Title

Weak convergence to equilibrium of statistical ensembles in integrable Hamiltonian systems

Journal

Journal of Mathematical Physics, 60(5)

ISSN

0022-2488

Author

Mitchell, Chad

Publication Date

2019-05-01

DOI

10.1063/1.5043419

Peer reviewed

Weak Convergence to Equilibrium of Statistical Ensembles in Integrable Hamiltonian Systems

Chad Mitchell*
Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA

(Dated: June 7, 2018)

This article explores the long-time behavior of the bounded orbits associated with an ensemble of initial conditions in a nondegenerate integrable Hamiltonian system. Such systems are inherently nonlinear and subject to highly regular phase space filamentation that can drive the ensemble of orbits toward a stationary state. Describing the statistical ensemble by a probability density on a neighborhood of a family of invariant tori, it is proved that the probability density describing the ensemble at time t converges weakly to an invariant density as $t \to \infty$. More generally, we provide sufficient conditions for convergence to equilibrium of a multiphase system in action-angle form. These ideas are applied to an illustrative exactly-soluble example. This work is relevant for understanding the statistical mechanics of integrable and near-integrable Hamiltonian systems.

I. INTRODUCTION

Establishing the precise conditions needed to ensure that a statistical ensemble of orbits relaxes asymptotically under a smooth flow (or map iteration) to an invariant steady-state is a central problem of ergodic theory and nonequilibrium statistical mechanics. In the case of Hamiltonian flows, an ergodic condition such as strong mixing [1]-[2] is needed to prove that an ensemble of orbits of fixed energy E relaxes (in an appropriate weak or coarse-grained sense) to the usual microcanonical ensemble of equilibrium statistical mechanics. In this case, the statistical behavior of orbits becomes asymptotically simple, as orbits are equally likely to be found anywhere on the level set of energy E.

At the other extreme, integrable Hamiltonian flows possess a large number of conserved quantities, resulting in highly regular orbits that are confined to lower-dimensional manifolds that foliate the system phase space. The study of integrable Hamiltonian systems is a rich area of mathematical research [3]-[6] that has recently found new applications to the physics of quantum statistical mechanics [7]-[9] and charged-particle dynamics in particle accelerators [10]-[11]. When the Hamiltonian is nondegenerate, an ensemble of orbits in such a system is subject to a process of *filamentation* or *phase mixing*, that drives the statistical system toward an invariant steady-state. Due to the presence of conserved quantities, the ensemble may relax into one of a large number of possible steady-states, depending on the detailed initial conditions of the ensemble.

The purpose of this article is to provide complete proofs of theorems regarding the long time behavior (as $t \to \infty$) of an ensemble of initial conditions evolving under the flow of such a system. More precisely, we consider the following random initial value problem in action-angle form:

$$\frac{d\vec{I}}{dt} = 0, \quad \frac{d\vec{\phi}}{dt} = \omega(\vec{I}), \qquad \vec{I}(0) = \vec{I}_0, \quad \vec{\phi}(0) = \vec{\phi}_0, \tag{1}$$

where $\vec{I} \in \Omega \subset \mathbb{R}^M$, Ω open, $\vec{\phi} \in \mathbb{T}^N$, $\omega : \Omega \to \mathbb{T}^N$, and the random initial condition $(\vec{I}_0, \vec{\phi}_0) \in \Omega \times \mathbb{T}^N$ is described by a probability density $f_0 = f_0(\vec{I}_0, \vec{\phi}_0)$. Given a continuous observable G, we give conditions for the convergence as $t \to \infty$ of the expected value of G along the orbit (1). Likewise, letting f_t denote the probability density describing the orbit (1) at time t, we give conditions for weak convergence of the associated probability measures P_t as $t \to \infty$.

* ChadMitchell@lbl.gov

The proofs make use of mathematical tools in harmonic analysis (Parseval's theorem, properties of multiple Fourier series, etc.) and measure-based integration (dominated convergence, Fubini's theorem), together with some basic results from functional analysis and probability, for which we use [12]-[15] as references. Some background in Hamiltonian mechanics [16]-[18] is beneficial, but no ergodic theory is used. We make no claim that the proofs given here are the simplest possible, but we have tried to include as many technical details as possible.

Section 2 provides basic results regarding the random initial value problem (1). In Section 3, we provide conditions sufficient to guarantee convergence as $t \to \infty$ of the expected values of the observables of such a system (Theorems 1-4). In Section 4, we demonstrate weak convergence as $t \to \infty$ of the one-parameter family of probability measures $P_t \Rightarrow P_{eq}$ describing the system at time t (Theorem 5). Section 5 provides a detailed and exactly-soluble numerical example. In Section 6, we conclude with a brief summary. Appendix A contains a summary of notation, while Appendix B provides useful background from the theory of integrable Hamiltonian systems.

II. PRELIMINARIES

A Hamiltonian dynamical system on a phase space of dimension $2N$ is said to be *completely integrable* if there exist N functionally independent integrals of motion $H = f_1, \ldots, f_N$ with $\{f_j, f_k\} = 0$ for $i, j = 1, \ldots, N$, where $\{\cdot, \cdot\}$ denotes the Poisson bracket. In such a system, the dynamics can locally be described in a particularly simple action-angle form. (See Appendix B for a precise statement.) As a result, it is sufficient to consider a problem of the following form.

Suppose M and N are positive integers. Let $\Omega \subset \mathbb{R}^M$ be nonempty and open, and consider a map $\omega : \Omega \to \mathbb{R}^N$, where we assume $\omega \in C^2(\Omega, \mathbb{R}^N)$. In this paper, we will consider the dynamical system on $\Omega \times \mathbb{T}^N$ given by $M + N$ equations of the form:

$$\frac{d\vec{I}}{dt} = 0, \qquad \frac{d\vec{\phi}}{dt} = \omega(\vec{I}), \qquad (\vec{I} \in \Omega, \quad \vec{\phi} \in \mathbb{T}^N). \tag{2}$$

In the special case that $M = N$, the coordinates $(\vec{I}, \vec{\phi})$ could play the role of local action-angle variables for an integrable Hamiltonian system. However, in the results that follow, nothing is lost by considering the more general system just described, which is of related interest [19].

We refer to the map $\omega : \Omega \to \mathbb{R}^N$ as the *frequency map* of the system (2). A point $\vec{I} \in \Omega$ is said to be a *regular point* of ω if the Jacobian matrix $D\omega$ has rank N at $\vec{I}$. This is equivalent to the condition that the set $\{D\omega_1(\vec{I}), \ldots, D\omega_N(\vec{I})\}$ is linearly independent in $\mathbb{R}^M$. Otherwise, $\vec{I} \in \Omega$ is said to be a *critical point* of ω.

The system of equations (2) defines a corresponding flow on $\Omega \times \mathbb{T}^N$ given by the one-parameter family of maps $\Psi_t(\vec{I}, \vec{\phi}) = (\vec{I}, \vec{\phi} + t\omega(\vec{I}))$, $t \in \mathbb{R}$. Note that the Jacobian matrix of Ψ_t satisfies:

$$D\Psi_t = \begin{pmatrix} \mathbf{1}_M & 0 \\ tD\omega & \mathbf{1}_N \end{pmatrix}, \qquad \det(D\Psi_t) = 1, \tag{3}$$

where $\mathbf{1}_M$ denotes the $M \times M$ identity matrix, with a similar definition for $\mathbf{1}_N$. Thus, the flow Ψ_t is volume-preserving.

We treat (2) as a *random initial value problem.* Suppose that the initial condition of (2) is modeled as a pair of (generally) dependent random vectors described by a joint probability density $f_0 \in L^1(\Omega \times \mathbb{T}^N)$. Then the point $\Psi_t(\vec{I}_0, \vec{\phi}_0)$ at time $t \in \mathbb{R}$ is described by the following probability density $f_t \in L^1(\Omega \times \mathbb{T}^N)$:

$$f_t(\vec{I}, \vec{\phi}) = f_0(\Psi_t^{-1}(\vec{I}, \vec{\phi})) = f_0(\vec{I}, \vec{\phi} - t\omega(\vec{I})), \qquad (\vec{I} \in \Omega, \quad \vec{\phi} \in \mathbb{T}^N). \tag{4}$$

This follows from the fact that Ψ_t is volume-preserving. By an *observable*, we mean a time-dependent random variable of the form:

$$G(\Psi_t(\vec{I}_0,\vec{\phi}_0)), \qquad G\in C(\Omega\times\mathbb{T}^N). \tag{5}$$

The *expected value* of G at any time $t\in\mathbb{R}$ is given by:

$$\langle G\rangle_t=\int_{\Omega\times\mathbb{T}^N}G(\Psi_t(\vec{I}_0,\vec{\phi}_0))f_0(\vec{I}_0,\vec{\phi}_0)d\vec{I}_0d\vec{\phi}_0=\int_{\Omega\times\mathbb{T}^N}G(\vec{I}_0,\vec{\phi}_0+t\omega(\vec{I}_0))f_0(\vec{I}_0,\vec{\phi}_0)d\vec{I}_0d\vec{\phi}_0, \tag{6}$$

provided this integral exists. In Section 4, we will consider the probability measures P_t ($t\in\mathbb{R}$) defined by:

$$P_t(A)=\int_A f_t(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi}, \qquad A\subset\Omega\times\mathbb{T}^N, \tag{7}$$

giving the probability that the orbit $\Psi_t(\vec{I}_0,\vec{\phi}_0)$ falls within the Borel set A at time t.

III. EXPECTED VALUES OF OBSERVABLES

In this section, we study observables of the random initial value problem (2). Given $G\in C(\Omega\times\mathbb{T}^N)$, we define the *angle average* $\bar{G}$ of G by:

$$\bar{G}(\vec{I})=\frac{1}{(2\pi)^N}\int_{\mathbb{T}^N}G(\vec{I},\vec{\phi})d\vec{\phi}, \qquad \vec{I}\in\Omega. \tag{8}$$

We give sufficient conditions for the following long time convergence of the expected value of G:

$$\lim_{t\to\infty}\langle G\rangle_t=\langle\bar{G}\rangle_0. \tag{9}$$

The main convergence result appears in Theorem 4, which we approach in a series of steps.

Theorem 1: In the system (2), suppose $\omega\in C^2(\Omega,\mathbb{R}^N)$ has no critical points. Suppose the initial condition is described by a probability density $f_0\in C_c(\Omega\times\mathbb{T}^N)$, continuous with compact support. For any continuous function $G\in C(\Omega\times\mathbb{T}^N)$, the expected value of the observable G exists and is bounded for all $t\in\mathbb{R}$, and

$$\lim_{t\to\infty}\langle G\rangle_t=\langle\bar{G}\rangle_0. \tag{10}$$

The key to the proof of this result is the following theorem regarding the asymptotic behavior of oscillatory integrals [14], which is a generalization of the Riemann-Lebesgue lemma. The proof is rather standard, but details have been included here for completeness.

Lemma 0: Let $\Omega\subset\mathbb{R}^M$ be open. Suppose that $a\in L^1(\Omega)$, and that $\phi\in C^2(\Omega)$ is real-valued with $\nabla\phi\neq0$. Then for $\lambda\in\mathbb{R}$,

$$I(\lambda)=\int_\Omega a(x)e^{i\lambda\phi(x)}dx\to0 \quad \text{as} \quad |\lambda|\to\infty. \tag{11}$$

Proof: We prove this result first for the special case that $a\in C_c^\infty(\Omega)$, and the general result will follow by a density argument. For any $\lambda\neq0$ we have on Ω that:

$$e^{i\lambda\phi}=\frac{1}{i\lambda}\frac{\nabla\phi}{|\nabla\phi|^2}\cdot\nabla e^{i\lambda\phi}. \tag{12}$$

Using this result in (11) gives

$$I(\lambda) = \frac{1}{i\lambda}\int_\Omega a(x)\frac{\nabla\phi(x)}{|\nabla\phi(x)|^2}\cdot\nabla e^{i\lambda\phi(x)}dx \tag{13}$$

$$= -\frac{1}{i\lambda}\int_\Omega \nabla\cdot\left(a(x)\frac{\nabla\phi(x)}{|\nabla\phi(x)|^2}\right)e^{i\lambda\phi(x)}dx \tag{14}$$

$$=: -\frac{1}{i\lambda}\int_\Omega u(x)e^{i\lambda\phi(x)}dx, \tag{15}$$

where the second equality follows by an application of integration by parts, since a vanishes on the boundary of Ω. It follows from the smoothness conditions on a and ϕ that u is continuous. Letting $K \subset \Omega$ denote the compact support of a, it follows that:

$$|I(\lambda)| \le \frac{1}{|\lambda|}\int_\Omega |u(x)|dx \le \frac{1}{|\lambda|}\max_{x\in K}|u(x)|m(K), \tag{16}$$

where m denotes Lebesgue measure on $\mathbb{R}^N$. The limit (11) now follows.

Now suppose that $a \in L^1(\Omega)$. Since $C_c^\infty(\Omega)$ is dense in $L^1(\Omega)$, there exists a sequence $a_n \in C_c^\infty(\Omega)$ $(n = 1, 2, \ldots)$ such that

$$||a - a_n||_1 \to 0 \quad \text{as} \quad n \to \infty. \tag{17}$$

Put

$$I(\lambda) = \int_\Omega a(x)e^{i\lambda\phi(x)}dx, \qquad I_n(\lambda) = \int_\Omega a_n(x)e^{i\lambda\phi(x)}dx. \tag{18}$$

Fix $\epsilon > 0$. By (17), there exists an integer $n_0 > 0$ such that $||a - a_{n_0}||_1 < \epsilon$. The first part of the proof shows that $I_{n_0}(\lambda)$ goes to zero for large λ. Thus, there exists $\lambda_0 > 0$ such that $|I_{n_0}(\lambda)| < \epsilon$ for all $|\lambda| > \lambda_0$. The triangle inequality and (18) together now imply that for all $|\lambda| > \lambda_0$,

$$|I(\lambda)| \le |I_{n_0}(\lambda)| + |I(\lambda) - I_{n_0}(\lambda)| \le |I_{n_0}(\lambda)| + ||a - a_{n_0}||_1 < 2\epsilon. \tag{19}$$

This proves the limit (11). □

The following two technical lemmas are needed to ensure that limiting operations can be exchanged freely, and that the above lemma can be applied to our specific case.

Lemma 1: Let $\Omega \subset \mathbb{R}^M$ be open, let $G \in C(\Omega \times \mathbb{T}^N)$, and define the function $\hat{G} : \Omega \times \mathbb{Z}^N \to \mathbb{C}$ by:

$$\hat{G}(\vec{I}, \vec{n}) = \frac{1}{(2\pi)^N}\int_{\mathbb{T}^N} G(\vec{I}, \vec{\phi})e^{-i\vec{n}\cdot\vec{\phi}}d\vec{\phi}, \qquad (\vec{I} \in \Omega, \quad \vec{n} \in \mathbb{Z}^N). \tag{20}$$

Then for each $\vec{n} \in \mathbb{Z}^N$ the function $\hat{G}(\cdot, \vec{n})$ lies in $C(\Omega)$.

Proof: Fix $\vec{I}_0 \in \Omega$, and let B be an open ball centered at $\vec{I}_0$ whose closure $\bar{B}$ satisfies $\bar{B} \subset \Omega$. Since $\bar{B} \times \mathbb{T}^N$ is compact and G is continuous, for all $\vec{I} \in \bar{B}$, $\vec{\phi} \in \mathbb{T}^N$ we have:

$$|G(\vec{I}, \vec{\phi})| \le \max_{\bar{B}\times\mathbb{T}^N}|G| < \infty. \tag{21}$$

Suppose $\{\vec{I}_m\}$ is any sequence in Ω such that $\vec{I}_m \to \vec{I}_0$ as $m \to \infty$. By choosing a tail of the sequence if necessary, we may assume that $\{\vec{I}_m\}$ lies in $\bar{B}$. It then follows from (21) and the dominated convergence theorem that

$$\lim_{m\to\infty}\hat{G}(\vec{I}_m, \vec{n}) = \frac{1}{(2\pi)^N}\int_{\mathbb{T}^N}\lim_{m\to\infty}G(\vec{I}_m, \vec{\phi})e^{-i\vec{n}\cdot\vec{\phi}}d\vec{\phi} = \hat{G}(\vec{I}_0, \vec{n}), \tag{22}$$

so $\hat{G}(\cdot,\vec{n})$ is continuous at $\vec{I}_0$. Since the same argument applies for any $\vec{I}_0 \in \Omega$, we have shown that $\hat{G}(\cdot,\vec{n}) \in C(\Omega)$. □

Lemma 2: Suppose $G \in C(\Omega \times \mathbb{T}^N)$ and $f_0 \in C_c(\Omega \times \mathbb{T}^N)$. Then

$$\int_\Omega d\vec{I} \sum_{\vec{n}\in\mathbb{Z}^N} |\hat{G}(\vec{I},\vec{n})\hat{f}_0(\vec{I},-\vec{n})| < \infty, \tag{23}$$

where $\hat{G}$ and $\hat{f}_0$ are defined as in Lemma 1.

Proof: Note that by the Cauchy-Schwarz inequality we have for each fixed $\vec{I} \in \Omega$ that:

$$\sum_{\vec{n}\in\mathbb{Z}^N} |\hat{G}(\vec{I},\vec{n})\hat{f}_0(\vec{I},-\vec{n})| \leq \left(\sum_{\vec{n}\in\mathbb{Z}^N} |\hat{G}(\vec{I},\vec{n})|^2\right)^{1/2} \left(\sum_{\vec{n}\in\mathbb{Z}^N} |\hat{f}_0(\vec{I},-\vec{n})|^2\right)^{1/2}. \tag{24}$$

Since the function $G(\vec{I},\cdot)$ is continuous for fixed $\vec{I} \in \Omega$ and therefore lies in $L^2(\mathbb{T}^N)$, we may apply Parseval's theorem to get:

$$\sum_{\vec{n}\in\mathbb{Z}^N} |\hat{G}(\vec{I},\vec{n})|^2 = \frac{1}{(2\pi)^N}\int_{\mathbb{T}^N} |G(\vec{I},\vec{\phi})|^2 d\vec{\phi} \equiv ||G(\vec{I},\cdot)||_2^2 < \infty. \tag{25}$$

An identical argument applies to the sum involving $\hat{f}_0$, so we have shown that for fixed $\vec{I} \in \Omega$:

$$\sum_{\vec{n}\in\mathbb{Z}^N} |\hat{G}(\vec{I},\vec{n})\hat{f}_0(\vec{I},-\vec{n})| \leq ||G(\vec{I},\cdot)||_2 ||f_0(\vec{I},\cdot)||_2 < \infty. \tag{26}$$

Consider the functions $g_1, g_2 : \Omega \to \mathbb{R}$ defined by $g_1(\vec{I}) = ||G(\vec{I},\cdot)||_2$ and $g_2(\vec{I}) = ||f_0(\vec{I},\cdot)||_2$ for all $\vec{I} \in \Omega$. Since g_1 and g_2 are defined by integrals of the form (25), an argument using the dominated convergence theorem (similar to Lemma 1) shows that g_1 and g_2 are each continuous on Ω. Since we have assumed that f_0 has compact support contained in $\Omega \times \mathbb{T}^N$, it follows that g_2 has compact support K contained in Ω. Since g_1 and g_2 each attain finite maxima on K, we have:

$$\int_\Omega d\vec{I} \sum_{\vec{n}\in\mathbb{Z}^N} |\hat{G}(\vec{I},\vec{n})\hat{f}_0(\vec{I},-\vec{n})| \leq \int_K g_1(\vec{I})g_2(\vec{I})d\vec{I} \leq m(K)\max_K |g_1 g_2| < \infty, \tag{27}$$

where m denotes Lebesgue measure on $\mathbb{R}^N$. This proves the Lemma. □

Proof of Theorem 1: From (6),

$$\langle G \rangle_t = \int_{\Omega\times\mathbb{T}^N} G(\vec{I},\vec{\phi} + t\vec{\omega}(\vec{I}))f_0(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi}. \tag{28}$$

Define functions $\hat{G} : \Omega \times \mathbb{Z}^N \to \mathbb{C}$ and $\hat{f}_0 : \Omega \times \mathbb{Z}^N \to \mathbb{C}$ describing the (action-dependent) Fourier coefficients of G and f_0:

$$\hat{G}(\vec{I},\vec{n}) = \frac{1}{(2\pi)^N}\int_{\mathbb{T}^N} G(\vec{I},\vec{\phi})e^{-i\vec{n}\cdot\vec{\phi}}d\vec{\phi}, \qquad \hat{f}_0(\vec{I},\vec{n}) = \frac{1}{(2\pi)^N}\int_{\mathbb{T}^N} f_0(\vec{I},\vec{\phi})e^{-i\vec{n}\cdot\vec{\phi}}d\vec{\phi}. \tag{29}$$

Since the functions $G(\vec{I},\cdot)$ and $f_0(\vec{I},\cdot)$ are continuous on $\mathbb{T}^N$ and therefore lie in $L^1(\mathbb{T}^N)$, the above integrals are defined for all $\vec{I} \in \Omega$. Furthermore, the Fourier coefficients of the function on $\mathbb{T}^N$ defined by $\vec{\phi} \mapsto G(\vec{I},\vec{\phi} + t\vec{\omega}(\vec{I}))$ for fixed $\vec{I}$ are given by:

$$\frac{1}{(2\pi)^N}\int_{\mathbb{T}^N} G(\vec{I},\vec{\phi} + t\vec{\omega}(\vec{I}))e^{-i\vec{n}\cdot\vec{\phi}}d\vec{\phi} = \hat{G}(\vec{I},\vec{n})e^{it\vec{n}\cdot\vec{\omega}(\vec{I})}. \tag{30}$$

Since the functions $\vec{\phi} \mapsto G(\vec{I}, \vec{\phi} + t\vec{\omega}(\vec{I}))$ and $f_0(\vec{I}, \cdot)$ lie in $L^2(\mathbb{T}^N)$ for fixed $\vec{I} \in \Omega$, Parseval's theorem implies that for each $\vec{I} \in \Omega$:

$$\frac{1}{(2\pi)^N} \int_{\mathbb{T}^N} G(\vec{I}, \vec{\phi} + t\vec{\omega}(\vec{I})) f_0^*(\vec{I}, \vec{\phi}) d\vec{\phi} = \sum_{\vec{n} \in \mathbb{Z}^N} \hat{G}(\vec{I}, \vec{n}) \hat{f}_0^*(\vec{I}, \vec{n}) e^{it\vec{n} \cdot \vec{\omega}(\vec{I})}. \tag{31}$$

Here the $*$ denotes complex conjugation. Using this result, together with the fact that f_0 is real, to evaluate (28) gives (noting that $\hat{f}_0^*(\vec{I}, \vec{n}) = \hat{f}_0(\vec{I}, -\vec{n})$):

$$\langle G \rangle_t = (2\pi)^N \int_\Omega \sum_{\vec{n} \in \mathbb{Z}^N} \hat{G}(\vec{I}, \vec{n}) \hat{f}_0(\vec{I}, -\vec{n}) e^{it\vec{n} \cdot \vec{\omega}(\vec{I})} d\vec{I}. \tag{32}$$

It follows from Lemma 2 that

$$|\langle G \rangle_t| \leq (2\pi)^N \int_\Omega \sum_{\vec{n} \in \mathbb{Z}^N} \left| \hat{G}(\vec{I}, \vec{n}) \hat{f}_0(\vec{I}, -\vec{n}) \right| d\vec{I} < \infty, \tag{33}$$

so $\langle G \rangle_t$ exists and is bounded for all $t \in \mathbb{R}$. It also follows from (33) that we may apply Fubini's theorem to interchange the summation and integral in (32), so that

$$\langle G \rangle_t = (2\pi)^N \sum_{\vec{n} \in \mathbb{Z}^N} \int_\Omega \hat{G}(\vec{I}, \vec{n}) \hat{f}_0(\vec{I}, -\vec{n}) e^{it\vec{n} \cdot \vec{\omega}(\vec{I})} d\vec{I}. \tag{34}$$

Again by (33), we may apply the dominated convergence theorem to conclude that:

$$\lim_{t \to \infty} \langle G \rangle_t = (2\pi)^N \sum_{\vec{n} \in \mathbb{Z}^N} \lim_{t \to \infty} \int_\Omega \hat{G}(\vec{I}, \vec{n}) \hat{f}_0(\vec{I}, -\vec{n}) e^{it\vec{n} \cdot \vec{\omega}(\vec{I})} d\vec{I}, \tag{35}$$

provided the limits appearing on the right-hand side of (35) exist for each $\vec{n} \in \mathbb{Z}^N$. It therefore remains to study the asymptotic behavior as $t \to \infty$ of the integrals

$$\int_\Omega \hat{G}(\vec{I}, \vec{n}) \hat{f}_0(\vec{I}, -\vec{n}) e^{it\vec{n} \cdot \vec{\omega}(\vec{I})} d\vec{I}. \tag{36}$$

By Lemma 1, the function $\hat{G}(\cdot, \vec{n}) \hat{f}_0(\cdot, -\vec{n})$ appearing in (36) is continuous on Ω. Furthermore, since f_0 has compact support in $\Omega \times \mathbb{T}^N$, this function has compact support in Ω, and therefore lies in $L^1(\Omega)$. Consider the exponent appearing in (36) for a single $\vec{n} \neq 0$. We have:

$$D(\vec{n} \cdot \vec{\omega}(\vec{I})) = n_1 D\omega_1(\vec{I}) + \ldots + n_N D\omega_N(\vec{I}). \tag{37}$$

Since $\{D\omega_1(\vec{I}), \ldots, D\omega_N(\vec{I})\}$ is linearly independent by hypothesis, it follows that $D(\vec{n} \cdot \vec{\omega}(\vec{I})) \neq 0$. It then follows from Lemma 0 that the integral (36) must vanish as $t \to \infty$ when $\vec{n} \neq 0$. Thus, only the $\vec{n} = 0$ contribution survives in this limit and (35) gives:

$$\lim_{t \to \infty} \langle G \rangle_t = (2\pi)^N \int_\Omega \hat{G}(\vec{I}, 0) \hat{f}_0(\vec{I}, 0) d\vec{I}. \tag{38}$$

But

$$\hat{G}(\vec{I}, 0) = \frac{1}{(2\pi)^N} \int_\Omega G(\vec{I}, \vec{\phi}) d\vec{\phi} = \bar{G}(\vec{I}), \tag{39}$$

so

$$\lim_{t \to \infty} \langle G \rangle_t = (2\pi)^N \int_\Omega \bar{G}(\vec{I}) \hat{f}_0(\vec{I}, 0) d\vec{I} = \int_\Omega \bar{G}(\vec{I}) f_0(\vec{I}, \vec{\phi}) d\vec{\phi} d\vec{I} = \langle \bar{G} \rangle_0, \tag{40}$$

which gives (10). This proves the theorem. □

The hypotheses $f_0 \in C_c(\Omega \times \mathbb{T}^N)$ and $G \in C(\Omega \times \mathbb{T}^N)$ stated in the previous theorem can be relaxed in various ways. The following formulation will play a central role in the following section.

Theorem 2: In the system (2), suppose $\omega \in C^2(\Omega, \mathbb{R}^N)$ has no critical points. Suppose the initial condition is described by a probability density $f_0 \in L^1(\Omega \times \mathbb{T}^N)$. For any bounded continuous function $G \in C_b(\Omega \times \mathbb{T}^N)$, the limit (10) holds.

Proof: This follows directly from Theorem 1 by a density argument. Suppose $f_0 \in L^1(\Omega \times \mathbb{T}^N)$. Since $C_c(\Omega \times \mathbb{T}^N)$ is dense in $L^1(\Omega \times \mathbb{T}^N)$, there exists a sequence $\{f_0^{(n)}\} \subset C_c(\Omega \times \mathbb{T}^N)$ such that

$$||f_0 - f_0^{(n)}||_1 \to 0 \quad \text{as} \quad n \to \infty. \tag{41}$$

Put

$$I(t) = \int_{\Omega \times \mathbb{T}^N} \left[G(\vec{I}, \vec{\phi}) - \bar{G}(\vec{I})\right] f_t(\vec{I}, \vec{\phi}) d\vec{I} d\vec{\phi}, \qquad I_n(t) = \int_{\Omega \times \mathbb{T}^N} \left[G(\vec{I}, \vec{\phi}) - \bar{G}(\vec{I})\right] f_t^{(n)}(\vec{I}, \vec{\phi}) d\vec{I} d\vec{\phi}. \tag{42}$$

Suppose $|G(\vec{I}, \vec{\phi})| \leq K$ for some $K > 0$. Using the triangle inequality, we see for each $(\vec{I}, \vec{\phi}) \in \Omega \times \mathbb{T}^N$ that

$$|G(\vec{I}, \vec{\phi}) - \bar{G}(\vec{I})| \leq |G(\vec{I}, \vec{\phi})| + |\bar{G}(\vec{I})| \leq 2K. \tag{43}$$

Fix $\epsilon > 0$. By (41), there exists $n_0 > 0$ such that

$$||f_0 - f_0^{(n_0)}||_1 < \epsilon. \tag{44}$$

Since Theorem 1 applies using G and $f_0^{(n_0)}$, we have:

$$\lim_{t \to \infty} I_{n_0}(t) = 0. \tag{45}$$

Thus, there exists $t_0 > 0$ such that for all $t > t_0$, $|I_{n_0}(t)| < \epsilon$. Combining the triangle inequality with (42-43) gives:

$$|I(t)| \leq |I_{n_0}(t)| + |I(t) - I_{n_0}(t)| \leq |I_{n_0}(t)| + 2K \left(||f_t - f_t^{(n_0)}||_1\right). \tag{46}$$

A change of variables using the map Ψ_t shows that

$$||f_t - f_t^{(n_0)}||_1 = ||f_0 - f_0^{(n_0)}||_1 < \epsilon. \tag{47}$$

Therefore, it follows from (46) that:

$$|I(t)| < \epsilon(1 + 2K). \tag{48}$$

This proves that $I(t) \to 0$ as $t \to \infty$, and the conclusion of Theorem 1 holds. This proves the theorem. □

When the domain Ω is unbounded, it is useful to consider observables G that may be unbounded. In this case, convergence can be proved by imposing sufficient decay and growth conditions on the functions f_0 and G, respectively. As an example, the following result follows easily from Theorem 2.

Theorem 3: In the system (2), suppose $\omega \in C^2(\Omega, \mathbb{R}^N)$ has no critical points. Suppose the initial condition is described by a probability density $f_0 \in L^1(\Omega \times \mathbb{T}^N)$, and let $G \in C(\Omega \times \mathbb{T}^N)$. If there exists $h \in C(\Omega)$ such that:

$$\langle h \rangle_0 < \infty, \qquad |G(\vec{I}, \vec{\phi})| < h(\vec{I}), \qquad (\vec{I} \in \Omega, \quad \vec{\phi} \in \mathbb{T}^N) \tag{49}$$

then the limit (10) holds.

Proof: It follows from (49) that $h > 0$. Put $R = \langle h \rangle_0 > 0$. Define the two functions:

$$f_0' = \frac{hf_0}{R}, \qquad G' = \frac{G}{h}. \tag{50}$$

Since h is continuous and nonvanishing, it follows from (49) that $G' \in C_b(\Omega \times \mathbb{T}^N)$. Note that $f_0' \geq 0$ and

$$\int_{\Omega\times\mathbb{T}^N} f_0'(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi} = \frac{1}{R}\int_{\Omega\times\mathbb{T}^N} h(\vec{I})f_0(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi} = \langle h \rangle_0/R = 1, \tag{51}$$

so $f_0' \in L^1(\Omega \times \mathbb{T}^N)$ defines a probability density. Note that we have:

$$\langle G \rangle_t = \int_{\Omega\times\mathbb{T}^N} G(\vec{I},\vec{\phi}+t\omega(\vec{I}))f_0(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi} = R\int_{\Omega\times\mathbb{T}^N} G'(\vec{I},\vec{\phi}+t\omega(\vec{I}))f_0'(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi}. \tag{52}$$

Applying Theorem 2 to the rightmost integral of (52) and noting that $\bar{G}' = \bar{G}/h$ gives:

$$\lim_{t\to\infty}\langle G \rangle_t = R\int_{\Omega\times\mathbb{T}^N} h(\vec{I})^{-1}\bar{G}(\vec{I})f_0'(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi} = \int_{\Omega\times\mathbb{T}} \bar{G}(\vec{I},\vec{\phi})f_0(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi}, \tag{53}$$

as claimed. □

For example, suppose G has the growth bound $|G| \leq K||\vec{I}||^m$ for constants $m,\ K > 0$, where $||\cdot||$ denotes the Euclidean norm on $\mathbb{R}^M$. Then it is sufficient to consider the majorizing function $h(\vec{I}) = 1 + K||\vec{I}||^m$.

We now consider the role played by the critical points of the frequency map $\omega : \Omega \to \mathbb{R}^N$. The following theorem shows that the set of critical points poses no difficulty if this set is sufficiently small.

Theorem 4: In the random initial value problem (2), let $\omega \in C^2(\Omega, \mathbb{R}^N)$ and suppose that the set of critical points of ω has Lebesgue measure zero in $\mathbb{R}^M$. Suppose that the pair (f_0, G) satisfies the hypotheses of any of Theorems 1-3. Then the limit (10) holds.

Proof: Let Ω_R and Ω_C denote the set of regular points and the set of critical points of the map ω, respectively. Then for all $t \in \mathbb{R}$ we have:

$$\langle G \rangle_t = \int_{\Omega_R\times\mathbb{T}^N} f_t(\vec{I},\vec{\phi})G(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi} + \int_{\Omega_C\times\mathbb{T}^N} f_t(\vec{I},\vec{\phi})G(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi}. \tag{54}$$

Since $m(\Omega_C) = 0$ by hypothesis, we have:

$$m(\Omega_C \times \mathbb{T}^N) = m(\Omega_C)m(\mathbb{T}^N) = 0. \tag{55}$$

It follows from (55) that the rightmost integral in (54) vanishes for all $t \in \mathbb{R}$. Now define the map $J : \Omega \to \mathbb{R}$ by:

$$J(\vec{I}) = \det\left[(D\omega(\vec{I}))^T(D\omega(\vec{I}))\right] \qquad (\vec{I} \in \Omega). \tag{56}$$

Note that $\vec{I} \in \Omega$ is a critical point of the map ω if and only if $J(\vec{I}) = 0$. Since J is continuous, it follows that Ω_R is open in $\mathbb{R}^M$. We may therefore apply Theorems 1-3 to the functions f_0 and G restricted to the smaller domain $\Omega_R \times \mathbb{T}^N \subset \Omega \times \mathbb{R}^N$ to conclude that

$$\lim_{t\to\infty}\int_{\Omega_R\times\mathbb{T}^N} f_t(\vec{I},\vec{\phi})G(\vec{I},\vec{\phi})d\vec{I}d\vec{\phi} = \int_{\Omega_R\times\mathbb{T}^N} f_0(\vec{I},\vec{\phi})\bar{G}(\vec{I})d\vec{I}d\vec{\phi}, \tag{57}$$

where $\bar{G}$ is the angle average of G, defined by (8). Combining (54) and (57), it follows that

$$\lim_{t\to\infty}\langle G \rangle_t = \int_{\Omega_R\times\mathbb{T}^N} f_0(\vec{I},\vec{\phi})\bar{G}(\vec{I})d\vec{I}d\vec{\phi} = \int_{\Omega\times\mathbb{T}^N} f_0(\vec{I},\vec{\phi})\bar{G}(\vec{I})d\vec{I}d\vec{\phi}, \tag{58}$$

where the second equality follows by again applying (55). This proves the theorem. □

The following two examples illustrate that the conclusion (10) guaranteed in Theorems 1-4 may fail if the set of critical points of the frequency map $\omega : \Omega \to \mathbb{R}^N$ is too large. In the simplest case, suppose that ω is constant, with

$$\omega(\vec{I}) = \vec{\nu}, \qquad (\vec{I} \in \Omega, \quad \vec{\nu} \in \mathbb{R}^N). \tag{59}$$

Then $D\omega = 0$, and so every point $\vec{p} \in \Omega$ is a critical point of ω. Suppose $f_0 \in L^1(\Omega \times \mathbb{T}^N)$, and take $G(\vec{I}, \vec{\phi}) = e^{i\vec{n}\cdot\vec{\phi}}$ for some nonzero $\vec{n} \in \mathbb{Z}^N$. Then $G \in C_b(\Omega \times \mathbb{T}^N)$ and $\bar{G} = 0$, but:

$$\langle G \rangle_t = \langle G \circ \Psi_t \rangle_0 = \langle e^{i\vec{n}\cdot(\vec{\phi}+t\vec{\nu})} \rangle_0 = e^{it\vec{n}\cdot\vec{\nu}} \langle G \rangle_0. \tag{60}$$

For any initial density f_0 chosen with:

$$\langle G \rangle_0 = \int_{\Omega \times \mathbb{T}^N} f_0(\vec{I}, \vec{\phi}) e^{i\vec{n}\cdot\vec{\phi}} d\vec{I} d\vec{\phi} \neq 0, \tag{61}$$

it follows that $\lim_{t\to\infty} \langle G \rangle_t$ does not exist unless $\vec{\nu} = 0$, in which case:

$$\lim_{t\to\infty} \langle G \rangle_t = \langle G \rangle_0 \neq 0 = \langle \bar{G} \rangle_0, \tag{62}$$

and the conclusion (10) fails. This can also occur for nonconstant ω. To see this, put $M = 1$, $N = 2$, $\Omega = \mathbb{R}$ and

$$\omega_1(I) = \omega_2(I) = I, \qquad (I \in \mathbb{R}). \tag{63}$$

Then $D\omega = (1, 1)^T$, so $\text{rank}(D\omega) = 1 < 2$, and every point of $\mathbb{R}$ is a critical point of ω. Put $G(I, \vec{\phi}) = e^{i\vec{n}\cdot\vec{\phi}}$ with $\vec{n} = (1, -1)^T$. Then $G \in C_b(\Omega \times \mathbb{T}^N)$ and $\bar{G} = 0$, but since $\vec{n} \cdot \omega(I) = (n_1 + n_2)I = 0$ we have:

$$\langle G \rangle_t = \langle G \circ \Psi_t \rangle_0 = \langle e^{i\vec{n}\cdot(\vec{\phi}+t\omega(I))} \rangle_0 = \langle e^{i\vec{n}\cdot\vec{\phi}} \rangle_0 = \langle G \rangle_0. \tag{64}$$

For any initial density f_0 satisfying (61), we again have (62), and the conclusion (10) fails.

IV. WEAK CONVERGENCE TO EQUILIBRIUM

In this section, we study the behavior as $t \to \infty$ of the probability density f_t describing the orbit $\Psi_t(\vec{I}_0, \vec{\phi}_0)$ of the initial value problem (2) with random initial condition $(\vec{I}_0, \vec{\phi}_0) \in \Omega \times \mathbb{T}^N$. Our main result appears as Theorem 5, which is most naturally expressed in terms of the corresponding probability measure P_t defined in (6).

Given a probability density $f_0 \in L^1(\Omega \times \mathbb{T}^N)$, we define the *angle average* of f_0 by:

$$\bar{f}_0(\vec{I}, \vec{\phi}) = \frac{1}{(2\pi)^N} \int_{\mathbb{T}^N} f_0(\vec{I}, \vec{\phi}') d\vec{\phi}' \qquad \vec{I} \in \Omega, \quad \vec{\phi} \in \mathbb{T}^N. \tag{65}$$

Note that $\bar{f}_0$ defines a probability density on $\Omega \times \mathbb{T}^N$ since $\bar{f}_0 \geq 0$ and by Fubini's theorem:

$$\int_{\Omega \times \mathbb{T}^N} \bar{f}_0(\vec{I}, \vec{\phi}) d\vec{I} d\vec{\phi} = \frac{1}{(2\pi)^N} \int_{\mathbb{T}^N} \left(\int_{\Omega \times \mathbb{T}^N} f_0(\vec{I}, \vec{\phi}') d\vec{I} d\vec{\phi}' \right) d\vec{\phi} = 1. \tag{66}$$

In addition, $\bar{f}_0$ is *invariant* under the flow Ψ_t associated with the system (2): for all $t \in \mathbb{R}$,

$$\bar{f}_0 \circ \Psi_t^{-1} = \bar{f}_0. \tag{67}$$

If $f_0 \in C_c(\Omega \times \mathbb{T}^N)$, Theorem 1 implies that for every $G \in C(\Omega \times \mathbb{T}^N)$:

$$\lim_{t\to\infty} \int_{\Omega\times\mathbb{T}^N} G(\vec{I},\vec{\phi}) f_t(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi} = \langle \bar{G} \rangle_0 = \int_{\Omega\times\mathbb{T}^N} G(\vec{I},\vec{\phi}) \bar{f}_0(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi}. \tag{68}$$

This suggests that $\bar{f}_0$ approximates f_t when t is large. We explore in what sense, if any, we have:

$$f_t \to \bar{f}_0 \quad \text{as} \quad t \to \infty. \tag{69}$$

First, we give a counterexample to demonstrate that (69) does *not* hold pointwise or in any of the standard L^p spaces.

Consider the following example for $M = N = 1$, with $\Omega = \mathbb{R}$:

$$\omega(I) = I, \quad f_0(I,\phi) = \frac{1}{\pi}\cos^2(\phi) B(I), \quad (I,\phi) \in \mathbb{R}\times\mathbb{T} \tag{70}$$

where B is any continuous, nonnegative function with compact support on $\mathbb{R}$, normalized so that:

$$\int_{-\infty}^{\infty} B(I) dI = 1. \tag{71}$$

Then $f_0 \in C_c(\mathbb{R}\times\mathbb{T})$ defines a probability density on $\mathbb{R}\times\mathbb{T}$, and

$$f_t(I,\phi) = f_0(I,\phi - tI) = \frac{1}{\pi}\cos^2(\phi - tI) B(I), \qquad \bar{f}_0(I,\phi) = \frac{1}{2\pi} B(I). \tag{72}$$

Clearly, at any point $(I,\phi) \in \mathbb{R}\times\mathbb{T}$ with $I \neq 0$ and $B(I) \neq 0$ we have that:

$$f_t(I,\phi) \nrightarrow \bar{f}_0(I,\phi) \quad \text{as} \quad t \to \infty. \tag{73}$$

In addition, this convergence fails to hold in every $L^p(\mathbb{R}\times\mathbb{T})$, $1 \leq p \leq \infty$. In fact, we have:

$$f_t(I,\phi) - \bar{f}_0(I,\phi) = \cos(2\phi - 2tI)\bar{f}_0(I,\phi), \tag{74}$$

so that for every $t \in \mathbb{R}$, using the fact that $\bar{f}_0$ is independent of the angle ϕ:

$$||f_t - \bar{f}_0||_\infty = \sup_{(I,\phi)\in\mathbb{R}\times\mathbb{T}} |\cos(2\phi - 2tI)|\, \bar{f}_0(I,\phi) = ||\bar{f}_0||_\infty > 0. \tag{75}$$

Likewise, since

$$\int_0^{2\pi} |\cos(2\phi - 2tI)|^p \, d\phi = 8 \int_0^{\pi/4} \cos^p(2\phi) d\phi = 2\pi \, (M_p)^p \tag{76}$$

for constants $M_p > 0$ ($p = 1, 2, \ldots$), independent of t, we have

$$||f_t - \bar{f}_0||_p = \left(\int_{-\infty}^{\infty} dI \left|\bar{f}_0(I,\phi)\right|^p \int_0^{2\pi} \frac{d\phi}{2\pi} |\cos(2\phi - 2tI)|^p \right)^{1/p} = M_p ||\bar{f}_0||_p > 0. \tag{77}$$

Both (75) and (77) hold independently of t, so convergence fails in every $L^p(\mathbb{R}\times\mathbb{T})$.

Below is the central result that we wish to prove regarding the behavior of the probability density f_t as $t \to \infty$. We first recall the following definitions [15].

Definitions:

- Let X be a metric space, and let $\mathcal{B}(X)$ denote the σ-algebra of Borel subsets of X. A sequence $\{P_n\}$ of probability measures defined on the measurable space $(X, \mathcal{B}(X))$ is said to *converge weakly* to a probability measure P, also defined on $(X, \mathcal{B}(X))$, if for any $g \in C_b(X)$ we have:

$$\lim_{n\to\infty} \int_X g dP_n = \int_X g dP. \tag{78}$$

In this case, we write $P_n \Rightarrow P$.

- Let P be a probability measure on $(X, \mathcal{B}(X))$. A $P-$*continuity set* is a set $A \in \mathcal{B}(X)$ such that $P(\partial A) = 0$, where ∂A denotes the boundary of A.

If a weak limit exists, it is unique: a sequence of probability measures cannot converge weakly to two distinct limits.

In the special case that $X = \Omega \times \mathbb{R}^N$, if P_n and P have densities f_n and f, respectively, then the condition (78) is that $P_n \Rightarrow P$ if and only if:

$$\lim_{n\to\infty} \int_{\Omega\times\mathbb{T}^N} g(\vec{I},\vec{\phi}) f_n(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi} = \int_{\Omega\times\mathbb{T}^N} g(\vec{I},\vec{\phi}) f(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi} \quad \text{for all} \quad g \in C_b(\Omega \times \mathbb{R}^N). \tag{79}$$

Theorem 5: In the system (2), let $\omega \in C^2(\Omega, \mathbb{R}^N)$ and suppose that the set of critical points of ω has Lebesgue measure zero in $\mathbb{R}^M$. Suppose that the random initial condition is described by a probability density $f_0 \in L^1(\Omega \times \mathbb{T}^N)$, and let f_t and $\bar{f}_0$ denote the probability densities defined in (4) and (65). Define corresponding measures P_t and P_{eq}, given by:

$$P_t(A) = \int_A f_t(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi}, \qquad P_{eq}(A) = \int_A \bar{f}_0(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi}, \qquad A \in \mathcal{B}(\Omega \times \mathbb{R}^N). \tag{80}$$

Then $P_t \Rightarrow P_{eq}$ as $t \to \infty$.

Proof: If $G \in C_b(\Omega \times \mathbb{T}^N)$, then Theorem 3 implies that as $t \to \infty$:

$$\int_{\Omega\times\mathbb{T}^N} G(\vec{I},\vec{\phi}) f_t(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi} = \langle G \rangle_t \to \langle \bar{G} \rangle_0. \tag{81}$$

It then follows from Fubini's theorem that:

$$\langle \bar{G} \rangle_0 = \int_{\Omega\times\mathbb{T}^N} \bar{G}(\vec{I}) f_0(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi} = \int_{\Omega\times\mathbb{T}^N} G(\vec{I},\vec{\phi}) \bar{f}_0(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi}. \tag{82}$$

Since this convergence holds for arbitrary $G \in C_b(\Omega \times \mathbb{T}^N)$, it follows that the probability measures P_t converge weakly to the probability measure P_{eq} as $t \to \infty$, as claimed. $\square$

Corollary: Suppose that the hypotheses of Theorem 5 hold. Consider any set $A \in \mathcal{B}(\Omega \times \mathbb{T}^N)$ whose boundary has measure zero, $m(\partial A) = 0$. Then:

$$\lim_{t\to\infty} P_t(A) = P_{eq}(A). \tag{83}$$

Proof: By Theorem 5, $P_t \Rightarrow P_{eq}$. Since P_{eq} is described by the density $\bar{f}_0 \in L^1(\Omega \times \mathbb{T}^N)$, it follows that $P_{eq}(\partial A) = 0$, so A is a $P_{eq}-$continuity set. The conclusion now follows by the Portmanteau theorem (Theorem 2.1 of [15]).

Informally, (83) states that for all sufficiently large t, the probability that the orbit $\Psi_t(\vec{I}_0, \vec{\phi}_0)$ lies in the set $A \subset \Omega \times \mathbb{T}^N$ is well-approximated by the quantity:

$$P_{eq}(A) = \int_A \bar{f}_0(\vec{I},\vec{\phi}) d\vec{I} d\vec{\phi}. \tag{84}$$

We have assumed that our random initial condition is modeled by a probability measure P_0 represented by a density on $\Omega \times \mathbb{T}^N$. (That is, P_0 is absolutely continuous with respect to the natural measure on $\Omega \times \mathbb{T}^N$.) If this is not the case, then the conclusion of Theorem 4 may fail. For example, let $M = N = 1$, $\Omega = \mathbb{R}$ and let $\omega : \mathbb{R} \to \mathbb{R}$ be the frequency map of (70). Let $\delta_{(I_0,\phi_0)}$ denote the unit mass concentrated at a point $(I_0, \phi_0) \in \Omega \times \mathbb{T}^1$, and let P_0 denote the probability measure on $\Omega \times \mathbb{T}^1$ given by:

$$P_0 = \delta_{(I_0,\phi_0)}, \qquad (I_0 \neq 0). \tag{85}$$

The probability that the orbit at time $t \in \mathbb{R}$ lies in a set A is modeled by the pushforward of P_0 under the flow Ψ_t of (2), defined in general by:

$$P_t(A) = P_0(\Psi_t^{-1}(A)), \qquad A \in \mathcal{B}(\Omega \times \mathbb{T}^N). \tag{86}$$

For this example, P_t is given explicitly by:

$$P_t = \delta_{(I_0,\phi_0+tI_0)}, \tag{87}$$

and for the function $G \in C_b(\Omega \times \mathbb{T}^1)$ given by $G(I, \phi) = e^{i\phi}$ we have:

$$\langle G \rangle_t = \int_{\Omega\times\mathbb{T}^1} G dP_t = e^{i(\phi_0+tI_0)}. \tag{88}$$

Since $I_0 \neq 0$, (88) fails to converge as $t \to \infty$. It follows that P_t does not converge weakly to any limiting probability measure P_{eq}.

V. A NUMERICAL EXAMPLE

In this section, we consider an exactly-soluble example in one degree of freedom ($M = N = 1$) that illustrates the ideas previously described. Let our symplectic manifold be $\mathcal{M} = \mathbb{R}^2$ with the standard coordinates (q, p) and the standard symplectic form $\Lambda = dq \wedge dp$, and consider the Hamiltonian given by:

$$H^{\mathcal{M}}(q, p) = \psi g + \frac{1}{2}\alpha g^2, \quad g = \frac{1}{2}(q^2 + p^2), \qquad (\psi > 0, \quad \alpha > 0). \tag{89}$$

Then $H^{\mathcal{M}}$ has one stable equilibrium point at the origin, where $dH^{\mathcal{M}} = 0$ (see footnote [21]), and the level sets defined by $H^{\mathcal{M}} = h$ are circles centered at the origin for all $h > 0$. Consider an initial probability density on $\mathcal{M}$ of the form:

$$f_0^{\mathcal{M}}(q, p) = \frac{1}{2\pi\epsilon_0} \exp\left(-\frac{(q - q_0)^2 + p^2}{2\epsilon_0}\right), \qquad (\epsilon_0 > 0, \quad q_0 > 0). \tag{90}$$

Figure 1 illustrates a set of 100K initial conditions randomly sampled from the density (90), together with the evolution of these initial conditions under the flow of (89). The visible filamentation of the density with increasing t is characteristic of nonlinear Hamiltonian flows, and we are interested in the limiting behavior of the ensemble of orbits as $t \to \infty$.

Let $\Omega = \{x \in \mathbb{R} : x > 0\}$. A canonical transformation to action-angle variables $\chi : \mathbb{R}^2 \setminus \{0\} \to \Omega \times \mathbb{T}^1$ is given by $\chi(q, p) = (I, \phi)$, with:

$$q + ip = \sqrt{2I} e^{-i\phi}, \qquad (\text{ so } dq \wedge dp = d\phi \wedge dI \ ). \tag{91}$$

The Hamiltonian (89) and the initial density (90) take the following forms in action-angle variables, for $(I, \phi) \in \Omega \times \mathbb{T}^1$:

$$H(I) = \psi I + \frac{1}{2}\alpha I^2, \qquad f_0(I, \phi) = \frac{1}{2\pi\epsilon_0} e^{-I/\epsilon_0} e^{-q_0^2/2\epsilon_0} \exp\left(\frac{q_0}{\epsilon_0}\sqrt{2I}\cos\phi\right). \tag{92}$$

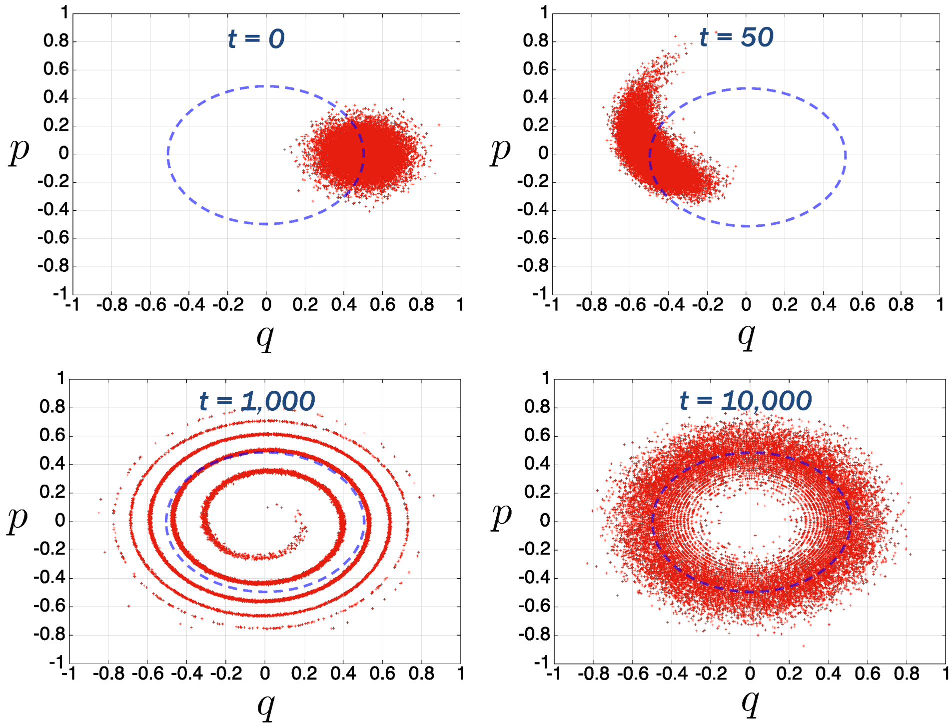

FIG. 1. An ensemble of 100K initial conditions sampled from the density (90) for parameters ϵ_0=0.01, $q_0 = 0.5$ is tracked under the Hamiltonian flow of (89) with ψ=0.3, α=0.1. The dashed circle illustrates the invariant torus defined by the level set where $H^M = 49/1280$.

The frequency map $\omega : \Omega \to \mathbb{R}$ is given by $\omega(I) = \partial H/\partial I = \psi + \alpha I$. Note that ω has no critical points since:

$$\det(D\omega(I)) = \frac{\partial \omega(I)}{\partial I} = \alpha \neq 0. \tag{93}$$

We first explore the evolution of the ensemble centroid, given by the pair of expected values $(\langle q\rangle_t, \langle p\rangle_t)$ for $t \in \mathbb{R}$. To do this, we take as our observable the function:

$$G = q + ip = \sqrt{2I}e^{-i\phi}, \qquad (I, \phi) \in \Omega \times \mathbb{T}^1. \tag{94}$$

The angle average $\bar{G}$ of the function G is given by:

$$\bar{G}(I) = \frac{1}{2\pi}\int_0^{2\pi} G(I,\phi)d\phi = \frac{1}{2\pi}\int_0^{2\pi} \sqrt{2I}e^{-i\phi}d\phi = 0. \tag{95}$$

We prove that the hypotheses of Theorem 3 hold. To see this, define $h : \Omega \to \mathbb{R}$ by:

$$h(I) = 1 + \sqrt{2I}, \qquad (I \in \Omega). \tag{96}$$

Clearly h is continuous on Ω and $|G| < h$. The expected value of h is given by:

$$\langle h\rangle_0 = \langle 1 + \sqrt{2I}\rangle_0 = 1 + \int_0^{2\pi}\int_0^{\infty} \sqrt{2I} f_0(I,\phi)dId\phi. \tag{97}$$

It is possible to show that the density in (92) satisfies the bound:

$$f_0(I,\phi) \leq \frac{1}{2\pi\epsilon_0}\exp\left(-\frac{q_0}{\epsilon_0}\sqrt{2I}\right), \qquad (I > 8q_0^2, \quad \phi \in \mathbb{T}^1). \tag{98}$$

Therefore,

$$\int_0^{2\pi}\int_{8q_0^2}^{\infty}\sqrt{2I}f_0(I,\phi)dId\phi \le \frac{1}{\epsilon_0}\int_0^{\infty}\sqrt{2I}\exp\left(-\frac{q_0}{\epsilon_0}\sqrt{2I}\right)dI = \frac{2\epsilon_0^2}{q_0^3} < \infty, \tag{99}$$

so $\langle h\rangle_0 < \infty$, and the hypotheses of Theorem 3 hold. The expected value of G at $t \in \mathbb{R}$ is given by:

$$\langle q+ip\rangle_t = \langle G\circ\Psi_t\rangle_0 = \int_0^{2\pi}\int_0^{\infty}\sqrt{2I}e^{-i(\phi+\omega(I)t)}f_0(I,\phi)dId\phi. \tag{100}$$

Using (92) and evaluating the integral in (100) explicitly gives:

$$\langle q+ip\rangle_t = -\frac{q_0}{(t\alpha\epsilon_0 - i)^2}\exp\left(-\frac{q_0^2 t\alpha}{2(t\alpha\epsilon_0 - i)} - it\psi\right). \tag{101}$$

In the limit $\alpha \to 0$, the Hamiltonian flow of (89) becomes linear, and (101) takes the simple form of a rotation $q_0\exp(-it\psi)$ of angular frequency ψ about the origin. In the general case, this rotation is modulated by a slowly-varying envelope described by the complex modulus of (101):

$$\left|\frac{\langle q+ip\rangle_t}{q_0}\right| = \frac{1}{(1+\tau^2)}\exp\left(-\frac{\tau^2\nu}{1+\tau^2}\right), \qquad \tau = t\alpha\epsilon_0, \qquad \nu = \frac{q_0^2}{2\epsilon_0}. \tag{102}$$

Taking the limit of (101) for large t gives:

$$\lim_{t\to\infty}\langle G\rangle_t = 0 = \langle\bar{G}\rangle_0. \tag{103}$$

This demonstrates that conclusion (10) of Theorem 3 holds, as expected. Figure 2 illustrates the quantity (101), together with the modulation envelope (102) and the centroid values obtained numerically from the ensemble of orbits shown in Fig. 1.

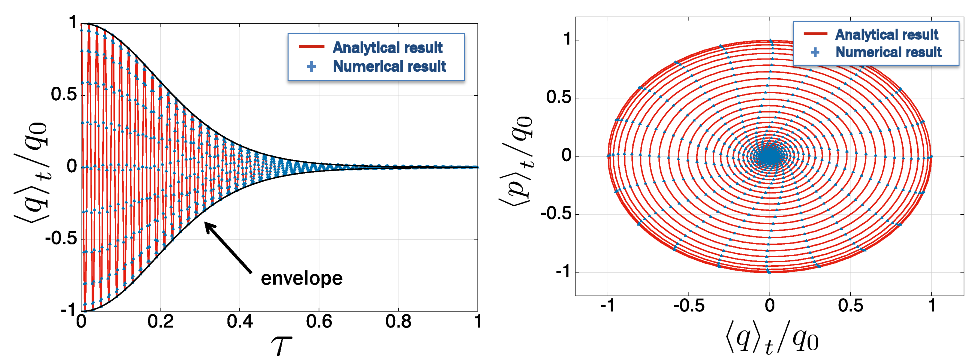

FIG. 2. The evolution of the ensemble centroid (101) is shown together with the modulating envelope (102). The parameters used are identical to those used to produce Fig. 1. The centroid must decay into the origin since the angle average of (94) vanishes.

Using the identity [23]

$$I_n(z) = \frac{1}{\pi}\int_0^{\pi}e^{z\cos\theta}\cos(n\theta)d\theta, \qquad (z\in\mathbb{C}, \quad n\in\mathbb{Z}) \tag{104}$$

where I_n denotes the modified Bessel function of order n, one may obtain the angle-averaged probability density (65) corresponding to the density in (92):

$$\bar{f}_0(I,\phi) = \frac{1}{2\pi\epsilon_0}e^{-I/\epsilon_0}e^{-q_0^2/2\epsilon_0}I_0\left(\frac{q_0}{\epsilon_0}\sqrt{2I}\right). \tag{105}$$

To gain insight into the convergence $f_t \to \bar{f}_0$, it is convenient to express the density in terms of its Fourier modes. To see this, define for each $t \in \mathbb{R}$ the *characteristic function* of the density f_t, given by:

$$b_t(k,n) = \langle e^{in\phi} e^{ikI} \rangle_t = \int_0^{2\pi} \int_0^{\infty} e^{in\phi} e^{ikI} f_t(I,\phi) dI d\phi \qquad (k \in \mathbb{R}, \quad n \in \mathbb{Z}). \tag{106}$$

Clearly, we have the relationship:

$$b_t(k,n) = \langle e^{in(\phi+t\omega)} e^{ikI} \rangle_0 = \langle e^{in\phi} e^{itn\psi} e^{i(k+tn\alpha)I} \rangle_0 = e^{itn\psi} b_0(k+tn\alpha, n), \tag{107}$$

which holds independently of the form of the initial density f_0.

Evaluating (106) at $t=0$ for the density in (92) gives explicitly:

$$b_0(k,n) = \frac{1}{2\mu}\sqrt{\frac{\pi\nu}{\mu}} \exp\left(\frac{\nu}{2\mu} - \nu\right) \left[I_{|n|/2-1/2}\left(\frac{\nu}{2\mu}\right) + I_{|n|/2+1/2}\left(\frac{\nu}{2\mu}\right) \right], \tag{108}$$

where ν was defined in (102) and μ is a complex parameter depending on k given by:

$$\mu = 1 - ik\epsilon_0. \tag{109}$$

Figure 3 illustrates the complex modulus of the function (108) for the values $n = 0, 1, 2, 3, 4$. It follows from (107) that the function $b_t(\cdot, n)$, representing the angular Fourier mode of index n, is simply translated in k to the left at the rate $n\alpha$ as t increases (Fig. 3, right panel). This reflects the fact that the filamented structure of the density visible in Fig. 1 does not disappear, but is shifted to increasingly shorter wavelengths as $t \to \infty$.

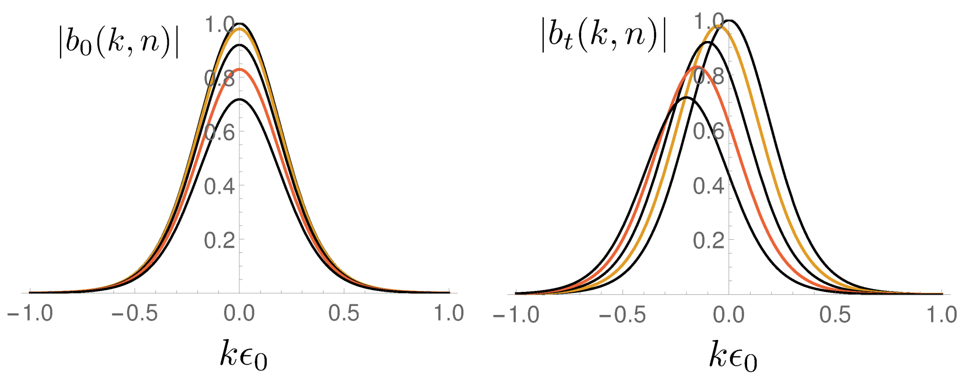

FIG. 3. The complex modulus of the characteristic function b_t of the density $f_t = f_0 \circ \Psi_t^{-1}$ for f_0 given by (92) at $t = 0$ (left) and at $t = 50$ (right) is shown for $n = 0, 1, 2, 3, 4$. The maxima decrease with increasing index n. The parameters used are identical to those used to produce Figs. 1-2. The curves corresponding to $n \neq 0$ are translated to the left with increasing t. In the weak limit $t \to \infty$, only the stationary $n = 0$ mode survives.

The stationary ($n = 0$) mode is given explicitly by:

$$b_0(k,0) = \frac{1}{\mu} \exp\left(\frac{\nu}{\mu} - \nu\right), \tag{110}$$

and one may verify that the characteristic function $\bar{b}_0$ of the angle-averaged density (105) is given by:

$$\bar{b}_0(k,n) = \delta_{n,0} b_0(k,0) \qquad (k \in \mathbb{R}, \quad n \in \mathbb{Z}). \tag{111}$$

Noting from (108) that $|b_0(k,n)| \to 0$ as $|k| \to \infty$ for each $n \in \mathbb{Z}$, one sees from (107) the pointwise convergence:

$$\lim_{t\to\infty} b_t(k,n) = \bar{b}_0(k,n), \qquad (k \in \mathbb{R}, \quad n \in \mathbb{Z}). \tag{112}$$

That is, the weak convergence $P_t \Rightarrow P_{eq}$ is reflected in the pointwise convergence $b_t \to \bar{b}_0$ of the Fourier modes of the probability density f_t.

Finally, let us return to the random initial value problem defined by the Hamiltonian flow of (89) on the original space $\mathcal{M} = \mathbb{R}^2$, described by the probability density $f_0^{\mathcal{M}} \in L^1(\mathbb{R}^2)$. Since weak convergence of probability measures is preserved under continuous mappings, Theorem 5 implies that $P_t^{\mathcal{M}} \Rightarrow P_{eq}^{\mathcal{M}}$, where $P_t^{\mathcal{M}} = P_t \circ \chi$ and $P_{eq}^{\mathcal{M}} = P_{eq} \circ \chi$. (See footnote [22].) For example, let $A \subset \mathbb{R}^2$ denote any Borel set whose boundary has zero measure (e.g., a rectangle). From the Corollary to Theorem 5, it follows that the probability $P_t^{\mathcal{M}}(A)$ that an orbit lies in A at time t converges, as $t \to \infty$, to the limit:

$$P_{eq}^{\mathcal{M}}(A) = P_{eq}(\chi(A)) = \int_{\chi(A)} \bar{f}_0(I, \phi) dI d\phi, \tag{113}$$

where $\bar{f}_0$ is given explicitly by (105). This provides a clear probabilistic picture of the limiting behavior of the system shown in Fig. 1.

VI. CONCLUSIONS

We have investigated the long time behavior (as $t \to \infty$) of the orbits of a dynamical system in the general action-angle form (1), with random initial conditions sampled from a probability density $f_0 \in L^1(\Omega \times \mathbb{T}^N)$. Given an observable G, Theorems 1-4 provide sufficient conditions on the triple (f_0, ω, G) to guarantee that the expected value of G converges to the expected value of its angle average $\bar{G}$ as $t \to \infty$. Likewise, Theorem 5 demonstrates that the probability density f_t describing the orbits at time t converges in an appropriate weak sense to the angle-averaged density $\bar{f}_0$ as $t \to \infty$. The conditions require: 1) that the frequency map $\omega : \Omega \to \mathbb{T}^N$ is of class C^2, and 2) that the critical points of ω form a set of measure zero. The exactly-soluble model in Section 5 provides an illustration of these ideas in the context of an integrable Hamiltonian system.

The results described here can be applied to study the global orbits of an integrable system on a symplectic manifold $\mathcal{M}$ (Appendix B), with random initial conditions sampled from a probability density $f_0 \in L^1(\mathcal{M})$. This is straightforward if the density f_0 is supported within the domain of a single transformation to local action-angle variables. If this is not the case, suppose that $\mathcal{M}$ is covered almost everywhere by a finite set $\{(V_j, \chi_j)\}_{j=1}^n$ of action-angle charts with disjoint open domains V_j ($j = 1, \ldots, n$) and maps $\chi_j : V_j \to \Omega_j \times \mathbb{T}^N$. Suppose that the frequency map $\omega^{(j)} : \Omega_j \to \mathbb{R}^N$ within each chart satisfies the hypotheses of Theorem 5. By applying Theorem 5 to the probability density on $\Omega_j \times \mathbb{T}^N$ given by $f_0^{(j)} = f_0|_{V_j} \circ \chi_j^{-1} / \int_{V_j} f_0 dm$, we obtain the desired weak convergence within each chart $j = 1, \ldots, n$.

The weak convergence to equilibrium guaranteed by Theorem 5 can be quantified by using the bounded Lipschitz metric d_{BL}, a metric characterizing the distance between Borel probability measures on a separable metric space. (See, for example, Section 11.3 of [24].) In particular, $P_n \Rightarrow P$ if and only if $d_{BL}(P_n, P) \to 0$ for a sequence of probability measures $\{P_n\}$. It is natural to seek estimates for the rate of convergence $d_{BL}(P_t, P_{eq}) \to 0$ as $t \to \infty$, expressed in terms of the density f_0 and the frequency map ω. This is a possible topic of future research.

VII. ACKNOWLEDGEMENT

The author thanks James Ellison, Robert Warnock, and Stephen Webb for helpful discussions. This work was supported by the Director, Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and made use of computer resources at the National Energy Research Scientific Computing Center.

APPENDIX A: SUMMARY OF NOTATION

For each nonempty open $\Omega \subset \mathbb{R}^M$, we denote by $C^k(\Omega)$ the space of real-valued functions on Ω with continuous partial derivatives up to and including order k. Similarly, $C^k(\Omega, \mathbb{R}^N)$ will denote the space of vector-valued functions $g : \Omega \to \mathbb{R}^N$ whose N components $(g_1, \ldots, g_N)$ each lie in $C^k(\Omega)$. If $\mathcal{M}$ is a smooth manifold, then $C^\infty(\mathcal{M})$ will denote the space of smooth real-valued functions on $\mathcal{M}$. $C(X)$, $C_b(X)$, and $C_c(X)$ denote the sets of complex-valued functions on a topological space X that are continuous, continuous and bounded, and continuous with compact support, respectively. If $A \subset X$, then $\bar{A}$ and ∂A denote the closure and boundary of A, respectively. Points on the N-torus $\mathbb{T}^N \cong \mathbb{R}^N / 2\pi\mathbb{Z}^N$ are denoted $\vec{\phi} = (\phi_1, \ldots, \phi_N) \mod 2\pi$. Integrals on subsets of $\mathbb{R}^N$ are taken with respect to the standard Lebesgue measure on $\mathbb{R}^N$ unless otherwise specified. Integrals on $\mathbb{T}^N$ are taken with respect to the measure that is induced by Lebesgue measure on $[0, 2\pi]^N$ so that:

$$\int_{\mathbb{T}^N} g(\vec{\phi}) d\vec{\phi} = \int_0^{2\pi} \ldots \int_0^{2\pi} g(\phi_1, \ldots, \phi_N) d\phi_1 \ldots d\phi_N. \tag{114}$$

By $L^p(\Omega \times \mathbb{T}^N)$ $(1 \leq p < \infty)$ we denote the space of Borel measurable functions $g : \Omega \times \mathbb{T}^N \to \mathbb{C}$ such that $||g||_p < \infty$ with the norm:

$$||g||_p = \left(\frac{1}{(2\pi)^N} \int_{\Omega \times \mathbb{T}^N} \left| g(\vec{I}, \vec{\phi}) \right|^p d\vec{I} d\vec{\phi} \right)^{1/p}, \tag{115}$$

and in the case of $L^2(\Omega \times \mathbb{T}^N)$ we define an inner product by:

$$\langle g_1, g_2 \rangle = \frac{1}{(2\pi)^N} \int_{\Omega \times \mathbb{T}^N} g_1(\vec{I}, \vec{\phi}) g_2^*(\vec{I}, \vec{\phi}) d\vec{I} d\vec{\phi}, \tag{116}$$

where * denotes complex conjugation. Similar definitions apply to the spaces $L^p(\Omega)$ and $L^p(\mathbb{T}^N)$. If X, Y, Z are sets and $g : X \times Y \to Z$, we let $g(\cdot, y) : X \to Z$ denote the function mapping $x \mapsto g(x, y)$ for fixed $y \in Y$, with a similar meaning for $g(x, \cdot)$. Finally, if G is a random variable depending on a time parameter t, the notation $\langle G \rangle_t$ will denote the expected value of G taken at time t.

APPENDIX B: INTEGRABLE HAMILTONIAN SYSTEMS

By a *completely integrable Hamiltonian system* with N degrees of freedom, we mean a symplectic manifold $(\mathcal{M}, \Lambda)$ of dimension $2N$, together with functions $H = f_1, f_2, \ldots, f_N \in C^\infty(\mathcal{M})$ such that [16–18]:

1. $\{f_j, f_k\} = 0$ for $j, k = 1, \ldots, N$, where $\{\cdot, \cdot\}$ denotes the Poisson bracket on $(\mathcal{M}, \Lambda)$,
2. The set $\{ df_1|_p, \ldots, df_N|_p\}$ is linearly independent for almost all $p \in \mathcal{M}$, with respect to the symplectic volume measure on $\mathcal{M}$.

The central result concerning such systems is as follows. Consider a single level set of the form:

$$\mathcal{M}_{\vec{c}} = \{p \in \mathcal{M} : f_1(p) = c_1, \ldots, f_N(p) = c_N\} \qquad (\vec{c} \in \mathbb{R}^N). \tag{117}$$

If $\{ df_1|_p, \ldots, df_N|_p\}$ is linearly independent at each $p \in \mathcal{M}_{\vec{c}}$, then $\mathcal{M}_{\vec{c}}$ forms a smooth submanifold of $\mathcal{M}$ invariant under the Hamiltonian flow of H. The theorem of Liouville-Arnold [16–18] states that if $\mathcal{M}_{\vec{c}}$ is compact and connected, then $\mathcal{M}_{\vec{c}}$ is diffeomorphic to the N-torus $\mathbb{T}^N$.

Furthermore, given a sufficiently small neighborhood $V \subset \mathcal{M}$ of such an invariant torus, there exists an open $\Omega \subset \mathbb{R}^N$ and a canonical transformation (see footnote [20]) $\chi : V \to \Omega \times \mathbb{T}^N$ to local action-angle variables $(\vec{I}, \vec{\phi}) \in \Omega \times \mathbb{T}^N$ such that the transformed Hamiltonian H is independent of $\vec{\phi}$ and the orbits therefore satisfy [16–18]:

$$\frac{d\vec{I}}{dt} = -\frac{\partial H(\vec{I})}{\partial \vec{\phi}} = 0, \qquad \frac{d\vec{\phi}}{dt} = \omega(\vec{I}) \equiv \frac{\partial H(\vec{I})}{\partial \vec{I}}, \qquad (\vec{I} \in \Omega, \quad \vec{\phi} \in \mathbb{T}^N). \tag{118}$$

It follows from (118) that the Hamiltonian flow on $\Omega \times \mathbb{T}^N$ is described by the one-parameter family of symplectic maps Ψ_t given by $\Psi_t(\vec{I}, \vec{\phi}) = (\vec{I}, \vec{\phi} + t\omega(\vec{I}))$, $t \in \mathbb{R}$.

The Hamiltonian H in (118) is said to be *nondegenerate* if the frequency map $\omega : \Omega \to \mathbb{R}^N$ satisfies:

$$\det(D\omega) \neq 0, \qquad [D\omega]_{jk} = \frac{\partial \omega_j}{\partial I_k} = \frac{\partial^2 H}{\partial I_j \partial I_k} \quad (j, k = 1, \ldots, N). \tag{119}$$

Note that a nondegenerate Hamiltonian must be nonlinear in the action variables $\vec{I}$.

In particular, the results of this paper apply to systems of the form (118) that satisfy the condition (119) on Ω, except possibly on a set of Lebesgue measure zero.

---

[1] J. Lebowitz and O. Penrose, "Modern ergodic theory," Physics Today **26**, 155-175, February 1973.
[2] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York (1982).
[3] A. V. Bolsinov and A. T. Fomenko, *Integrable Hamiltonian Systems: Geometry, Topology, Classification*, CRC Press, Boca Raton (2004).
[4] M. Audin, *Hamiltonian Systems and Their Integrability*, American Mathematical Society, 2008.
[5] A. Pelayo and San Ngoc, "Symplectic theory of completely integrable Hamiltonian systems," Bulletin of the American Mathematical Society **48**, 409-455, July 2011.
[6] E. Fiorani and G. Sardanashvily, "Global action-angle coordinates for completely integrable systems with non-compact invariant submanifolds," Journal of Mathematical Physics **48**, 032901 (2007).
[7] E. A. Yuzbashyan, "Generalized microcanonical and Gibbs ensembles in classical and quantum integrable dynamics," Annals of Physics **367**, 288-296 (2016).
[8] M. Rigol *et al*, "Relaxation in a Completely Integrable Many-Body Quantum System: An *Ab Initio* Study of the Dynamics of the Highly Excited States of 1D Lattice Hard-Core Bosons," Physical Review Letters **98**, 050405 (2007).
[9] R. Mathew and E. Tiesinga, "Phase-space mixing in dynamically unstable, integrable few-mode quantum systems," Physical Review A **96**, 013604 (2017).
[10] V. Danilov and S. Nagaitsev, "Nonlinear accelerator lattices with one and two analytic invariants," Physical Review Special Topics - Accelerators and Beams **13**, 084002 (2010).
[11] S. Antipov *et al*, "IOTA (Integrable Optics Test Accelerator: facility and experimental beam physics program," Journal of Instrumentation **12**, T03002 (2017).
[12] W. Rudin, *Real and Complex Analysis*, 2nd ed., McGraw-Hill, Inc., New York, 1974.
[13] Y. Katznelson, *An Introduction to Harmonic Analysis*, 3rd ed., Cambridge University Press, Cambridge, UK, 2004.
[14] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993.
[15] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., John Wiley & Sons, Inc., New York, 1999.
[16] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer, New York, 1989.
[17] R. Abraham and J. Marsden, *Foundations of Mechanics*, 2nd ed., Addison-Wesley Publishing Company, Inc., Redwood City, CA, 1978.

[18] J. Moser and E. Zehnder, *Notes on Dynamical Systems*, American Mathematical Society, Courant Institute of Mathematical Sciences, 2005.

[19] P. Lochak and C. Meunier, *Multiphase Averaging for Classical Systems: with Applications to Adiabatic Theorems*, Springer-Verlag, New York, 1988.

[20] More precisely, χ is a symplectomorphism between (V, Λ) and the smooth manifold $\Omega \times \mathbb{T}^N$ with its standard symplectic form $\sum_{j=1}^N d\phi_j \wedge dI_j$. By abuse of notation, we use H to denote both the Hamiltonian on V and the Hamiltonian $H \circ \chi^{-1}$ on $\Omega \times \mathbb{T}^N$.

[21] Note that $H^{\mathcal{M}}$ is in Birkhoff normal form [16] with respect to this equilibrium, and $H^{\mathcal{M}}$ is the simplest Hamiltonian of this type containing terms of degree > 2.

[22] Although the origin in $\mathbb{R}^2$ lies outside the domain of χ, this point is an invariant set of Lebesgue measure zero. Thus, we must have $P_{eq}^{\mathcal{M}}(\{0\}) = P_t^{\mathcal{M}}(\{0\}) = 0$ for all $t \in \mathbb{R}$, and we lose nothing by working on $\mathbb{R}^2 \setminus \{0\}$.

[23] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York (1965).

[24] R. M. Dudley, *Real Analysis and Probability*, 2nd ed., Cambridge University Press, Cambridge, UK (2002).