Theoretical and Mathematical Physics, **134**(3): 339–350 (2003)

WEAK CONVERGENCE OF SOLUTIONS OF THE LIOUVILLE EQUATION FOR NONLINEAR HAMILTONIAN SYSTEMS

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

We suggest sufficient conditions for the existence of weak limits of solutions of the Liouville equation as time increases indefinitely. The presence of the weak limit of the probability distribution density leads to a new interpretation of the second law of thermodynamics for entropy increase.

Keywords: Hamiltonian system, Liouville equation, weak convergence, entropy

1. Introduction

Let $\Gamma = T^*M$ be the phase space of an autonomous Hamiltonian system, let $M = \{x_1, \dots, x_n\}$ be the configuration space, let $H(x, y)$ be the Hamiltonian function, and let $y = (y_1, \dots, y_n)$ be the canonical momenta.

Following Gibbs [1], we introduce a probability measure μ ($\mu(\Gamma) = 1$) in the phase space Γ at the initial instant $t = 0$. It is assumed that the measure μ has a density $\rho(x, y)$. It is clear that ρ is a function of class $L_1(\Gamma)$. This measure is transferred by the phase flow g^t of the Hamiltonian system. Therefore, its density ρ_t depends on time and satisfies the *Liouville equation*, i.e., if $\rho_t \in C^1(\Gamma)$, then

$$\frac{\partial \rho_t}{\partial t} + \sum_{i=1}^{n} \left[\frac{\partial}{\partial x_i} \left(\rho_t \frac{\partial H}{\partial y_i} \right) - \frac{\partial}{\partial y_i} \left(\rho_t \frac{\partial H}{\partial x_i} \right) \right] = 0.$$

The function $\rho_0 = \rho$ serves as the Cauchy datum.

Liouville equations play a key role throughout statistical mechanics. In this connection, it suffices to mention the chain of Bogoliubov equations [2] (see, e.g., [3] for a survey of Bogoliubov's ideas). But some difficult questions of a fundamental character arise here. First, how should the initial probability distribution density ρ_0 be chosen? Second, does the function ρ_t have a limit as $t \to \pm\infty$?

See [2] and [3] for a detailed discussion of choosing ρ_0. Regarding the second question, Gibbs already tried to show that as $t \to \infty$, the density ρ_t tends (in some sense) to the density of the stationary distribution corresponding to the thermal equilibrium. For this, he introduced a microcanonical probability distribution whose density depends only on the total energy H. But, as a rule, according to the Poincaré recursion theorem, ρ_t has no limit at all in the usual sense as $t \to \pm\infty$.

We study the question of *weak convergence* for solutions of the Liouville equation. This approach is natural from the standpoint of justifying thermodynamics in the sense of the transition to a macroscopic description of the evolution of a dynamical system because the probability measure density manifests itself in calculating means (mathematical expectations) of dynamical characteristics rather than "exists" in itself. This standpoint, which was neither understood nor used by specialists in statistical mechanics, was already used by Poincaré in [4] (see Sec. 6 for more detail on this question).

*Moscow State University, Moscow, Russia.

Translated from Teoreticheskaya i Matematicheskaya Fizika, Vol. 134, No. 3, pp. 388–400, March, 2003. Original article submitted July 5, 2002.

We confine ourselves to a narrower class of initial distributions and assume that ρ is a square-integrable function (i.e., ρ belongs to $L_2(\Gamma)$). This assumption is natural from the standpoint of the possibility of calculating means of functions on the phase space. Let z be a point in the phase space Γ (it is determined by a set of canonical variables x, y), and let $\{g^t\}$ be the phase flow of the Hamiltonian system with the Hamiltonian H. We assume that all solutions of this Hamiltonian system can be extended throughout the time axis $\mathbb{R} = \{t\}$. For example, it can be assumed that all energy manifolds $\{H = \mathrm{const}\} \subset \Gamma$ are compact. In this case, the transformations g^t are defined for all $t \in \mathbb{R}$.

Because the transformations g^t preserve the phase volume in Γ (according to the Liouville theorem), ρ_t is a first integral of the Hamilton equations with Hamiltonian H. This simple fact allows writing the general form of solutions of the Liouville equation. Let z_0 be the initial state of the system. Then $t \mapsto g^t(z_0)$ is a solution of the Hamilton equations. Every first integral is a function of the initial data $\rho(z_0)$. Because $z_0 = g^{-t}(z)$, we have

$$\rho_t(z) = \rho\bigl(g^{-t}(z)\bigr).$$

As is well known, the transfer of functions belonging to L_2 by the phase flow of a dynamical system with an invariant measure is equivalent to the action of a one-parameter group of unitary operators U^t (see, e.g., [5]), $U^t\rho(z) = \rho\bigl(g^t(z)\bigr)$, whence $\rho_t = U^{-t}\rho$. The operator U is often called the Koopman operator.

Let φ be another function in $L_2(\Gamma)$. Then the time function

$$K(t) = (U^{-t}\rho, \varphi) = \int_\Gamma \varphi U^{-t}\rho \, d^{2n}z$$

is well defined. It has a simple meaning, namely, if φ is a characteristic function of a measurable domain Φ in Γ, then $K(t)$ is the fraction of Hamiltonian systems belonging to the Gibbs ensemble that are in the domain Φ at time t. If

$$\lim_{t\to\infty} K(t) = (\bar{\rho}, \varphi)$$

for each function $\varphi \in L_2$, then ρ_t is *weakly convergent* to $\bar{\rho}$.

Our objective is to establish conditions for weak convergence and a method for calculating weak limits of probability measure densities.

This paper develops and complements the results in [6]. In Sec. 2, a formula is derived for the weak limit $\bar{\rho}$ of the probability distribution density ρ_t on the condition that this limit exists. In Sec. 3, the main investigation object, a class of systems (of form (3.1)), is singled out. We present the necessary motivations and examples and also discuss the main technical apparatus for further analysis, namely, a generalization of the von Neumann statistical ergodic theorem with the averaging over time replaced with averaging over some probability measure. Section 4 describes a class of systems (with the so-called layered flows) for which we manage to prove the weak convergence of solutions of the Liouville equation as time increases indefinitely. Typical representatives of this class of dynamical systems are geodesic flows and some other Hamiltonian quasihomogeneous systems. The main results are proved in Sec. 5. Concluding the paper in Sec. 6, we prove that the entropy of the limit measure $\bar{\rho}\, d^{2n}z$ is not less than that of the original measure $\rho\, d^{2n}z$ and also discuss the relation between this result and the second law of thermodynamics.

2. Weak limit

Theorem 1. *Let the limit*

$$\lim_{t\to\infty} K(t) = K_\infty$$

exist for a function $\varphi \in L_2(\Gamma)$. Then $K_\infty = (\bar{\rho}, \varphi)$, where

$$\bar{\rho}(z) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \rho\big(g^t(z)\big)\,dt. \tag{1}$$

Formula (1) holds to within a set of points $z \in \Gamma$ of measure zero (this is a generally accepted approach in measure theory). Because $\rho \in L_1$, it follows (by the Birkhoff–Khinchin theorem) that the function $\bar{\rho}$ is defined almost everywhere, is nonnegative, and serves as an integral of the Hamilton equations (it is invariant with respect to g^t), and (if the energy surfaces $H = \text{const}$ are compact) the relation

$$\int_\Gamma \bar{\rho}\, d^{2n}z = 1$$

holds. Consequently, $\bar{\rho}$ is the density of a stationary probability measure.

If the limit $\lim K(t)$ exists for all $\varphi \in L_2$, then the function $\bar{\rho}$ satisfying the relation $K_\infty = (\bar{\rho}, \varphi)$ is unique. This implies the following corollary.

Corollary 1. *If ρ_t is weakly convergent to $\bar{\rho}$, then $\bar{\rho}$ is defined by formula* (1).

Proof of Theorem 1. Let $(\rho_t, \varphi) \to K_\infty$ as $t \to \infty$. Then (by the Cauchy theorem) the limit relation

$$\frac{1}{T}\int_0^T (\rho_t, \varphi)\,dt \to K_\infty \quad \text{as } T \to \infty$$

holds. We note that the integral in this relation exists for all T if ρ and φ belong to L_2 (see, e.g., [7]). By the Fubini theorem,

$$\frac{1}{T}\int_0^T (\rho_t, \varphi)\,dt = \int_\Gamma \tilde{\rho}_T(z)\varphi(z)\,d^{2n}z,$$

where

$$\tilde{\rho}_T = \frac{1}{T}\int_0^T \rho\big(g^t(z)\big)\,dt.$$

Furthermore, by the von Neumann theorem, we have

$$\int_\Gamma (\tilde{\rho} - \bar{\rho})\,d^{2n}z \to 0 \quad \text{as } T \to \infty.$$

It follows that

$$\frac{1}{T}\int_0^T (\rho_t, \varphi)\,dt \to (\bar{\rho}, \varphi) \quad \text{as } T \to \infty.$$

Indeed, in view of the von Neumann theorem,

$$\left[\int_\Gamma (\tilde{\rho}_T - \bar{\rho})\varphi\,d^{2n}z\right]^2 \le \int_\Gamma (\tilde{\rho}_T - \bar{\rho})^2\,d^{2n}z \int_\Gamma \varphi^2\,d^{2n}z \to 0 \quad \text{as } T \to \infty,$$

which implies what we had to prove.

Theorem 1 is a priori in the sense that if the weak limit of the probability density exists, then it serves as the well-known object of the ergodic theorem.

The question under consideration now reduces to finding conditions guaranteeing the existence of the weak limit of the density ρ_t as $t \to \pm\infty$. We note that weak convergence does not always occur. For example, *linear* Hamiltonian systems are an exception here.

Remark 1. The Cesàro mean in (1) can be replaced by a more general mean. For instance, we can set

$$\bar{\rho}(z) = \lim_{T\to\infty} \left(\int_0^T \alpha(t)\rho\big(g^t(z)\big)\,dt \Big/ \int_0^T \alpha(t)\,dt \right), \tag{2}$$

where $\alpha(t) > 0$ and the integral

$$\int_0^\infty \alpha(t)\,dt$$

is divergent. If limit (2) exists, then it coincides with (1) [8].

3. Generalization of the statistical ergodic theorem

The question of the existence of the weak limit of the density ρ_t is considered for dynamical systems determined by differential equations of the form

$$\dot{z} = v(z,\omega), \qquad \dot{\omega} = 0. \tag{3}$$

The phase space Γ is a direct product $\Lambda \times D$, where $\Lambda = \{z_1, \dots, z_n\}$ is a smooth n-dimensional manifold and D is a domain in $\mathbb{R}^m = \{\omega_1, \dots, \omega_m\}$. The coordinates ω serve as first integrals. We assume that the system on Λ has an invariant measure $d\nu = \lambda(z,\omega)\,d^{2n}z$ for fixed values of ω,

$$\sum \frac{\partial(v_i\lambda)}{\partial z_i} = 0.$$

In particular, Hamiltonian systems have this form. Here, $m = 1$ and Λ is an energy surface. The role of the coordinate ω is played by the total energy. The phase space Γ of the Hamiltonian system is split into cells $h_1 \le H \le h_2$ such that the interval (h_1, h_2) contains no critical values of the Hamiltonian function H.

We begin by considering a special case of Eq. (3) in which the field $v(z,\omega)$ has the form of a product $\omega v(z)$, where $z \mapsto v(z)$ is a smooth vector field on the manifold Λ. The phase flow of this system is the family of transformations $\{g^{\omega t}\}$, where $\{g^t\}$ is the flow of the dynamical system

$$\dot{z} = v(z), \quad z \in \Lambda.$$

Its invariant measure ν is independent of the parameter ω.

Theorem 2. *Let f_1 and f_2 be some functions belonging to $L_2(\Lambda,\nu)$, let $\nu(\Lambda) < \infty$, and let h be an integrable function on a measurable set $I \subset \mathbb{R} = \{\omega\}$ (i.e., $h \in L_1(I, d\omega)$). Then*

$$\lim_{t\to\infty} \int_I h(\omega)(U^{\omega t} f_1, f_2)\,d\omega = (\bar{f}_1, f_2) \int_I h(\omega)\,d\omega. \tag{4}$$

The classical von Neumann theorem [7] is a special case of Theorem 2. Indeed, if h is the characteristic function of the interval $[0,1]$, then

$$\int_{-\infty}^{+\infty} h(\omega)\varphi(\omega t)\,d\omega = \frac{1}{t}\int_0^t \varphi(s)\,ds$$

for an arbitrary integrable function $\varphi\colon \mathbb{R} \to \mathbb{R}$.

Formula (4) seems particularly simple in the case of an ergodic flow. If h is a probability measure density on $\mathbb{R} = \{\omega\}$, then

$$\lim_{t\to\infty} \int_{-\infty}^{+\infty} \int_\Lambda h(\omega) f_1\bigl(g^{\omega t}(x)\bigr) f_2(x)\, d\nu\, d\omega = \nu(\Lambda) \int_\Lambda f_1\, d\nu \int_\Lambda f_2\, d\nu. \tag{5}$$

Thus, the functons $f_1\bigl(g^{\omega t}(x)\bigr)$ and $f_2(x)$ are statistically independent in the mean for large values of t, namely, the integral of their product is equal to the product of the integrals of the functions. Some special cases of formula (5) were previously given in [6].

We mention another interesting consequence of formula (5), namely, if the variance σ of the normal distribution increases indefinitely, then

$$\frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} \int_\Lambda e^{-t^2/2\sigma^2} f_1\bigl(g^t(x)\bigr) f_2(x)\, d\nu\, dt \to \nu(\Lambda) \int_\Lambda f_1\, d\nu \int_\Lambda f_2\, d\nu.$$

In Theorem 2, the proof itself uses the von Neumann ergodic theorem. For an arbitrary $\varepsilon > 0$, there is a piecewise constant function $h_\varepsilon\colon \mathbb{R} \to \mathbb{R}$ such that

1. $h_\varepsilon(\omega) = c_k = \text{const}$ on the intervals (ω_k, ω_{k+1}), $k = 1, \dots, N$ (it is possible that $\omega_1 = -\infty$ and $\omega_{N+1} = +\infty$),
2. $I \subset (\omega_1, \omega_{N+1})$, and
3. $\int_I |h - h_\varepsilon|\, d\omega < \varepsilon$.

Therefore (in view of the isometry of the operator U), we have

$$\left| \int_I h(\omega)(U^{\omega t} f_1, f_2)\, d\omega - \int_I h_\varepsilon(\omega)(U^{\omega t} f_1, f_2)\, d\omega \right| \le \int_I |h - h_\varepsilon|\, d\omega\, \|f_1\|\, \|f_2\| \le \varepsilon \|f_1\|\, \|f_2\|,$$

where $\|\cdot\|$ is the norm in L_2. Hence, it suffices to establish the convergence of the integrals

$$J_k(t) = \int_{\omega_k}^{\omega_{k+1}} h_\varepsilon(\omega)(U^{\omega t} f_1, f_2)\, d\omega.$$

By the von Neumann theorem,

$$J_k(t) = \frac{c_k}{t} \int_{\omega_k t}^{\omega_{k+1} t} (U^s f_1, f_2)\, ds \to c_k(\omega_{k+1} - \omega_k)(\bar f_1, f_2) \quad \text{as } t \to \infty.$$

It remains to note that

$$\sum_1^N c_k(\omega_{k+1} - \omega_k) = \int_I h_\varepsilon(\omega)\, d\omega = \int_I h(\omega)\, d\omega + \delta,$$

where $|\delta| \le \varepsilon$, which implies what we had to prove.

Remark 2. As outlined in [6], we used the Stone formula for spectral expansion of the group of unitary operators in the original proof of Theorem 2 (and in that of main Theorem 3 in Sec. 4). We note that the starting point in von Neumann's proof of the statistical ergodic theorem is also the Stone formula (see, e.g., [9].) Modern proofs of ergodic theorems use some other technique.

4. Limit measures of layered flows

We apply the method used in the preceding section to dynamical systems with *layered* phase flows. These systems are a special case of systems (3) if there is only one variable ω ($m = 1$). They are characterized by a distinctive property: the phase flows on Λ at different values of ω turn out to be conjugate (after a suitable change of time). *Geodesic flows* on smooth manifolds provide an important example of layered flows. We now proceed to exact definitions.

Let I be a (possibly infinite) interval on the number axis $\mathbb{R}$, let Λ be a smooth manifold, and let dynamical system (3) with $z \in \Lambda$ and $\omega \in I$ be defined in the phase space $\Gamma = \Lambda \times I$.

We set $P_\gamma = \{(z, \omega) \in \Gamma \colon \omega = \gamma\}$. These are n-dimensional integral manifolds of system (3). The map $\psi_\omega \colon (z, \omega) \to z$ defines a natural diffeomorphism between P_ω and Λ. It is clear that the vector field v of the dynamical system under consideration is tangent to P_ω at the points $(z, \omega) \in \Gamma$. Let v_ω denote the restriction of v to P_ω.

Furthermore, let $\{g^t\}$ be a phase flow on Γ generated by system (3), and let $\{g^t_\omega\}$ be the restriction of $\{g^t\}$ to P_ω. Because all manifolds P_ω are diffeomorphic to Λ, we can assume that $\{g^t_\omega\}$ is a one-parameter family (with ω as the family parameter) of transformation groups on Λ.

Definition 1. A flow g^t is said to be *layered* if there is a smooth function $\alpha \colon I \to (0, \infty)$ and a flow $g^s_* \colon \Lambda \to \Lambda$ such that the diagram

$$\begin{array}{ccc} P_\omega & \xrightarrow{g^t_\omega} & P_\omega \\ {\scriptstyle\psi_\omega}\big\downarrow & & \big\downarrow{\scriptstyle\psi_\omega} \\ \Lambda & \xrightarrow{g_*^{\alpha(\omega)t}} & \Lambda \end{array} \tag{6}$$

is commutative for all $\omega \in I$ and all $t \in \mathbb{R}$. A layered flow is said to be *nonsingular* if the function $\omega \mapsto \alpha(\omega)$ has only isolated critical points.

Using the diffeomorphism ψ_ω to identify P_ω and Λ, we can represent the commutativity property of diagram (6) in the form of the relation

$$g^t_\omega = g_*^{\alpha(\omega)t}. \tag{7}$$

The following proposition is obvious.

Proposition 1. *Let the flow g^t_* preserve a measure ν_* on the manifold Λ, and let σ be an arbitrary measure on the interval I. Then the flow g^t on $\Lambda \times I$ preserves the measure $\mu = \nu_* \times \sigma$.*

We comment on this assertion for the case where the measures ν_* and σ have smooth densities, namely,

$$d\nu_* = \lambda(z)\, d^n z, \qquad d\sigma = \varphi(\omega)\, d\omega.$$

Let the flow g^t_* generate a vector field on Λ

$$v_*(z) = \frac{d}{dt}\bigg|_{t=0} \big(g^t_*(z)\big)$$

such that g^t_* is the phase flow of the system of differential equations $\dot z = v_*(z)$ on Λ. Condition (7) means that system (3) has the form

$$\dot z = \alpha(\omega) v_*(z), \qquad \dot\omega = 0 \tag{8}$$

for the layered flow.

The assumption of invariance of the measure v_* with respect to the action of the flow g_*^t means that the density λ of v_* satisfies the Liouville equation

$$\sum_{i=1}^{n} \frac{\partial \lambda (v_*)_i}{\partial z_i} = 0.$$

Consequently, the density $\rho = \lambda(z)\varphi(\omega)$ of the measure μ satisfies the equation

$$\sum_{i=1}^{n} \frac{\partial \rho v_i}{\partial z_i} = 0,$$

where $v = \lambda v_*$. In view of the relation $\dot{\omega} = 0$, this formula gives a criterion for the invariance of the measure μ with respect to the phase flow of system (8).

Let U^t be a family of unitary Koopman operators on $L_2(\Gamma, \mu)$ generated by a flow g^t.

Theorem 3. *Let g^t be a nonsingular layered flow on $\Gamma = \Lambda \times I$, let the measure ν_* be absolutely continuous with respect to the measure defined by a Riemannian metric on Λ, let the measure σ be absolutely continuous with respect to the usual Lebesgue measure on $\mathbb{R}$, and let $\nu_*(\Lambda) = \sigma(I) = 1$. Then the limit $\lim_{t\to\infty}(U^t f', f'')$ exists for all $f', f'' \in L_2(\Gamma, \mu)$.*

Theorems 1 and 3 imply the following corollary.

Corollary 2. *For a dynamical system of form* (3) *with a nonsingular layered phase flow, the probability distribution density ρ_t has weak limits as $t \to +\infty$ and $t \to -\infty$, and these limits coincide and can be calculated by formula* (1).

Theorem 3 is proved in the next section. Before proceeding to its proof, we present some examples of dynamical systems with layered phase flows.

Let P_ω, $\omega > 0$, be a one-parameter family of smooth manifolds, and let $\varphi_\omega \colon P_1 \to P_\omega$ be a family of diffeomorphisms. The union $\Gamma = \bigcup_{\omega>0} P_\omega$ has the structure of the direct product $\Lambda \times I$, where $\Lambda = P_1$ and I is the half-infinite interval $(0, +\infty)$. We endow Γ with the smooth structure of the direct product $\Lambda \times I$.

Definition 2. A vector field v on Γ is said to be φ_ω-*homogeneous of degree* k if there is a vector field φ_1 on P_1 such that

$$v \circ \varphi_\omega = \omega^k (D\varphi_\omega) v_1 \tag{9}$$

for all $\omega > 0$. Here, D is the differential of the map.

Because v_1 is tangent to P_1 and we have $\varphi_\omega(P_1) = P_\omega$, the field v is tangent to the layers P_ω, $\omega > 0$.

Example 1. Let $\big(M, \langle\,,\,\rangle\big)$ be a smooth Riemannian manifold, and let $\langle\,,\,\rangle^*$ be the metric conjugate to the metric $\langle\,,\,\rangle$. Let p be an element of the conjugate space T_q^*M (i.e., p is a momentum of a mechanical system). We set $\|p\|^2 = \langle p, p\rangle^*$ and $P_s = \big\{p \in T_q^*M :\ q \in M,\ \|p\| = s\big\}$. It is clear that $\Gamma = T^*M \setminus P_0$.

We consider the Hamiltonian vector field v on Γ defined by the standard symplectic structure $\sum dp_i \wedge dq_i$ and by the Hamiltonian $H = \|p\|^2/2$. This field generates a dynamical system on Γ, which is called a *geodesic flow*. The corresponding phase flow g^t is defined for all $t \in \mathbb{R}$ if $\big(M, \langle\,,\,\rangle\big)$ is a complete Riemannian manifold (i.e., all geodesics have an infinite length).

Let (q, p) be a point in Γ. We set $\varphi_\omega(q, p) = (q, \omega p)$, $\omega > 0$. Then the Hamiltonian vector field v determined by the Hamilton equations

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q} \tag{10}$$

is φ_ω-homogeneous of degree $k = 1$. Indeed, the structure of equations of geodesics (10) is such that $\dot{p}$ is quadratic with respect to p and $\dot{q}$ is linear with respect to p.

Example 2. We consider Hamiltonian system (10) in $\mathbb{R}^{2n} = \{q_1, \ldots, q_n, p_1, \ldots, p_n\}$ with the Hamiltonian

$$H = \frac{1}{2}\sum_{i=1}^{n}(p_i^2 + \alpha_i^2 q_i^2), \quad \alpha_i > 0. \tag{11}$$

(It corresponds to a multiharmonic oscillator with the frequencies $\alpha_1, \ldots, \alpha_n$.) We again set

$$P_\omega = \big\{(p, q) \in \mathbb{R}^{2n} : H(p, q) = \omega\big\}, \quad \omega > 0,$$

and $\varphi_\omega \colon (q, p) \to (\omega q, \omega p)$. It is easy to understand that the Hamiltonian vector field v on $\Gamma = \bigcup_{\omega>0} P_\omega$ generated by quadratic Hamiltonian (11) is φ_ω-homogeneous of degree zero (i.e., $k = 0$).

Proposition 2. *Let v be a φ_ω-homogeneous vector field of degree k on Γ. Then v generates a layered flow g^t on Γ for which $\alpha(\omega) = \omega^k$.*

Corollary 3. *If $k \neq 0$, then the flow g^t is nonsingular.*

Therefore, the geodesic flow under consideration is nonsingular, whereas the flows generated by linear Hamiltonian systems (see Example 2) are singular.

Proof of Proposition 2. We set $\Lambda = P_1$ and note that every point $z \in \Gamma$ can be uniquely represented in the form

$$z = \varphi_\omega(z_1), \quad \omega > 0, \quad z_1 \in P_1. \tag{12}$$

Because the map $\Lambda \times (0, +\infty) \to \Gamma$ defined by (12) is a diffeomorphism, it suffices to verify that

$$g^t \circ \varphi_\omega = \varphi_\omega \circ g_1^{\omega^k t},$$

where g_1^t is the restriction of the flow g^t to P_1.

The above relation obviously holds for $t = 0$. We verify the equality of the derivatives with respect to t. Differentiating with respect to t at the point $t = 0$, we obtain

$$\frac{d}{dt}\bigg|_{t=0} g^t \circ \varphi_\omega = v \circ \varphi_\omega,$$

$$\frac{d}{dt}\bigg|_{t=0} \varphi_\omega \circ g_1^{\omega^k t} = \omega^k (D\varphi_v) v_1,$$

where v_1 is the restriction of v to P_1. The equality of these derivatives follows from definition (9) of a φ_ω-homogeneous field of degree k. The equality of the derivatives for all values of t follows from the group property of phase flows. Proposition 2 is proved.

5. Koopman operators for layered flows

In this section, we prove Theorem 3. Let U^t, U^t_ω, and U^t_* be the Koopman unitary operators corresponding to the flows g^t, g^t_ω, and g^t_*.

Proposition 3. *We have the relation*

$$U^t_\omega = U_*^{\alpha(\omega) t}.$$

Indeed, for an arbitrary function f in $L_2(\Lambda, \nu_*)$, we have the chain of relations

$$U_\omega^t f = f \circ g_\omega^t = f \circ g_*^{\alpha(\omega)t} = U_*^{\alpha(\omega)t} f.$$

To derive the second relation in this chain, we use the diffeomorphism ψ_ω to identify the manifolds Λ and P_ω and then apply relation (7). Proposition 3 is proved.

Let f' and f'' be some functions belonging to $L_2(\Gamma, \mu)$, $\Gamma = \Lambda \times I$. We set

$$f'_\omega(\,\cdot\,) = f'(\,\cdot\,, \omega), \qquad f''_\omega(\,\cdot\,) = f''(\,\cdot\,, \omega).$$

Applying the Fubini theorem and Proposition 3, we obtain

$$(U^t f', f'') = \int_I (U_\omega^t f'_\omega, f''_\omega)\, d\sigma = \int_I \Big(U_*^{\alpha(\omega)t} f'_\omega, f''_\omega\Big)\, d\sigma. \tag{13}$$

Let A' and A'' be some measurable subsets of Λ, and let I' and I'' be intervals lying in I. Let $\chi' \colon \Gamma \to \mathbb{R}$ and $\chi'' \colon \Gamma \to \mathbb{R}$ be the characteristic (indicator) functions of the respective sets $A' \times I'$ and $A'' \times I''$. We set $J(t) = (U^t \chi', \chi'')$.

Main lemma. *The limits* $\lim_{t\to+\infty} J(t)$ *and* $\lim_{t\to-\infty} J(t)$ *exist.*

Theorem 3 follows from the main lemma. Indeed, because ν_* is absolutely continuous with respect to the measure defined on Λ by a Riemannian metric and σ is absolutely continuous with respect to the Lebesgue measure, the space of compactly supported continuous functions on Γ is everywhere dense in $L_2(\Gamma, \mu)$. In turn, the linear space of functions that are linear combinations of indicator functions of μ-measurable subsets in Γ is everywhere dense (even with respect to the C^0-norm) in the abovementioned space of compactly supported continuous functions on Γ.

Next, let f' and f'' be two arbitrary functions in $L_2(\Gamma, \mu)$. To prove the existence of the limit $\lim_{t\to+\infty}(U^t f', f'')$, we use the Cauchy test for convergence, i.e., it is necessary to prove that the difference

$$(U^{t_1} f', f'') - (U^{t_2} f', f'') \tag{14}$$

is smaller than an arbitrarily chosen $\varepsilon > 0$ for all $t_1, t_2 > T(\varepsilon)$. For this, we approximate f' and f'' with functions g' and g'' that are linear combinations of the indicator functions χ' and χ'', namely, for an arbitrary $\varepsilon > 0$, there are some functions g' and g'' such that

$$\|f' - g'\| < \varepsilon, \qquad \|f'' - g''\| < \varepsilon, \tag{15}$$

where $\|\cdot\|$ is the norm in L_2. After this remark, difference (14) should be represented as

$$\begin{aligned}
&(U^{t_1} g', g'') - (U^{t_2} g', g'') + \\
&\qquad + \big(U^{t_1}(f' - g'), f'' - g''\big) + \big(U^{t_1}(f' - g'), g''\big) + (U^{t_1} f', f'' - g'') - \\
&\qquad - \big(U^{t_2}(f' - g'), f'' - g''\big) - \big(U^{t_2}(f' - g'), g''\big) - (U^{t_2} f', f'' - g'').
\end{aligned} \tag{16}$$

According to the main lemma, the difference in the first line in formula (16) can be made arbitrarily small for sufficiently large values of t_1 and t_2. In view of the Cauchy–Schwarz inequality and by the unitarity of the Koopman operator U and inequalities (15), the other terms in (16) tend to zero uniformly with respect to t_1 and t_2 as $\varepsilon \to 0$.

The remaining part of this section is devoted to proving the main lemma.

We set $A_0 = A' \cap A''$ and $I_0 = I' \cap I''$. Let $\chi_0 \colon \Lambda \times I \to \mathbb{R}$ be the characteristic (indicator) function of the measurable set $A_0 \times I_0$, and let $\tilde{\chi}_0 \colon \Lambda \to \mathbb{R}$ be that of the set A_0. Clearly,

$$J(t) = \int_{I_0} (U_*^{\alpha(\omega)t} \tilde{\chi}_0, \tilde{\chi}_0)\, d\sigma.$$

Let $D_\gamma = \{\omega \in I_0 \colon |\alpha'(\omega)| > \gamma\}$, where $\alpha' = d\alpha/d\omega$. By the hypothesis, the critical points of the function $\omega \mapsto \alpha(\omega)$ are isolated. Consequently, the σ-measure of the set $I \setminus D_\gamma$ tends to zero as $\gamma \to 0$. Moreover, it can be assumed that D_γ is a union of finitely many intervals. Let (ω_1, ω_2) be one of the intervals composing D_γ. Then α can be regarded as a coordinate on (ω_1, ω_2). Indeed, the inverse function $\omega(\alpha)$ of $\alpha \colon (\omega_1, \omega_2) \to \mathbb{R}$ exists and is smooth. We set $d\sigma(\omega) = h(\omega)\, d\omega$. By the assumptions in Theorem 3, $\omega \to h(\omega)$ is an integrable function, i.e., $h \in L_1(I, d\omega)$. Hence,

$$\int_{\omega_1}^{\omega_2} (U_*^{\alpha(\omega)t} \tilde{\chi}_0, \tilde{\chi}_0) h(\omega)\, d\omega = \int_{\alpha(\omega_1)}^{\alpha(\omega_2)} (U_*^{\alpha t} \tilde{\chi}_0, \tilde{\chi}_0) h\big(\omega(\alpha)\big)\omega'(\alpha)\, d\alpha. \tag{17}$$

Because $h\big(\omega(\alpha)\big)\omega'(\alpha) \in L_1\big(\big(\alpha(\omega_2), \alpha(\omega_1)\big), d\alpha\big)$, Theorem 2 implies that integral (17) has a limit as $t \to \infty$. The proof of the main lemma is complete.

6. Entropy increase

As is known, entropy is defined in statistical mechanics as the integral

$$S_t = -\int_\Gamma \rho_t \log \rho_t\, d\mu, \qquad d\mu = d^{2n} z.$$

Because we have $\rho_t(z) = \rho\big(g^{-t}(z)\big)$ and the flow g^t preserves the measure μ, it is clear that $S_t = \text{const}$. This remark is a special case of the general result by Poincaré concerning the constancy of the *fine* entropy of dynamical systems [4].

On the other hand, ρ_t weakly converges to $\bar{\rho}$ as $t \to \pm\infty$. Generalizing the considerations of Gibbs and Poincaré, we can assume that the stationary probability distribution density $\bar{\rho}$ corresponds to the *thermal equilibrium* of the dynamical system in question. Therefore, it is natural to introduce the entropy of the equilibrium,

$$S_\infty = -\int_\Gamma \bar{\rho} \log \bar{\rho}\, d\mu.$$

Theorem 4. *The relation*

$$S_t \le S_\infty \tag{18}$$

holds.

We prove the theorem using the property of concavity of the function $h(x) = -x \log x$ for positive values of x. Because $S_t = \text{const}$, the Fubini theorem implies the formula

$$S_t = \frac{1}{T} \int_0^T \int_\Gamma h(\rho_t)\, d\mu\, dt = \int_\Gamma \left[\frac{1}{T} \int_0^T h(\rho_t)\, dt\right] d\mu.$$

In view of the Jensen inequality (see [10]), we have

$$\frac{1}{T}\int_0^T h(\rho_t)\,dt \le h\biggl(\frac{1}{T}\int_0^T \rho_t\,dt\biggr), \quad T > 0.$$

Consequently,

$$S_t\text{‘}\int_\Gamma h\biggl(\frac{1}{T}\int_0^T \rho_t\,dt\biggr)\,d\mu.$$

Passing to the limit as $T \to \infty$ and using Theorem 1, we obtain what we had to prove.

Remark 3. Inequality (18) was previously established in [11] for a collisionless medium in a rectangular parallelepiped. In some special cases, it was already indicated by Poincaré [4].

Because ρ_t weakly converges to the same function $\bar{\rho}$ both as $t \to +\infty$ and as $t \to -\infty$, the conclusion of Theorem 4 about entropy increase is invariant with respect to reversal of the time t. We recall that the Boltzmann kinetic equation implies a *monotonic* increase of entropy with increasing time and that the famous Loschmidt paradox is related to this property (see, e.g., [12] for a discussion of this paradox from various standpoints).

In conclusion, we comment on Krylov's criticism [13] (pp. 51–52 in the Russian edition) of Poincaré's result concerning entropy increase under perturbations in a collisionless medium [4]. Poincaré considers an equilibrium of a one-dimensional ideal gas uniformly filling an interval. A gravitating body approaches the interval from infinity, the gas is allowed to attain a new equilibrium, after which the body recedes back to infinity. The related entropy increase, which turns out to be positive, is calculated in [4].

Krylov notes that this conclusion contradicts the result in Poincaré's work on the constancy of the fine entropy. (The fact that the equations of motion are nonautonomous does not play any role here.) But, in reality, there is no contradiction here at all. Poincaré, in fact, replaced the initial probability density (without stipulating this explicitly) with its weak imit. As the body comes closer to the interval with the gas, the corresponding dynamical system with one degree of freedom changes, namely, the potential energy of gravitational interaction is added to the Hamiltonian. The weak limit ρ_- for this new nonlinear system turns out to be a function of the total energy. After that the body recedes to infinity, the function ρ_- becomes the initial distribution density, and the Hamiltonian again coincides with the kinetic energy. Therefore, the weak limit ρ_+ of the solution of the Liouville equation for the new system with the initial datum ρ_- becomes a function depending only on the kinetic energy. By Theorem 4, $S_+ \ge S_-$. As is shown by Poincaré's calculations, we in fact have $S_+ > S_-$.

Acknowledgements. The authors are grateful to V. Ten for the valuable discussion.

This work was supported by the Russian Foundation for Basic Research (Grant Nos. 02-01-00400, 02-01-01059, and 00-15-99269) and INTAS (Grant No. 00-221).

REFERENCES

1. J. W. Gibbs, "Elementary principles in statistical mechanics," in: *The Collected Works of J. W. Gibbs*, Vol. II, Part 1, Yale Univ. Press, New Haven, Conn. (1931).
2. N. N. Bogoliubov, *Problems of a Dynamical Theory in Statistical Physics* [in Russian], Gostekhizdat, Moscow (1946); English transl., North-Holland, Amsterdam (1962).
3. G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics* (Lectures in Applied Mathematics. Proceedings of the Summer Seminar, Boulder, Colorado, 1960, Vol. 1), Amer. Math. Soc., Providence, R. I. (1963).
4. H. Poincaré, *J. Phys. Théor. Appl., 4 Série*, **5**, 369–403 (1906).

5. P. R. Halmos, *Lectures on Ergodic Theory*, The Mathematical Society of Japan, Tokyo (1956); A. Lasota and M. C. Mackey, *Probabilistic Properties of Deterministic Systems*, Cambridge Univ. Press, Cambridge (1985).
6. V. V. Kozlov, *Dokl. Rossiiskoi Akad. Nauk*, **382**, 593 (2002).
7. V. V. Nemytskii and V. V. Stepanov, *Qualitative Theory of Differential Equations* [in Russian], Gostekhizdat, Moscow (1949); English transl. (Princeton Math. Series), Princeton Univ. Press, Princeton, N. J. (1960).
8. G. H. Hardy, *Divergent Series*, Clarendon, Oxford (1949).
9. F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionnelle* (6th Ed.), Akadémiai Kiadó, Budapest (1972); B. Sz.-Nagy, "Unitary dilations of Hilbert space operators and related topic," in: *Expository Lectures from the CBMS Regional Conference Held at the University of New Hampshire, June 7–11, 1971*, Amer. Math. Soc., Providence, R. I. (1974).
10. G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, Cambridge Univ. Press, Cambridge (1934).
11. V. V. Kozlov, *Reg. Chaotic Dyn.*, **6**, 235 (2001).
12. M. Kac, *Probability and Related Questions in Physics* [in Russian], Mir, Moscow (1951); I. Prigogine and I. Stengers, *Order out of Chaos*, Hienemann, London (1984).
13. N. S. Krylov, *Works on the Foundations of Statistical Physics* [in Russian], USSR Acad. Sci. Press, Moscow (1950); English transl. (Princeton Series in Physics), Princeton Univ. Press, Princeton, N. J. (1979).