Theoretical and Mathematical Physics, **151**(1): 539–555 (2007)

FINE-GRAINED AND COARSE-GRAINED ENTROPY IN PROBLEMS OF STATISTICAL MECHANICS

V. V. Kozlov* **and D. V. Treshchev*†**

We consider dynamical systems with a phase space Γ that preserve a measure μ. A partition of Γ into parts of finite μ-measure generates the coarse-grained entropy, a functional that is defined on the space of probability measures on Γ and generalizes the usual (ordinary or fine-grained) Gibbs entropy. We study the approximation properties of the coarse-grained entropy under refinement of the partition and also the properties of the coarse-grained entropy as a function of time.

Keywords: invariant measure, Gibbs entropy, coarse-grained entropy

1. Definitions

Let (Γ, dist) be a metric space with a measure μ, and let ν be a probability measure with a measurable density $\rho > 0$,

$$d\nu = \rho\, d\mu, \qquad \nu(\Gamma) = \int_\Gamma \rho\, d\mu = 1.$$

In particular, $\rho \in L_1(\Gamma, \mu)$.

Let $\{\Gamma_j\}_{j\in J}$ be a partition of Γ into measurable subsets,

$$\gamma_j = \mu(\Gamma_j), \quad 0 < \mu(\Gamma_j) < \infty, \quad j \in J.$$

The set J is assumed to be finite or countable. The expression

$$\sup_{j\in J}(\mathrm{diam}\,\Gamma_j) \le \infty$$

(the diameter $\mathrm{diam}\,\Gamma_j$ is taken in the metric of the space Γ) is called the diameter of the partition $\{\Gamma_j\}$. We set

$$\rho_j = \frac{\lambda_j}{\gamma_j}, \qquad \lambda_j = \int_{\Gamma_j} \rho\, d\mu, \qquad \sum_{j\in J} \lambda_j = 1,$$

and consider a new density $\bar\rho\colon \Gamma \to \mathbb{R}$ such that $\bar\rho|_{\Gamma_j} = \rho_j$, $j \in J$. We call $\bar\rho$ the *coarse-grained* density. The corresponding measure $\bar\nu$, $d\bar\nu = \bar\rho\, d\mu$, is also a probability measure because

$$\int_\Gamma \bar\rho\, d\mu = \sum_{j\in J} \int_{\Gamma_j} \rho_j\, d\mu = \sum_{j\in J} \rho_j \gamma_j = 1.$$

*Steklov Mathematical Institute, RAS, Moscow, Russia, e-mail: kozlov@pran.ru.

†Moscow State University, Moscow, Russia, e-mail: treschev@mi.ras.ru.

Translated from Teoreticheskaya i Matematicheskaya Fizika, Vol. 151, No. 1, pp. 120–137, April, 2007. Original article submitted July 24, 2006.

We define a functional $\mathbf{S}$ such that

$$\mathbf{S}(\alpha) = -\int_\Gamma \alpha \log \alpha \, d\mu$$

for an arbitrary nonnegative μ-measurable function $\alpha\colon \Gamma \to \mathbb{R}$ under the condition that the integral converges to a finite or infinite value. As usual, the function $\alpha \log \alpha$ is additionally defined as zero at $\alpha = 0$.

We define the *fine-grained* entropy $s = \mathbf{S}(\rho)$ and also the *coarse-grained* entropy $\bar{s} = \mathbf{S}(\bar{\rho})$. The relation

$$\bar{s} = -\sum_{j\in J} \gamma_j \rho_j \log \rho_j = -\sum_{j\in J} \lambda_j \log \lambda_j + \sum_{j\in J} \lambda_j \log \gamma_j \tag{1.1}$$

holds. In particular, if all γ_j are equal to each other, then we have

$$\bar{s} = -\sum_{j\in J} \lambda_j \log \lambda_j + \log \gamma, \qquad \gamma = \gamma_j, \quad j \in J,$$

and the coarse-grained entropy hence coincides with the *information* entropy $-\sum \lambda_j \log \lambda_j$ up to an additive constant (depending only on γ). We note that in the case of a discrete probability distribution, a formula of form (1.1) for the entropy is used in the theory of equilibriums (see, e.g., [1]).

For fixed values of γ_j, the maximum value of $\bar{s}$ in (1.1) is attained for

$$\lambda_j = \frac{\gamma_j}{\sum_{i\in J} \gamma_i}$$

under the condition $\mu(\Gamma) < \infty$. If the measure μ of the phase space is infinite and all γ_j are bounded, then $\sup \bar{s} = +\infty$.

As was established by Gibbs, the inequality

$$s \le \bar{s} \tag{1.2}$$

holds. It is a simple consequence of the Jensen inequality for the convex function $\rho \log \rho$ (see [2] for an instructive discussion).

2. An example of the absence of approximation

If $\mu(\Gamma) = \infty$, then the coarse-grained entropy, generally speaking, does not approximate the fine-grained entropy even if the partition diameter is arbitrarily small.

Example 1. Let Γ be the space $\mathbb{R}$, $\Gamma = \mathbb{R}$, with the Lebesgue measure. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence such that

$$0 \le a_n < 1, \qquad \sum_{n=1}^{\infty} a_n = 1, \qquad \sum_{n=1}^{\infty} a_n \log a_n = -\infty.$$

For example, $c/(n \log^{1+\varepsilon} n)$, $0 < \varepsilon < 1$, can be taken as a_n.

We consider a probability measure ν with the density

$$\rho(x) = \begin{cases} 1 & \text{for } x \in [n, n + a_n] \text{ with some } n \in \mathbb{N}, \\ 0 & \text{for the other values of } x. \end{cases}$$

The fine-grained entropy s for this measure is obviously zero.

We consider a partition of the form

$$\Gamma_j = \left\{ x \in \mathbb{R}\colon \frac{j}{K} \le x < \frac{(j+1)}{K} \right\}, \quad j \in \mathbb{Z},$$

where K is an arbitrary positive integer. The diameter of the partition $\{\Gamma_j\}$ is equal to $1/K$.

Proposition. *For every positive integer K, the coarse-grained entropy $\bar{s}$ is equal to $+\infty$.*

Proof. We have the relation

$$\bar{s} = -\sum_{j\in\mathbb{Z}} \frac{1}{K}\rho_j \log \rho_j, \qquad \rho_j = K\int_{j/K}^{(j+1)/K} \rho(x)\,dx \le 1.$$

Let there be an $N \in \mathbb{N}$ such that the inequality $a_n < 1/K$ holds for all $n > N$. We then obtain $\rho_{Kn} = Ka_n$ for $n > N$.

Because $\rho_j \log \rho_j < 0$ for all $j \in \mathbb{Z}$, the estimate

$$\bar{s} \ge -\sum_{n>N} \frac{1}{K}\rho_{Kn} \log \rho_{Kn} = -\sum_{n>N} a_n \log(Ka_n) = +\infty$$

holds, which was to be proved.

This assertion disproves the widespread opinion that the coarse-grained entropy tends to the fine-grained entropy as the partition is refined (cf. [3] and [4]). We also note that if $\mu(\Gamma) = \infty$, then the coarse-grained density, generally speaking, does not tend to the fine-grained density (with respect to the norm in $L_1 = L_1(\Gamma, \mu)$) as the partition diameter decreases indefinitely. Furthermore, the coarse-grained density does not tend to the fine-grained density even in the weak sense. Weak convergence of a sequence of functions ρ_n in L_1 to a function $\rho \in L_1$ means that

$$\int_\Gamma \rho_n \varphi\,d\mu \to \int_\Gamma \rho\varphi\,d\mu$$

for every test function $\varphi \in L_\infty(\Gamma, \mu)$.

3. Approximation theorems

3.1. Compact case. As a rule, the coarse-grained entropy (density) approximates the fine-grained entropy (density) in the compact case. Here, the main requirement is that the structures of the metric and measurable spaces on Γ be compatible. Namely, we assume that the space $C^0(\Gamma)$ of continuous functions on Γ is dense in $L_1(\Gamma, \mu)$.

Theorem 1. *Let Γ be a compact space, let $\mu(\Gamma) = 1$, and let $C^0(\Gamma)$ be dense in $L_1(\Gamma, \mu)$. Then the density $\bar{\rho}$ approximates ρ in the $L_1(\Gamma, \mu)$ metric with an arbitrary prescribed accuracy as the diameter of the partition $\{\Gamma_j\}$ decreases indefinitely.*

In other words, as the partition diameter decreases indefinitely, the coarse-grained density converges weakly to the fine-grained density. This property proves essential in passing from the micro- to the macro-description (i.e., in investigating the evolution of means of dynamical parameters).

The proof of Theorem 1 in given in Sec. A.1 in the appendix.

Theorem 2. *Let Γ be a compact space, let $\mu(\Gamma) = 1$, let $C^0(\Gamma)$ be dense in $L_1(\Gamma, \mu)$, and let $|s| < \infty$. Then the entropy $\bar{s}$ approximates s with an arbitrary prescribed accuracy as the diameter of the partition $\{\Gamma_j\}$ decreases indefinitely.*

The proof of Theorem 2 is based on the following two lemmas.

Lemma 1. *Theorem 2 holds under the additional assumption that $\rho < \Delta$ for some $\Delta > 1$.*

As usual, let e be the base of the natural logarithms.

Lemma 2. *Theorem 2 holds under the additional assumption that* $\delta < \rho < \Delta$ *for some* $\delta \in (0, 1/e)$ *and* $\Delta > 1$.

To prove Theorem 2, we prove

that Lemma 1 implies Theorem 2 (see Sec. A.2 in the appendix),

that Lemma 2 implies Lemma 1 (see Sec. A.3 in the appendix), and

Lemma 2 itself (see Sec. A.4 in the appendix).

We note that the condition $\sup_{j\in J} \operatorname{diam}(\Gamma_j) \to 0$ in Theorems 1 and 2 cannot be replaced by the weaker condition $\sup_{j\in J} \mu(\Gamma_j) \to 0$. In this connection, we present a simple example. Let Γ be the square

$$\{(x, y) \in \mathbb{R}^2 \colon 0 \le x \le 1,\ 0 \le y \le 1\},$$

let μ be the standard Lebesgue measure on Γ, let $\rho(x, y) = y$, and let the measurable parts Γ_j in the partition of Γ be the strips

$$\left\{\frac{j-1}{N} \le x \le \frac{j}{N},\ 0 \le y \le 1\right\}, \quad j = 1, \dots, N.$$

Then $\lim_{N\to\infty} \mu(\Gamma_j) = 0$, but the diameter of Γ_j does not tend to zero. It can be easily shown that for a typical density $\rho(x, y)$ (in an arbitrary reasonable sense), we have

$$\lim_{N\to\infty} \bar{\rho} \ne \rho, \qquad \lim_{N\to\infty} \bar{s} \ne s.$$

3.2. Noncompact case. If the density ρ tends to zero at infinity sufficiently fast, then the coarse-grained entropy also approximates the fine-grained entropy in the noncompact case. To state the exact result, we need some definitions.

Definition 1. The space $(\Gamma, \mu, \operatorname{dist})$ is said to be of type n if there is a sequence of compact sets $K_0, K_1, \dots$ such that

a. $K_0 \subset K_1 \subset \dots \subset \Gamma$,

b. $\Gamma = \bigcup_{l=0}^{\infty} K_l$,

c. $\mu(K_0) < \infty$ and $\mu(K_{l+1} \setminus K_l) \le C l^{n-1}$ for some constant $C > 0$, and

d. the inequality $\operatorname{dist}(x, y) < 1$ is possible for arbitrary points $x \in K_l$ and $y \in K_s$ only if $|l - s| \le 1$.

The simplest example of a space of type n is the space $\mathbb{R}^n$ with a Lebesgue or Euclidean metric. Here, the balls of radii j with a common center can be taken as the compact sets K_j.

Theorem 3. *Let*

1. *the space* $(\Gamma, \mu, \operatorname{dist})$ *be of type* n,

2. $C^0(K_l)$ *be dense in* $L_1(K_l, \mu)$ *for all* $l = 0, 1, \dots,$

3. $\rho|_{K_l} < c_\rho l^{-n-\delta}$, $l = 0, 1, \dots,$ *where* c_ρ *and* δ *are some positive constants, and*

4. $|s| < \infty$.

Then the entropy $\bar{s}$ *approximates* s *with an arbitrary prescribed accuracy as the diameter of the partition* $\{\Gamma_j\}$ *decreases indefinitely.*

The proof of Theorem 3 is given in Sec. A.5 in the appendix.

4. Density stabilization problem

Now let Γ be the phase space of a dynamical system determined by a flow (a one-parameter transformation group) g^t, $t \in \mathbb{R}$, preserving a measure μ. We consider a probability measure $\nu = \nu^t$ with a μ-measurable density $\rho = \rho^t \geq 0$, $d\nu^t = \rho^t \, d\mu$. By definition, for an arbitrary $t \in \mathbb{R}$ and for every measurable set $D \subset \Gamma$, we have

$$\nu^t(D) = \nu^0(g^{-t}(D)).$$

It follows that $\rho^t = \rho^0 \circ g^t$. If Γ is a smooth manifold and g^t is a flow determined by a vector field v, then in local coordinates, ρ^t satisfies the Liouville equation[1]

$$\frac{\partial \rho^t}{\partial t} - \operatorname{div}(\rho^t v) = 0.$$

The coarse-grained density $\bar{\rho}^t$ is defined as

$$\bar{\rho}^t|_{\Gamma_j} = \rho_j(t) = \frac{1}{\gamma_j} \int_{\Gamma_j} \rho^t \, d\mu.$$

We are primarily interested in Hamiltonian systems, although many results also hold in the case of systems of a more general form. In the Hamiltonian case, μ is an invariant Liouville measure, i.e., the volume element of the phase space. Of course, $\mu(\Gamma) = \infty$ for natural systems, but the restriction of μ to an energy level may turn out to be a finite measure.

In the constructions related to the Gibbs entropy, an important role is played by the existence of the limits

$$\lim_{t \to +\infty} \rho^t, \qquad \lim_{t \to -\infty} \rho^t. \tag{4.1}$$

These limits should be understood in the sense of weak convergence in one of the function spaces, of which the spaces $L_1(\Gamma, \mu)$ and $L_2(\Gamma, \mu)$ are most important in the given context. The existence of similar limits for the coarse-grained densities $\bar{\rho}^t$ is also of considerable interest.

As is well known, the density ρ^t, generally speaking, has no limit as $t \to \infty$. This is the major obstacle to justifying the "zeroth" law of thermodynamics in the theory of Gibbs ensembles. One attempt to overcome this difficulty is to introduce the coarse-grained density $\bar{\rho}^t$ generated by a partition $\{\Gamma_j\}$ of the phase space. Gibbs tried to prove (see Chap. 12 in [5]) that in a typical case as $t \to \infty$, the coarse-grained density $\bar{\rho}^t$ converges to a function depending only on the total energy of the Hamiltonian system. In this connection, it was written in [2] that it is almost hopeless to try to prove this assertion, that it is stronger than the ergodic theorem, and that the well-known argument by Gibbs himself (based on analogies with the mixing of liquids), even if the essential errors involved are discarded, at best serves only to indicate the plausibility of this "theorem" (see Chap. 3 in [2]).

But Gibbs himself noted that in the case of *linear* Hamiltonian systems, the coarse-grained density $\bar{\rho}^t$ oscillates and has no limit at all with indefinitely increasing time. On the other hand, the Gibbs assumption certainly holds for Hamiltonian systems with mixing on isoenergy surfaces. This observation belongs to Krylov [3], who did not however note an important fact: the limits of the coarse-gained density $\bar{\rho}^t$ as $t \to -\infty$ and $t \to +\infty$ *coincide* for the systems with mixing. This future–past symmetry (which is implied by the reversibility of natural Hamiltonian systems) contradicts the traditional idea of a one-sided approach of an isolated system to thermal equilibrium.

[1] Traditionally, the Liouville equation has a plus sign before the second term. The equation we use here results from the change of variable $t \mapsto -t$ in the standard Liouville equation.

Below, we discuss the Gibbs assumption on the approach to thermal equilibrium (in a somewhat weakened statement) for *quasihomogeneous* Hamiltonian systems. We recall that a Hamiltonian system

$$\dot{x}_j = \frac{\partial H}{\partial y_j}, \qquad \dot{y}_j = -\frac{\partial H}{\partial x_j}, \quad j = 1, \dots, n, \tag{4.2}$$

is said to be quasihomogeneous if there are some real constants (homogeneity weights) α, β, and γ, $\alpha+\beta+1 = \gamma$, such that

$$H(\lambda^\alpha x, \lambda^\beta y) = \lambda^\gamma H(x, y)$$

for all $\lambda > 0$. In other words, the Hamiltonian equations are invariant under the substitutions

$$t \mapsto \frac{t}{\lambda}, \qquad x \mapsto \lambda^\alpha x, \qquad y \mapsto \lambda^\beta y.$$

As usual, the Liouville measure having a unit density in the coordinates (x, y) is taken as the invariant measure μ.

We present two examples.

Example 2. We consider systems with a homogeneous potential,

$$H = \frac{1}{2} \sum y_j^2 + V_m(x),$$

where m is the degree of homogeneity of the potential energy V_m. Here,

$$\alpha = \frac{2}{m-2}, \qquad \beta = \frac{m}{m-2}, \qquad \gamma = \frac{2m}{m-2}.$$

In particular, this includes the n-body problem with the Newtonian potential ($\alpha = -2/3$, $\beta = 1/3$, and $\gamma = 2/3$ in this case). The exceptional case $m = 2$ corresponds to linear systems, which are not quasihomogeneous.

Example 3. We consider the case of inertial motion,

$$H = \frac{1}{2} \sum a_{jk}(x) y_j y_k, \quad x \in M,$$

where M is a smooth Riemannian manifold. Here, $\alpha = 0$, $\beta = 1$, and $\gamma = 2$. This also includes the billiard system in which M is a manifold with a boundary and there is elastic reflection from the boundary.

Theorem 4. *Let Hamiltonian system* (4.2) *be quasihomogeneous, and let the initial density* ρ *be a function belonging to* $L_p(\Gamma, \mu)$, $1 \le p \le \infty$. *Then the limits in* (4.1) *exist* (*in the sense of weak* $L_p(\Gamma, \mu)$ *convergence*) *and coincide.*

We make some remarks.

1. Theorem 4 was proved in [6]. It holds for a more general class of dynamical systems (not necessarily Hamiltonian), the so-called foliated flows introduced in [7].

2. Let $\mathcal{A}^k$ be the class of systems of smoothness k in which limits (4.1) exist for every density $\rho \in L_p(\Gamma, \mu)$.

Hypothesis. *For an arbitrary sufficiently large k (including $k = \infty$ and $k = \omega$), the set $\mathcal{A}^k$ consists of systems in general position in the space of systems of smoothness k.*

Rather little is known about the validity of this hypothesis. If there is a mixing hyperbolic system (an Anosov system) on (Γ, μ), then the interior of $\mathcal{A}^k$ in the C^k topology is nonempty. Indeed, hyperbolic mixing systems obviously lie in $\mathcal{A}^k$; they, as is known, are structurally stable.

Unfortunately, it should not be expected that the condition of general position in the hypothesis can be understood in the sense that $\mathcal{A}^k$ is open and everywhere dense for $k < \omega$. The reason is that a smooth system with homoclinic tangency, according to [8], can be approximated arbitrarily accurately in the C^∞ topology by a system having an invariant set on which the dynamics are a rigid rotation. No limits (4.1) exist in such a system for the majority of initial densities ρ. (Strictly speaking, the results in [8] are obtained for maps, but the application of similar methods is undoubtedly also possible in the case of flows.) It therefore seems that the assumption of general position in the hypothesis should be understood in the sense that $\mathcal{A}^k$ is a subset of the second category according to Baire.

3. If a Hamiltonian system is ergodic on isoenergy manifolds, then the static equilibrium determined by the stationary invariant measure $d\nu^\infty = \rho^\infty \, d\mu$ is microcanonical (the density ρ^∞ depends only on the total energy). Generally speaking, the measure $\bar{\nu}^\infty$ does not have this property, namely, it is not even invariant under the phase flow.

4. The constructions are similar in the case of discrete dynamical systems. Indeed, let $g\colon \Gamma \to \Gamma$ be an automorphism (or an endomorphism) of a measurable space (Γ, μ), i.e., let

$$\mu(D) = \mu(g^{-1}(D))$$

for an arbitrary μ-measurable set $D \subset \Gamma$. As usual, $g^{-1}(D)$ is the full preimage of D under the map g.

If $\rho^0\colon \Gamma \to \mathbb{R}$ is the probability density of a measure ν^0, $d\nu^0 = \rho^0 \, d\mu$, then we have the measure ν^n for integer values of $n \geq 0$,

$$d\nu^n = \rho^n \, d\mu, \qquad \rho^n = \rho^0 \circ g^n.$$

In this case,

$$\nu^n(D) = \nu^0(g^{-n}(D))$$

for every μ-measurable set $D \subset \Gamma$.

5. Stabilization of the coarse-grained density

Gibbs tried to prove that the coarse-grained entropy $\bar{s}_t$ increases with time, but his argument also proved incorrect (see [3] for the related analysis and for some other references). It is interesting that this incorrect result was first taken seriously by many authors. For example, Poincaré [4] wrote about it as if it were a well-known fact. The coarse-grained entropy should be distinguished from the Boltzmann entropy related to statistics in the μ space. Some information about the behavior of the coarse-grained entropy $\bar{\rho}^t$ as $t \to \infty$ is given by the following theorem.

Theorem 5. *Let (4.2) be a quasihomogeneous Hamiltonian system, let the initial density ρ be a function in $L_p(\Gamma, \mu)$, $1 \leq p \leq \infty$, and let $\{\Gamma_j\}$ be a partition of the phase space into parts of finite Liouville measure. Then $\lim_{t\to+\infty} \rho_j(t)$ and $\lim_{t\to-\infty} \rho_j(t)$ exist in the sense of weak L_p convergence, and they coincide.*

Theorem 5 can be easily deduced from Theorem 4. Indeed, let g^t be the phase flow in system (4.2). Then $\rho^t = \rho \circ g^t \in L_p(\Gamma, \mu)$ for all $t \in \mathbb{R}$. Let φ_j be the characteristic function of a measurable domain Γ_j. Because ρ^t is weakly convergent to ρ^∞ as $t \to \pm\infty$ and $\rho^\infty \in L_p(\Gamma, \mu)$, we have

$$\rho_j(t) = \frac{1}{\gamma_j} \int_\Gamma \rho^t \varphi_j \, d\mu \to \frac{1}{\gamma_j} \int_\Gamma \rho^\infty \varphi_j \, d\mu. \tag{5.1}$$

It remains to note that $\int_\Gamma \bar{\rho}^\infty \, d\mu = \int_\Gamma \rho^\infty \, d\mu$ and $\bar{\rho}^\infty \in L_p(\Gamma)$.

As was shown in [7], if the energy levels of a quasihomogeneous Hamiltonian system are *compact*, then ρ^∞ is the density of a probability measure,

$$\int_\Gamma \rho^\infty \, d\mu = 1.$$

Consequently, in this case, the function $\bar{\rho}^\infty$ in Theorem 5 also determines some probability measure.

We consider quasihomogeneous Hamiltonian system (4.2) on an invariant part

$$\Gamma = \{(x, y) \colon h_1 \le H(x, y) \le h_2\}, \quad h_1 < h_2. \tag{5.2}$$

Let Γ be compact. Because $C^0(\Gamma)$ is dense in $L_1(\Gamma, \mu)$ (μ is the Liouville measure $d^n x \, d^n y$), Theorems 1 and 5 imply the following corollary.

Corollary. *Let the initial density* $\rho \colon \Gamma \to \mathbb{R}$ *be a function belonging to* $L_1(\Gamma, \mu)$, *and let* $\{\Gamma_j\}$ *be a partition of the set in* (5.2). *Then as* $\sup(\operatorname{diam} \Gamma_j) \to 0$, *the density* $\bar{\rho}^\infty$ *approximates the weak limit* ρ^∞ *in the* $L_1(\Gamma, \mu)$ *metric with an arbitrary prescribed accuracy.*

Indeed, the right-hand side of limit relation (5.1) coincides with $(\bar{\rho}^\infty)_j$. Hence, the density $\bar{\rho}^\infty$ results from averaging the density ρ^∞ over the partition cells in $\Gamma = \cup \Gamma_j$.

6. A growth theorem for the coarse-grained entropy

Let ρ^∞ be the weak limit of the density ρ^t as $t \to \pm\infty$ ($\rho_0 = \rho$). Then (as was established in [7]) the inequality

$$S(\rho^\infty) \ge S(\rho) \tag{6.1}$$

holds. It turns out that the coarse-grained entropy does not always increase, contrary to the widespread opinion (first advanced by Gibbs and supported by Poincaré). This is illustrated by the following simple example.

Example 4. We consider a vertical line segment of length l in a gravity field and an ensemble of particles that are elastically reflected from the ends of the segment. If the square of the velocity of a particle exceeds $2gl$ (g is the acceleration of gravity), then the particle periodically collides with the ends of the segment. Let the initial density ρ be constant in the direct product Γ of the line segment $0 \le x \le l$ and the domain $V = \{v \colon 2gl \le v^2 \le c, \ c = \text{const}\}$ on the velocity axis. We consider the partition of Γ into two equal parts,

$$\Gamma_1 = \left\{0 \le x \le \frac{l}{2}\right\}, \qquad \Gamma_1 = \left\{\frac{l}{2} \le x \le l\right\}.$$

The initial entropy is calculated by formula (1.1) with $\lambda_1 = \lambda_2 = 1/2$ and $\gamma_j = \mu(\Gamma_1) = \mu(\Gamma_2)$. It can be easily understood that in a stationary state (in which the density ρ^t is replaced by its weak limit), the majority of the particles in the ensemble are located in the upper half of the segment (because the particles move with a lower velocity in this half). Consequently, for $\rho = \rho^{\infty}$ in formula (1.1), we already have $\lambda_1 \neq \lambda_2$. Therefore, as is known, $S(\rho^{\infty}) < S(\rho)$. It is easy to understand that the same conclusion also holds in the more general case of a partition Γ generated by the division of the line segment into n ($n \geq 2$) equal parts.

To state *sufficient* growth conditions for the coarse-grained entropy, we once again consider a quasihomogeneous system of Hamiltonian equations in (4.2) that is restricted to compact invariant domain (5.2).

Theorem 6. *Let the initial density* $\rho\colon \Gamma \to \mathbb{R}$ *be a function in* $L_1(\Gamma, \mu)$, *and let* $\{\Gamma_j\}$ *be a partition of* Γ. *If* (6.1) *is a strong inequality,* $S(\bar{\rho}^{\infty}) < \infty$, *then the inequality* $\mathbf{S}(\bar{\rho}^{\infty}) > \mathbf{S}(\bar{\rho})$ *holds for sufficiently small values of* $\sup(\operatorname{diam}\Gamma_j)$.

This assertion can be proved straightforwardly using the corollary given in the preceding section and Theorem 2 on the approximation. Namely, because $C^0(\Gamma)$ is dense in $L_1(\Gamma, \mu)$ in the case under consideration and we have $\mathbf{S}(\rho) < \mathbf{S}(\rho^{\infty}) < \infty$, the differences between $S(\bar{\rho})$ and $\mathbf{S}(\rho)$ and also between $\mathbf{S}(\bar{\rho}^{\infty})$ and $\mathbf{S}(\rho^{\infty})$ are arbitrarily small as $\sup(\operatorname{diam}\Gamma_j) \to 0$.

7. Integrable systems

We consider a Liouville-integrable Hamiltonian system with compact compatible levels of first integrals. In a domain containing no critical integral levels, such a system can be written in the "action–angle" variables,

$$\dot{x} = \omega(y), \quad \dot{y} = 0, \qquad \mathbb{T}^n = \{x = (x_1, \dots, x_n) \bmod 1\}, \quad y \in D \subset \mathbb{R}^n.$$

If the dependence of the frequencies ω on the actions y is nonsingular (i.e., $\det(\partial\omega/\partial y) \neq 0$), then ω can be taken as phase coordinates instead of y, at least locally. Then the system becomes

$$\dot{x} = \omega, \qquad \dot{\omega} = 0. \tag{7.1}$$

We note that (7.1) are quasihomogeneous Hamiltonian equations, i.e., x_j and ω_j are conjugate canonical variables and $H = (\omega_1^2 + \cdots + \omega_n^2)/2$ is the Hamilton function.

We can obtain Eqs. (7.1) proceeding from some other considerations. Let there be a collisionless continuous medium enclosed in an n-dimensional vessel in the form of a rectilinear parallelepiped. We assume that the particles are elastically reflected from walls of the vessel, i.e., from the boundary of the parallelepiped and do not collide; therefore, this medium can be called an *ideal gas*. This model of a one-dimensional ideal gas was first considered by Poincaré [4]. Of course, all this is a particular case of the general theory of Gibbs ensembles. As noted by Poincaré, after the passage to a 2^n-sheeted covering of the parallelepiped by a torus $\mathbb{T}^n$, the equations of motion coincide with system (7.1).

Equations (7.1) have an invariant measure $d\mu = dx\,d\omega$ in the phase space $\Gamma = \mathbb{T}^n \times \mathbb{R}^n$. Let $\rho(x, \omega)$ be the density of a probability measure ν, $d\nu = \rho\,d\mu$. For an arbitrary pair N, M of positive integers, we consider the partition of the phase space Γ into parts Γ_{jk}, $j \in \mathbb{Z}^n/N\mathbb{Z}^n$, $k \in \mathbb{Z}^n$,

$$\begin{aligned}
&\Gamma_{jk} = \Gamma_j^x \times \Gamma_k^\omega, \\
&\Gamma_j^x = \left\{x \in \mathbb{T}^n\colon \frac{j_l}{N} \leq x_l \leq \frac{j_l + 1}{N},\ l = 1, \dots, n\right\}, \\
&\Gamma_k^\omega = \left\{\omega \in \mathbb{R}^n\colon \frac{k_l}{M} \leq \omega_l \leq \frac{k_l + 1}{M},\ l = 1, \dots, n\right\}.
\end{aligned}$$

The measure μ of every partition part Γ_{jk} is equal to $\mu(\Gamma_{jk}) = (NM)^{-n}$.

We calculate the value of the coarse-grained density on Γ_{jk},

$$\rho_{jk}(t) = (NM)^n \int_{\Gamma_{jk}} \rho(x + \omega t, \omega)\, dx\, d\omega.$$

We set

$$\langle\rho\rangle(\omega) = \int_{\mathbb{T}^n} \rho(x, \omega)\, dx, \qquad \langle\rho\rangle_k = \int_{\Gamma_k^\omega} \langle\rho\rangle(\omega)\, d\omega.$$

It is clear that $\langle\rho\rangle$ is a density of some probability measure on Γ, and we have $\langle\rho\rangle_{jk}(t) = \langle\rho\rangle_k$ for an arbitrary value of t.

Theorem 7. *Let ρ be a bounded function on Γ, and let it be Lipschitzian with respect to the variables ω. Then*

$$|\rho_{jk}(t) - \langle\rho\rangle_k| \le \frac{nMN^n}{t}\left(\frac{\|\rho\|_1}{M} + 2\|\rho\|_0\right),$$

$$\|\rho\|_0 = \sup_\Gamma |\rho|, \qquad \|\rho\|_1 = \sup \frac{|\rho(x,\omega') - \rho(x,\omega'')|}{|\omega' - \omega''|}$$

for $t > 0$, where $(x,\omega'), (x,\omega'') \in \Gamma$, $\omega' \ne \omega''$.

Theorem 7 is proved in Sec. A.6 in the appendix.

Let ρ be a compactly supported function and let $\bar\rho$ and $\overline{\langle\rho\rangle}$ be the coarse-grained densities corresponding to the partition Γ_{jk} and to the respective densities ρ and $\langle\rho\rangle$. Then

$$\mathbf{S}(\bar\rho) - \mathbf{S}\big(\overline{\langle\rho\rangle}\big) = -\frac{1}{(MN)^n} \sum_{j,k} (\rho_{jk} \log \rho_{jk} - \langle\rho\rangle_k \log\langle\rho\rangle_k),$$

where only finitely many terms under the summation sign are nonzero. It seems that in a typical situation, it should be expected that the difference $\mathbf{S}(\bar\rho) - \mathbf{S}\big(\overline{\langle\rho\rangle}\big)$ is of the order of $1/t$ as $t \to \infty$, although it is easy to construct some examples in which $\mathbf{S}(\bar\rho) - \mathbf{S}\big(\overline{\langle\rho\rangle}\big) \sim t^{-1}\log t$.

8. Mixing systems

The dynamics in systems of the type considered in Sec. 7 are commonly said to be regular; the antipode is chaotic dynamics. Here, the mixing systems are primarily meant. We recall the related definition.

Definition 2. Let a flow g^t on a phase space Γ preserve a probability measure μ. The flow g^t is called a mixing flow if

$$\lim_{t\to\infty} \int_\Gamma \varphi \circ g^t\, \psi\, d\mu = \int_\Gamma \varphi\, d\mu \int_\Gamma \psi\, d\mu \tag{8.1}$$

for an arbitrary pair of functions $\varphi, \psi\colon \Gamma \to \mathbb{R}$ that belong to a sufficiently extensive function space. In the case of hyperbolic systems (Anosov systems), the correlations decrease exponentially, i.e.,

$$\left|\lim_{t\to\infty} \int_\Gamma \varphi \circ g^t\, \psi\, d\mu - \int_\Gamma \varphi\, d\mu \int_\Gamma \psi\, d\mu\right| < C\tau^{|t|} \tag{8.2}$$

for some constants $C > 0$ and $\tau \in (0,1)$.

In this case, the coarse-grained density always tends to unity as $t \to \infty$. Indeed, we have

$$\rho_j(t) = \frac{1}{\gamma_j} \int_{\Gamma_j} \varphi \circ g^t \, d\mu = \frac{1}{\gamma_j} \int_{\Gamma} \varphi \circ g^t \, \chi_{\Gamma_j} \, d\mu,$$

where χ_{Γ_j} is the characteristic function of the set Γ_j. It follows from (8.1) that

$$\rho_j \to \frac{1}{\gamma_j} \int_{\Gamma} \varphi \, d\mu \int_{\Gamma} \chi_{\Gamma_j} \, d\mu = 1.$$

Moreover, if condition (8.2) holds, then ρ_j tends to unity at an exponential rate. The coarse-grained entropy should be expected to have the same behavior, i.e., to have the property that $\mathbf{S}(\bar{\rho}^t)$ tends rapidly to $\mathbf{S}(1) = 0$ at an exponential rate.

Nevertheless, we note that these assertions should be relied on somewhat cautiously because the function spaces for which relations (8.1) and (8.2) can be proved are to some extent narrower than the space of continuous functions on Γ, and they do not contain the characteristic function of the sets Γ_j. The main reason for what has been said seems to be the property that the functions χ_{Γ_j} are difficult to include in a suitable function space (from the standpoint of verifying (8.1) and (8.2)) rather than that they are too "bad." Incidentally, in concrete examples (for instance, for linear hyperbolic automorphisms of a torus), the assertion about the behavior of the coarse-grained density and of the coarse-grained entropy presented in this section can be verified straightforwardly.

Appendix

A.1. Proof of Theorem 1. Let $\varepsilon > 0$ be an arbitrary number. Because $C^0(\Gamma)$ is dense in $L_1(\Gamma, \mu)$, there is a function $\rho \in C^0(\Gamma)$ such that

$$\|\rho - \rho_c\| < \varepsilon,$$

where $\|\cdot\|$ denotes the L_1 norm. A simple inequality of the form

$$\|\bar{\rho} - \bar{\rho}_c\| \le \|\rho - \rho_c\| < \varepsilon \tag{A.1}$$

also holds. The function ρ_c is uniformly continuous on Γ. Therefore, there exists $\delta > 0$ such that

$$|\rho_c(z_1) - \rho_c(z_2)| < \varepsilon$$

for arbitrary $z_1, z_2 \in \Gamma$, $\operatorname{dist}(z_1, z_2) < \delta$. Hence, $|\bar{\rho}_c - \rho_c| < \varepsilon$, whence it follows that

$$\|\bar{\rho}_c - \rho_c\| < \varepsilon. \tag{A.2}$$

The combination of formulas (A.1) and (A.2) gives $\|\rho - \bar{\rho}\| < 3\varepsilon$.

A.2. Proof of Theorem 2, Part 1. In this section, we deduce Theorem 2 from Lemma 1. We set

$$\rho_0(x) = \begin{cases} \rho(x) & \text{for } \rho(x) \le \Delta, \\ \Delta & \text{for } \rho(x) > \Delta, \end{cases}$$

$$\rho_{0j} := \bar{\rho}_0|_{\Gamma_j} = \frac{1}{\gamma_j} \int_{\Gamma_j} \rho_0 \, d\mu, \qquad s_0 = \mathbf{S}(\rho_0), \qquad \bar{s}_0 = \mathbf{S}(\bar{\rho}_0).$$

It follows from formula (A.2) that $s_0 \le \bar{s}_0$.

We fix an arbitrary $\varepsilon > 0$. If Δ is sufficiently large, then

$$0 \le s - s_0 \le \varepsilon. \tag{A.3}$$

If the diameter of the partition $\{\Gamma_j\}$ is sufficiently small, then according to Lemma 1, the inequalities

$$0 < \bar{s}_0 - s_0 \le \varepsilon \tag{A.4}$$

hold. Thus, it suffices to prove that

$$\bar{s} < \bar{s}_0 + c\,\varepsilon|\log\varepsilon|, \tag{A.5}$$

where $c > 0$ is a constant that is independent of ε. Indeed, according to formulas (A.3) and (A.4), we have $|s - \bar{s}_0| < 2\varepsilon$ in this case. Consequently, formulas (A.2) and (A.5) imply that

$$s \le \bar{s} \le \bar{s}_0 + c\varepsilon\log\varepsilon \le s + c\,\varepsilon\log\varepsilon + 2\varepsilon,$$

whence it follows that $|s - \bar{s}| < c\varepsilon\log\varepsilon + 2\varepsilon$.

It remains to prove estimate (A.5). We write formula (A.3) in greater detail,

$$0 < \int_{\Gamma\cap\{\rho\ge\Delta\}} (\rho\log\rho - \rho_0\log\rho_0)\,d\mu < \varepsilon.$$

It follows that

$$0 < \int_{\Gamma\cap\{\rho\ge\Delta\}} (\rho\log\Delta - \rho_0\log\Delta)\,d\mu < \varepsilon,$$

whence

$$0 < \int_{\Gamma} (\rho - \rho_0)\,d\mu < \frac{\varepsilon}{\log\Delta}. \tag{A.6}$$

We have the inequalities

$$0 \le \rho_{0j} \le \rho_j, \qquad \sum_j (\rho_j - \rho_{0j})\gamma_j < \frac{\varepsilon}{\log\Delta}. \tag{A.7}$$

We prove an auxiliary assertion: for arbitrary a and σ, $0 \le a \le b$, $\sigma \in (0, 1/e)$, the inequality

$$a\log a \le b\log b + |\sigma\log\sigma| + |1 + \log\sigma|\,(b - a) \tag{A.8}$$

holds. Indeed, if $a > \sigma$, then using the obvious inequality $\min_{\rho\ge\sigma}(\rho\log\rho)' = 1 + \log\sigma$, we obtain

$$a\log a - b\log b \le |1 + \log\sigma|(b - a).$$

And if $a \in (0, \sigma)$, then

$$a\log a - b\log b \le |a\log a - \sigma\log\sigma| + |\sigma\log\sigma - b\log b| \le |\sigma\log\sigma| + |1 + \log\sigma|(b - a).$$

Inequality (A.8) immediately implies

$$\sum_j \gamma_j\rho_{0j}\log\rho_{0j} \le \sum_j \gamma_j\rho_j\log\rho_j + |\sigma\log\sigma|\sum_j \gamma_j + \sum_j (|1 + \log\sigma|)(\rho_j - \rho_{0j})\gamma_j,$$

which, in view of formula (A.7), can be rewritten in the form

$$\bar{s} \le \bar{s}_0 + |\sigma\log\sigma| + \frac{|1 + \log\sigma|}{\log\Delta}\varepsilon.$$

It now suffices to set $\sigma = \varepsilon$.

A.3. Proof of Theorem 2, Part 2. In this section, we deduce Lemma 1 from Lemma 2. Let $\rho < \Delta$. We set

$$\rho_*(x) = \begin{cases} \rho(x) & \text{for } \rho(x) \geq \delta, \\ \delta & \text{for } \rho(x) < \delta, \end{cases}$$

$$\rho_{*j} := \bar{\rho}_*|_{\Gamma_j} = \frac{1}{\gamma_j} \int_{\Gamma_j} \rho_0 \, d\mu, \qquad s_* = \mathbf{S}(\rho_*), \qquad \bar{s}_* = \mathbf{S}(\bar{\rho}_*).$$

We have

$$0 \leq s - s_* \leq \delta \log \delta. \tag{A.9}$$

If the diameter of the partition $\{\Gamma_j\}$ is sufficiently small, then according to Lemma 2, the inequalities

$$0 < \bar{s}_* - s_* \leq \delta \tag{A.10}$$

hold. Hence, it suffices to show that

$$\bar{s} < \bar{s}_* + c\delta, \tag{A.11}$$

where $c > 0$ is a constant that is independent of δ. Certainly, in this case, in view of formulas (A.9) and (A.10), we have $|s - \bar{s}_*| < \delta + \delta \log \delta$. Consequently, it follows from formulas (A.2) and (A.11) that $s \leq \bar{s} \leq \bar{s}_* + c\delta \leq s + (1 + c)\delta + \delta \log \delta$, whence we obtain $|s - \bar{s}| < (1 + c)\delta + \delta \log \delta$.

We verify estimate (A.11). According to the definition of the function ρ_*, we have

$$\delta \leq \rho_* \leq \rho + \delta \leq \Delta + \delta.$$

Therefore, $\delta \leq \bar{\rho}_* \leq \bar{\rho} + \delta \leq \Delta + \delta$. It follows from the inequality

$$\sup_{\rho \in (0, \Delta+\delta)} (\rho \log \rho)' = 1 + \log(\Delta + \delta)$$

that $\bar{\rho}_* \log \bar{\rho}_* \leq \bar{\rho} \log \bar{\rho} + (1 + \log(\Delta + \delta))\delta$, whence we conclude that

$$\bar{s} \leq \bar{s}_* + (1 + \log(\Delta + 1))\delta.$$

A.4. Proof of Theorem 2, Part 3. In this section, we prove Lemma 2. We fix an arbitrary $\varepsilon > 0$. Because $C^0(\Gamma)$ is dense in $L_1(\Gamma, \mu)$, there is a function $\rho_c \in C^0(\Gamma)$ such that

$$\delta < \rho_c < \Delta, \qquad \rho = \rho_c(1 + \rho_l), \qquad \|\rho_l\| < \varepsilon, \tag{A.12}$$

where $\|\cdot\|$ denotes the L_1 norm. It follows that

$$s = -\int_\Gamma \rho_c(1 + \rho_l) \log(\rho_c(1 + \rho_l)) \, d\mu = s_c + A_1 + A_2,$$

$$s_c := \mathbf{S}(\rho_c), \qquad A_1 = -\int_\Gamma \rho \log(1 + \rho_l) \, d\mu, \qquad A_2 = -\int_\Gamma \rho_c \rho_l \log \rho_c \, d\mu.$$

Hence, it remains to verify that $|\bar{s} - s_c|$, A_1, and A_2 are small.

Because $|\rho_c \log \rho_c| \le \Delta \log \Delta$, we have

$$|A_2| \le \Delta \log \Delta \int_\Gamma |\rho_l| \, d\mu < \varepsilon \Delta \log \Delta. \tag{A.13}$$

According to (A.12), $\rho_l = (\rho - \rho_c)/\rho_c \in I$, where $I = [-1+\delta/\Delta,\, 1+\Delta/\delta]$. It can be easily established that

$$\max_{\rho \in I} \left| \frac{\log(1+\rho)}{\rho} \right| \le \log \frac{\Delta}{\delta}.$$

Therefore,

$$|A_1| \le \Delta \int_\Gamma |\log(1+\rho_l)| \, d\mu \le \Delta \log \frac{\Delta}{\delta} \int_\Gamma |\rho_l| \, d\mu \le \varepsilon \Delta \log \frac{\Delta}{\delta}. \tag{A.14}$$

The function ρ_c is continuous on the compact set Γ and is consequently uniformly continuous, i.e., there exists $\sigma > 0$ such that $|\rho_c(z_1) - \rho_c(z_2)| < \varepsilon$ for all $z_1, z_2 \in \Gamma$ such that $\operatorname{dist}(z_1, z_2) < \sigma$.

Let $\operatorname{diam} \Gamma_j < \sigma$. Then $|\bar{\rho}_c - \rho_c| < \varepsilon$. We have the estimate

$$|\bar{s}_c - s_c| = \left| \int_\Gamma (\rho_c \log \rho_c - \bar{\rho}_c \log \bar{\rho}_c) \, d\mu \right| \le$$
$$\le \left(1 + \log \frac{\Delta}{\delta}\right) \int_\Gamma |\rho_c - \bar{\rho}_c| \, d\mu \le \left(1 + \log \frac{\Delta}{\delta}\right) \varepsilon. \tag{A.15}$$

We set $\rho_{cj} = \bar{\rho}_c|_{\Gamma_j}$. Then

$$|\rho_j - \rho_{cj}| = \left| \frac{1}{\gamma_j} \int_{\Gamma_j} \rho_c \rho_l \, d\mu \right| \le \frac{\Delta}{\gamma_j} r_j,$$

$$r_j = \int_{\Gamma_j} |\rho_l| \, d\mu, \qquad \sum_{j \in J} r_j < \varepsilon.$$

Consequently, for every $j \in J$, we have

$$|\rho_j \log \rho_j - \rho_{cj} \log \rho_{cj}| \le \left(1 + \log \frac{\Delta}{\delta}\right) |\rho_j - \rho_{cj}| \le \frac{\Delta}{\gamma_j} \left(1 + \log \frac{\Delta}{\delta}\right) r_j.$$

We thus obtain the estimate for the difference of entropies

$$|\bar{s}_c - \bar{s}| = \left| \sum_{j \in J} (\rho_j \log \rho_j - \rho_{cj} \log \rho_{cj}) \gamma_j \right| \le \Delta \left(1 + \log \frac{\Delta}{\delta}\right) \varepsilon. \tag{A.16}$$

From formulas (A.15) and (A.16), we derive

$$|\bar{s} - s_c| \le (1 + \Delta) \left(1 + \log \frac{\Delta}{\delta}\right) \varepsilon.$$

A.5. Proof of Theorem 3. Obviously, it can be assumed that $d < 1$. Let $\varepsilon > 0$ be arbitrary. We set

$$\hat{J}_l = \{j \in J \colon \Gamma_j \subset K_l\}, \qquad \widehat{K}_l = \bigcup_{j \in \hat{J}_l} \Gamma_j.$$

Then the inequalities

$$\left|s - \int_{\widehat{K}_N} \rho \log \rho \, d\mu\right| < \varepsilon \tag{A.17}$$

$$\left|\bar{s} - \int_{\widehat{K}_N} \bar{\rho} \log \bar{\rho} \, d\mu\right| < \varepsilon \tag{A.18}$$

hold for sufficiently large N. Indeed, we verify inequality (A.17),

$$\left|s - \int_{\widehat{K}_N} \rho \log \rho \, d\mu\right| \le A_N + A_{N+1} + \dots,$$

$$A_l = \int_{\widehat{K}_{l+1} \setminus \widehat{K}_l} |\rho \log \rho| \, d\mu.$$

For every $l > 1$, the inclusions $K_{l-1} \subset \widehat{K}_l \subset K_l$ hold (the first follows from Definition 1d and from the inequality $d < 1$, and the second follows from the definition of the set K_l). Thus,

$$\widehat{K}_{l+1} \setminus \widehat{K}_l \subset K_{l+1} \setminus K_{l-1}.$$

Consequently, by Definition 1c, we have

$$\mu(\widehat{K}_{l+1} \setminus \widehat{K}_l) \le \mu(K_{l+1} \setminus K_l) + \mu(K_l \setminus K_{l-1}) \le 2Cl^{n-1}.$$

It follows from assumption 3 in the theorem that

$$\rho|_{\widehat{K}_{l+1} \setminus \widehat{K}_l} \le \rho|_{K_{l+1}} < c_\rho (l+1)^{-n-\delta}.$$

Because we can assume that $c_\rho (N+1)^{-n-\delta} < 1/e$, we have

$$A_l \le 2Cl^{n-1} c_\rho (l+1)^{-n-\delta} |\log(c_\rho (l+1)^{-n-\delta})| \tag{A.19}$$

for all $l \ge N$. For sufficiently large values of N, inequality (A.17) follows from estimate (A.19). Inequality (A.18) is proved similarly.

According to Definition 1c, $\mu(\widehat{K}_N) < \infty$. Therefore, it follows from Theorem 1 that

$$\left|\int_{\widehat{K}_N} \rho \log \rho \, d\mu - \int_{\widehat{K}_N} \bar{\rho} \log \bar{\rho} \, d\mu\right| < \varepsilon \tag{A.20}$$

for sufficiently small values of $d > 0$. Inequalities (A.17), (A.18), and (A.20) imply that $|s - \bar{s}| < 3\varepsilon$ for sufficiently small $d > 0$, which was to be proved.

A.6. Proof of Theorem 7. We note that $\rho \ge 0$ is not used anywhere in the proof. Therefore, replacing ρ by $\rho - \langle \rho \rangle$, we see that we can confine ourself to considering the case $\langle \rho \rangle = 0$.

We consider the functions

$$\rho_k(t, x) = M^n \int_{\Gamma_k^\omega} \rho(x + \omega t, \omega) \, d\omega = \frac{M^n}{t^n} \int_{t\Gamma_k^\omega} \rho\left(x + \beta, \frac{\beta}{t}\right) d\beta$$

for large values of t, where

$$t\Gamma_k^\omega = \left\{\beta \in \mathbb{R}^n \colon \frac{tk_l}{M} \le \beta_l \le \frac{t(k_l + 1)}{M}, \ l = 1, \dots, n\right\}.$$

Lemma A.1. *If* $\langle\rho\rangle = 0$, *then*

$$|\rho_k(t,x)| \le \frac{nM}{t}\biggl(\frac{\|\rho\|_1}{M} + 2\|\rho\|_0\biggr).$$

Theorem 7 follows straightforwardly from Lemma A.1 because

$$|\rho_{jk}(t)| = \biggl|\int_{\mathbb{T}^n} \rho_k(t,x)\,dx\biggr|$$

for $\langle\rho\rangle = 0$.

Proof of Lemma A.1**.** We represent the set $t\Gamma_k^\omega$ as a union of unit cubes $\mathcal{C}_m$, $m \in \mathcal{Z}$, and a remainder R. Here, $\mathcal{Z} = \mathcal{Z}(t)$ is a finite subset in $\mathbb{Z}^n$,

$$\mathcal{C}_m = \{\beta \in \mathbb{R}^n \colon m_l \le \beta_l \le m_l + 1,\ l = 1,\dots,n\},$$

and $R = t\Gamma_k^\omega \setminus \bigcup_{m\in\mathcal{Z}} \mathcal{C}_m$. We assume that R does not contain any of the cubes $\mathcal{C}_m$ entirely. Then

$$\int_R d\beta \le \frac{2nt^{n-1}}{M^{n-1}}, \qquad \#\mathcal{Z} \le \frac{t^n}{M^n}. \tag{A.21}$$

We have

$$\begin{aligned} \rho_k &= \frac{1}{M^n t^n}\biggl(\sum_{m\in\mathcal{Z}} I_m + I_R\biggr), \\ I_m &= \int_{\mathcal{C}_m} \rho\Bigl(x+\beta, \frac{\beta}{t}\Bigr)\,d\beta, \qquad I_R = \int_R \rho\Bigl(x+\beta, \frac{\beta}{t}\Bigr)\,d\beta. \end{aligned} \tag{A.22}$$

Because $\langle\rho\rangle = 0$, the inequality

$$\int_{\mathcal{C}_m} \rho\Bigl(x+\beta, \frac{\beta}{t}\Bigr)\,d\beta = 0$$

holds for every $\beta_0 \in \mathbb{R}^n$. Therefore,

$$|I_m| = \biggl|\int_{\mathcal{C}_m} \Bigl(\rho\Bigl(x+\beta, \frac{\beta}{t}\Bigr) - \rho\Bigl(x+\beta, \frac{m}{t}\Bigr)\Bigr)\,d\beta\biggr| \le \int_{\mathcal{C}_m} \frac{n\|\rho\|_1}{t}\,d\beta = \frac{n\|\rho\|_1}{t}.$$

On the other hand, $|R| \le \int_R \|\rho\|_0\,d\beta$. Consequently, using formulas (A.21) and (A.22), we obtain

$$|\rho_k(t,x)| \le \frac{M^n}{t^n}\biggl(n\|\rho\|_1\frac{t^{n-1}}{M^n} + 2n\|\rho\|_0\frac{t^{n-1}}{M^{n-1}}\biggr) = \frac{nM}{t}\biggl(\frac{\|\rho\|_1}{M} + 2\|\rho\|_0\biggr).$$

The lemma is proved.

Acknowledgments. The authors are grateful to V. V. Sidorenko and O. G. Smolyanov for the valuable discussion.

This work was supported by the Russian Foundation for Basic Research (Grant Nos. 05-01-01058 and 05-01-01119), the Program for Supporting Leading Scientific Schools (Grant No. NSh-1312.2006.1), and the program of the Presidium of the Russian Academy of Sciences "Nonlinear Dynamics."

REFERENCES

1. R. Bowen, *Methods in Symbolic Dynamics* [Russian transl. of selected works], Mir, Moscow (1979).
2. M. Kac, *Probability and Related Topics in Physical Sciences* (Lect. Appl. Math., Vol. 1), Interscience, London (1959).
3. N. S. Krylov, *Works on the Foundations of Statistical Physics* [in Russian], Izdat. Akad. Nauk SSSR, Moscow (1950); English transl., Princeton Univ. Press, Princeton, N. J. (1979).
4. H. Poincaré, *Journal de Physique Théorique et Appliquée, 4 Série*, **5**, 369–403 (1906).
5. J. W. Gibbs, *Thermodynamics*: *Statistical Mechanics* [Russian transl. of selected works], Nauka, Moscow (1982).
6. V. V. Kozlov and D. V. Treshchev, *Theor. Math. Phys.*, **136**, 1325–1335 (2003).
7. V. V. Kozlov and D. V. Treshchev, *Theor. Math. Phys.*, **134**, 339–350 (2003).
8. S. V. Goncharenko, D. V. Turaev, and L. P. Shil'nikov, *Dokl. Rossiiskoi Akad. Nauk*, **407**, 299–303 (2006).