 45 (1990) 105–152.
[8] L. Bunimovich, Ya. Sinai and N. Chernov, Russian Math. Surveys 46 (4) (1991) 47–106.
[9] R. Caflisch, Commun. Pure Appl. Math. 33 (1980) 651–666.
[10] C. Cercignani, R. Illner and M. Pulvirenti, The Mathematical Theory of Dilute Gases (Springer, New York, 1994).
[11] J. Crawford and J. Cary, Physica D 6 (1982–1983) 223–232.

[12] R. Dautray, Méthodes probabilistes pour les équations de la physique (Masson, Paris, 1989)
[13] A. De Masi, R. Esposito and J. Lebowitz, Commun. Pure Appl. Math. 42 (1990) 1189–1214.
[14] M. Denker and W. Philipp, Ergodic Theory and Dyn. Syst. 4 (1984) 541–552.
[15] N. Dunford and J. Schwartz, Linear Operators, Part I: General Theory (Interscience, New York, 1958).
[16] A. Garrett, in: From Statistical Physics to Statistical Inference and Back, eds. P. Grassberger and J.-P. Nadal (1994) 45–75.
[17] R. Illner and M. Pulvirenti, Commun. Math. Phys. 105 (1986) 189–203; 121 (1989) 143–146.
[18] Y. Katznelson, Israel J. Math. 10 (1971) 186–195.
[19] O. Lanford, in: Lecture Notes in Physics, ed. J. Moser, Vol. 38 (Springer, Berlin, 1975).
[20] J. Lebowitz, Physica A 194 (1993) 1–27.
[21] M. Ratner, Israel J. Math. 16 (1973) 181–197.
[22] H. Babovski, Limit theorems for deterministic Knudsen flows between two plates, preprint No. 147, Weierstrass Institut AAS, Berli (1995).