Cantor spectra and scaling of gap widths in deterministic aperiodic systems

J. M. Luck
Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette CEDEX, France;
Department of Theoretical Physics, Oxford University, 1 Keble Road, Oxford OX1 3NP, United Kingdom;
Department of Physics, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom
(Received 28 September 1988)

The relationship between geometry and physical properties of aperiodic structures is investigated by considering the example of the tight-binding Schrödinger equation in one dimension, where the site potentials are given by an arbitrary deterministic aperiodic sequence. In a perturbative analysis of the integrated density of states, the gaps in the energy spectrum can be "labeled" by the singularities of the Fourier transform of the sequence of potentials. This approach confirms known properties of quasiperiodic and almost-periodic systems, and suggests an extension of them to more general sequences, such as those with a singular continuous Fourier transform. There is strong evidence that the spectrum is a Cantor set with zero measure for a much larger class of models than quasiperiodic ones. The dependence of the widths of various gaps on the potential strength is also determined: several different kinds of behavior are obtained, such as a power law with a nontrivial exponent, or an essential singularity. These general results are compared with those of various other approaches for four self-similar sequences generated by substitution, namely the Thue-Morse sequence, the period-doubling sequence, a "circle sequence," and the Rudin-Shapiro sequence.

I. INTRODUCTION

The discovery of quasicrystals[1] has revived the interest in quasiperiodic systems, which are the first class of structures on the way from periodic (crystalline) to random (amorphous, or glassy) matter.

More recent works have been devoted to structures beyond quasiperiodicity, either deterministic or partially random,[2–4] which seem relevant to the physics of quasicrystals. Focusing our attention to one dimension, we mention the investigation of structures with an unbounded density fluctuation,[5–7] or a singular continuous geometrical Fourier transform.[8,9] Some physical properties of another kind of model with a singular continuous Fourier spectrum, namely the Thue-Morse sequence, have also been studied.[10–13] Furthermore, any aperiodic sequence may now have a physical realization, in the form of epitaxially grown semiconductor superlattices.[14–16]

The general goal of this paper is a better understanding of the relationship between basic geometrical characteristics of one-dimensional aperiodic structures, such as those coded in their Fourier transform, and their physical properties. We restrict ourselves to linear propagation problems, the prototype of which is the discrete Schrödinger equation, describing electrons in the tight-binding approximation

$$-\psi_{n+1}-\psi_{n-1}+V_n\psi_n=E\psi_n \ . \tag{1.1}$$

The (diagonal) site potentials V_n form an arbitrary *deterministic aperiodic sequence,* with sufficient homogeneity properties. More precisely, we assume adequate clustering properties of the correlations which enter the expansion of Sec. II B. Such a sequence will be called *homogeneous.* The results of this paper can be extended, *mutatis mutandis,* to the whole class of linear problems which can be formulated as discrete Laplace equations, where the aperiodicity enters either in diagonal (site) variables, or in nondiagonal (bond) variables. We mention, in particular, the study of harmonic vibrations (phonons), spin-wave propagation, and diffusion, as well as some quantum spin chains which are equivalent to free fermionic fields.

We will mostly consider binary sequences, where the potentials V_n take only two different values. By shifting the zero of energies, we can choose, without loss of generality, these two values of the potentials to be opposite, namely

$$V_n=V\varepsilon_n \ , \tag{1.2}$$

where V, either positive or negative, it is the potential strength, and $\varepsilon_n=\pm1$. The questions of most physical interest on Eq. (1.1) concern the energy spectrum, and the nature of the eigenstates (in the following, we will speak of the energy, or electronic, spectrum, in order to distinguish it from the Fourier spectrum of the underlying structure, i.e., of the sequence of potentials).

We first point out a simple symmetry property of the problem. The Schrödinger equation (1.1) is invariant under the simultaneous transformation

$$\psi_n\mapsto(-1)^n\psi_n, \quad V_n\mapsto-V_n, \quad E\mapsto-E \ . \tag{1.3}$$

This property shows that the spectra corresponding to opposite values of the potential strength V are mirror images of each other with respect to $E=0$. We assume therefore from now on, again without loss of generality, that V is positive.

A considerable amount of work has been devoted to the study of Eq. (1.1) for both random and quasiperiodic sequences V_n. If the potential is disordered, i.e., the V_n are independent random variables, the main feature of the system is Anderson localization: all eigenstates are exponentially localized. The energy spectrum itself is al-

ways a regular object, with at most a finite number of bands (for a bounded potential). Indeed, it is known rigorously[17] to be the convolution of the interval $[-2,+2]$ by the support of the distribution of the V_n. In the case of a binary potential distribution, the spectrum has one or two bands, but the density of states inside that spectrum may be very singular. It is then advantageous to use the *integrated density of states* (IDOS) at energy E, $H(E)$, defined as the fraction of eigenvalues less than E. This quantity is always well defined,[17] and is convenient to describe different types of spectral singularities.[18,19]

Among quasiperiodic binary sequences, the Fibonacci sequence has been the most extensively studied, since the pioneering work of two groups.[20–24] The existence of a trace map has played a crucial role in the subject. The energy spectrum is known to be a Cantor set of zero (Lebesgue) measure: there is an infinity of gaps, and the total "bandwidth" vanishes. Moreover, the eigenstates are neither extended nor localized, but exhibit a rather intricate kind of intermediate behavior. Statistical properties of the energy spectrum and of the eigenstates have been described using the formalism of multifractals,[25–28] in analogy with other complex objects, such as strange attractors.[29]

More recently, several authors have studied the electronic spectra associated with other deterministic sequences. References 30–32 deal with quasiperiodic binary sequences, generated by substitutions associated with some classes of quadratic numbers, just as the Fibonacci sequence is associated with the golden mean τ $[\equiv(\sqrt{5}+1)/2]$. As mentioned above, Refs. 10–13 deal with the Thue-Morse sequence, long known in the mathematical literature, but not to physicists. This sequence has a singular continuous Fourier transform. We will come back to this example in detail in Sec. IV.

The striking common feature of these models is that their spectrum seems to be always a *Cantor set of zero measure.* This characteristic, discovered first in the study of the Fibonacci chain, turns out to be universal within a very large class of aperiodic deterministic sequences. This is in sharp contrast with the fact that the spectrum associated with a random sequence consists of one or two bands for a binary distribution, as described above.

The present work sheds some new light on the universality of zero-measure Cantor sets among energy spectra associated with deterministic aperiodic sequences, well beyond the quasiperiodic class. We consider, in particular, the location of gaps, and the dependence of gap widths on the potential strength.

The setup of this paper is as follows. A first part (Sec. II) presents a general perturbative analysis of the Lyapunov exponent and the IDOS, for a small potential strength V. The location and widths of the gaps are related, in a general sense, to the singularities of the Fourier intensity. For a quasiperiodic (or almost-periodic) sequence, each delta peak of the Fourier transform, i.e., each point of the frequency module, corresponds to a gap in the energy spectrum; the width of this gap behaves linearly in V. This result has been derived in a rigorous nonperturbative way for a smooth enough quasiperiodic (or almost-periodic) potential: this is the "*gap labeling theorem.*"[33,34] It is expected to hold even if the potential is not a smooth quasiperiodic perturbation, as in the example of the Fibonacci sequence. The concept of gap labeling is generalized to other types of sequences, such as those having a singular continuous Fourier transform. The singularities of this Fourier transform, provided they are strong enough, also generate gaps, with widths that scale as V^{β}, with a nontrivial exponent $\beta>1$.

A second part (Secs. III–VI) is devoted to a detailed analysis of the energy spectra of four self-similar sequences generated by substitution, namely the Thue-Morse sequence (Sec. III), already considered in Refs. 10–13; the period-doubling sequence (Sec. IV), originating in the study of dynamical systems;[35,36] the "circle sequence" (Sec. V), which shows up in the study of some structures beyond quasiperiodicity;[7–9] and the Rudin-Shapiro sequence[37] (Sec. VI). For each of these examples, the general results of the first part are compared with those of other approaches, such as the exact determination of the spectral gaps in the large-V limit, and numerical data. In the first two cases, the sequences are generated by a binary substitution, and hence there exists a polynomial trace map.[38] Section VII presents a short discussion.

II. GENERAL PERTURBATIVE RESULTS

A. Fourier transforms

In this section, we recall some useful concepts and notations concerning the Fourier spectrum of deterministic sequences. The Fourier transform of an arbitrary homogeneous sequence ε_n $(n\geq 1)$ is defined as follows. For $0\leq q\leq 2\pi$, we consider the partial Fourier amplitudes

$$G_N(q)=\sum_{1\leq n\leq N}\varepsilon_n e^{inq} \tag{2.1}$$

and the corresponding intensities

$$S_N(q)=\frac{1}{N}|G_N(q)|^2 . \tag{2.2}$$

In the limit of an infinite system, the only meaningful object to consider, in general, is the *Fourier intensity measure* of the sequence, defined as

$$d\mu(q)=\lim_{N\to\infty}\left[S_N(q)\frac{dq}{2\pi}\right] . \tag{2.3}$$

This measure is the Fourier transform of the two-point correlation function of the sequence

$$\begin{aligned} S_a &= \lim_{N\to\infty}\left[\frac{1}{N}\sum_{1\leq n\leq N}\varepsilon_n\varepsilon_{n+a}\right] \\ &= \langle\varepsilon_n\varepsilon_{n+a}\rangle=\int d\mu(q)e^{iqa} . \end{aligned} \tag{2.4}$$

Throughout this paper, the angular brackets will denote such Cesàro averages (defined by averaging over N sites, and then letting N go to infinity).

Just as any positive measure, $d\mu$ can have three different components, namely a discrete one, an absolutely continuous one, and a singular continuous one. This decomposition is made clearer by considering the distribution function, or "integrated density," $\mu(q)$, defined by

$$\mu(q)=\lim_{N\to\infty}\left[\int_0^q S_N(q')\frac{dq'}{2\pi}\right] . \tag{2.5}$$

(a) The *discrete (or atomic)* part is coded in the discontinuities of $\mu(q)$: any delta peak of the form $C(q_0)\delta(q-q_0)$ in the Fourier amplitude $G(q)$ corresponds to a discontinuity at $q=q_0$ of strength $|C(q_0)|^2$ in the function $\mu(q)$. The partial sums defined in Eq. (2.1) then behave as

$$G_N(q)\sim C(q_0)N . \tag{2.6}$$

(b) The *absolutely continuous* part $S(q)$ is coded in the derivative of μ (at points where it exists): $S(q)=2\pi\, d\mu/dq$.

(c) The *singular continuous* part is the most difficult to capture; in the mathematical literature, it is essentially defined as "what remains" when the previous two components are subtracted.

We will show in the following that the spectral gaps of Eq. (1.1) are related to the singularities of the Fourier distribution function $\mu(q)$. For quasiperiodic and almost-periodic sequences, the only singularities are discontinuities, which occur at a dense set of values of q, usually referred to as the "frequency module." For singular continuous sequences, there is usually a dense set of values q_0 of q around which $\mu(q)$ has power-law singularities

$$\mu(q)-\mu(q_0)\sim\pm A_\pm|q-q_0|^\alpha \quad \text{as } q\to q_0^\pm , \tag{2.7}$$

with a q_0-dependent exponent α. Reference 9 presents a detailed study of such singularities in a geometrical Fourier transform. Reference 13 contains a similar analysis of the Fourier spectrum of the Thue-Morse sequence; this study will be summarized in Sec. IV.

B. Perturbative expansion of the Lyapunov exponent and the IDOS

This section is devoted to a perturbative expansion of the Lyapunov exponent and the integrated density of states of the tight-binding model (1.1) in powers of the site potentials V_n for an arbitrary *homogeneous* sequence of potentials. This concept will become clearer at the end of this section. Without any loss of generality, the sequence V_n can be assumed to have zero Cesàro average: $\langle V_n\rangle=0$.

The following approach is a generalization of a method extensively used in the study of disordered systems in one dimension.[39,40] Starting from the tight-binding equation (1.1), we define the ratios (Riccati variables)

$$R_n=\frac{\psi_{n+1}}{\psi_n} , \tag{2.8}$$

which obey the recursion relation

$$R_n=V_n-E-\frac{1}{R_{n-1}} . \tag{2.9}$$

We consider a complex energy E, and define the complex Lyapunov exponent $\Omega(E)$ by

$$\Omega(E)=\lim_{N\to\infty}\left[\frac{1}{N}\sum_{1\le n\le N}\ln R_n\right]=\langle\ln R_n\rangle . \tag{2.10}$$

We assume that this definition yields a well-behaved quantity for any homogeneous sequence of site potentials. We use the principal determination of the logarithm, with a cut along the negative real axis. Therefore, when E goes to a real value, the usual Lyapunov exponent $\gamma(E)$ and the integrated density of states (IDOS) $H(E)$ are recovered as

$$\lim_{\varepsilon\to0}\Omega(E\pm i\varepsilon)=\gamma(E)\pm i\pi H(E) . \tag{2.11}$$

We illustrate now in some detail how this formalism works in the (trivial) case where there are no site potentials ($V_n=0$), since this will be the starting point of our perturbative expansion. The transform (2.9) acting on the R_n becomes an n-independent Möbius map: $R\to\phi(R)=-E-1/R$. We perform the change of variable,

$$E=-2\cosh t , \tag{2.12}$$

which maps the complement of the spectrum $\mathcal{S}=[-2,+2]$, onto the semi-infinite strip ($\mathrm{Re}\,t>0$, $-\pi<\mathrm{Im}\,t<\pi$). The fixed points of the map ϕ are e^t, attractive, and e^{-t}, repulsive. Hence the R_n converge exponentially towards e^t, and we have

$$\Omega^{(0)}=t , \tag{2.13}$$

where the superscript (0) denotes the absence of site potentials. For E inside the spectrum $\mathcal{S}$, we have

$$t(E\pm i\varepsilon)=\pm iQ, \quad 0\le Q\le\pi \tag{2.14}$$

where Q is the wave vector of the extended states. Hence Eq. (2.11) yields the dispersion relation $E=-2\cos Q$, and Eq. (2.10) shows that the Lyapunov exponent γ vanishes, whereas the IDOS reads $H(E)=Q/\pi$, as expected.

The systematic expansion of the complex Lyapunov exponent in powers of the potentials proceeds as follows. It is advantageous to define new variables Y_n by

$$Y_n=\frac{1-e^{-t}R_n}{1-e^tR_n}, \quad R_n=e^t\frac{1-Y_n}{1-e^{2t}Y_n} . \tag{2.15}$$

The recursion relation (2.9) then becomes

$$Y_n=e^{-2t}\frac{Y_{n-1}+g_n(1-Y_{n-1})}{1+g_n(1-Y_{n-1})} \tag{2.16}$$

with the notation

$$g_n=\frac{V_n}{2\sinh t} \tag{2.17}$$

and we obtain the following expression for Ω:

$$\Omega=t+\left\langle\ln\frac{1-Y_n}{1-e^{2t}Y_n}\right\rangle . \tag{2.18}$$

This equation can be advantageously rearranged by replacing Y_n by Y_{n-1} in the numerator, and using the recursion relation (2.16). This yields

$$\Omega = t + \langle \ln[1 + g_n(1 - Y_{n-1})] \rangle \ . \tag{2.19}$$

The interesting point of this expression is that it contains an explicit factor of g_n. Hence it will be sufficient to expand the quantities Y_n in powers of g_n to the $(k-1)$th order to obtain the expansion of Ω to the kth order. Indeed, if $\Omega^{(k)}$ and $Y_n^{(k)}$ denote the terms of order k, i.e., which contain k potentials V_m, in the expansions of Ω and Y_n, we have

$$\begin{aligned} &\Omega^{(0)} = t, \quad \Omega^{(1)} = \langle g_n \rangle = 0 \ , \\ &\Omega^{(2)} = -\langle g_n Y_{n-1}^{(1)} \rangle - \tfrac{1}{2}\langle g_n^2 \rangle \ , \\ &\Omega^{(3)} = -\langle g_n Y_{n-1}^{(2)} \rangle + \langle g_n^2 Y_{n-1}^{(1)} \rangle + \tfrac{1}{3}\langle g_n^3 \rangle \ . \end{aligned} \tag{2.20}$$

The expression of $\Omega^{(0)}$ is just our previous result (2.13). The term $\Omega^{(1)}$ vanishes, as a consequence of the assumption that the potential has zero Cesàro average. We will compute the next two terms explicitly, although only $\Omega^{(2)}$ will be used in the following.

The building blocks of Eq. (2.20) can be calculated by expanding Eq. (2.16), and solving the obtained equations in a recursive way.

1. Calculation of $Y_n^{(1)}$ and $\Omega^{(2)}$

Equation (2.16) yields the following recursion relation for $Y_n^{(1)}$:

$$e^{2t} Y_n^{(1)} = Y_{n-1}^{(1)} + g_n \ . \tag{2.21}$$

The solution of this equation reads

$$Y_n^{(1)} = \sum_{a\ (\geq 1)} e^{-2ta} g_{n-a+1} \ . \tag{2.22}$$

Equation (2.20) now leads to

$$\Omega^{(2)} = -\sum_{a\ (\geq 1)} e^{-2ta} \langle g_n g_{n-a} \rangle - \tfrac{1}{2}\langle g_n^2 \rangle \ . \tag{2.23}$$

2. Calculation of $Y_n^{(2)}$ and $\Omega^{(3)}$

We deduce from Eq. (2.16) the following recursion formula for $Y_n^{(2)}$:

$$e^{2t} Y_n^{(2)} = Y_{n-1}^{(2)} + S_n^{(2)} \quad \text{with } S_n^{(2)} = -2 g_n Y_{n-1}^{(1)} - g_n^2 \ . \tag{2.24}$$

In analogy with Eq. (2.22), the solution of this equation reads

$$\begin{aligned} Y_n^{(2)} &= \sum_{a\ (\geq 1)} e^{-2ta} S_{n-a+1}^{(2)} \\ &= -2 \sum_{a,b\ (\geq 1)} e^{-2t(a+b)} g_{n-a+1} g_{n-a-b+1} \\ &\quad - \sum_{a\ (\geq 1)} e^{-2ta} g_{n-a+1}^2 \ . \end{aligned} \tag{2.25}$$

Equation (2.20) finally leads to

$$\begin{aligned} \Omega^{(3)} = 2 &\sum_{a,b\ (\geq 1)} e^{-2t(a+b)} \langle g_n g_{n-a} g_{n-a-b} \rangle \\ &+ \sum_{a\ (\geq 1)} e^{-2ta} (\langle g_n^2 g_{n-a} \rangle + \langle g_n g_{n-a}^2 \rangle) \\ &+ \tfrac{1}{3}\langle g_n^3 \rangle \ . \end{aligned} \tag{2.26}$$

The general structure of the perturbative expansion shows up clearly from $\Omega^{(2)}$ and $\Omega^{(3)}$. The generic term $\Omega^{(k)}$ involves k-point correlations of the reduced potentials g_n. These correlations occur in sums over up to k positive integers $a, b, \ldots$ with exponentially damped propagators, since the complex variable t has a strictly positive real part outside the spectrum $\mathcal{S}$.

We are therefore now able to define in a precise way that we mean in this paper by a *homogeneous* sequence V_n: it is a sequence such that all k-point correlations, defined as Cesàro averages $\langle V_n V_{n-a} V_{n-b} \cdots \rangle$ exist, and obey the usual clustering conditions of statistical mechanics, for all values of k. Then the complex Lyapunov exponent $\Omega(E)$ has a perturbative expansion to all orders.

This expansion holds for an arbitrary sequence of potentials V_n, and has therefore numerous possible applications. In particular, much effort has been devoted to the disordered case, where the potentials are independent random variables. Besides Refs. 39 and 40, let us mention the study of "anomalies"[40–42] which occur at any rational value of the unperturbed IDOS Q/π. In a related problem, namely the vibration spectrum of harmonic chains with random masses, Ref. 43 presents a low-frequency expansion of the complex Lyapunov exponent up to the tenth order, using the same approach as above.

C. Application to spectral gaps

We now investigate the consequences of the results of the previous section for an arbitrary sequence, with, e.g., a quasiperiodic or singular continuous Fourier spectrum.

By virtue of Eq. (2.4), we have

$$\langle g_n g_{n-a} \rangle = \frac{V^2}{4 \sinh^2 t} \int d\mu(q)\, e^{iqa} \ . \tag{2.27}$$

Equation (2.23) then yields

$$\Omega^{(2)} = -\frac{V^2}{8 \sinh^2 t} \int d\mu(q) \frac{e^{2t - iq} + 1}{e^{2t - iq} - 1} \ . \tag{2.28}$$

As mentioned in the previous section [see Eq. (2.14)], the properties along the spectrum are obtained by letting t go to the imaginary axis: $t \to iQ$, where $0 \leq Q \leq \pi$ is the wave vector of the unperturbed states. For each value of Q, the integral in Eq. (2.28) will be eventually dominated by the structure of the Fourier intensity at $q = 2Q$.

In particular, to each power-law singularity (2.7) of the Fourier intensity at $q = q_0$, there corresponds the following power-law singularity of the characteristic function at $Q = Q_0 = q_0/2$:

$$\Omega^{(2)} \approx B V^2 (t - iQ_0)^{\alpha - 1} \tag{2.29}$$

with

$$B=\frac{2^{\alpha}}{8\sin^2 Q_0}\frac{\pi\alpha}{\sin(\pi\alpha)}(A_+e^{i\pi\alpha/2}+A_-e^{-i\pi\alpha/2}) .$$

The occurrence of a singularity in the perturbative expansion of $\Omega(E)$ is very likely to be the signature of a spectral gap. We are thus led to predict that each power-law singularity (2.7) in the Fourier intensity generates a gap in the spectrum, and that the IDOS inside that gap has the following limit value for $V\to 0$:

$$H_0=Q_0/\pi=q_0/(2\pi) . \tag{2.30}$$

Moreover, we can determine the scaling law between the gap width Δ and the strength V of the potential, by means of the following argument. The beginning of the expansion of the IDOS for $V\to 0$ and $Q\to Q_0$ reads

$$\pi H=Q_0+(Q-Q_0)+C_{\pm}|Q-Q_0|^{\alpha-1}V^2+\cdots$$
$$\text{with } C_{\pm}=\mathrm{Im}(Be^{\pm i\pi(\alpha-1)/2}) . \tag{2.31}$$

This singular expansion is reminiscent of the expansion around mean-field theory for critical phenomena. Below the critical dimension, the successive terms are more and more singular for small $(T-T_c)$, which plays a role similar to that of $(Q-Q_0)$. Hence we are tempted to evaluate the gap width along the lines of the Ginzburg criterion[44] for evaluating the size of the critical region, namely by requiring that both terms of the expansion are of the same order of magnitude at the gap edges. We thus obtain

$$\Delta\sim V^{\beta} \text{ as } V\to 0 \text{ with } \beta=2/(2-\alpha) . \tag{2.32}$$

We are therefore led to conclude that a gap opens in the spectrum only for strong enough singularities, namely for $\alpha\le 2$.

The above results are in agreement with known properties of spectra of periodic and quasiperiodic sequences. In both cases, the Fourier spectrum is made of delta peaks, corresponding to $\alpha=0$. Hence we predict $\Delta\sim V$. For a periodic sequence, with P atoms per cell, the spectrum generically consists of P bands. The band edges are determined by a polynomial "secular" equation, such that indeed $\Delta\sim V$ under generic circumstances.

Consider now a quasiperiodic or almost-periodic sequence of potentials V_n. Then the Fourier spectrum is made of delta peaks on a dense set of points q_0, called the frequency module of the spectrum. Equation (2.30) asserts that the value for $V\to 0$ of the IDOS inside a gap belongs to the frequency module of the potential. This perturbative result has been proven to hold independently of the potential strength (at least for a smooth enough potential). This is the celebrated gap-labeling theorem.[33,34] Equation (2.32) predicts that the gap widths are linear in the potential strength, in agreement with the standard perturbation theory, which also yields the amplitude $A(q_0)$ such that $\Delta\sim A(q_0)V$. This linear dependence has been verified quantitatively with a high accuracy in the case of the Fibonacci chain.[45]

For a more general sequence of potentials V_n, the main predictions of this section are the following.

(a) Each strong enough singularity in the Fourier intensity of the sequence (such that $\alpha\le 2$) generates a spectral gap.

(b) The IDOS inside this gap is given by Eq. (2.30), at least in the $V\to 0$ limit. We are tempted to conjecture that the gap-labeling theorem can be widely generalized, and that Eq. (2.30) gives actually the exact value of the IDOS, independently of the potential strength.

(c) The gap width Δ scales as the power β of the potential strength V, where the exponent β is related to the exponent α of the Fourier intensity by Eq. (2.32).

III. THE THUE-MORSE SEQUENCE

A. Definition

The Thue-Morse sequence[46,47] has been extensively studied in the mathematical literature,[37] as the prototype of a sequence generated by substitution, with highly nontrivial properties. Some physical properties of this sequence have been already explored in Refs. 10–13.

The Thue-Morse sequence ε_n has the following arithmetic definition. $\varepsilon_n=+1$ ($\varepsilon_n=-1$) if the number of digits "1" in the binary representation of n is even (odd). This sequence has been shown to be generated by a substitution, acting on two letters, A and B, such that

$$T_{\mathrm{TM}}:\begin{cases} A\mapsto AB , & \text{(3.1a)}\\ B\mapsto BA . & \text{(3.1b)}\end{cases}$$

Let A_k and B_k be the transforms of the letters A and B under k iterations of the substitution T_{TM}. These words have 2^k letters, and obey the recursion relations

$$A_{k+1}=A_kB_k, \quad k\ge 0 \tag{3.2a}$$

$$B_{k+1}=B_kA_k, \quad k\ge 0 . \tag{3.2b}$$

The Thue-Morse sequence is the limit of the words A_k. It is turned into a numerical sequence, such as that of Eqs. (1.2), by the following transcription:

$$\begin{cases}\text{letter } A\Longrightarrow \varepsilon=+1 , & \text{(3.3a)}\\ \text{letter } B\Longrightarrow \varepsilon=-1 . & \text{(3.3b)}\end{cases}$$

B. Fourier transform

The Fourier transform of the Thue-Morse sequence has long been known to be a singular continuous measure.[37] Here we study the Fourier measure in an elementary way, with emphasis on the local singularities of the form (2.7) of the Fourier distribution function. A similar analysis has been presented in Ref. 13, in a different language.

Let $g_k^A(q)$ and $g_k^B(q)$ denote the partial Fourier amplitudes associated with the words A_k and B_k, according to Eq. (2.1). These amplitudes obey the recursion relations

$$\begin{cases} g_{k+1}^A=g_k^A+e^{iq2^k}g_k^B , & \text{(3.4a)}\\ g_{k+1}^B=g_k^B+e^{iq2^k}g_k^A , & \text{(3.4b)}\end{cases}$$

with initial values $g_0^A=1$ and $g_0^B=-1$. The recursion relations (3.4) are explicitly soluble. The intensities $S_k(q)=2^{-k}|g_k^A|^2=2^{-k}|g_k^B|^2$ are given by

$$S_k(q)=\prod_{0\le l\le k-1}[2\sin^2(2^{l-1}q)]$$
$$=2\sin^2\left[\frac{q}{2}\right]S_{k-1}(2q)\ . \tag{3.5}$$

We refer the reader to Ref. 13 for a detailed study of this expression.

It can be deduced from Eq. (3.5) that the Fourier measure has singularities of the form (2.7) at *each rational* (nondyadic) value of q, in units of 2π.

Indeed, consider first "primitive" wave vectors $q_0=2\pi m/n$, with odd n. Then the sequence $2^{l-1}q \pmod{2\pi}$ is eventually periodic (i.e., periodic for $l\ge l_0$), with period $P\le n$. So are the factors of the product (3.5). Let F denote the product along one period,

$$F=\prod_{l_0+1\le l\le l_0+P}[2\sin^2(2^{l-1}q)]\ . \tag{3.6}$$

Then the asymptotic growth of the intensities reads

$$S_k\underset{k\to\infty}{\sim}F^{k/P}\sim L^{\gamma}\ \text{ with }\ \gamma=\frac{\ln F}{P\ln 2}\ . \tag{3.7}$$

We have replaced the label k by the word length (number of letters) $L=2^k$. A scaling argument[8,9] then shows that the Fourier distribution function $\mu(q)$ has a power-law singularity at $q=q_0$, of the form (2.7),

$$\mu(q)-\mu(q_0)\sim|q-q_0|^{\alpha}\ \text{ as } q\to q_0,\ \text{with } \alpha=1-\gamma\ . \tag{3.8}$$

It is clear from Eq. (3.5) that the exponent α of a given primitive wave vector q_0 (with odd denominator) also characterizes the singularities at wave vectors of the form $q_1=2^{-N}(q_0+2\pi M)$, with M,N integers. The smallest exponent, i.e., the strongest singularity, occurs for $q/(2\pi)=\frac{1}{3}$ and $\frac{2}{3}$, and equals $\alpha=2-\ln 3/\ln 2=0.415\,04\ldots$. A numerical investigation suggests that an infinity of primitive values q_0 have an exponent $\alpha<2$. The analysis of Sec. II hence predicts the existence of an infinity of spectral gaps at small enough potential strength V, and the gap widths are expected to scale according to Eq. (2.32).

We end this section with the behavior of the Fourier distribution function around the origin. It is clear from Eq. (3.5) that the intensity S_N vanishes as q^{2N} for small q, and hence the limit distribution function $\mu(q)$ vanishes more rapidly than any power of q as $q\to 0$. In a more quantitative way, a rough estimate of $\mu(q)$ is obtained by deriving from the exact expression (3.5) the approximate equation $\mu(q)\sim Kq^2\mu(2q)$, where K is some constant. This equation yields

$$\mu(q)\sim\exp\left[-\frac{(\ln q)^2}{\ln 2}+O(\ln q)\right]\ . \tag{3.9}$$

A similar singularity is present at all dyadic values of the wave vector: $q_0=2\pi M2^{-N}$.

C. The trace map

The discovery of a trace map in the case of the Fibonacci sequence[20–24] has permitted a considerably deeper understanding of the problem. It has been shown[38] that there exists a polynomial trace map for any sequence generated by a binary substitution. In the case of the Morse sequence, this map has been used in Refs. 10 and 11. Trace maps have also been used[30,31] for quasiperiodic sequences related to some classes of quadratic irrationals.

We present here the transfer matrix formalism and the derivation of the trace map in the present case, for sake of completeness. Equation (1.1) can be recast in matrix form

$$\begin{bmatrix}\psi_{n+1}\\ \psi_n\end{bmatrix}=\underline{T}_n\begin{bmatrix}\psi_n\\ \psi_{n-1}\end{bmatrix}\ \text{ with }\ \underline{T}_n=\begin{bmatrix}V_n-E & -1\\ 1 & 0\end{bmatrix} \tag{3.10}$$

in such a way that we have

$$\begin{bmatrix}\psi_{n+1}\\ \psi_n\end{bmatrix}=\underline{T}_n\underline{T}_{n-1}\cdots\underline{T}_2\underline{T}_1\begin{bmatrix}\psi_1\\ \psi_0\end{bmatrix}$$

so that the propagation along any finite part of the chain is described by the ordered product of the $\underline{T}_n$ along that segment.

Let $\underline{\mathcal{A}}_k$ and $\underline{\mathcal{B}}_k$ denote the products of the 2^k matrices $\underline{T}_n$ corresponding to the letters of the words A_k and B_k defined in Sec. III A. Since the elementary matrices $\underline{T}_n$ defined in Eq. (3.10) have a determinant equal to unity, the matrices $\underline{\mathcal{A}}_k$ and $\underline{\mathcal{B}}_k$ also share that property. By virtue of Eq. (3.2), these matrices obey an analogous recursion relation, namely

$$\underline{\mathcal{A}}_{k+1}=\underline{\mathcal{B}}_k\underline{\mathcal{A}}_k,\quad k\ge 0 \tag{3.11a}$$

$$\underline{\mathcal{B}}_{k+1}=\underline{\mathcal{A}}_k\underline{\mathcal{B}}_k,\quad k\ge 0\ . \tag{3.11b}$$

Let us now introduce the notation

$$\alpha_k=\mathrm{tr}\underline{\mathcal{A}}_k,\quad \mathcal{B}_k=\mathrm{tr}\underline{\mathcal{B}}_k,\quad \gamma_k=\mathrm{tr}(\underline{\mathcal{A}}_k\underline{\mathcal{B}}_k)\ . \tag{3.12}$$

It follows directly from Eq. (3.11) that $\alpha_{k+1}=\beta_{k+1}=\gamma_k$ $(k\ge 0)$. Moreover, $\beta_{k+2}=\mathrm{tr}(\underline{\mathcal{A}}_k^2\underline{\mathcal{B}}_k^2)$. To evaluate this quantity, we use the Cayley-Hamilton theorem, which results in the following identity, $\underline{M}$ being an arbitrary 2×2 matrix:

$$\underline{M}^2=(\mathrm{tr}\underline{M})\underline{M}-(\det\underline{M})\underline{\mathbb{1}}\ . \tag{3.13}$$

We end up with $\beta_{k+2}=\alpha_k\beta_k\gamma_k-\alpha_k^2-\beta_k^2+2$. Going to the usual parametrization,

$$x_k=\tfrac{1}{2}\alpha_k=\tfrac{1}{2}\mathrm{tr}\underline{\mathcal{A}}_k\ , \tag{3.14}$$

we obtain a closed-form recursion relation for the x_k,

$$x_{k+2}=4x_k^2(x_{k+1}-1)+1\quad(k\ge 1)\ . \tag{3.15}$$

This equation is already given in Refs. 10 and 11. It has to be completed with the appropriate initial conditions, namely,

$$x_1=\tfrac{1}{2}(E^2-V^2)-1\ , \tag{3.16a}$$

$$x_2=\tfrac{1}{2}(E^4+V^4)-E^2V^2-2E^2+1\ . \tag{3.16b}$$

For $V=0$, we have

$$x_k=\cos(2^kQ) \quad \text{with } E=-2\cos Q \tag{3.17}$$

as expected, since we just have a uniform chain, described as blocks of length $L=2^k$.

The trace map contains, in principle, enough information to describe the whole energy spectrum. Consider indeed the periodic approximants to the Thue-Morse sequence, defined as infinite periodic repetitions of the words A_k. These periodic chains have 2^k atoms per cell, and hence their spectrum is made of 2^k bands. Since propagation along one cell A_k is described by the transfer matrix $\underline{\mathcal{A}}_k$, the dispersion relation between E and the wave vector Q which labels the Bloch eigenstates is obtained by noticing that the eigenvalues of $\underline{\mathcal{A}}_k$ are $e^{\pm iQ}$. We have therefore $\mathrm{tr}\underline{\mathcal{A}}_k=2\cos Q$, i.e., $x_k=\cos Q$. Since x_k is a polynomial of degree 2^k in E, the dispersion relation has indeed 2^k branches, as expected.

In the following, we will mostly use trace maps to study the behavior of some gaps widths for small potential strength. The perturbative analysis for $V\to 0$ of a trace map such as Eq. (3.15) is indeed much more explicit than the general case exposed in Sec. II. The corresponding analysis for the Fibonacci sequence has been reported in Ref. 45.

D. The energy spectrum

The trace map provides a very easy way of studying symmetry properties of the energy spectrum. The initial values x_1 and x_2 of Eq. (3.16) are even functions of E and V, separately. The trace map ensures that this property is valid for all the x_k. Going to the limit of an infinite system size, we deduce the following properties. (a) The energy spectrum does not depend on the sign of the potential strength V. (b) At fixed V, the spectrum is symmetric with respect to the origin of energies. One has in particular the relation $H(E)+H(-E)=1$. Let us notice that properties (a) and (b) are equivalent, because of the symmetry (1.3).

1. *Location of gaps*

Figure 1 shows a plot of the IDOS against energy for $V=0.5$. All the most visible gaps can be labeled by either one of the following classes of values of IDOS. (a) As has been suggested in Refs. 10 and 11, there is a spectral gap at any *dyadic* value of the IDOS: $H=M2^{-N}$, with integer M (odd) and N. (b) The perturbative analysis of Sec. II also implies the presence of gaps for some *nondyadic rational* values of H, namely those corresponding to $\alpha<2$. We are therefore led to conjecture that all the spectral gaps of the Thue-Morse sequence are located at those two sets of values of the IDOS.

In the limit where the potential strength goes to infinity, the spectrum can be determined exactly. Indeed, to leading order in V, the nondiagonal terms can be neglected in Eq. (1.1), in such a way that the eigenstates

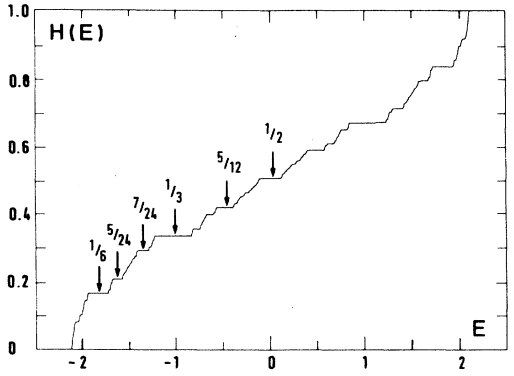

FIG. 1. Plot of the IDOS $H(E)$, against energy E, for the tight-binding equation (1.1), where the potentials V_n are given by the Thue-Morse sequence, with a potential strength $V=0.5$. The indicated values of the IDOS inside gaps correspond to the labeling discussed in the text.

are localized onto one site, and have an energy $E\approx V_n$. Since both letters A and B occur with the same density $\frac{1}{2}$, we predict half the states around $E=V$, and half around $E=-V$. The next approximation consists in taking into account whether the letters A or B are isolated, or belong to words of two, or more, identical letters, and in calculating the energy levels of those words, with Dirichlet boundary conditions. It can be argued that the error made is of the order of $1/V$. The n levels of a word of n letters A read

$$E_m=V+2\cos[m\pi/(n+1)],\quad 1\le m\le n\ . \tag{3.18}$$

Consider for the sake of definiteness the letters A. They occur in words of one or two letters. Let ρ_A and ρ_{AA} denote the densities of these words, so that $\rho_A+2\rho_{AA}=\frac{1}{2}$. These quantities can be determined by considering the substitution (3.1) which leaves the infinite sequence invariant. Since a word AA comes from a word BA, we have $\rho_{AA}=\rho_{BA}/2$. Moreover, a word BA comes either from a letter B or from a word AA, hence $\rho_{BA}=(\frac{1}{2}+\rho_{AA})/2$. We have therefore $\rho_A=\rho_{AA}=\frac{1}{6}$. The same result holds for the letters B. To this approximation, the six allowed values of the energy, $-V-1$, $-V$, $-V+1$, $V-1$, V, and $V+1$, occur with equal weights $\frac{1}{6}$. More precisely, the spectrum is included in intervals of width $O(1/V)$ centered around these six values of energy. The gaps with a nonvanishing width for $V\to\infty$ are therefore located at $H=k/6$ $(1\le k\le 5)$.

The rest of this section is devoted to a detailed study of two of the spectral gaps, namely those at $H=\frac{1}{2}$ and $\frac{2}{3}$.

2. *The gap at $H=\frac{1}{2}$*

The spectral gap at $H=\frac{1}{2}$, located at the center of the symmetric energy spectrum, is the best example of the first kind of gaps, namely those occurring at dyadic values of the IDOS. The edges of this central gap can, in fact, be determined analytically. Indeed, we have $x_1=1$ for

$$E_\pm=\pm[(1+V^2)^{1/2}-1]\ . \tag{3.19}$$

Equation (3.15) then implies that $x_k=1$ for all $k\geq 1$, and hence Eq. (3.19) gives the exact gap edges. This is a very particular feature of the trace map of the Thue-Morse sequence.

For large V, Eq. (3.19) has the expansion $E_\pm=\pm(V-1+1/V+\cdots)$. This expression agrees with the above analysis of the $V\to\infty$ limit. For small V, the gap width Δ vanishes as $\Delta=V^2-V^4/8+\cdots$. This gap has, therefore, $\alpha=2$, in the language of Sec. II. On the other hand, the associated singularity at $q=\pi$ of the Fourier distribution function is of the form (3.9). This essential singularity corresponds formally to $\beta=\infty$. Hence the gap under consideration is an exception to the rule (2.32).

3. *The gap at $H=\frac{2}{3}$*

The spectral gap at $H=\frac{2}{3}$ is the best example of a gap occurring at a nondyadic rational value of the IDOS, associated with a singularity of the form (2.7) in the Fourier transform, with $\alpha<2$. The above analysis of the large-V limit shows that the width of this gap goes to 1 in the $V\to\infty$ limit. The small-V regime can be studied by using the trace map (3.15). This approach confirms in an independent way the general analysis of Sec. II.

The value $H=\frac{2}{3}$ corresponds to $E_0=1$, in absence of potential. We are therefore led to expand the trace map for that value of energy. Since the x_k are even functions of V, the first nontrivial terms in their small-V expansion are proportional to V^2. If we set

$$x_k=-\tfrac{1}{2}+a_kV^2+\cdots\ , \tag{3.20}$$

the recursion relation (3.15) yields $a_k=a_{k-1}+6a_{k-2}$. The solution which satisfies the initial values $a_1=-\frac{1}{2}$, $a_2=-1$ reads

$$a_k=-\tfrac{2}{15}3^k+\tfrac{1}{20}(-2)^k\ . \tag{3.21}$$

On the other hand, the behavior of the x_k in absence of potential, for small $\delta E=E-E_0$, can be obtained by expanding Eq. (3.17). Keeping only the leading terms for large k, we get

$$x_k=-\tfrac{1}{2}-\tfrac{1}{2}(-2)^k\delta E-\tfrac{2}{15}3^kV^2+\cdots\ . \tag{3.22}$$

The behavior of the gap width Δ for small V can be extracted from this expansion by means of the following argument. The expression (3.22) is invariant under the simultaneous change of variable

$$k\mapsto k+1,\quad \delta E\mapsto\frac{\delta E}{2},\quad V^2\mapsto\frac{V^2}{3}\ . \tag{3.23}$$

The system size $L=2^k$ is multiplied by two by this operation, which can therefore be considered as a discrete *renormalization-group* transformation. Physical quantities related to the infinite system, such as the gap width Δ, have to be invariant under this transform. The only possible invariant relation is

$$\Delta\sim V^\beta\ \text{ with }\beta=\frac{2\ln 2}{\ln 3}=1.261\,68\ . \tag{3.24}$$

This expression agrees with Eq. (2.32), the value of α being given in Sec. III B. A similar perturbative renormalization-group analysis of the trace map can be performed for the gaps located at all the rational values of the IDOS, associated with the power-law singularities (3.8). The values thus obtained of the exponents β coincide numerically with those obtained from Eq. (2.32). We have not found in a closed form the trigonometric identities needed to check the agreement between both results in the general case.

Figure 2 shows a log-log plot of the width Δ of the gap located at $H=\frac{2}{3}$, against the potential strength V. Both the power law with exponent β at small V, and the limit $\Delta\to 1$ at large V, are very clearly seen.

IV. THE PERIOD-DOUBLING SEQUENCE

A. Definition

This sequence originates in the theory of dynamical systems. It describes the behavior of any system at the Myrberg point, which is the accumulation point of the period-doubling cascade. As will be discussed below, this sequence is generated by a binary substitution, and is almost-periodic.

The reader is referred to Refs. 35 and 36 for a detailed review of the theory of period doubling. We just recall here the basic result concerning the period-doubling sequence. Consider the one-parameter family of maps, e.g.,

$$f_\mu:\ x\mapsto 1-\mu x^2\ .$$

There exists an increasing sequence of values μ_n of the parameter such that the map f_μ has a superstable cycle of length 2^n. These values have an accumulation point μ_∞ usually called the Myrberg point. The orbit of the origin $x=0$ under the corresponding map is aperiodic. The period-doubling sequence is an infinite sequence of

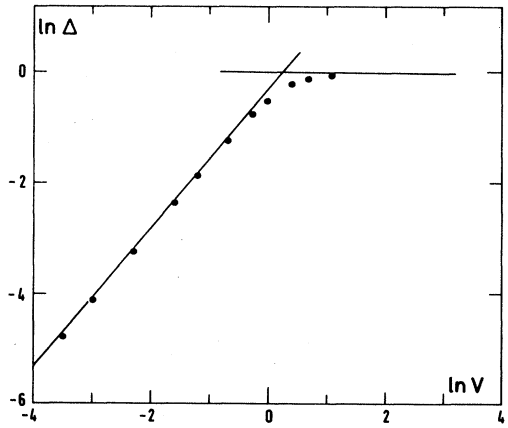

FIG. 2. log-log plot of the width Δ of the gap located at $H=\frac{2}{3}$, for the Thue-Morse sequence, against the potential strength V. The horizontal line shows the large-V limit $\Delta=1$. The other straight line has the slope β, according to the prediction (2.32, 3.24), and a fitted intercept.

letters R and L (symbolic dynamics) defined as follows: the nth letter is an R if the nth iterate $(f_{\mu_\infty})^n(0)$ is positive (i.e., to the *right* of 0), and an L if that number is negative (i.e., to the *left* of 0). It has been shown to be generated by the following binary substitution:

$$T_{\mathrm{PD}}:\begin{cases} R\mapsto RL\ , & (4.1a)\\ L\mapsto RR\ . & (4.1b)\end{cases}$$

In analogy with the Thue-Morse sequence, we define the words R_k and L_k as the transforms of R and L under k iterations of the substitution T_{PD}. These words have 2^k letters, and obey the recursion relations

$$R_k=R_{k-1}L_{k-1},\quad k\geq 1 \qquad (4.2a)$$

$$L_k=R_{k-1}R_{k-1},\quad k\geq 1\ . \qquad (4.2b)$$

The period-doubling sequence is the limit of the words R_k. We will consider it as a numerical sequence through the convention

$$\begin{cases}\text{letter } R\Longrightarrow \varepsilon=+1\ , & (4.3a)\\ \text{letter } L\Longrightarrow \varepsilon=-1\ . & (4.3b)\end{cases}$$

B. Fourier transform

The Fourier transform of the period-doubling sequence can be explicitly determined as follows. Let $g_k(q)$ denote the partial Fourier amplitudes associated with the words R_k so that

$$g_k(q)=\sum_{1\leq n\leq 2^k}\varepsilon_n e^{inq}\ . \qquad (4.4)$$

Since the beginning of the word R_k is R_{k-1}, the ε_n are the first 2^k letters of the infinite sequence. The form (4.1) of the substitution is such that these letters obey $\varepsilon_{2m-1}=1$, and $\varepsilon_{2m}=-\varepsilon_m$, for every $m\geq 1$. These relations imply the following functional equation:

$$g_k(q)+g_{k-1}(2q)=\frac{e^{i2^kq}-1}{2i\sin q}\ . \qquad (4.5)$$

For $q=0$, the right-hand side of Eq. (4.5) equals 2^{k-1}, and a summation of these terms yields

$$g_k(0)=\tfrac{1}{3}[2^k+2(-1)^k]\ . \qquad (4.6)$$

This result expresses that, among the 2^k letters of the word R_k, there are $N_k^R=\frac{1}{3}[2\times 2^k+(-1)^k]$ letters R, and $N_k^1=\frac{1}{3}[2^k-(-1)^k]$ letters L. In virtue of Eq. (2.6), Eq. (4.6) also shows that the Fourier transform of the infinite period-doubling sequence contains a δ-function peak at $q=0$, with an amplitude $C(0)=\frac{1}{3}$.

For $q=\pi$, the right-hand side of Eq. (4.5) equals -2^{k-1}, and a similar analysis demonstrates that there is also a δ-function peak, with an amplitude $C(\pi)=-\frac{2}{3}$. The iteration of Eq. (4.5) then shows that the Fourier transform contains δ-function peaks at all *dyadic* values of $q/(2\pi)$, and allows a recursive calculation of their amplitudes: the peak situated at $q=2\pi M\times 2^{-N}$ [$N\geq 1$, $1\leq M\ (\text{odd})<2^N$] has an amplitude

$$C(2\pi M2^{-N})=\frac{1}{3(-2)^{N-2}}\ . \qquad (4.7)$$

The period-doubling sequence is therefore almost-periodic. The distribution function $\mu(q)$ of the Fourier intensity can be derived in a closed form from Eq. (4.7). We obtain after some manipulations, with the notation $x=q/(2\pi)$,

$$\mu(q)=-\tfrac{1}{18}+x+\tfrac{5}{3}\mathrm{Frac}(x)-\tfrac{4}{3}f(x)$$
$$\text{with } f(x)=\sum_{N\geq 0}2^{-2N}\mathrm{Frac}(2^Nx)\ . \qquad (4.8)$$

$\mathrm{Frac}(x)$ denotes the fractional part of x, i.e., the difference between x and $\mathrm{Int}(x)$, the integer part of x. The function μ is by construction discontinuous at each dyadic value of x. The accumulation of these discontinuities has the result that it also develops weaker singularities at nondyadic values of x. Consider the example of $x=\frac{1}{3}$, which will be used hereafter in the study of spectral gaps. It is easily realized that the function f defined in Eq. (4.8) obeys the following functional equation, for $|y|<\frac{1}{3}$:

$$f(\tfrac{1}{3}+y)=\tfrac{2}{3}+y-\tfrac{1}{4}f(\tfrac{1}{3}-2y)\ . \qquad (4.9)$$

This equation implies the following exact behavior:

$$f(\tfrac{1}{3}+y)=\tfrac{8}{15}+2y+y^2P_{\pm}\left[\frac{\ln y}{\ln 4}\right]\ . \qquad (4.10)$$

The subscript $\pm$ refers to the sign of y, and $P_{\pm}$ are two periodic functions of their argument, with unit period. The Fourier distribution function μ has, therefore, a singular part at $q=2\pi/3$, with index $\alpha=2$, in the language of Sec. II A. This is precisely the marginal value which corresponds to an infinite gap exponent β. The actual behavior of the width of the associated gap will be determined in Sec. IV D.

C. The trace map

The following derivation of the trace map associated with the period-doubling sequence will proceed in analogy with the case of the Thue-Morse sequence, presented in Sec. III C.

Let $\underline{\mathcal{R}}_k$ and $\underline{\mathcal{L}}_k$ denote the products of the 2^k transfer matrices corresponding to the letters of the words R_k and L_k, defined in Sec. IV A. These matrices obey the following recursion relation:

$$\underline{\mathcal{R}}_{k+1}=\underline{\mathcal{L}}_k\underline{\mathcal{R}}_k\ ,\quad k\geq 0 \qquad (4.11a)$$

$$\underline{\mathcal{L}}_{k+1}=\underline{\mathcal{R}}_k^2,\quad k\geq 0\ . \qquad (4.11b)$$

If we introduce the notation

$$\alpha_k=\mathrm{tr}\underline{\mathcal{R}}_k,\quad \beta_k=\mathrm{tr}\underline{\mathcal{L}}_k,\quad \gamma_k=\mathrm{tr}(\underline{\mathcal{R}}_k\underline{\mathcal{L}}_k)\ , \qquad (4.12)$$

the recursion relation (4.11) implies $\alpha_{k+1}=\gamma_k$, and, by virtue of Eq. (3.13), $\beta_{k+1}=\alpha_k^2-2$. The calculation of $\gamma_{k+1}=\mathrm{tr}(\underline{\mathcal{L}}_k\underline{\mathcal{R}}_k^3)$ needs the following identity for the cube of an arbitrary 2×2 matrix $\underline{M}$:

$$\underline{M}^3=[(\mathrm{tr}\underline{M})^2-\det\underline{M}]\underline{M}-(\mathrm{tr}\underline{M})(\det\underline{M})\underline{\mathbb{1}}\ , \tag{4.13}$$

which is simply derived by iterating Eq. (3.13). We thus obtain $\gamma_{k+1}=(\alpha_k^2-1)\gamma_k-\alpha_k\beta_k$. With the usual normalization

$$x_k=\tfrac{1}{2}\alpha_k=\tfrac{1}{2}\mathrm{tr}\underline{\mathcal{R}}_k\ , \tag{4.14}$$

we obtain the following recursion relation:

$$x_{k+3}=(4x_{k+1}^2-1)x_{k+2}-2(2x_k^2-1)x_{k+1} \tag{4.15}$$

together with initial conditions

$$x_0=\tfrac{1}{2}(V-E)\ , \tag{4.16a}$$

$$x_1=\tfrac{1}{2}(E^2-V^2)-1\ , \tag{4.16b}$$

$$x_2=\tfrac{1}{2}(E^4-V^4)+EV(V^2-E^2)+2E(V-E)+1\ . \tag{4.16c}$$

It turns out that the trace map (4.15) has the following invariant:

$$\mathcal{I}-x_{k+2}-(4x_k^2-2)x_{k+1}+1\ . \tag{4.17}$$

Namely, Eq. (4.15) is such that $\mathcal{I}$ is invariant under the replacement of k by $k+1$. The initial values (4.16) are such that the quantity $\mathcal{I}$ vanishes. Hence we have

$$x_{k+2}=(4x_k^2-2)x_{k+1}-1\ . \tag{4.18}$$

Equation (4.18) is the final form of the trace map, which will be used hereafter.

D. The energy spectrum

1. Location of gaps

Let us first notice that, apart from the symmetry (1.3), the spectrum associated with the period-doubling sequence has no further symmetry property. Since that sequence is almost-periodic, with the set of dyadic numbers as a frequency module, a naive application of the gap-labeling theorem would suggest that all the gaps of the energy spectrum are located at dyadic values of the IDOS.

The following analysis of the large-potential-strength limit shows that this result is not correct. In analogy with the case of the Thue-Morse sequence, we have to determine the densities of the words made of only one type of letters, either R or L. We have seen in the previous section that the letters L have a density of $\frac{1}{3}$; they are always isolated. The letters R occur either isolated or in words RRR. Let ρ_R and ρ_{RRR} denote the densities of these occurrences, so that $\rho_R+\rho_{RRR}=\frac{2}{3}$. Under the substitution (4.1), an isolated R is generated from every R. Hence we have $\rho_R=\rho_{RRR}=\frac{1}{3}$. Therefore, up to terms of order $1/V$, the allowed values of the energy are $-V$, $V-\sqrt{2}$, V, and $V+\sqrt{2}$, with respective weights $\frac{1}{3}$, $\frac{1}{6}$, $\frac{1}{3}$, and $\frac{1}{6}$. The gaps with a nonvanishing width for $V\to\infty$ are located at the values $\frac{1}{3}$, $\frac{1}{2}$, and $\frac{5}{6}$ of the IDOS.

Figure 3 shows a plot of the IDOS against energy for $V=0.5$. All the most visible gaps can be labeled by values H of the IDOS such that $3H$ is a dyadic number. In other words, H is a rational number, and its denominator in irreducible form is either a power of 2, or three times a power of 2. We are led to conjecture that all the spectral gaps can be labeled in such a way, for any finite value of V. The gaps at dyadic values of the IDOS have a width linear in V for small potential strength V, just as for any standard quasiperiodic or almost-periodic system. The behavior of the other gaps is studied in the next subsection, with the example of the largest of these gaps, located at $H=\frac{1}{3}$.

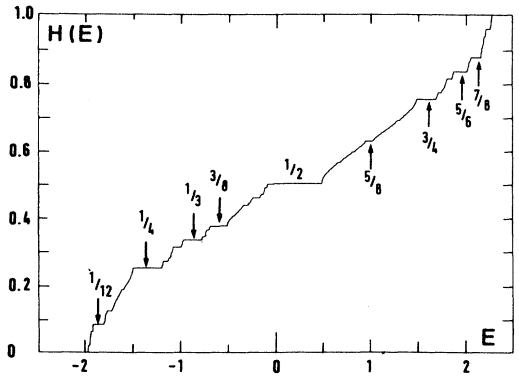

FIG. 3. Same as Fig. 1, for the period-doubling sequence, with $V=0.5$.

2. The gap at $H=\frac{1}{3}$

The above analysis of the large-V regime shows the existence of a spectral gap located at $H=\frac{1}{3}$. The corresponding singularity in the Fourier transform of the sequence is described by Eq. (4.10). It has an exponent $\alpha=2$, which corresponds through Eq. (2.32) to an infinite value of the gap exponent β. The actual small-V behavior of the width of this gap can be extracted from the trace map (4.18) as follows.

The value $H=\frac{1}{3}$ corresponds to $E_0=1$, in the absence of potential. For $E=E_0$, the x_k have the following small-V expansion:

$$x_k=-\tfrac{1}{2}+a_kV+b_kV^2+\cdots\quad(k\geq 1)\ . \tag{4.19}$$

The coefficients a_k obey the recursion relation $a_{k+2}=2a_k-a_{k+1}$, for $k\geq 1$, with initial values $a_1=0$, and $a_2=-1$. Hence we have

$$a_k=-\tfrac{1}{6}(-2)^k-\tfrac{1}{3}\ . \tag{4.20}$$

The coefficients b_k obey the recursion relation $b_{k+2}=2b_k-b_{k+1}-2a_k^2-4a_ka_{k+1}$, for $k\geq 1$, with the initial values $b_1=-\frac{1}{2}$ and $b_2=0$. Hence we have

$$b_k=\tfrac{1}{108}4^k+\tfrac{11}{108}(-2)^k-\frac{2k+1}{9}\ . \tag{4.21}$$

On the other hand, the expansion of the x_k for small $\delta E=E-E_0$ can be extracted from Eq. (3.17), which is indeed valid for any substitution of constant length 2. We finally obtain, keeping only the leading terms,

$$x_k = -\tfrac{1}{2} + \tfrac{1}{2}(-2)^k \delta E' - \tfrac{1}{3}\left[V + \frac{2k+1}{3} V^2 \right] + \cdots$$

$$\text{with } \delta E' = \delta E - \tfrac{1}{3} V + \tfrac{11}{54} V^2 + \cdots . \quad (4.22)$$

The occurrence of the combination $\delta E'$ has the following interpretation: δE is shifted by a quantity which is regular in V, and represents the shift of the center of the gap. This is confirmed by the fact that, among the higher-order terms which have not been written in Eq. (4.22), the terms in 4^k are proportional to $(\delta E')^2$.

In analogy with the analysis of Sec. III D 3, we notice that Eq. (4.22) is invariant under the following renormalization-group transform:

$$k \mapsto k+1, \quad \delta E' \mapsto \frac{\delta E'}{2}, \quad \frac{1}{V} \mapsto \frac{1}{V} + \frac{2}{3} . \quad (4.23)$$

The functional relation between the gap width Δ and V has to be invariant under this transformation. The only possibility is that of an essential singularity

$$\Delta \sim \exp\left[-\frac{3 \ln 2}{2V} \right] . \quad (4.24)$$

It is remarkable that this exponentially small gap width indeed corresponds to an infinite exponent β, as expected from the analysis of the Fourier transform performed in Sec. IV B. The perturbative result (4.22) only predicts the leading dependence of Δ on V. In analogy with the exponential singularities which occur, e.g., in the mass gap of nonlinear sigma models in two dimensions, we expect that Eq. (4.24) is modified by a power law

$$\Delta \approx V^{\varphi} \exp\left[-\frac{3 \ln 2}{2V} \right] . \quad (4.25)$$

We have tested this assumption numerically. With the notation $x = 1/V$, Eq. (4.25) implies

$$Y = \ln\left[\frac{\Delta(x)}{\Delta(x+1)} \right] \approx \frac{3 \ln 2}{2} + \varphi V . \quad (4.26)$$

Figure 4 shows a plot of this quantity against V. The intercept agrees perfectly with the expected value; the linear dependence on V is also very clear, and suggests the value $\varphi = \frac{1}{2}$ with a very good accuracy.

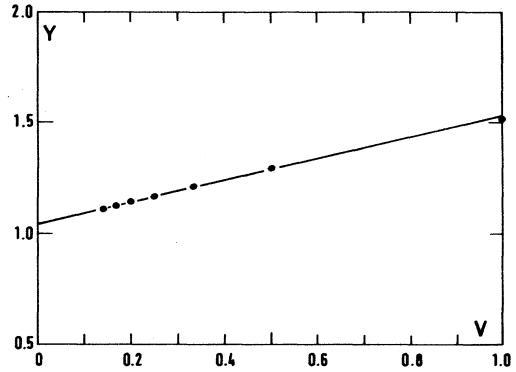

FIG. 4. Study of the gap located at $H = \frac{1}{3}$ for the period-doubling sequence: plot of the quantity Y defined in Eq. (4.26) against the potential strength V. The straight line has the exactly known intercept $(3 \ln 2)/2$; its fitted slope suggests the value $\varphi = \frac{1}{2}$.

V. THE "CIRCLE SEQUENCE"

A. Definition

As we have discussed in the Introduction, several recent works have been devoted to one-dimensional structures "beyond quasiperiodicity," with either an unbounded density fluctuation,[5–7] or a singular continuous Fourier transform.[8,9] The examples considered explicitly in these references were built from the following sequence:

$$\varepsilon_n = 2\chi_{\Delta}(nl) - 1 . \quad (5.1)$$

l and Δ are arbitrary numbers between 0 and 1, and $\chi_{\Delta}(x)$ denotes the characteristic function of the interval $[0, \Delta]$ (mod 1). In other words, we have $\varepsilon_n = +1$ ($\varepsilon_n = -1$) if the fractional part of nl lies between 0 and Δ (Δ and 1). The reader is referred to the above papers for a discussion of the possible physical origins of such a sequence. Since the definition (5.1) involves the multiples of a given number l (mod 1), it can be visualized as the result of a uniform rotation around the unit circle, with a rotation angle l. This is the origin of the name "circle sequence."

The circle sequence can be shown to be quasiperiodic for any values of l (irrational) and Δ. The sequence has nevertheless different kinds of behavior, according to number-theoretic properties l and Δ. In order to discuss this point, the simplest quantity to consider is the fluctuation of the sequence around its average $\langle \varepsilon_n \rangle = 2\Delta - 1$, defined by

$$\Sigma_N = \sum_{1 \le n \le N} \varepsilon_n - (2\Delta - 1)N . \quad (5.2)$$

It turns out that, when Δ and l are related by the Kesten condition[48]

$$\Delta = rl \pmod 1 , \quad (5.3)$$

where r is any (positive or negative) integer, the fluctuation Σ_N is bounded, and the sequence ε_n is a well-behaved quasiperiodic sequence. For instance, the Fibonacci sequence is recovered for $r = 1$, $\Delta = l = \tau^{-2}$, where $\tau = (\sqrt{5}+1)/2$ denotes the golden mean.

Whenever the condition (5.3) is not fulfilled, the fluctuation Σ_N is not bounded. Among those generic values of the parameters, there exist classes of self-similar sequences,[7–9] generated by substitutions. They correspond to quadratic algebraic values of the rotation angle l. The fluctuation Σ_N then grows as $\ln N$, with a periodic modulation.[7] The simplest of those sequences is obtained for

$$l = \tau^{-2}, \quad \Delta = \tfrac{1}{2} . \quad (5.4)$$

From now on, the circle sequence will refer to this particular choice of parameters. It has been shown[9] that it can be generated by the following substitutions acting on three letters A, B, and C:

$$T_C:\begin{cases} A\mapsto CAC\ , & (5.5a)\\ B\mapsto ACCAC\ , & (5.5b)\\ C\mapsto ABCAC\ . & (5.5c)\end{cases}$$

Let us define, in analogy with the previous sequences, the words A_k, B_k, and C_k as the transforms of the letters A, B, and C under k iterations of the substitution T_C. It is easily shown by induction that the lengths (numbers of letters) of these words are $l_k^A=F_{3k+1}$, $l_k^B=l_k^C=F_{3k+2}$, where the F_k is the kth Fibonacci number, defined by

$$F_k=F_{k-1}+F_{k-2} \text{ for } k\geq 2, \text{ with } F_0=0 \text{ and } F_1=1\ . \tag{5.6}$$

Since the F_k grow as τ^k, the lengths of these words grow as τ^{3k}. The sequence ε_n is then recovered by applying the identification

$$\begin{cases}\text{letter } A \text{ or } B \Longrightarrow \varepsilon=+1\ , & (5.7a)\\ \text{letter } C \Longrightarrow \varepsilon=-1\ . & (5.7b)\end{cases}$$

to the infinite word obtained as the limit of the A_k, for odd k, forgetting about its first letter.

B. Fourier transform

The definition (5.1) yields in a straightforward way the Fourier transform of the circle sequence. We limit ourselves to the values (5.4) of the parameters. Indeed, since the periodic function entering Eq. (5.1) has the Fourier series

$$2\chi_{1/2}(x)-1=\sum_{k \text{ odd}}\frac{2i}{\pi k}e^{-2\pi ik}\ , \tag{5.8}$$

the Fourier transform of the sequence ε_n reads

$$G(q)=\sum_{k \text{ odd},\, j}\frac{2i}{\pi k}\delta\left[\frac{q}{2\pi}-j+k\tau^{-2}\right]\ . \tag{5.9}$$

The circle sequence is therefore quasiperiodic. The δ-function peaks which lie between $q=0$ and $q=2\pi$ are located at $q/(2\pi)=-\mathrm{Frac}(k\tau^{-2})=\mathrm{Frac}(k\tau)$. It follows from Eq. (5.9) that the intensity distribution function $\mu(q)$ has the following expression (with the notation $q=2\pi x$):

$$\mu(q)=\sum_{\text{odd } k(\geq 1)}\frac{4}{(\pi k)^2}\{1+\mathrm{Int}[x-\mathrm{Frac}(k\tau)]+\mathrm{Int}[x+\mathrm{Frac}(k\tau)]\}\ . \tag{5.10}$$

Just as the expression (4.8) in the case of the period-doubling sequence, the function $\mu(q)$ given by Eq. (5.10) has, besides its discontinuities at $x=\mathrm{Frac}(k\tau)$, weaker singularities at other values of x. We consider the example of $x=\frac{1}{2}$, since it will be shown in the next section that it corresponds to a gap located at the center of the spectrum. The analysis of the singularity at $x=\frac{1}{2}$ is a rather intricate question. We do not pretend any rigor, and will content ourselves with the following rough argument. The singularity is clearly due to odd values of k such that $\mathrm{Frac}(k\tau)$ is closest to $\frac{1}{2}$. There exists a particular sequence ρ_m, which corresponds to the best possible approximation of $\Delta=\frac{1}{2}$ by multiples of $l=\tau^{-2}$ (mod 1). This problem is considered in Appendix A of Ref. 9 for arbitrary values of l and Δ. In the present case, the ρ_m are defined by

$$\rho_m-\rho_{m-1}=F_{3m+1}\quad(\rho_0=0)\ . \tag{5.11}$$

We have $\rho_1=1$, $\rho_2=4$, $\rho_3=17$, etc. Since the integers m and ρ_m have the same parity, only odd values of m have to be considered. It can be checked that these numbers grow as $\rho_m\sim\tau^{3m}$, and that $|\mathrm{Frac}(\rho_m\tau)-\frac{1}{2}|\sim\tau^{-3m}$. The contribution of these numbers to the sum in Eq. (5.10) yields therefore

$$\mu(q=2\pi x)-\mu(q=\pi)\sim(x-\tfrac{1}{2})^2P_{\pm}\left[\frac{\ln|x-\frac{1}{2}|}{3\ln\tau}\right]\ . \tag{5.12}$$

We claim that Eq. (5.12) gives the correct form of the scaling behavior of the function μ around $x=\frac{1}{2}$. This result is fully analogous to Eq. (4.10). The corresponding spectral gap will be studied in the next section.

C. The energy spectrum

Since the substitution (5.5) which generates the circle sequence acts on three letters, even though the sequence itself is binary, the result of Ref. 38 on the existence of a polynomial trace map does not apply. We have therefore to use other approaches to the study of the associated energy spectrum. In order to explore its symmetry properties, we notice that the sequence $\varepsilon_n'=-\varepsilon_n$ is also generated by Eq. (1.1), if the constant $\frac{1}{2}$ is added to the argument nl of the function $\chi_{1/2}$. This property of the value $\Delta=\frac{1}{2}$ holds independently of l. Hence both sequences ε_n and ε_n' have clearly the same energy spectrum. Using the symmetry property (1.3), we conclude that opposite values V and $-V$ of the potential strength lead to the same spectrum, and that this spectrum is symmetric with respect to the origin of energies: $H(E)+H(-E)=1$.

1. Location of gaps

In analogy with the previous sections, we can study the energy spectrum in an exact way in the limit of a large potential strength. To do so, it is easier to consider the circle sequence as made up of $+$ and $-$, and to use the definition (5.1) itself, rather than the substitution (5.5). We first notice that the symmetry just above implies that the clusters of $+$ and those of $-$ have the same densities. We consider the $+$ for sake of definiteness. Since $2\tau^{-2}>\frac{1}{2}$, the $+$ occur either isolated, or in clusters $++$. More precisely, ε_n is an isolated $+$ if and only if $\mathrm{Frac}(n\tau^{-2})$ lies in the interval $[\frac{1}{2}-\tau^{-2},\tau^{-2}]$. Since the length of this interval is $\lambda=2\tau^{-2}-\frac{1}{2}$, the allowed values of the energy, up to terms of order $1/V$, are the following: $-V$ and V, with weights λ, $-V-1$, $-V+1$, $V-1$, and $V+1$, with weights $(\frac{1}{2}-\lambda)/2=\frac{1}{2}-\tau^{-2}$. The gaps

with a nonvanishing width for $V \to \infty$ correspond therefore to the values $\frac{1}{2}-\tau^{-2}$, τ^{-2}, $\frac{1}{2}$, τ^{-1}, and $\frac{1}{2}+\tau^{-2}$, of the IDOS.

Among these five values of the IDOS, only two, namely τ^{-2} and τ^{-1}, belong to the frequency module of the sequence. Thus the gap-labeling theorem does not apply to the circle sequence. In analogy with the period-doubling sequence, the module has to be extended in order to describe all spectral gaps. Figure 5 shows a plot of the IDOS against energy, for $V=0.75$. All the most visible gaps can be labeled either as (a) $H=\mathrm{Frac}(k\tau)$, or as (b) $H=\frac{1}{2}+\mathrm{Frac}(k\tau)$, where k is an odd, positive, or negative integer. We conjecture that these numbers correspond to all the gaps, for any finite value of the potential strength. The gaps of class (a) are associated with δ-function peaks in the Fourier transform of the sequence; their width is expected to vanish linearly for small V. The behavior of the gaps of class (b) is a more delicate question. The next section presents accurate numerical data concerning the largest of them, located at $H=\frac{1}{2}$.

2. *The gap at $H=\frac{1}{2}$*

We have shown in the preceding subsection that there is a gap located at the center of the spectrum, with the value $H=\frac{1}{2}$ of the IDOS. This gap is the largest one for large V, since its width behaves as $\Delta=2V-2+O(1/V)$. The corresponding singularity of the Fourier transform is described by Eq. (5.12): it has an exponent $\alpha=2$, which yields an infinite gap exponent β. The very same behavior is observed in the case of the period-doubling sequence. It seems therefore reasonable to expect the same exponentially small behavior of the gap width Δ, namely

$$\Delta \sim e^{-C/V} \quad \text{as } V \to 0 \, . \tag{5.13}$$

Figure 6 shows a plot of the logarithm of the width Δ of the central gap, against $1/V$. Owing to the obvious fact that the value $E=0$ is the center of this gap, we have been able to "follow" it down to widths where we were limited by the accuracy of the computer. The quantity $2/V$ has been added to the y axis, in order to show in a better way how accurately the numerical data agree with Eq. (5.13), even for values of $1/V$ as small as four. The slope of the graph is 0.110. We predict therefore $C=1.890$, with an error bar of a few 10^{-3}.

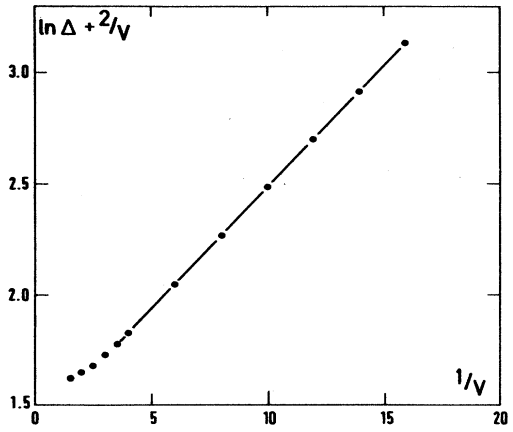

FIG. 6. Study of the central gap, located at $H=\frac{1}{2}$, for the circle sequence: plot of the quantity $\ln\Delta+2/V$, against $1/V$. The fitted straight line has a slope 0.110, yielding the value 1.890 for the constant C of the exponential singularity (5.13).

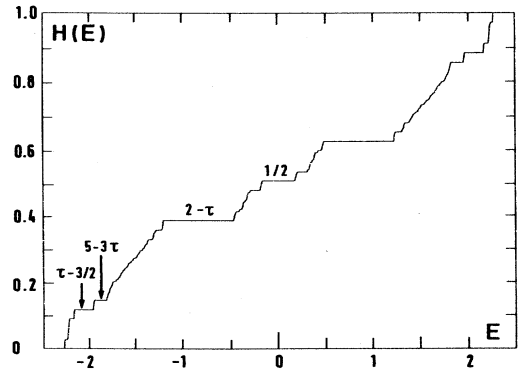

FIG. 5. Same as Fig. 1, for the circle sequence, with $V=0.75$.

VI. THE RUDIN-SHAPIRO SEQUENCE

A. Definition

The Rudin-Shapiro sequence[49,50] is one of the most classical arithmetic sequences,[37] with the Thue-Morse sequence studied in Sec. III. We have chosen to include this example in the present study, because the triviality of its Fourier transform does not exclude the presence of a dense set of gaps in the energy spectrum.

The Rudin-Shapiro sequence ε_n has the following arithmetic definition: $\varepsilon_n=+1$ (-1) if the number of times the word 11 occurs in the binary digit representation of n is even (odd). This sequence can be shown to be generated by the following substitution, acting on four letters, usually denoted by figures from 0 to 3, such that

$$T_{\mathrm{RS}}: \begin{cases} 0 \mapsto 02 \, , & \text{(6.1a)} \\ 1 \mapsto 32 \, , & \text{(6.1b)} \\ 2 \mapsto 01 \, , & \text{(6.1c)} \\ 3 \mapsto 31 \, . & \text{(6.1d)} \end{cases}$$

Let W_k denote the transform of the letter 0 under k iterations of the substitution T_{RS}. The Rudin-Shapiro sequence is the limit of the W_k, converted into a numerical sequence by the following transcription:

$$\begin{cases} \text{letter 0 or 2} \Longrightarrow \varepsilon=+1 \, , & \text{(6.2a)} \\ \text{letter 1 or 3} \Longrightarrow \varepsilon=-1 \, . & \text{(6.2b)} \end{cases}$$

B. Fourier transform

The study of the Fourier transform of the Rudin-Shapiro sequence involves some elaborate substitution

theory, which we do not want to tackle here. The reader is referred to Ref. 37 for a complete review of this area. The result of the following: the Fourier measure is absolutely continuous, with a uniform density

$$S(q)=1 \ . \tag{6.3}$$

Equation (2.4) then implies that the two-point correlation function S_a of the sequence is trivial:

$$S_a=\langle \varepsilon_n \varepsilon_{n+a} \rangle=\delta_{a,0}=\begin{cases}1 & \text{if } a=0\\ 0 & \text{otherwise} \ .\end{cases} \tag{6.4}$$

C. The energy spectrum

The perturbative analysis of Sec. II does not yield much information about the spectrum associated with the Rudin-Shapiro sequence, since the second-order term (2.28) is trivial, while the k-point functions ($k \geq 2$) involved in the higher-order terms of the expansion have not been investigated, as far as we know. Since the substitution (6.1) acts on four letters, there is also no trace map.

The substitution (6.1) is left invariant under the following interchange of letters:

$$\begin{cases}0 \leftrightarrow 3 \ , \\ 1 \leftrightarrow 2 \ .\end{cases} \tag{6.5}$$

Since this transform exchanges the letters which yield $\varepsilon=1$ and $\varepsilon=-1$, we conclude, by the same argument as in Sec. V C, that opposite values V and $-V$ of the potential strength correspond to the same energy spectrum, which is symmetric with respect to the origin of energies.

1. Location of gaps

In analogy with the previous sections, the energy spectrum can be determined exactly in the limit of a large potential strength V. As in Sec. V, we use the shorter notation $+$ and $-$ for the ε_n. By virtue of the above symmetry, the clusters of $+$ and $-$ have the same densities: we consider the $+$ for the sake of definiteness. They occur in clusters of length 1–4. Let ρ_n ($1 \leq n \leq 4$) denote the densities of these words. These quantities can be determined by the same method as in the previous examples, namely by listing which words generate the clusters of $+$ under the substitution (6.1), which words generate these words, etc. A closed set of relations is obtained after considering two iterations. The calculation is lengthy but simple; the result reads $\rho_1=\frac{1}{8}$, $\rho_2=\rho_4=\frac{1}{32}$, and $\rho_3=\frac{1}{16}$. The associated values of energy are given by Eq. (3.18), up to terms of order $1/V$. By sorting these values in increasing order, and adding up the corresponding densities, we obtain the set of values of the IDOS where there is a gap, at least for large enough V:

$$H \text{ or } 1-H=K/32 \ ,$$
$$\text{with } K=1,3,4,5,11,12,13,15,16 \ . \tag{6.6}$$

Figure 7 shows a plot of the IDOS, against energy, for a potential strength $V=1.25$. Besides the largest gaps, located at values of the IDOS with a denominator of $2^5=32$, smaller gaps can be seen, which are likely to correspond to dyadic values, with a higher power of 2 as denominator. We are led to conjecture that all the spectral gaps can be labeled by dyadic numbers.

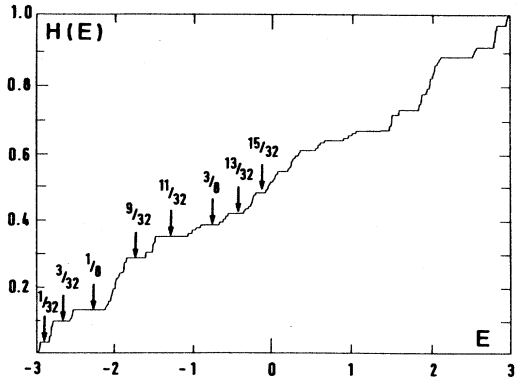

FIG. 7. Same as Fig. 1, for the Rudin-Shapiro sequence, with $V=1.25$.

2. The gap at $H=\frac{1}{2}$

The analysis of the previous subsection shows that there is a spectral gap at the value $H=\frac{1}{2}$ of the IDOS, at least for large values of the potential strength V. This gap is symmetric with respect to $E=0$, as the whole spectrum. Its width behaves as $\Delta=2V-2\tau+O(1/V)$ for large V, with $\tau=2\cos(\pi/5)$ (this number is equal to the golden mean introduced in sec. V).

Figure 8 shows a plot of the width Δ of this central gap, against V. This figure exhibits a striking feature that has not been met at all in the previous examples: the gap width has a highly nonmonotonic V dependence. The data correspond to an extrapolation of the exact gap edges of periodic systems with a cell W_k of length 2^k. The convergence properties are not good enough to explore smaller values of V than those shown on the plot. This complicated behavior is certainly due to this very particular feature of the Rudin-Shapiro sequence: since the Fourier transform is trivial, the origin of spectral

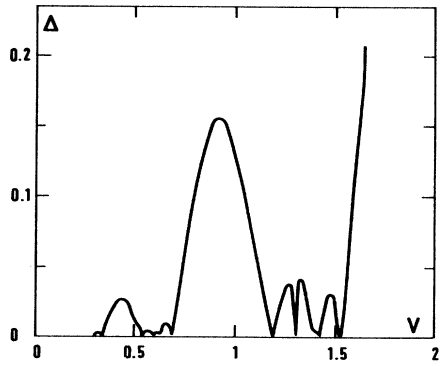

FIG. 8. Plot of the width of the central gap, located at $H=\frac{1}{2}$, for the Rudin-Shapiro sequence, against the potential strength V.

gaps is concealed in higher-order correlations in an intricate way.

VII. DISCUSSION

This paper contains two different kinds of results concerning the spectrum of the Schrödinger equation with an arbitrary homogeneous binary sequence of aperiodic potentials, namely those of a general perturbative analysis for small potential strength, and of the specific study of four example of self-similar sequences.

The perturbation analysis exposed in Sec. II demonstrates—in a nonrigorous way—the mechanism responsible for the universality of Cantor spectra. In the case of a quasiperiodic or almost-periodic potential, it is clear, at the level of perturbation theory, that each δ-function peak in the Fourier transform is responsible for a spectral gap. The gap-labeling theorem shows on a rigorous basis that this result is stable, beyond perturbation theory, at least for a smooth quasiperiodic potential. The value of the IDOS inside a gap belongs to the frequency module of the potential; it is strictly independent of the potential strength. In other words, no (extensive number of) states cross gaps when parameters are smoothly varied. Hence the IDOS can be used to "label" spectral gaps. Since the module is a dense set in nontrivial situations, the spectrum possesses a dense set of gaps. The fact that its total Lebesgue measure, or "bandwidth," vanishes is a much more difficult question, which needs a global analysis of the problem.

The results of Sec. II suggest that the concept of gap labeling can be extended to a much more general class of sequences, like, e.g., those having a singular continuous Fourier transform. The general idea is that *spectral gaps are labeled by the singularities of the Fourier intensity of the potential,* independently of parameters such as the potential strength. This statement has to be made more precise in several respects. Only the stronger singularities, namely the power laws with $\alpha<2$, with the notation of Eq. (2.7), do correspond to a gap; the gap width scales with the exponent β, given by Eq. (2.32). The gaps thus generated usually cover a dense set of values of the IDOS.

Sections III–VI, summarized in Table I, are devoted to the more specific study of four aperiodic but self-similar sequences, generated by substitution. These examples come either from mathematics—the (singular continuous) Thue-Morse sequence and the (absolutely continuous) Rudin-Shapiro sequence are among the most well-known "arithmetic" sequences, or from physics—dynamical systems, with the (almost-periodic) period-doubling sequence, or incommensurate structures, with the (quasiperiodic) circle sequence. This study makes use of various techniques. The large-potential regime can be studied exactly, by considering clusters of identical letters, for any sequence. This approach yields the extensive list of gaps which do not shrink to a vanishing width in the limit of an infinite potential strength. In the case of a binary substitution, there exists a polynomial trace map; its perturbative renormalization-group analysis permits an analytic determination of the scaling of some gap widths. Numerical data are in excellent agreement with the different kinds of analytical results.

The study of two of these specific cases has illustrated one of the difficult points in the relationship between singularities in the Fourier transform and spectral gaps, namely the violation of the gap-labeling theorem in the spectrum of the (almost-periodic) period-doubling sequence, and the (quasiperiodic) circle sequence. The spectra associated with both sequences have gaps outside the frequency module of the potential. Since the theorem has only been proved under rather restrictive conditions of smoothness of the potential, the observed effect is no paradox. On a heuristic basis, we have proposed to label the gaps with elements of an extended frequency module. For the period-doubling sequence, the module is the set of dyadic numbers; we claim that the values of the IDOS inside gaps are such that $3H$ is a dyadic. For the circle sequence, the module is made of numbers $j+k\tau$, with

TABLE I. A summary of the study, presented in Secs. III–VI, of the spectra associated with four self-similar sequences, generated by substitutions. For each sequence, the table gives the nature of its Fourier transform, its frequency module (when applicable), and the conjectured "labeling" of gaps in the energy spectrum, in terms of the integrated density of states.

Sequence	Fourier transform	location of gaps (values of the IDOS)
Thue-Morse (Sec. III)	singular continuous	(a) H dyadic (b) H rational such that $\alpha<2$
period doubling (Sec. IV)	almost-periodic module: dyadics	(a) $H\in$ module (b) $3H\in$ module
"circle" (Sec. V)	quasiperiodic module: $j+k\tau$ (k odd)	(a) $H\in$ module (b) $H-1/2\in$ module
Rudin-Shapiro (Sec. VI)	absolutely continuous $S(q)=1$	H dyadic

odd k; we claim that either H or $H-\frac{1}{2}$ belongs to the module. This conjecture has been guided by the exact results in the limit of the large potential strength showing, in particular, the existence of a gap at $H=\frac{1}{2}$ for the circle sequence, and at $H=\frac{1}{3}$ for the Rudin-Shapiro sequence. Both of these gaps correspond to a marginal ($\alpha=2$) singularity in the Fourier transform, and their width has been shown to be exponentially small in the potential strength. This behavior is likely to be common to all gaps such that the IDOS is not in the module. Loosely speaking, at least in the examples considered here, the gap-labeling theorem is only violated by transcendental terms at small potential strength. Because of these effects, gap labeling remains a subtle question in general.

The present work has been limited on purpose to the *local* study of the energy spectrum, with emphasis on spectral gaps, the key concept of gap stability and labeling, and the analysis of the scaling behavior of gap widths for a small potential strength. Another class of physically relevant questions, such as the vanishing Lebesgue measure of the spectrum, or its fractal dimension(s), involves the understanding of its *global* properties. The obtainment of general analytical results in this area, even at the level of perturbation theory, remains a difficult open subject.

ACKNOWLEDGMENTS

A major part of this work has been done during visits to the Universities of Oxford and Edinburgh. J. Ashraff and R. Stinchcombe are especially acknowledged for fruitful discussions. It is also a pleasure to thank C. Godrèche and Th. M. Nieuwenhuizen for useful comments.

[1]P. J. Steinhardt and S. Ostlund, *The Physics of Quasicrystals* (World Scientific, Singpore, 1987).
[2]B. Minchau, K. Y. Szeto, and J. Villain, Phys. Rev. Lett. **58**, 1960 (1987).
[3]C. Henley, J. Phys. A **21**, 1649 (1988).
[4]M. Perreau and J. C. S. Lévy, (unpublished).
[5]S. Aubry and C. Godrèche, J. Phys. (Paris) Colloq. **47**, C3-187 (1986).
[6]S. Aubry, C. Godrèche, and F. Vallet, J. Phys. (Paris) **48**, 327 (1987).
[7]C. Godrèche, J. M. Luck, and F. Vallet, J. Phys. A **20**, 4483 (1987).
[8]S. Aubry, C. Godrèche, and J. M. Luck, Europhys. Lett. **4**, 639 (1987).
[9]S. Aubry, C. Godrèche, and J. M. Luck, J. Stat. Phys. **51**, 1033 (1988).
[10]F. Axel, J. P. Allouche, M. Kléman, and M. Mendès-France, J. Phys. (Paris) Colloq. **47**, C3-181 (1986).
[11]F. Axel and J. Peyrière, C. R. Acad. Sci. (Ser. II) **306**, 179 (1988).
[12]R. Riklund, M. Severin, and Y. Liu, Int. J. Mod. Phys. B **1**, 121 (1987).
[13]Z. Cheng, R. Savit, and R. Merlin, Phys. Rev. B **37**, 4375 (1988).
[14]R. Merlin, K. Bajema, R. Clarke, F. Y. Juang, and P. K. Bhattacharya, Phys. Rev. Lett. **55**, 1768 (1985).
[15]J. Todd, R. Merlin, R. Clarke, K. M. Mohanty, and J. D. Axe, Phys. Rev. Lett. **57**, 1157 (1986).
[16]M. W. C. Dharma-wardana, A. H. MacDonald, D. J. Lockwood, J. M. Baribeau, and D. C. Houghton, Phys. Rev. Lett. **58**, 1761 (1987).
[17]L. Pastur, Commun. Math. Phys. **75**, 179 (1980).
[18]Th. M. Nieuwenhuizen and J. M. Luck, J. Stat. Phys. **41**, 745 (1985).
[19]Th. M. Nieuwenhuizen, J. M. Luck, J. Canisius, J. L. van Hemmen, and W. J. Ventevogel, J. Stat. Phys. **45**, 395 (1986).
[20]M. Kohmoto, L. P. Kadanoff, and C. Tang, Phys. Rev. Lett. **50**, 1870 (1983).
[21]M. Kohmoto, Phys. Rev. Lett. **51**, 1198 (1983).
[22]M. Kohmoto and Y. Oono, Phys. Lett. **102A**, 145 (1984).
[23]S. Ostlund, D. Pandit, D. Rand, H. J. Schellnhuber, and E. D. Siggia, Phys. Rev. Lett. **50**, 1873 (1983).
[24]S. Ostlund and R. Pandit, Phys. Rev. B **29**, 1394 (1984).
[25]C. Tang and M. Kohmoto, Phys. Rev. B **34**, 2041 (1986).
[26]S. N. Evangelou, J. Phys. C **20**, L295 (1987).
[27]G. Mato and A. Caro, J. Phys. C **20**, L717 (1987).
[28]M. Kohmoto, B. Sutherland, and C. Tang, Phys. Rev. B **35**, 1020 (1987).
[29]G. Paladin and A. Vulpiani, Phys. Rep. **156**, 147 (1987).
[30]G. Gumbs and M. K. Ali, Phys. Rev. Lett. **60**, 1081 (1988).
[31]M. Holzer, Phys. Rev. B **38**, 1709 (1988).
[32]T. Odagaki and H. Aoyama, Phys. Rev. Lett. **61**, 775 (1988).
[33]B. Simon, Adv. Appl. Math. **3**, 463 (1982).
[34]J. Bellissard, in *Statistical Mechanics and Field Theory: Mathematical Aspects, Proceedings of the International Conference, Groningen,* No. 257 of *Lecture Notes in Physics* (Springer, Berlin, 1986).
[35]P. Collet and J. P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems* (Birkhäuser, Boston, 1980).
[36]J. P. Eckmann, in *Chaotic Behavior of Deterministic Systems,* Proceedings of the Les Houches Summer School, 1981 (North-Holland, Amsterdam, 1983).
[37]M. Queffélec, *Dynamical Systems-Spectral Analysis, Lecture Notes in Mathematics,* (Springer, Berlin, 1987), Vol. 1294.
[38]J. P. Allouche and J. Peyrière, C. R. Acad. Sci. (Ser. II) **302**, 1135 (1986).
[39]Th. M. Nieuwenhuizen, Physica **113A**, 173 (1982).
[40]B. Derrida and E. Gardner, J. Phys. (Paris) **45**, 1283 (1984).
[41]M. Kappus and F. Wegner, Z. Phys. B **45**, 15 (1981).
[42]C. J. Lambert, Phys. Rev. B **29**, 1091 (1984); J. Phys. C **17**, 2401 (1984).
[43]Th. M. Nieuwenhuizen, J. Phys. A **17**, 1111 (1984).
[44]V. L. Ginzburg, Fiz. Tverd. Tela. (Leningrad) **2**, 2031 (1960) [Sov. Phys.—Solid State **2**, 1824 (1960)].
[45]R. B. Stinchcombe (unpublished).
[46]A. Thue, Selected Math. Papers, Oslo (1977).
[47]M. Morse, Trans. Am. Math. Soc. **22**, 84 (1921).
[48]H. Kesten, Acta Arith. **12**, 193 (1966).
[49]W. Rudin, Proc. Am. Math. Soc. **10**, 855 (1959).
[50]H. S. Shapiro, Master's Thesis, Massachusetts Institute of Technology 1951.