JOURNAL OF FUNCTIONAL ANALYSIS **48**, 408–419 (1982)

Cantor Spectrum for the Almost Mathieu Equation

J. BELLISSARD

University of Provence and CNRS, Centre de Physique Théorique, Luminy-Case 907, F-13288 Marseille Cedex 09, France

AND

B. SIMON*

Departments of Mathematics and Physics, California Institute of Technology, Pasadena, California 91125

Communicated by P. Lax

Received June 7, 1982

For a dense G_δ of pairs (λ, α) in R^2, we prove that the operator $(Hu)(n) = u(n+1) + u(n-1) + \lambda \cos(2\pi\alpha n + \theta)\, u(n)$ has a nowhere dense spectrum. Along the way we prove several interesting results about the case $\alpha = p/q$ of which we mention: (a) If $q\theta$ is not an integral multiple of π, then all gaps are open, and (b) If q is even and θ is chosen suitably, then the middle gap is closed for all λ.

1. INTRODUCTION

Recently, there has been an explosion of interest in the study of Schrödinger operators and Jacobi matrices with almost periodic potential (see, e.g., the review [16]). The general belief is that generically the spectrum is a Cantor set, i.e., a nowhere dense perfect, closed set. Since it is easy to prove that the spectrum is closed and perfect (see, e.g., [2]), the key is to prove that the spectrum is nowhere dense. This is definitely not always true: There are very special finite gap potentials, i.e., V's for which $-d^2/dx^2 + V(x)$ has a spectrum which is the finite union of closed intervals [6]. For the special class of limit periodic potentials, Chulaevsky [5], Moser [13] and Avron and Simon [2] have proven that generically (in the sense of dense G_δ) the spectrum is nowhere dense. Our goal in this paper is to prove that the spectrum is nowhere dense for another special class of potentials, specifically for the operators on l_2,

$$(Hu)(n) = u(n+1) + u(n-1) + \lambda \cos(2\pi\alpha n + \theta)\, u(n), \tag{1}$$

* Research partially supported by USNSF Grant MCS-81-19979.

0022-1236/82/120408–12$02.00/0

where λ, α, θ are parameters which label distinct H's. (1) is called the *almost Mathieu equation* because of its similarity to the Mathieu equation on L^2,

$$(Hu)(x) = -u''(x) + 2\lambda \cos(2x)\, u(x). \tag{2}$$

Our main goal in this paper is to prove (solving, at least in a weak sense, "the ten martini problem" of [16])

THEOREM 1. *For a set of pairs (λ, α) in R^2 which is a dense G_δ, (1) has a nowhere dense spectrum.*

We caution the reader that in the context of rational approximation, the two notions of "generic," namely, dense G_δ and full measure, are often distinct. For example, $S = \{\alpha \in [0, 1] \,|\, |\alpha - p/q| \geqslant Cq^{-k}$ some $C, k\}$ has measure 1 but its complement is a dense G_δ! The α's in our proof will be precisely those sufficiently well approximated by rationals and indeed the proof is so soft that we do not even have any estimates that tell us how well approximated α needs to be; we exploit the Baire category theorem to guarantee that the α's remaining are dense and uncountable.

For the Mathieu equation (2), it is an ancient result that *every* gap allowed by the theory of periodic potentials is open. For this reason, one may well expect that all the gaps allowed (in a sense we will make precise below) in the almost Mathieu equation are there. One of the results we prove in Section 3 partly supports this, viz.,

THEOREM 2. *If $\alpha = p/q$ is rational and $\theta q/\pi$ is not an integer, then all gaps of (1) are open for all $\lambda \neq 0$, i.e., the spectrum consists of exactly q disjoint closed intervals.*

However, Section 3 will also show that

THEOREM 3. *If $\alpha = p/q$ with q even and p odd and $\theta = 0$ [resp. π/q] if $q \equiv 0 (\mathrm{mod}\ 4)$ [resp. $q \equiv 2 (\mathrm{mod}\ 4)$], then one gap of (1) is closed for all λ.*

We find this result doubly surprising, because first it is contrary to the intuition from the Mathieu equation and because it seems very likely that a single closed gap for V fixed and all λ cannot occur in the Schrödinger case.

An important notion in our considerations is the ids (≡"integrated density of states"), $k(E)$, discussed from distinct viewpoints in [3, 4, 12]. On any gap of spec(H), $k(E)$ is constant, and we will prove Theorem 1 by showing that for any open interval J, $\{(\alpha, \lambda) \,|\, \mathrm{spec}(H(\alpha, \lambda))$ has a "gap" where $k(E) \in J\}$ is a dense open set. The word "gap" is in quotations since for α rational, gap is defined in a different way than usual (see Section 2). In Section 2, we prove that the set in question is open, and in Sections 3 and 4 that for any $\alpha = p/q$ rational, with q sufficiently large, $\{\lambda \,|\, \mathrm{spec}(H(\alpha, \lambda))$ has a "gap" where $k(E) \in J\}$ is dense in R.

It is a result of Bellissard *et al.* [4] (see also Johnson and Moser [12] for the Schrödinger case) that in any gap for (1) (or any a.p. Jacobi matrix), the value of $k(E)$ in a gap is always $(m\alpha)$, where $(\ ) \equiv$ fractional part of $\cdot$ and m is an integer. We believe that for a dense G_δ of (α, λ), one has that for every $m \neq 0$, there is a gap where $k(E) = (m\alpha)$. In Section 5, we describe briefly why we have not succeeded in various ways of trying to prove this stronger theorem.

We note that Hofstader [9] has pictures of $\text{spec}(H(\alpha, \lambda))$ for $\theta = 0$, $\lambda = 2$ and $\alpha = p/q$ computed numerically. The resulting picture shows Cantor behavior quite clearly. This is actually a very special picture; there is numerical evidence [17] that for $\lambda = 2$ and α irrational, $\text{spec}(H(\alpha, \lambda))$ has measure 0.

2. Reduction to the Rational Case

For any α, λ fixed, we define

$$S(\alpha, \lambda) = \bigcup_\theta \text{spec}(H(\alpha, \lambda, \theta)),$$

where $H(\alpha, \lambda, \theta)$ is the operator of (1). It is not hard to see that S is closed. For α irrational $\text{spec}(H(\alpha, \lambda, \theta))$ has no θ-dependency but there is nontrivial θ dependence for α rational. The integrated density of states $k(E; \alpha, \lambda, \theta)$ has nontrivial dependence on θ for α rational and general E but

LEMMA 2.1. *If (a, b) is disjoint from $S(\alpha, \lambda)$, then $k(E; \alpha, \lambda, \theta)$ is the same for all θ and all $E \in (a, b)$.*

Proof. Let $\alpha = p/q$ be rational. (a, b) must lie in a gap for each $H(\alpha, \lambda; \theta)$ $(\theta \in (0, 2\pi))$, and so on (a, b), $k(E; \alpha, \lambda, \theta)$ is independent of E and has a value m/q with $m = 0, 1,..., q$. Since k is also continuous in θ for E fixed (see [3]), this quantization of values implies that k is constant. Independence in θ for α irrational is true for any E[3]. ∎

This lemma says that it makes sense to talk about the value of $k(E)$ in a gap of $S(\alpha, \lambda)$. For any rational $a < b$ we define

$$A_{a,b} = \{(\alpha, \lambda) \mid S(\alpha, \lambda) \text{ has a gap on which } a < k(E) < b\}.$$

In this section we will show

THEOREM 1a. *$A_{a,b}$ is open*

and in Section 4 that

THEOREM 1b. *$A_{a,b}$ is dense if $(a, b) \cap [0, 1] \neq \phi$.*

Thus $\bigcap_{a<b \text{ rational}} A_{a,b}$ is a dense G_δ by the Baire category theorem. Theorem 1 follows from Theorems 1a and b and

LEMMA 2.2. *If $(\alpha, \lambda) \in \bigcap_{a,b} A_{a,b}$, then* $\operatorname{spec}(H(\alpha, \lambda))$ *is nowhere dense.*

Remark. Because of the mentioned discrete quantization of $k(E)$ on gaps where α is rational, $(\alpha, \lambda) \in \bigcap_{a,b} A_{a,b}$ implies that α is irrational, so $\operatorname{spec}(H(\alpha, \lambda))$ is independent of θ.

Proof. Let $E \in \operatorname{spec}(H(\alpha, \lambda))$. Then, by general principles [3], $k(E)$ is non-constant in every interval $(E-\delta, E+\delta)$, so for every such interval, we can find a, b rational with $k(E-\delta) < a < b < k(E+\delta)$ and $0 < a < b < 1$. Since $\operatorname{spec}(H(\alpha, \lambda))$ has a gap where $k \in (a, b)$, there is a gap in $(E-\delta, E+\delta)$. ∎

To prove Theorem 1a we require the following result of Avron and Simon [3] (Elliott [7] has independently obtained a related result):

LEMMA 2.3 [3, 7]. *Let $\alpha_n \to \alpha$, $\lambda_n \to \lambda, E_n \to E$ and suppose $E_n \in S(\alpha_n, \lambda_n)$. Then $E \in S(\alpha, \lambda)$.*

For completeness, we sketch the proof of this Lemma at the conclusion of this section.

Proof of Theorem 1a. Let $(\alpha, \lambda) \in A_{a,b}$. Then there exist $(e_0, e_1) \subset R \backslash S(\alpha, \lambda)$ and $k(E) \in (a, b)$ for $e_0 < E < e_1$. Pick e_0' and e_1' with $e_0 < e_0' < e_1' < e_1$. By Lemma 2.3, for (α', λ') sufficiently near to (α, λ), we must have $[e_0', e_1'] \subset R \backslash S(\alpha', \lambda')$. For all such (α', λ'), $k(e_0'; \alpha', \lambda', \theta)$ is independent of θ (by Lemma 2.1), so since $\int d\theta\, k(e_0', \alpha', \lambda; \theta)$ is continuous, $k(e_0'; \alpha', \lambda') \in (a, b)$ for (α', λ') near (α, λ). Thus, for (α', λ') near (α, λ), we have $(\alpha', \lambda') \in A_{a,b}$, i.e., $A_{a,b}$ is open. ∎

We will prove Theorem 1b by finding enough (α, λ) with α rational lying in $A_{a,b}$. Thus we have succeeded in reducing things to the rational case which we study in the next two sections.

Sketch of the Proof of Lemma 2.3 *(following Avron and Simon* [3])

(1) By the standard Floquet analysis of the periodic case, if α is rational, and $E \in \operatorname{spec}(H(\alpha, \lambda, \theta))$, then $Hu = Eu$ has a bounded solution.

(2) If α is irrational and $E \in \operatorname{spec}(H(\alpha, \lambda))$ we can find α_n rational and $E_n \to E$ with $E_n \in \operatorname{spec}(H(\alpha_n, \lambda, \theta = 0))$ since spectrum cannot suddenly appear under strong limits. By using (1) and shifting θ from 0 to a suitable point, we find u_n with $\|u_n\|_\infty \leqslant 1$ and $u_n(0) = 1$ so $H(\alpha_n, \lambda, \theta_n)\, u_n = E_n u_n$. Taking limits ($\theta_n$ has a limit point, θ_∞, by compactness of the circle) we find

a function u with $u(0)=1=\|u\|_\infty$ and $H(\alpha,\lambda,\theta_\infty)\,u=Eu$. Thus, if $E\in\operatorname{spec}(H(\alpha,\lambda))$ we have a bounded solution of $Hu=Eu$ for some θ. (Remarks: (2) is a quick proof of a result originally proven by Johnson [10].)

(3) If $E_n\in S(\alpha_n,\lambda_n)$, then by (2) there is u_n and θ_n so $H(\alpha_n,\lambda_n,\theta_n)\,u_n=E_n u_n$. By shifting θ_n we can suppose $u_n(0)=\frac{1}{2}$, $\|u_n\|_\infty\leqslant 1$ and then by taking a suitable limit point of θ and u, $H(\alpha,\lambda,\theta)\,u=Eu$ with $\|u\|_\infty\leqslant 1$ and $u\not\equiv 0$. If $u\in L^2$, then obviously $E\in\operatorname{spec}(H(\alpha,\lambda,\theta))\subset S(\alpha,\lambda)$. If $u\notin L^2$, let $u_m(n)=u(n)$ (if $|n|\leqslant m$), $=0$ (if $|n|>m$). Then $[(H-E)\,u_m](n)=0$ (unless $n=\pm m$, $\pm(m+1)$) and so $\|(H-E)\,u_m\|$ is bounded as $m\to\infty$. Since $\|u_m\|\to\infty$ as $m\to\infty$, $\|(H-E)\,u_m\|/\|u_m\|\to 0$ so $E\in\operatorname{spec}(H(\alpha,\lambda,\theta))\subset S(\alpha,\lambda)$. Thus $E\in S(\alpha,\lambda)$, i.e., Lemma 2.3 is proven. ∎

3. Analysis of the Rational Case, I: Dependence on θ

In the analysis of periodic Jacobi matrices, a basic role is played by the discrete analog of the discrimant, i.e., given $V(n)$ of period q, we set

$$f(E)=\operatorname{Tr}\left[\begin{pmatrix}E-V(0) & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}E-V(1) & -1\\ 1 & 0\end{pmatrix}\cdots\begin{pmatrix}E-V(q-1) & -1\\ 1 & 0\end{pmatrix}\right];$$

we use $f_{p/q,\lambda}(E,\theta)$ for this object when $V(n)=\lambda\cos(2\pi pq^{-1}n+\theta)$. The basic fact we will require to analyze θ dependence is

THEOREM 4. *Let p and q be relatively prime. Then*

$$f_{p/q,\lambda}(E,\theta)=g_{p/q,\lambda}(E)+2(-\lambda/2)^q\cos(q\theta). \tag{3}$$

Proof. Since $V(j)=\lambda/2[e^{2\pi i\alpha j}e^{i\theta}+e^{-2\pi i\alpha j}e^{-i\theta}]$, f is obviously a polynomial in $e^{i\theta}$ and $e^{-i\theta}$ of degree at most q. By cyclicity of trace, f is invariant under the transformation, $\theta\to\theta+2\pi p/q$. Since p and q are relatively prime, this implies that f must have a Fourier expansion containing only $e^{im\theta}$ with $m\equiv 0\ (\mathrm{mod}\ q)$. Thus

$$f_{p/q,\lambda}(E)=g_{p/q,\lambda}(E)+h_{+,p/q,\lambda}(E)\,e^{iq\theta}+h_{-,p/q,\lambda}(E)\,e^{-iq\theta}.$$

But, since each matrix has only one piece of order $e^{i\theta}$ or of order $e^{-i\theta}$, it is easy to read off the terms $h_\pm$. ∎

Recall the basic facts about how $f(E)$ behaves and how it relates to spec(H) (see [15] but note that since $-\Delta$ is approximated by $-u(n+1)-u(-1)+2u(n)$, the sign of E is opposite here to that in [15]).

For E very large, $f \to +\infty$. As E decreases, $f(E)$ passes through the value 2, down to where its value is -2. It is strictly monotone on the interval where $-2 < f(E) < 2$. Call E_1, the maximal point where $f(E) = -2$. It can happen that $f'(E_1) = 0$, in which case we say the top gap is closed and then just below E_1, f increases monotonically as E decreases until it reaches $+2$. If $f'(E_1) \neq 0$, then eventually as E is decreased there is a next point E_2 with $f(E_2) = -2$ and we say the top gap is (E_2, E_1). Below E_2,f is initially monotone till we reach the next point where $f(E) = 2$. This repeats so that there are points

$$E_0 > E_1 \geqslant E_2 > E_3 \geqslant E_4 > \cdots > E_{2j-1} \geqslant E_{2j} > E_{2j+1} \cdots > E_{2q-1}$$

so that the only points where $f(E) = 2$ are precisely the points E_l with $l \equiv 0, 3 (\mathrm{mod}\ 4)$ and the only points where $f(E) = -2$ are the E_l with $l \equiv 1, 2 (\mathrm{mod}\ 4)$. If $l \equiv 1$ [resp. 3] (mod 4(, then $f(E) < -2$ [resp. $f(E) > 2$] on (E_{l+1}, E_l) and if $l \equiv 2$ [resp. 0] (mod 4), then on (E_{l+1}, E_l), f lies in $(-2, 2)$ and f is strictly monotone increasing [resp. decreasing] as E decreases. The spectrum is precisely $\bigcup_{j=0}^{q-1} [E_{2j+1}, E_{2j}]$ and the "lth gap" is $(E_{2q-2l}, E_{2q-2l-1}) \equiv \Delta_l (l = 1,...; q-1)$. If $E_{2q-2l} = E_{2q-2l-1}$, we say the lth gap is closed. On the lth gap $k(E)$ has the value l/q.

A critical fact is that all the critical points of f occur in regions where $|f(E)| \geqslant 2$. We therefore have that all the critical points of the function g of (3) have $|g + 2(-\lambda/2)^q \cos(q\theta)| \geqslant 2$ for *all* θ, and thus

LEMMA 3.1. *All the critical points of g occur when* $|g(E)| \geqslant 2 + 2^{1-q} |\lambda|^q$.

We are now prepared for

Proof of Theorem 2. For a gap to be closed, $f_{p/q}(\cdot, \theta)$ must have a critical point where $|f| = 2$, so g must have a critical point where $|g(E)| = 2 \pm 2^{1-q} |\lambda| \cos(q\theta)$. By the lemma, this can only happen if $q\theta/\pi$ is an integer. ∎

The gaps are determined by where the lines $y = \pm 2 - 2(-\lambda/2)^q \cos(q\theta)$ intersect the curve $g(E) = y$. Depending on the value of $(-\lambda)^q/|\lambda|^q$, the lines move up or down as θ varies. The gaps open or close at both sides of a gap as the lines move up or down (if we look at the $-2 - \cdots$ line and down or up if we look at the $+2 - \cdots$ line). The minimal gap occurs at an extreme value of θ; either 0, π/q. Thus

PROPOSITION 3.2. (a) *If q is even or q is odd and $\lambda > 0$, then the gaps for $l = 1, 3,...$ have the property that $\bigcap_\theta \Delta_l(\theta) = \Delta_l(0)$ and for $l = 2, 4,...$ $\bigcap_\theta \Delta_l(\theta) = \Delta_l(\pi/q)$.*

(b) *If q is odd and $\lambda < 0$, then the gaps for $l = 1, 3, \ldots$ have the property that $\bigcap_\theta \Delta_l(\theta) = \Delta_l(\pi/q)$ and for $l = 2, 4, \ldots$ $\bigcap_\theta \Delta_l(\theta) = \Delta_l(0)$.*

Note. After completing this work, we received a preprint from Herman [18] which also states and proves Theorem 4.

4. Analysis of the Rational Case, II: Large λ

Our primary goal in this section will be to prove

THEOREM 5. *Fix q odd and $\theta = 0$ or π/q. Then for sufficiently large $|\lambda|$, all gaps $\Delta_l(\theta)$, $l = 1, \ldots, q-1$ are open.*

Remark. By Aubry duality, the result is also true for $|\lambda|$ sufficiently small.

Assuming this result, we conclude the proof of Theorem 1 by

Proof of Theorem 1b (*see Section* 2). We will show that for a dense set of pairs (α, λ) with $\alpha = p/q$, we have $\bigcap_\theta \Delta_l(\theta) \neq \phi$ for an l with $l/q \in (a, b)$. Fix any p/q with q odd, p relatively prime to q and q so large that for some l among $1, \ldots, q-1$, $l/q \in (a, b)$. Since the set of p/q is dense, we need only show $\{\lambda \mid \bigcap_\theta \Delta_l(\theta) \neq \phi\}$ is dense. But, by Proposition 3.2, for $\theta_0 = 0$ or π/q appropriately, this is just $\{\lambda \mid \Delta_l(\theta_0) \neq \phi\}$. The set of λ for which this gap is closed is given by the vanishing of an analytic function (the difference of two eigenvalues which can only cross each other and no other eigenvalues) and by Theorem 5 this function has no zeros for $|\lambda|$ large and real. Thus the set of possible bad λ's is finite and thus its complement is dense as we required. ∎

Our calculation for λ large is essentially a perturbation calculation made easier by the fact that it is very similar to one done by Avron and Simon [1], who in turn relied on the realization that they had a "tunnelling" problem (as we do) so that certain ideas of Harrell [8] can be used. As a preliminary, we note

LEMMA 4.1. *It suffices to prove Theorem* 5 *for the case $\theta = 0$ and $l = 1, 3, \ldots, q-2$.*

Proof. The transformation $u(n)$ to $(-1)^n u(n)$ maps H_0 (the operator (1) where $\lambda = 0$) to $-H_0$ and commutes with potentials, and so it takes $H_0 + \lambda \cos(2\pi\alpha n)$ to $-(H_0 - \lambda \cos(2\pi\alpha n))$ and thus the gaps $\Delta_1, \Delta_2, \ldots, \Delta_{q-1}$ for the case λ, $\theta = 0$ and those $\tilde{\Delta}_1, \ldots, \tilde{\Delta}_{q-1}$ for $-\lambda, \theta = 0$ are related by $\Delta_l = \tilde{\Delta}_{q-l}$ and so control of $\Delta_1, \ldots, \Delta_{q-2}$ for both λ and $-\lambda$ yields control of $\Delta_2, \ldots, \Delta_{q-1}$ (put in physicists' language the gaps for l odd are determined by periodic B.C. and those for l even by antiperiodic B.C. $u(n) \mapsto (-1)^n u(n)$ for q odd interchanges the two types of B.C.).

Since q and p are relatively prime, we can find A and B integers with $pA + qB = 1$. By changing (A, B) to $(A - q, B + p)$ we can suppose that A is even in which case B must be odd. The translation $n \to n + \frac{1}{2}A$ takes $\cos(2\pi\alpha n)$ to $\cos(2\pi\alpha n + \pi(1 - qB)/q) = -\cos(2\pi\alpha n + \pi q^{-1})$ since B is odd. Thus, the operator with λ, $\theta = \pi/q$ is unitarily equivalent to that for $-\lambda$, $\theta = 0$ so we need only study that case. ∎

We now use the fact that a closed gap when $f(E) = 2$ can only occur if an eigenvalue of the problem (1) with periodic B.C., i.e., $u(n+q) = u(n)$ is degenerate. Theorem 5 is thus reduced to

PROPOSITION 4.2. *Fix q = odd and take $\theta = 0$. Then all sufficiently large λ, the eigenvalues of Eq.* (1) *with the boundary condition $u(n+q) = u(n)$ are non-degenerate.*

Proof. Instead of (1) we study

$$\cos(2\pi\alpha n)\, u(n) + \lambda^{-1}(u(n+1) + u(n-1)) = E(\lambda^{-1})\, u(n). \tag{1'}$$

One notes that for $\lambda^{-1} = 0$, this equation has many doubly degenerate eigenvalues and we need only show that for $\lambda^{-1} \neq 0$ and small this degeneracy is broken. Note also the symmetry which is preserved by the boundary condition $n \to -n$. Thus, for any λ, we can look at even and odd eigenvalues. Suppose $q = 2k + 1$. Then for $\lambda = 0$ we have eigenvalues $E_l(0) \equiv \cos(2\pi\alpha l)$, $l = 0, 1, \ldots, k$. For $l \neq 0$, there are two eigenvalues $E_l^{\pm}(\lambda)$ near $E_l(0)$ whose eigenfunctions $u_l^{\pm}(\lambda; n)$ obey $u_l^{\pm}(\lambda, -n) = \pm u_l^{\pm}(\lambda, n)$. We must show $E_l^+ - E_l^- \neq 0$ for all $\lambda^{-1} \neq 0$ and small and each l. We use the following device borrowed from [1] (a discrete version of an idea of [8]): multiply (1′) for E^+ by $u^-(n)$, and subtract the product of (1′) for E^- by $u^+(n)$ and sum from $n \equiv 0$ to $n = k$:

$$E_l^+ - E_l^- \left(\sum_0^k u_l^+(\lambda, k)\, u_l^-(\lambda, k)\right) = \lambda^{-1}[u_l^+(\lambda, 0)\, u_l^-(\lambda, 1) + 2u_l^+(\lambda, k)\, u_l^-(\lambda, k)], \tag{4}$$

where we use the fact that the sum on the right telescopes and we have used $u_l^-(\lambda, 0) = 0$, $u_l^-(\lambda, -1) = -u_l^-(\lambda, 1)$, $u_l^{\pm}(\lambda, k+1) = u_l^{\pm}(\lambda, -k) = \pm u_l^{\pm}(\lambda, k)$. We can normalize $u_l^{\pm}$ by requiring $u_l^{\pm}(\lambda, l) = 1$, so that by an elementary application of perturbation theory,

$$u_l^{\pm}(\lambda, m) = \delta_{l,m} \pm \delta_{-l,m} + 0(\lambda^{-1}). \tag{5}$$

Thus the left side of (4) is $(E_l^+ - E_l^-)(1 + 0(\lambda^{-1}))$. Look at the eigenvalue equation (1′) at $m = l \pm 1$. By (5), we find

$$u_l^{\pm}(\lambda, l \pm 1) = \lambda^{-1}(E_l(0) - E_{l+1}(0))^{-1} + 0(\lambda^{-2}), \tag{6}$$

where we have used $E_l(\lambda) = E_l(0) + 0(\lambda^{-1})$. The equation at $m \neq l,\ l \pm 1$, shows

$$u_l^{\pm}(\lambda, m) = 0(\lambda^2), \qquad m \neq \pm l,\ \pm(l+1), \pm(l-1).$$

This argument is easily iterated. The result is

$$u_l^{\pm}(\lambda, m) = a_{l,m}^{\pm}(\lambda^{-1})^{|l-m|} + 0((\lambda^{-1})^{|l-m|+1})$$

for $m = 0, 1,..., k$, where

$$a_{l,m}^{+} = a_{l,m}^{-} \neq 0 \qquad \text{if} \quad m \neq 0,$$

$a_{l,m=0}^{+} \neq 0.$

Thus the two terms on the right of (4) are $c_1 \lambda^{-2l} + 0(\lambda^{-(2l+1)})$ and $c_2 \lambda^{2l-2k-1} + 0(\lambda^{-2l-2k-2})$ with $c_1 \neq 0 \neq c_2$. Since $2l$ is even and $2k + 1 - 2l$ is odd, the two terms cannot cancel and thus the right side of (4) is non-zero for λ^{-1} small. ∎

The same argument would work for q even (but we have to check it separately for $\theta = 0$ and $\theta = \pi/q$ and separately for periodic and antiperiodic B.C.) except that for one gap, the two terms that occur on the right side of (4) have the same order so there might be a cancellation. In fact, when we examined the case $q = 4$, we were somewhat surprised to discover by explicit calculation that this gap was permanently closed. This led to our discovering Theorem 3 whose proof we now give.

Proof of Theorem 3. The closed gap is the one with $l = \frac{1}{2}q$, i.e., the "middle" gap. Indeed, we will show that $E_{q-1} = E_q = 0$ in the language of Section 3. Consider first the case $q \equiv 0 (\text{mod}\ 4)$ and $\theta = 0$. The relevant boundary conditions are periodic. There are two subspaces $V^{\pm}$ where $u(m) = \pm u(-m)$ and since 0 and $\pm(\frac{1}{2}q)$ (which are equivalent mod q) are left invariant by $m \to -m$, $\dim(V^+) = \frac{1}{2}q + 1$, $\dim(V^-) = \frac{1}{2}q - 1$. Both are odd. Consider the covariance Lemma 4.1. $u(m) \to (-1)^m u(m)$ now leaves periodic B.C. invariant and takes $H_0 + \lambda \cos(2\pi\alpha n)$ to $-H_0 + \lambda \cos(2\pi\alpha n)$. Translating n by $q/2$ and using the fact that p is odd $\cos(2\pi\alpha n) \to \cos(2\pi\alpha n + \pi p) = -\cos(2\pi\alpha n)$ so by combining the two transformations, we see that $H_0 + \lambda \cos(2\pi\alpha n)$ with periodic B.C. is unitarily equivalent to its negative. Moreover, both transformations and thus their composition leave $V^{\pm}$ invariant. Since both subspaces are odd dimensional, the middle eigenvalue on each must be 0.

Consider next the case $q \equiv 2 (\mathrm{mod}\, 4)$, $\theta = \pi/q$. By a suitable translation, this is equivalent to looking at (1) with $\theta = 0$, but where n now takes the values $\pm\frac{1}{2}, \pm\frac{3}{2}, \ldots$. In this realization, define $V^{\pm}$ as before; they each have dimension $\frac{1}{2}q$ which is odd. The same transformation as above shows $H \restriction V^{+}$ and $H \restriction V^{-}$ are each unitarily equivalent to their negatives and again their middle eigenvalues are 0. ∎

5. Attempts at a Stronger Theorem

We would like to prove a stronger result than Theorem 1, namely (recall-(x) = fractional part of the real number x),

Conjectured result. For each integer $m \neq 0$, there exists a dense open set A_m of pairs (λ, α) so that if $(\lambda, \alpha) \in A_m$, then $S(\alpha, \lambda)$ has a gap in which $k(E) = (m\alpha)$.

In this final section, we *briefly* describe some ideas connected with this conjecture. Given our approach above, the natural approach is to do two things:

(a) Associate an integer, m, to a gap in $S(\alpha, \lambda)$ so that $k(E) = (m\alpha)$ in the gap and so that the gap in $S(\alpha', \lambda')$ which is there for $|\alpha - \alpha'| + |\lambda - \lambda'|$ small has the same m associated to it.

(b) For rational $\alpha = p/q$, and λ so all gaps are open, prove that the gaps at $k(E) = l/q$ have the associated m with $m = \pm 1, \pm 2, \ldots, \pm[\frac{1}{2}q]$ if q is odd.

The main problem is to verify (b). For suppose we succeed in (a) but there is never a gap with $m = +1$. This could happen if the gap with $k(E) = p/q$ has $m = q + 1$ rather than 1. The two approaches we have tried follow.

K-theory approach. If for all (α, λ) near (α_0, λ_0) there is a gap containing $(E_0 - \delta, E_0 + \delta)$ we find f, a continuous function which is 1 on $(-\infty, E_0 - \delta]$ and 0 on $[E_0 + \delta_0 \infty)$. Then $f(H(\alpha, \lambda))$ is, for (α, λ) near (α_0, λ_0), a projection in the C^*-algebra M_α associated to that α and $\mathrm{tr}_\alpha(f(H(\alpha, \lambda))) = (m\alpha)$ by the basic results of K-theory [14]. (This is how Bellissard *et al.* [4] prove the gap labeling theorem.) There is an explicit formula for m which one can show is continuous in (α, λ) so (a) can be proven. However, we do not see how to verify (b), that is, to show the possibility we describe after (b) does not occur.

Homotopy approach. We have another approach that is connected with an argument that Johnson [11] used as an alternate proof of the gap labeling theorem in the continuous case. If $E \notin \cup\ S(\alpha, \lambda)$, then for any θ, there is a unique point in $RP(1)$, the set of pairs (a, b) modulo multiplication by constants, so that the solution of $(Hu)(n) = Eu(n)$ with $(u(0), u(1)) = (a, b)$ decays as $n \to +\infty$. The point $p(\theta)$ in $RP(1)$ is easily seen to be continuous in θ. The map p has a winding number $\tilde{m}$. $\tilde{m}$ is easily seen to be independent of E as E varies in a gap and to be continuous in (α, λ) as long as a gap persists. On the basis of Johnson's proof in the continuous case, we believe that $k(E) = (-\tilde{m}\alpha)$ in the gap (the $-$ sign comes from the usual lack of sign in the finite difference operator). This would provide a new proof of the gap labeling theorem but we have not found a general proof of the fact.

By a perturbation argument, we have proven that for $\alpha = p/q$, q odd and $|\lambda|$ small, the gap where $k(E) = l/q$ has $\tilde{m} = a$, where a is the solution of $ap \equiv l (\text{mod } q)$ with $|a|$ minimal. Thus, for λ small, we have succeeded in step (b). If we could show that for $\alpha = p/q$ no gap closes for any λ, then we would by a continuity argument have a proof of the conjecture given at the beginning of this section.

Acknowledgments

J. Bellissard would like to thank W. Luxemburg and R. Vogt for the hospitality of Caltech, where some of this work was done, and we would like to thank J. Avron, W. Craig, A. Formoso, A. Katok, D. Lima, and D. Testard for valuable discussions.

References

1. J. Avron and B. Simon, The asymptotics of the gap in the Mathieu equation, *Ann. Phys.* **134** (1981), 76–84.
2. J. Avron and B. Simon, Almost periodic Schrödinger operators. I. Limit periodic potentials, *Commun. Math. Phys.* **82** (1981), 101–120.
3. J. Avron and B. Simon, Almost periodic Schrödinger operators. II. The density of states, Caltech preprint.
4. J. Bellissard, R. Lima, and D. Testard, Almost random operators: K-theory and spectral properties, in preparation; see also J. Bellissard, Schrödinger operators with almost periodic potentials: An overview, *in* "Springer Lecture Notes in Physics, 153."
5. V. Chulaevsky, *Russian Math. surveys*, in press.
6. B. Dubrovin, V. Matveev, and S. Novikov, Nonlinear equations of *KdV* type, finite zone linear operators and abelian varieties, *Russ. Math. Surveys* **31** (1976), 59 – 146.
7. G. Elliott, Gaps in the spectrum of an almost periodic Schrödinger operator, Warwick preprint.
8. E. Harrell, Double wells, *Commun. Math. Phys.* **75** (1980), 239–261.
9. D. Hofstader, Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields, *Phys. Rev. B* **14** (1976), 2239–2249.
10. R. Johnson, The recurrent Hill's equation, *J. Differential Equations*, in press.
11. R. Johnson, private communication.
12. R. Johnson and J. Moser, The rotation number for almost periodic potentials, *Commun. Math. Phys.* **84** (1982), 403–438.

13. J. MOSER, An example of a Schrödinger operator with almost periodic potential and nowhere dense spectrum, *Comment. Math. Helv.* **56** (1981), 198–224.
14. M. PIMSNER AND D. VOICULESCU, Imbedding the irrational rotation C^*-algebra in an AF-algebra, *J. Operator Th.* **4** (1980), 201–218.
15. M. REED AND B. SIMON, "Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators, Academic Press, New York 1978.
16. B. SIMON, Almost periodic Schrödinger operators, a review, *Adv. Appl. Math.*, in press.
17. D. THOULESS, private communication; G. ANDRE AND S. AUBRY, private communication.
18. M. HERMAN, A method for minorizing the Lyaponov exponent and several examples showing the local character of a theorem of Arnold and Moser on the torus of dimension 2, École Polytechnique preprint.