Persistence of gaps in the spectrum of certain almost periodic operators

Norbert Riedel

Department of Mathematics, Tulane University, New Orleans, LA 70118, USA
nriedel@tulane.edu

Abstract

It is shown that for any irrational rotation number and any admissible gap labelling number the almost Mathieu operator (also known as Harper’s operator) has a gap in its spectrum with that labelling number. This answers the strong version of the so-called “Ten Martini Problem”. When specialized to the particular case where the coupling constant is equal to one, it follows that the “Hofstadter butterfly” has for any quantum Hall conductance the exact number of components prescribed by the recursive scheme to build this fractal structure.

Introduction

The present work is concerned with the spectral properties of the simplest kind of discrete Schrödinger-type operators with an almost periodic potential. These operators form a self-dual class with respect to the Fourier transform, with exactly one operator being invariant. More specifically, a
precise one-to-one relationship between gaps and admissible gap labelling numbers will be established. While relegating the technical formulation of the problem and the outline of its proof to the first section, we are going to dwell for the rest of the introduction on the significance of this result in the special case of the self-dual operator for physics and mathematics.

Recent progress that has occurred in solid state physics through the development of new experiments which led to improved measurements, has made it possible to obtain improved physical evidence for the presence of the butterfly fractal spectrum. In [4] portions of the “Hofstadter butterfly” have been observed in lateral superlattices patterned on GaAs/AlGaAs heterostructure by exploiting the quantum Hall conductance as a diagnostic tool. One way to view the significance of the present paper is the recognition, that this diagnostic approach to detecting the fractal structure experimentally is on solid theoretical ground, at least as long as the “Hofstadter butterfly” is accepted as a paradigm for the quantum Hall effect: Knowing the specific value of the quantum Hall conductance, it is possible to allocate the corresponding gap components in the fractal structure in a prescribed way, without running the risk of missing one. This follows from the considerations in [8], where it was shown that, assuming that the strong version of the “Ten Martini problem” holds, the components associated with the same quantum Hall conductance can be counted by means of a specific combinatorial formula. In another experiment cold neutral atoms in an optical lattice were used to exhibit salient features of the butterfly fractal [6]. The employment of lasers allows the simulation of a magnetic flux through the lattice. The success of this approach for specific rotation numbers depends on the visibility of the particle density, which in turn depends on its periodicity. As shown by a specific experiment in [6] (figure 4(b)), the visibility decreases for the irrational rotation number \( \alpha = \frac{1}{2\pi} \). Again, as the present work will show, it is reassuring to know, that the butterfly fractal does not allow for unexpected transitions at irrational rotation numbers to occur, which could not be picked up through a suitably designed experiment.

We turn now to the mathematical significance of the present work, which at this point is somewhat speculative in nature. The set of quantum Hall numbers is a cyclic subgroup of the additive group of real numbers. If \( g \) is the positive generator of this group, then each quantum Hall number \( q \) can be written as \( q = kg \), for a suitable integer \( k \). As was shown in [8] under the assumption that the strong version of the “Ten Martini conjecture” holds, the number of components in the butterfly fractal with a common positive quantum Hall number \( q \) is equal to

\[
\Phi(2k), \quad \text{where} \quad \Phi(n) = \sum_{j=1}^{n} \varphi(j),
\]
and $\varphi$ is Euler’s totient function. On the other hand, an observation by J. Franel from the 1920s asserts that the Riemann hypothesis is true if and only if the relation

$$\sum_{j \leq \Phi(n)} \left( r_j^{(n)} - \frac{j}{\Phi(n)} \right)^2 = O(n^{-1+\varepsilon})$$

holds for any $\varepsilon > 0$. Here $r_j^{(n)}$ is the $j$th Farey fraction of order $n$. For a detailed exposition of this subject see Edmund Landau’s lectures on number theory [7], Band 2, Kapitel 13, or [3], Section 12.2. Note that the Farey fractions, in the order as they appear in this asymptotic formula, occupy a prominent position in the butterfly fractal. So it appears that the butterfly fractal holds information about the Riemann hypothesis, the exact nature of which remains yet to be determined. Inspired by the colour coding of the butterfly fractal according to the Hall conductance, which was also introduced in [8], and which has become quite popular in recent years, it is tempting to take a cue from Marc Kac, who is not only alleged to have offered ten martinis as a reward for the solution of the eponymous problem which is the subject of this work, but who also famously asked “Can one hear the shape of a drum?”, and pose the question, “Can one see the hue of the Riemann hypothesis?”

Finally, we turn to a description of the organization of the paper. In Section 1 the problem will be introduced in a form that is conducive to the employment of tools which are needed to solve it. Based on three propositions, two of which are known results, while the third one still needs to be established, the short proof of the major result will be given in that section. In Section 2, material from this author’s previous work will be assembled in a fashion that facilitates its usage in the present context, and a number of preparatory results will be established. In Section 3, the proof for the outstanding proposition will be provided.

Section 1

The observation that gaps tend to open up readily for small coupling constants due to basic perturbations of the degenerate case, that is when the coupling constant is equal to zero, naturally leads to a search for an argument that allows one to show that those gaps can not close as the coupling constant increases. The advantage of such an approach goes beyond mere expediency. By establishing the persistence of gaps as opposed to proving the existence of gaps for specific parameters, one can reach through to the
elusive self-dual case “from the outside”. It is the purpose of the present work to develop such an argument. Earlier attempts can be readily traced in the literature. For instance the major thrust in [5] is along the same line: If one substitutes the conjecture C3 in that paper, which at this point remains unproven, by the combination of Proposition 2 and Proposition 3 below, then one can simply follow the argument provided in “Remarque 3.2.4” in [5] to obtain the desired result. In conclusion our proof will require two legs to stand on: The first leg establishes that gaps with prescribed labels open up for sufficiently small (but by no means uniform!) coupling constants. This has been rigorously shown in [5]. The second leg guarantees that (open) gaps can’t close, as the coupling constant increases. Our argument to accomplish this has two components. First, we need to show that if a gap closes, then it follows that the Lyapunov exponent, considered as a function of two real parameters, namely the coupling constant and the spectral parameter, has a local maximum in one of the resolvent sets. Second, we need to prove that the Lyapunov exponent does not have any critical points. It is the proof of the second component that will occupy the main body of the paper.

We turn now to outlining the setting for the proof of that particular component. Let \( \alpha \) be an irrational number, let \( u \) and \( v \) be unitary operators satisfying the relation

\[
uv = e^{2\pi i} vu, \tag{1.1}
\]

and let \( A_\alpha \) be the (abstract) \( C^* \)-algebra generated by \( u \) and \( v \). Furthermore, let \( \tau \) be the unique tracial state on \( A_\alpha \). This is a positive linear functional, standardized by setting \( \tau(e) = 1 \), where \( e \) denotes the unit element in \( A_\alpha \), such that \( \tau(ab) = \tau(ba) \) holds for all elements in \( a, b \in A_\alpha \). In this setting we define for any positive coupling constant \( \beta \) the almost Mathieu operator as follows,

\[
h(\beta) = u + u^* + \beta(v + v^*). \tag{1.2}
\]

As usual, the upper right asterisk denotes the adjoint of an operator. The integrated density of states can now be identified with the restriction of the functional \( \tau \) to the abelian \( C^* \)-algebra generated by \( h(\beta) \). To obtain the integrated density of states proper, all one has to do is to represent this restricted functional by a probability measure on the spectrum of \( h(\beta) \). In light of the comments made above we can now formulate our first proposition, which establishes the first leg of the argument.

**Proposition 1.1** ([5]). For any number \( r \in [0, 1] \cap \mathbb{Z} + \alpha \mathbb{Z} \) there exists a positive number \( \beta_0 \) with the property that for any \( \beta \in (0, \beta_0] \), the operator
$h(\beta)$ has a gap in its spectrum with the label $r$. More specifically, there exists a real number $s_\beta$ in the resolvent set of $h(\beta)$, such that the spectral projection $p$ associated with the interval $(-\infty, s_\beta]$ has the property $\tau(p) = r$.

We turn now to the second leg. First, we need to define the Lyapunov exponent in a way that is compatible with the present settings.

$$L : \mathbb{R} \times \mathbb{C} \to \mathbb{R}, L(\beta, z) = \tau(\log |h(\beta) - z|).$$

While the operator $\log |h(\beta) - z| = \log[(h(\beta) - z)(h(\beta) - \bar{z})]^{\frac{1}{2}}$ is an element of $A_\alpha$ for $z \in \mathbb{C}\setminus\text{Sp}(h(\beta))$, $\text{Sp}(h(\beta))$ denoting the spectrum of $h(\beta)$, this is not the case for $z \in \text{Sp}(h(\beta))$. However, for any complex number this operator is contained in $L^1(A, \tau)$, the space of “integrable operators” associated with $A_\alpha$ and $\tau$. By virtue of the so-called Thouless formula the number $L(\beta, z)$ is seen to coincide with the usual definition of the Lyapunov exponent for $h(\beta)$ at $z$. We can now formulate our next statement.

**Proposition 1.2.** The function $L$ is jointly continuous in both variables. Moreover, for $\beta \leq 1$, $L(\beta, z) = 0$ for every $z \in \text{Sp}(h(\beta))$.

While this follows for rotation numbers satisfying a diophantine condition from the present author’s earlier work in conjunction with the semicontinuity of the spectrum (for badly approximable numbers see [12], Proposition 2.12, and for sufficiently well approximable numbers see [10], Corollary 2.3), a proof which does not rely on any diophantine condition is implicit in [2]. Indeed, the crucial Proposition 9 towards the end of that paper is valid for any pair of sufficiently smooth potentials, not just those which differ by the spectral parameter $E$. The remaining argument then carries through essentially without change.

The third statement, which will be proved in Sections 2 and 3, is as follows.

**Proposition 1.3.** For every $z \in \mathbb{R}\setminus\text{Sp}(h(\beta))$,

$$\tau((h(\beta) - z)^{-1}), \tau((h(\beta) - z)^{-1}v)) \neq (0, 0).$$

The following theorem, whose proof is the main objective in the sequel, has many precursors. The most up-to-date partial result appears to be [1], where the stated claim is that the strong version of the “Ten Martini Problem” holds for a set of badly approximable rotation numbers, and for all coupling constants other than 0 and 1. For a survey of contributions that preceded this partial result the reader is referred to that paper.
Theorem 1.4. For every $r \in [0, 1] \cap (\mathbb{Z} + \alpha \mathbb{Z})$ there exists $s \in \mathbb{R} \setminus \text{Sp}(h(\beta))$ such that the spectral projection $p$ associated with the interval $(-\infty, s]$ has the property $\tau(p) = r$.

Proof. By duality it suffices to consider the case $0 < \beta \leq 1$ only. Let $r \in [0, 1] \cap (\mathbb{Z} + \alpha \mathbb{Z})$, and choose $\beta_0$ as in Proposition 1.1. For some $\tilde{\beta} \in (0, \beta_0]$ let $s_{\tilde{\beta}}$ be as in Proposition 1.1. Let $\Omega_r$ be the connected component of $R = \{(\beta, t) \in \mathbb{R}^+ \times \mathbb{R} / t \in \mathbb{R} \setminus \text{Sp}(h(\beta))\}$ containing the point $(\tilde{\beta}, s_{\tilde{\beta}})$. Since $\tilde{\beta}$ can be chosen arbitrarily close to zero, it suffices to show that $\Omega_r$ contains a point whose first coordinate is equal to 1. Suppose this were not true. Then $\Omega_r \subset [0, 1] \times [-4, 4]$, and therefore $\overline{\Omega}_r$ is compact. By Proposition 1.2 the function $L$ is continuous on $\overline{\Omega}_r$, and it takes the value zero on the boundary $\partial \overline{\Omega}_r$. It follows that $L$ has a local maximum at some point $(\beta_e, s_e)$ in $\Omega_r$. Since $L$ is infinitely differentiable in $\Omega_r$, the gradient of $L$ at $(\beta_e, s_e)$ vanishes.

Thus

$$\frac{\partial L}{\partial \beta}(\beta_e, s_e) = \tau((h(\beta_e) - s_e)^{-1}(v + v^*) = 2\tau((h(\beta_e) - s_e)^{-1}v) = 0.$$

By Proposition 1.3 this is impossible. \hfill \Box

Remarks. 1) While the first of the two partial derivatives occurring in the proof is obvious, the second one warrants a few words of explanation. Differentiating the first of the above partial derivatives with respect to $\beta$, by invoking the formula

$$\frac{\partial}{\partial \beta}((z - h(\beta))^{-1} = (z - h(\beta))^{-1}(v + v^*)(z - h(\beta))^{-1},$$

and then antidifferentiating the result with respect to $z$ yields the claimed formula plus a function, which depends on $\beta$ only, $f(\beta)$ say. In order to show that $f(\beta)$ is actually zero, one would simply like to let $z$ approach infinity, because $\frac{\partial L}{\partial \beta}$ then approaches zero. This is of course impossible, since $z$ is confined to a (bounded) gap. Therefore, one has to replace the real parameter $z$ by a complex one, $z = t + i\varepsilon$, for a small positive number $\varepsilon$. Repeating the steps just outlined in this particular situation, and then letting $\varepsilon$ approach zero, yields the claimed formula for the partial derivative $\frac{\partial L}{\partial \beta}$.

2) It is worthwhile mentioning, that at this stage it is already clear that any possible critical point for the function $L$ has to be a local maximum.
Indeed, taking the second partial derivatives
\[
\frac{\partial^2 L}{\partial z^2}(\beta, z) = -\tau((z - h(\beta))^{-2}),
\]
\[
\frac{\partial^2 L}{\partial \beta \partial z}(\beta, z) = \tau((z - h(\beta))^{-1}(v + v^*)(z - h(\beta))^{-1}) = \tau((z - h(\beta))^{-2}(v + v^*)),
\]
\[
\frac{\partial^2 L}{\partial \beta^2}(\beta, z) = -\tau([(z - h(\beta))^{-1}(v + v^*)]^2),
\]
and applying the Cauchy–Schwarz inequality shows that the determinant of the Hessian of the function \(L\) is strictly positive. Since the diagonal entries of the Hessian are negative numbers, the claim follows.

**Section 2**

We now turn to the expansion and refinement of the settings introduced in Section 1. In the following, we assume throughout that \(\alpha\) is a fixed irrational number, and that \(\beta\) is a positive number less than 1. For \(p, q \in \mathbb{Z}\), we define the standardized monomials
\[
w_{pq} = e^{-pq\pi\alpha i}u^pv^q,
\]
and for \(z \in \mathbb{C}\setminus Sp(h(\beta))\),
\[
c_{pq}(z) = \tau((h(\beta) - z)^{-1})w_{pq}.
\]
The standardization ensures that these numbers are real valued whenever \(z\) is a real number. The double sequence \(\{c_{pq}(z)\}\) solves the following system linear equations for \(s = z\),
\[
\cos \pi\alpha q(x_{p+1,q} + x_{p-1,q}) + \beta \cos \pi\alpha p(x_{p,q+1} + x_{p,q-1}) = sx_{pq},
\]
\[
\sin \pi\alpha q(x_{p+1,q} - x_{p-1,q}) - \beta \sin \pi\alpha p(x_{p,q+1} - x_{p,q-1}) = 0,
\]
for all \(p, q \in \mathbb{Z}\), except \(p = q = 0\). We shall refer to the system (2.1) with the case \(p = q = 0\) exempted by (2.1)*.

Before we proceed with our objective, we are going to dwell a little on the linear system (2.1). First, if one multiplies the second equation by the imaginary unit \(i\), and adds the result to the first equation, then one obtains Harper’s equation on the two dimensional lattice, which is so common in
physics, and there are no algebraic complications attached to this equation. By contrast, the system (2.1) and its truncated version (2.1)*, combines two features which call for a specifically designed approach to generate and analyze its solutions (see [11]). On the one hand, the system is largely overdetermined, giving rise to redundancies: Asymptotically there are roughly twice as many equations than variables. On the other hand, the system is degenerate along the diagonals \( p = q \) and \( p = -q \): a recursion involving \( 4 \times 4 \) matrices to generate the solutions of this system collapses as one tries to cross either one of those two axes. The tenet underlying the proof of Proposition 1.3 is that these two features, reflecting intrinsic properties of the operator \( h(\beta) \), must also hold the key to its more elusive spectral properties. Generally speaking, for numbers \( s \) in the spectrum of \( h(\beta) \) the system (2.1) yields uniformly bounded solutions, which are obtained by evaluating certain states defined on the \( C^* \)-algebra \( A_\alpha \) at the standardized monomials \( w_{pq} \). (In [9], where it was shown that the linear dimension of the space of uniformly bounded solutions is always either equal to one or to two, they were referred to as “eigenstates”.) For \( s = z \) in the resolvent set of \( h(\beta) \), the double sequence \( \{ c_{pq}(z) \} \) is always exponentially decaying as \( |p| \to \infty \) and \( |q| \to \infty \). This is the crucial property we shall exploit in the proof. However, in order to take advantage of it, we first need to “homogenize” the double sequence. In other words, we have to find a double sequence solving the homogeneous system (2.1), which preserves some measure of that exponential decay, but also shares the vanishing conditions that hail from a possible critical point for the function \( L \). To this effect, we need to delve a bit deeper into system (2.1)*.

The solutions to this system form a linear space of dimension 6. Up to a scaling factor to be determined below, there are four solutions with the property that each of those has non-vanishing coefficients only in exactly one of the four sectors separated by the lines \( p = q \) and \( p = -q \) in the two-dimensional lattice. The components of these four solutions are nothing but the “Fourier coefficients” of the resolvent of perturbations of \( h(\beta) \), which are obtained by multiplying the generators \( u \) and \( v \) with suitable complex numbers of modulus larger than one or less than one. Put in technical terms, expanding these resolvents in terms of the standardized monomials \( w_{pq} \) yields a “non-commutative” multiple Laurent series whose coefficients are exactly those solutions. To assign a symbol to each of the four solutions, let us say that \( R_{pq}^{(1,0)}(s) \) vanishes for \( p \leq 0 \), \( R_{pq}^{(-1,0)}(s) \) vanishes for \( p \geq 0 \), \( R_{pq}^{(0,1)}(s) \) vanishes for \( q \leq 0 \), and \( R_{pq}^{(0,-1)}(s) \) vanishes for \( q \geq 0 \). All four of these solutions can be computed by means of a two component recursion of the type (3.13) in [11]. Since we assumed \( \beta \) to be less than one, it follows that \( R^{(1,0)}(s) \) and \( R^{(-1,0)}(s) \) decay exponentially of order \( \beta \) along any line in the lattice with slope 1 or -1, and that \( R^{(0,1)}(s) \) as
well as \( R^{(0,-1)}(s) \) grow exponentially of order \( \beta^{-1} \) along the four lines with slope 1 or \(-1\) through the points \((0,1)\) or \((0,-1)\) in exactly one direction. Taking the arithmetic mean of \( R^{(1,0)}(z) \) and \( R^{(-1,0)}(z) \), both suitably scaled, yields a solution \( \{d_{pq}(z)\} \) of (2.1)* which has the following properties:

\[
d_{pq}(z) = d_{|p||q|}(z), \text{ for all } p, q \in \mathbb{Z}; \quad d_{1,0}(z) = \frac{1}{2}; \quad d_{pq}(z) = 0 \text{ for } |q| \geq |p|. \tag{2.2}
\]

\[
\lim_{|p| \to \infty} \beta^{-|p|}|d_{p,k+p}(z)| < \infty, \quad \lim_{|p| \to \infty} \beta^{-|p|}|d_{p,k-p}(z)| < \infty, \text{ for all } k \in \mathbb{Z}. \tag{2.3}
\]

Returning to our objective, it now follows that the double sequence \( \phi_{pq}(z) = c_{pq}(z) - d_{pq}(z) \) solves system (2.1) and has the following additional properties:

\[
\{\phi_{pq}(z)\} \text{ decays two-sided exponentially along any line in the lattice with slope 1 or } -1 \quad (2.4)
\]

\[
\text{If } c_{00}(z) = c_{01}(z) = 0, \quad \text{ then } \phi_{pq}(z) = 0 \text{ for } |p|, |q| \leq 1. \tag{2.5}
\]

We are going to shelve this for a while, and turn to the discussion of a certain subalgebra of the \( C^\ast \)-algebra \( A_\alpha \). For two elements \( a, b \in A_\alpha \) we denote by \( \text{alg}^\ast(a, b) \) the \( \ast \)-algebra generated by \( a \) and \( b \). Next, we define two distinguished elements.

\[
U = \beta^{-\frac{1}{2}}u + \beta^\frac{1}{2}v, \quad V = w_{-1,1},
\]

which satisfy the relations

\[
UU = \lambda^{-2}VV, \quad U^\ast V = \lambda^2VU^\ast, \text{ where } \lambda = e^{\pi \alpha i}. \tag{2.6}
\]

One way to look at these two elements is that, while mimicking the generators for the rotation algebra \( A_\alpha \), they also allow for the representation of the element \( h(\beta) \) in a form that resembles representing the degenerate element \( h(0) \) in terms of \( u \) and \( v \),

\[
h(\beta) = U + U^\ast.
\]

The next step is to complete the mimicry, rendering \( \text{alg}^\ast(U, V) \) as much as possible a look-alike of \( \text{alg}^\ast(u, v) \). What’s missing from the picture is a basic
symmetry, a conjugate linear involutive automorphism that assigns to one of the generators its adjoint, while fixing the other one. Such a symmetry is readily available for \( \text{alg}^*(u, v) \),

\[
\sigma(u) = u^*, \sigma(v) = v.
\] (2.7)

Reversal of the roles of \( u \) and \( v \) leads to another symmetry, which is equivalent to the one just defined. But due to the asymmetric nature of the elements \( U \) and \( V \), only one of them survives the mimicry. To obtain such a symmetry for \( \text{alg}^*(U, V) \), we first introduce an automorphism of \( A_\alpha \) that appeared for the first time in [12],

\[
\rho_\beta(u) = vuv(uv + \beta)^{-1}(v^*u^* + \beta), \quad \rho_\beta(v) = v(uv + \beta)^{-1}(v^*u^* + \beta).
\] (2.8)

Note that, due to the general properties of \( A_\alpha \), it is quite easy to define automorphisms of \( A_\alpha \). All one has to do is to assign unitary elements in \( A_\alpha \) to the two generators which preserve the fundamental commutation relation (1.1). Any assignment of this kind extends automatically to an automorphism. Since it can be shown that \( U \) is a generator for \( A_\alpha \), in other words the set of all (non-commutative) polynomials in \( U \) and \( U^* \) is norm dense in \( A_\alpha \), the automorphism \( \rho_\beta \) is uniquely determined by the identity,

\[
\rho_\beta(u + \beta v) = u^* + \beta v.
\] (2.9)

Thus, composition of \( \sigma \) and \( \rho_\beta \),

\[
\sigma_\beta = \sigma \circ \rho_\beta,
\]

yields a conjugate linear automorphism that is uniquely determined by the assignments

\[
\sigma_\beta(u) = vu^*v(u^*v + \beta)^{-1}(v^*u + \beta), \quad \sigma_\beta(v) = v(u^*v + \beta)^{-1}(v^*u + \beta).
\] (2.10)

The restriction of this symmetry to the algebra \( \text{alg}^*(U, V) \) is exactly what we need,

\[
\sigma_\beta(U) = U, \quad \sigma_\beta(V) = V^*.
\] (2.11)

While the first relation is obvious, the second one can be checked through straightforward manipulations. Our next objective is to show that this
symmetry, when evaluated at the standardized monomials, yields elements whose expansion in the standardized monomials have desirable exponential decay properties. First, we observe that

\[(u^*v + \beta)^{-1} = \sum_{n=0}^{\infty} (-\beta)^n (v^*u)^{n+1},\]

and taking the adjoint on both sides yields of course a similar expansion. This shows that \(\sigma_{\beta}(w_{pq})\) is a product of \(v^{p+q}\) and an element that has a power series expansion in the monomial \(w_{1-1}\) whose radius of convergence is equal to \(\beta^{-1}\). In conclusion, we obtain the following representation,

\[\sigma_{\beta}(w_{pq}) = \sum_{m \in r(p,q)} m w_{m,p+q-m}, \text{ where } \lim_{|m| \to \infty} |r^{(p,q)}_m|^{1/|m|} \leq \beta. \quad (2.12)\]

In preparation for the proof of Proposition 1.3 in the next section we introduce two linear functional \(\varphi_z\) and \(\varphi_z \cdot \sigma_{\beta}\), which are defined for \(z \in \mathbb{R} \setminus Sp(h(\beta))\) on the algebra \(alg^*(u,v)\) by the assignments

\[\varphi_z(w_{pq}) = R^{(1,0)}_{pq}(z), \quad \varphi_z \cdot \sigma_{\beta}(w_{pq}) = \sum_{m \in r(p,q)} r^{(p,q)}_m \varphi_z(w_{m,p+q-m}). \quad (2.14)\]

Notice that, by (2.4) and (2.12) the terms in the sum on the right-hand side of the second formula decay exponentially of an order less than or equal to \(\beta\). The second formula defines essentially the composition of the first functional with the symmetry \(\sigma_{\beta}\). Since \(\sigma_{\beta}\) is conjugate linear, we have to conjugate the terms in the sum in order to render the resulting functional linear. We are now going to show that the two functionals are actually equal.

\[\varphi_z = \varphi_z \cdot \sigma_{\beta}. \quad (2.13)\]

To see this, we need to return briefly to the settings at the beginning of the present section. First, since \(\sigma_{\beta}(h(\beta)) = h(\beta)\), we also have \(\sigma_{\beta}((h(\beta) - z)^{-1}) = (h(\beta) - z)^{-1}\). Note that, by our assumption, \(z\) is a real number. This means that all we need to show is, that the functionals \(\vartheta_z\) and \(\vartheta_z \cdot \sigma_{\beta}\), which are defined below, are equal.

\[\vartheta_z(w_{pq}) = R^{(1,0)}_{pq}(z), \quad \vartheta_z \cdot \sigma_{\beta}(w_{pq}) = \sum_{m \in r(p,q)} r^{(p,q)}_m \vartheta_z(w_{m,p+q-m}). \quad (2.14)\]
and that a similar statement holds for $R^{(-1,0)}(z)$. Again, since $\sigma_\beta(h(\beta)) = h(\beta)$, and since $R^{(1,0)}(z)$ solves the system (2.1)*, the second functional solves the system (2.1)* as well. However, since by the representation in (2.12) and the vanishing properties of $R^{(1,0)}(z)$, $\vartheta_z \circ \sigma_\beta(w_{pq})$ vanishes for all indices located below or on the line $p = -q$ in the two dimensional lattice, the double sequence $\{\vartheta_z \circ \sigma_\beta(w_{pq})\}$ must be a linear combination of $R^{(1,0)}(z)$ and $R^{(0,1)}(z)$. Since $R^{(0,1)}(z)$ grows exponentially of order $\beta^{-1}$ along the two lines with slope 1 or $-1$ through the point (0,1), while $\vartheta_z \circ \sigma_\beta(w_{pq})$ and $R^{(1,0)}_{pq}(z)$ vanish or decay exponentially along these two lines, it follows that $\{\vartheta_z \circ \sigma_\beta(w_{pq})\}$ is just a scalar multiple of $R^{(1,0)}(z)$. But since $\sigma_\beta$ is unital, the two double sequences must actually be equal. A similar statement can now be obtained for $R^{(-1,0)}(z)$ along a similar line of reasoning. This concludes our argument establishing the validity of (2.13).

Remark. Restricting the solutions of the system (2.1)*, viewed as linear functionals on the linear space $\text{alg}^*(u, v)$, to the subspace $\text{alg}^*(U, V)$, reduces the linear dimension from 6 to 5.

Section 3

We turn now to the proof of Proposition 1.3. Henceforth, we shall simply write $\varphi$ for the functional $\varphi_z$ introduced in Section 2, because we shall assume that $z$ is a fixed real number in the resolvent set of $h(\beta)$. The idea of the proof is to exploit the decay conditions of $\varphi(w_{pq})$ along lines with slope $-1$ in the two dimensional lattice, to construct a functional $\psi$ on $\text{alg}^*(u, v)$, with the property that $\varphi$ can be recovered from $\psi$ by the identity $\psi(aU) = \varphi(a)$ for all $a \in \text{alg}^*(u, v)$, and then showing that such a functional can not exist, in case the function $L$ in Section 1 has a critical point. Of course, there are infinitely many functionals with this property. One simply has to implement a separate elementary recursion along every single line with slope $-1$. The crux is to impose constraints which restrict the availability of such functionals severely. More specifically, we shall prove the following.

Lemma 3.1. There exists a linear functional $\psi$ on $\text{alg}^*(u, v)$, having the properties,

(i)\forall a \in \text{alg}^*(u, v) : \psi(aU) = \varphi(a).
(ii)\forall p \in \mathbb{N}_0, \forall q \in \mathbb{Z} : \psi(U^pV^q) = \overline{\psi(U^pV^{-q})}, \psi((U^*)^pV^q) = \overline{\psi((U^*)^pV^{-q})}.
(iii)\forall a \in \text{alg}^*(u, v) : \psi((h(\beta) - z)a) = \psi(a(h(\beta) - z)U) = 0.
Proof. First, we extend slightly the settings of Section 2. Let \( \mathcal{B} \) be the set of elements \( a \) in \( A_\alpha \) which can be written in the form

\[
a = \sum_{n=-N}^{N} \sum_{m \in \mathbb{R}} k_{m,n-m} w_{m,n-m}, \text{ where } \lim_{|m| \to \infty} \left| k_{m,n-m} \right| \frac{1}{|m|} < \infty
\]

for \( -N \leq n \leq N \).

This set is an involutive subalgebra of \( A_\alpha \). Also, (2.12) implies

\[
\sigma_\beta(\mathcal{B}) = \mathcal{B}.
\]

Furthermore,

\[
\mathbb{U}^{-1} = \beta^{-\frac{1}{2}} \sum_{n=0}^{\infty} (-\beta)^n (u^*)^n u^* \epsilon \mathcal{B}.
\]

Now (2.3) allows us to extend the definition of the functional \( \varphi \) to elements of the form (3.1) as follows:

\[
\varphi(a) = \sum_{n=-N}^{N} \sum_{m \in \mathbb{R}} k_{m,n-m} \varphi(w_{m,n-m}).
\]

Obviously, this extended functional is also linear. Moreover, by (2.13)

\[
\varphi(\sigma_\beta(a)) = \overline{\varphi(a)}, \quad a \in \mathcal{B},
\]

and since the double sequence \( \{ \phi_{pq}(z) \} \) in Section 2 solves the system (2.1),

\[
\varphi((h(\beta) - z)a) = \varphi(a(h(\beta) - z)) = 0, \quad a \in \mathcal{B}.
\]

We are now going to define the functional \( \psi \) as follows,

\[
\psi(a) = \varphi(a \mathbb{U}^{-1}), \quad a \in \mathcal{B}.
\]

We need to check that \( \psi \) has the claimed properties. By construction this is obvious for (i). Next, if \( p \) is a non-negative integer, and \( q \) is an arbitrary
integer, then (2.11) and (3.3) yield,
\[ \psi(U_p V_q) = \varphi(U_p V_q U^{-1}) = \frac{\varphi(\sigma_\beta(U_p V_q U^{-1}))}{\psi(U_p V^{-q})}, \]
which establishes the first identity in (ii). The second identity can be shown in the same way. Finally, (3.4) yields,
\[ \psi((h(\beta) - z)a) = \varphi((h(\beta) - z)a U^{-1}) = 0, \forall a \in \mathfrak{B}, \]
and also,
\[ \psi(a(h(\beta) - z)U) = \varphi(a(h(\beta) - z)U U^{-1}) = \varphi(a(h(\beta) - z)) = 0, \]
which establishes (iii) as well. \( \square \)

Remark. Tracking the significance of the decay condition (2.4) through the discussion so far, one observes that this condition is far stronger than what is needed to make the arguments work. It would be enough to assume that the double sequence in (2.4) does not increase exponentially of an order larger than or equal to \( \beta^{-1} \) along any line with slope \(-1\) in the two-dimensional lattice. All one has to do is to impose a stronger exponential decay condition on the elements in the algebra \( \mathfrak{B} \).

In the proof of the following lemma, we shall be using nothing but the relation (2.6), as well as the following:
\[ U^* U = \lambda V + \lambda^{-1} V^* + \gamma e, \text{ where } \gamma = \beta + \beta^{-1} \quad (3.5) \]

Lemma 3.2. If \( \psi \) is a linear functional defined on \( \text{alg}^*(U, V) \) such that for some non-zero real number \( t \) the following two conditions hold,
\[ (i) \forall a \in \text{alg}^*(U, V) : \psi((U + U^* - te)a) = 0, \psi(a(U + U^* - te)U) = 0, \]
\[ (ii) \psi(U) = \psi(U^*) = \psi(V) = \psi(V^*) = \psi(e) = 0, \]
then \( \psi(UV) = 0. \)
Proof. Since
\[
(U + U^* - te)UV^q = U^2V^q + U^*UV^q - tUV^q,
\]
\[
V^q(U^2 + U^*U - tU) = \lambda^4qU^2V^q + U^*UV^q - \lambda^2q tUV^q,
\]
(i) yields
\[
\psi(U^*UV^q) = t\psi(UV^q) - \psi(U^2V^q),
\]
\[
\psi(U^*UV^q) = t\lambda^2q\psi(UV^q) - \lambda^4q\psi(U^2V^q),
\]
hence,
\[
t\psi(UV^q) = (1 + \lambda^2q)\psi(U^2V^q), \quad \text{for } q \neq 0. \tag{3.6}
\]
Next, we derive two elementary identities involving \( \psi(UV) \) and \( \psi(UV^2) \). Employing the second identity in (i) with \( a = VU^* \), yields
\[
\psi(\psi(UV^2 + U^*U - tU)) = 0,
\]
hence
\[
\psi(V(U^*U)) + \psi(V(U^*U)) - t\psi(V(U^*U)) = 0,
\]
which by virtue of (3.5) yields
\[
\psi(V(\lambda V + \lambda^{-1}V^* + \gamma e)U) + \psi(V(U^*(\lambda V + \lambda^{-1}V^* + \gamma e)))
\]
\[
- t\psi(V(\lambda V + \lambda^{-1}V^* + \gamma e)) = 0,
\]
and so by (ii),
\[
\lambda\psi(V^2U) + \gamma\psi(VU) + \gamma\psi(VU^*V) + \gamma\psi(VU^*) - t\lambda\psi(V^*) = 0,
\]
which finally implies by (2.6)
\[
\lambda^5\psi(UV^2) + \gamma\lambda^2\psi(UV) + \lambda\psi(U^*V^2) + \gamma\lambda^{-2}\psi(U^*V) - t\lambda\psi(V^2) = 0.
\]
Combining this with
\[
\psi((U + U^*)V) = t\psi(V) = 0,
\]
which follows from the first identity in (i) for \( a = V \), and from (ii), we conclude
\[
\lambda^5\psi(UV^2) + \gamma(\lambda^2 - \lambda^{-2})\psi(UV) + \lambda\psi(U^*V^2) - t\lambda\psi(V^2) = 0.
\]
Combining this with
\[ \psi((U + U^*)V^2) = t\psi(V^2), \]
which is true by the first identity in (i) for \( a = V^2 \), yields
\[ \lambda^5\psi(\mathbb{U}\mathbb{V}^2) + \gamma(\lambda^2 - \lambda^{-2})\psi(\mathbb{U}\mathbb{V}) + \lambda(t\psi(V^2) - \psi(U^2\mathbb{V})) - t\lambda\psi(V^2) = 0. \] (3.7)

Next, since
\[ \psi((U + U^* - te)U^2V) = 0, \]
which holds by the first identity in (i) for \( a = \mathbb{U}\mathbb{V} \), properties (3.5) and (ii) yield
\[ \psi(\mathbb{U}^2V) + \lambda\psi(V^2) - t\psi(\mathbb{U}\mathbb{V}) = 0. \]
Combining this with (3.6) for \( q = 1 \) yields,
\[ \frac{t}{1 + \lambda^2}\psi(\mathbb{U}\mathbb{V}) + \lambda\psi(V^2) - t\psi(\mathbb{U}\mathbb{V}) = 0, \]
which in turn simplifies to
\[ \psi(V^2) = \frac{\lambda t}{1 + \lambda^2}\psi(\mathbb{U}\mathbb{V}). \]
Combining this with (3.7) yields
\[ \lambda^5\psi(\mathbb{U}\mathbb{V}^2) + \gamma(\lambda^2 - \lambda^{-2})\psi(\mathbb{U}\mathbb{V}) + \frac{\lambda^2 t^2}{1 + \lambda^2}\psi(\mathbb{U}\mathbb{V}) - \lambda\psi(\mathbb{U}\mathbb{V}^2) \]
\[ - \frac{\lambda^2 t^2}{1 + \lambda^2}\psi(\mathbb{U}\mathbb{V}) = 0, \]
which simplifies to
\[ (\lambda^5 - \lambda)\psi(\mathbb{U}\mathbb{V}^2) + \gamma(\lambda^2 - \lambda^{-2})\psi(\mathbb{U}\mathbb{V}) = 0. \] (3.8)

Next, starting over again, using the first identity in (i) with \( a = \mathbb{U}^2\mathbb{V} \),
\[ \psi((U + U^* - te)U^2\mathbb{V}) = 0, \]
or equivalently,
\[ \psi(\mathbb{U}^3\mathbb{V}) + \psi((U^*)U\mathbb{V}) - t\psi(\mathbb{U}^2\mathbb{V}) = 0, \]
yielding together with (3.5),
\[ \psi(U^3V) + \psi((\lambda V + \lambda^{-1}V^* + \gamma e)UV) - t\psi(U^2V) = 0, \]
which by (ii) implies,
\[ \psi(U^3V) + \lambda^3\psi(UV^2) + \gamma\psi(UV) - t\psi(U^2V) = 0. \]
Employing (3.6) to this for \( q = 1 \), we obtain,
\[ \psi(U^3V) + \lambda^3\psi(UV^2) + \left(\gamma - \frac{t^2}{1+\lambda^2}\right)\psi(UV) = 0. \]  
(3.9)
On the other hand, the second identity in (i) with \( a = UV \) yields,
\[ \psi(UV(U^2 + U^*U - tU)) = 0, \]
or equivalently
\[ \psi(UVU^2) + \psi(UV(U^*U)) - t\psi(UVU) = 0, \]
hence by (3.5),
\[ \psi(UVU^2) + \psi(UV(U^*U)) - t\psi(UVU) = 0, \]
which by (ii) and (2.6) yields,
\[ \lambda^4\psi(U^3V) + \lambda\psi(UV^2) + \gamma\psi(UV) - t\lambda^2\psi(U^2V) = 0, \]
and after invoking (3.6) for \( q = 1 \) again,
\[ \lambda^4\psi(U^3V) + \lambda\psi(UV^2) + \left(\gamma - \frac{\lambda^2t^2}{1+\lambda^2}\right)\psi(UV) = 0. \]
Finally, multiplying this by \( \lambda^{-4} \) and subtracting it from (3.9) we obtain,
\[ (\lambda^3 - \lambda^{-3})\psi(UV^2) + \left[(1 - \lambda^{-4})\gamma - \frac{1-\lambda^{-2}}{1+\lambda^2 t^2}\right]\psi(UV) = 0. \]  
(3.10)
Now suppose that \( \psi(UV) \neq 0 \). Then comparison of (3.8) with (3.10) yields,
\[ (\lambda^3 - \lambda^{-3})(\lambda^2 - \lambda^{-2})\gamma = (\lambda^5 - \lambda) \left[(1 - \lambda^{-4})\gamma - \frac{1-\lambda^{-2}}{1+\lambda^2 t^2}\right], \]
which turns into,
\[(\lambda^{-6} - \lambda^{-4} - \lambda^{-2} + 1)\gamma = (2 - \lambda^2 - \lambda^{-2})t^2.\]

Since the number on the right-hand side of this equation is real, it follows that \(\lambda^{-6} - \lambda^{-4} - \lambda^{-2}\) must be real as well or equivalently
\[\lambda^{-6} - \lambda^{-4} - \lambda^{-2} = \lambda^6 - \lambda^4 - \lambda^2.\]

This in turn implies that
\[\lambda^{12} - \lambda^{10} - \lambda^8 + \lambda^4 + \lambda^2 - 1 = 0.\]

This conflicts with the fact that, \(\alpha\) being irrational, the set \(\{\lambda^{2n}/n\in\mathbb{Z}\}\) is dense in the unit circle. Therefore, \(\psi(UV) = 0\), as claimed. \(\square\)

**Remark.** Performing the kind of manipulations in the proof of Lemma 2.2 for more general terms of the form \(U^pV^q\) and \((U^*p)V^q\) one can actually show that the functional \(\psi\) “almost” vanishes. If the condition \(\psi(U^2) = 0\) is added, then \(\psi\) vanishes completely.

**Proof of Proposition 1.3.** First note that the functional \(\varphi\) cannot be zero. This is true because the double sequence \(\{c_{pq}(z)\}\) decays exponentially as \(|p| \to \infty\) and \(|q| \to \infty\), while this is not true for \(\{d_{pq}(z)\}\). If \(\{d_{pq}(z)\}\) were decaying exponentially as \(|p| \to \infty\) and \(|q| \to \infty\), then the two solutions \(R_1,0(z)\) and \(R_{-1,0}(z)\) of the system (2.1)* would give rise to two distinct inverses of the same element \(h(\beta) - z\), which is of course impossible. Now suppose that
\[(\tau((h(\beta) - z)^{-1}) = \tau((h(\beta) - z)^{-1}v)) = 0.\]

It follows from (2.5),
\[\varphi(u) = \varphi(u^*) = \varphi(v) = \varphi(v^*) = 0\]

Since the double sequence \(\{\varphi(w_{pq})\}\) solves the system (2.1), this implies that \(\varphi\) is non-zero if and only if
\[\varphi(V) \neq 0.\]

The vanishing conditions for \(\varphi\) translate into several vanishing conditions for \(\psi\). First, by Lemma 2.1,
\[\psi(U) = \varphi(e) = 0.\]
Furthermore,

\[ \beta^{-\frac{1}{2}} \psi(e) + \beta^{\frac{1}{2}} \psi(u^*v) = \psi(u^*U) = \varphi(u^*) = 0, \]
\[ \beta^{-\frac{1}{2}} \psi(v^*u) + \beta^{\frac{1}{2}} \psi(e) = \psi(v^*U) = \varphi(v^*) = 0, \]

or, equivalently,

\[ \beta^{-\frac{1}{2}} \psi(e) + \beta^{\frac{1}{2}} \lambda^{-1} \psi(V^*) = 0, \]
\[ \beta^{-\frac{1}{2}} \lambda \psi(V) + \beta^{\frac{1}{2}} \psi(e) = 0. \]

By Lemma 2.1(ii),

\[ \psi(e) = \overline{\psi(e)}, \psi(V^*) = \overline{\psi(V)}. \]

This yields a linear system for \( \psi(e) \) and \( \psi(V) \),

\[ \beta^{-\frac{1}{2}} \psi(e) + \beta^{\frac{1}{2}} \lambda \psi(V) = 0, \]
\[ \beta^{\frac{1}{2}} \psi(e) + \beta^{-\frac{1}{2}} \lambda \psi(V) = 0. \]

Since the determinant of this system,

\[ \begin{vmatrix} \beta^{-\frac{1}{2}} & \beta^{\frac{1}{2}} \lambda \\ \beta^{\frac{1}{2}} & \beta^{-\frac{1}{2}} \lambda \end{vmatrix} = \lambda(\beta^{-1} - \beta) \]

is non-zero for \( \beta \neq 1 \), it follows that

\[ \psi(e) = \psi(V) = \psi(V^*) = 0. \]

Finally, by Lemma 3.1(iii), since \( \psi(e) = 0 \), and \( \psi(U) = \varphi(e) = 0 \)

\[ \psi(U^*) = \psi(U + U^* - ze) = 0. \]

In conclusion we have shown, that \( \psi \) (more precisely its restriction to \( \text{alg}^*(U, V) \)) satisfies the conditions (i) and (ii) of Lemma 2.2, letting \( t = z \beta^{-\frac{1}{2}} \). Therefore,

\[ \varphi(V) = \psi(VU) = \lambda^2 \psi(UV) = 0. \]

Since we observed at the beginning of the proof that \( \varphi(V) \) cannot be zero, we have reached a contradiction.
References


