REGULARITY OF THE SPECTRUM
FOR THE ALMOST MATHIEU OPERATOR

NORBERT RIEDEL

(Communicated by David R. Larson)

Abstract. It is shown that the logarithmic potential associated with the
integrated density of states is constant on the spectrum of the almost Mathieu
operator in case the irrational frequency is sufficiently well approximable by
rationals in terms of a diophantine condition.

1. Introduction

The discovery that for discrete one-dimensional random Schrödinger operators
the Lyapunov index and the integrated density of states are intimately related, as
expressed by the Thouless formula, has given logarithmic potentials a prominent
place in the spectral analysis of these operators. In the special case of an almost
Mathieu operator with irrational frequency \( \alpha \), coupling constant \( \beta \) and integrated
density of states \( \mu \), this observation leads up to the fundamental inequality

\[
\int \log |z-s|d\mu(s) \geq \max\{0, \log \|\beta\|\} \quad \text{in } \mathbb{C}.
\]

The purpose in this paper is to show that equality holds in (1.1) whenever \( z \) is in
the spectrum \( \text{Sp}(\alpha, \beta) \) of this operator, in case \( \alpha \) is sufficiently well approximable
by rationals,

\[
\int \log |z-s|d\mu(s) = \max\{0, \log \|\beta\|\} \quad \text{for } z \in \text{Sp}(\alpha, \beta).
\]

This implies that \( \text{Sp}(\alpha, \beta) \) is a regular compactum and \( \mu \) is the equilibrium distribu-
tion of \( \text{Sp}(\alpha, \beta) \). In the proof of this statement we will employ the techniques which
were also used in [R3]. However, in a striking departure from the approach taken
in [R3], we will invoke an argument regarding the invertibility of polynomials in the
irrational rotation algebra under certain small perturbations, which will replace the
considerations in [R3] pertaining to the localization of generalized eigenfunctions.

It is tempting to ask whether some kind of inverse spectral theory for almost
Mathieu operators (or some more general class of operators) can be developed. In
such a theory the spectrum and the integrated density of states should be derived
from a specified (infinite) set of parameters. We shall take up this issue in section
4. In section 2 we give the proof of the statement announced above, and in section
3 we exhibit an observation that may be conducive to finding a proof for a broader
class of parameters \( \alpha, \beta \).
As in [R3], \( A = A_\alpha \) denotes the irrational rotation \( C^*\)-algebra associated with the irrational frequency \( \alpha \), which is generated by unitaries \( u, v \) satisfying the relation \( uv = e^{2 \pi i \alpha} vu \). We set \( w_{pq} = e^{-pq \pi i \alpha} u^p v^q \). For given real \( \beta \) let

\[ h(\alpha, \beta) = u + u^* + \beta(v + v^*) , \]

and for non-zero real numbers \( \gamma, \delta \) let

\[ h(\gamma, \delta) = \gamma^{-1} u + \gamma u^* + \beta(\delta^{-1} v + \delta v^*) . \]

Since for \( \beta \neq 0 \) the operators \( h(\alpha, \beta) \) and \( \beta h(\alpha, \beta^{-1}) \) have the same spectrum we suffer no loss in generality by confining our discussion to the case \( \beta \geq 1 \) only. In order to avoid combinatorial ramifications we will impose additional conditions on the parameters \( \gamma, \delta \):

\[ \gamma, \delta \geq 1 , \ \gamma \leq \beta \delta . \]

We consider the Fourier expansion of the resolvent of \( h(\gamma, \delta) \),

\[ (h(\gamma, \delta) - z)^{-1} = \sum_{p,q \in \mathbb{Z}} c_{pq}^{(\gamma, \delta)}(z) w_{pq} . \]

According to [R2] or [R3], section 4, we distinguish two types of points in the resolvent set of \( h(\gamma, \delta) \). A point \( z \) is of type I if the series

\[ \sum_{p,q \in \mathbb{Z}} \gamma^p \delta^q c_{pq}^{(\gamma, \delta)}(z) w_{pq} \]

is absolutely convergent and its limit equals the inverse of \( h(\alpha, \beta) - z \). A point is of type II if

\[ c_{pq}^{(\gamma, \delta)}(z) = \gamma^{-p} \delta^{-q} d_{pq}^{(+)}(z) \quad \text{for all} \quad p, q \in \mathbb{Z} , \]

where \( \{d_{pq}^{(+)}(z)\} \) is a double sequence with the properties

\[ d_{pq}^{(+)}(z) = 0 \quad \text{for} \quad q \leq |p| , \quad d_{pq}^{(+)}(z) = d_{-p,q}^{(+)} \quad \text{for all} \quad p, q \in \mathbb{Z} , \]

\[ d_{p, p+1}^{(+)}(z) = -(-\beta)^{-p-1} \quad \text{for} \quad p \geq 0 . \]

Distinct points in the same component of the resolvent set are of the same type. So it makes sense to distinguish between components of type I and components of type II. As in [R3], section 5, we define subharmonic functions \( \rho_I \) and \( \rho_{II} \) measuring the exponential order of decay or growth, respectively, for the Fourier expansions of the resolvent of the operators \( h(1, \delta) \):

\[ \rho_I(z) = \inf \{- \log \delta / \delta \geq 1 ; \quad z \quad \text{is contained in a component of} \quad \text{type I of the resolvent set of} \quad h(1, \delta) \} , \quad \text{for} \quad z \in \mathbb{C} \setminus \text{Sp}(\alpha, \beta) ; \]

\[ \rho_{II}(z) = \inf \{ \log \delta / \delta \geq 1 ; \quad z \quad \text{is contained in a component of} \quad \text{type II of the resolvent set of} \quad h(1, \delta) \} , \quad \text{for} \quad z \in \mathbb{C} . \]

2.1. Lemma. The following two statements are equivalent for some \( \chi \in \text{Sp}(\alpha, \beta) \):

(2.1.1) For every \( \varepsilon > 0 \) there are \( \gamma, \delta \in [1-\varepsilon, 1+\varepsilon] \) such that \( h(\gamma, \delta) - \chi \) is invertible in \( A \).

(2.1.2) \( \rho_{II}(\chi) = 0 \).
Proof. (2.1.1) $\implies$ (2.1.2): Let $\{(\gamma_n, \delta_n)\}$ be a sequence of pairs of positive numbers converging to (1,1) such that $h(\gamma_n, \delta_n) - \chi$ is invertible for all $n \in \mathbb{N}$. Since $h(\gamma, \delta), h(\gamma, -\delta), h(\gamma, \delta^{-1}), h(\gamma, -\delta^{-1})$ always have the same spectrum and $h(\gamma, \delta), h(\delta, \gamma)$ have the same spectrum in case $\beta = 1$, we may assume that $\gamma_n, \delta_n$ satisfy the condition (2.1.1) for all $n \in \mathbb{N}$. By [R2], Theorem 2.6 (c), or [R3], section 4, the point $\chi$ is of type II for $h(\gamma_n, \delta_n)$. Whence

$$(h(\gamma_n, \delta_n) - \chi)^{-1} = \sum_{p \in \mathbb{Z}, q \geq 0} \gamma_n^{-p} \delta_n^{-q} d_{pq}^{(+)}(\chi) u_{pq},$$

where the series on the right hand side is absolutely convergent. Since $d_{pq}^{(+)}(\chi) = d_{-p,q}(\chi)$, the series

$$\sum_{p \in \mathbb{Z}, q \geq 0} \delta_n^{-q} d_{pq}^{(+)}(\chi) u_{pq}$$

is also absolutely convergent. Thus $\chi$ is of type II for $h(1, \delta_n)$. Since $\delta_n \to 1$ as $n \to \infty$, this entails $\rho_{II}(\chi) = 0$.

(2.1.2) $\implies$ (2.1.1): If $\rho_{II}(\chi) = 0$, then for every $\varepsilon > 0$ there exists $\delta \in (1, 1 + \varepsilon)$ such that $h(1, \delta) - \chi$ is invertible, which in turn implies (2.1.1).

2.2. Theorem. The following statements are equivalent.

(2.2.1) For every $\chi \in \text{Sp}(\alpha, \beta)$ and every $\varepsilon > 0$ there are $\gamma, \delta \in [1 - \varepsilon, 1 + \varepsilon]$, such that $h(\gamma, \delta) - \chi$ is invertible.

(2.2.2) $\text{Sp}(h(1, \delta)) = \{z \in \mathbb{C} / \int \log |z - s| \, d\mu(s) = \log(\beta \delta)\}$ for $\delta \geq 1$.

If any of these statements holds true, so does 1.2. In particular, $\text{Sp}(\alpha, \beta)$ is a regular compactum and $\mu$ is its equilibrium distribution.

Proof. (2.2.1) $\implies$ (2.2.2): By Lemma 2.1 $\rho_{II}(\chi) = 0$ for every $\chi \in \text{Sp}(\alpha, \beta)$. By [R3], Lemma 5.3, we know that $\rho_{I}(z) < 0$, $\rho_{I}(z) + \rho_{II}(z) \geq 0$ for $z \in \mathbb{C}\setminus\text{Sp}(\alpha, \beta)$; $\rho_{II}(z) \geq 0$ for $z \in \mathbb{C}$. Since $\rho_{II}$ is upper semi-continuous, it follows that

$$\lim_{z \to \chi^{-}} \rho_{I}(z) = \lim_{z \to \chi^{+}} \rho_{II}(z) = \rho_{II}(\chi) = 0$$

for every $\chi \in \text{Sp}(\alpha, \beta)$. Whence the function

$$\rho_{0}(z) = \begin{cases} 
\rho_{I}(z) + \rho_{II}(z) & \text{for } z \in \mathbb{C}\setminus\text{Sp}(\alpha, \beta), \\
0 & \text{for } z \in \text{Sp}(\alpha, \beta)
\end{cases}$$

is upper semi-continuous in $\mathbb{C}$ and subharmonic in $\mathbb{C}\setminus\text{Sp}(\alpha, \beta)$. Since $\rho_{0}$ is non-negative in $\mathbb{C}$, we have

$$0 = \rho_{0}(\chi) \leq \frac{1}{2\pi} \int_{0}^{2\pi} \rho_{0}(\chi + re^{it}) \, dt,$$

for every $\chi \in \text{Sp}(\alpha, \beta)$ and $r > 0$. It follows that $\rho_{0}$ is subharmonic in $\mathbb{C}$. By [R3], Lemma 5.3 (iii)

$$\rho_{I}(z) = -\log |z/\beta| + o(1), \quad \rho_{II}(z) = \log |z/\beta| + o(1),$$

whence $\rho_{0}$ vanishes at infinity. By [R3], A4, this entails that $\rho_{0} = 0$. Thus we have shown that $\rho_{II}$ is a non-negative subharmonic function which is harmonic in $\mathbb{C}\setminus\text{Sp}(\alpha, \beta)$, which vanishes on $\text{Sp}(\alpha, \beta)$, and which, as $|z|$ approaches infinity, has the asymptotic property stated above. It follows from [R3], A5, that $\text{Sp}(\alpha, \beta)$ is
a regular compactum and $\rho_{\Pi} + \log \beta$ is its conductor potential. This entails that 
$\log \beta$ is Robin’s constant and the corresponding equilibrium distribution is the only probability measure $\sigma$ on $\text{Sp}(\alpha, \beta)$ having the property

$$
\int \int \log |t - s| d\sigma(s) d\sigma(t) = \log \beta.
$$

Since, on the other hand,

$$
\int \int \log |t - s| d\nu(s) d\nu(t) \leq \log \beta
$$

for any probability measure $\nu$ on $\text{Sp}(\alpha, \beta)$, and on the other hand, by (1.1),

$$
\int \int \log |t - s| d\mu(s) d\mu(t) \geq \log \beta,
$$

we conclude that $\mu$ is the equilibrium distribution of $\text{Sp}(\alpha, \beta)$ ([R3], A1). This shows that $\rho_{\Pi}(z) + \log \beta = \int \log |z - s| d\mu(s)$ in $\mathbb{C}$.

Since $\text{Sp}(h(1, \delta)) = \{z \in \mathbb{C} \setminus \text{Sp}(\alpha, \beta) \mid \rho_{\Pi}(z) \geq -\log \delta, \rho_{\Pi}(z) \geq \log \delta\}$ for $\delta > 1$, and the function $\rho_{\Pi}$ vanishes, we conclude that (2.2.2) holds true.

(2.2.2) $\implies$ (2.2.1): The validity of (2.2.2) implies that $\text{Sp}(\alpha, \beta)$ is contained in the resolvent set of $h(1, \delta)$ for $\delta > 1$. Since points in $\text{Sp}(\alpha, \beta)$ are of type II ([R2], 2.6(c), or [R3], section 4) it follows that $\rho_{\Pi}(\chi) = 0$ on $\text{Sp}(\alpha, \beta)$. Whence, due to Lemma 2.1, (2.2.1) holds true.

Proposition 2.1 in [R1] applied to $h(\alpha, \beta)$ yields that (2.2.1) holds whenever $\alpha$ satisfies a certain diophantine condition. We do not reproduce this condition here because it is technical in nature. This gives rise to the following conclusion.

**2.3. Corollary.** Whenever $\alpha$ is sufficiently well approximable by rationals in terms of a diophantine condition,

$$
\text{Sp}(h(1, \delta)) = \{z \in \mathbb{C} \mid \int \log |z - s| d\mu(s) = \log(\beta\delta)\}
$$

for $\delta \geq 1$.

Pooling the diophantine conditions in [CEY] and [RT] into one yields a class of regular Cantor sets in $\mathbb{R}$ having the additional property that the equilibrium distribution $\mu$ provides a complete gap labeling scheme: The map, which to every class of $\mu$-equivalent sets of the form $(-\infty, a] \cap \text{Sp}(\alpha, \beta)$ with $a \in \mathbb{R} \setminus \text{Sp}(\alpha, \beta)$, assigns the number $\mu\left((-\infty, a] \cap \text{Sp}(\alpha, \beta)\right)$, is a bijection onto the set $[0, 1] \cap (\mathbb{Z} + \alpha\mathbb{Z})$. Earlier examples of regular Cantor sets with Lebesgue measure zero were constructed by Nevanlinna [N] (ordinary linear Cantor set) and Tsuji [T] (general linear Cantor set). The occurrence of regular Cantor sets admitting a gap labeling scheme of the type indicated above will be looked at more closely in section 4.

If $K \subset \mathbb{C}$ is a regular compactum, then a general version of Nevanlinna’s Two Constant Theorem ([W], §A3) shows that the fibration of the level curves of the conductor potential of $K$ gives rise to a structure which we will call a scale of function algebras.
3.1. Definition. Let \( \{K_t\}_{t \geq 0} \) be a family of compact subsets of \( \mathbb{C} \) having the properties:

\begin{enumerate}
  \item \( s \leq t \implies K_s \subset K_t \).
  \item \( \bigcap_{s \leq t} K_s = K_t \), \( \bigcup_{t \geq 0} K_t = \mathbb{C} \).
  \item For every polynomial \( f \), \( 0 \leq t_1 \leq t_0 \leq t_2 \) implies
    \[ \|f\|_{t_0} \leq \frac{\|f\|_{t_2}}{2^{t_2-t_1}} \cdot \|f\|_{t_1} \],
\end{enumerate}
where \( \|f\|_t \) denotes the maximum norm of \( f \) on \( K_t \).

Then we call \( (\mathcal{P}, \|\cdot\|_t) \) an incomplete scale of function algebras (\( \mathcal{P} = \) algebra of complex polynomials in one variable).

Remark. a) Taking the completion of \( \mathcal{P} \) with respect to each norm \( \|\cdot\|_t \) in an incomplete scale of function algebras yields what we call a (complete) scale of function algebras.

b) Using some basic theory of function algebras it is not hard to see that condition (3.1.3) forces the family \( \{K_t\}_{t \geq 0} \) to be continuous in the following sense: \( \bigcup_{s \leq t} K_s = K_t \).

3.2. Theorem. For every \( t \geq 0 \) let \( K_t = \text{Sp}(h(1, \varepsilon t)) \cup (\text{interior of } \text{Sp}(h(1, \varepsilon t))) \). Then the family \( \{K_t\}_{t \geq 0} \) gives rise to a scale of function algebras.

Proof. Let \( f \) be a polynomial. Then \( z \mapsto f(u + u^* + \beta(z^{-1}v + zv^*)) \) is an analytic function on \( \mathbb{C} \setminus \{0\} \) which takes its values in the \( C^* \)-algebra \( \mathcal{A} \). Therefore, if \( \rho(a) \) denotes the spectral radius of \( a \),
\[ z \mapsto \rho(f(u + u^* + \beta(z^{-1}v + zv^*)) ) \]
is a subharmonic function on \( \mathbb{C} \setminus \{0\} \) ([1], p. 9, Theorem 1). Since the spectrum of \( f(u + u^* + \beta(z^{-1}v + zv^*)) \) depends on the modulus of \( z \) only, it follows from the “Three Circles Theorem” for subharmonic functions ([1], Theorem II.9),
\[ \rho(f(h(1, \delta_0))) \leq \rho(f(h(1, \delta_1)))^t \cdot \rho(f(h(1, \delta_2)))^{1-t} \]
whenever \( \log \delta_0 = t \log \delta_1 + (1-t) \log \delta_2 \) for some \( t \in [0,1] \). This establishes the validity of (3.1.3). The first part of (3.1.2) is an immediate consequence of the upper semi-continuity of spectral radius. The validity of the remaining properties is evident.

Even though Theorem 3.2 seems to be supporting the view to some extent that the equivalent properties of Theorem 2.2 may hold for all irrational numbers \( \alpha \), it is not sufficient to this end. For instance, if \( \{K_t\}_{t \geq 0} \) and \( \{K'_t\}_{t \geq 0} \) are families of compact subsets of \( \mathbb{C} \) which give rise to scales of function algebras, then the family \( \{K_t \cup K'_t\}_{t \geq 0} \) also gives rise to a scale of function algebras. This fact permits the construction of a large variety of function algebras which are not related to conductor potentials in the sense indicated earlier.

We now give a generic construction of all regular compact subsets of \( \mathbb{R} \) with a given positive capacity \( \beta \), such that the corresponding equilibrium distribution gives rise to a gap labeling scheme as indicated in section 2:
Choose a sequence $r_1, r_2, \ldots$ of numbers in the interval $[0, \beta^{-1})$, such that $\lim_{n \to \infty} r_n = 0$, let $r_0 = \beta^{-1}$, and define an open simply connected set

$$\Omega = \{ z \in \mathbb{C} | |z| < \beta^{-1}, \ I m \ z < 0 \} \setminus \bigcup_{n=-\infty}^{\infty} \{re^{-\pi[n\alpha]i} | \beta^{-1} - r_n \leq r \leq \beta^{-1} \},$$

where $[n\alpha]$ denotes the unique number in $[0, 1]$ having the form $m + n\alpha$. The Riemann mapping theorem yields an analytic univalent function $\phi$ from the upper half-plane $\Gamma$ onto $\Omega$. This map can be extended continuously on the closed upper half-plane including the point at infinity ($\mathbb{P}$, §9.3). We denote that extended map by $\phi$ again. By composing $\phi$ with a suitable fractional transformation of $\Gamma$ we can arrange that $\phi$ has the following additional property:

$$(+) \quad \lim_{|z| \to \infty} \phi(z) = 0.$$

If $\tilde{\phi}$ is another analytic univalent function from $\Gamma$ onto $\Omega$ having the property $(+)$, then there are real constants $a, b$ such that $\tilde{\phi}(z) = \phi(az + b)$. We define

$$\varphi(t) = -\frac{1}{\pi} \arg \phi(t) \quad \text{whenever} \quad \phi(t) \not\in (-\infty, 0),$$

and extend $\varphi$ continuously on $\mathbb{R}$, by assigning $\varphi(t)$ either the value 0 or 1 for those $t \in \mathbb{R}$ where it has not been defined already. Then $\varphi$ is either monotone increasing or monotone decreasing. We may adjust the map $\phi$ further to the effect that $\varphi$ becomes increasing. Since $\varphi(t) = 0$ for sufficiently small $t$ and $\varphi(t) = 1$ for sufficiently large $t$ we can conclude that there is a (unique) probability measure $\mu$ with compact support $K$, such that

$$\varphi(t) = \int_{-\infty}^{t} d\mu(s).$$

It follows that there is a real constant $c$ such that

$$\log |\phi(z)| = -\int \log |z - s|d\mu(s) + c, \ z \in \mathbb{P}.$$

By implementing a final adjustment of $\phi$ we may assume that $c = 0$. We thus obtain the representation

$$\phi(z) = \exp \left( -\int \log(z - s)d\mu(s) \right).$$

By construction, $K$ is a regular compactum with capacity $\beta$ and equilibrium distribution $\mu$. Furthermore, if $t \in \mathbb{R}\setminus K$, then

$$\mu(((-\infty, t]) \in \mathbb{Z} + \alpha \mathbb{Z}.$$

The possible absence of a gap in $K$ with the label $m + n\alpha \in [0, 1]$ is brought about by letting $r_n = 0$.

If $\tilde{\phi}$ is another analytic univalent function from $\Gamma$ onto $\Omega$ such that there is a probability measure $\tilde{\mu}$ on $\mathbb{R}$ giving rise to the representation

$$\tilde{\phi}(z) = \exp \left( -\int \log(z - s)d\tilde{\mu}(s) \right),$$

then $\tilde{\mu}$ is nothing but a translation of $\mu$. Thus the regular compactum $K$ is essentially uniquely determined by the domain $\Omega$ and the construction given above. If we are looking for regular compacta in $\mathbb{R}$ which are symmetric with respect to
the origin (as it happens to be the case with the spectrum of the almost Mathieu operator), the construction can be refined as to yield a unique regular compactum.

Conversely, suppose that $K \subset \mathbb{R}$ is a regular compactum with positive capacity $c$, let $\mu$ be its equilibrium distribution having the property that

$$\mu((-\infty, t] \cap K) \in \mathbb{Z} + \alpha \mathbb{Z} \quad \text{whenever} \quad t \in \mathbb{R} \setminus K.$$  

Let

$$\phi(z) = \exp\left(- \int \log(z - s) \, d\mu(s)\right), \quad z \in \Gamma,$$

and let $I_n$ be the (possibly empty) maximal interval in $\mathbb{R} \setminus K$ representing the possible gap with label $m + n\alpha$. We define

$$r_0 = c^{-1}, \quad r_n = \sup\{|\phi(t)| \mid t \in I_n\} \quad \text{for} \quad n \geq 1.$$

Here it is tacitly understood that $r_n$ is zero in case $I_n$ is empty. Then with $\Omega$ being defined as before, we can recover the measure $\mu$ by the construction above. In conclusion, the sequence $r_0, r_1, r_2, \ldots$ provides an essentially complete parametrization of regular compacta in $\mathbb{R}$ admitting the desired gap labeling scheme. The question which of those sequences give rise to the spectrum of an almost Mathieu operator needs further exploration. Finally, we would like to mention some recent work by Sodin and Yuditski which tackles the problem to develop a general inverse spectral theory for one-dimensional almost periodic Schrödinger operators along the lines the inverse spectral theory for periodic operators has been built.

REFERENCES


[N] R. Nevanlinna; Eindeutige analytische Funktionen; Springer Verlag (1953). MR 15:208c


[R2] N. Riedel; Spectra and resolvents of certain polynomials in the irrational rotation algebra; Indiana J. Math. 39 (1990), 937-945. MR 92a:47045

[R3] N. Riedel; The spectrum of a class of almost periodic operators; preprint.


DEPARTMENT OF MATHEMATICS, TULANE UNIVERSITY, NEW ORLEANS, LOUISIANA 70118