Journal of Statistical Physics, Vol. 42, Nos. 3/4, 1986

Localization in Disordered, Nonlinear Dynamical Systems

Jürg Fröhlich,[1] Thomas Spencer,[2] and C. Eugene Wayne[3]

Received September 16, 1985

We study localization and wave trapping in disordered, nonlinear dynamical systems. For some models of classical, disordered anharmonic crystal lattices, we prove that, with large probability, there are quasiperiodic lattice vibrations of finite total energy which lie on some infinite-dimensional, compact invariant tori in phase space. Such vibrations remain localized, for all times, and there is no transport of energy through the lattice. Our general concepts and techniques extend to other systems, such as disordered, nonlinear Schrödinger equations, or randomly coupled rotors.

KEY WORDS: Localization; invariant tori; KAM theory; disordered systems.

1. INTRODUCTION: GENERAL IDEAS ON LOCALIZATION IN LINEAR AND NONLINEAR SYSTEMS

Localization is a physical phenomenon observed in the context of wave propagation through disordered media. When the disorder is sufficiently large certain types of waves get trapped, and wave propagation is anomalously slow or impossible. Localization was first analyzed by Anderson.[1] He studied the propagation of the wave functions of electrons in a random array of scatterers, the impurities or defects in a metal. This problem is relevant for a theory of electrical conductivity in disordered metals.

[1] Theoretical Physics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.
[2] Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012.
[3] Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania 16802.

0022-4715/86/0200-0247$05.00/0

In this paper, we analyze a related problem. We study the classical mechanics of a vibrating crystal lattice.[2] We assume that the crystal lattice is disordered, in the sense that masses of and/or couplings between oscillators are random. We study the propagation of vibrations, i.e., sound and heat propagation, through such crystals. In the approximation, where the crystal is harmonic, our problem is mathematically closely related to the one studied by Anderson; see Ref. 3. However, crystals are never strictly harmonic. The main point of our paper is to introduce mathematical concepts and prove some results which we hope will lead toward an understanding of localization in disordered, nonlinear dynamical systems, such as disordered, anharmonic crystals, nonlinear Schrödinger equations with a random potential, etc.

Next, we describe the systems which we wish to study in more detail. At each site j of a crystal lattice, here chosen to be $\mathbb{Z}^d$, a classical oscillator (an atom or ion) is attached. Its configurations are described by vectors $\mathbf{q}_j \in \mathbb{R}^n$. The equations of motion of these oscillators are given by

$$\ddot{\mathbf{q}}_j = -\overset{0}{\omega}{}_j^2 \mathbf{q}_j + \sum_i f_{ij}(\mathbf{q}_i - \mathbf{q}_j) - \frac{\partial U}{\partial \mathbf{q}_j}(\mathbf{q}) \tag{1.1}$$

where f_{ij} and $\partial U/\partial \mathbf{q}_j$ are of short range, e.g.,

$$f_{ij} = 0 = \frac{\partial^2 U}{\partial \mathbf{q}_i \partial \mathbf{q}_j}, \qquad \text{unless } |i-j| = 1 \tag{1.2}$$

The matrices $\overset{0}{\omega}{}_j^2$ and f_{ij} are independent, identically distributed (i.i.d.) $n \times n$ random matrices with smooth distributions of fast decay at infinity which have their support on positive matrices. The most interesting case for crystal physics corresponds to $\overset{0}{\omega}{}_j \equiv 0$. The potential U of the harmonic forces is assumed to be analytic in the position variables $(\mathbf{q}_j)_{j \in \mathbb{Z}^d}$ and satisfy a bound

$$0 \leqslant U(\mathbf{q}) \leqslant 0(|\mathbf{q}|^4) \tag{1.3}$$

as

$$|\mathbf{q}| = \left(\sum_j |\mathbf{q}_j|^2 \right)^{1/2} \to 0$$

For ease of notation, we henceforth set

$$n = 1 \tag{1.4}$$

rather than $n = d (= 2 \text{ or } 3)$ as would be appropriate physically. But our methods and results clearly extend to the more general case described above.

First, we review some of the main results for harmonic crystals, with $U(\mathbf{q}) = 0$. Equation (1.1) may be rewritten, in more compact notation, as follows:

$$\ddot{q}_j = -(\Omega^2 q)_j, \qquad j \in \mathbb{Z}^d \qquad (U = 0) \tag{1.5}$$

where Ω^2 is the Jacobi matrix given by

$$(\Omega^2)_{ij} = \begin{cases} -f_{ij}, & \text{if } |i-j| = 1 \\ \overset{0}{\omega}{}_i^2 + \sum_{k:\,|k-i|=1} f_{ki}, & \text{if } j = i \\ 0, & \text{otherwise} \end{cases} \tag{1.6}$$

Here $f_{ij} = f_{ji}$ and $\overset{0}{\omega}{}_i^2$ are i.i.d. non-negative random variables with smooth distribution of fast decrease at infinity. It follows that Ω^2 is almost surely a positive, self-adjoint operator on $l^2(\mathbb{Z}^d)$, the Hilbert space of square-summable sequences on $\mathbb{Z}^d$. If we define $-\Delta_f$ to be the off-diagonal part of Ω^2, and

$$v(j) \equiv \overset{0}{\omega}{}_j^2 + \sum_{k:\,|k-j|=1} f_{kj} \tag{1.7}$$

then

$$\Omega^2 = H \equiv -\Delta_f + v \tag{1.8}$$

is the Hamiltonian of Anderson's tight binding model describing the motion of a quantum mechanical electron through a crystal lattice, $\mathbb{Z}^d$. This is the system which Anderson considered in Ref. 1. For rigorous results see Refs. 4–6.

Under the assumptions described above, the general solution of the equations of motion (1.5) is given by

$$q(t) = \cos(\Omega t) \cdot q_0 + \sin(\Omega t)\, \Omega^{-1} \cdot p_0 \tag{1.9}$$

where $q = (q_j)_{j \in \mathbb{Z}^d}$, $q_0 = q(0)$, $p = \dot{q}$, $p_0 = \dot{q}(t = 0)$.

We see that the analysis of the time evolution of a disordered harmonic crystal can be reduced to the spectral analysis of the Jacobi matrix Ω^2. Let us therefore review the main results of that analysis.

(a) If $\overset{0}{\omega}{}_j \equiv 0$ (the case of main interest in crystal physics) and the disorder, δ, in the distribution of f_{ij} is large enough (where

$\delta = [\sup(d\phi/df)(\cdot)]^{-1}$, and $d\phi(\cdot)$ is the distribution of f_{ij}), then there exists a finite ω_c^2 such that the spectrum of Ω^2 in the interval (ω_c^2, ∞) is almost surely pure point. More precisely, the spectrum of Ω^2 in (ω_c^2, ∞) consists of simple eigenvalues (the squares of the *normal frequencies* of the crystal) which correspond to *normal modes*, q_j^i, that decay exponentially, as $|j| \to \infty$.[4]

(b) If $\overset{0}{\omega}_j$ has a smooth distribution of sufficiently large disorder, the entire spectrum of Ω^2 is pure point, consisting of exponentially decaying eigenfunctions, $q_{\cdot}^i$. See Refs. 4–6 for proofs.

We remark that, in *one dimension*, arbitrarily weak disorder suffices to arrive at the conclusions in (a) and (b), and $\omega_c^2 = 0$.

In this paper we consider the situation described in (b), but the following remarks also hold if Ω^2 is as in (a), but we restrict the states of the system to belong to the spectral subspace of Ω^2 associated with the interval $[\omega_c^2 + \varepsilon, \infty)$, for some $\varepsilon > 0$. Then we have the following:

(1) Equation (1.5) has a complete, orthonormal system of periodic solutions,

$$\cos(\omega_i t)\, q_j^i, \qquad \sin(\omega_i t)\, q_j^i, \qquad i \in \mathbb{Z}^d \tag{1.10}$$

with periods $2\pi/\omega_i$. The normal modes, q_j^i, are eigenfunctions of Ω^2 localized near $i \in \mathbb{Z}^d$ and decreasing *exponentially*, as $|j| \to \infty$, corresponding to the normal frequencies ω_i. Moreover,

$$\sum_j q_j^i q_j^k = \delta_{ik} \qquad \left(\text{and } \sum_i q_j^i q_l^i = \delta_{jl}\right) \tag{1.11}$$

(2) Every solution, $q(t)$, of (1.5) is a linear combination of the solutions (1.10), i.e.,

$$q_j(t) = \sum_{i \in \mathbb{Z}^d} \left[Q_i \cos(\omega_i t) + \frac{P_i}{\omega_i} \sin(\omega_i t) \right] q_j^i \tag{1.12}$$

for some finite, real coefficients Q_i, P_i.

The conserved total energy of the trajectory $q(t)$ is given by

$$E = \frac{1}{2} \sum_{i \in \mathbb{Z}^d} (P_i^2 + \omega_i^2 Q_i^2) \tag{1.13}$$

[4] For rigorous results see F. Delyon, H. Kunz, and B. Souillard, Ref. 4, for $d = 1$, and Refs. 5 and 6 for $d \geqslant 2$.

We define a torus $\mathbb{T}^\infty(\varepsilon)$ by

$$\mathbb{T}^\infty(\varepsilon) = \{(Q_i, P_i): P_i^2 + \omega_i^2 Q_i^2 \equiv 2\varepsilon_i < \infty, i \in \mathbb{Z}^d\} \tag{1.14}$$

Then *every* $q(t)$, given by (1.12), describes a quasiperiodic motion on an infinite dimensional torus, $\mathbb{T}^\infty(\varepsilon)$, and, for $E < \infty$, $\mathbb{T}^\infty(\varepsilon)$ is *compact* (i.e., essentially "localized" in a finite region of $\mathbb{Z}^d$).

(3) We propose to measure the *spreading of the energy of lattice vibrations*, originally localized near the origin of $\mathbb{Z}^d$, by the quantity

$$R^2(t) \equiv \frac{1}{2} \sum_{j \in \mathbb{Z}^d} |j|^2 \{p_j^2(t) + \sum_i q_i(t)\, \Omega_{ij}^2 q_j(t)\} \tag{1.15}$$

where $p_j(t) \equiv \dot{q}_j(t)$. We may introduce a diffusion constant

$$D(q_0, p_0) \equiv \overline{\lim_{t \to \infty}}\ t^{-1} R^2(t) \tag{1.16}$$

with $q_0 = q(0)$, $p_0 = \dot{q}(0)$. This quantity is analogous to the diffusion constant introduced in Anderson's tight binding model,[1] where it is proportional to the electrical conductivity; see Ref. 5 for results.

More generally, one might study the time-dependence of arbitrary moments of the energy density, $p_j^2(t) + \sum_i q_i(t)\, \Omega_{ij}^2 q_j(t)$, which we denote by $R^{(n)}(t)$, $n = 2, 4, \ldots$; $R^{(2)}(t) \equiv R^2(t)$. For disordered, harmonic crystals, under the hypotheses specified above, one can easily show, using the superposition principle (1.12), that for initial conditions localized in a bounded region, i.e., $q_{0j} = p_{0j} = 0$, for sufficiently large values of $|j|$ ($j \in \mathbb{Z}^d$), all moments $R^{(n)}(t)$ remain uniformly bounded in t, as $|t| \to \infty$, and hence $D(q_0, p_0) = 0$ (absence of diffusion). Thus there is no transport of energy (heat, sound) in such systems.

In nature, crystal lattices are never perfectly harmonic. Realistic equations of motion contain anharmonic forces, i.e., $U(q) \neq 0$, with U as in (1.2), (1.3). As an example of a typical, anharmonic potential used in crystal physics we mention

$$U(q) = \sum_j \lambda_j u_j(q_j) + \sum_{\langle i,j \rangle} f_{\langle i,j \rangle}(q_i - q_j) \tag{1.17}$$

where $u_j(q)$ and $f_{\langle i,j \rangle}(q)$ are bounded by const $|q|^4$, as $|q| \to 0$, $\lambda_j > 0$, and $\langle i, j \rangle$ ranges over nearest-neighbor pairs of $\mathbb{Z}^d$. The equations of motion (1.1), i.e.,

$$\begin{aligned} \dot{p}_j &= -\sum_i \Omega_{ji}^2 q_i - \frac{\partial U}{\partial q_j}(q) \\ \dot{q}_j &= p_j, \qquad j \in \mathbb{Z}^d \end{aligned} \tag{1.18}$$

are Hamiltonian equations of motion derived from the Hamilton function

$$H(q, p) = \frac{1}{2}\left(\sum_j p_j^2 + \sum_{ij} q_i \Omega_{ij}^2 q_j\right) + U(q) \tag{1.19}$$

We are interested in understanding the qualitative behaviour of the solutions of (1.18) when the disorder of Ω^2 is large, specifically in situation (b) described above. We want to know whether the behavior of the solutions of Eq. (1.18) is qualitatively similar to that of the solutions of the *linear* equations of motion (1.5) which we have described in points (1)–(3), above. If there were only finitely many degrees of freedom (replace $\mathbb{Z}^d$ by a finite lattice) we could use well-known theorems to show the following:

(1′) (In the vicinity of an isolated, stable equilibrium point) there are periodic solutions of the Hamiltonian equations of motion derived from a Hamilton function of the form of (1.19).[7]

(2′) If the initial conditions q_0, p_0 are sufficiently small and the eigenvalues of Ω^2 satisfy some nonresonance condition (depending on q_0, p_0) then the solutions of Eqs. (1.18) (for $\mathbb{Z}^d$ replaced by a finite lattice) are *quasiperiodic motions* on *invariant tori*. This can be proven with the help of the Kolmogorov–Arnol'd–Moser (KAM) theory.[8]

The questions of interest to us in this paper are whether results (1′) and (2′) can be extended to infinitely many degrees of freedom, as described above for linear systems. Our *main result* concerns the construction of quasiperiodic motions on *infinite-dimensional, compact tori* for Hamiltonian systems with Hamilton functions related to that in (1.19), for a "large" set of matrices, Ω^2. Our construction is perturbative and is based on an extension of the KAM techniques.

In addition to the questions concerning localized, periodic solutions and compact invariant tori we should ask the physically more important question concerning *transport of energy* in *disordered, classical, nonlinear crystals*, for a large set of localized initial conditions of small total energy and for almost all Ω^2. More precisely, we propose the following:

(3′) Consider solutions of the equations of motion (1.18) of *finite total energy*, with initial conditions localized near the origin. Does the quantity $R^2(t)$, introduced in (1.15), remain bounded in $|t|$ or diverge more slowly than $|t|$, [i.e., $D(q_0, p_0) = 0$, "absence of diffusion"], as $|t| \to \infty$, for almost all Ω^2, provided the disorder is large?

For nonlinear systems, we have no such result which is valid for arbitrary initial conditions of sufficiently small total energy. We will show that, for a class of models similar to those described above, near any localized state there is a state which remains localized for all time, for most choices of Ω^2. The localized states we will construct lie on *infinite-dimen-*

sional, compact invariant tori. However, such tori presumably form a rather sparse subset of (finite-energy) phase space. As in the finite-dimensional case, our KAM theorem makes no statement about the motions corresponding to remaining initial conditions, and, indeed, at least some of these trajectories are expected to drift far away in the phase space (Arnol'd diffusion; Ref. 9) if the number of degrees of freedom in the system exceeds 2.

An important step in understanding these motions in the finite-dimensional case was provided by Nekhoroshev,[9] who showed that, even if these motions do not lie on invariant tori, their rate of drift is so slow that they might well be mistaken for quasiperiodic motions experimentally.

We hope that a similar result will hold in infinite dimensions. More precisely, given an arbitrary localized initial state and an arbitrary random matrix, Ω^2, we would like to show that if the state is not localized for all times, it takes a *very* long time to escape from some large finite region, in the sense that $R^2(t)$ grows *very slowly* in $|t|$.

In the case of one-dimensional systems, our hopes are based on the following considerations. The random matrices, Ω^2, to which our theorem does not apply are excluded because they contain severe resonances. These resonances occur, for instance, when the diagonal part of Ω^2 which we call the "random potential" has nearly identical values at different sites. In elementary quantum mechanics one finds that resonances between the eigenvalue spectra of subsystems localized near two distinct potential wells lead to a much enhanced tunneling rate when the systems are coupled. Similarly, successive resonances between the eigenvalue spectra of matrices, Ω^2, associated with distinct subsystems of oscillators lead to rapid transfer of energy from one subsystem to another when they are coupled. However, as in the Anderson tight binding model,[1,5,6] the spectra of random matrices, Ω^2, associated with finite subsystems will be *resonance* free over arbitrarily large regions localized far away from the origin of $\mathbb{Z}^d$, with probability one. Such regions suppress tunneling and act as barriers against energy transport. For, in resonance-free regions the system can be canonically transformed into an integrable one, up to a very high degree of approximation.[10] Since the motions in these regions lie very nearly on invariant tori (i.e., remain localized), they form indeed strong barriers against the transport of energy from one part of the system to another. If these nonresonant regions are sufficiently dense, one hopes to show that the motion is (approximately) localized in some increasing sequence of finite regions, for longer and longer times, leading to very slow growth of $R^2(t)$ in $|t|$.

There are unfortunately several gaps in this picture. First, it is not clear exactly how one should define the nonresonant regions. While it

seems likely that a large region containing only a few very sparse resonances still forms a very effective barrier to energy transport, as is the case for linear systems,[3–6] we cannot yet prove this fact. If, on the other hand, we attempt to construct our barriers using only regions where the potential is entirely free of resonances, the density of such regions becomes much smaller, and we have a much harder time restricting the region in which the state is localized to a reasonable size, as $|t|$ becomes very large. (For large, but finite $|t|$, see, however, Ref. 10.) A second problem arises from the fact that, although a region of the system may be nonresonant initially, resonances may develop within the region as the trajectory evolves, making the barrier more porous.

Our paper is organized as follows.

In Section 2, we introduce a class of models of disordered, anharmonic crystal lattices which we are able to analyze rigorously, and we state our main results on the existence of some infinite-dimensional, compact invariant tori.

In Section 3, the proofs of our main results are outlined. (Some mathematical details on these and related issues will appear elsewhere.)

In Section 4, extensions of our techniques and results to other disordered systems, such as the nonlinear Schrödinger equation, nonlinear equations for waves, e.g., classical spin waves in disordered magnets, or systems of coupled rotors, are described, and some open problems are discussed.

2. STATEMENT OF THE MAIN RESULTS

Before we give precise statements of our results we introduce and motivate a special class of Hamiltonian systems with infinitely many degrees of freedom.

We recall that, under hypotheses described in Section 1 [see case (b)], a linear disordered system, with equations of motion given by (1.5) can be solved explicitly in terms of a discrete infinity of normal modes, q_j^i, which are given in (1.10). The normal modes, q_j^i, are the eigenvectors of the random matrix Ω^2:

$$\Omega^2 q_j^i = \omega_i^2 q_j^i, \qquad \omega_i^2 > 0 \tag{2.1}$$

They decrease exponentially in $|j|$ and are localized near $i \in \mathbb{Z}^d$. Moreover, they are orthonormal and complete, i.e.,

$$\sum_{j \in \mathbb{Z}^d} q_j^i q_j^k = \delta_{ik} \qquad \text{and} \qquad \sum_{i \in \mathbb{Z}^d} q_j^i q_l^i = \delta_{jl} \tag{2.2}$$

Therefore, the change of variables,

$$(q, p) \to (Q, P)$$

given by

$$\begin{aligned} q_j &= \sum_{i \in \mathbb{Z}^d} Q_i q_j^i \\ p_j &= \sum_{i \in \mathbb{Z}^d} P_i q_j^i \end{aligned} \tag{2.3}$$

is $1-1$ and symplectic, as follows directly from (2.2). In terms of the (Q, P) variables the quadratic Hamiltonian,

$$\frac{1}{2}\left(\sum_j p_j^2 + \sum_{i,j} q_i \Omega_{ij}^2 q_j\right)$$

is given by

$$H_0(Q, P) = \frac{1}{2} \sum_{i \in \mathbb{Z}^d} (P_i^2 + \omega_i^2 Q_i^2) \tag{2.4}$$

The higher-order potential $U(q)$ is given by a function $V(Q)$ which depends, of course, on Ω^2. More generally, we consider perturbations $V(Q, P)$ which may also depend on P. If $U(q)$ is of finite range, e.g., as in Section 1, (1.17), then $V(Q, P)$ is *short range*, in the sense that

$$\left|\frac{\partial^2 V}{\partial Z_i \partial Z_k}\right| \leqslant \text{const } e^{-m|i-k|} \tag{2.5}$$

for some $m > 0$, where $Z_i = Q_i$ or P_i, provided (Q, P) are in the subset of phase space given by

$$\{(Q, P): P_i^2 + \omega_i^2 Q_i^2 \leqslant 1, \quad \text{for all } i\}$$

Next, we note that, the larger the disorder (i.e., the more random the matrix Ω^2 is), the larger the decay rate m will be.[4,5] Hence, for very large disorder, it is quite reasonable to assume that $V(Q, P)$ is of *finite range*, i.e.,

$$\frac{\partial^2 V}{\partial Z_i \partial Z_k}(Q, P) \equiv 0, \qquad \text{unless} \quad |i-k| \leqslant \tilde{\rho} \tag{2.6}$$

for some finite, positive $\tilde{\rho}$.

Note that, by (2.1) and (2.3), the frequencies, ω_i, and the functional $V(Q, P)$ are really functions of the random matrix Ω^2, i.e., random

variables. In order to simplify our notations we wish to assume, however, that the basic random variables are the frequencies ω_i, that they are only weakly correlated, for large separation, and that $V(Q, P)$ is essentially independent of $(\omega_i)_{i \in \mathbb{Z}^d}$. While this assumption is merely made for convenience, the assumption (2.6) is much more severe, and our present methods do not cover interactions, V, for which $\partial^2 V/\partial Z_i \partial Z_k$ decays only exponentially, or more slowly, as the case may be in a large class of physical systems. (Such systems display much more "rigidity" than we can handle at present.) However, the class of Hamilton functions to which our methods apply describes physical systems, like (i) thin surface layers of atoms deposited on a (disordered) crystalline surface, (ii) systems of "bed springs," etc.

From now on, the contents of this paper are mathematically rigorous. First, we give a precise statement of our results for the Hamilton function

$$H(Q, P) = \frac{1}{2} \sum_{i \in \mathbb{Z}^d} (P_i^2 + \omega_i^2 Q_i^2) + \varepsilon \sum_{\langle i,j \rangle} f_{\langle i,j \rangle}(Q_i, Q_j; P_i, P_j) \quad (2.7)$$

introduced above. The interactions $f_{\langle i,j \rangle}$ are assumed to be analytic on some domain specified below and are $O(Q^4 + P^4)$.[5] We also assume the frequencies, ω_i, to be weakly correlated, e.g., to be i.i.d. random variables, with a smooth distribution, such as

$$d\rho(\omega_i) = \begin{cases} \frac{2}{\sqrt{\pi}} \exp[-\omega_i^2], & \text{for} \quad \omega_i > 0 \\ 0, & \text{if} \quad \omega_i \leqslant 0 \end{cases} \quad (2.8)$$

We can, however, handle a general class of smooth distributions with correlations of short range between different ω_i's.

Introducing the action-angle variables for the harmonic oscillators through the equations

$$Q_i = (I_i/\omega_i)^{1/2} \cos \phi_i, \qquad P_i = -(\omega_i I_i)^{1/2} \sin \phi_i \quad (2.9)$$

the Hamilton function $H(Q, P)$ of (2.7) takes the form

$$H(I, \phi) = \sum_i \omega_i I_i + \varepsilon \sum_{\langle i,j \rangle} f_{\langle i,j \rangle}(I_i, I_j; \phi_i, \phi_j) \quad (2.10)$$

Since $f_{\langle i,j \rangle}(Q_i, Q_j; P_i, P_j)$ is $O(Q^4 + P^4)$, for all $\langle i, j \rangle$, we conclude from (2.9) that

$$f_{\langle i,j \rangle}(I, \phi) \sim O(I^2), \qquad \text{for all} \quad \langle i, j \rangle \quad (2.11)$$

[5] It is enough to assume that the $f_{\langle i,j \rangle}$ are $O(|Q|^\delta + |P|^\delta)$, for any $\delta > 2$.

We must, however, avoid points $I_i = 0$, since the change of variables (2.9) and hence $f_{\langle i,j\rangle}(I, \phi)$ fail to be analytic at such points. Finally, we note that $f_{\langle i,j\rangle}(I, \phi)$ will, in general, be a function of (ω_i), as discussed above [see also (2.9)], but we ignore this dependence. Our reason for doing so comes from the observation that the only place this dependence may cause trouble is at $\omega_i = 0$, for some i, and in the course of our construction a neighborhood of these points will be excluded anyway.

We shall consider initial conditions strongly localized in space. So we define

$$I_j^0 = \exp[-|j|^{d+\alpha}] \tag{2.12}$$

with $\alpha > 0$ (arbitrarily small), and $|j| \equiv \sum_{\alpha=1}^{d} |j_\alpha|$, $j \in \mathbb{Z}^d$. The exact choice of the sequence $(I_j^0)_{j \in \mathbb{Z}^d}$ is not important—inside any finite region in $\mathbb{Z}^d$ it could be chosen quite arbitrarily—but what is necessary for the present techniques to work is the very strong decay with distance from the origin that our choice exhibits. As with the interactions, we would like to choose the decay of $(I_j^0)_{j \in \mathbb{Z}^d}$ to be only exponential in $|j|$, but our methods do not encompass this case, yet. (The best we might hope to understand, at present, is a decay like $\exp[-|j|^{1+\alpha}]$, $\alpha > 0$.)

Given some $\rho = (\rho_j)_{j \in \mathbb{Z}^d}$ in $(\mathbb{R}^+)^{\mathbb{Z}^d}$, w and δ in $\mathbb{R}^+$, and some subset N of $\mathbb{R}^{\mathbb{Z}^d}$, we define

$$W(\rho, \delta, w; N) \equiv \{(I, \phi, \omega) \in \mathbb{C}^{\mathbb{Z}^d} \times \mathbb{C}^{\mathbb{Z}^d} \times \mathbb{C}^{\mathbb{Z}^d} : |I_j - I_j^0| < \rho_j, |\operatorname{Im} \phi_j| < \delta,$$
$$\text{and } |\omega_j - \omega_j'| < w, \text{ for some } \omega' \in N, \text{ for all } j \text{ in } \mathbb{Z}^d\} \tag{2.13}$$

This defines suitable complex neighborhoods of tori in the product of phase and frequency space. Let $\rho^0 \overset{\text{e.g.}}{=} \frac{1}{4} I^0$, and $N^0 = \mathbb{R}^{\mathbb{Z}^d}$. We assume that there exist w^0 and δ^0 such that the Hamiltonian (2.10) is analytic on $W(\rho^0, \delta^0, w^0; N^0)$. With our assumption that $f_{\langle i,j\rangle}$ does not depend on ω, we can choose w^0 to be anything we like. Note that the domain $W(\rho^0, \delta^0, w^0; N^0)$ avoids the points $\{I : I_j = 0, \text{ for some } j\}$, where $f_{\langle i,j\rangle}(I_i, I_j; \phi_i, \phi_j)$ may fail to be analytic. Also note that, since

$$f_{\langle i,j\rangle}(I, \phi) \sim O(I^2)$$

we have

$$\sup_W |f_{\langle i,j\rangle}(I, \phi)| \leqslant K \exp[-2(|i| - 1)^{d+\alpha}] \tag{2.14}$$

for some $K > 0$. We can assume K has any finite value we want, just by redefining the constant ε in (2.10) to include a factor of K. For later convenience, we set

$$K = (2^4 d^2)^{-1}$$

Our principal result is as follows:

Theorem T. There exists $\varepsilon_0 > 0$ such that, for $\varepsilon < \varepsilon_0$, there is a set, $\Omega(I^0)$, of ("nonresonant") frequencies, ω, with Prob$(\Omega(I^0))$ arbitrarily close to one (depending on ε and I^0) such that if $\omega \in \Omega(I^0)$ then there is a sequence, $(\tilde{I}_j)_{j \in \mathbb{Z}^d}$, of action variables with the properties that $|\tilde{I}_j - I_j^0| < \rho_j^0$ and that the trajectory of the canonical flow generated by the Hamiltonian, H, of (2.10) with initial conditions $(\tilde{I}, \tilde{\phi})$ (for some $\tilde{\phi}_j \in [0, 2\pi)$, $\forall j \in \mathbb{Z}^d$) lies on an infinite dimensional invariant torus, $\mathbb{T}(\tilde{I})$.

Remarks. (1) We see from the proof below that for *any* point I' on the infinite-dimensional torus, $|I'_j - I_j^0| < \rho_j^0$.

(2) While this theorem tells us that, given I^0, we can find an invariant torus nearby, for most choices of frequencies, ω, it unfortunately does not tell us explicitly how the sets $\Omega(I^0)$ are related to each other, for different choices of I^0, nor is it easy to determine whether, for some fixed I^0, a given ω belongs to $\Omega(I^0)$ or not. (Infinitely many nonresonance conditions must be checked.)

(3) Using the ergodic theorem it is easy to extend Theorem T to construct infinitely many infinite-dimensional, compact invariant tori with probability 1, for small enough ε. However, we have no good idea of the structure of the set of all invariant tori in phase space, for typical frequencies ω. For this reason, Theorem T does not tell us how the quantity $R^2(t)$, introduced in (1.15), depends on t, as $|t| \to \infty$, for a *large* class of initial conditions of finite energy, but only for very special initial conditions, $(\tilde{I}, \tilde{\phi}) \in \mathbb{T}(\tilde{I})$, for which it remains bounded.

(4) Results of the form of Theorem T can also be proven for systems of coupled rotors; see Section 4. In such models we can, in principle, construct many invariant tori, since the unperturbed frequencies depend on I. (J. Bellissard has informed us that M. Vittot and he have independently obtained similar results.[11])

3. A SKETCH OF THE PROOF

The proof of Theorem T uses the method of sequential canonical changes of variables developed by Kolmogorov, Arnol'd, and Moser,[8] with two novel additions:

First, we introduce a sequence of length scales, $L^k \nearrow \infty$, and at the kth stage of our iterative procedure we consider only sites j whose distance from the origin is less than L^k; [for a definition of all inductive constants see a list further in this section]. Secondly, although the integrable part of our Hamiltonian does not satisfy the usual "anisochronicity" assumptions

of the KAM theory, we are able to control the small denominators we encounter by adjusting the *frequencies* themselves, rather than the initial conditions. (This would be different if we studied systems of coupled rotors, for example.)

A standard way of showing that a trajectory of a Hamiltonian, H, lies on an invariant torus is to construct a canonical transformation, C, such that the Hamiltonian $H \circ C$ is a function *only* of the *action variables*. We construct this transformation inductively, attempting, at the kth stage of the inductive process, to "kill" only those parts of the interaction, $f_{\langle i,j\rangle}$, with the points i and j lying inside the box B_{L^k}, which consists of all sites j with $|j| < L^k$. More precisely, one has the following:

Proposition 3.1. There exists $\varepsilon_0 > 0$ such that if $\varepsilon < \varepsilon_0$ then, for every $k \geqslant 0$, there exists a set $N^k \subseteq \mathbb{R}^{\mathbb{Z}^d}$, and a canonical transformation, C_k, which is analytic and invertible on $W(\rho^k, \delta^k, w^k; N^k)$ and maps this set into $W(\rho^0, \delta^0, w^0; N^0)$. The Hamiltonian $H^k = H^0 \circ C_k$ has the form

$$\sum_{i \in \mathbb{Z}^d} \omega_i I_i + h^k(I, \omega) + f^k(I, \phi, \omega) + \varepsilon \sum_{\substack{\langle i,j\rangle | \\ \operatorname{dist}(\langle i,j\rangle, 0) \geqslant L^k}} f_{\langle i,j\rangle}(I_i, I_j, \phi_i, \phi_j) \tag{3.1}$$

where h^k and f^k depend only on (I_j, ϕ_j, ω_j), with $|j| \leqslant L^k$. Furthermore,

$$\begin{aligned} &\sup |f^k| < \varepsilon^k \\ &\sup |\partial h^k / \partial I_j| < \hat{\omega}_j^k \\ &\sup |\partial^2 h^k / \partial I_i \partial \omega_j| < \mu_{ij}^k \\ &\sup |\partial^2 h^k / \partial I_i \partial I_j| < \zeta^k \end{aligned} \tag{3.2}$$

where in each case the supremum is over $W(\rho^k, \phi^k, w^k; N^k)$. Writing $C_k(I, \phi) = [I + \Psi^k(I, \phi),\ \phi + \Phi^k(I, \phi)]$, we have

$$|\Psi_l^k| \leqslant \sum_{m=m(l)}^{k-1} (\varepsilon^m)^{7/8} \quad \text{and} \quad |\Phi_l^k| \leqslant \sum_{m=m(l)}^{k-1} (\varepsilon^m)^{3/8} \tag{3.3}$$

[Here $m(l)$ is defined by the inequalities $L^{m(l)} \leqslant |l| < L^{m(l)+1}$.] Furthermore, $C_k =$ identity at sites, j, with $|j| > L^k$. Finally,

$$\text{Prob}\,(N^k) \geqslant 1 - \sum_{j=0}^{k} (\varepsilon^j)^\kappa, \text{ for some } \kappa > 0$$

The various inductive constants appearing in the Proposition have the values displayed in the following list:

(a) $\varepsilon^k = \varepsilon^{(4/3)^k}$; ε^k bounds the size of the interaction after k iterations.
(b) $\delta^{k+1} = \delta^k - b^k \equiv \delta^k - \delta^0/[64(k+1)^2]$; δ^k measures the size of the

analyticity domain in the angular variables after k iterations, and b^k the amount by which the domain shrinks in the $(k+1)$st step.

(c) $w^k = (\varepsilon^k)^{-2\gamma}$; w^k measures the size of the analyticity domain in the frequency space. γ is a small positive constant—the same one appearing in the definition of the resonant sets, R^k.

(d) $\rho_j^{k+1} = 2^{-3}\rho_j^k$ if $|j| > L^{k+1}$,
$= 2^{-3}\rho_i^k$ if $|j| \leqslant L^{k+1}$, where i is any site with $|i| = L^{k+1}$. (Note that this definition is independent of the site i chosen.) ρ^k measures the size of the analyticity domain for the action variables. Recall that $\rho^0 \equiv I^0/4$, where I^0 is given by (2.12).

(e) $L^k \equiv \{(1+\beta)\,|\ln \varepsilon^k|/2\}^{1/d+\alpha}$; L^k determines the size of the region we must consider at the kth iterative step. Here β is a small positive constant, α is the constant in (2.12) which determines how much faster than exponentially our initial conditions decay, and ε^k is defined in (a).

(f) $M^k = 2(b^k)^{-1}\,|\ln \varepsilon^k|$; M^k determines the number of Fourier coefficients we must consider at the kth step of the iteration. b^k is defined in (b).

(g) $\hat{\omega}_j^k = 0$ if $|j| \geqslant L^k$,
$= \sum_{m=m(j)}^{k-1} (\varepsilon^m)^{3/8}$ if $|j| < L^k$ [Here, $m(j)$ is defined by $L^{m(j)} \leqslant |j| < L^{m(j)+1}$.]

(h) $\mu_{jl}^k = \min[\sum_{m=m(j)}^{k-1} (\varepsilon^m)^{1/4}, \sum_{m=m(l)}^{k-1} (\varepsilon^m)^{1/4}]$ if $|j| < L^k$ *and* $|l| < L^k$,
$= 0$ otherwise

(i) $\zeta^0 = 0$; $\zeta^k = C\sum_{j=0}^{k-1} 2^{6j}(\varepsilon^k)^{-\beta}\ \exp[2m(d+\alpha)(L^{j+1})^{d+\alpha-1}]$, for some $C < 0$.

We also collect several other frequently used symbols, so that the reader may consult this list whenever they occur.

(j) N^k: At each stage of the iterative process we are forced to exclude a small set of resonant frequencies, ω. N^k is the set of *nonresonant* frequencies remaining after k steps. (We recall that $N^0 = \mathbb{R}^{\mathbb{Z}^d}$.)

(k) R^k: The *resonant* frequencies excluded at the kth iterative step. For the definition of R^k see (3.8). Note that $N^k = \mathbb{R}^{\mathbb{Z}^d} \backslash \bigcup_{j=1}^k R^j$.

(l) $\mathbb{X}^k$: In constructing the canonical transformation, C_k, in Proposition 3.1 we will consider only a finite set of Fourier coefficients of the interaction terms in (3.1). These Fourier coefficients are labeled by vectors, ν, in a finite subset $\mathbb{X}^k$ of $\mathbb{Z}^{\mathbb{Z}^d}$.

(m) $\tilde{\omega}^k$: These are the frequencies of the motion generated by the "integrable piece" of the Hamiltonian H^k. See (3.4) for a definition.

(n) B_{L^k}: The box of size L^k, i.e., all sites $j \in \mathbb{Z}^d$, with $|j| < L^k$.

Theorem T follows easily from Proposition 3.1. Let $(I^k(t), \phi^k(t))$ be the solutions of Hamilton's equations with Hamiltonian, H^k, and initial conditions (I^0, ϕ^0) (ϕ^0 an arbitrary element of $S_1^{\mathbb{Z}^d}$). The bounds in (3.2) imply that the points $(I^k(t), \phi^k(t))$ remain in $W(\rho^k, \delta^k, w^k; N^k)$, for $0 \leqslant t \leqslant T^k \equiv (\varepsilon^k)^{-1/8}$, and that they obey, for the same time interval, $I_j^k(t) = I_j^0 + 0((\varepsilon^k)^{3/4})$, and $\phi_j^k(t) = \tilde{\omega}_j^k(I^0, \omega)\, t + \phi_j^0 + O((\varepsilon^k)^{3/8})$, where

$$\tilde{\omega}_j^k(I, \omega) \equiv \omega_j + \frac{\partial h^k}{\partial I_j}(I, \omega) \tag{3.4}$$

These results are shown by using Cauchy's theorem to bound derivatives like $|\partial f^k/\partial I_j|$ by $C \cdot (\rho_j^k)^{-1} \varepsilon^k$ on any domain contained in $W(\rho^k, \delta^k, w^k; N^k)$, say $W(\rho^k/2, \delta^k, w^k; N^k)$. If one inserts the values of ρ^k and ε^k given in (d) and (a) above and derives analogous bounds on $\partial f^k/\partial \phi_j$ one arrives at the stated expressions.

Let $(I(t), \phi(t))$ be the solution of Hamilton's equations with Hamiltonian $H(I, \phi)$ and initial conditions $C_k(I^0, \phi^0)$, and $(\hat{I}^k(t), \hat{\phi}^k(t)) = C_k(I^0, \tilde{\omega}(I^0, \omega)\, t + \phi^0)$. The bounds in (3.3), and the fact that canonical transformations take solutions of Hamilton's equations to solutions of Hamilton's equations imply $|I_j(t) - \hat{I}_j^k(t)| < O((\varepsilon^k)^{1/8})$ and $|\phi_j(t) - \hat{\phi}_j^k(t)| < O((\varepsilon^k)^{1/8})$, for all $0 \leqslant t \leqslant T^k = (\varepsilon^k)^{-1/8}$. We show below that

$$\lim_{k \to \infty} C_k(I^0, \tilde{\omega}^k(I^0, \omega)\, t + \phi^0)$$

exists for all t, and all $\omega \in \bigcap_k N^k$, thereby obtaining an invariant torus, and completing the proof of Theorem T.

The proof of Proposition 3.1 begins by noting that we can choose H^0 to be our initial Hamiltonian and $C^0 =$ identity. [The peculiar choice of the constant K in (2.14) was to insure that $|f^0| < \varepsilon$, here.] Assuming the proposition holds for integers less than or equal to k, we sketch the construction of C_{k+1}.

The transformation C_{k+1} is of the form $C_k \circ C^k$, and since C_k is known, we need only construct C^k. If H^k were integrable, then all its trajectories would lie on invariant tori and we would be done. To show a system is integrable we attempt to solve the Hamilton–Jacobi equation

$$H^k\left(\frac{\partial \Sigma}{\partial \phi}(I', \phi), \phi\right) = \tilde{h}(I')$$

for functions $\Sigma(I', \phi)$ and $h(I')$. If such a solution exists (globally) we obtain a canonical transformation by inverting the equations

$$I = \frac{\partial \Sigma}{\partial \phi} \qquad \text{and} \qquad \phi' = \frac{\partial \Sigma}{\partial I'}$$

to express I and ϕ in terms of I' and ϕ' (or vice-versa), and this canonical transformation gives a set of action-angle variables which transform H^k into $\tilde{h}$. While we cannot solve the Hamilton–Jacobi equation exactly we obtain a good approximate solution by noting that since H^k is close to integrable we expect the desired change of variables to be close to the identity, and thus attempt to write

$$\Sigma(I', \phi) = I' \cdot \phi + S(I', \phi)$$

(remember that $I' \cdot \phi$ is the generating function of the identity transformation) and assume that S is $O(\varepsilon^k)$, the amount by which H^k differs from an integrable system.

Now the Hamilton–Jacobi equation takes the form

$$H^k\left(I' + \frac{\partial S}{\partial \phi}, \phi\right) = \tilde{h}(I')$$

and if we expand the left-hand side and discard quantities of second order in the small quantities ε^k and S we obtain

$$\sum_i \omega_i I_i' + \sum_i \tilde{\omega}_i^k(I', \omega) \cdot \frac{\partial S}{\partial \phi_i}(I', \phi) + f^k(I', \phi, \omega) + \sum_{\substack{\langle i,j\rangle \\ \mathrm{dist}(\langle i,j\rangle, 0) \geqslant L^k}} f_{\langle i,j\rangle}(I_i', I_j'; \phi_i, \phi_j) = \tilde{h}(I') \quad (3.5)$$

with $\tilde{\omega}^k$ defined by (3.4). We note that, for $\mathrm{dist}(\langle i, j\rangle, 0) \geqslant L^k$, $\varepsilon f_{\langle i,j\rangle} \sim 0(\varepsilon^k)$, so we are justified in discarding terms like $\varepsilon(\partial f_{\langle i,j\rangle}/\partial I') \cdot \partial S/\partial \phi$. Furthermore, by (2.14) we see that if $|i|$ or $|j|$ is greater than or equal to $L^{k+1} \equiv \mathrm{const} \cdot |\ln \varepsilon^k|^{1/d+\alpha}$, the constribution of the terms $f_{\langle i,j\rangle}$ to the sum on the left-hand side of (3.5) will be of higher order in ε^k and thus may be discarded without worsening the approximation. Since f^k and $f_{\langle i,j\rangle}$ are periodic functions of ϕ, we can solve (3.5) by means of Fourier series and we find

$$S(I', \phi, \omega) = i \sum_{\nu \neq 0} \frac{[f_\nu(I', \omega) + \varepsilon \tilde{f}_\nu(I', \omega)]\, e^{i\nu \cdot \phi}}{\tilde{\omega}(I', \omega) \cdot \nu} \quad (3.6)$$

Here, the f_ν's are the Fourier coefficients of f^k, the $\tilde{f}_\nu$'s are the Fourier coefficients of $\sum_{\langle i,j\rangle, L_{k+1} > |i|, |j| \geqslant L_k} f_{\langle i,j\rangle}$, and the sum runs over all nonzero vectors, ν, in $\mathbb{Z}^{\mathbb{Z}^d}$. Also note that the generating function will depend on the frequencies, and we denote this dependence explicitly in (3.6).

In general the sum in (3.6) will diverge. To cure this problem, we first reduce the infinite sum to a finite one. Note that, since f^k is analytic on

$W(\rho^k, \delta^k, w^k; N^k)$, Cauchy's theorem implies $|f_\nu^k(I')| < \varepsilon^k e^{-\delta^k|\nu|}$. A similar remark holds for $|\tilde{f}_\nu^k|$. Thus, terms with $|\nu|$ large (here $|\nu| = \sum_i |\nu_i|$) contribute little to the sum, and since we have already discarded terms of $O((\varepsilon^k)^2)$ when linearizing the Hamilton–Jacobi equation, we can discard all terms in (3.6) with $|\nu| \geqslant M^{k+1} \equiv \text{const} \, |\ln \varepsilon^k|$, and hope not to worsen our approximation.

Because we have also discarded those parts of the interaction coming from distances more than L^{k+1} from the origin we can restrict the sum in (3.6) to vectors $\nu \in \mathbb{X}^{k+1}$, with

$$\mathbb{X}^{k+1} \equiv \{\nu \in \mathbb{Z}^{\mathbb{Z}^d} : 0 < |\nu| < M^{k+1}, \quad \text{and} \quad \nu_i = 0 \quad \text{if} \quad |i| \geqslant L^{k+1}\} \tag{3.7}$$

With these restrictions, the sum in (3.6) contains only a finite number of terms, and a simple estimate shows this number is bounded by $(2M^{k+1})^{(2L^{k+1})^d}$.

The sum in (3.6) can now fail to be well defined only if the denominator in one of the terms vanishes. To prevent this, we exclude from consideration the (resonant) frequencies

$$R^k = \{\omega \in N^k : \text{there exists } \nu \in \mathbb{X}^{k+1} \text{ such that } |\tilde{\omega}^k(I^0, \omega) \cdot \nu| < (\varepsilon^k)^\gamma\} \tag{3.8}$$

Here γ is the same constant that appears in (c) of the list of inductive constants. We now set $N^{k+1} = N^k \backslash R^k$ (the nonresonant frequencies at step $k+1$) and attempt to define our canonical transformation on $W(\rho^{k+1}, \delta^{k+1}, w^{k+1}; N^{k+1})$.

Condition (3.8) bounds the denominators in (3.6) only if $I' = I^0$ and $\omega \in N^{k+1}$. However, note that for any (I', ω') in $W(\rho^k, \delta^k, w^k; N^{k+1})$ we can write

$$\begin{aligned}\tilde{\omega}^k(\tilde{I}', \omega') \cdot \nu &= \tilde{\omega}^k(I^0, \omega) \cdot \nu \\ &\quad \cdot \{1 - [(\tilde{\omega}^k(I^0, \omega) \cdot \nu)^{-1}(\tilde{\omega}^k(I^0, \omega) \cdot \nu - \tilde{\omega}^k(I', \omega') \cdot \nu)]\end{aligned} \tag{3.9}$$

A little calculation using the fundamental theorem of calculus, and the bounds on derivatives of $\tilde{\omega}^k$ given by (3.2) shows that the term in square brackets in (3.9) is bounded in magnitude by 1/2, for ε sufficiently small. Thus, the denominator in (3.6) is larger than $C \cdot (\varepsilon^k)^\gamma$ on $W(\rho^k, \delta^k, w^k; N^{k+1})$.

The summand in (3.6) is now defined and bounded by $C \cdot (\varepsilon^k)^{1-\gamma}$, and we bound derivatives of S on domains smaller than $W(\rho^k, \delta^k, w^k; N^{k+1})$ by combining Cauchy's theorem with our previous estimate on the number of terms in the sum. In this way we find that $|\partial S/\partial I_j|$ and $|\partial S/\partial \phi_j|$ are bounded by $C(\rho_j^k)^{-1}(\varepsilon^k)^{1-\gamma}(2M^{k+1})^{(2L^{k+1})^d}$ and $C \cdot (\delta^k)^{-1}(\varepsilon^k)^{1-\gamma}(2M^{k+1})^{(2L^{k+1})^d}$, respectively.

We can do even better than this if the site j satisfies $|j| > L^k$. In this case, when one differentiates the summand in (3.6), $f_\nu(I')\, e^{i\nu\cdot\phi}$ makes no contribution, since $f^k(I, \phi, \omega)$ depends only on (I_j, ϕ_j), with $|j| \leqslant L^k$. The derivatives of $\tilde{f}^k_\nu(I')$ can, on the other hand, be bounded by $D^k_j \exp[-2m(|j|-1)^{d+\alpha} - \delta^k |\nu|]$, using (2.14), where D^k_j is either $C\cdot(\rho^k_j)^{-1}$ or $C(\delta^k)^{-1}$, depending on whether we differentiate with respect to I_j or ϕ_j. Using this estimate we see that when $|j| > L^k$ the factor of $(\varepsilon^k)^{1-\gamma}$ in the bounds on derivatives of S can be replaced by $\exp[-2m(|j|-1)^{d+\alpha} - \delta^k |\nu|]$, resulting in much better control of these quantities.

By the analytic inverse function theorems,[12] the equations $I = I' + (\partial S/\partial\phi)(I', \phi, \omega)$ and $\phi' = \phi + (\partial S/\partial I')(I', \phi, \omega)$ can be inverted to yield an analytic and invertible canonical transformation on $W(\rho^{k+1}, \delta^{k+1}, w^{k+1}; N^{k+1})$, provided $|\partial S/\partial\phi_j| \ll \rho^k_j$ and $|\partial S/\partial I'_j| \ll \delta^k$. The bounds on derivatives of S in the preceding paragraphs, and the explicit forms of ρ^k and δ^k in the list of inductive constants immediately imply these inequalities, so we get

$$C^k(I', \phi') = (I' + \Xi(I', \phi', \omega), \phi' + \Delta(I', \phi', \omega)) \tag{3.10}$$

Note that $\Xi(I', \phi', \omega) = (\partial S/\partial\phi)(I', \phi', \omega)$ and $\Delta(I', \phi, \omega) = -(\partial S/\partial I')(I', \phi, \omega)$. So our bounds on the derivatives of S give bounds on the canonical transformation. Furthermore, since $C_{k+1} = C_k \circ C^k$, the bounds on Ξ and Δ [and the observation that C^k maps $W(\rho^{k+1}, \delta^{k+1}, w^{k+1}; N^{k+1})$ into $W(\rho^k, \delta^k, w^k; N^k)$] imply the bounds on C^{k+1} stated in (3.3). If we in addition use the fact that S does not depend on (I'_j, ϕ_j) with $|j| \geqslant L_{k+1}$, we see that C^k, and hence C_{k+1} will reduce to the identity at these sites.

These bounds also imply

$$|C_{k+1}(I^0, \phi) - C_k(I^0, \phi)| < O((\varepsilon^k)^{1/4})$$

so that these transformations converge, as $k \to \infty$.

The bounds on the probability that $\omega \in N^{k+1}$ follow by noting that $\omega \in N^k \backslash N^{k+1}$ only if there is $\nu \in \mathbb{X}^{k+1}$ such that $|\tilde{\omega}^k(I^0, \omega)\cdot\nu| < (\varepsilon^k)^\gamma$. If we could replace $\tilde{\omega}^k$ in this expression by ω, the unperturbed frequency, the probability that ω satisfied this inequality, for some fixed ν, could easily be estimated by bounding a Gaussian integral. The estimates on derivatives of h^k in the proposition are sufficient to show that on the set of ω's where $|\tilde{\omega}^k(I^0, \omega)\cdot\nu| < (\varepsilon^k)^\gamma$, the transformation $\omega \to \tilde{\omega}^k$ is one to one and has Jacobian bounded by a constant independent of k and ν. Thus, just by

changing variables in the Gaussian integral defining our probability measure, we find

$$\text{Prob}(\{\omega: |\tilde{\omega}^k(I^0, \omega)\cdot \nu| < (\varepsilon^k)^\gamma\}) \leqslant C^{(2L^{k+1})^d}\varepsilon_k^\gamma \tag{3.11}$$

and since there are at most $(2M^{k+1})^{(2L^{k+1})^d}$ vectors in $\mathbb{X}^k$, we find that $\text{Prob}(N^k \backslash N^{k+1})$ is bounded by $(\varepsilon^k)^\kappa$ for some $0<\kappa<\gamma$, and the bound on $\text{Prob}(N^{k+1})$ follows.

It remains to verify (3.2). Write $H^{k+1}(I', \phi') = H^k \circ C^k(I', \phi') = H^k(I' + \Xi, \phi' + \Delta)$ which is in turn equal to

$$\sum_i \omega_i(I_i' + \Xi_i(I', \phi', \omega)) + h^k(I' + \Xi(I', \phi', \omega), \omega)$$
$$+ f^k(I' + \Xi, \phi' + \Delta, \omega) + \sum_{\substack{\langle i,j\rangle \\ \text{dist}(\langle i,j\rangle, 0) \geqslant L^k}} f_{\langle i,j\rangle}(I' + \Xi, \phi' + \Delta) \tag{3.12}$$

If we now expand the last three terms in (3.12), using the fundamental theorem of calculus, use the fact that $\Xi_i = \Delta_i = 0$ if $|i| \geqslant L^{k+1}$ (because $C^k =$ identity at these sites) and also note that $\sum_i \omega_i \Xi_i' = \sum_i \omega_i(\partial S/\partial \phi_i)$ and will thus cancel many Fourier modes of the interaction terms in (3.12), we find (3.12) takes the form

$$H^{k+1}(I', \phi') = \sum_i \omega_i I_i' + h^{k+1}(I', \omega) + f^{k+1}(I', \phi', \omega)$$
$$+ \sum_{\substack{\langle i,j\rangle \\ \text{dist}(\langle i,j\rangle, 0) \geqslant L^{k+1}}} f_{\langle i,j\rangle}(I_i', I_j', \phi_i', \phi_j') \tag{3.13}$$

with $h^{k+1}(I', \omega) = h^k(I', \omega) + f_0(I', \omega) + \tilde{f}_0(I', \omega)$, and $f^{k+1}(I', \phi', \omega)$ having the rather complicated form

$$\sum_{|\nu| \geqslant M^{k+1}} [f_\nu(I', \omega) + \tilde{f}_\nu(I')]\, e^{i\nu\cdot\phi} + \sum_{\substack{\langle i,j\rangle \\ \text{either } |i| \text{ or } |j| \geqslant L^{k+1} \\ \text{but not both}}} f_{\langle i,j\rangle}(I' + \Xi, \phi)$$
$$+ \sum_j \int_0^1 ds\, \frac{\partial f^k}{\partial I_j}(I' + s\Xi, \phi, \omega)\cdot \Xi_j + \sum_{i,j} \int_0^1 ds \int_0^s dt$$
$$\times \frac{\partial^2 h^k}{\partial I_i \partial I_j}(I' + t\Xi)\, \Xi_i \cdot \Xi_j + \sum_{\substack{\langle i,j\rangle \\ L^k \leqslant \text{dist}(0, \langle i,j\rangle) < L^{k+1}}} \int_0^1 ds$$
$$\times \left[\frac{\partial f_{\langle i,j\rangle}}{\partial I_j}(I' + s\Xi, \phi)\cdot \Xi_j + \frac{\partial f_{\langle i,j\rangle}}{\partial I_i}(I' + s\Xi, \phi)\cdot \Xi_i\right] \tag{3.14}$$

with $\phi = \phi(I', \phi', \omega)$.

The bounds on derivatives of h^{k+1} are easily obtained from the bounds on $f^k, f_{\langle i,j\rangle}$ and Cauchy estimates. [Note that it is clear that $h^{k+1}(I', \omega)$ depends on I'_i only if $|i| < L^{k+1}$, since $\tilde{f}_0$ is a Fourier coefficient of a sum of terms $f_{\langle i,j\rangle}$ with $|i|$ and $|j| < L^{k+1}$.]
Note that the bounds on $\tilde{\omega}^{k+1}$ imply

$$|\tilde{\omega}^{k+1}(I^0, \omega) - \tilde{\omega}^k(I^0, \omega)| \leqslant O((\varepsilon^k)^{3/4}) \tag{3.15}$$

so $\lim_{k\to\infty} \tilde{\omega}^k(I^0, \omega)$ exists if $\omega \in \bigcap_k N^k$.

All that remains is to verify the first inequality in (3.2). Although the expression for f^{k+1} looks complicated it contains nothing unexpected—it consists of terms arising from the fact that in linearizing the Hamilton–Jacobi equation we discarded things of higher order in ε^k [this accounts for the last four terms in (3.14)], terms with $|\nu| \geqslant M^{k+1}$ [the first term in (3.14)], and terms that arose too far from the origin (the second term in (3.14)].

Using our observation that the Fourier coefficients of a function analytic on $W(\rho^k, \delta^k, w^k; N^k)$ decay as $e^{-\delta^k|\nu|}$, and the fact that there are at most $(2M)^{(2L^{k+1})^d}$ vectors ν with $\operatorname{supp} \nu \subset B_{L^{k+1}}$ and with $|\nu| = M$, the first term in (3.14) is bounded on $W(\rho^{k+1}, w^{k+1}; N^{k+1})$ by

$$\sum_{M \geqslant M^{k+1}} \varepsilon^k \cdot M^{(2L^{k+1})^d} \cdot e^{-(\delta^k - \delta^{k+1})M} < O((\varepsilon^k)^2) \tag{3.16}$$

The second term in (3.14) is bounded by noting that there are at most $(CL^{k+1})^{d-1}$ terms in the sum, each of which is bounded by $\varepsilon \cdot \exp[-2(L^{k+1}-2)^{d+\alpha}]$, so the contribution of the sum is $(\varepsilon \cdot \varepsilon^{k+1})$.

The third and fifth terms in (3.14) are bounded by noting that a dimensional estimate bounds $|\partial f^k/\partial I_j|$ and $|\partial f_{\langle i,j\rangle}/\partial I_j|$ by $C \cdot (\rho_j^k)^{-1} \varepsilon^k$ and $C \cdot (\rho_j^k)^{-1} \exp[-2(|j|-1)^{d+\alpha}]$, respectively. The factors of Ξ_j are bounded, using our observation that $\Xi_j = \partial S/\partial \phi_j$, and we find that each of these terms is less than $O((\varepsilon^k)^{3/2-\eta})$, where η is a small, positive constant that can easily be computed from the list containing definitions of the inductive constants.

Finally, the fourth term is bounded by using ζ^k, to bound the derivative of h^k, and the usual bounds on Ξ_j, and we find it is less than $O((\varepsilon^k)^{2-\eta})$.

Thus, if we choose $\varepsilon^{k+1} = (\varepsilon^k)^{4/3}$, (3.14) will be bounded by ε^{k+1}, on $W(\rho^{k+1}, \phi^{k+1}, w^{k+1}; N^{k+1})$, completing the verification of (3.2) and the proof of the proposition.

4. GENERALIZATIONS, CONCLUSIONS

The techniques developed in this paper have some rather straightforward extensions to other nonlinear dynamical systems. They per-

mit us, in principle, to construct examples of invariant tori for the flows of some Hamiltonian systems with infinitely many degrees of freedom. We should like to mention some examples.

4.1. Coupled Rotors

In action-angle variables, the Hamilton function of a system of uncoupled rotors located at the sites of a lattice, $\mathbb{Z}^d$, is given by

$$h_0(I) = \sum_{j \in \mathbb{Z}^d} I_j^2 \tag{4.1}$$

It is obvious that all motions of this system are quasiperiodic motions on infinite-dimensional tori. It is an interesting problem to study what happens if an angle-dependent perturbation is added to h_0. Thus, consider a Hamiltonian of the form

$$H(I, \phi) = h_0(I) + \varepsilon U(I, \phi) \tag{4.2}$$

where

$$\frac{\partial^2 U}{\partial Z_i \partial Z_j} \equiv 0, \qquad \text{if} \quad |i-j| > \tilde{\rho} \tag{4.3}$$

for some finite $\tilde{\rho}$. Here $Z = I$ or ϕ. Condition (4.3) expresses the circumstance that interactions between different rotors are of finite range. We also assume that $U(I, \phi)$ is analytic on a complex neighborhood, W_0, of

$$\{I, \phi: 0 < I_j < I_0,\ 0 \leqslant \phi_j < 2\pi, \text{ for all } j\}$$

for some finite I_0 (arbitrarily large), and that on W_0

$$|U(I, \phi)| \leqslant |I|^{\delta} \tag{4.4}$$

for some sufficiently large $\delta > 2$.

We note that h_0 satisfies a strong *anisochronicity condition*, namely,

$$\frac{\partial \omega_i}{\partial I_j} = \delta_{ij}, \qquad \text{where } \omega_i \equiv \frac{\partial h_0}{\partial I_i} \tag{4.5}$$

and that it is useful to assume that

$$\begin{aligned} \left| \left(\omega_i^{-1} \frac{\partial \bar{U}}{\partial I_i} \right)(I) \right| &< |I|^{\delta'} \\ \left\| \frac{\partial^2 \bar{U}}{\partial I_i \partial I_j}(I) \right\| &< |I|^{\delta''} \end{aligned} \tag{4.6}$$

for some positive δ' and δ''. Here

$$\bar{U}(I) = \left(\prod_{j \in \mathbb{Z}^d} \frac{1}{2\pi} \int_0^{2\pi} d\phi_j \right) U(I, \phi)$$

and $\|A\|$ is the norm of the infinite matrix $A = (a_{ij})_{i,j \in \mathbb{Z}^d}$. These conditions show that it is possible to avoid resonances in a perturbative solution of the Hamilton–Jacobi equation to first order in ε by restricting the action variables, I, to a suitable (resonance-free) subset of phase space which fills up a "large" subset of

$$\{(I, \phi): 0 < |I_j| < e^{-m|j|^{d+\alpha}}\} \tag{4.7}$$

for some $\alpha > 0$. This observation permits us to start a KAM iteration of the type described in Section 3 and, under some additional assumptions on U, construct infinite-dimensional, compact invariant tori for the flow corresponding to the Hamilton function H defined in (4.2). In fact, for such systems one may be able to construct a family of invariant tori contained in the subset defined in (4.7) which has positive measure, with respect to the Lebesgue measures $\prod_{j \in \Lambda} dI_j \, d\phi_j$, where Λ is any finite subset of $\mathbb{Z}^d$, (this would be a result analogous to those in Refs. 12 and 13 which concern Hamiltonian systems with finitely many degrees of freedom). We have, however, not carried out the details of all the calculations required to check this. [We feel that conditions (4.4), (4.6) on U are a bit unphysical and that the systems studied in Sections 2 and 3 are somewhat more interesting and more realistic.] J. Bellissard has kindly informed us that he and M. Vittot have independently studied systems of infinitely many coupled rotors which admit infinite-dimensional invariant tori.[11]

4.2. Nonlinear Schrödinger Equation with Random Potential

The discretized nonlinear Schrödinger equation with a random potential has the form

$$i \frac{\partial}{\partial t} \psi_t(x) = [(-\Delta + v) \psi_t](x) + \lambda \psi_t(x) \sum_{y \in \mathbb{Z}^d} V(|x - y|) \, |\psi_t(y)|^2 \tag{4.8}$$

Here $t \in \mathbb{R}$ is time, x, y,... are sites in $\mathbb{Z}^d$, Δ is the finite difference Laplacian on $l_2(\mathbb{Z}^d)$, $v = (v(x))_{x \in \mathbb{Z}^d}$ is a random potential [the $v(x)$'s being, e.g., i.i.d. random variables], V is some positive-definite, deterministic "potential,"

e.g., $V(|x-y|)=\delta_{xy}$, and $\lambda>0$ some coupling constant. Identifying $\psi(x)$ with $(1/\sqrt{2})(q_x+ip_x)$ we see that

$$\{\psi(x), \psi(y)\} = \{\psi^*(x), \psi^*(y)\} = 0$$
$$\{\psi(x), \psi^*(y)\} = -i\delta_{xy}, \quad (\psi^* \equiv \bar{\psi})$$

and that (4.8) are Hamiltonian equations of motion, with Hamilton function

$$H = \sum_{x,y \in \mathbb{Z}^d} \{\psi^*(x)(-\Delta_{xy} + v(x)\,\delta_{xy})\,\psi(y) + \lambda\,|\psi(x)|^2\,V(x-y)\,|\psi(y)|^2\} \tag{4.9}$$

Equation (4.8) arises in the theory of vortices in boson systems, where it is known as the Gross–Pitaevskii equation,[15,16] in Hartree–Fock theory (see, e.g., Ref. 17), etc. The problem of analyzing the properties of solutions of (4.8) is a direct, "nonlinear" generalization of the Anderson problem.[1] It is interesting to study the following questions.

(1) *Existence of stationary solutions.* To find solutions of (4.8), one may make the ansatz

$$\psi_t(x) = e^{-itE}u_E(x) \tag{4.10}$$

which solves (4.8) iff

$$Eu_E(x) = [(-\Delta + v)\,u_E](x) + \lambda u_E(x) \sum_{y \in \mathbb{Z}^d} V(|x-y|)\,|u_E(y)|^2 \tag{4.11}$$

Similar equations also arise in the theory of superconductivity.[18] Let us look for solutions, u_E, of the nonlinear eigenvalue problem (4.11) which have finite l_2 norm. Since we have introduced a coupling constant λ, we may then normalize u_E to satisfy

$$\sum_{x \in \mathbb{Z}^d} |u_E(x)|^2 = 1 \tag{4.12}$$

We have constructed solutions of (4.11) satisfying (4.12) for a class of potentials, v, for which the Schrödinger operator $-\Delta + v$ has bound states, for small enough λ. The case where v is a random potential such that $(v(x))_{x \in \mathbb{Z}^d}$ are i.i.d. (e.g., Gaussian) random variables, with large disorder, is more difficult. Extending ideas in Refs. 5 and 6 and adapting the strategy in Section 3, it appears very promising that, for small λ, one may construct solutions of (4.11) for countably many values of E which are "close" to

eigenfunctions of $-\Delta + v$. We have, however, not checked all the details of this rather complicated construction (based on a Newton method).

It would be interesting to estimate reflection and transmission coefficients for solutions of (4.11), for finite one-dimensional systems $(\mathbb{Z}^d \to [-l, -l+1, \ldots, 0, 1, 2, \ldots, l])$. B. Souillard informs us that he has some results in that direction.

Concepts more general than stationary solutions are the following:

(1′) *Time-periodic solutions.* It is clear that, in general, Eq. (4.8) will have periodic-in-time solutions which are not stationary, but we have no interesting results on that kind of solutions. The same comment applies to the following.

(2) *Quasiperiodic solutions, invariant tori.* So far, we have not been able to construct invariant tori for the flow corresponding to the Hamilton function H defined in (4.9). However, if in (4.9) we replace the term $\sum_{x,y} \psi^*(x)[-\Delta_{xy} + v(x)\,\delta_{xy}]\,\psi(y)$ by $\sum_x \omega_x^2 |\psi(x)|^2$, where $(\omega_x)_{x \in \mathbb{Z}^d}$ are i.i.d. (e.g., Gaussian) random variables, and if $V(|x-y|)$ is of finite range then one can extend the KAM techniques of Section 3 to the present system to construct infinite-dimensional, compact invariant tori for a set of ω_x's of large probability. Solutions of (4.8) which lie on such tori remain localized, for all times, e.g., in the sense that

$$R^{(n)}(t) \equiv \left[\sum_{x \in \mathbb{Z}^d} |\psi_t(x)|^2 \right]^{-1} \sum_{x \in \mathbb{Z}^d} |x|^n |\psi_t(x)|^2 \tag{4.13}$$

is uniformly bounded in t, for arbitrary finite n. We may introduce a diffusion constant

$$D \equiv \overline{\lim_{t \to \infty}}\, t^{-1} R^{(2)}(t) \tag{4.14}$$

By rescaling λ we may always normalize initial conditions in (4.8) by requiring, e.g.,

$$\sum_{x \in \mathbb{Z}^d} |\psi_{t=0}(x)|^2 = 1$$

We conjecture the following:

(3) *Absence of diffusion, for small λ and large disorder.* If the random potential v in (4.8), (4.9) has a distribution with sufficiently large disorder and λ is small enough then, for "most" initial conditions,[5] ψ_0, of finite support

$$D \equiv D(\psi_0) = 0 \tag{4.15}$$

with $D(\psi_0)$ as in (4.13), (4.14).

[6] "Most," e.g., with respect to the uniform measure on finite-dimensional unit spheres.

Unfortunately we are missing some good ideas of how to prove a result like (4.15) (but see the discussion at the end of Section 1).

Our last example concerns the following.

4.3. Nonlinear, Classical Spin Waves

Consider some material doped with magnetic ions of very large spin S which are ferromagnetically coupled. The Hamilton operator of this system is then given by

$$H_\Lambda = -\sum_{i,j\in\Lambda} J_{ij}\mathbf{S}_i\cdot\mathbf{S}_j - \sum_{j\in\Lambda} h_j S_j^z \tag{4.16}$$

where the exchange couplings, J_{ij}, are positive, h_j is an external magnetic field in the z direction, and $\Lambda\subseteq\mathbb{Z}^d$. If the magnetic ions are diluted by non-magnetic material the J_{ij} and the h_j may be assumed to be random variables. Specifically we assume that

$$SJ_{ij} = \mathscr{J}_{ij}, \qquad |i-j| = 1, \ =0, \text{ otherwise}$$

where $(\mathscr{J}_{ij})_{\langle i,j\rangle\in\mathbb{Z}^d}$ are i.i.d. random variables with smooth distributions supported on $[0,\infty)$. Since we have assumed that the spin, S, given by

$$\mathbf{S}_i^2 = S(S+1), \qquad \text{for all } i$$

is very large, the dynamics of this systems is approximately *classical*,[19] i.e., we may view the spins, $\mathbf{S}_i$, as *unit vectors* in $\mathbb{R}^3$ which satisfy the equations of motion

$$\dot{\mathbf{S}}_i = \mathbf{S}_i\wedge\left[\sum_{j\in\mathbb{Z}^d}(SJ_{ij})\,\mathbf{S}_j + \mathbf{h}_i\right] \tag{4.17}$$

where $\mathbf{h}_j = (0,0,h_j)$. These are the so-called Landau–Lifshitz equations[20] describing a Larmor precession of the spins, $\mathbf{S}_i$, in the effective magnetic field $\sum_j (SJ_{ij})\,\mathbf{S}_j + \mathbf{h}_i$. They are Hamiltonian equations of motion, the Hamilton function being

$$H = -\sum_{i,j}(SJ_{ij})\,\mathbf{S}_i\cdot\mathbf{S}_j - \sum_j h_j S_j^z \tag{4.18}$$

The phase space of this system is given by

$$\mathsf{X}_{j\in\mathbb{Z}^d} S_{(j)}^2$$

where $S^2_{(j)}$ is a copy of the unit sphere in $\mathbb{R}^3$, for every j. If ϕ_j and θ_j are the standard polar angles on $S^2_{(j)}$ the symplectic form can be written as

$$\sum_j d(\cos\theta_j) \wedge d\phi_j \qquad (\theta_j \neq 0)$$

and the Liouville measure as

$$d\mu = \prod_j \frac{1}{4\pi} d(\cos\theta_j)\, d\phi_j$$

For large h_j,

$$\mathbf{S}_j \approx (0, 0, 1)$$

and we may expand all quantities in powers of the small angles θ_j. Setting

$$\psi(j) = \theta_j e^{i\phi_j} \in \mathbb{C}$$

$0 < \theta_j < \infty$, we obtain, to second order in θ_j,

$$\mathbf{S}_i \cdot \mathbf{S}_j \simeq 1 - \tfrac{1}{2} |\psi(i) - \psi(j)|^2$$

The symplectic structure is given by

$$\sum_j dq_j \wedge dp_j$$

with $q_j = \theta_j \cos\phi_j$, $p_j = \theta_j \sin\phi_j$. To second order in θ, the Hamiltonian is given by

$$H^{(2)} = (S/2) \sum_{ij} J_{ij} |\psi(i) - \psi(j)|^2 + \sum_j h_j |\psi(j)|^2 \tag{4.19}$$

The problem of finding the eigenmodes and eigenfrequencies for the dynamics given by $H^{(2)}$ is identical to the problem of determining eigenfunctions and spectrum of the Jacobi matrix Ω^2, with

$$\begin{aligned} \Omega^2_{ij} &= -\mathscr{J}_{ij}, && |i-j| = 1 \\ &= \left(\sum_k \mathscr{J}_{ik}\right) + h_i, && i = j \end{aligned}$$

This problem has been analyzed in Section 1; see Refs. 4–6. We recall that if, for example, the disorder in h is large enough there exists a complete orthonormal system of localized (standing) spin waves.

In complete analogy with our analysis in Sections 2, 3, and 4.2 we ought to address the problem of analyzing the propagation properties of spin waves when the nonlinear terms in the equations of motion are not neglected. This problem is especially challenging if the only randomness is in the J_{ij}, $h_j = h \searrow 0$, and there is spontaneous magnetization. Although our general ideas and concepts can be applied to this problem, we have no interesting rigorous results, at present.

Finally, we wish to mention that our techniques appear to have interesting applications to the theory of electrical networks with inductances and capacitances that are slightly random. For such systems, our basic assumptions of sharply localized initial conditions and finite energy are actually quite realistic.

SUMMARY

In this paper we have formulated and discussed the general problem of localization in nonlinear dynamical systems. We have succeeded in extending the KAM iteration techniques to some Hamiltonian systems with infinitely many degrees of freedom. Although our strategy involves some novel ideas, its main technical ingredients are quite standard and somewhat crude. It would be most interesting and important to try to relax the very strong compactness condition (2.12), [i.e., $I_j^0 \sim \exp(-|j|^{d+\alpha})$, $\alpha > 0$], by means of a more careful analysis of the locality properties of the terms $h^k(I, \omega)$ and $f^k(I, \phi, \omega)$ (with respect to subsets of B_{L^k} and $\mathbb{X}^k$) in the Hamiltonian (3.1), and to show that the interaction terms generate anisochronicity.

Another problem for which we have partial but not entirely satisfactory results concerns the existence of periodic orbits in Hamiltonian systems with infinitely many degrees of freedom. But physically (and probably mathematically) the most interesting problem would be to prove the absence of diffusion (i.e., no energy transport) in nonlinear, disordered dynamical systems. In the Hamiltonian case we have some hopes that a combination of ideas developed in Refs. 5 and 10 will lead to such results, as discussed at the end of Section 1, but good general ideas and methods are still missing.

ACKNOWLEDGMENTS

We gratefully acknowledge instructive discussions with J. Bellissard, G. Gallavotti, R. de la Llave, and especially with R. Joynt, J. Pöschel, and E. Trubowitz.

Two of us (J. F. and T. S.) thank the Institute for Advanced Study for hospitality during a period when some of the work presented here was carried out.

C.E.W. wishes to thank the Theoretical Physics department of the ETH—Zürich for its hospitality during the preparation of this manuscript.

The work of J. Fröhlich was supported in part by the U.S. National Science Foundation under Grant No. DMS-8405264 during the author's stay at the Institute for Advanced Study.

T. Spencer was supported in part by NSF Grant No. DMR-8401225.

C. E. Wayne was supported in part by NSF Grant No. DMS-8403664.

REFERENCES

1. P. W. Anderson, *Phys. Rev.* **109**:1492 (1958).
2. Ph. Choquard, *The Anharmonic Crystal* (W. A. Benjamin, New York, 1967).
3. K. Ishii, *Suppl. Progr. Theor. Phys.* **53**:77 (1977).
4. F. Wegner, *Z. Phys. B* **22**:273 (1975); Ya. Gol'dsheid, S. Molchanov, and L. Pastur, *Funct. Anal. Appl.* **11**:1 (1977); H. Kunz and B. Souillard, *Commun. Math. Phys.* **78**:201 (1980); F. Delyon, H. Kunz, and B. Souillard, *J. Phys. A* **16**(1), 25 (1983); R. Carmona, *Duke Math. J* **49**:191 (1982).
5. J. Fröhlich and T. Spencer, *Commun. Math. Phys.* **88**:151 (1983).
6. J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer, *Commun. Math. Phys.* **101**:21 (1985); B. Simon and T. Wolff, Comm. Pure Appl. Math., to appear; F. Delyon, Y. Lévy and B. Souillard, Preprint, 1985.
7. A. Weinstein, *Ann. Math.* **98**:377 (1973); *Inv. Math.* **20**:47, (1973); J. Moser, *Commun. Pure Appl. Math.* **29**:727 (1976).
8. N. Kolmogorov, *Dokl. Akad. Nauk.* **98**:527 (1954); V. Arnol'd, *Russian Math. Survey* **18**:9, 85 (1963); J. Moser, *Nachr. Akad. Wiss. Göttingen I* **1a**, (1962).
9. N. Nekhoroshev, *Russian Math. Survey* **32**:1 (1977).
10. C. E. Wayne, On the elimination of non-resonance harmonics, and Bounds on the trajectories of a system of weakly coupled rotators, preprints, 1984/85.
11. J. Bellissard and M. Vittot, preprint, Marseille, 1985.
12. G. Gallavotti, *The Elements of Mechanics*, (Springer, New York, 1983), appendices and references.
13. L. Chierchia and G. Gallavotti, *Nuovo Cimento B* **67**:277, 1982; see also G. Gallavotti, in *Scaling and Self-Similarity in Physics*, Jürg Fröhlich, Ed. (Birkhäuser, Boston, 1983).
14. J. Pöschel, Ph. D. thesis, ETH—Zürich, 1981.
15. E. P. Gross, *Nuovo Cimento* **20**:454 (1961).
16. L. P. Pitaevskii, *Sov. Phys.-JETP* **13**:451 (1961).
17. A. W. Overhauser, *Phys. Rev. Lett.* **4**:415 (1960); E. H. Lieb and M. de Llano, *J. Math. Phys.* **19**:860 (1978); B. Sriram Shastry, *Phys. Rev. Lett.* **50**:633 (1983).
18. J. R. Schrieffer, *Theory of Superconductivity* (W. A. Benjamin, New York, 1964); P. G. DeGennes, *Superconductivity of Metals and Alloys*, (W. A. Benjamin, New York, 1964).
19. P. A. Vuillermot, *Commun. Math. Phys.* **76**:1 (1980).
20. L. D. Landau and E. M. Lifshitz, *Phys. Z. Sowj. Union* **8**:153 (1935).