Renormalization-group analysis of the discrete quasiperiodic Schrödinger equation

Stellan Ostlund*
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106

Rahul Pandit
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853
(Received 11 July 1983)

Recently developed scaling concepts in the theory of quasiperiodic dynamical systems are used to develop an *exact* renormalization group applicable to the discrete, quasiperiodic Schrödinger equation. To illustrate the power of the method, we calculate the universal scaling properties of the states and eigenvalue spectrum at and below the localization transition for an energy which corresponds to an integrated density of states of $\frac{1}{2}$. The modulating potential has a frequency $\frac{1}{2}(\sqrt{5}-1)$ relative to the underlying lattice for the example we work out in greatest detail.

I. INTRODUCTION

A. Overview

Schrödinger equations with quasiperiodic potentials[1-3] have been studied for many years. The motivation for this study is both physical and mathematical.[4] Quasiperiodic potentials arise naturally in incommensurate structures,[5] in nonstoichiometric intergrowth compounds,[6] such as $Hg_{3-\delta}AsF_6$, in the calculation of band structures of periodic crystals in magnetic fields,[7-16] and in superconducting lattices in magnetic fields.[17] Mathematically, quasiperiodic potentials are interesting because Bloch's theorem is inapplicable. These potentials lead to rich spectra and wave functions because they are, in some sense, intermediate between periodic and random. Periodic potentials lead to absolutely continuous spectra[18] and extended eigenstates, whereas random potentials lead to pure-point spectra and localized eigenstates in one dimension. Though there are no rigorous proofs, the general belief is[4] that generic quasiperiodic potentials lead to spectra that are Cantor sets, have both absolutely continuous and pure-point components, and, in addition, a singular continuous component.[18-20] The wave functions can be extended, localized, or "critical" in a sense that we will specify below.

Quasiperiodic potentials pose mathematical problems that have a fundamental connection with the small-divisor problems that were investigated by Kolmogorov, Arnol'd, and Moser[19] (KAM) in classical mechanics. If the effect of a quasiperiodic potential is calculated within perturbation theory,[5,20,21] then high-order terms in the perturbation expansion (for, say, the wave function) have small denominators. However, Dinaburg and Sinai[22,19] showed that for a sufficiently weak quasiperiodic potential, a large part of the spectrum is still absolutely continuous. These proofs use the ideas of the KAM theory and the results are the analog of the KAM theorem, which states that most of the invariant tori in the phase space of an integrable Hamiltonian system are not destroyed by a sufficiently weak nonintegrable perturbation.

Recently, nonperturbative renormalization-group (RG) methods have been used to study scaling phenomena in a variety of small-divisor problems. These include invariant circles of area-preserving maps,[23] circle maps,[24,25] and invariant curves in mappings of the complex plane onto itself.[26] Recent work[27] has shown that similar RG methods can be used to study the scaling properties of the spectra and wave functions of a discrete, one-dimensional Schrödinger equation with a special nonanalytic, quasiperiodic potential with two incommensurate frequencies. These ideas are extended to analytic, quasiperiodic potentials with two incommensurate frequencies in the present work.

We believe scaling properties of the spectra and wave functions of quasiperiodic Schrödinger operators can be studied most conveniently by RG methods. The exponents that characterize scaling behavior (which we define in Sec. II) are simply related to the eigenvalues of the linearized RG transformation in the vicinity of its fixed points.[28] The universality of these exponents follows naturally because different quasiperiodic Schrödinger operators that flow to the same fixed point under successive iterations of the RG transformation share the same scaling behavior. Thus, even though we study a specific quasiperiodic potential, we can predict the scaling properties of the spectra and wave functions for a large class of quasiperiodic potentials. In addition to these fundamental insights, the exact RG that we construct in Sec. IV gives a very efficient numerical algorithm for computing various properties of quasiperiodic operators.

There have been some previous attempts to explain properties of quasiperiodic operators using RG ideas.[20,29,30] These investigations either use approximations or do not carry out a fixed-point analysis with which scaling could be investigated. By contrast, we analyze the fixed point of an exact renormalization group which governs the scaling properties of the states and spectra. One of the most powerful aspects of the present analysis is that the exact formulation can be implemented using a nu-

merical algorithm which allows identification and calculation of scaling properties to any desired accuracy. A combination of an exact RG theory and a numerically stable algorithm with which it can be implemented has formed the basis of a rigorous mathematical investigation of similar problems.[31] We believe that the present ideas could be developed into a rigorous mathematical theory as well. In addition, our RG analysis complements rigorous mathematical work[4] on quasiperiodic potentials that has not focused on scaling properties and calculations of scaling indices.

The remainder of this paper is organized as follows. In Sec. I B we define the model we work with and present our principal results. In Sec. II we use transfer-matrix methods to obtain empirical scaling results for the spectra and wave functions of our model. Section III reviews the renormalization group used to study circle maps.[25] In Sec. IV we use the results of Sec. III to construct a RG for a Schrödinger equation with a quasiperiodic potential. In Sec. V we show how to find and analyze the fixed points of this RG. Section VI is devoted to the results of this fixed-point analysis. These results are related to the empirical results of Sec. II in Sec. VII. The latter section also contains some concluding remarks. In the Appendix we give some definitions and some results of number theory which we use in this paper.

B. Model and principal results

The eigenvalue equation we analyze is

$$-(\psi_{n+1}-2\psi_n+\psi_{n-1})+\epsilon V(x_n)\psi_n=E\psi_n\ , \tag{1.1a}$$

where $V(x)=V(x+1)$ and x_n satisfies the recursion formula

$$x_{n+1}=f(x_n)\ . \tag{1.1b}$$

The "circle map" f satisfies $f(x+1)=1+f(x)$. Note that if $x_0=\varphi_0$, $f(x)=x+\sigma$, and $V(x)=\cos(2\pi x)$, we get Harper's equation,[7] also called the "almost Mathieu" equation,[32] for which

$$V(x_n)=\epsilon\cos[2\pi(\sigma n+\varphi_0)]\ . \tag{1.2}$$

We discuss this potential throughout much of this paper. Although our results are, in principle, more general, we shall usually assume $f(x)=x+\sigma$.

If $\sigma=p/q$, with p and q relatively prime integers, $V(x_n)$ is periodic. The unit cell has length q and the spectrum consists of $q-1$ gaps and q bands. We normalize the Bloch index κ in an extended-zone scheme by $2\pi q$, so $0\le\kappa\le\frac{1}{2}$. The gaps are at $\kappa=\frac{1}{2}[n\sigma(\mathrm{mod}\ 1)]\equiv\frac{1}{2}\langle n\sigma\rangle$ and $\kappa=\frac{1}{2}(1-\langle n\sigma\rangle)$ (see Ref. 33). For sufficiently small ϵ in Eq. (1.1), the size of the gaps decreases[21] with increasing n, for $1\le n\le\frac{1}{2}q$.

In the bands, there are, of course, extended states, whose normalization scales with system size. For energies in the gap one usually says that because $|\psi_n|$ diverges exponentially at $+\infty$ or $-\infty$ (or both), there are no states. However, for a finite system, there may be a pair of states which are exponentially localized at the ends. Precisely *at* the gap edge, there is one extended state. There is also a state whose amplitude grows (or decays) *linearly*. This linear solution is usually discarded by imposing periodic boundary conditions, but in a *finite* system it has to be considered.[34]

If σ is irrational, $V(x_n)$ is quasiperiodic.[1] It has either been proved or is believed to be true that generic quasiperiodic Schrödinger operators lead to[4,22,32,35–38] the following.

(1) A Cantor-set spectrum with a dense set of gaps (see, e.g., Fig. 1): the spectrum can have absolutely-continuous, singular-continuous, and/or pure-point components.[18]

(2) It has been proved that the gaps can be labeled by the Bloch index $\kappa(E)$ which (suitably defined as in Sec. II and Ref. 39) continues to exist for all energies E. All evidence indicates that the principal gaps are at $\kappa=\frac{1}{2}\langle n\sigma\rangle$ and $\kappa=\frac{1}{2}(1-\langle n\sigma\rangle)$ and their size generally decreases with increasing n.

(3) States can be localized, extended, or neither localized nor extended (we call these "critical" states); whether a state is localized, extended, or critical in general depends on the strength of the potential ϵ, the Bloch index $\kappa(E)$, and σ.

For the special case of Harper's equation (1.2) a variety of additional results are known. Some of these results are listed below. More detail may be found in Refs. 4, 5, and 32.

(a) If we substitute

$$\psi_n=e^{2\pi i\kappa(E)n}\sum_{m=-\infty}^{\infty}f_m e^{2\pi i m(\sigma n+\varphi_0)}$$

in Harper's equation (1.2), we obtain its "dual" representation

$$-(f_{m+1}-2f_m+f_{m-1})+\epsilon_D\cos[2\pi(\sigma m+\tilde{\varphi}_0]f_m=\tilde{E}f_m\ , \tag{1.3}$$

where the dual variables are

$$\epsilon_D=4/\epsilon,\quad \tilde{\varphi}_0=\kappa(E)(\mathrm{mod}\ \sigma),\quad \tilde{E}=2E/\epsilon\ . \tag{1.4}$$

Note that the models (1.2) and (1.3) are self-dual for $\epsilon=2$. This duality was first noted by Aubry and André.[5]

(b) For σ a Roth number (see the Appendix), ϵ sufficiently large, and for almost all φ_0, Harper's equation (1.2) has a pure-point spectrum and exponentially localized eigenstates.[32]

(c) For fixed φ_0, the measure of the bands approaches zero as ϵ approaches its self-dual value 2 from below.[5] This approach is linear[40] in $\epsilon-2$.

(d) There is good evidence[41] that, for σ a Roth number and any φ_0, all states are extended for $\epsilon<2$ and localized for $\epsilon>2$. However, as far as we know, there is no rigorous proof of this statement. In the remainder of this paper we shall assume that the localization transition does occur at the self-dual point $\epsilon=2$. All our results are consistent with this.

New results have been obtained by two methods. Using standard numerical experiments, we have calculated scaling properties of the spectrum and eigenstates of Harper's equation (1.2). More importantly, we have also derived these scaling properties from a RG analysis. Our princi-

pal results, some of which are contained in Ref. 27, are summarized below.

(i) If σ is a quadratic irrational, we find that the local scale invariance of the spectrum around point $E(\kappa)$ is determined by the equation[42]

$$|\alpha|^{p\tilde{\gamma}}=\lim_{n\to\infty}\frac{E(|\alpha|^{-p(n-1)}x+\kappa)-E(\kappa)}{E(|\alpha|^{-pn}x+\kappa)-E(\kappa)}, \tag{1.5}$$

where the limit exists independent of x for some finite integer $p(\kappa)$ and a scaling exponent $\tilde{\gamma}(\kappa)$ at certain values of κ. α is related to σ (see Sec. III). For $\epsilon<2$, $\tilde{\gamma}$ is independent of φ_0 and is 1 for all σ and κ. For $\epsilon=2$, the scaling described by Eq. (1.5) may hold only for special values of φ_0 (Secs. II and VII). Our most extensive calculations have been done for $\sigma=\sigma_G=\frac{1}{2}(\sqrt{5}-1)$ and $\kappa=\frac{1}{4}$. In this case, $-\alpha=\sigma_G^{-1}$, and for $\epsilon=2$ and $\varphi_0=-\frac{1}{4}+\frac{1}{2}\sigma_G$, $\tilde{\gamma}(\frac{1}{4})=1.829\pm0.001$ and $p=3$.

(ii) The degree of localization of a state can be characterized by the exponent

$$-\beta=\frac{1}{p\ln|\alpha|}\lim_{n\to\infty}\left[\ln\left[\frac{L_n}{L_{n+p}}\right]\right], \tag{1.6a}$$

where

$$L_n=\frac{1}{q_n}\sum_{j=-q_n}^{q_n}|\psi_j|^2 \tag{1.6b}$$

and q_n is the denominator of the nth rational approximant to σ (see the Appendix) and p is defined in Eq. (1.5). The exponent β is 0 for an extended state and -1 for a localized state. We find that $\beta=0$ ($\beta=-1$) for $\epsilon<2$ (for $\epsilon>2$). For $\epsilon=2$ we find $-1<\beta<0$, i.e., the states are neither localized nor extended, but what we call "critical." These critical states do not decay to zero and are in fact scale invariant. This is made precise in Sec. II.

(iii) We construct an exact RG that allows us to map the model (1.1), for any σ, onto a similar, but coarse-grained, model. The structure of this RG gives fixed points only when σ is a quadratic irrational and then only for certain values of κ. For Harper's equation with $\sigma=\sigma_G$ and $\kappa=\frac{1}{4}$ we show explicitly how our empirical scaling results emerge from this RG: The $\epsilon<2$ behavior is governed by a "weak-coupling" fixed point with one relevant eigenvalue,[28] whereas the $\epsilon=2$ behavior is governed by a "critical" fixed point with two physically relevant eigenvalues. It is also shown that a large class of quasiperiodic potentials that can be expressed as sums of harmonics of Harper's potential lie in the same universality class, i.e., their spectra and wave functions scale with the same exponents as those of Harper's equation (1.2).

Interesting questions we *cannot* address directly follow. What is the total measure of the energy for which states are localized? How do the energies for which states are extended, localized, and critical fill up the spectrum, respectively? These questions can be answered numerically for particular models by using the RG as a tool to multiply matrices efficiently. However, this is not the most powerful application of RG theory, since it does not derive scaling and universality from eigenvalues of the transformation at a fixed point, and there is therefore no rigorous way to argue for the universal character of the properties calculated in this manner.

II. EMPIRICAL RESULTS

A. Transfer-matrix method

Equation (1.1) can be recast into the recursion formula

$$\begin{bmatrix}\psi_{n+1}\\ \psi_n\end{bmatrix}=\underline{M}_1(x_n)\begin{bmatrix}\psi_n\\ \psi_{n-1}\end{bmatrix}, \tag{2.1a}$$

$$x_{n+1}=f(x_n), \tag{2.1b}$$

where

$$\underline{M}_1(x)=\begin{bmatrix}2-E+V(x) & -1\\ 1 & 0\end{bmatrix}. \tag{2.2}$$

If the sequence $\{x_n\}_{n=0}^{n=\infty}$ is of period q, so that $x_q=x_0+p$ where p and q are integers, the band structure is obtained by investigating the matrix $\underline{M}_q(x_0)$ defined by

$$\underline{M}_q(x_0)\equiv\prod_{n=0}^{q-1}\underline{M}_1(x_n) \tag{2.3}$$

(see Ref. 43).

Since $\det\underline{M}_q=1$, the condition that a band exists at energy E is

$$|\mathrm{Tr}\underline{M}_q(E)|<2, \tag{2.4}$$

which makes the eigenvalues of $\underline{M}_q(E)$ complex. Our normalization of the Bloch index κ in an extended-zone scheme enables us to write

$$2\cos[2\pi q\kappa(E)]=\mathrm{Tr}\underline{M}_q(E). \tag{2.5}$$

This definition of the Bloch index is independent of the choice of length of the unit cell, and we can sensibly discuss the irrational limit.

It is important to realize that this definition of the Bloch index makes $\kappa(E)$ equivalent to half the integrated density of states, i.e.,

$$\kappa(E)=\tfrac{1}{2}\int_{-\infty}^{E}dE'\rho(E'), \tag{2.6}$$

and to the average rotation rate of the phase of the wave function of an eigenstate[39]

$$\kappa(E)=\frac{1}{2\pi}\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathrm{Im}\ln\left[\frac{\psi_{n+1}}{\psi_n}\right], \tag{2.7}$$

where we define the complex logarithm to have a branch cut below the negative real axis. The latter two definitions remain well defined in the limit when σ becomes irrational.

In the remainder of this section we give certain scaling properties of the spectra, wave functions, and the matrices $\underline{M}_{q_k}$ which we have obtained by numerical experiments. These results are related to our RG results in Secs. V–VII. Since we have done *explicit* RG calculations only for $\kappa=\frac{1}{4}$, we present empirical results only for this value

of κ. Related results for other values of κ, both at band edges and in the bands, are summarized in Ref. 27. Thouless and Niu[20] and Kohmoto[47] have also studied some scaling properties of the spectra and wave function for Harper's potential (1.2), but not for fixed κ. (See also *Note added.*)

B. Empirical scaling results for spectra

Using formula (2.7) we calculated $\kappa(E)$ vs E for $\epsilon \leq 2$ when $V(x)=\cos(2\pi x)$ and $f(x)=x+\sigma_G$, where $\sigma_G=(\sqrt{5}-1)/2=$ "golden mean." Results for $\epsilon=1.5$ and $\epsilon=2$ are shown in Figs. 1 and 2 (see the Appendix).

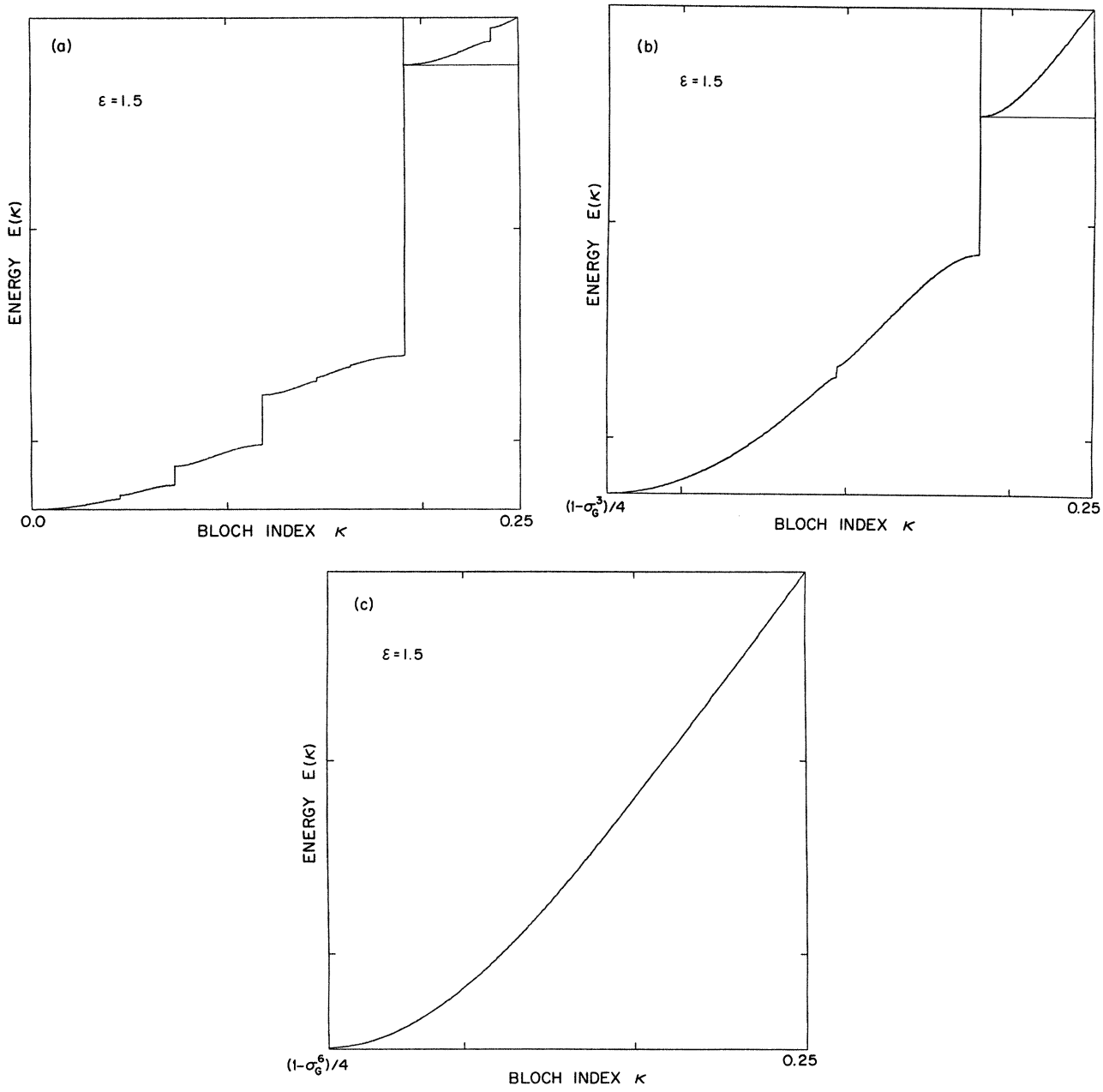


FIG. 1. (a) The energy $E(\kappa)$ is plotted as a function of the Bloch index κ for Harper's equation, when the incommensurability σ is the golden mean. This figure shows the spectrum when $\epsilon=1.5$. The normalization corresponds to the full band extending from $0\leq\kappa\leq\frac{1}{2}$, so that half of the spectrum is shown. The full spectrum from $\kappa=0$ to $\kappa=\frac{1}{2}$ is antisymmetric about the point $E=2$ and $\kappa=\frac{1}{4}$. There is a dense set of gaps, but only about a dozen gaps can be resolved in this picture. The spectrum in the box in the upper right corner is expanded to full scale in the next figure. (b) The spectrum from $\kappa=\frac{1}{2}\sigma^2=\frac{1}{4}(1-\sigma^3)$ to $\kappa=\frac{1}{4}$ is shown. This corresponds to the box in the upper right hand corner of (a) and is obtained by rescaling the domain of κ by a factor σ_G^3. (c) The spectrum from $\kappa=\frac{1}{4}(1-\sigma^6)$ to $\kappa=\frac{1}{4}$ is shown. This is the box in the upper right corner of (b) but the domain is magnified by a factor σ^{-3}. On this scale the gaps have all but vanished. By continuing these rescalings one obtains a linear function, which is trivially scale invariant. The construction illustrates the scaling index $\tilde{\gamma}=1$, which occurs below the localization threshold for this value of κ.

All the results for $\epsilon<2$ are qualitatively similar to Fig. 1(a). The band structure in the vicinity of $\kappa=\frac{1}{4}$, which by symmetry is at $E(\kappa)=2$, is found to have a scale invariance determined by a scaling index $\tilde{\gamma}(\kappa)$ and an integer $p(\kappa)=3$ defined by ($\sigma=\sigma_G$)

$$\gamma(\kappa)\equiv\sigma^{-p\tilde{\gamma}}=\lim_{k\to\infty}\frac{E(\sigma^{p(k-1)}x+\kappa)-E(\kappa)}{E(\sigma^{pk}x+\kappa)-E(\kappa)}\ . \quad (2.8)$$

The index $\tilde{\gamma}$ is independent of x, but depends on whether or not ϵ is less than or equal to 2. We make explicit the dependence on κ because the form of the scaling holds for other values of κ, although $p(\kappa)$ and $\tilde{\gamma}(\kappa)$ may vary.

The advantage of using Harper's equation is that it has a reflection summetry in the band structure which yields $E=2$ for $\kappa=\frac{1}{4}$. Furthermore (see Sec. I B), the localiza-

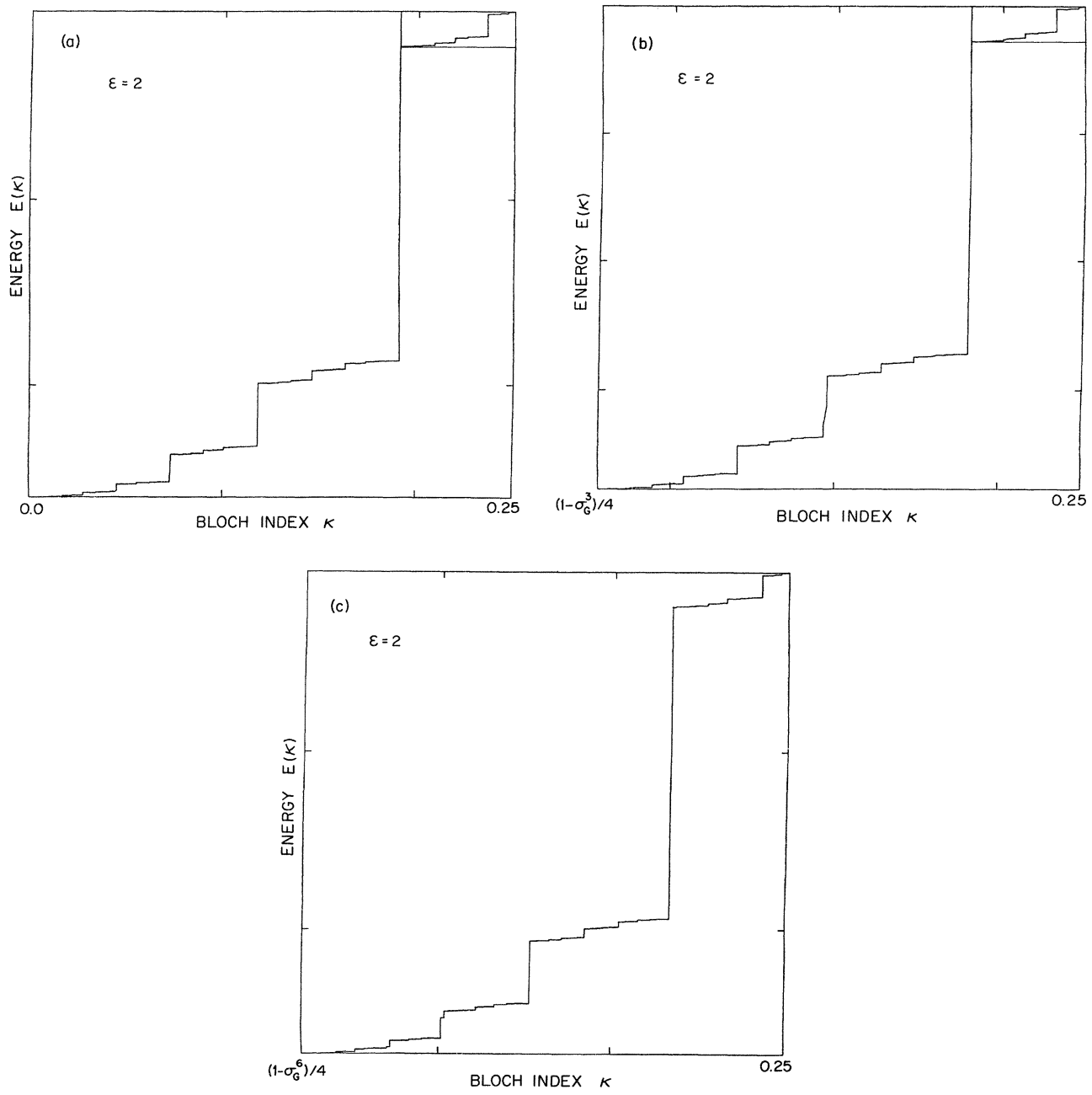


FIG. 2. The same construction as in Figs. 1(a)–1(c) is shown, but for $\epsilon=2$ corresponding to the localization threshold. In this case, the gaps remain on the scale of the plots. It is seen that these plots are scale invariant, since there is no visible difference between the three figures. Each successive plot is obtained by magnifying the E by a factor $\sigma^{-3\tilde{\gamma}}=14.01\ldots$ while rescaling κ by a factor $\sigma^{-3}=4.23\ldots$ in the previous plot. These plots are scale invariant in the limit of large magnification about $E(\frac{1}{4})$ and $\kappa=\frac{1}{4}$, a feature which we show is universal for a large class of potentials with $\sigma=\sigma_G$. The fact that these plots show this scale invariance so precisely for relatively large energy scales is a nonuniversal property of the cosine potential.

tion transition occurs at $\epsilon=2$. This simplifies the calculations considerably since, in contrast to the case of a general potential, we do not have to search for the ϵ and E at which the localization transition occurs. Of course, we run the risk of using a potential which is so restrictive that we cannot reach any conclusions about the universality of our results. However, we show in Sec. V that this is not the case. The addition of higher harmonics to Harper's potential turns out to be an irrelevant perturbation at the fixed point we investigate.

The two cases $\epsilon<2$ and $\epsilon=2$ behave very differently. For $\epsilon<2$, we find that $\tilde{\gamma}(\frac{1}{4})=1$. This is illustrated in Figs. 1(a)–1(c). In Fig. 1(a) we show $E(\kappa)$ for $0\leq\kappa\leq\frac{1}{4}$. The portion of the spectrum obtained by rescaling κ by σ^3 in the vicinity of $\kappa=\frac{1}{4}$ is shown in the upper right-hand corner of Fig. 1(a), and shown expanded to full scale in Fig. 1(b). A similar construction generates Fig. 1(c) from Fig. 1(b). Thus Fig. 1(c) is a scaled version of the second tiny box just inside the upper right-hand corner of Fig. 1(b). The *size* of the gaps at values of κ in the vicinity of $\kappa^*=\frac{1}{4}$ *vanish* on an energy scale $[E(\kappa)-E(\kappa^*)]$ as $\kappa\rightarrow\kappa^*$. In this limit, the spectrum in units of energy which scale with $E(\kappa)$ becomes linear. (This does *not* imply that the spectrum becomes differentiable in this limit; hence the *density* of states can still have structure.)

The spectrum for $\epsilon=2$ is very different. In this case, we find[47] $\sigma^{-p\tilde{\gamma}}=14.017\ldots$ which results in $\tilde{\gamma}=1.829\pm0.001$. The sequence of rescalings described for Figs. 1(a)–1(c) is repeated in Fig. 2 for this case. The scale factor $\tilde{\gamma}$ converges remarkably fast as a function of k in (2.8), and to the accuracy of the width of the *lines* drawn in Fig. 2, we see no difference between Figs. 2(a)–2(c). In fact, even for $k=1$, we have $\tilde{\gamma}$ to 2–3 decimal places, and we cannot rule out the possibility that Eq. (2.8) may hold *exactly* for finite κ. This very rapid "crossover" is surely an artifact of the special choice of our starting model. In Sec. V this value of $\tilde{\gamma}$, in the *limit* of formula (2.8), is derived from our RG and shown to be "universal."

C. Scaling of the states for Harper's equation with $\sigma=\sigma_G$

1. General definitions

We measure the degree of localization around the site where the wave function attains a maximum (say site 0) by the exponent β defined by[27]

$$\beta=\lim_{m\rightarrow\infty}\left[\frac{1}{p\ln\sigma}\ln\frac{L_m}{L_{m+p}}\right], \tag{2.9}$$

where the norm L_m is defined by

$$L_m=\frac{1}{q_m}\sum_{n=-q_m}^{q_m}|\psi_n|^2, \tag{2.10}$$

p is some integer which may depend on κ, and q_m represents the denominator of successive rational approximants to σ. An ordinary extended state has $\beta=0$, while a localized state has $\beta=-1$. *A priori* β may depend on the boundary condition $[\psi_1,\psi_0]$ and possibly on $x_0\equiv\varphi_0$.

2. $\epsilon<2$

Throughout this discussion, we set $E=2$, so that we are working at $\kappa=\frac{1}{4}$ and $p=3$. When $\epsilon<2$, the states are clearly "extended." By this we mean that for *any* initial condition chosen, i.e., $[\psi_1,\psi_0]$ and $x_0\equiv\varphi_0$, the successive ψ_n neither diverge nor decay to zero and the exponent β is 0 *for all boundary conditions.* Setting $[\psi_0,\psi_1]=[1,0]$ and $\varphi_0\equiv0.1234$ we plot ψ_n vs n in Fig. 3(a).

The rapid oscillations in $|\psi_n|$ make it hard to interpret, so it is useful to look at the *hull* function $\chi(x)$ defined as

$$\psi_n^{\pm}=e^{\pm2\pi in\kappa}\chi(n\sigma), \tag{2.11}$$

where $\chi(x)$ is of period 1. Since each of the states at $E=2$ is a superposition of states with $\kappa=\pm\frac{1}{4}$, we plot $(-1)^n\psi_{2n}$ vs $\langle2n\sigma\rangle$ to recover $\chi(x)$. This function is shown in Fig. 3(b) and is clearly continuous. The continuity of the hull function implies $\beta=0$ and is independent of φ_0 and boundary conditions.

3. $\epsilon=2$

Let us repeat the above exercise for $\epsilon=2$. In Fig. 4(a) we plot ψ_n vs n for the arbitrary initial condition $[\psi_0,\psi_1]=[1,0]$ and $\varphi_0=0.1234$. We see that ψ_n grows and in fact keeps growing as $n\rightarrow\infty$. Furthermore, the hull function in Fig. 4(b) is discontinuous.

However, we can still extract the following important fact. As $n\rightarrow\infty$, the *local* structure of ψ_n around successive maxima starts looking similar. [This is not apparent from Fig. 4(a) for two reasons: Not enough points have been plotted, and to see the scale invariance the length must be expanded around the maximum to exclude all other local maxima where $|\psi_n|$ is larger than the central value.] Let us define N_k to be the location of the largest value of $|\psi_n|$ up to $n=q_k$, i.e.,

$$|\psi_{N_k}|\geq|\psi_n|,\quad n<q_k. \tag{2.12}$$

Let us further define φ_n to be the phase at site n:[33]

$$\varphi_n\equiv\langle n\sigma+\varphi_0\rangle. \tag{2.13}$$

We then find that, as $k\rightarrow\infty$, φ_{N_k} tend to any one of the values $\frac{1}{4}+\frac{1}{2}n+\sigma m=\varphi^*$ where n and m are low-order integers: $n,m=0,\pm1,\pm2$. Since a lattice translation by m shifts φ_n by $m\sigma$, the σm is unimportant. Changing the *sign* of the potential corresponds to adding a half-integer to φ_0. This is equivalent to transforming $\psi_n\rightarrow(-1)\psi_n$ in the case $E=2$. Thus we can set $\varphi_0=\frac{1}{4}$ and investigate the structure around the site of the maximum by investigating the structure of ψ_n around the *origin* given the boundary condition that $\psi_{\pm\infty}<\infty$ and $\varphi_0=\frac{1}{4}$. With these boundary conditions, we obtained Figs. 5(a)–5(c). In Figs. 5(a), 5(b), and 5(c) we plot $|\psi_n|$ at the $q_{17}=4181$,

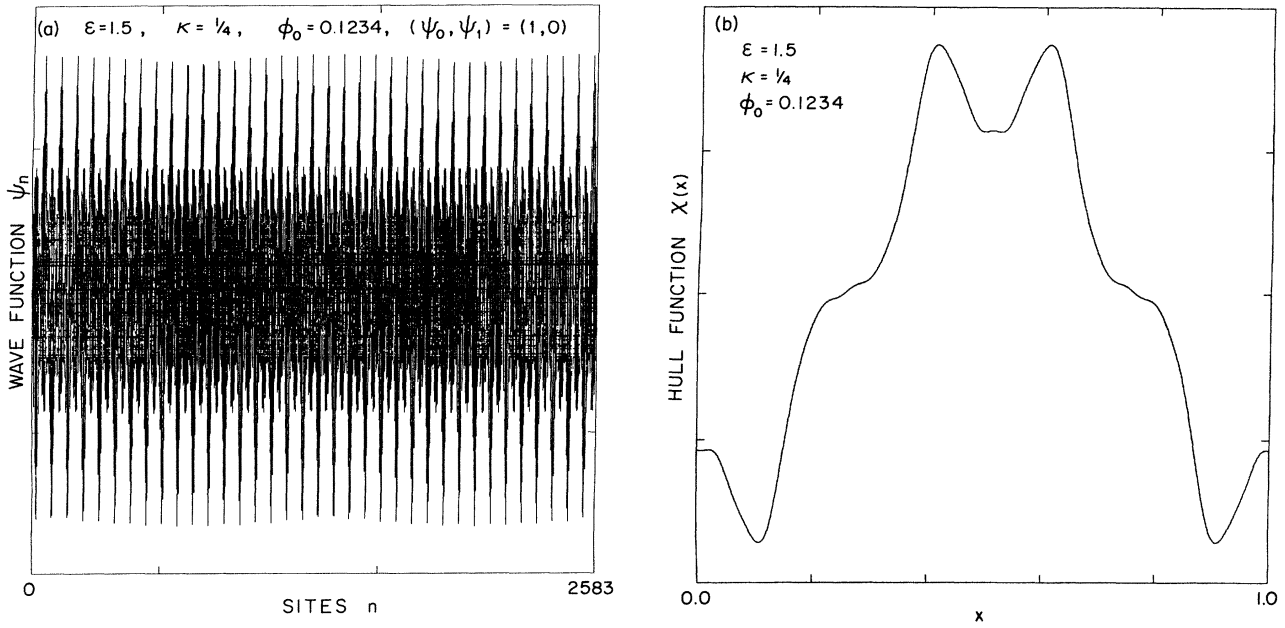


FIG. 3. (a) The wave function ψ_n at 2584 successive points is shown for arbitrary initial conditions on $[\psi_1,\psi_0]$, an arbitrary choice of $\varphi_0=0.1234$, $\epsilon=1.5$, and $E=2$. It appears to be extended. (b) The hull function $\chi(x)$ is plotted for the wave function in (a). Since $\kappa=\frac{1}{4}$ we recover the hull function by plotting $(-1)^n\psi_{2n}$ vs $\langle 2n\sigma_G\rangle$. The plot therefore includes half the points in (a), but the order of the points are rearranged by this transformation. The continuity of the hull function is further proof that the wave function is extended.

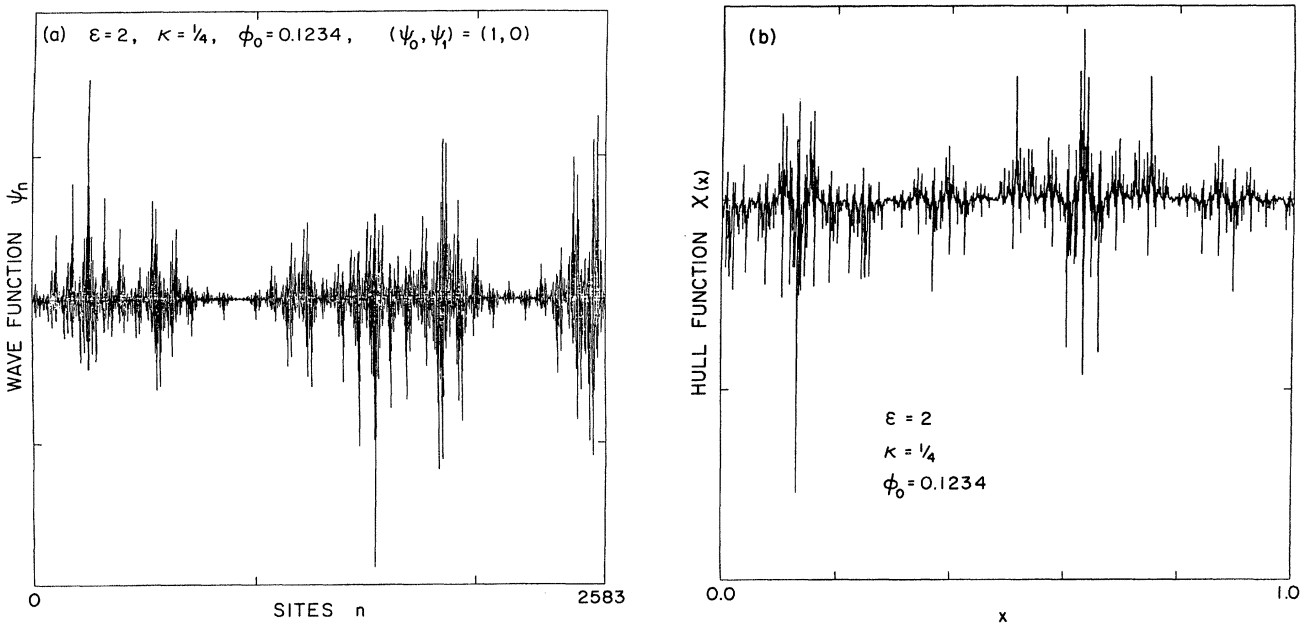


FIG. 4. (a) The wave function is shown for arbitrary initial conditions at the localization threshold $\epsilon=2$. Parameters are identical to those which generated Fig. 3(a) except for the value of ϵ. The wave function has structure which cannot be easily identified as localized or extended. The structure around successive maxima becomes self-similar for very long lattices, but in this figure this is not apparent since we have only included 2584 lattice sites. The "phase" $\langle n\sigma_G+\varphi_0\rangle$ at successive maxima approaches one of three well-defined constants $\frac{1}{4}+n\sigma_G$ where $n=0\pm1$. (b) The hull function for (a) appears to be highly discontinuous.

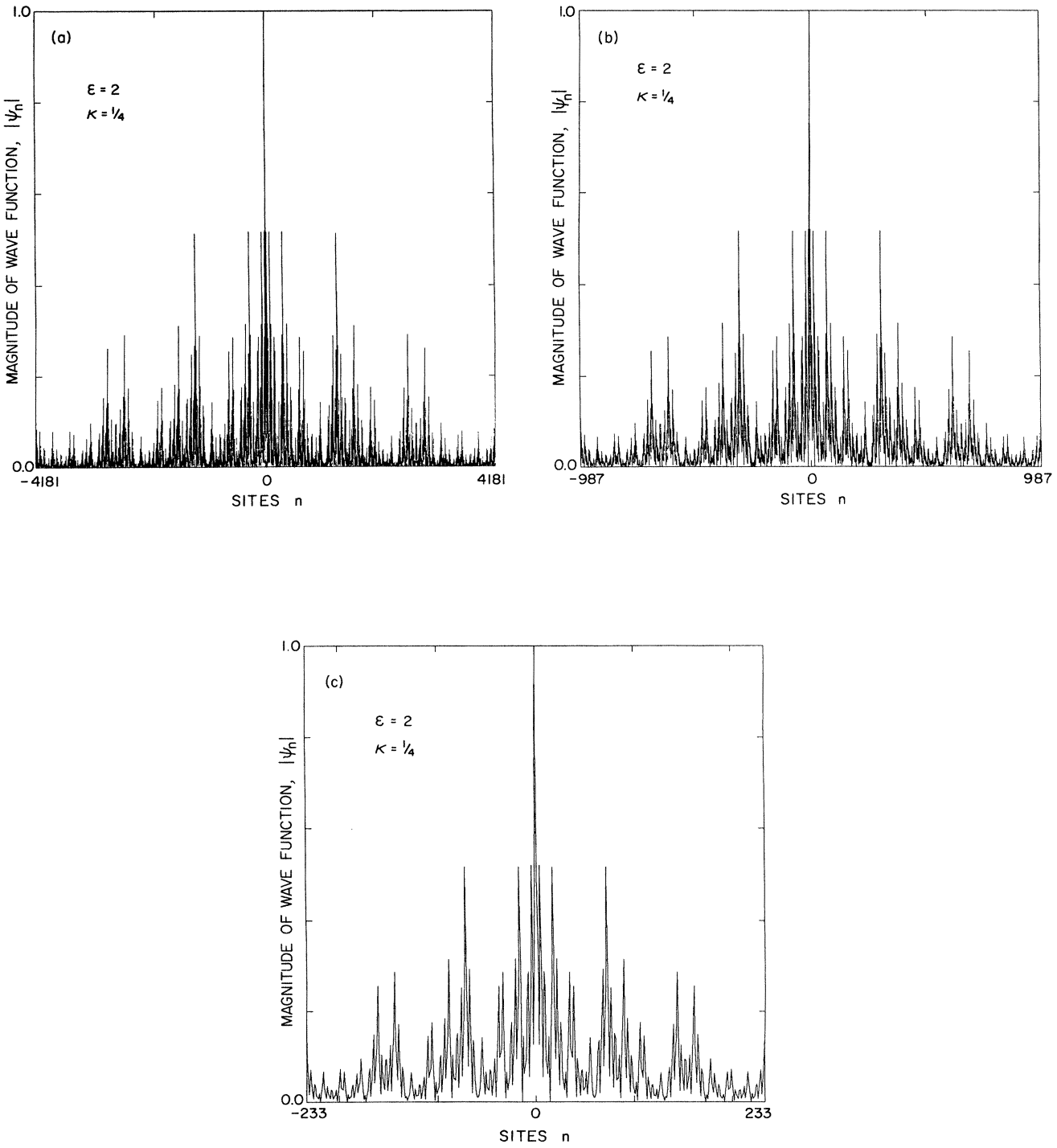


FIG. 5. (a) The magnitude of the wave function at $2F_{17}=8362$ sites around the site of a maximum of an extremely large (infinite) system is shown at the localization threshold $\epsilon=2$. Since this maximum occurs at the phase $\frac{1}{4}-\sigma_G$ we simply chose $\varphi_0=\frac{1}{4}-\sigma_G$ and plotted the magnitude of the wave function subject to the restriction that ψ_n remains finite at $\pm\infty$. The heights of the wave-function subpeaks far from the center approach a well-defined fraction ζ of the main peak height. Thus, the wave function never decays to zero, although the points where there is a large amplitude are further and further apart from the origin. The overall normalization of the wave function scales as a nonintegral but universal power β of the number of lattice points included in the sum. (b) The central portion ($2F_{14}=1974$ sites) of (a) is replotted. This displays the self-similarity of the wave function, since this picture corresponds to rescaling lengths by σ_G^3 relative to (a). (c) The central portions (466 sites) of (a) and (b) further illustrate the scale invariance.

$q_{14}=987$, and $q_{11}=233$ points, respectively, on either side of the origin. Clearly $|\psi_n|$ is scale invariant, since these pictures appear identical.

Actually, it is a little tricky to precisely define scale invariance on a discrete lattice, and closer inspection reveals that only the dominant features (large amplitudes) are in fact scale invariant. We define the scale invariance seen in these pictures by the following formula[33] ($p=3$):

$$\psi_n \to \psi_{[\sigma^p n+1/2]} \text{ as } \langle \sigma^p n \rangle \to 0 . \tag{2.14}$$

Thus, lattice sites which fall closely on top of other lattice sites under rescaling also have amplitudes which are approximately equal.

We note the following important property of $|\psi_n|$: The wave function does *not* vanish as $n \to \infty$ and, in fact, local maxima approach a constant fraction ζ of the peak height at $n=0$ for special points which are spaced further and further apart. The structure around the *subpeaks* approaches a length-rescaled version of the structure around zero, in the limit that the peak is infinitely far from zero. Thus we define the scale invariance by the formula

$$\lim_{k\to\infty} \psi_{\tilde{N}_k+n} = \psi_n \zeta , \tag{2.15}$$

where $\tilde{N}_k$ is the location of the kth subpeak and n is finite. The scaling relations (2.14) and (2.15) define the scale invariance of the structure, and there is thus a whole hierarchy of subpeaks with peak height ζ^m of the central peak. We find empirically that the scaling indices ζ and β (2.10) are $\zeta=0.315\pm0.005$ and $-\beta=0.639\pm0.005$.

4. $\epsilon \gtrsim 2$

When $\epsilon \gtrsim 2$, the states are localized. We retain the phase $\varphi_0=\frac{1}{4}$ and investigate the state in the vicinity of the origin with finite boundary conditions. The localization length l obeys[5] (see also *Note added*)

$$l \propto |\epsilon-\epsilon_c|^{-\nu} , \tag{2.16}$$

where $\nu=1$ and $\epsilon_c=2$. In particular, defining $\tilde{\epsilon}=\epsilon-\epsilon_c$, we find the following stronger scaling formula:[5]

$$\psi_n(\tilde{\epsilon}) \to \psi_{[\sigma^3 n+1/2]}(\tilde{\epsilon}/\sigma^3) \text{ as } \langle \sigma^3 n \rangle \to 0 . \tag{2.17}$$

Thus, if we rescale the length by σ^3 and move away from the localization potential by a factor of $1/\sigma^3$ in the parameter $\tilde{\epsilon}$, the plot of the wave function is unchanged. In Figs. 6(a), 6(b), and 6(c) we show $|\psi_n|$ on $233=q_{11}$ sites on either side of the origin for $\tilde{\epsilon}=0.05$, on 987 sites for $\tilde{\epsilon}=0.05\sigma^3$, and on 4181 sites for $\tilde{\epsilon}=0.05\sigma^6$, respectively, to illustrate this point.

D. Scaling of the matrices $\underline{M}_{q_k}(\varphi_0)$ for $\epsilon=2$

To understand the previous results, it is natural to inquire about the behavior of $\underline{M}_{q_k}(\varphi_0)$ [Eq. (2.3)], with $x_0 \equiv \varphi_0$ as usual. When $\epsilon=2$, we find that $\underline{M}_{q_k}(\varphi_0)$ varies greatly with φ_0 for most values of φ_0. For $\varphi_0=\frac{1}{4}+\frac{1}{2}n+\sigma m$, where the wave function for $\kappa=\frac{1}{4}$ attains a maximum, all entries of $\underline{M}_{q_k}$ go to infinity (as $k\to\infty$) at the same rate, i.e., their ratios are finite when n is a small integer. Thus, in contrast to the situation for $\epsilon<2$ when $\underline{M}_{q_k}$ remains bounded for all φ_0 and κ, φ_0 plays an *important* role in determining the asymptotic properties of $\underline{M}_{q_k}(\varphi_0)$ when $\epsilon=2$.

The divergence of $\underline{M}_{q_k}(\varphi_0)$ makes the analysis of these matrices particularly difficult. Thus, we ask another question: Is there a φ_0^* such that $\underline{M}_{q_k}(\varphi_0^*)$ does *not* diverge? To answer this question we simply search for the φ_0^* which minimizes

$$\lim_{k\to\infty} \sum_{i,j=1}^{2} \{[\underline{M}_{q_k}(\varphi_0^*)]^2\}_{ij}$$

and find, for $\kappa=\frac{1}{4}$, that $\varphi_0^*=-\frac{1}{4}+\frac{1}{2}\sigma$. We do not yet understand *why* this phase is singled out by this prescription, and present this as an experimental fact. Numerically, using this value of φ_0^* we find that

$$\lim_{k\to\infty} \underline{M}_{q_{6k}}(\varphi_0^*) = \underline{M}^* .$$

so that the matrices go to a six-cycle, with all members of the six-cycle having finite entries.

III. RENORMALIZATION OF THE CIRCLE MAP

A. Preliminaries

Since $V(x)$ is of period 1, only the value of x modulo 1 enters the equations, and $f(x)$ in Eq. (1.1) defines a mapping of the unit circle onto itself. We make some elementary observations and definitions based on this. The *winding number* $\rho(f)$ is defined to be

$$\rho(f) = \lim_{n\to\infty} \frac{x_n - x_0}{n} , \tag{3.1}$$

where $x_{n+1}=f(x_n)$. The special case $f(x)=x+\sigma$ gives $\rho=\sigma$. It is known that for invertible maps, $\rho(f)$ is independent of x_0. (See, for instance, Ref. 25.) The map f has a *cycle of length* q if we can find integers p and q and an x_0 such that $x_q=x_0+p$. (We only consider p and q relatively prime.)

We next define the *reduced map* $[\xi_0,\eta_0](x)$ which is equivalent to $\langle f(x) \rangle$ by the following formulas:

$$\begin{aligned} &\xi_0(x)=f(x) \text{ when } f(0)-1 \le x < 0 , \\ &\eta_0(x)=f(x)-1 \text{ when } 0 \le x < f(1) . \end{aligned} \tag{3.2}$$

The value of an arbitrary reduced map $[\xi,\eta](x)$ is defined by

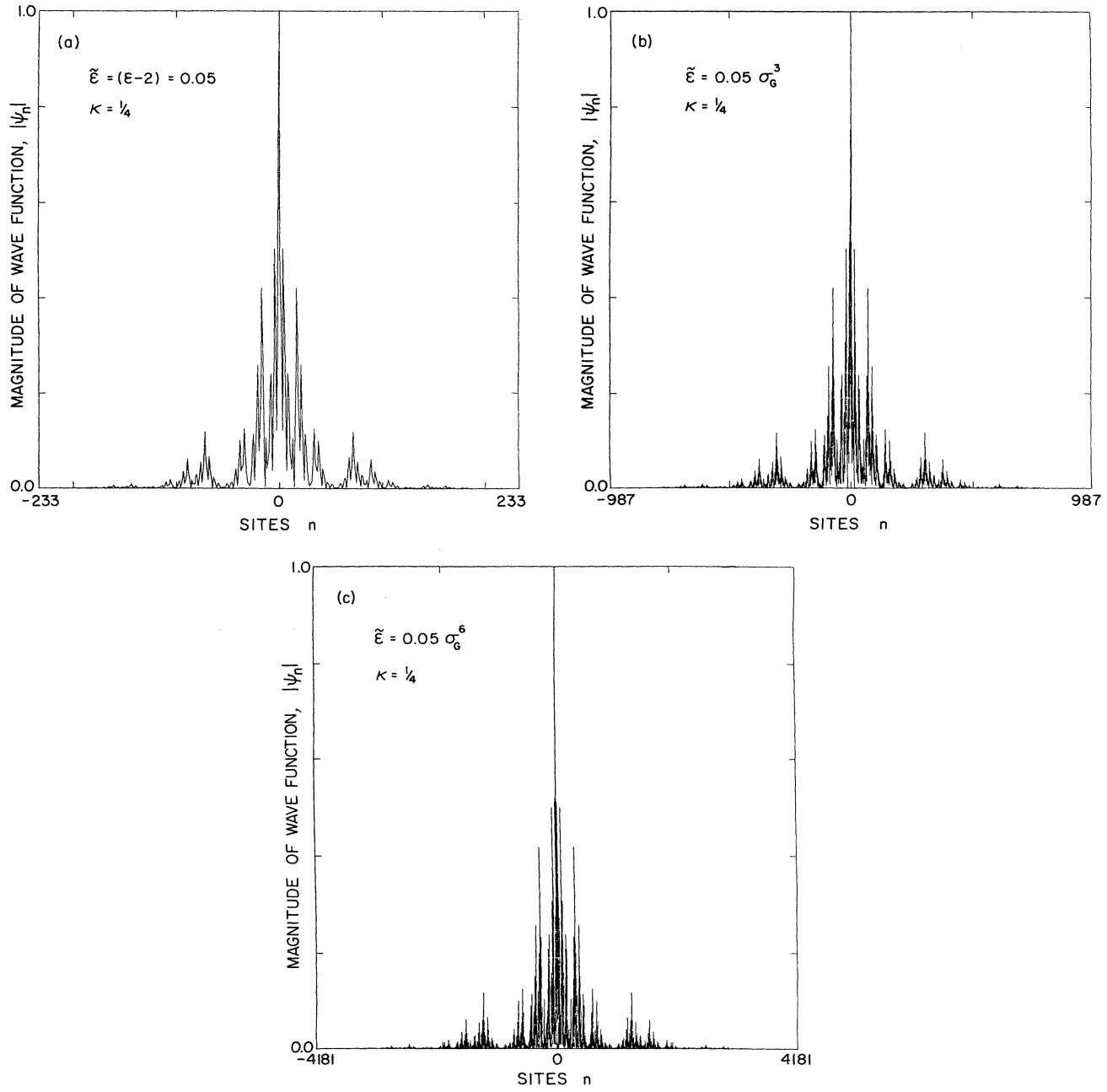


FIG. 6. The magnitude of the wave function around the site of localization is shown for (a) $\tilde{\epsilon}=(\epsilon-2)=0.05$ and $2F_{11}=466$ sites, (b) $\tilde{\epsilon}=0.05\sigma_G^3$ and $2F_{14}=1974$ sites, and (c) $\tilde{\epsilon}=0.05\sigma_G^6$ and $2F_{17}=8362$ sites. The phase φ_0 is again $\frac{1}{4}-\sigma_G$. Clearly graphs (a), (b), and (c) are essentially the same. This illustrates the similarity of the wave function at $\tilde{\epsilon}$ and the wave function at $\tilde{\epsilon}\sigma^{-3}$ after rescaling the length by σ^3. This scaling near the localization point is a universal of the localization transition.

$$[\xi,\eta](x)=\begin{cases}\xi(x) & \text{when } \eta(0)<x\leq 0,\\ \eta(x) & \text{when } 0<x\leq\xi(0).\end{cases} \tag{3.3}$$

The construction is shown in Fig. 7 for $f(x)=x+\sigma_G$ where $\sigma_G=(\sqrt{5}-1)/2$. We next define the functional commutator $C[\xi,\eta](x)$ as[44]

$$C[\xi,\eta](x)=\xi\eta(x)-\eta\xi(x)\ . \tag{3.4}$$

According to (3.2), $C[\xi_0,\eta_0]=0$ (see Ref. 45).

B. Renormalization transformation

We review in this section the renormalization group for the circle map as described by Rand, Ostlund, Sethna, and Siggia[25] (ROSS). We refer the reader to the original references for further details.

Given an arbitrary reduced map $[\xi,\eta]$, we define the integer $n[\xi,\eta]\equiv n$ by the *smallest* n such that the number α defined by

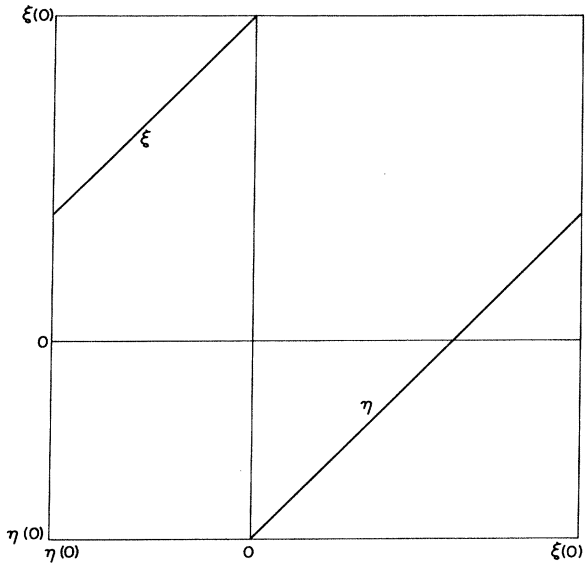


FIG. 7. Construction of the reduced circle map $[\xi,\eta](x)$ [Eqs. (3.2) and (3.3)]. In this figure $f(x)=x+\sigma_G$.

$$\alpha^{-1}=\xi^{n-1}\eta(0)-\xi^{n-1}\eta\xi(0) \tag{3.5}$$

is not greater than zero. The integer n is therefore the smallest number $n>0$ such that $[\xi,\eta]^n$ applied to the open-closed interval $I_0\equiv(0,\xi(0)]$ contains the origin. The *length* of the image of this interval is precisely $|\alpha^{-1}|$.

Given these values of n and α, we define the renormalization transformation T on $[\xi,\eta]$ by

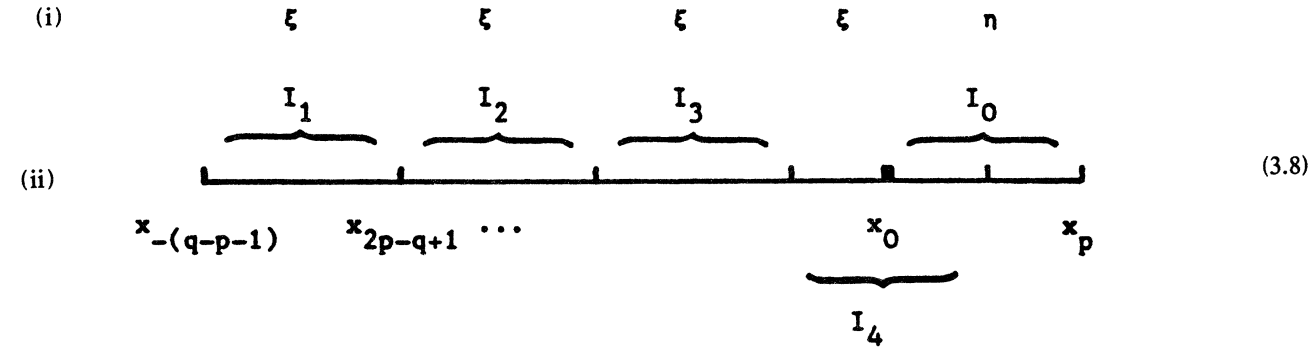
$$T[\xi,\eta](x)=\alpha[\xi^{n-1}\eta,\xi^{n-1}\eta\xi](x/\alpha)\ . \tag{3.6}$$

We remark that the space of reduced maps $[\xi,\eta]$ which obey $C([\xi,\eta])=0$ is closed under T, so that

$$C([\xi,\eta])=0 \Rightarrow C(T[\xi,\eta])=0\ . \tag{3.7}$$

We now explain what the transformation (3.6) accomplishes.

Let us suppose that $[\xi,\eta]$ has a cycle of length q and winding number p/q containing the origin. We define $x_0\equiv 0$ and label the points in the cycle by $\{x_{-(q-p-1)},\ldots,x_{-1},0,x_1,\ldots,x_p\}$. The function η is defined on the points $\{0,\ldots,x_p\}$, while ξ is defined on $\{x_{-(q-p-1)},\ldots,0\}$. (Note that *both* η and ξ are defined at 0). It is clear that $x_p=\xi(0)$, since on this set of points, $[\xi,\eta](x_m)=x_{(m+p)\,(\mathrm{mod}q)}$ must be valid for $[\xi,\eta]$ to have rotation number p/q. To illustrate this let us choose $n[\xi,\eta]=4$ for definiteness in the rest of this section. The intervals relevant to the problem are shown below.



(3.8)

As discussed by ROSS[25(b)] each interval I_0,I_1,I_2,I_3 gets mapped into I_1,I_2,I_3,I_4, respectively, where I_4 is the first interval which contains the origin. I_4 is the image of I_0 under $[\xi,\eta]^n$ and, according to (3.5) has length α^{-1}. Furthermore, according to (3.6) and (3.3), I_4 corresponds to the renormalized *domain* of $T[\xi,\eta]$, since the left-hand point is $\xi^3\eta(0)$, while the right-hand point is $\xi^3\eta\xi(0)$.

Note that under $T[\xi,\eta]$, the portion of the domain $I_4\cap I_0$ is mapped by $\xi^{n-1}\eta$ back into I_4, while I_4-I_0 is mapped by $\xi^{n-1}\eta\xi$ into I_4. Thus, $T[\xi,\eta]$ certainly maps I_4 *into* I_4. Furthermore, since each ξ,η was monotonic, so is $T[\xi,\eta]$, and the new mapping is in fact one-one on the interval I_4.

Since we have not destroyed any cycle structure, we can easily find out how the winding number has changed. The image of the origin under $\xi^{n-1}\eta$ is now the $(q-np)$th point to the *left* of the origin since $[\xi,\eta](x_m)=x_{(m+p)\,(\mathrm{mod}\,q)}$. [The fact that α is negative causes the left and right sides of the domain to be switched when rescaled by α in Eq. (3.6). This is important for the winding number to be positive.] The number of points in the complete interval I_4 is p, since it is the image of I_0, so the renormalization integers p' and q', which define the p'/q' cycle of $T[\xi,\eta]$, must obey

$$p'/q'=(q-np)/p\ .$$

Taking the limit where p and q may be large, we find

$$\rho'=\frac{1}{\rho}-n\ . \tag{3.9}$$

Thus, according to the Appendix, the *transformation removes exactly one integer from the continued-fraction expansion of* ρ. We observe that this construction does not require prior knowledge of ρ. We can find what ρ is by using the n calculated recursively by iterating the renormalization transformation. Thus defining

$$[\xi_k,\eta_k]=T^k[\xi_0,\eta_0] \tag{3.10}$$

we see that $\rho=1/\{n_1+1/[n_2+1/(n_3+\cdots)]\}$

Let us ask what the sequence of functions is which maps a point in the image interval I_4 through the lattice. Starting on the right at the origin of the lattice and working toward the left through the lattice points, each time we move to the left one lattice point we skip over p points ordered as (3.8). Each application of the reduced map acting on any lattice site maps x_m to $x_{(m+p)\,(\mathrm{mod}\,q)}$. Thus *if* η acts on x_m, the composition with η must be followed by composition with at least three ξ's. After the third ξ we *may* have another ξ and then an η must occur. Thus, denoting by a subscript ? the ξ's which may be present or absent from the string of functional compositions required to get successive x_m's, we find the following sequence of functional compositions (spaces have been put in to guide the eye only):

$$\cdots \xi\xi\xi\eta\xi_?\ \xi\xi\xi\eta\xi_?\ \xi\xi\xi\eta\xi_?\ \xi\xi\xi\eta\xi_?\cdots . \tag{3.11}$$

Any sequence of compositions is made of a string of blocks each one of which can be either $\xi\xi\xi\eta$ or $\xi\xi\xi\eta\xi$. These groupings are precisely the sequences in the renormalization transformation in (3.6) which define the renormalized ξ and η. This was first pointed by Feigenbaum and Hasslacher.[46]

IV. RENORMALIZATION OF THE EIGENVALUE PROBLEM

A. Transformation

Analogous to our definition of a reduced map constructed from a function pair $[\xi,\eta]$ we define an associated *matrix pair* whose entries are functions of x:

$$[\underline{B},\underline{A}](x)=\begin{cases}\underline{B}(x) & \text{when } \eta(0)<x\le 0,\\ \underline{A}(x) & \text{when } 0<x\le\xi(0).\end{cases} \tag{4.1}$$

Since we want $[\underline{B}_0,\underline{A}_0](x)$ to be equivalent to $\underline{M}_1(x)$, we choose $\underline{A}_0(x)=\underline{M}_1(x)$ and $\underline{B}_0(x)=\underline{M}_1(x)$ according to Eq. (2.2). The matrix $\underline{B}$ gets used at site m to map $[\psi,\psi_{m-1}]$ to $[\psi_{m+1},\psi_m]$ whenever the function ξ gets used to map x_m to x_{m+1}, and $\underline{A}$ gets used whenever η is used to map x_m to x_{m+1}. Thus, it follows by arguments *identical* to those used when we derived (3.1) that the string of matrices that maps $[\psi_1,\psi_0]$ into $[\psi_N,\psi_{N-1}]$ has the form [Ref. 25(b), Sec. IX]

$$\begin{bmatrix}\psi_N\\ \psi_{N-1}\end{bmatrix}=\cdots \underline{B}\,\underline{B}\,\underline{B}\,\underline{A}\,\underline{B}_?\,\underline{B}\,\underline{B}\,\underline{B}\,\underline{A}\,\underline{B}_?\,\underline{B}\,\underline{B}\,\underline{B}\,\underline{A}\,\underline{B}_?\cdots\begin{bmatrix}\psi_1\\ \psi_0\end{bmatrix} \tag{4.2}$$

when $n=4$. Here again the subscript ? denotes a matrix that may or may not be absent from each block. This string of $\underline{A}$'s and $\underline{B}$'s is derived by substituting an $\underline{A}$ for an η and a $\underline{B}$ for a ξ in (3.10), since $\underline{B}$ acts whenever ξ is used to map x_m and $\underline{A}$ acts whenever η is used. By considering how the argument of each matrix gets passed by $[\xi,\eta]$ to the argument of the next matrix, we find that the renormalization transformation induced by $T[\xi,\eta]$ is[45]

$$T\begin{bmatrix}\underline{B}\\ \underline{A}\end{bmatrix}(x)\equiv\begin{bmatrix}\underline{B}(\xi^{n-2}\eta)\cdots\underline{B}(\eta)\underline{A}(\underset{\sim}{1})\\ \underline{B}(\xi^{n-2}\eta\xi)\cdots\underline{B}(\eta\xi)\underline{A}(\xi)\underline{B}(\underset{\sim}{1})\end{bmatrix}(x/\alpha) . \tag{4.3}$$

The argument of the innermost function that occurs as the argument of each matrix is taken from the right-hand side of the bracket and as usual $\underset{\sim}{1}$ is the identity function $\underset{\sim}{1}(x)=x$. When $n=2$ or 1, Eq. (4.3) becomes

$$T_{n=1}[\underline{B},\underline{A}]=[\underline{A}(\underset{\sim}{1}),\underline{A}(\xi)\underline{B}(\underset{\sim}{1})](x/\alpha) , \tag{4.4a}$$

$$T_{n=2}[\underline{B},\underline{A}](x)=[\underline{B}(\eta)\underline{A}(1),\underline{B}(\eta\xi)\underline{A}(\xi)\underline{B}(\underset{\sim}{1})](x/\alpha) . \tag{4.4b}$$

As usual we denote $T^k[\underline{B}_0,\underline{A}_0]=[\underline{B}_k,\underline{A}_k]$.

There are two length scales $L_{\underline{A}}$ and $L_{\underline{B}}$ corresponding to the number of matrix multiplications of the original matrix $\underline{M}(x)$ that have been absorbed into $\underline{A}$ and $\underline{B}$. Thus

$$[L_{\underline{B}_{k+1}},L_{\underline{A}_{k+1}}]=[(n_k-1)L_{\underline{B}_k}+L_{\underline{A}_k},nL_{\underline{B}_k}+L_{\underline{A}_k}] . \tag{4.5}$$

We define the *renormalized* Bloch index pair normalized relative to the length scale of the *renormalized* lattice by $[\kappa_{\underline{B}_k},\kappa_{\underline{A}_k}]$. For these Bloch indices we find[33]

$$[\kappa_{\underline{B}_{k+1}},\kappa_{\underline{A}_{k+1}}]=[\langle(n_k-1)\kappa_{\underline{B}_k}+\kappa_{\underline{A}_k}\rangle,\langle n_k\kappa_{\underline{B}_k}+\kappa_{\underline{A}_k}\rangle] . \tag{4.6}$$

B. Conditions under which we can hope to find a fixed point

A necessary condition for a fixed point to exist for our recursion relation is that the sequence n_k defined by Eq. (3.5) eventually becomes periodic.[27] This is equivalent to having a periodic tail in the continued-fraction expansion of $\rho(f)$, which in turn is equivalent to ρ satisfying a quadratic equation with integer coefficients (see the Appendix). Let us assume that for $k>N$ we find that $n_k=n_{k+p_\rho}$.

This condition is necessary (but *not* sufficient) for $[\xi_k,\eta_k]$ to approach a cycle of length p_ρ as $k\to\infty$. However, the recursion of the complete eigenvalue problem cannot have a fixed point unless $[\underline{B}_k,\underline{A}_k]$ *also* converges to a cycle, which cannot be shorter than p_ρ. In order that we find a *matrix* cycle structure of length, say p_κ, it is necessary that the renormalized Bloch index $[\kappa_{\underline{A}_k},\kappa_{\underline{B}_k}]$ obeys

$$0=\lim_{k\to\infty}[\kappa_{\underline{B}_k},\kappa_{\underline{A}_k}]-[\kappa_{\underline{B}_{k+p_\kappa}},\kappa_{\underline{A}_{k+p_\kappa}}] .$$

It is then possible to have an overall cycle structure of length p where p is a common multiple of p_κ and p_ρ. The *existence* of a fixed point is by no means guaranteed, and this is simply a necessary condition in order that a cycle structure of length p to exist.

As an example, let us assume $\rho=\sigma_G$. In this case $n_k=1$ so that $p_\rho=1$ for all k. Since every sixth Fibonacci number is divisible by four, $p_{\kappa=1/4}=6$. Indeed, we shall see in Sec. V that we do find a six-cycle in this case, although if there are additional symmetries present, a subcycle of the matrix cycle may determine the scaling of the spectrum. This is discussed in more detail in Sec. VII A.

C. Commutators and conservation laws

The space in which we do our renormalization group includes (in principle) all 2×2 matrix pairs of functions $[\underline{B},\underline{A}]$. This space is much larger than the subspace of matrix pairs generated by $\underline{M}_1(x)$ [Eq. (2.1)]. There turns out to be *marginal and relevant* directions at the fixed points of our RG that take us out of the subspace of matrix pairs accessible by varying $\underline{M}_1(x)$. Using various constraints and commutation laws that must be obeyed by the space of matrices generated by $\underline{M}_1(x)=[\underline{B}_0,\underline{A}_0](x)$, we can eliminate these "unphysical" or "spurious" eigenvectors. This must be done in two steps. We must first identify the constraints and conservation laws, then show how their violation leads to relevant and marginal eigenvectors. These eigenvalues can be discarded because they represent variations which are inaccessible by using $\underline{M}_1(x)$ to define $[\underline{B}_0,\underline{A}_0]$.

D. Matrix-functional commutators

Analogous to (3.4), we define

$$\underline{C}([\underline{B},\underline{A}],[\xi,\eta])=\underline{A}(\xi)\underline{B}(\underset{\sim}{1})-\underline{B}(\eta)\underline{A}(\underset{\sim}{1}) . \tag{4.7}$$

It is straightforward to verify that

$$\underline{C}([\underline{B}_0,\underline{A}_0],[\xi_0,\eta_0])=0 \tag{4.8}$$

and somewhat tedious to prove that if $C([\xi,\eta])=0$ in Eq. (3.4) and $\underline{C}([\underline{B},\underline{A}][\xi,\eta])=0$ then

$$\underline{C}(T[\underline{B},\underline{A}],T[\xi,\eta])=0 .$$

It therefore follows that renormalized matrices remain in the subspace where

$$\underline{C}([\underline{B}_k,\underline{A}_k],[\xi_k,\eta_k])=0 . \tag{4.9}$$

Since $\det\underline{A}_0=\det\underline{B}_0=1$, it follows that

$$\det\underline{A}_k=\det\underline{B}_k=1 . \tag{4.10}$$

This constraint also determines a closed space. Note that (4.9) represents four functional constraints, one for each matrix entry, and (4.10) represents two constraints all of which may or may not be independent constraints. There are therefore three closed spaces of interest:

(i) S_{full} is the space of all pairs of matrices $[\underline{B}_k,\underline{A}_k]$ whose entries are functions;

(ii) $S_{\det}\subset S_{\text{full}}$ is the subspace of S_{full} consisting of unimodular matrix pairs;

(iii) $S_{\text{restrict}}\subset S_{\det}\subset S_{\text{full}}$ is the subspace of S_{full} obeying (4.9), containing the space generated by $\underline{M}_1(x)$.

The commutation relations (4.9) are closely related to *four invariants* in the subspace $S_{\det}$. Let us define arbitrary constant matrices $\underline{C}$ and $\underline{D}$ and the functions

$$I_{\underline{C}\,\underline{D}}([\underline{B},\underline{A}])=\det[\underline{C}\,\underline{A}(\xi)\underline{B}(\underset{\sim}{1})-\underline{D}\,\underline{B}(\eta)\underline{A}(\underset{\sim}{1})] . \tag{4.11}$$

Further, define m^{ij} by

$$(m^{ij})_{kl}=\delta_{ik}\delta_{jl} . \tag{4.12}$$

In the space $S_{\det}$, a rather tedious calculation shows that there are four independent functions $I_{1,1}$, $I_{m^{11}m^{21}}$, $I_{m^{11}m^{22}}$, and $I_{m^{12}m^{22}}$. A straightforward but somewhat tedious calculation shows that if $C[\xi,\eta]=0$ [(Eq. (3.7)], then

$$I_{\underline{C}\,\underline{D}}T^2[\underline{B},\underline{A}](x)=I_{\underline{C}\,\underline{D}}[\underline{B},\underline{A}](x/\alpha^2) , \tag{4.13}$$

even if Eq. (4.9) is violated. α^2 is the product of the two successive α's.

Given that we are in the space S_{restrict}, it follows that $I_{\underline{C}\,\underline{D}}[\underline{B},\underline{A}]=\det(\underline{C}-\underline{D})$ from which it follows that $I_{11}=I_{m^{11}m^{21}}=I_{m^{12}m^{22}}=0$ and $I_{m^{11}m^{22}}=-1$.

Equation (4.13) shows that there are *four* invariants in the *larger* space $S_{\det}$, namely

$$I_{\underline{C}\,\underline{D}}(T^2[\underline{B},\underline{A}])(0)=I_{\underline{C}\,\underline{D}}[\underline{B},\underline{A}](0) . \tag{4.14}$$

This equation is important, because it generates a *conservation* law in the space $S_{\det}$ under the operation T^2. This, in turn, generates marginal operators in $S_{\det}$. It is important to keep track of which space one is dealing with. Being constant, $I_{\underline{C}\,\underline{D}}(x)$ is conserved for all x in S_{restrict}, but not $S_{\det}$, while $I_{\underline{C}\,\underline{D}}(0)$ is conserved in $S_{\det}$ but not S_{full}.

V. IMPLEMENTATION OF THE RG

A. Setting up the transformation

Working with Harper's equation $V(x_n)=\cos[2\pi(\sigma_G n+\varphi_0)]$, $f(x)=\sigma_G+x$ and $x_0=\varphi_0$ we rederive the results of Sec. II within the RG framework. In this special case, it can be verified that the following equalities hold for all k:

$$\begin{aligned}&\alpha_k=\alpha=-\sigma_G^{-1} ,\\ &n_k\equiv n=1 ,\\ &\xi_k(x)=\xi(x)\equiv x+\sigma_G ,\\ &\eta_k(x)=\eta(x)\equiv x-\sigma_G^2 .\end{aligned} \tag{5.1}$$

Furthermore, the Fibonacci integers (see the Appendix) define the length scales $[L_{\underline{B}_k},L_{\underline{A}_k}]$ to be

$$L_{\underline{B}_k}=F_{k-1},\quad L_{\underline{A}_k}=F_k\ . \tag{5.2}$$

Equations (4.3) and (4.4a) reduce to

$$T\begin{bmatrix}\underline{B}\\ \underline{A}\end{bmatrix}(x)=\begin{bmatrix}\underline{A}(-\sigma_G x)\\ \underline{A}(-\sigma_G x+\sigma_G)\underline{B}(-\sigma_G x)\end{bmatrix}\ . \tag{5.3}$$

We represent the matrices $\underline{A}$ and $\underline{B}$ as power series in x,

$$\begin{aligned}B_{ij}(x)&=\sum_{n=0}^{N-1}b_{ij}(n)x^n\ ,\\ A_{ij}(x)&=\sum_{n=0}^{N-1}a_{ij}(n)x^n\ ,\end{aligned} \tag{5.4}$$

and perform the operations (5.3) exactly using (5.4), each time truncating to the basis (5.4).

B. Finding the fixed point

To start our renormalization procedure, we choose N and define $\underline{A}_0(x)=\underline{B}_0(x)$ according to Eq. (5.4) by keeping the first N terms in a Taylor-series expansion of $\epsilon\cos[2\pi(x+\varphi_0)]$. As expected by the discussion in Sec. II, we choose $\varphi_0=\frac{1}{4}-\frac{1}{2}\sigma$ in order that $\underline{A}_k(x)$ and $\underline{B}_k(x)$ remain finite for $\epsilon=2$ as $k\to\infty$. We observe that $[\underline{B}_k,\underline{A}_k]$ approaches a six-cycle for large k, and we monitor numerically quantity Δ_k defined by

$$\begin{aligned}\Delta_k^2=\frac{1}{8N}\sum_{i,j=1}^{2}\sum_{n=0}^{N-1}&[a_{ij}^k(n)-a_{ij}^{k-6}(n)]^2\\ &+[b_{ij}^k(n)-b_{ij}^{k-6}(n)]^2\ .\end{aligned} \tag{5.5}$$

We find that Δ_k attains a minimum at, say, $k=k_{\rm opt}$, because of the approximations introduced by numerical roundoff and the truncation to finite N. Using 16-figure numerical operations and $N=30$, we find $k_{\rm opt}=31$ and $\Delta_{k_{\rm opt}}\simeq 10^{-6}$, whereas if quadruple precision (32 figures) is used with $N=60$, $\Delta_{k_{\rm opt}}\simeq 10^{-9}$ at $k_{\rm opt}=60$.

We use the value $[\underline{B}_{k_{\rm opt}},\underline{A}_{k_{\rm opt}}]$ as the best approximation to the fixed point of T^6. This is accurate to nine decimal places. We have *not* used a Newton-Raphson method for finding the solution to the equation $T^6[\underline{B}^*,\underline{A}^*]=[\underline{B}^*,\underline{A}^*]$ because severe complications are introduced by the many marginal eigenvalues and the associated many-dimensional surface of solutions of the fixed-point equations. Previous methods that have used projections[23–26] to get rid of unphysical marginal and relevant directions have been unsuccessful because of the complexity of the commutation constraints.

As an aside, we note that since the sixtieth Fibonacci number $F_{60}\simeq 4\times 10^{12}$, with a few seconds of IBM 370 computer CPU (central processing unit) time, we have multiplied together 4×10^{12} matrices by iterating T 60 times. By representing an arbitrary integer $q\lesssim 4\times 10^{12}$ as a sum of Fibonacci numbers, we can compute $\underline{M}_q(x_0)$ with a similar amount of computer time. Multiplying so many matrices would be impossible by the ordinary transfer-matrix methods described in Sec. II. The renormalization transformation is thus a powerful tool simply to multiply together large numbers of matrices, even in the absence of a fixed-point analysis.

C. Numerical analysis of the fixed point

1. *Analyzing the Jacobian in S_{full}*

Given our value of $[\underline{B}^*,\underline{A}^*](x)$, we must calculate the Jacobian $D_{T^6}[\underline{B}^*,\underline{A}^*]$ in order to obtain its eigenvalues. We do this using the basis provided by Eq. (5.4) and again truncate at Nth order. Since T is fairly simple, we can calculate D_T analytically as a function of $[\underline{B},\underline{A}]$ and do not have to resort to numerical differentiation. There are eight functions which define $[\underline{B},\underline{A}]$ so, when $N=30$, we get a 240×240 Jacobian matrix (real, general) to diagonalize. We use standard numerical routines to perform the diagonalization. This task is not trivial since we are interested in *all* eigenvalues whose magnitudes are in the vicinity of, or greater than, one. Although we use the fixed point obtained from $T^{60}[\underline{B}_0,\underline{A}_0]$ using quadruple precision with $N=60$, we truncate the fixed point to $N=30$ and use double precision to compute and diagonalize the Jacobian. The reason for doing this is that the Jacobian obtained for $N=60$ is too large to diagonalize. Also, the difference between this approximation to the fixed point and the true fixed point is of the order of the errors involved in the truncation.

2. *Analyzing the Jacobian in $S_{restrict}$*

In Sec. IV, we showed that the values of certain commutators were conserved under T. Since many of the eigenvectors calculated in $S_{\rm full}$ correspond to a violation of these conservation laws, we must eliminate these in order to understand physically relevant variations in the starting parameters. Conceptually, the simplest way to do this is to somehow parametrize the space $S_{\rm restrict}$, since it is closed under T, and use this parametrization to compute the Jacobian. However, we have not been able to discover a convenient parametrization. Instead, we compute D_{T^6} in $S_{\rm full}$, and then locally form a new orthonormal basis which separates the tangent space at the fixed point into basis vectors parallel to and orthogonal to the tangent space of $S_{\rm restrict}$.

In order to perform this change of basis, define the six functionals $\mathcal{F}_1,\ldots,\mathcal{F}_6$ by

$$\mathcal{F}_1([\underline{B},\underline{A}])=[\underline{A}(\xi)\underline{B}(\underset{\sim}{1})-\underline{B}(\xi)\underline{A}(\underset{\sim}{1})]_{11}\ , \tag{5.6a}$$

$$\vdots$$

$$\mathcal{F}_4([\underline{B},\underline{A}])=[\underline{A}(\xi)\underline{B}(\underset{\sim}{1})-\underline{B}(\xi)\underline{A}(\underset{\sim}{1})]_{22}\ . \tag{5.6b}$$

The definitions for $\mathcal{F}_2$ and $\mathcal{F}_3$ are analogous, and $\mathcal{F}_5$ and $\mathcal{F}_6$ are given by

$$\mathcal{F}_5([\underline{B},\underline{A}])=\det\underline{B}-1\ , \tag{5.6c}$$

$$\mathcal{F}_6([\underline{B},\underline{A}])=\det\underline{A}-1\ . \tag{5.6d}$$

The six functional constraints (4.9) and (4.10) can be written in the basis (5.4) as

$$\mathcal{F}_n([\underline{B},\underline{A}])=\sum_{l=0}^{N-1}\mathcal{F}_n^l(\{a_{ij}(m),b_{ij}(m)\})x^l=0\ , \tag{5.7}$$

where $1\le n\le 6$ and $0\le l<N$. These equalities can be expanded into $6N$ simple equations, which we can write as

$$0=\mathcal{F}_n^l(\{a_{ij}(n),b_{ij}(n)\}) . \tag{5.8}$$

A necessary and sufficient condition that an eigenvector $(\underline{Y},\underline{X})$ of $D_{T^6}[\underline{B}^*,\underline{A}^*]$ be tangent to the space S_{restrict} is that

$$D\mathcal{F}_n[\underline{B}^*,\underline{A}^*]\cdot(\underline{Y},\underline{X})=0 \tag{5.9}$$

for all n. In the polynomial basis, if we write $X(x)=\sum_n X_n x^n$ and a similar expression for Y, this is equivalent to the $6N$ constraints

$$0=\sum_{m=0}^{N}\left[\frac{\partial\mathcal{F}_n^l}{\partial a_{ij}(m)}X_m+\frac{\partial\mathcal{F}_n^l}{\partial b_{ij}(m)}Y_m\right]$$

$$\equiv D_{\mathcal{F}_n^l}[\underline{B}^*,\underline{A}^*]\cdot(\underline{Y},\underline{X}) \tag{5.10}$$

for each $1\le n\le 6$, $0\le l<N$. Since

$$D_{\mathcal{F}_n^l}([\underline{B}^*,\underline{A}^*])\cdot D_{\mathcal{F}_n^l}([\underline{B}^*,\underline{A}^*])>0 \tag{5.11}$$

the $6N$ vectors $D_{\mathcal{F}_n^l}$ are necessarily *not* in the tangent space of S_{restrict}.

We now use the Gram-Schmidt orthogonalization procedure, starting with the $6N$ vectors $D_{\mathcal{F}_n^l}$ as basis vectors of S_{full}. This set of vectors has rank L, where $L\le 6N$, is not in the tangent space of S_{restrict}, and does not form a complete basis. By choosing further basis vectors at random, while continuing to use the Gram-Schmidt procedure, we complete the space with $8N-L$ vectors which *are* in the tangent space of S_{restrict}. By construction, these last $8N-L$ vectors satisfy Eq. (5.9). We denote the complete basis of S_{full} by $\{t_s\}_{s=0}^{s=8N}$ where each t_s is a linear combination of the $a_{ij}(n)$ and $b_{ij}(n)$ of Eq. (5.9). [It turns out that the restrictions (5.10) are not independent and in fact the set of equations have at most rank $5N$. This does not change the Gram-Schmidt procedure as long as we are careful not to include vectors which have length zero after projections into the previous vectors are removed by the orthogonalization procedure. The linear dependence of these equations stems from the fact that the four commutators, Eqs. (5.6a) and (5.6b), are not independent in the space of all pairs of unimodular matrices.] By construction we then find

$$D_{\mathcal{F}_n^l}(t_s)\begin{cases}\neq 0 & \text{if } 1\le s\le L ,\\ =0 & \text{if } L<s<8N .\end{cases} \tag{5.12}$$

Since the constraints $\mathcal{F}_n([\underline{B},\underline{A}])=0$ are closed under T^6, (i.e., $\mathcal{F}_n([\underline{B},\underline{A}])=0$ for all n implies $\mathcal{F}_n(T[\underline{B},\underline{A}])=0$ for all n) the Jacobian matrix $\underline{J}_{st}$ in the basis $\{t_a\}$ takes the block form

$$\begin{bmatrix}(\underline{J}_{st})_{L\times L} & \underline{0}_{L\times(8N-L)}\\ (\underline{J}_{st})_{(8N-L)\times N} & (\underline{J}_{st})_{(8N-L)\times(8N-L)}\end{bmatrix} .$$

Since $\{t_n\}_{n=L}^{n=8N}$ is a basis of S_{restrict}, it is sufficient to diagonalize the sub-block J_{st}, where $s>L$ and $t>L$, in order to pull out the eigenvectors and eigenvalues in the restricted space. The eigenvectors which are not in S_{restrict} we call "unphysical" or "spurious." As a consistency check on the numerical calculations and to further understand these results, we justify by *analytic* arguments in the next section, which eigenvectors of S_{full} are eliminated by restricting the vectors to lie in the tangent space of S_{restrict} rather than S_{full}.

D. Analytic investigation of spurious eigenvectors

We first show how a spurious marginal and a spurious relevant eigenvector are eliminated by restricting the space S_{full} to $S_{\text{det}}\supset S_{\text{restrict}}$ (Sec. IV C). Assume that $[\underline{B}_0^*,\underline{A}_0^*]$ is a fixed point of T^p for some integer p (i.e., $\underline{B}_0^*=\underline{B}_p^*$, etc.) and that $[\underline{B}_0^*,\underline{A}_0^*]\in S_{\text{det}}$ and $T([\underline{B}_k^*,\underline{A}_k^*])=[\underline{B}_{k+1}^*,\underline{A}_{k+1}^*]$. If $\delta_{\underline{A}}(x)$ and $\delta_{\underline{B}}(x)$ are infinitesimal arbitrary functions, to first order,

$$T\begin{bmatrix}[1+\delta_{\underline{B}}(x)]\underline{B}_k^*\\ [1+\delta_{\underline{A}}(x)]\underline{A}_k^*\end{bmatrix}=\begin{bmatrix}[1+\delta_{\underline{A}}(\underset{\sim}{\alpha}^{-1})]\underline{B}_{k+1}^*(\underset{\sim}{1})\\ [1+\delta_{\underline{A}}(\underset{\sim}{\xi}\underset{\sim}{\alpha}^{-1})+\delta_{\underline{B}}(\underset{\sim}{\alpha}^{-1})]\underline{A}_{k+1}^*(\underset{\sim}{1})\end{bmatrix} . \tag{5.13}$$

Hence, T separates into two parts, with T acting on the infinitesimal function piece as

$$T\begin{bmatrix}\delta_{\underline{B}}\\ \delta_{\underline{A}}\end{bmatrix}=\begin{bmatrix}\delta_{\underline{A}}(\underset{\sim}{\alpha}^{-1})\\ \delta_{\underline{A}}(\underset{\sim}{\xi}\underset{\sim}{\alpha}^{-1})+\delta_{\underline{B}}(\underset{\sim}{\alpha}^{-1})\end{bmatrix} \tag{5.14}$$

and T acting on the $[\underline{B}_k^*,\underline{A}_k^*]$ piece as usual. Let us assume that

$$\frac{d}{dx^m}\delta_{\underline{A}}(x)=\frac{d}{dx^m}\delta_{\underline{B}}(x)=0 \tag{5.15}$$

for every $m>m_0$, such that $\delta_{\underline{A}}(x)=\sum_{n=0}^{m_0}(\delta_{\underline{A}})_n x^n$. We ask if there is a solution to the equation

$$T\begin{bmatrix}\delta_{\underline{B}}\\ \delta_{\underline{A}}\end{bmatrix}=\tilde{\lambda}\begin{bmatrix}\delta_{\underline{B}}\\ \delta_{\underline{A}}\end{bmatrix} . \tag{5.16}$$

This equation reduces to[45]

$$\delta_{\underline{A}}(\underset{\sim}{\xi}\underset{\sim}{\alpha}^{-1})+\frac{1}{\tilde{\lambda}}\delta_{\underline{A}}(\underset{\sim}{\alpha}^{-2})=\tilde{\lambda}\delta_{\underline{A}}(\underset{\sim}{1}) . \tag{5.17}$$

By differentiating this equation m_0 times, at $x=0$ we find that there are two solutions

$$\tilde{\lambda}=(\alpha)^{-m_0}\frac{1}{\sigma_G}$$

and

$$\tilde{\lambda}=(\alpha)^{-m_0}(-\sigma_G) \tag{5.18}$$

with $\alpha=-(\sigma_G)^{-1}$. This means that there exist functions $(\delta_{\underline{B}},\delta_{\underline{A}})$ such that $(\delta_{\underline{B}}(x)\underline{B}^*_0(x),\delta_{\underline{A}}(x)\underline{A}^*_0(x))$ are eigenvectors of $D_{T^p}[\underline{B}^*_0,\underline{A}^*_0]$ with eigenvalue λ^p. There are therefore two series of spurious eigenvalues equal to

$$\lambda=\begin{cases}(-1)^p(-\sigma_G)^{p(m_0-1)}\\(-\sigma_G)^{p(m_0+1)}\end{cases} \tag{5.19}$$

Of these, there are two that are marginal or relevant: $(\sigma_G)^{-p}$ and $(-1)^p$. In our case $p=6$, so we find $+1$ and σ_G^{-6} as spurious eigenvalues.

Let us now consider a point $[\underline{B}^*_0,\underline{A}^*_0]\in S_{\text{restrict}}$ and let $(\underline{Y},\underline{X})$ be an eigenvector of eigenvalue λ of T^6 in the tangent space of S_{det}. By Eq. (4.13) it follows that

$$(D_{I_{\underline{C}\underline{D}}}[\underline{B}^*_0,\underline{A}^*_0]\cdot(\underline{Y},\underline{X}))(x)$$
$$=(D_{I_{\underline{C}\underline{D}}}T^6[\underline{B}^*_0,\underline{A}^*_0]\cdot(\underline{Y},\underline{X}))(x/\alpha^6)\ . \tag{5.20}$$

Using the chain rule, the fact that $T^6[\underline{B}^*_0,\underline{A}^*_0]=[\underline{B}^*_0,\underline{A}^*_0]$, and our assumption that $(\underline{Y},\underline{X})$ is an eigenvector of D_{T^6}, we rewrite this as

$$D_{I_{\underline{C}\underline{D}}}[\underline{B}^*_0,\underline{A}^*_0)\cdot(\underline{Y},\underline{X})(x)$$
$$=\lambda(D_{I_{\underline{C}\underline{D}}}[\underline{B}^*_0,\underline{A}^*_0]\cdot(\underline{Y},\underline{X}))(x/\alpha^6)\ . \tag{5.21}$$

Let us now assume that $(\underline{Y},\underline{X})$ is *not* in the tangent space to S_{restrict}. It follows that

$$D_{I_{\underline{C}\underline{D}}}[\underline{B}^*_0,\underline{A}^*_0](\underline{Y},\underline{X})\neq 0\ . \tag{5.22}$$

There is, therefore, some m_0 where

$$\frac{d^{m_0}}{dx^{m_0}}D_{I_{\underline{C}\underline{D}}}[\underline{B}^*_0,\underline{A}^*_0](\underline{Y},\underline{X})\bigg|_{x=0}\neq 0\ . \tag{5.23}$$

Differentiating (5.21) m_0 times and using (5.23), it follows that

$$\lambda=\alpha^{-6m_0}\ . \tag{5.24}$$

Thus any eigenvector which violates (4.14) is at worst marginal. We can, therefore, expect to find a set of spurious marginal eigenvectors in S_{det} that violate the commutativity restrictions in Sec. IV.

E. Similarity transformations and redundant eigenvalues

Let us assume that $[\underline{B}^*_0,\underline{A}^*_0]$ solves the fixed-point equation for T^p. Clearly, if $\underline{S}$ is a constant invertible matrix, then $\underline{S}^{-1}[\underline{B}^*_0,\underline{A}^*_0]\underline{S}$ solves the fixed-point equations as well. To each independent generator of simultaneous similarity transformations there corresponds a marginal eigenvector. A change induced by this eigenvector can be transformed away by a change of the basis that represents the wave function. Such a marginal eigenvector is called redundant since it cannot affect the exponents or spectrum. It is therefore necessary to understand the dimension of the space of redundant eigenvectors so we know how many marginal eigenvectors can be eliminated when we obtain a list of relevant and marginal eigenvalues in the tangent space of S_{full}.

For clarity, instead of considering arbitrary matrix functions, let us consider the moment pairs of constant matrices $(\underline{\tilde{b}},\underline{\tilde{a}})$. Let us assume that $\underline{\tilde{a}}$ can be diagonalized. We can then without further loss of generality, assume that $\underline{S}_0$ diagonalizes $\underline{\tilde{a}}$ and replace $(\underline{\tilde{b}},\underline{\tilde{a}})$ by $\underline{S}_0^{-1}(\underline{\tilde{b}},\underline{\tilde{a}})S_0$. The dimension of the space of matrix pairs similar to $(\underline{\tilde{b}},\underline{\tilde{a}})$ is the same as the dimension of the space

$$\underline{S}^{-1}\left[\underline{b},\begin{bmatrix}\lambda_1&0\\0&\lambda_2\end{bmatrix}\right]\underline{S}\ , \tag{5.25}$$

where $\underline{S}$ is arbitrary. The space of generators of such similarity transformations consists of all traceless matrices $\underline{M}$, where we write $\underline{S}_\delta\equiv\underline{1}+\delta\underline{M}$. Then for each matrix $\underline{M}$, $(\underline{Y},\underline{X})$ defined by

$$(\underline{Y},\underline{X})\equiv\frac{d}{d\delta}\underline{S}_\delta^{-1}(\underline{b},\underline{a})\underline{S}_\delta\bigg|_{\delta=0} \tag{5.26}$$

lies in the tangent space of the space of matrix pairs similar to $(\underline{b},\underline{a})$. It is trivial to check that

$$(\underline{Y},\underline{X})=([\underline{b},\underline{M}]_c,[\underline{a},\underline{M}]_c)\ , \tag{5.27}$$

where $[\ ]_c$ means ordinary matrix commutator. Using the explicit form for $(\underline{b},\underline{a})$ from Eq. (5.25) and parametrizing $\underline{M}$ by

$$\underline{M}=\begin{bmatrix}x&y\\z&-x\end{bmatrix}\ , \tag{5.28}$$

we find

$$\underline{Y}=\begin{bmatrix}b_{12}z-b_{21}y&(b_{11}-b_{22})y-2b_{21}x\\(b_{22}-b_{11})z+2b_{12}x&(b_{21}y-b_{12}z)\end{bmatrix},$$
$$\underline{X}=(\lambda_1-\lambda_2)\begin{bmatrix}0&y\\-z&0\end{bmatrix}. \tag{5.29}$$

Therefore, the redundant space is

(a) zero-dimensional if both $\underline{b}$ and $\underline{a}$ are multiples of the identity,

(b) two-dimensional if $[\underline{b},\underline{a}]_c=0$ which holds if $b_{12}=0$ and $b_{21}=0$ (Ref. 48), or

(c) three-dimensional if $[\underline{b},\underline{a}]_c\neq 0$ which holds if b_{12} or $b_{21}\neq 0$.

For our problem, we must extend the above argument to the two infinite sets of matrix pairs $\{\underline{b}^*_0(n),\underline{a}^*_0(m)\}_{n,m=0}^{n,m=\infty}$ defined by Eq. (5.4). Since $\underline{S}^{-1}[\underline{B},\underline{A}]\underline{S}$ acts on each element in $\underline{b}^*_0(n)$, $\underline{a}^*_0(m)$ separately, it follows that if there is *any* $\underline{a}^*(m)$ or $\underline{b}^*(n)$ that can be diagonalized [say $\underline{a}^*_0(0)$], unless $0=[\underline{b}^*_0(n),\underline{a}^*_0(0)]_c=[\underline{a}^*_0(n),\underline{a}^*_0(0)]_c$ for *all* n, the redundant space is three-dimensional. A likely exception to this arises if the fixed point consists of *constant* matrices, which then must commute according to (4.7) and (4.9). In that case the redundant space is two-dimensional.

F. Coupling of starting parameters to eigenvectors

Let us assume that $[\underline{B}_0,\underline{A}_0](x;\tau)$, as a function of τ, represents a variation in the functional form of $[\underline{B}_0,\underline{A}_0]$ and that

$$\lim_{k\to\infty} T^{kp}[\underline{B}_0,\underline{A}_0](x;0)=[\underline{B}_0^*,\underline{A}_0^*](x)\ . \tag{5.30}$$

In the RG description $[\underline{B}_0,\underline{A}_0](x;0)$ flows to the fixed point under applications of T and $\tau=0$ represents the critical coupling. Let us now vary τ slightly, so that

$$[\underline{B}_0,\underline{A}_0](x;\tau)=[\underline{B}_0,\underline{A}_0](x;0)+\tau(\underline{Y},\underline{X})_{,\tau}+O(\tau^2)+\cdots\ , \tag{5.31}$$

where of course

$$(\underline{Y},\underline{X})_{,\tau}=\frac{d}{d\tau}[\underline{B}_0,\underline{A}_0](x;\tau)\bigg|_{\tau=0}\ . \tag{5.32}$$

In particular, for our initial parametrization, $V(x)=\epsilon\cos[2\pi(x+\varphi_0)]$, $(\underline{Y},\underline{X})$ would be any one of

$$\begin{aligned}(\underline{Y},\underline{X})_{,E}&=\frac{d}{dE}[\underline{B}_0,\underline{A}_0]\bigg|_{E=2}\ ,\\ (\underline{Y},\underline{X})_{,\varphi_0}&=\frac{d}{d\varphi_0}[\underline{B}_0,\underline{A}_0]\bigg|_{\varphi_0=\varphi_0^*}\ ,\\ (\underline{Y},\underline{X})_{,\epsilon}&=\frac{d}{d\epsilon}[\underline{B}_0,\underline{A}_0]\ ,\end{aligned} \tag{5.33}$$

where $[\underline{B}_0,\underline{A}_0](x;\varphi_0,E,\epsilon)$ is given by (2.2).

By repeated applications of the chain rule it is easy to see that for $k\gg 1$

$$\begin{aligned}\frac{d}{d\tau}T^{pk}[\underline{B}_0,\underline{A}_0](x,\tau)\bigg|_{\tau=0}&=D_{T^p}[\underline{B}_{pk},\underline{A}_{pk}]\\ &\quad\times D_{T^p}[\underline{B}_{p(k-1)},\underline{A}_{p(k-1)}]\\ &\quad\times D_{T^p}[\underline{B}_0,\underline{A}_0]\cdot(\underline{Y},\underline{X})_{,\tau}\ ,\end{aligned} \tag{5.34}$$

where $[\underline{B}_k,\underline{A}_k]\equiv T^k[\underline{B}_0,\underline{A}_0]$. Using the fact that $[\underline{B}_{pk},\underline{A}_{pk}]\to[\underline{B}_0^*,\underline{A}_0^*]$ for large k, the right-hand side is dominated by the largest eigenvalue, say λ_τ into which successive Jacobians applied to $(\underline{Y},\underline{X})_{,\tau}$ project. Generally, in the absence of symmetries or conservation laws which $(\underline{Y},\underline{X})_{,\tau}$ might satisfy, $\lambda_\tau=\lambda_1$, is the *largest* eigenvalue of $D_{T^p}[\underline{B}_0^*,\underline{A}_0^*]$.

From standard RG arguments, we find

$$\begin{aligned}T^{kp}[\underline{B}_0,\underline{A}_0](x;\tau)&=[\underline{B}_0^*,\underline{A}_0^*]+C_1\tau\lambda_{j_1}^k(\underline{Y}_0^*,\underline{X}_0^*)_{j_i}\\ &\quad+C_2\tau\lambda_{j_2}^k(\underline{Y}_0^*,\underline{X}_0^*)_{j_2}+\cdots\ ,\end{aligned} \tag{5.35}$$

where C_j are ("nonuniversal") constants. $(\underline{Y}_0^*,\underline{X}_0^*)_{j_1}$ is the eigenvector corresponding to the eigenvalue $\lambda_\tau\equiv\lambda_{j_1}$, and λ_{j_2} is the next-largest eigenvector into which $(\underline{Y},\underline{X})_{,\tau}$ has a projection. We will say that variations in the starting parameter τ *project* onto the eigenvector $(\underline{Y},\underline{X})_j$ *if* λ_j dominates Eq. (5.35).

By computing $(\underline{Y},\underline{X})_{,E}$, $(\underline{Y},\underline{X})_{,\epsilon}$ and $(\underline{Y},\underline{X})_{,\varphi_0}$ from the Taylor series of $[\underline{B}_0,\underline{A}_0]$ via Eq. (5.33), we calculate numerically the quantity

$$(\underline{Y},\underline{X})_{,\tau}^n\equiv\prod_{k=0}^{n}(D_{T^p}[\underline{B}_{kp},\underline{A}_{kp}])\cdot(\underline{Y},\underline{X})_{,\tau} \tag{5.36}$$

for n large. We can then find λ_τ by using the formula

$$\lim_{n\to\infty}[(\underline{Y},\underline{X})_{,\tau}^{n+1}-\lambda_\tau(\underline{Y},\underline{X})_{,\tau}^n]=0\ . \tag{5.37}$$

Numerically, for finite $n\gg 1$, the λ_τ obtained in this way will reproduce λ_j (calculated by diagonalizing the Jacobian at the fixed point) to two or three decimal places. We then identify this value of λ_j as the true λ_τ. This correspondence is also valuable as a consistency check on our numerics.

VI. RESULTS

A. Weak-coupling fixed point

Let us first consider the limit $\epsilon=0$. It is easy to verify that

$$[\underline{B}_0^*,\underline{A}_0^*]=\left[\begin{bmatrix}0&-1\\1&0\end{bmatrix},\begin{bmatrix}0&-1\\1&0\end{bmatrix}\right] \tag{6.1}$$

is an element of a six-cycle (a fixed point of T^6). By calculating the Jacobian in S_{full} at this trivial fixed point, we find, in order of decreasing magnitude, the relevant eigenvalues in S_{full} to be

$$\begin{aligned}&\lambda_1^{\text{full}}=\lambda_2^{\text{full}}=(\sigma)^{-6}\ ,\\ &\lambda_3^{\text{full}}=\lambda_4^{\text{full}}=\cdots=\lambda_8^{\text{full}}=1\ .\end{aligned} \tag{6.2}$$

As discussed in Sec. V, one relevant and one marginal eigenvector violate the condition (4.10), while three marginal eigenvectors violate (4.14). We are therefore left with one relevant eigenvalue $\lambda_1=\sigma^{-6}$ and the two redundant marginal eigenvalues $\lambda_3=\lambda_4=1$, expected from the arguments in Sec. V E.

When $\epsilon<2$, we find that $[\underline{B}_k,\underline{A}_k]$ approaches a pair of constant matrices which are similarity transforms of $[\underline{B}_0^*,\underline{A}_0^*]$ in (6.1). This similarity transformation $\underline{S}$, defined by

$$[\underline{B}_0^*,\underline{A}_0^*](x,\epsilon)=\underline{S}^{-1}(\epsilon,\varphi_0)\{[\underline{B}_0^*,\underline{A}_0^*](x,\epsilon=0)\}\underline{S}(\epsilon,\varphi_0) \tag{6.3}$$

is a function of ϵ and φ_0. Thus λ_ϵ and λ_{φ_0} are both marginal and redundant. We find, using the method of Sec. V F, that $\lambda_E=\lambda_1=(\sigma)^{-6}$, so that the only relevant parameter for $\epsilon<2$ is the energy E.

B. Results: Critical-coupling fixed point

When $\epsilon=2$ and $p=6$, we find that $[\underline{B}_0^*,\underline{A}_0^*](x)$ in Eq. (5.30) is a matrix of nontrivial, finite functions whenever

$\varphi_0=\frac{1}{4}+(n-\frac{1}{2})\sigma$. By calculating $D_{T^6}[\underline{B}^*_0,\underline{A}^*_0]$, we identify the following eigenvalues in the tangent space of S_{full}:

$$\lambda_1=196.296\ldots,\quad \lambda_2=\lambda_3=\lambda_4=\sigma^{-6},$$
$$\lambda_5=\cdots=\lambda_{11}=1.$$

All these are computed to six significant figures. According to the discussion in Sec. V D, one relevant and one marginal direction corresponds to the spurious eigenvalues that violate unimodularity of B and A. According to Sec. V E three marginal directions are redundant. It follows from the analysis in Sec. V C that all other marginal directions are spurious since they violate Eq. (5.9).

Using the formula (5.37), we identify the three remaining relevant eigenvalues as

$$\lambda_E=\lambda_1=196.296\ldots,$$
$$\lambda_{\varphi_0}=\lambda_3=\sigma^{-6},$$
$$\lambda_\epsilon=\lambda_4=\sigma^{-6}.$$

Since $\lambda_{\varphi 0}$ corresponds to a shift in the overall phase, it can be transformed away by translating the origin. Hence, only two relevant eigenvalues, λ_E and λ_ϵ, control the spectrum.

VII. RELATION BETWEEN RG RESULTS AND EMPIRICAL RESULTS

A. Subcycle structure of the renormalization group

The definitions of the critical indices β, $\tilde{\gamma}$, and ν in Sec. II are independent of the value of p. In particular $p=3$ gives a well-defined limit in Eqs. (2.8) and Eq. (2.9), and therefore $p=6$ also gives a well-defined limit with the same values for the scaling indices. We have, however, found that $p=6$ apparently governs the matrix cycle structure of the fixed point. The fact that $p=3$ is the minimal p which yields scaling comes from the hidden symmetry discussed briefly in Sec. II.

Let us write the elements of the six-cycle as $[\underline{B}^*_k,\underline{A}^*_k]$, where $T[\underline{B}^*_k,\underline{A}^*_k]=[\underline{B}^*_{k+1},\underline{A}^*_{k+1}]$ and $[\underline{B}^*_k,\underline{A}^*_k]=[\underline{B}^*_{k+6},\underline{A}^*_{k+6}]$. It turns out that we can find a *fixed* matrix $\underline{S}_0$ such that

$$\underline{S}_0^{-1}[\underline{B}^*_0(x),\underline{A}^*_0(x)]\underline{S}_0=(-1)[\underline{B}^*_3,\underline{A}^*_3](x). \tag{7.1}$$

This is a very powerful functional generalization of the similarity operation discussed in Sec. II since here $[\underline{B}^*_0,\underline{A}^*_0]$ are functions, whereas $\underline{S}$ is a 2×2 fixed matrix. Therefore, up to an overall sign change and similarity transformation, the underlying cycle structure of a basis-independent quantity is governed by the three-subcycle of the six-cycle. Since our definitions of the scaling indices are independent of the choice of the length of the cycle, we will continue working with the six-cycle of the matrices and invoke this hidden symmetry to explain why the observed minimal p in the scaling equations is 3.

The existence of the six-cycle comes about because we are working with $\kappa=\frac{1}{4}$ and the Fibonacci integers modulo 4 exhibit a six-cycle. However, the Fibonacci integers modulo 2 show the three-cycle. This is relevant as long as we ignore the sign of the wave function. An example of this point is illustrated by the fact that $\cos(2\pi\kappa F_k)$ shows a six-cycle, while $|\cos(2\pi\kappa F_k)|$ has a three-cycle.

B. The equality $\lambda_1=\gamma=\sigma^{-p\tilde{\gamma}}$ when $\epsilon\leq 2$

Let us define $\kappa_k[\underline{B}_k,\underline{A}_k]$ to be the Bloch index of the wave function $[\psi_n,\psi_{n-1}]$ on the *decimated* lattice (at the kth level of renormalization). The matrix which takes $[\psi_n,\psi_{n-1}]$ to the next lattice site, is therefore, $[\underline{B}_k,\underline{A}_k](x_n)$. Inside the block of sites which $\underline{A}_k$ or $\underline{B}_k$ represent, the phase of $[\psi_n,\psi_{n-1}]$ may rotate by integer multiples (say $N_{\underline{A}_k},N_{\underline{B}_k}$) of 2π. The integers $N_{\underline{A}_k},N_{\underline{B}_k}$ cannot be deduced directly from $[\underline{B}_k,\underline{A}_k]$, since these matrices only relate wave functions at the boundaries of the blocks.

Let us assume that the infinite lattice contains a proportion $f_{\underline{A}_k}$ of blocks of length $L_{\underline{A}_k}$ and a proportion $f_{\underline{B}_k}=1-f_{\underline{A}_k}$ of blocks of length $L_{\underline{B}_k}$. Therefore, the average rotation of the phase $\kappa_0\equiv\kappa_0([\underline{B}_0,\underline{A}_0])$ per unit length *in the original* lattice is given by

$$\kappa_0=\frac{\kappa_k[\underline{B}_k,\underline{A}_k]+f_{\underline{A}_k}N_{\underline{A}_k}+f_{\underline{B}_k}N_{\underline{B}_k}}{f_{\underline{A}_k}N_{\underline{A}_k}+f_{\underline{B}_k}N_{\underline{B}_k}}, \tag{7.2}$$

where the second and third terms represent the average integral contribution to the rotation per unit length *inside* each block. The first term is the nonintegral part of the rotation represented by $[\underline{B}_k,\underline{A}_k]$.

The integers $N_{\underline{A}_k}$ and $N_{\underline{B}_k}$ can be written in terms of κ_0 as

$$N_{\underline{A}_k}=[L_{\underline{A}_k}\kappa_0];\quad N_{\underline{B}_k}=[L_{\underline{B}_k}\kappa_0] \tag{7.3}$$

(see Ref. 33). The fraction $f_{\underline{A}_k}$ is equal to σ_k when $f(x)=x+\sigma$, and $f_{\underline{B}_k}$ is $1-\sigma_k$. Putting these formulas together, we find

$$\kappa_0=\frac{\kappa_k[\underline{B}_k,\underline{A}_k]+\sigma_k[\kappa_0 L_{\underline{A}_k}]+(1-\sigma_k)[\kappa_0 L_{\underline{B}_k}]}{\sigma_k L_{\underline{A}_k}+(1-\sigma_k)L_{\underline{B}_k}} \tag{7.4}$$

In the special case $\sigma=\sigma_G$ (F_k are the Fibonacci numbers) this reduces to

$$\kappa_0=\frac{\kappa_0[\underline{B}_k,\underline{A}_k+\sigma[\kappa_0 F_k]+\sigma^2[\kappa_0 F_{k-1}]}{\sigma F_k+\sigma^2 F_{k-1}} \tag{7.5}$$

Let us now consider a set of one-parameter variations as in Eq. (5.31). We can write

$$\kappa(\tau)\equiv\kappa_0\{[\underline{B}_0,\underline{A}_0]+\tau(\underline{Y},\underline{X})_\tau\}.$$

For sufficiently small τ (how small depends on κ) it follows that

$$\kappa_0(\tau)-\kappa_0(0)=\frac{\kappa_k(T^k[\underline{B}_k,\underline{A}_k](\tau))-\kappa_k(T^k[\underline{B}_k,\underline{A}_k](0))}{\sigma F_k+\sigma^2 F_{k-1}} \tag{7.6}$$

since if τ is sufficiently small

$$[\kappa_0(\tau)F_k]=[\kappa_0(0)F_k]. \tag{7.7}$$

To leading order we can thus substitute (5.35) in Eq. (7.5) to find

$$\lim_{\tau\to 0}\frac{\kappa_0(\tau)-\kappa_0(0)}{\kappa_0(\tau\lambda_\tau^{-1})-\kappa_0(0)}=\sigma^{-p}\ . \tag{7.8}$$

We have used the fact that increasing k by p and decreasing τ by a factor λ_τ^{-1} leaves the argument of κ_k unchanged in (7.6); however, the prefactor changes by a factor σ^{-p}. The limit $\tau\to 0$ is necessary because the above limits require k large, which in turn requires τ small for (6.2) to be valid. We note that no differentiability of $\kappa_0(\tau)$ is assumed (nor expected). Having identified γ with λ_E in this manner (see Refs. 27 and 47), we find [see Eq. (2.8)]

$$\tilde{\gamma}=1.8286\ldots \text{ for } \epsilon=2,$$

$$\tilde{\gamma}=1 \text{ for } \epsilon<2,$$

by using

$$\tilde{\gamma}=\frac{-1}{p}\frac{\ln\gamma}{\ln\sigma}\ .$$

These are precisely the exponents found in Sec. II.

C. Localization length exponent for $\epsilon>2$

Let us assume that the one-parameter variation τ projects into $(Y_0^*,X_0^*)_{,\epsilon}$. This occurs in Harper's model when we move along a trajectory in the parameter space (ϵ, φ_0, and E) by changing the *potential* strength only. However, in a more general model, all three parameters (ϵ, φ_0, and E) must be controlled to stay on such a trajectory. In the RG description, the localization transition is a multicritical point, and we are now investigating changes *along* the critical lines which connect to the localization point.

Our numerical investigations and the work of Aubry and André[5] indicate that there is a localization length denoted by $l(\tau)$, which defines the envelope of Fig. 6. Using the equality (5.35) and the fact that the l scales like a length, we find that

$$l(\tau)=\sigma^{-p}l(\tau\lambda_\epsilon)\ . \tag{7.9}$$

Therefore, by associating $\tau=\epsilon-2$, we find that $\nu=1$ in Eq. (2.16). (This result has also been obtained in Ref. 5 in a more general framework. See also the Note added.)

Actually, Eq. (5.35) gives even stronger results. As long as the *undecimated* sites determine the visible structure (i.e., peaks) of ψ_n, then at these sites n_k, the values of the wave function $\{|\psi_{n_k}|\}$ at a particular value of $\tau=\tau_0$ must be equal to the values of the wave function $\{|\psi_{n_{p+k}}|\}$ on the *decimated* sites n_{k+p} lattice at a value $\tau=\tau_0\lambda_\epsilon$. Then since the average decimated lattice constant is simultaneously increased by a factor σ^p, it follows that

$$\psi_n(\tau)=\psi_{[\sigma^p n+1/2]}(\tau\sigma^{-p})\ . \tag{7.10}$$

This confirms Eq. (2.14) when $\tau=0$ and explains the accurate scaling of the features in Figs. 6(a) and 6(b).

D. Summary

We have recovered most of the scaling results in Sec. II with the renormalization group and have shown in what sense $\tilde{\gamma}$ and ν are universal. We have *not* been able to obtain β from the fixed-point analysis, although preliminary work indicates that ζ can be derived from the fixed point. We believe that both of these quantities *are* universal, but in the absence of a direct derivation of this fact from the fixed-point analysis, this conclusion is still speculative. We are still puzzled about the role of the special phases $\varphi_0=\frac{1}{4}-\frac{1}{2}\sigma_G+n\sigma$ which are required to get a fixed point out of our recursion relations at $\kappa=\frac{1}{4}$. We do not know why these phases are singled out.

We have worked out the simplest application of the general renormalization-group theory discussed in Secs. II–V, and the application to various other points in the spectrum appears to be fairly straightforward. In contrast to these direct extensions of the present work, there are also more subtle questions which we believe could be in principle answered. What is the behavior of the states and spectrum near gap edges? This problem is addressed in Ref. 27 using the model of Refs. 30 and 27, but with an analytic potential the problem is more complicated. What happens for arbitrary choices of σ and κ when the renormalization group does not flow to a fixed point?

It is also worth noting that in our RG it is necessary to work at a fixed value of the Bloch index. The disadvantage of this approach is that it is not clear how it can be used to calculate global properties (e.g., the measure of the gaps in the spectrum.[30,47]) However, the strength of our method lies in the *universality* of its predictions. For a given value of the Bloch index, a large class of quasiperiodic potentials made up of harmonics of Harper's potential should share scaling properties characterized by the exponents $\tilde{\gamma}$ and ν. It is not clear to us whether such a strong statement can be made for the global properties of the spectrum of a general quasiperiodic potential.[47] For such a potential, different states may localize at different values of the potential strength (unlike what happens in Harper's model), so it seems improbable that the measure of the gaps would show any universal scaling behavior.

Note added. After this paper had been written, we came across a recent paper by Suslov[49] in which the result (2.16) is obtained for Harper's potential (1.2) by using a renormalization group. Also, Kohmoto and Oono[50] have recently carried out a fixed-point analysis of the RG equations of Ref. 30. J. Wilkinson has also carried out an approximate RG on Harper's model.[51]

ACKNOWLEDGMENTS

We would like to thank Anders Carlsson, David Rand, Hans Schellnhuber, and Eric Siggia for useful discussions. We are especially grateful to Eric Siggia for a critical reading of this paper. This research was supported by National Science Foundation Grants Nos. DMR-80-20429 and PHY-77-27084.

APPENDIX: CONTINUED FRACTIONS

Let ρ be an arbitrary rational or irrational number. We can write ρ in a continued-fraction expansion as

$$\cfrac{1}{n_1+\cfrac{1}{n_2+\cfrac{1}{n_3+\cdots}}} \equiv [n_1,n_2,\ldots,n_k,\ldots] , \quad \text{(A1)}$$

where n_k are integers obeying $1 \le n_k \le \infty$. The integers n_k are generated by the recursion formula

$$\rho_0=\rho, \quad n_k=[1/\rho_k 1], \quad \rho_{k+1}=\frac{1}{\rho_k}-n_k , \quad \text{(A2)}$$

where [ ] denotes the integer part thereof. We say an irrational number has a periodic tail in the continued fraction if $n_k=n_{k+m}$ for every k larger than some finite N. This is true if and only if ρ solves a quadratic equation with integer coefficients.

The kth rational approximant to ρ is given by the ratio p_k/q_k defined by

$$p_k/q_k=[n_1,\ldots,n_k,\infty] . \quad \text{(A3)}$$

These are the "best" rational approximants to ρ in the sense that there is no rational number with a smaller denominator that approximates the irrational number better.

After a bit of algebra, it is possible to show that the integers p_k and q_k obey the recursion formulas

$$p_{k+1}=n_k p_k+q_{k-1}, \quad q_{k+1}=n_k q_k+q_{k-1} , \quad \text{(A4)}$$

where $p_{-1}=0$, $q_{-1}=1$; $p_0=1$; $q_0=0$. The quadratic irrational $\rho=\sigma_G=\frac{1}{2}(\sqrt{5}-1)$, which is the reciprocal of the golden mean, has $n_k=1$ for all k. For this case $p_k=F_k$ and $q_k=F_{k-1}$, where the Fibonacci integers F_k satisfy $F_{k+1}=F_k+F_{k-1}$ with $F_{-1}=1$ and $F_0=1$.

A Liouville number α is an irrational number for which $|\alpha-p_n/q_n| < 1/(q_n)^n$, where p_n and $q_n \to \infty$ as $n \to \infty$. Only for the subclass of Liouville numbers for which there is a constant c such that $|\alpha-p_n/q_n| \le c/n^{(q_n)}$, has it been shown that quasiperiodic Schrödinger operators with $\sigma=\alpha$ in Eq. (1.2) have no localized states.[4,32]

An irrational number ϑ is called a Roth number if, for every $\epsilon>0$, there is a constant $c_\epsilon>0$ such that $|\vartheta-p/q| > c_\epsilon/q^{2+\epsilon}$ for all positive integers p and q. The class of Roth numbers is of full Lebesque measure.[32]

---

*Present address: Department of Physics, University of Pennsylvania, Philadelphia, PA 19104.

[1]The potential $V(x)$ in the Schrödinger equation

$$\left[-\frac{d^2}{dx^2}+V(x)\right]\psi(x)=E\psi(x)$$

is quasiperiodic on $(-\infty,\infty)$ if it admits the representation

$$V(x)=\sum_{(m_1,\ldots,m_j)} a_{m_1,\ldots,m_j}\exp\left[2\pi i\sum_{k=1}^{j} m_k\omega_k x\right]$$

where $m_k \in Z$ (the integers) and $\omega_k \in R$ (real numbers) for all $1 \le k \le j$, $j \ge 2$, $a_{m_1,\ldots,m_j}$ are such that $V(x)$ is well defined, and the ω_k's are incommensurate, i.e., for r_k rational,

$$\sum_{k=1}^{j} r_k\omega_k=0$$

if and only if $\tau_k=0$ for all k. A simple example (with $j=2$) is $V(x)=\cos(2\pi x)+\cos(2\pi\omega x)$, with ω irrational. In the discrete case

$$-(\psi_{n+1}-2\psi_n+\psi_{n-1})+V_n\psi_n=E\psi_n ,$$

V_n should be quasiperiodic on the integers. An example is $V_n=\cos(2\pi\omega n)$ with ω irrational. [Note that the lattice provides one of the frequencies, so a continuum limit of the model with $V_n=\cos(2\pi\omega n)$ does not yield a Schrödinger equation with a quasiperiodic potential.] Some authors prefer to use the expression almost periodic (Refs. 4, 32, and 36) instead of quasiperiodic.

[2]For a general reference on quasiperiodic functions, see A. S. Besicovitch, *Almost Periodic Functions* (Cambridge, London, 1932).

[3]Some properties of quasiperiodic functions that are particularly relevant for our purposes are contained in M. V. Romerio, J. Math. Phys. 12, 552 (1971).

[4]For a recent review of the rigorous mathematical literature on quasiperiodic Schrödinger operators, see B. Simon, Adv. Appl. Math. 3, 463 (1982).

[5]S. Aubry and G. Andre, in *Proceedings of the Israel Physical Society,* edited by C. G. Kuper (Hilger, Bristol, 1979), Vol. 3, p. 133.

[6]E. Batalla, F. S. Razavi, and W. R. Datars, Phys. Rev. B 25, 2109 (1982), and references cited therein. The parameter δ that governs the incommensurability can be varied by the application of pressure.

[7]P. G. Harper, Proc. Phys. Soc. London Sect. A 68, 874 (1955); G. E. Zilberman, Zh. Eksp. Teor. Fiz. 30, 1092 (1956) [Sov. Phys.—JETP 3, 835 (1956)].

[8]E. I. Blount, Phys. Rev. 126, 1636 (1962).

[9]E. Brown in *Solid State Physics,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1968), Vol. 22 , p. 3313, and references cited therein.

[10]M. Ya Azbel', Zh. Eksp. Teor. Fiz. 46, 939 (1964) [Sov. Phys.—JETP 19, 634 (1964)]; Phys. Rev. Lett. 43, 1954 (1979).

[11]A. Rauh, G. H. Wannier, and G. M. Obermair, Phys. Status Solidi B 63, 215 (1974).

[12]G. M. Obermair and H. J. Schellnhuber, Phys. Rev. B 23, 5185 (1981); H. J. Schellnhuber, G. M. Obermair, and A. Rauh, *ibid.* 23, 5191 (1981), and refernces cited therein.

[13]D. R. Hofstadter, Phys. Rev. B 14, 2239 (1976).

[14]G. H. Wannier, Phys. Status Solidi B 88, 757 (1978).

[15]C. de Lange and T. Janssen [J. Phys. C 14, 5269 (1981)] study related problems in the lattice dynamics of incommensurate crystals.

[16]C. M. Soukoulis and E. N. Economou, Phys. Rev. Lett. 48, 1043 (1982).

[17]R. Rammal, T. C. Lubensky, and G. Toulouse, Phys. Rev. B 27, 2820 (1983); S. Alexander, *ibid.* 27, 1541 (1983).

[18]See. M. Reed and B. Simon, *Methods of Modern Mathematical Physics, I, Functional Analysis* (Academic, New York, 1975), p. 231, for definitions of pure-point, absolutely continuous, and singular continuous spectra.

[19]A. N. Kolmogorov, Dokl. Akad. Nauk USSR 98, 527 (1954); V. I. Arnol'd, Russ. Math. Surveys 18:5 (1963); *ibid.* 18:6 (1963); J. Moser, *Stable and Random Motions in Dynamical Systems* (Princeton University Press, Princeton, NJ, 1973).

[20]D. J. Thouless and Q. Niu, J. Phys. A 16, 1911 (1983).

[21]D. J. Thouless, M. Kohmoto, P. Nightingale, M. den Nijs, Phys. Rev. Lett. 49, 405 (1982).

[22]E. I. Dinaburg and Ya. G. Sinai, Func. Anal. Appl. 9, 279 (1976). For an extension of these results see Refs. 35 and 32.

[23]S. J. Shenker and L. P. Kadanoff, J. Stat. Phys. 27, 631 (1982); L. P. Kadanoff, Phys. Rev. Lett. 47, 1641 (1981).

[24]S. J. Shenker, Physica D 5, 405 (1982); M. J. Feigenbaum, L. P. Kadanoff, and S. J. Shenker, *ibid.* 5, 370 (1982).

[25](a) D. Rand, S. Ostlund, J. Sethna, and E. D. Siggia, Phys. Rev. Lett. 49, 132 (1982); (b) S. Ostlund, D. Rand, J. Sethna, and E. D. Siggia, Physica D 8, 303 (1983).

[26]N. S. Manton and M. Nauenberg, Commun. Math. Phys. 89, 555 (1983); M. Widom, D. Bensimon, L. P. Kadanoff, and S. J. Shenker, J. Stat. Phys. (in press).

[27]S. Ostlund, R. Pandit, D. Rand, H. Schellnhuber, and E. D. Siggia, Phys. Rev. Lett. 50, 1873 (1983).

[28]K. G. Wilson and J. Kogut, Phys. Rept. 12C, 75 (1974); M. J. Feigenbaum, J. Stat. Phys. 19, 25 (1978); 21, 669 (1979).

[29]J. B. Sokoloff and J. V. Jose, Phys. Rev. Lett. 49, 3341 (1982).

[30]M. Kohmoto, L. P. Kadanoff, and C. Tang, Phys. Rev. Lett. 50, 1870 (1983).

[31]P. Collet, J. P. Eckmann, and O. E. Lanford, Commun. Math. Phys. 76, 211 (1980).

[32]J. Belissard, R. Lima, and D. Testard, Commun. Math. Phys. 88, 207 (1983); J. Belissard, D. Bessis, and P. Moussa, Phys. Rev. Lett. 49, 701 (1982).

[33]The notation $[x]$ means integer truncation of x and $\langle x \rangle$ means the fractional portion. Thus $x=[x]+\langle x \rangle$.

[34]F. Seitz, *Modern Theory of Solids* (McGraw-Hill, New York, 1940).

[35]H. Rüssmann, Ann. N.Y. Acad. Sci. 357, 90 (1980).

[36]J. Avron and B. Simon, Commun. Math. Phys. 82, 101 (1981); Math. J. 50, 369 (1983).

[37]R. Johnson and J. Moser, Commun. Math. Phys. 84, 403 (1982); F. Delyon and B. Souillard, *ibid.* 89, 415 (1983).

[38]D. R. Grempel, S. Fishman, and R. E. Prange, Phys. Rev. Lett. 49, 833 (1982); R. E. Prange, D. R. Grempel, and S. Fishman (unpublished).

[39]What we call the Bloch index is equivalent to the rotation number of Johnson and Moser (Ref. 37) and is half the integrated density of states used by Avron and Simon (Refs. 4 and 36).

[40]This may be proved rigorously. (D.J. Thouless, private communication and unpublished). See also Ref. 47.

[41]K. S. Dy and T. C. Ma, J. Phys. C 15, 6971 (1982). See also Refs. 5 and 27.

[42]In Ref. 27 two exponents $\delta(\kappa)$ and $\gamma(\kappa)$ were defined to characterize the scaling properties of the spectrum. Their definitions relied on the approximation of a quasiperiodic system by periodic ones with progressive larger unit cells. Both these exponents are trivially related to the exponent $\tilde{\gamma}(\kappa)$ defined here by $\delta=\gamma\sigma^{-p\tilde{\gamma}}$.

[43]Of course, $\underline{M}_q(x)$ also depends on E and ϵ. If we want to consider the dependence on any other variable, say ϵ and E, we shall simply list these in the argument, i.e., $\underline{M}_q(x;\epsilon,E) \equiv \underline{M}_q(x) \equiv \underline{M}_q$. Thus, the arguments will vary according to context.

[44]We use the following notation convention throughout this paper. Unless an argument to a function is explicitly given, composition with the object to the right is implied. Thus $fg(x) \equiv f(g(x))$. Furthermore, $f^p(x)$ is $fff\cdots f(x)$ p times. We also defer the writing of an argument by deferring the parentheses. Thus

$$[fg-gf](x/\alpha) \equiv fg(x/\alpha)-gf(x/\alpha) .$$

[45]We use the convention that $\underset{\sim}{0}(x)=\underset{\sim}{0}$ and $\underset{\sim}{1}(x)=x$ for all x. In general, we use bold constants to denote the functions $\underset{\sim}{\alpha}(x) \equiv \alpha x$.

[46]M. J. Feigenbaum and B. Hasslacher, Phys. Rev. Lett. 49, 605 (1982).

[47]M. Kohmoto, [Phys. Rev. Lett. 51, 1198 (1983)] defines an exponent α similar to our exponent $\gamma=\sigma^{-p\tilde{\gamma}}$, finds $\alpha=14.01\ldots$ ($\alpha=\sigma_G^{-s}=4.2\ldots$) for Harper's equation in agreement with our value for γ for $\kappa=\frac{1}{4}$ and $\epsilon=2$ ($\epsilon<2$). His definition of α does not explicitly account for the dependence on the Bloch index κ and cannot be generalized to other models in which different states get localized at different values of ϵ.

[48]If $\lambda_1=\lambda_2$, then $[\underline{b},\underline{a}]_c=0$. However, in this case, $\underline{S}$ can be chosen to diagonalize $\underline{b}$, i.e., $b_{12}=b_{21}=0$. Thus, the redundant space is again two dimensional.

[49]I. M. Suslov, Zh. Eksp. Teor. Fiz. 83, 1079 (1982) [Sov. Phys.—JETP 56, 612 (1982)].

[50]M. Kohmoto and Oono (unpublished).

[51]J. Wilkinson (unpublished).