DISCRETE BETHE–SOMMERFELD CONJECTURE FOR
TRIANGULAR, SQUARE, AND HEXAGONAL LATTICES

By
JAKE FILLMAN AND RUI HAN

Abstract. We prove the discrete Bethe–Sommerfeld conjecture on the
graphene lattice, on its dual lattice (the triangular lattice), and on the extended
Harper lattice. For each of these lattice geometries, we analyze the behavior of
small periodic potentials. In particular, we provide sharp bounds on the number
of gaps that may perturbatively open, we describe sharp arithmetic criteria on the
periods that ensure that no gaps open, and we characterize those energies at which
gaps may open in the perturbative regime. In all three cases, we provide examples
that open the maximal number of gaps and estimate the scaling behavior of the gap
lengths as the coupling constant goes to zero.

Contents

1 Introduction 271
2 Floquet theory for periodic Schrödinger operators on graphs 278
3 Triangular Laplacian 280
4 Hexagonal Laplacian 295
5 Square Laplacian with next-nearest-neighbor interactions 305

1 Introduction

The Bethe–Sommerfeld conjecture asserts that: for any \( d \geq 2 \) and any periodic
function \( V : \mathbb{R}^d \rightarrow \mathbb{R} \), the spectrum of the Schrödinger operator

\[
L_V := -\nabla^2 + V
\]

has only finitely many gaps. This was studied by many people with important ad-
vances in [24, 26, 35, 37, 38, 39, 41], and culminating in the paper of Parnovski [32].
Another way to think about the (continuum) Bethe–Sommerfeld conjecture is that the interval spectrum of the free Laplacian is preserved in the regime for which the periodic potential is relatively small compared to the Laplacian, and this precisely happens in the high-energy region. Since discrete Schrödinger operators are bounded, the appropriate analogy to the high-energy region is the region of small $V$. Note that in the discrete setting, the number of gaps is always finite; if a potential is $(p_1, p_2)$-periodic, then the spectrum consists of $P = p_1p_2$ bands and hence has at most $P - 1$ gaps. Thus, the questions are: if the number of gaps is much smaller than $P - 1$, and further, the exact number of possible gaps, the locations at which gaps may open, and the size of the gaps when they do open. These questions have been answered for the square lattice $\mathbb{Z}^d$ in recent papers [28, 12, 21]. Motivated by prominent physical models, most notably graphene, the aim of the present work is to prove the discrete Bethe–Sommerfeld conjecture on lattices of relevance to physical investigations: the hexagonal lattice, its dual lattice (the triangular lattice), and the square lattice with next-nearest-neighbor interactions (which arises in the extended Harper’s model).

1.1 Main results. Let us now describe more precisely the setting in which we work and the results that we prove. By a graph, we shall mean a pair $\Gamma = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is a nonempty set and $\mathcal{E}$ is a nonempty subset of $\mathcal{V} \times \mathcal{V}$ with the following properties:

- For no $v \in \mathcal{V}$ does one have $(v, v) \in \mathcal{E}$.
- If $(u, v) \in \mathcal{E}$, then $(v, u) \in \mathcal{E}$.

If $(u, v) \in \mathcal{E}$, we write $u \sim v$ and we say that $u$ and $v$ are neighbors or neighboring vertices. We think of $\mathcal{E}$ as the set of directed edges; $(u, v)$ represents the edge that originates at $u$ and terminates at $v$.

Given such a graph, we consider $\mathcal{H}_\Gamma = \ell^2(\mathcal{V})$ and the associated graph Laplacian$^1$ $\Delta_\Gamma : \mathcal{H}_\Gamma \to \mathcal{H}_\Gamma$, which acts via

$$[\Delta_\Gamma \psi]_u = \sum_{v \sim u} \psi_v, \quad u \in \mathcal{V}, \quad \psi \in \mathcal{H}_\Gamma.$$ 

By a Schrödinger operator on $\Gamma$, we mean an operator of the form

$$H_Q = H_{\Gamma, Q} = \Delta_\Gamma + Q,$$

$^1$Technically, this is the adjacency operator of the graph. Other authors use $\psi_v - \psi_u$ where we have only $\psi_v$. Our convention is slightly more natural for the setting in which we wish to work. Concretely, all of the graphs that we consider in the present work have uniform degree (all vertices in a given graph have the same number of incident edges), and hence leaving off the $-\psi_u$ term merely costs us a multiple of the identity operator, and it simplifies the appearance of a few calculations.
where $Q : \mathcal{V} \rightarrow \mathbb{R}$ is a bounded function that acts on $\mathcal{H}_\Gamma$ by multiplication:

$$[Q\psi]_u = Q(u)\psi_u, \quad u \in \mathcal{V}, \ \psi \in \mathcal{H}_\Gamma.$$
Figure 2. A portion of the hexagonal lattice. A fundamental domain is highlighted in red (color figure available online).

**Theorem 1.1** (Bethe–Sommerfeld for the hexagonal lattice). For all \( p = (p_1, p_2) \) in \( \mathbb{Z}^2_+ \), there is a constant \( c = c_p > 0 \) such that, if \( Q : \mathcal{V}_{\text{hex}} \to \mathbb{R} \) is \( p \)-periodic and \( \|Q\|_{\infty} \leq c \), the following statements hold true for \( H_Q = \Delta_{\text{hex}} + Q \):

1. \( \sigma(H_Q) \) consists of no more than four intervals.
2. If at least one of \( p_1 \) or \( p_2 \) is odd, then \( \sigma(H_Q) \) consists of no more than two intervals.

Moreover, gaps may only open at 0 and \( \pm 1 \) in the first case, and only at zero in the second case.

Furthermore, this theorem is sharp in the following sense: there exists a \((1, 1)\)-periodic potential \( Q_1 \) which infinitesimally opens a gap at zero, and there is a \((2, 2)\)-periodic potential \( Q_2 \) which infinitesimally opens gaps at \(-1, 0, \) and \( 1 \).

**Theorem 1.2.**

1. There exists \( Q_1 : \mathcal{V}_{\text{hex}} \to \mathbb{R}^2 \) which is \((1, 1)\)-periodic such that \( \sigma(H_{\lambda Q_1}) \) has exactly two connected components for all \( \lambda > 0 \). Furthermore, the gap size at 0 is of order \( \lambda \).

2. There exists \( Q_2 : \mathcal{V}_{\text{hex}} \to \mathbb{R}^2 \) which is \((2, 2)\) periodic such that \( \sigma(H_{\lambda Q_2}) \) has exactly four connected components for any sufficiently small \( \lambda > 0 \). Furthermore, the gap size at 0 is of order \( \lambda \), and the gap sizes at \( \pm 1 \) are of order \( \lambda^2 \).

Let us remark that Theorem 1.2.(1) is well-known; we merely list it for completeness. The example in Theorem 1.2.(2) is novel.
1.3 The triangular lattice. The next graph that we consider is the triangular lattice, which is dual to the hexagonal lattice. The graph has vertices

\[ V_{\text{tri}} = \{ n\mathbf{a}_1 + m\mathbf{a}_2 : n, m \in \mathbb{Z} \}, \]

where \( \mathbf{a}_j \) are as before. One then declares \( v \sim w \) for \( v, w \in V \) if \( \|v - w\| = 1 \). Thus, after identifying \( n\mathbf{a}_1 + m\mathbf{a}_2 \) with the point \( (n, m) \in \mathbb{Z}^2 \), we may view the Laplacian on the triangular lattice as an operator on \( \ell^2(\mathbb{Z}^2) \) via

\[ \Delta_{\text{tri}} \psi_{n,m} = [\Delta_{\text{sq}} \psi]_{n,m} + \psi_{n-1,m+1} + \psi_{n+1,m-1}. \]

This correspondence amounts to shearing and stretching the triangular lattice, and identifies the triangular lattice with the square lattice equipped with additional skewed next-nearest-neighbor interactions. See Figures 3 and 4.

**Theorem 1.3** (Bethe–Sommerfeld for the triangular lattice). For all \( p = (p_1, p_2) \in \mathbb{Z}_+^2 \),

there is a constant \( c = c_p > 0 \) such that, if \( Q : V_{\text{tri}} \to \mathbb{R} \) is \( p \)-periodic and \( \|Q\|_\infty \leq c \), the following hold true for \( H_Q = \Delta_{\text{tri}} + Q \):

1. \( \sigma(H_Q) \) consists of no more than two intervals.
2. If at least one of \( p_1 \) or \( p_2 \) is odd, then \( \sigma(H_Q) \) consists of a single interval.

Moreover, the gap in the first setting may only open at the energy \( E = -2 \).

This theorem is sharp vis-à-vis the number of intervals in the spectrum and the arithmetic restrictions on the periods. Concretely, we exhibit a \( (2, 2) \)-periodic potential that perturbatively opens a gap at \(-2\).

**Theorem 1.4.** There exists \( Q : V_{\text{tri}} \to \mathbb{R} \) which is \( (2, 2) \)-periodic, such that \( \sigma(H_{\lambda Q}) \) has exactly two connected components for any sufficiently small \( \lambda > 0 \). Furthermore, the gap size at \(-2\) is of order \( \lambda \).
1.4 The EHM lattice. In addition to the hexagonal and triangular lattices, we also study the square lattice with next-nearest-neighbor interactions, which is motivated by the extended Harper’s model (EHM). The EHM was proposed by Thouless [40] and has also led to a lot of study in mathematics and physics [1, 18, 19, 20, 22, 25]; it corresponds to an electron in a square lattice that interacts not only with its nearest neighbors but also its next-nearest-neighbors. In the following, we will refer to a square lattice with next-nearest-neighbor interactions as the EHM lattice, in order to distinguish it from the standard square lattice.

The EHM lattice also has vertex set $V_{sqn} = \mathbb{Z}^2$. However, now, one connects $n$ and $n'$ if and only if they are nearest neighbors or next-nearest neighbors in the square lattice. Equivalently, one declares

$$n \sim n' \iff \|n - n'\|_\infty = 1.$$ 

The associated Laplacian acts on $\ell^2(\mathbb{Z}^2)$ via

$$[\Delta_{sqn}\psi]_{n,m} = [\Delta_{sq}]_{n,m} + \psi_{n-1,m-1} + \psi_{n-1,m+1} + \psi_{n+1,m-1} + \psi_{n+1,m+1}.$$ 

See Figure 5.

**Theorem 1.5** (Bethe–Sommerfeld for the EHM lattice). For all $p = (p_1, p_2) \in \mathbb{Z}_+^2$, there is a constant $c = c_p > 0$ such that, if $Q : V_{sqn} \to \mathbb{R}$ is $p$-periodic and $\|Q\|_\infty \leq c$, the following hold true for $H_Q = \Delta_{sqn} + Q$:

1. $\sigma(H_Q)$ consists of no more than two intervals.
2. If at least one of $p_1$ or $p_2$ is not divisible by three, then $\sigma(H_Q)$ consists of a single interval.

Moreover, the gap in the first setting may only open at the energy $E = -1$. 
This theorem is also sharp:

**Theorem 1.6.** There exists $Q : \mathbb{Z}^2 \to \mathbb{R}$ which is $(3, 3)$-periodic such that $\sigma(H_{1Q})$ has exactly two connected components for any sufficiently small $\lambda > 0$. Furthermore, the gap size at $-1$ is of order $\lambda$.

It is also reasonable to consider an EHM lattice in which the diagonal hopping terms are different from the cardinal direction hopping terms; for example, one might consider the one-parameter family

$$[H_\mu \psi]_{n,m} = [\Delta_{sq}]_{n,m} + \mu(\psi_{n-1,m-1} + \psi_{n-1,m+1} + \psi_{n+1,m-1} + \psi_{n+1,m+1})$$

with $\mu \in [0, \infty)$; $\mu = 0$ yields the square Laplacian and $\mu = 1$ yields $\Delta_{sqn}$. The authors plan to address this in a forthcoming work [16].

1.5 Further remarks. Let us mention a closely related work [23]. In [23], Helffer, Kerdelhué and Royo-Letelier developed a Chambers analysis for magnetic Laplacians on the hexagonal lattice (and its dual, the triangular lattice) with rational flux. They showed that for a non-trivial rational flux $p/q \notin \mathbb{Z}$, the magnetic Laplacians on hexagonal and triangular lattices have non-overlapping (possibly touching) bands. This recovers a similar feature of the square lattice [4]. However, unlike the square lattice, which has no touching bands except at the center for $q$ even [30], they were able to give an explicit example of non-trivial touching bands for hexagonal and triangular lattices. Indeed they showed that the triangular Laplacian has touching bands at energy $E = -\sqrt{3}$ for $p/q = 1/6$, and the hexagonal Laplacian has touching bands at energies $E = \pm \sqrt{3}$ and 0 for $p/q = 1/2$. Therefore, the underlying geometry is responsible for the formation of non-overlapping bands. But it has remained unclear whether there will be other touching bands for different fluxes (and if any, what are the locations).

In our work we are able to give a sharp criterion of the formation of touching bands for the free Laplacians on these lattices and the EHM lattice. Although the general strategy follows that of [21], there are several challenges to overcome in the present work:

- The Floquet parameters and perturbation directions that we choose in the perturb-and-count technique are strongly model-dependent in a subtle fashion. For example, at non-exceptional energies, we locate Floquet parameters and a perturbation direction in a way such that the Floquet eigenvalues with vanishing linear terms have quadratic terms of the same sign along this direction. At the exceptional energy of the triangular lattice, we choose two directions such that the eigenvalues with vanishing gradients have quadratic
terms of different signs along the two directions; for a more detailed discussion, see Remark 3.4. This is similar to what was done in [21] for the square lattice case. However, for the EHM lattice, any direction will lead to the same number of positive and negative quadratic terms; see Remark 5.5. This issue is resolved by a new construction: we find a direction that moves approximately 2/3 of the Floquet eigenvalues up while the other 1/3 move down. All these constructions depend heavily on the Floquet representation of the eigenvalues, and thus get more difficult as the underlying geometry gets more complicated.

- Applying the perturb-and-count ideas directly to the hexagonal lattice is quite difficult, due to the fact that the Floquet eigenvalues do not have simple expressions; compare (4.3). However, one can relate Laplacians and Floquet matrices for the triangular and hexagonal lattices in a fairly elegant fashion (see [23] and our (4.2)). Thus, we prove the Bethe–Sommerfeld conjecture directly for the triangular lattice and then derive the corresponding statement for the hexagonal lattice via duality.

- Because of the more complicated structure of the lattices involved, constructing potentials that open gaps at the exceptional energies is substantially more difficult than in the square lattice. In particular, we need to construct (2,2)-periodic potentials that live on eight vertices for the hexagonal lattice, and (3,3)-periodic potential for the EHM lattice. In this paper we develop a robust technique to study these finite volume problems in a sharp way. Indeed, we can not only prove that a gap exists, but also estimate its size up to a constant factor (see Theorems 3.5, 4.2, and 5.6). In the case of the triangular lattice, we are even able to use our technique to exactly compute the gap, not only to estimate its size (Theorem 3.5).

The remainder of the paper is organized as follows. Section 2 recalls Floquet theory for \( \mathbb{Z}^2 \)-periodic graphs. We work with the triangular lattice in Section 3, proving Theorems 1.3 and 1.4. We then work with the hexagonal lattice in Section 4, proving Theorems 1.1 and 1.2. Finally, we conclude with the EHM lattice in Section 5, proving Theorems 1.5 and 1.6.

2 Floquet theory for periodic Schrödinger operators on graphs

Let \( \Gamma = (\mathcal{V}, \mathcal{E}) \) be a \( \mathbb{Z}^2 \)-periodic graph with translation symmetries \( \mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2 \), and suppose \( Q : \mathcal{V} \to \mathbb{R} \) is \( \mathbf{p} = (p_1, p_2) \)-periodic, that is,

\[
Q(u + p_j \mathbf{a}_j) = Q(u), \quad u \in \mathcal{V}, \; j = 1, 2.
\]
We will briefly describe Floquet theory for \( H_Q = \Delta_1 + Q \), following [27]. The main purpose of this section is to establish notation, so we do not give any proofs. One may write \( H_Q \) as a constant-fiber direct integral over the fundamental domain. Concretely, let

\[
V_f = V \cap \{ sa_1 + ta_2 : 0 \leq s < p_1, \; 0 \leq t < p_2 \}.
\]

By periodicity, \( |V_f| = P := p_0p_1p_2 \), where

\[
p_0 = |V \cap \{ sa_1 + ta_2 : 0 \leq s < p_1, \; 0 \leq t < 1 \}|.
\]

Here, and throughout the paper, we use \( |S| \) to denote the cardinality of the set \( S \). For each edge \( (u, v) \in E \) there exist unique vertices \( u_f, v_f \in V_f \) and unique integers \( n, m, n', m' \in \mathbb{Z} \) with

\[
u = u_f + np_1a_1 + mp_2a_2, \quad v = v_f + n'p_1a_1 + m'p_2a_2,
\]

We then define the index of \( (u, v) \) by \( \tau(u, v) = (n' - n, m' - m) \). Finally, for \( u, v \in V_f \), we define \( B(u, v) \) to be the set of all translates of \( v \) that connect to \( u \) via an edge of \( \Gamma \):

\[
B(u, v) = \{ w \in V : w \sim u \text{ and } w = v + np_1a_1 + mp_2a_2 \text{ for some } n, m \in \mathbb{Z} \}.
\]

Then, for each \( \theta = (\theta_1, \theta_2) \in \mathbb{R}^2 \), the corresponding Floquet matrix is a self-adjoint operator on \( \mathcal{H}_f := \ell^2(V_f) = \mathbb{C}^{V_f} \) defined by

\[
\langle \delta_u, H_Q(\theta)\delta_v \rangle = \sum_{w \in B(u, v)} \exp(i\langle \tau(u, w), \theta \rangle).
\]

In the event that the sum in (2.1) is empty, \( \langle \delta_u, H_Q(\theta)\delta_v \rangle = 0 \). Clearly, if \( \theta'_j - \theta_j \in 2\pi \mathbb{Z} \) for \( j = 1, 2 \), then

\[
H_Q(\theta) = H_Q(\theta'),
\]

so \( H_Q(\theta) \) descends to a well-defined function of \( \theta \in \mathbb{T}^2 := \mathbb{R}^2/(2\pi \mathbb{Z})^2 \cong [0, 2\pi)^2 \). We will freely use \( \theta \in \mathbb{R}^2 \) or \( \theta \in \mathbb{T}^2 \) depending on which is more convenient in a given setting.

Informally, (2.1) represents the restriction of \( H_Q \) to the discrete torus

\[
(\mathbb{Z}a_1 \oplus \mathbb{Z}a_2)/(p_1\mathbb{Z}a_1 \oplus p_2\mathbb{Z}a_2) \cong \mathbb{Z}_{p_1} \oplus \mathbb{Z}_{p_2}
\]

with the following boundary conditions: wrapping once around the torus in the positive \( a_1 \) direction accrues a phase \( e^{i\theta_1} \) and wrapping around once in the positive \( a_2 \) direction accrues a phase \( e^{i\theta_2} \). More precisely, we may view \( H_Q(\theta) \) in the following
manner. The operator $H_Q$ acts on the space $\mathbb{C}^\mathbb{V}$ of arbitrary (not necessarily square-summable) functions $\mathbb{V} \rightarrow \mathbb{C}$. When $Q$ is $(p_1, p_2)$-periodic, then for each $\theta \in \mathbb{T}^2$, $H_Q$ preserves the subspace

$$\mathcal{H}(\theta) = \{ \psi \in \mathbb{C}^\mathbb{V} : \psi(u + p_j \alpha_j) = e^{i\theta_j} \psi(u) \}.$$ 

Then, $H_Q(\theta)$ is equivalent to the restriction of $H_Q$ to $\mathcal{H}(\theta)$.

For each $\theta$, order the eigenvalues of $H_Q(\theta)$ as

$$E_1(\theta) \leq \cdots \leq E_P(\theta)$$

with each eigenvalue listed according to its multiplicity. Then, for $1 \leq j \leq P$, the $j$-th spectral band of $H_Q$ is defined by

$$F_j = F_j(Q) := \text{ran}(E_j) = \{ E_j(\theta) : \theta \in \mathbb{T}^2 \} = \{ E_j(\theta) : \theta \in \mathbb{R}^2 \}.$$ 

**Theorem 2.1.** With notation as above,

$$\sigma(H_Q) = \bigcup_{\theta \in \mathbb{T}^2} H_Q(\theta) = \bigcup_{j=1}^P F_j.$$

We will use Theorem 2.1 in the following way. Making the dependence on the potential $Q$ explicit, one may write

$$F_j = F_j(Q) = [E_j^-(Q), E_j^+(Q)].$$

The key fact is the following: by standard perturbation theory for self-adjoint operators, $E_j^\pm(Q)$ are 1-Lipschitz functions of $Q$. Here, one views $Q$ as an element of $\mathbb{R}^P$ and the perturbation is measured with respect to the uniform metric thereupon. In particular, if an energy $E$ satisfies $E \in \text{int}(F_j(Q))$, then $(E - \delta, E + \delta) \subseteq F_j(Q)$ for some positive $\delta$, and it follows that $E \in F_j(Q') \subseteq \sigma(H_{Q'})$ for any $(p_1, p_2)$-periodic $Q'$ with $\|Q - Q'\|_\infty < \delta$. Note that here it is very important that one views the periods as fixed: one may only perturb within $\mathbb{R}^P$ for a fixed $P$. Thus, our analysis revolves around determining for a given energy $E$, whether $E$ belongs to the interior of some band of the Laplacian, where the Laplacian is viewed as a degenerate $(p_1, p_2)$-periodic operator.

### 3 Triangular Laplacian

We view the triangular Laplacian as acting on the square lattice $\ell^2(\mathbb{Z}^2)$, but with extra connections as in (1.1):

$$\left[ \Delta_{\text{tri}} u \right]_{n,m} = u_{n-1,m} + u_{n+1,m} + u_{n,m-1} + u_{n,m+1} + u_{n-1,m+1} + u_{n+1,m-1}$$

$$= \left[ \Delta_{\text{sq}} u \right]_{n,m} + u_{n-1,m+1} + u_{n+1,m-1}.$$
Now, given \(p_1, p_2 \in \mathbb{Z}_+\), we view \(\Lambda_m\) as a \(\mathbf{p}\)-periodic operator and perform the Floquet decomposition. Define \(P := p_1p_2\) as in Section 2, and put
\[
\Lambda := \mathbb{Z}^2 \cap ([0, p_1) \times [0, p_2)).
\]
For \(\mathbf{\theta} = (\theta_1, \theta_2) \in \mathbb{R}^2\), it is straightforward to check that
\[
\sigma(H(\mathbf{\theta})) = \{ e_\ell(\mathbf{\theta}) : \ell \in \Lambda \},
\]
where \(\ell = (\ell_1, \ell_2)\) and
\[
e_\ell(\mathbf{\theta}) = 2 \cos \left( \frac{\theta_1 + 2\pi \ell_1}{p_1} \right) + 2 \cos \left( \frac{\theta_2 + 2\pi \ell_2}{p_2} \right) + 2 \cos \left( \frac{\theta_1 + 2\pi \ell_1}{p_1} - \frac{\theta_2 + 2\pi \ell_2}{p_2} \right).
\]
Let us point out that one needs to be somewhat careful at this point; namely, \(e_\ell(\mathbf{\theta})\) is not a well-defined function of \(\mathbf{\theta} \in \mathbb{T}^2\). However, the error incurred in using a different coset representative of \(\mathbf{\theta} \in \mathbb{T}^2\) is simply a change in the index \(\ell\), and one can check that the family \(\{ e_\ell(\mathbf{\theta}) : \ell \in \Lambda \}\) is a well-defined function on \(\mathbb{T}^2\) (as well it should, since the operator \(H(\mathbf{\theta})\) is itself a well-defined function of \(\mathbf{\theta} \in \mathbb{T}^2\)).

In any case, the ambiguity disappears when one considers the covering space \(\mathbb{R}^2\), which we do for most of the paper. One could also use the minimal covering space \(\mathbb{R}^2/(p_1\mathbb{Z} \oplus p_2\mathbb{Z})\) on which the \(e_\ell(\mathbf{\theta})\) are well-defined, but this does not accrue any benefits vis-à-vis the present work, so we simply use \(\mathbb{R}^2\).

As in Section 2, we label these eigenvalues in increasing order according to multiplicity by
\[
E_1(\mathbf{\theta}) \leq E_2(\mathbf{\theta}) \leq \cdots \leq E_P(\mathbf{\theta})
\]
and denote the \(P\) spectral bands by
\[
F_k^\Lambda = \{ E_k^\Lambda(\mathbf{\theta}) : \mathbf{\theta} \in \mathbb{R}^2 \}, \quad 1 \leq k \leq P.
\]
Straightforward computation shows that \(\sigma(\Lambda_m) = [-3, 6)\), and thus
\[
\bigcup_{k=1}^{P} F_k^\Lambda = [-3, 6].
\]
Henceforth, we view \(p_1\) and \(p_2\) as fixed and so we drop \(\Lambda\) from the superscripts. Our main theorem of this section is the following.

**Theorem 3.1.** Let \(p_1, p_2 \in \mathbb{Z}_+\) be given.

1. Each \(E \in (-3, 6) \setminus \{-2\}\) belongs to \(\text{int}(F_k)\) for some \(1 \leq k \leq P\).
2. If one of the periods \(p_1, p_2\) is odd, then \(E = -2\) belongs to \(\text{int}(F_k)\) for some \(1 \leq k \leq P\).

**Proof of Theorem 1.3.** As already discussed, this follows immediately from Theorem 3.1. \(\square\)
3.1 Proof of Theorem 3.1. We will divide the proof into two different cases: \( E \neq -2 \) and \( E = -2 \). Our general strategy is to argue by contradiction. More specifically, we assume \( E = \min F_{k+1} = \max F_k \) for some \( 1 \leq k \leq P - 1 \), and show that this leads to a contradiction. We will use the following two lemmas, whose proofs we provide at the end of the present section.

**Lemma 3.2.** For any \( E \in [-3, 6] \), there exist \( x, y \in [0, 2\pi) \) such that

\[
\begin{align*}
\cos(x) + \cos(y) + \cos(x - y) &= \frac{E}{2}, \\
\sin(x) + \sin(y) &= 0.
\end{align*}
\]

Furthermore, if \( E \in [-3, -2) \), we have

\[
\cos(x) + \cos(y) = -1 \pm \sqrt{E + 3} < 0
\]

for any \( x, y \) that satisfy conditions (3.1) and (3.2), and, if \( E \in (-2, 6] \), then we have

\[
\cos(x) + \cos(y) = -1 + \sqrt{E + 3} > 0
\]

for any \( x, y \) that satisfy conditions (3.1) and (3.2).

**Lemma 3.3.** Consider the following system:

\[
\begin{cases}
\cos(x) + \cos(y) + \cos(x - y) = \frac{E}{2}, \\
\sin(x) + \sin(x - y) = 0, \\
\sin(y) - \sin(x - y) = 0.
\end{cases}
\]

For any \( E \in (-3, 6) \setminus \{-2\} \), the solution set of (3.5) is empty. For \( E = -2 \), the solutions of (3.5) in \([0, 2\pi)^2\) are \((0, \pi), (\pi, 0)\) and \((\pi, \pi)\).

We will use Lemma 3.2 in the \( E \neq -2 \) case, and Lemma 3.3 in the \( E = -2 \) case.

3.1.1 \( E \neq -2 \).

**Proof of Theorem 3.1.** Let \( E \in (-3, 6) \setminus \{-2\} \) be given and suppose for the purpose of establishing a contradiction that \( E = \max F_k = \min F_{k+1} \) for some \( 1 \leq k < P \). Let \((x, y)\) denote a solution to (3.1) and (3.2) from Lemma 3.2, and take

\[
\tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2) \in [0, 2\pi)^2 \quad \text{and} \quad \ell^{(1)} = (\ell_1^{(1)}, \ell_2^{(1)}) \in \Lambda
\]

such that

\[
p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1^{(1)}) = x \quad \text{and} \quad p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2^{(1)}) = y.
\]
It is clear that $\tilde{\theta}$ and $\ell^{(1)}$ are uniquely determined by $x$ and $y$. Let us also note that (3.1) is equivalent to
\[ e_{\ell^{(1)}}(\tilde{\theta}) = E. \]
Define $\Lambda_E(\tilde{\theta}) \subseteq \Lambda$ to be the set of all $\ell \in \Lambda$ such that $e_\ell(\tilde{\theta}) = E$. Then $r := |\Lambda_E(\tilde{\theta})|$ is the multiplicity of $E \in \sigma(H(\tilde{\theta}))$ and clearly $\ell^{(1)} \in \Lambda_E(\tilde{\theta})$.

Since $E \in F_k$ by assumption, let $s \in \mathbb{Z} \cap [1, r]$ be chosen so that
\[ E_{k-s}(\tilde{\theta}) < E_{k-s+1}(\tilde{\theta}) = \cdots = E_k(\tilde{\theta}) = \cdots = E_{k+r-s}(\tilde{\theta}) < E_{k+r-s+1}(\tilde{\theta}). \]
Since all the eigenvalues are continuous in $\theta$, we can take $\varepsilon > 0$ small enough such that
\[ E_{k-s}(\theta) < E_{k-s+1}(\theta) \quad \text{and} \quad E_{k+r-s}(\theta) < E_{k+r-s+1}(\theta) \]
hold whenever $\|\theta - \tilde{\theta}\|_{\mathbb{R}^2} < \varepsilon$. Our goal is to perturb about the point $\tilde{\theta}$ in two directions, one of which is “generic” and one of which is carefully chosen. The generic perturbation moves half of the eigenvalues to the right and half to the left, which we shall use to conclude that $r = 2s$. The non-generic perturbation is carefully chosen to contradict this.

Given $\ell \in \Lambda$ and a unit vector $\beta = (\beta_1, \beta_2)$, we have
\[ e_\ell(\tilde{\theta}) + t\beta = e_\ell(\tilde{\theta}) + t\beta \cdot \nabla e_\ell(\tilde{\theta}) + O(t^2) \]
and
\[ e_\ell(\tilde{\theta}) + t\beta = e_\ell(\tilde{\theta}) + t\beta \cdot \nabla e_\ell(\tilde{\theta}) \]
\[ = e_\ell(\tilde{\theta}) + t\beta \cdot \nabla e_\ell(\tilde{\theta}) \]
\[ - \frac{t^2}{2} \left[ \frac{\beta_1^2}{p_1^2} \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) + \frac{\beta_2^2}{p_2^2} \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right. \]
\[ + \left. \left( \frac{\beta_1}{p_1} - \frac{\beta_2}{p_2} \right)^2 \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} - \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right] + O(t^3). \]
In particular, we will use (3.6) if $\beta \cdot \nabla e_\ell(\tilde{\theta}) \neq 0$, and (3.7) otherwise.

For any vector $\beta \in \mathbb{R}^2 \setminus \{0\}$, let
\[ \mathcal{J}_\beta^0 = \mathcal{J}_\beta^0(\tilde{\theta}) := \{ \ell \in \Lambda_E(\tilde{\theta}) : \beta \cdot \nabla e_\ell(\tilde{\theta}) = 0 \}, \]
\[ \mathcal{J}_\beta^\pm = \mathcal{J}_\beta^\pm(\tilde{\theta}) := \{ \ell \in \Lambda_E(\tilde{\theta}) : \pm \beta \cdot \nabla e_\ell(\tilde{\theta}) > 0 \}. \]
Consequently, we always have
\[ |\mathcal{J}_\beta^0| + |\mathcal{J}_\beta^+| + |\mathcal{J}_\beta^-| = r. \]
We also define $\mathcal{J}_0$ as follows
\[ \mathcal{J}_0 = \mathcal{J}_0(\tilde{\theta}) := \{ \ell \in \Lambda_E(\tilde{\theta}) : \nabla e_\ell(\tilde{\theta}) = 0 \}. \]
Since $E \neq -2$, Lemma 3.3 clearly implies $\mathcal{J}_0 = \emptyset$. 

We choose \( \beta_1 = (\beta_{1,1}, \beta_{1,2}) = (p_1, p_2)/\sqrt{p_1^2 + p_2^2} \). Then (3.2) is equivalent to

\[
\beta_1 \cdot \nabla e_{\ell_0}(\tilde{\theta}) = 0,
\]

hence \( \mathcal{J}^0_{\beta_1} \neq \emptyset \).

Next we are going to perturb the point \( \tilde{\theta} \) and count the eigenvalues. Since \( \mathcal{J}_0 = \emptyset \), we can choose a unit vector \( \beta_2 \) such that

\[
(3.11) \quad \beta_2 \cdot \nabla e_{\ell}(\tilde{\theta}) \neq 0
\]

holds for any \( \ell \in \Lambda(\tilde{\theta}) \). Thus, \( \mathcal{J}^0_{\beta_2} = \emptyset \), so one concludes

\[
(3.12) \quad |\mathcal{J}^+_{\beta_2}| + |\mathcal{J}^-_{\beta_2}| = r.
\]

**Perturbation along \( \beta_2 \).** We first perturb the eigenvalues along the \( \beta_2 \) direction. Since \( \mathcal{J}^0_{\beta_2} = \emptyset \), we will always employ (3.6).

For \( t > 0 \) small enough, we have the following:

- If \( \ell \in \mathcal{J}^+_{\beta_2} \), we have
  \[
  E_{k+r-s+1}(\tilde{\theta} + t\beta_2) > e_{\ell}(\tilde{\theta} + t\beta_2) > E = \max F_k,
  \]
  which implies
  \[
  (3.13) \quad |\mathcal{J}^+_{\beta_2}| \leq r - s.
  \]

- If \( \ell \in \mathcal{J}^-_{\beta_2} \), we have
  \[
  E_{k-s}(\tilde{\theta} + t\beta_2) < e_{\ell}(\tilde{\theta} + t\beta_2) < E = \min F_{k+1},
  \]
  which implies
  \[
  (3.14) \quad |\mathcal{J}^-_{\beta_2}| \leq s.
  \]

In view of (3.12), Equations (3.13) and (3.14) imply

\[
(3.15) \quad |\mathcal{J}^-_{\beta_2}| = s.
\]

Upon realizing that \( \mathcal{J}^0_{-\beta_2} = \emptyset \) and \( \mathcal{J}^\pm_{-\beta_2} = \mathcal{J}^\mp_{\beta_2} \), we may apply the analysis above with \( \beta_2 \) replaced by \( -\beta_2 \) and conclude that

\[
(3.16) \quad |\mathcal{J}^+_{-\beta_2}| = |\mathcal{J}^-_{-\beta_2}| = s.
\]

In particular, (3.15) and (3.16) imply

\[
(3.17) \quad r = 2s.
\]
Perturbation along $\beta_1$. Now we perturb the eigenvalues along $\beta_1$. Without loss of generality, we assume $E \in (-2, 6)$. The other case can be handled similarly. The case when $\ell \in \partial_{\beta_1}^{\pm}$ is similar to that of $\beta_2$. The difference here is $\partial_{\beta_1}^{0} \neq \emptyset$.

By Lemma 3.2, we have

$$
\cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) + \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) = -1 + \sqrt{E + 3} \neq 0
$$

for $\ell = (\ell_1, \ell_2) \in \partial_{\beta_1}^{0}$. Thus, by employing (3.7), we obtain

$$
e_e(\tilde{\theta} + t\beta_1)
$$

(3.19)

$$
= E - \frac{t^2}{2(p_1^2 + p_2^2)} \left( \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) + \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right) + O(t^3)
$$

Notice that the choice of $\beta_1$ causes the third $t^2$ term of (3.7) to drop out.

Since $E \in (-2, 6)$, (3.19) implies that

$$
e_e(\tilde{\theta} + t\beta_1) < E = \min F_{k+1}
$$

(3.20)

holds for $|t| > 0$ small enough and for any $\ell \in \partial_{\beta_1}^{0}$.

Combining (3.20) with (3.6), we have the following:

For $t > 0$ small enough:

- If $\ell \in \partial_{\beta_1}^{+}$, we have

$$
E_{k+r-s+1}(\tilde{\theta} + t\beta_1) > e_\ell(\tilde{\theta} + t\beta_1) > E = \max F_k,
$$

which implies

$$
|\partial_{\beta_1}^{+}| \leq r - s = s,
$$

(3.21)

where the equality follows from (3.17).

- If $\ell \in \partial_{\beta_1}^{0} \cup \partial_{\beta_1}^{-}$, we have

$$
E_{k-s-1}(\tilde{\theta} + t\beta_1) < e_\ell(\tilde{\theta} + t\beta_1) < E = \min F_{k+1},
$$

which implies

$$
|\partial_{\beta_1}^{0}| + |\partial_{\beta_1}^{-}| \leq s.
$$

(3.22)

In view of (3.9) and (3.17), equations (3.21) and (3.22) yield

$$
|\partial_{\beta_1}^{+}| = |\partial_{\beta_1}^{0}| + |\partial_{\beta_1}^{-}| = s.
$$

(3.23)
As before, we may observe that \( J_0 - \beta_1 = J_0^* \beta_1 \) and \( J_{-\beta_1} = J_{-\beta_1}^+ \). Then, the analysis above applied with \( \beta_1 \) replaced by \(-\beta_1\) forces

\[
|J_{-\beta_1}| = |J_0^* \beta_1| + |J_{-\beta_1}^+| = s.
\]

Taken together, (3.23) and (3.24) imply \( |J_0^* \beta_1| = 0 \), which contradicts \( J_0^* \beta_1 \neq \emptyset \). \( \square \)

3.1.2 \( E = -2 \). First, we would like to make a remark on our strategy in the proof of the \( E = -2 \) case, and on the importance of one of the periods being odd.

Remark 3.4. We will choose \( \tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2) \) and \( \ell^{(1)} = (\ell_1^{(1)}, \ell_2^{(1)}) \) such that \( e_{\ell^{(1)}}(\tilde{\theta}) = -2 \) and \( \nabla e_{\ell^{(1)}}(\tilde{\theta}) = 0 \). Lemma 3.3 yields three possibilities

\[
(p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1^{(1)}), p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2^{(1)})) = (0, \pi), \quad (\pi, 0) \quad \text{or} \quad (\pi, \pi).
\]

The choice of \( \tilde{\theta} \) depends on which one of \( p_1, p_2 \) is odd; we will choose \((0, \pi)\) if \( p_1 \) is odd and \((\pi, 0)\) if \( p_2 \) is odd. This choice guarantees that the only eigenvalue located at \(-2\) with vanishing gradient is \( e_{\ell^{(1)}}(\tilde{\theta}) \). Consequently, it suffices to control the second-order perturbation of (a single eigenvalue) \( e_{\ell^{(1)}}(\tilde{\theta}) \) along a given direction \((\beta_1, \beta_2)\). When \( p_1 \) is odd, this is equivalent to controlling the sign of the following expression (compare (3.28)):

\[
-\beta_2 \left( \frac{\beta_1}{p_1} - \frac{\beta_2}{p_2} \right).
\]

We can easily choose two directions such that the expression above has different signs, which leads to un-even eigenvalue counts and hence to the desired contradiction.

A posteriori, the existence of a \((2, 2)\)-periodic potential satisfying the conclusion of Theorem 1.4 implies that this argument must fail if both \( p_1 \) and \( p_2 \) are even; let us briefly describe why this must be the case. If both \( p_1, p_2 \) are even, there will be three eigenvalues at \(-2\) with vanishing gradients, corresponding to all three solutions \((0, \pi), (\pi, 0), (\pi, \pi)\). Trying to control the second-order perturbations of all these three eigenvalues along \((\beta_1, \beta_2)\) is equivalent to controlling the signs of the following three expressions simultaneously

\[
-\beta_2 \left( \frac{\beta_1}{p_1} - \frac{\beta_2}{p_2} \right), \quad \beta_1 \left( \frac{\beta_1}{p_1} - \frac{\beta_2}{p_2} \right), \quad \text{and} \quad \beta_1 \beta_2.
\]

A simple inspection of these three expressions yields that two of them are always non-negative with the other one being non-positive. Therefore we can never choose two different directions that lead to un-even eigenvalue counts. This explains why at least one of the periods must be odd for our argument to work.
Proof of Theorem 3.1.2. Now let us give a detailed proof. Without loss of
generality, assume \( p_1 \) is odd, let \( E = -2 \), and assume for the sake of contradiction
that \( E = \max F_k = \min F_{k+1} \) for some \( k \). We choose \( \theta \) and \( \ell^{(1)} \) via
\[
\bar{\ell}_1 = 0, \quad \ell_1^{(1)} = 0, \quad (\bar{\ell}_2, \ell_2^{(1)}) = \begin{cases} (0, \frac{p_2}{2}), & \text{if } p_2 \text{ is even}, \\ (\pi, \frac{p_2-1}{2}), & \text{if } p_2 \text{ is odd}. \end{cases}
\]

With these choices of \( \ell^{(1)} \) and \( \bar{\theta} \), one can check that \( e_{\ell^{(1)}}(\bar{\theta}) = -2 = E \). As
before, let \( r \) denote the multiplicity of \( E \) and let \( \Lambda_E(\bar{\theta}) \) denote the set of \( \ell \in \Lambda \) with
\( e_\ell(\bar{\theta}) = -2 \). Note that we also have \( \nabla e_{\ell^{(1)}}(\bar{\theta}) = 0 \), and thus \( \mathcal{J}_0 \neq \emptyset \). Moreover, we
claim that \( \mathcal{J}_0 = \{ \ell^{(1)} \} \). To see this, suppose there exists \( \ell \neq \ell^{(1)} \) in \( \mathcal{J}_0 \). In view of
Lemma 3.3, we must have
\[
\frac{\bar{\ell}_1 + 2\pi \ell_1}{p_1} = \pi,
\]
which implies \( p_1 = 2\ell_1 \), which is impossible, since \( p_1 \) is odd. Consequently,
\( \mathcal{J}_0 = \{ \ell^{(1)} \} \).

Let us choose \( \beta_1 = (\beta_{1,1}, \beta_{1,2}) = (0, 1) \) and a unit vector
\[
\beta_2 = (\beta_{2,1}, \beta_{2,2}) \sim (2p_1, p_2)/\sqrt{4p_1^2 + p_2^2}
\]
such that
\[
(3.25) \quad \beta_{2,2} \left( \frac{\beta_{2,1}}{p_1} - \frac{\beta_{2,2}}{p_2} \right) > 0,
\]
and
\[
(3.26) \quad \beta_2 \cdot \nabla e_{\ell^{(1)}}(\bar{\theta}) \neq 0 \quad \text{holds for any } \ell \in \Lambda_E(\bar{\theta}) \setminus \{ \ell^{(1)} \}.
\]

We will use (3.25) to control the perturbation of \( e_{\ell^{(1)}}(\bar{\theta}) \) along the \( \beta_2 \) direction.
We also note that (3.26) simply says
\[
(3.27) \quad \mathcal{J}_{\beta_2}^0 = \mathcal{J}_0 = \{ \ell^{(1)} \}.
\]

Perturbation along \( \beta_2 \). We first perturb the eigenvalues along the \( \beta_2 \) direction.

By (3.27), we need only consider first-order perturbation theory as in (3.6) for
\( \ell \in \Lambda_E(\bar{\theta}) \setminus \{ \ell^{(1)} \} \). Since \( \ell^{(1)} \in \mathcal{J}_0 \), we need to employ (3.7) for \( e_{\ell^{(1)}} \). Indeed, by
(3.7), we have for \( |t| > 0 \) small enough,
\[
(3.28) \quad e_{\ell^{(1)}}(\bar{\theta} + t\beta_2) = e_{\ell^{(1)}}(\bar{\theta}) - \frac{t^2}{2} \left[ \frac{\beta_{2,1}^2}{p_1^2} - \frac{\beta_{2,2}^2}{p_2^2} - \left( \frac{\beta_{2,1}}{p_1} - \frac{\beta_{2,2}}{p_2} \right)^2 \right] + O(t^3)
\]
\[
= -2 - \frac{\beta_{2,2}}{p_2} \left( \frac{\beta_{2,1}}{p_1} - \frac{\beta_{2,2}}{p_2} \right) t^2 + O(t^3) < -2 = \min F_{k+1},
\]
where we used (3.25) in the last inequality.
For \( t > 0 \) small enough, we then have the following:

- If \( \ell \in J_{\beta_2}^+ \), we have
  \[
  E_{k+r-s+1}(\tilde{\theta} + t\beta_2) > e_\ell(\tilde{\theta} + t\beta_2) > E = \max F_k,
  \]
  which implies
  \[
  (3.29) \quad |J_{\beta_2}^+| \leq r - s.
  \]

- If \( \ell \in J_{\beta_2}^- \cup J_0 \), we have
  \[
  E_{k-s}(\tilde{\theta} + t\beta_2) < e_\ell(\tilde{\theta} + t\beta_2) < E = \min F_{k+1},
  \]
  which implies
  \[
  (3.30) \quad |J_{\beta_2}^-| + |J_0| \leq s.
  \]

Taking (3.9), (3.27), (3.29), and (3.30) into account, we have

\[
(3.31) \quad |J_{\beta_2}^-| = s - 1.
\]

Replacing \( \beta_2 \) by \(-\beta_2\) as in previous phases of the argument, we arrive at

\[
(3.32) \quad |J_{\beta_2}^+| = s - 1.
\]

Combining (3.31) with (3.32), we arrive at

\[
(3.33) \quad r = |J_{\beta_2}^+| + |J_{\beta_2}^-| + |J_0| = 2s - 1.
\]

**Perturbation along \( \beta_1 \).** Now we perturb the eigenvalues along \( \beta_1 = (0, 1) \). The case when \( \ell \in J_{\beta_1}^+ \) is similar to that of \( \beta_2 \). The difference here is the behavior of perturbations of \( e_{\ell(1)} \) in the direction \( \beta_1 \). Indeed, by (3.7), we have

\[
e_{\ell(1)}(\tilde{\theta} + t\beta_1) = e_{\ell(1)}(\tilde{\theta}) - \frac{t^2}{2} \left[ \frac{\beta_{1.1}^2}{p_1^2} - \frac{\beta_{1.2}^2}{p_2^2} - \left( \frac{\beta_{1.1}}{p_1} - \frac{\beta_{1.2}}{p_2} \right)^2 \right] + O(t^3)
\]

\[
= -2 + \frac{t^2}{p_2^2} + O(t^3) > -2 = \max F_k.
\]

Thus, the perturbations of \( e_{\ell(1)} \) in the direction \( \beta_1 \) always move up.

For \( t > 0 \) small enough:

- If \( \ell \in J_{\beta_1}^+ \cup J_0 \), we have
  \[
  E_{k+r-s+1}(\tilde{\theta} + t\beta_1) > e_\ell(\tilde{\theta} + t\beta_1) > E = \max F_k,
  \]
  which implies
  \[
  (3.34) \quad |J_{\beta_1}^+| + |J_0| \leq r - s.
  \]
• If $\ell \in \mathcal{J}_{-\beta_1}$, we have

$$E_{k-s-1}(\tilde{\theta} + t\beta_1) < e\ell(\tilde{\theta} + t\beta_1) < E = \min F_{k+1},$$

which implies

(3.35) \hspace{1cm} |\mathcal{J}_{-\beta_1}| \leq s.

In view of (3.9), Equations (3.34) and (3.35) yield

(3.36) \hspace{1cm} |\mathcal{J}_{\beta_1}| = s.

Applying the usual symmetry argument, we also arrive at $|\mathcal{J}_{\beta_1}| = s$, which leads to

$$r = |\mathcal{J}_{\beta_1}| + |\mathcal{J}_{-\beta_1}| + |\mathcal{J}_0| = 2s + 1,$$

which in turn contradicts (3.33).

3.2 Proof of Lemmas 3.2 and 3.3.

Proof of Lemma 3.2. Let $E \in [-3, 6]$ be given, let $x$ be as-yet-unspecified, set $y = 2\pi - x$, and note that (3.2) holds. Then, using $y = 2\pi - x$, we note that

$$\cos(x) + \cos(y) + \cos(x - y) = 2\cos(x) + \cos(2x) = 2\cos(x) + 2\cos^2(x) - 1.$$

Setting $z = \cos(x)$, we seek to solve $2z + 2z^2 - 1 = E/2$, which gives

$$z^2 + z - \frac{1}{2} - \frac{E}{4} = 0 \implies z = \frac{-1 \pm \sqrt{3 + E}}{2}.$$

Thus, we may take $x$ so that

$$\cos(x) = \frac{-1 + \sqrt{3 + E}}{2}.$$

In fact, since $-3 \leq E \leq 6$, we may take $0 \leq x \leq 2\pi/3$. Thus, with this choice of $x$ (and $y = 2\pi - x$), we get (3.1).

Finally, suppose $x$ and $y$ solve (3.1) and (3.2) for $E \neq -2$. From (3.2), we deduce that either $x + y = 2\pi$ or $|x - y| = \pi$. The second option leads to $E = -2$, so we must have $y = 2\pi - x$. Solving for $\cos x$ as before yields (3.3) when $E < -2$ and (3.4) when $E > -2$.

Proof of Lemma 3.3. Suppose that $x$ and $y$ solve

(3.37) \hspace{1cm} \cos(x) + \cos(y) + \cos(x - y) = \lambda,

(3.38) \hspace{1cm} \sin(x) + \sin(x - y) = 0,

(3.39) \hspace{1cm} \sin(y) - \sin(x - y) = 0,
for some \( \lambda \in (-3/2, 3) \). Adding (3.38) and (3.39), we arrive at

\[
\sin(x) = -\sin(y).
\]

For \((x, y) \in [0, 2\pi)^2\), this forces either \(|x-y| = \pi\) or \(x+y = 2\pi\). In the case \(|x-y| = \pi\), substituting into (3.38) and (3.39) gives \(\sin(x) = \sin(y) = 0\), forcing \(x, y \in \{0, \pi\}\). Plugging the various possibilities into (3.37), one either gets \(\lambda = 3 \notin (-3/2, 3)\) (when \(x = y = 0\)) or \(\lambda = -1\) (when at least one of \(x\) or \(y\) is \(\pi\)).

Alternatively, if \(x = 2\pi - y\), (3.38) yields \(\sin(x) + \sin(2x) = 0\), which leads to

\[
\sin(x)(1 + 2 \cos(x)) = 0.
\]

Setting \(\sin(x) = 0\) yields \(x \in \{0, \pi\}\) which leads to the same solutions as before. Setting \(1 + 2 \cos(x) = 0\) yields \((x, y) = (2\pi/3, 4\pi/3)\) or \((x, y) = (4\pi/3, 2\pi/3)\). Plugging in either possibility into (3.37) yields

\[
\cos(x) + \cos(y) + \cos(x-y) = -\frac{3}{2} \notin (-3/2, 3),
\]

as claimed.

### 3.3 Opening a gap at \(-2\).

Let us exhibit a \((2, 2)\)-periodic potential that perturbatively opens a gap at energy \(E = -2\) for the triangular lattice (see Figure 6).

![Figure 6. A (2, 2)-periodic potential on the triangular lattice with a gap at \(E = -2\) for all positive coupling constants (color figure available online).](image-url)
Theorem 3.5. Define
\[ Q_{n,m} = (-1)^{mn} = \begin{cases} 1 & \text{if } m \text{ or } n \text{ is even,} \\ -1 & \text{if both } m \text{ and } n \text{ are odd,} \end{cases} \]
and denote \( H_\lambda = \Delta_{\text{tri}} + \lambda Q \). For all \( \lambda > 0 \), \( \sigma(H_\lambda) \) has two connected components. Moreover, for all \( \lambda > 0 \) sufficiently small, the gap that opens about \( E = -2 \) is precisely equal to
\[ g_\lambda = (-\sqrt{4 + \lambda^2}, -2 + \lambda). \]
In particular,
\[ |g_\lambda| = \lambda + (\sqrt{4 + \lambda^2} - 2) \sim \lambda + \frac{\lambda^2}{4}, \]
so the gap opens linearly as \( \lambda \downarrow 0 \).

The following lemma will be used:

Lemma 3.6. For all \( \theta \in \mathbb{T}^2 \) and all \( 0 \leq a \leq 54, \)
\[ 4(\sin \theta_1 + \sin \theta_2 - \sin(\theta_1 + \theta_2))^2 + a(1 + \cos \theta_1 + \cos \theta_2 + \cos(\theta_1 + \theta_2)) \geq 0. \]

Proof. Define
\[ g(\theta_1, \theta_2, a) = 4(\sin \theta_1 + \sin \theta_2 - \sin(\theta_1 + \theta_2))^2 + a(1 + \cos \theta_1 + \cos \theta_2 + \cos(\theta_1 + \theta_2)). \]
We begin by checking the boundary of \( \mathbb{T}^2 \times [0, 54] \). It is easy to see that \( g \geq 0 \) if \( a = 0 \). For \( a = 54 \), define \( h(\theta) = g(\theta, 54) \). Using the identities
\[ \sin x + \sin y = 2 \sin \left( \frac{x}{2} \right) \sin \left( \frac{y}{2} \right) \sin \left( \frac{x+y}{2} \right), \]
\[ \cos x - \cos(x+y) = 2 \sin \left( \frac{y}{2} \right) \sin \left( \frac{x+y}{2} \right), \]
\[ \sin x + \sin(x+y) = 2 \cos \left( \frac{y}{2} \right) \sin \left( \frac{x+y}{2} \right), \]
we may simplify \( \nabla h \) to get
\begin{align*}
\frac{\partial h}{\partial \theta_1} &= 4 \sin \left( \frac{\theta_1 + \theta_2}{2} \right) \\
&\quad \times \left[ 16 \sin \left( \frac{\theta_1}{2} \right) \sin^2 \left( \frac{\theta_2}{2} \right) \sin \left( \frac{\theta_1 + \theta_2}{2} \right) - 27 \cos \left( \frac{\theta_2}{2} \right) \right], \\
\frac{\partial h}{\partial \theta_2} &= 4 \sin \left( \frac{\theta_2 + \theta_1}{2} \right) \\
&\quad \times \left[ 16 \sin^2 \left( \frac{\theta_1}{2} \right) \sin \left( \frac{\theta_2}{2} \right) \sin \left( \frac{\theta_1 + \theta_2}{2} \right) - 27 \cos \left( \frac{\theta_1}{2} \right) \right].
\end{align*}
Consequently, setting $\nabla h = 0$ leads to four cases. For notational convenience, define 
\[ \alpha = \arcsin \sqrt{\frac{27}{32}}. \]

**Case 1.**
\[ \sin \left( \frac{\theta_1 + \theta_2}{2} \right) = \sin \left( \frac{\theta_2 + \theta_1}{2} \right) = 0. \]
This implies $\theta_1 + \frac{1}{2} \theta_2 \in \pi \mathbb{Z}$ and $\theta_2 + \frac{1}{2} \theta_1 \in \pi \mathbb{Z}$. Solving the resulting systems for solutions in $[0, 2\pi)$ yields three points:
\[ \boldsymbol{\theta} = (0, 0), \left( \frac{2\pi}{3}, \frac{2\pi}{3} \right), \left( \frac{4\pi}{3}, \frac{4\pi}{3} \right). \]

**Case 2.**
\[ \sin \left( \frac{\theta_1 + \theta_2}{2} \right) = 16 \sin^2 \left( \frac{\theta_1}{2} \right) \sin \left( \frac{\theta_2}{2} \right) \sin \left( \frac{\theta_1 + \theta_2}{2} \right) - 27 \cos \left( \frac{\theta_1}{2} \right) = 0. \]
As before, the first condition forces $\theta_1 + \frac{1}{2} \theta_2 \in \pi \mathbb{Z}$. Plugging the various possibilities that this yields into the second condition gives three solutions:
\[ \boldsymbol{\theta} = (\pi, 0), \left( 2\alpha, 2\pi - 4\alpha \right), \left( 2\pi - 2\alpha, 4\alpha \right). \]

**Case 3.**
\[ \sin \left( \frac{\theta_2 + \theta_1}{2} \right) = 16 \sin \left( \frac{\theta_1}{2} \right) \sin^2 \left( \frac{\theta_2}{2} \right) \sin \left( \frac{\theta_1 + \theta_2}{2} \right) - 27 \cos \left( \frac{\theta_2}{2} \right) = 0. \]
Arguing as in Case 2, there are three solutions:
\[ \boldsymbol{\theta} = (0, \pi), \left( 2\pi - 4\alpha, 2\alpha \right), \left( 4\alpha, 2\pi - 2\alpha \right). \]

**Case 4.**
\begin{align*}
(3.42) & \quad 16 \sin^2 \left( \frac{\theta_1}{2} \right) \sin \left( \frac{\theta_2}{2} \right) \sin \left( \frac{\theta_1 + \theta_2}{2} \right) - 27 \cos \left( \frac{\theta_1}{2} \right) = 0, \\
(3.43) & \quad 16 \sin \left( \frac{\theta_1}{2} \right) \sin^2 \left( \frac{\theta_2}{2} \right) \sin \left( \frac{\theta_1 + \theta_2}{2} \right) - 27 \cos \left( \frac{\theta_2}{2} \right) = 0.
\end{align*}
Multiply (3.42) by $\sin(\theta_2/2)$, multiply (3.43) by $\sin(\theta_1/2)$, and subtract the results to obtain
\[ \sin \left( \frac{\theta_1 - \theta_2}{2} \right) = 0. \]
Using this, we see that the solutions are
\[ \boldsymbol{\theta} = (\pi, \pi), \left( 2\alpha, 2\alpha \right), \left( 2\pi - 2\alpha, 2\pi - 2\alpha \right). \]
Evaluating \( g \) at these points, we find out that \( \max h(\theta) = 216 \) attained at \((0, 0)\) and \( \min h(\theta) = 0 \), attained at \((\pi, 0), (0, \pi), (\pi, \pi)\).

Finally, we need to look at critical points of \( g \) in the interior of \( \mathbb{T}^2 \times [0, 54] \). However, this is easy. Any zero of \( \nabla g \) must in particular satisfy

\[
\partial g = 0, \quad \text{which forces} \quad 1 + \cos \theta_1 + \cos \theta_2 + \cos(\theta_1 + \theta_2) = 0,
\]

which clearly implies \( g \geq 0 \). \( \square \)

**Proof of Theorem 3.5.** For \( \theta = (\theta_1, \theta_2) \in \mathbb{T}^2 \), denote by \( H_\lambda(\theta) \) the Floquet matrix corresponding to \( H_\lambda \). Ordering the vertices of the \( 2 \times 2 \) fundamental domain as shown in Figure 6, we obtain

\[
H_\lambda(\theta) - (-2 + \varepsilon) = \begin{bmatrix}
2 + \lambda - \varepsilon & 1 + e^{-i\theta_1} & 1 + e^{-i\theta_2} & 1 + e^{-i(\theta_1 + \theta_2)} \\
1 + e^{i\theta_1} & 2 + \lambda - \varepsilon & e^{i\theta_1} + e^{-i\theta_2} & 1 + e^{-i\theta_2} \\
1 + e^{i\theta_2} & e^{-i\theta_1} + e^{i\theta_2} & 2 + \lambda - \varepsilon & 1 + e^{-i\theta_1} \\
1 + e^{i(\theta_1 + \theta_2)} & 1 + e^{i\theta_2} & 1 + e^{i\theta_1} & 2 + \lambda - \varepsilon
\end{bmatrix}.
\]

For \( \theta \in \mathbb{T}^2, \lambda > 0 \), and \( \varepsilon \in \mathbb{R} \), define

\[
p(\theta, \lambda, \varepsilon) = \det(H_\lambda(\theta) - (-2 + \varepsilon)) = \det(H_\lambda(\theta) - 2) \]

After some calculations, one observes that

\[
p(\theta, \lambda, \varepsilon) = -\lambda^4 - 4\lambda^3 + X(\theta) - 4\varepsilon\lambda \left(3 - \frac{\lambda^2}{2} - \cos \theta_1 - \cos \theta_2 - \cos(\theta_1 + \theta_2)\right) + 4\varepsilon^2(3 + 3\lambda - \cos \theta_1 - \cos \theta_2 - \cos(\theta_1 + \theta_2)) - 2\varepsilon^3(4 + \lambda) + \varepsilon^4,\]

where

\[
X(\theta) = -4(\sin \theta_1 + \sin \theta_2 - \sin(\theta_1 + \theta_2))^2.
\]

Clearly \( X(\theta) \leq 0 \) for all \( \theta \), so we have

\[
det(H_\lambda(\theta) + 2) = p(\theta, \lambda, 0) \leq -\lambda^4 - 4\lambda^3 < 0
\]

for all \( \lambda > 0 \); consequently \(-2 \notin \sigma(H_\lambda)\) for all \( \lambda > 0 \), which proves the first claim of the theorem. Introducing

\[
W_1(\lambda, \varepsilon) := -\lambda^4 - 4\lambda^3 + 2\varepsilon\lambda^3 + 12\varepsilon^2\lambda - 2\varepsilon^3(4 + \lambda) + \varepsilon^4,
\]
we may rewrite $p$ as

\[(3.44)\quad p(\theta, \lambda, \varepsilon) = X(\theta) - 4\varepsilon(\lambda - \varepsilon)(3 - \cos \theta_1 - \cos \theta_2 - \cos(\theta_1 + \theta_2)) + W_1(\lambda, \varepsilon).\]

By standard eigenvalue perturbation theory, we know that $|g_\lambda^2 + 2| \leq \lambda$, so we need only concern ourselves with $|\varepsilon| \leq \lambda$. Since $X(\theta) \leq 0$ for all $\theta$ and the second term of (3.44) is non-positive whenever $0 \leq \varepsilon \leq \lambda$, we arrive at

\[p(\theta, \lambda, \varepsilon) \leq -\lambda^4 - 4\lambda^3 + 2\varepsilon\lambda^2 - 2\varepsilon^3(4 + \lambda) + \varepsilon^4 = W_1(\lambda, \varepsilon)\]

for all $\theta \in \mathbb{T}^2$, all $\lambda > 0$, and all $0 \leq \varepsilon \leq \lambda$. Moreover, we observe that $p(0, \lambda, \varepsilon) = W_1(\lambda, \varepsilon)$, so this bound is sharp. Factoring $W_1$, we arrive at

\[W_1(\lambda, \varepsilon) = (\lambda - \varepsilon)^2(\varepsilon^2 - 8\varepsilon - \lambda^2 - 4\lambda).\]

Consequently, we see that $W_1(\lambda, \varepsilon) < 0$ for $\varepsilon \in [0, \lambda)$, which implies that $p(\theta, \lambda, \varepsilon) < 0$ for all $\theta \in \mathbb{T}^2$, all $\lambda > 0$, and all $0 \leq \varepsilon < \lambda$; consequently, $[-2, -2 + \lambda) \cap \sigma(H_\lambda) = \emptyset$, which is to say

\[(3.45)\quad [-2, -2 + \lambda) \subseteq \sigma_1.\]

On the other hand, $p(0, \lambda, \lambda) = 0$, so

\[(3.46)\quad -2 + \lambda \in \sigma(H_\lambda(0)) \subseteq \sigma(H_\lambda).\]

Alternatively, $-2 + \lambda \in \sigma(H_\lambda)$ is clear from eigenvalue perturbation theory as soon as one has $[-2, 2 + \lambda) \cap \sigma(H_\lambda) = \emptyset$.

Now, for $-\lambda \leq \varepsilon \leq 0$, we have to be more careful with the term

\[q(\theta, \lambda, \varepsilon) := -4\varepsilon(\lambda - \varepsilon)(3 - \cos \theta_1 - \cos \theta_2 - \cos(\theta_1 + \theta_2)),\]

as $q$ can be positive when $-\lambda < \varepsilon < 0$. Naively, one can bound

\[3 - \cos \theta_1 - \cos \theta_2 - \cos(\theta_1 + \theta_2) \leq \frac{9}{2},\]

which leads to the upper bound of $X(\theta) + q(\theta, \lambda, \varepsilon) \leq -18\varepsilon(\lambda - \varepsilon)$. However, the maximum of $q$ occurs at the global minimum of $X$, so we can do better. Indeed, for $\lambda > 0$ small and $-\lambda \leq \varepsilon \leq 0$, we have

\[(3.47)\quad X(\theta) + q(\theta, \lambda, \varepsilon) \leq -16\varepsilon(\lambda - \varepsilon).\]

In particular, by Lemma 3.6, the bound in (3.47) holds for all $\varepsilon$ such that $-\lambda \leq \varepsilon \leq 0$ as long as $8\lambda^2 < 54$, i.e., $0 < \lambda < \frac{3\sqrt{3}}{2}$. This then leads us to

\[p(\theta, \lambda, \varepsilon) \leq W_2(\lambda, \varepsilon) := -\lambda^4 - 4\lambda^3 + 2\varepsilon\lambda^2 + 12\varepsilon^2\lambda - 2\varepsilon^3(4 + \lambda) + \varepsilon^4 - 16\varepsilon(\lambda - \varepsilon) = W_1(\lambda, \varepsilon) - 16\varepsilon(\lambda - \varepsilon)\]
for $\lambda > 0$ small and $-\lambda \leq \varepsilon \leq 0$. Factoring $W_2$ yields

$$p(\theta, \lambda, \varepsilon) \leq W_2(\lambda, \varepsilon) = (\varepsilon - \lambda)(\varepsilon - \lambda - 4)(\varepsilon^2 - 4\varepsilon - \lambda^2)$$

for $\lambda > 0$ small and $-\lambda \leq \varepsilon \leq 0$. It is straightforward to find the roots of $W_2$ and to observe that $W_2(\lambda, \varepsilon) < 0$ when

$$2 - \sqrt{4 + \lambda^2} < \varepsilon \leq 0.$$

As a result, this implies $p(\theta, \lambda, \varepsilon) < 0$ for all $\theta$, all $\lambda > 0$ small, and all $\varepsilon \in (2 - \sqrt{4 + \lambda^2}, 0]$, which in turn yields

$$(-\sqrt{4 + \lambda^2}, -2] \subseteq g_\lambda.$$

On the other hand,

$$p((\pi, \pi), \lambda, 2 - \sqrt{4 + \lambda^2}) = W_2(\lambda, 2 - \sqrt{4 + \lambda^2}) = 0,$$

which leads us to conclude that

$$(-\sqrt{4 + \lambda^2}, -2 + \lambda)$$

for small $\lambda$, as promised. □

The effort involved in proving Lemma 3.6 in order to improve the constant “18” to “16” is non-trivial, but worthwhile. In particular, this is exactly what enables the exact factorization of $W_2$ and hence the ability to exactly compute the gap edges.

4 Hexagonal Laplacian

We now continue with the Laplacian on the hexagonal lattice. Let

$$\Gamma_\text{hex} = (\mathcal{V}_\text{hex}, \mathcal{E}_\text{hex})$$

and

$$b_\pm = \frac{1}{2} \left[ \begin{array}{c} 3 \\ \pm \sqrt{3} \end{array} \right]$$

be as in the introduction. It is not hard to check that $\{0, a_1\}$ is a fundamental set of vertices and hence every $v \in \mathcal{V}_\text{hex}$ may be written uniquely as either $nb_+ + mb_-$ or $a_1 + nb_+ + mb_-$ for integers $n, m$, so we have

$$\mathcal{V}_\text{hex} = \{nb_+ + mb_- : n, m \in \mathbb{Z}\} \cup \{a_1 + nb_+ + mb_- : n, m \in \mathbb{Z}\}.$$
Recall that $u \sim v$ for $u, v \in \mathcal{V}_{\text{hex}}$ if and only if $\|u - v\|_2 = 1$. After some calculations, we see that

$$[\Delta_{\text{hex}} \Psi]_{n, m, +} = \psi_{a_1 + n b_s + m b_-} + \psi_{a_1 + (n-1) b_s + mb_-},$$

$$[\Delta_{\text{hex}} \Psi]_{a_1 + n b_s + mb_-} = \psi_{b_s + mb_-} + \psi_{b_s + (m+1)b_-} + \psi_{(m+1)b_s + mb_-}.$$

The formula for $\Delta_{\text{hex}}$ can be made more compact if we view the associated Hilbert space as

$$\ell^2(\mathbb{Z}^2, \mathbb{C}^2) = \left\{ \Psi : \mathbb{Z}^2 \to \mathbb{C}^2 : \sum_{n, m} \|\Psi_{n, m}\|^2 < \infty \right\},$$

where the standard basis of $\mathbb{C}^2$ corresponds to the left and right vertices of the fundamental domain, respectively. More precisely, given $\psi \in \ell^2(\mathcal{V}_{\text{hex}})$, define

$$\Psi_{n, m} = \begin{bmatrix} \psi_{n b_s + m b_-} \\ \psi_{a_1 + n b_s + m b_-} \end{bmatrix}.$$

Identifying $\ell^2(\mathcal{V}_{\text{hex}})$ and $\ell^2(\mathbb{Z}^2, \mathbb{C}^2)$ in this fashion, the Laplacian for the hexagonal lattice is given by

$$[\Delta_{\text{hex}} \Psi]_{n, m} = U(\Psi_{n, m-1} + \Psi_{n-1, m}) + L(\Psi_{n, m+1} + \Psi_{n+1, m}) + J\Psi_{n, m},$$

where

$$U = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad L = U^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad J = U + L = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Equivalently, if we denote by $S_1, S_2 : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ the shift operators

$$[S_1 \psi]_{n, m} = \psi_{n+1, m}, \quad [S_2 \psi]_{n, m} = \psi_{n, m+1},$$

we have

$$\Delta_{\text{hex}} \Psi = \begin{bmatrix} (S_1^* + S_2^* + \mathbb{I}) \psi^- \\ (S_1 + S_2 + \mathbb{I}) \psi^+ \end{bmatrix} \quad \text{for any } \Psi = \begin{bmatrix} \psi^+ \\ \psi^- \end{bmatrix} \in \ell^2(\mathbb{Z}^2, \mathbb{C}^2).$$

Abbreviating somewhat, we write

\begin{equation}
(4.1) \quad \Delta_{\text{hex}} = \begin{bmatrix} 0 & S_1^* + S_2^* + \mathbb{I} \\ S_1 + S_2 + \mathbb{I} & 0 \end{bmatrix}.
\end{equation}

Now, let periods $p_1, p_2 \in \mathbb{Z}_+$ be given, and view $H = \Delta_{\text{hex}}$ as a $(p_1, p_2)$-periodic operator. For this setting, there are two vertices of $\mathcal{V}_{\text{hex}}$ in $\{ s b_s + t b_- : 0 \leq s, t < 1 \}$, so our Floquet operator $H(\theta)$ will be a $P \times P$ matrix with $P = 2p_1p_2$. As usual, define $\Lambda = ((0, p_1) \times (0, p_2)) \cap \mathbb{Z}^2$, denote the eigenvalues of $H(\theta)$ by

$$E_1^\Lambda(\theta) \leq \cdots \leq E_P^\Lambda(\theta),$$
and let $F_k^\Lambda$ for $1 \leq k \leq P$ denote the bands of the spectrum. Our main theorem in this section is the following result.

**Theorem 4.1.** Let $p_1, p_2 \in \mathbb{Z}_+$ be given.

(1) Every $E \in (-3, 3) \setminus \{-1, 0, 1\}$ belongs to $\text{int}(F_j)$ for some $1 \leq j \leq P$.

(2) If at least one of $p_1$ or $p_2$ is odd, then $-1 \in \text{int}(F_k)$ and $+1 \in \text{int}(F_\ell)$ for some $1 \leq k \leq \ell \leq P$.

**Proof of Theorem 1.1.** This follows immediately from Theorem 4.1. □

**Proof of Theorem 4.1.** From (4.1), we have

$$\Delta_{\text{hex}} = \begin{bmatrix} 0 & S_1^* + S_2^* + I \\ S_1 + S_2 + I & 0 \end{bmatrix},$$

where $S_j : \ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$ denote the shifts

$$[S_1 \psi]_{n,m} = \psi_{n+1,m}, \quad [S_2 \psi]_{n,m} = \psi_{n,m+1}.$$

It is easy to see that

$$S_1 + S_1^* + S_2 + S_2^* + S_1S_2^* + S_1^*S_2 = \Delta_{\text{tri}}$$

is the triangular Laplacian. Thus, a simple calculation shows that

$$(4.2) \quad [\Delta_{\text{hex}}^2 \Psi]_n \begin{bmatrix} [\Delta_{\text{tri}} \psi^+]_n + 3 \psi^+_n \\ [\Delta_{\text{tri}} \psi^-]_n + 3 \psi^-_n \end{bmatrix} \quad \text{for} \quad \Psi = \begin{bmatrix} \psi^+ \\ \psi^- \end{bmatrix} \in \ell^2(\mathbb{Z}^2, \mathbb{C}^2).$$

This calculation extends to the Floquet matrices, so we see that for each $1 \leq k \leq P$, the bands of $H = \Delta_{\text{hex}}$ obey

$$F_{k, \text{hex}}^\Lambda = -F_{P+1-k, \text{hex}}^\Lambda$$

and

$$(4.3) \quad F_{k, \text{hex}}^\Lambda = \begin{cases} \sqrt{F_{k-\frac{P}{2}, \text{tri}}^\Lambda + 3}, & \frac{P}{2} < k \leq P, \\
-\sqrt{F_{\frac{P}{2}+1-k, \text{tri}}^\Lambda + 3}, & 1 \leq k \leq \frac{P}{2}. \end{cases}$$

From this, we deduce that $E \in (-3, 3)$ lies in the interior of some $F_{k, \text{hex}}$ if and only if $E^2 - 3$ lies in the interior of some $F_{\ell, \text{tri}}$. For $E \in (-3, 3) \setminus \{-1, 0, 1\}$, $E^2 - 3 \in (-3, 6) \setminus \{-2\}$, while $(\pm 1)^2 - 3 = -2$. Thus, the conclusions of the theorem follow from Theorem 3.1. □
4.1 Opening gaps at 0 and ±1. Define the (1, 1)-periodic potential $Q_1$ on $\mathcal{V}_{\text{hex}}$ by $Q_1(0) = 1$ and $Q_1(a_1) = -1$, that is,

$$Q_1(nb_+ + mb_-) = 1, \quad Q_1(a_1 + nb_+ + mb_-) = -1, \quad n, m \in \mathbb{Z}.$$ 

After identifying $\ell^2(\mathcal{V}_{\text{hex}})$ with $\ell^2(\mathbb{Z}^2, \mathbb{C}^2)$ in the usual way, we get (as an operator) $[Q_1 \Psi]_n = Z^n \Psi_n$, where

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$ 

From the calculations $ZU = U = -UZ$ and $ZL = -L = -LZ$, we deduce that

$$Q_1 \Delta_{\text{hex}} + \Delta_{\text{hex}} Q_1 = 0,$$

and hence

$$(\Delta_{\text{hex}} + \lambda Q_1)^2 = \Delta_{\text{hex}}^2 + \lambda^2 \geq \lambda^2.$$ 

Consequently, $(-\lambda, \lambda) \cap \sigma(\Delta_{\text{hex}} + \lambda Q_1) = \emptyset$ and there is a gap at zero. In particular, the gap is precisely $(-\lambda, \lambda)$, and so opens linearly at the maximal possible rate.

Let us consider the (2, 2)-periodic case. We parameterize our potential as $\mathbf{q} = (q_1, \ldots, q_8) \in \mathbb{R}^8$ as shown in Figure 7.

![Figure 7. A portion of the hexagonal lattice. A fundamental domain for a (2, 2)-periodic potential is highlighted in red (color figure available online).](image)

We now turn to the construction of a potential that opens gaps at 0, 1, and −1 simultaneously. We show that it opens gaps linearly at zero, quadratically at ±1. Later on, we will show that one cannot open gaps linearly at ±1 on both sides.

**Theorem 4.2.** Order the vertices of a $2 \times 2$ fundamental cell of the hexagonal lattice as shown in Figure 7, define a (2, 2)-periodic potential $Q$ by

$$(q_1, \ldots, q_8) = (1, -1, 1, 2, -2, -1, 1, -1),$$

...
and denote $H_\lambda = \Delta_{\text{hex}} + \lambda Q$. Then, for $|\lambda| > 0$ sufficiently small, $\sigma(H_\lambda)$ consists of four connected components. Moreover, if $g_{E,\lambda} = (g_{E,\lambda}^+, g_{E,\lambda}^-)$ denote the gaps of $\sigma(H_\lambda)$ that open at $E = 0, \pm 1$, one has

$$
\left( \pm 1 - \frac{\lambda^2}{20}, \pm 1 + \frac{\lambda^2}{20} \right) \subset g_{\pm 1,\lambda} \subseteq \left( \pm 1 - \frac{1}{2} \lambda^2, \pm 1 + \frac{1}{2} \lambda^2 \right)
$$

and

$$
\left( - \frac{\lambda}{5}, \frac{\lambda}{5} \right) \subset g_{0,\lambda} \subset \left( - \frac{\lambda}{4}, \frac{\lambda}{4} \right)
$$

for all $|\lambda| > 0$ sufficiently small.

Let us point out that we do not carefully optimize the constants in the upper and lower bounds of the gaps; it is possible to get better constants than $1/20, 1/2, 1/5,$ and $1/4$.

**Proof.** For $\theta = (\theta_1, \theta_2) \in \mathbb{T}^2$, let $H_\lambda(\theta)$ denote the Floquet matrix corresponding to $H_\lambda$. Ordering the vertices of the fundamental domain as in Figure 7, we obtain

$$
H_\lambda(\theta) = \begin{bmatrix}
\lambda & 1 & 0 & e^{-i\theta_1} & 0 & e^{-i\theta_2} & 0 & 0 \\
1 & -\lambda & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & \lambda & 1 & 0 & 0 & 0 & e^{-i\theta_2} \\
e^{i\theta_1} & 0 & 1 & 2\lambda & 0 & 0 & 1 & 0 \\
e^{i\theta_2} & 0 & 0 & 0 & 1 & \lambda & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & \lambda & 1 \\
0 & 0 & e^{i\theta_2} & 0 & e^{i\theta_1} & 0 & 1 & -\lambda
\end{bmatrix}.
$$

(4.4)

First, let us consider the gaps at $E = \pm 1$. Calculations yield

$$
\det(H_\lambda(\theta) - (\pm 1 + s \lambda^2)I) = X_{\theta}^\pm(\theta) + X_{\theta}^\pm(\theta, s)\lambda^4 + X_{\theta}^\pm(\theta, s)\lambda^6 + O(\lambda^8),
$$

in which

$$
X_{\theta}^\pm(\theta) = -4(-\sin(\theta_1) + \sin(\theta_1 - \theta_2) + \sin(\theta_2))^2,
$$

$$
X_{\theta}^\pm(\theta, s) = 8(s \pm 1)(2s \mp 1)(3 - \cos(\theta_1) - \cos(\theta_1 - \theta_2) - \cos(\theta_2)),
$$

$$
X_{\theta}^\pm(\theta, s) = -1 \mp 12s + 72s^2 \mp 16s^3
- 4s^2(\pm 4s + 1)(\cos(\theta_1) + \cos(\theta_1 - \theta_2) + \cos(\theta_2)).
$$

It is clear that

$$
X_{\theta}^\pm(\theta) \leq 0 \quad \text{for all } \theta \in \mathbb{T}^2.
$$

(4.6)
Since \( \cos(\theta_1) + \cos(\theta_1 - \theta_2) + \cos(\theta_2) \leq 3 \), we also have
\[
\begin{align*}
X_4^\pm(\theta, s) &\leq 0 \quad \text{for all } \theta \in \mathbb{T}^2, \ |s| \leq 1/2. 
\end{align*}
\]
We also have for \(|s| \leq 1/4\),
\[
X_6^+(\theta, s) \leq -1 - 12s + 72s^2 - 16s^3 + 12s^2(4s + 1) =: T(s)
\]
and
\[
X_6^-(\theta, s) = X_6^+(\theta, -s) \leq T(-s).
\]
One easily checks that \( T(s) \) is decreasing on \([-0.05, 0.05]\), and
\[
T(-0.05) = -0.194.
\]
Hence for \(|s| \leq 0.05\),
\[
\begin{align*}
X_6^\pm(\theta, s) &\leq -0.194. 
\end{align*}
\]
Combining (4.6), (4.7), and (4.8), we obtain that for \(|\lambda| > 0\) sufficiently small, and \(|s| \leq 1/20\),
\[
\det(H_\lambda(\theta) - (\pm 1 + s\lambda^2)I)) \leq -0.1\lambda^6 < 0.
\]
This proves the claimed lower bound on the gaps at \( \pm 1 \).

On the other hand, let us note that \( X_0^\pm(0, 0) = X_0^\pm(\pi, \pi) = 0 \), while
\[
\begin{align*}
X_4^+(\theta, 0.5) &= 0 \quad \text{and} \quad X_4^+(\pi, \pi, 0.5) = 12, \\
X_4^+(0, 0, s) &= 0 \quad \text{and} \quad X_4^+(0, 0, -0.5) = 28, \\
X_4^-((0, 0), s) &= 0 \quad \text{and} \quad X_4^-((0, 0), 0.5) = 28, \\
X_4^-((0, 0), -0.5) &= 0 \quad \text{and} \quad X_6^-((\pi, \pi), -0.5) = 12.
\end{align*}
\]
Thus for small \( \lambda > 0 \), we have
\[
\begin{align*}
\det(H_\lambda(\pi, \pi) - (1 + 0.5\lambda^2)I)) &> 0, \\
\det(H_\lambda(0, 0) - (1 - 0.5\lambda^2)I)) &> 0, \\
\det(H_\lambda(0, 0) - (-1 + 0.5\lambda^2)I)) &> 0, \\
\det(H_\lambda(\pi, \pi) - (-1 - 0.5\lambda^2)I)) &> 0.
\end{align*}
\]
We also easily check that
\[
X_0^\pm(\frac{\pi}{2}, \pi) = -16,
\]
which implies that for small \( \lambda > 0 \) we have
\[
\det \left( H_\lambda \left( \frac{\pi}{2}, \pi \right) - (\pm 1 \pm 0.5\lambda^2)I)) \right) < 0.
\]
We therefore conclude that
\[
\pm 1 + 0.5\lambda^2 \in \sigma(H_\lambda) \quad \text{and} \quad \pm 1 - 0.5\lambda^2 \in \sigma(H_\lambda),
\]
which proves the upper bounds on the gaps at \( \pm 1 \).
Now let us consider the gap at $E = 0$. After calculations, we have

\begin{equation}
\det(H_\lambda(\theta) - s\lambda I) = Y_0(\theta) + Y_2(\theta, s)\lambda^2 + Y_4(\theta, s) + O(\lambda^6),
\end{equation}

where

\begin{align*}
Y_0(\theta) &= 15 + 2\cos(2\theta_1) - 4\cos(\theta_1 - 2\theta_2) + 2\cos(2\theta_1 - 2\theta_2) - 4\cos(2\theta_1 - \theta_2) \\
&\quad + 2\cos(2\theta_2) - 4\cos(\theta_1 + \theta_2), \\
Y_2(\theta, s) &= 2[5 - 26s^2 + (2 + 4s^2)(\cos(\theta_1) + \cos(\theta_1 - \theta_2) + \cos(\theta_2))],
\end{align*}

and

\begin{align*}
Y_4(\theta, s) &= (1 - s^2)[-3 - 42s^2 + 4(2 + s^2)(\cos(\theta_1) + \cos(\theta_1 - \theta_2) + \cos(\theta_2))].
\end{align*}

We claim that
\begin{equation}
Y_0(\theta) \geq 0 \quad \text{for all } \theta \in \mathbb{T}^2.
\end{equation}

Let us see how to use (4.10) to prove the claimed gap at zero and defer the proof of (4.10) for a moment. Using

\[
\cos(\theta_1) + \cos(\theta_1 - \theta_2) + \cos(\theta_2) \in \left[ -\frac{3}{2}, 3 \right],
\]

we obtain that for $|s| < 1/5$

\begin{equation}
Y_2(\theta, s) \geq 2(5 - 26s^2 - 3(1 + 2s^2)) = 4(1 - 16s^2) > \frac{36}{25}.
\end{equation}

Combining (4.9) with (4.11), we obtain that for $|\lambda| > 0$ sufficiently small

\[
\det(H_\lambda(\theta) - s\lambda I) > \lambda^2.
\]

This proves the claimed lower bound of the gap at 0, modulo the claim that $Y_0(\theta) \geq 0$ for all $\theta \in \mathbb{T}^2$.

To prove the upper bound, we compute

\[
\begin{cases}
Y_0\left(\frac{2\pi}{3}, \frac{4\pi}{3}\right) = 0, \\
Y_2\left(\left(\frac{2\pi}{3}, \frac{4\pi}{3}\right), s\right) = 4(1 - 16s^2), \\
Y_4\left(\left(\frac{2\pi}{3}, \frac{4\pi}{3}\right), s\right) = 3(s^2 - 1)(16s^2 + 5),
\end{cases}
\]

which implies that for small $\lambda > 0$,

\[
\det\left(H_\lambda\left(\frac{2\pi}{3}, \frac{4\pi}{3}\right) \pm 0.25\lambda I \right) < 0.
\]
We also compute that $Y_0(0,0) = 9$, which shows that for small $\lambda > 0$, 

$$\det(H_\lambda(0,0) \pm 0.25\lambda \mathbb{I}) > 0.$$ 

Thus we conclude that 

$$\pm 0.25\lambda \in \sigma(H_\lambda),$$

which proves the claimed upper bound of the gap at 0.

To complete the argument, all that remains is to show $Y_0(\theta) \geq 0$ for all $\theta \in \mathbb{T}^2$. To that end, introduce two auxiliary variables 

$$z := \cos\left(\frac{\theta_1 - \theta_2}{2}\right), \quad w := \cos\left(\frac{\theta_1 + \theta_2}{2}\right),$$

and write $g(z, w)$ to mean $Y_0(\theta)$ in the variables $z$ and $w$. Thus, to optimize $Y_0(\theta)$ on $\mathbb{T}^2$, it suffices to optimize $g(z, w)$ on the square $[-1, 1]^2$. To execute this change of variables, first note the following simple consequences of standard identities:

$$\cos(2\theta_1) + \cos(2\theta_2) = 2(2z^2 - 1)(2w^2 - 1),$$

$$\cos(2\theta_1 - 2\theta_2) = 2(2z^2 - 1) - 1,$$

$$\cos(\theta_1 + \theta_2) = 2w^2 - 1,$$

$$\cos(\theta_1 - 2\theta_2) + \cos(2\theta_1 - \theta_2) = 2zw(4z^2 - 3).$$

Putting all this together, 

$$g(z, w) = 15 + 4(2z^2 - 1)(2w^2 - 1) - 8zw(4z^2 - 3) + 2(2(2z^2 - 1)^2 - 1) - 4(2w^2 - 1).$$

It is easy to check that $g \geq 0$ holds on the boundary; concretely, 

$$g(\pm 1, w) = 15 + 4(2w^2 - 1) \mp 8w + 2 - 4(2w^2 - 1)$$

$$= 17 \mp 8w \geq 17 - 9 > 0$$

and 

$$g(z, \pm 1) = 15 + 4(2z^2 - 1) \mp 8z(4z^2 - 3) + (16z^4 - 16z^2 + 2) - 4$$

$$= 16z^4 + 32z^3 - 8z^2 \pm 24z + 9 = (3 \pm 4z - 4z^2)^2 \geq 0.$$ 

So, we now seek zeros of $\nabla g$ for $|z| < 1$ and $|w| < 1$. One easily computes $\partial_z g$ and $\partial_w g$:

$$\partial_z g = 8(w - 2z)(3 + 4z(w - z)),$$

$$\partial_w g = 8(3z - 4z^3 + 4w(z^2 - 1)).$$
Setting $\partial_w g = 0$ yields

\begin{equation}
\label{eq:4.12}
w = \frac{4z^3 - 3z}{4(z^2 - 1)}.
\end{equation}

Since we are working on the interior of $[-1, 1]^2$, $z \not\in \pm 1$ and the denominator does not vanish. Substituting this expression for $w$ into $\partial_z g$ and simplifying, we get

$$
\partial_z g \left( z, \frac{4z^3 - 3z}{z^2 - 1} \right) = 2z \left( \frac{1}{(z^2 - 1)^2} - 16 \right).
$$

Setting this equal to zero, we obtain three values of $z$ with $|z| < 1$: 0 and $\pm \sqrt{3}/2$. Inserting these $z$ values into (4.12), the corresponding $w$ values are all readily seen to be zero. Plugging in the three critical points $(0, 0)$ and $(\pm \sqrt{3}/2, 0)$ into $g$ yields 25 and 16, respectively, which concludes the proof that $g \geq 0$ and hence

$$
Y_0(\theta) \geq 0
$$

for all $\theta \in \mathbb{T}^2$, proving (4.10).

Next, we show that for any $(2, 2)$-periodic potential, it is impossible that it opens linear order gaps on both sides of $E = \pm 1$ simultaneously.

**Theorem 4.3.** For any $(2, 2)$-periodic potential $Q$ and any constant $c > 0$, the following holds for all sufficiently small $\lambda > 0$:

$$
((-1 - c\lambda, -1 + c\lambda) \cup (1 - c\lambda, 1 + c\lambda)) \cap \sigma(H_\lambda) \neq \emptyset.
$$

**Proof.** Let $(q_1, q_2, \ldots, q_8)$ be the potential on a $2 \times 2$ fundamental cell, as shown in Figure 7. The corresponding Floquet matrix $H_\lambda(\theta)$ is

$$
H_\lambda(\theta) =
\begin{bmatrix}
\lambda q_1 & 1 & 0 & e^{-i\theta_1} & 0 & e^{-i\theta_2} & 0 & 0 \\
1 & \lambda q_2 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & \lambda q_3 & 1 & 0 & 0 & 0 & e^{-i\theta_2} \\
e^{i\theta_1} & 0 & 1 & \lambda q_4 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & \lambda q_5 & 1 & 0 & e^{-i\theta_1} \\
e^{i\theta_2} & 0 & 0 & 0 & 1 & \lambda q_6 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & \lambda q_7 & 1 \\
0 & 0 & e^{i\theta_2} & 0 & e^{i\theta_1} & 0 & 1 & \lambda q_8
\end{bmatrix}.
$$

For $0 < |s| < c$, let us consider

$$
\det(H_\lambda(\theta) - (\pm 1 + s\lambda)I) = \sum_{k=0}^{8} X_k^\pm(\theta, s)\lambda^k.
$$
After a calculation, we obtain

\begin{equation}
X_0^+(0, s) = X_1^+(0, s) = X_2^+(0, s) = 0 \quad \text{for all } s
\end{equation}

and

\begin{equation}
X_3^+(0, s) = -X_3^-(0, s) = a_0 + a_2 s^2 + 64 s^3,
\end{equation}

where

\[
a_0 = -2[(q_1 + q_2 + q_7 + q_8)(q_4 + q_5)(q_3 + q_6) + (q_1 + q_8)(q_2 + q_7)(q_3 + q_4 + q_5 + q_6)]
+ 8[(q_1 + q_8)(q_2 + q_7) + (q_3 + q_6)(q_4 + q_5) + (q_1 + q_2 + q_7 + q_8)(q_3 + q_4 + q_5 + q_6)],
\]
\[
a_2 = -24 \sum_{k=1}^{8} q_k.
\]

By (4.14), we have

\[
X_3^+(0, s_0) = -X_3^-(0, s_0) \neq 0
\]

for some \( s_0 \) such that \( 0 < |s_0| < c \). Without loss of generality, we assume

\[
X_3^+(0, s_0) > 0 > X_3^-(0, s_0).
\]

Combining this with (4.13), we obtain

\begin{equation}
\det(H_\lambda(0) - (1 + s_0 \lambda)I) > 0
\end{equation}

for small \( \lambda > 0 \). We also have

\begin{equation}
X_0^+(\pi/4, 3\pi/4, s_0) = -4.
\end{equation}

In particular, (4.16) implies that

\begin{equation}
\det(H_\lambda(\pi/4, 3\pi/4)) - (1 + s_0 \lambda)I < 0
\end{equation}

for all \( \lambda \geq 0 \) small.

Combining (4.17) with (4.15), for any sufficiently small \( \lambda > 0 \), there exists \( \theta \) such that

\[
\det(H_\lambda(\theta) - (1 + s_0 \lambda)I) = 0.
\]

Hence

\[
(1 - c\lambda, 1 + c\lambda) \cap \sigma(H_\lambda) \neq \emptyset
\]

as claimed. \( \Box \)
5 Square Laplacian with next-nearest-neighbor interactions

We now turn our attention to the EHM lattice, whose Laplacian is given by

\[
\begin{align*}
\Delta_{sqn} u_{n,m} & = u_{n-1,m} + u_{n+1,m} + u_{n,m-1} + u_{n,m+1} + u_{n-1,m+1} \\
& \quad + u_{n+1,m-1} + u_{n+1,m+1} \\
& = \Delta_{sqn} u_{n,m} + u_{n-1,m-1} + u_{n-1,m+1} + u_{n+1,m-1} + u_{n+1,m+1} \\
& = \Delta_{tri} u_{n,m} + u_{n-1,m-1} + u_{n+1,m+1}.
\end{align*}
\]

Now, given \( p_1, p_2 \in \mathbb{Z}_+ \), we define \( P = p_1 p_2 \) and \( \Lambda = \mathbb{Z}^2 \cap ([0, p_1) \times [0, p_2)) \) as before and view \( \Delta_{sqn} \) as a \((p_1, p_2)\)-periodic operator and perform the Floquet decomposition. For \( \theta = (\theta_1, \theta_2) \in \mathbb{R}^2 \), it is straightforward to check that

\[
\sigma(H(\theta)) = \{ e_\ell(\theta) : \ell \in \Lambda \},
\]

where \( \ell = (\ell_1, \ell_2) \) and

\[
e_\ell(\theta) = 2 \cos \left( \frac{\theta_1 + 2\pi \ell_1}{p_1} \right) + 2 \cos \left( \frac{\theta_2 + 2\pi \ell_2}{p_2} \right) + 2 \cos \left( \frac{\theta_1 + 2\pi \ell_1}{p_1} - \frac{\theta_2 + 2\pi \ell_2}{p_2} \right) + 2 \cos \left( \frac{\theta_1 + 2\pi \ell_1}{p_1} + \frac{\theta_2 + 2\pi \ell_2}{p_2} \right).
\]

As in Section 2, we label these eigenvalues in increasing order according to multiplicity by

\[
E_1(\theta) \leq E_2(\theta) \leq \cdots \leq E_P(\theta)
\]

and denote the \( P \) spectral bands by

\[
F_k = \{ E_k(\theta) : \theta \in \mathbb{R}^2 \}, \quad 1 \leq k \leq P.
\]

Straightforward computation shows that \( \sigma(\Delta_{sqn}) = [-4, 8] \), hence

\[
\bigcup_{k=1}^P F_k = [-4, 8].
\]

Our main theorem of this section is

**Theorem 5.1.** Let \( p_1, p_2 \in \mathbb{Z}_+ \) be given.

(1) Each \( E \in (-4, 8) \setminus \{-1\} \) belongs to \( \text{int}(F_k) \) for some \( 1 \leq k \leq P \).

(2) If one of the periods \( p_1, p_2 \) is not divisible by three, then \( E = -1 \) belongs to \( \text{int}(F_k) \) for some \( 1 \leq k \leq P \).

**Proof of Theorem 1.5.** This follows immediately from Theorem 5.1. \( \square \)
5.1 Proof of Theorem 5.1. As with the proof of Theorem 3.1, we will divide the proof into two different cases: \(E \neq -1\) and \(E = -1\) and argue by contradiction. To that end, assume for the sake of establishing a contradiction that

\[
E = \min F_{k+1} = \max F_k \quad \text{for some } 1 \leq k \leq P - 1.
\]

We will use the following lemmas, whose proofs we provide at the end of the present section.

**Lemma 5.2.** Let us consider the following system:

\[
\begin{align*}
\cos(x) + \cos(y) + \cos(x - y) + \cos(x + y) &= \frac{E}{2}, \\
\sin(x) + \sin(x - y) + \sin(x + y) &= 0.
\end{align*}
\]

(5.1)

For any \(E \in (-4, 8) \setminus \{-1\}\), the solution set of (5.1) in \([0, 2\pi]^2\) satisfies

\[
(5.2) \quad x = 0, \quad 1 + 2 \cos(y) = \frac{E + 1}{3},
\]

or

\[
(5.3) \quad x = \pi, \quad 1 + 2 \cos(y) = -(E + 1).
\]

**Lemma 5.3.** Consider the following system:

\[
\begin{align*}
\cos(x) + \cos(y) + \cos(x + y) + \cos(x - y) &= \frac{E}{2}, \\
\sin(x) + \sin(x - y) + \sin(x + y) &= 0, \\
\sin(y) - \sin(x - y) + \sin(x + y) &= 0.
\end{align*}
\]

(5.4)

For any \(E \in (-4, 8) \setminus \{0, -1\}\), the solution set of (5.4) is empty. For \(E = 0\), the unique solution of (5.4) in \([0, 2\pi]^2\) is \((\pi, \pi)\). For \(E = -1\), the solutions of (5.4) in \([0, 2\pi]^2\) are \((2\pi/3, 2\pi/3), (2\pi/3, 4\pi/3), (4\pi/3, 2\pi/3)\) and \((4\pi/3, 4\pi/3)\).

We will use Lemma 5.2 in the \(E \neq -1\) case, and Lemma 5.3 in the \(E = -1\) case.

5.1.1 \(E \neq -1\).

**Proof of Theorem 5.1.1.** Let \(E \in (-4, 8) \setminus \{-1\}\) be given. Define

\[
\begin{align*}
&\tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2) \in [0, 2\pi)^2 \quad \text{and} \quad \ell^{(1)} = (\ell^{(1)}_1, \ell^{(1)}_2) \in \Lambda \\
&\text{via}
\end{align*}
\]

\[
\begin{align*}
\tilde{\theta}_1 &= 0, \quad \ell^{(1)}_1 = 0, \quad \frac{\tilde{\theta}_2 + 2\pi\ell^{(1)}_2}{p_2} = \arccos \left( \frac{E - 2}{6} \right) \in (0, \pi).
\end{align*}
\]

(5.5)
Note that since $E \in (-4, 8)$, we have $\frac{E-2}{6} \in (-1, 1)$, hence $\arccos\left(\frac{E-2}{6}\right)$ is always well-defined. Note also that $\tilde{\theta_2}$ and $\ell^{(1)}_2$ are uniquely determined. Using (5.5), one easily checks that

$$e_{\ell^{(1)}}(\tilde{\theta}) = E,$$

and

$$\text{(5.6) } (1, 0) \cdot \nabla e_{\ell^{(1)}}(\tilde{\theta}) = 0.$$ 

As in the proof of Theorem 3.1, denote $\Lambda_E(\tilde{\theta}) = \{ \ell \in \Lambda : e_\ell(\tilde{\theta}) = E \}$, let $r := |\Lambda_E(\tilde{\theta})|$ be the multiplicity of $E$ as an eigenvalue of $H(\tilde{\theta})$, and choose $s \in \mathbb{Z} \cap [1, r]$ such that

$$E_{k-s}(\tilde{\theta}) < E_{k-s+1}(\tilde{\theta}) = \cdots = E_k(\tilde{\theta}) = \cdots = E_{k+r-s}(\tilde{\theta}) < E_{k+r-s+1}(\tilde{\theta}).$$

Since all the eigenvalues are continuous in $\theta$, we can take $\varepsilon > 0$ small enough such that

$$E_{k-s}(\theta) < E_{k-s+1}(\theta), \text{ and } E_{k+r-s}(\theta) < E_{k+r-s+1}(\theta)$$

hold whenever $\|\theta - \tilde{\theta}\|_{\mathbb{R}^2} < \varepsilon$. Given $\ell \in \Lambda$ and a unit vector $\beta = (\beta_1, \beta_2)$, we have

$$\text{(5.7) } e_{\ell}(\tilde{\theta} + t\beta) = e_{\ell}(\tilde{\theta}) + t\beta \cdot \nabla e_{\ell}(\tilde{\theta}) + O(t^2)$$

$$\text{(5.8) } = e_{\ell}(\tilde{\theta}) + t\beta \cdot \nabla e_{\ell}(\tilde{\theta})$$

$$- \frac{r^2}{2} \left( \frac{\beta_1^2}{p_1} \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) + \frac{\beta_2^2}{p_2} \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right)$$

$$+ \left( \frac{\beta_1}{p_1} - \frac{\beta_2}{p_2} \right)^2 \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} - \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right)$$

$$+ \left( \frac{\beta_1}{p_1} + \frac{\beta_2}{p_2} \right)^2 \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} + \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right] + O(t^3).$$

In particular, we will use (5.7) if $\beta \cdot \nabla e_{\ell}(\tilde{\theta}) \neq 0$ and (5.8) otherwise.

For any vector $\beta \in \mathbb{R}^2 \setminus \{0\}$, let

$$\mathcal{J}_0^0 = \mathcal{J}_0^0(\tilde{\theta}) := \{ \ell \in \Lambda_E(\tilde{\theta}) : \beta \cdot \nabla e_{\ell}(\tilde{\theta}) = 0 \},$$

$$\mathcal{J}_0^\pm = \mathcal{J}_0^\pm(\tilde{\theta}) := \{ \ell \in \Lambda_E(\tilde{\theta}) : \pm \beta \cdot \nabla e_{\ell}(\tilde{\theta}) > 0 \}.$$ 

By definition, we must have

$$\text{(5.10) } |\mathcal{J}_0^0| + |\mathcal{J}_0^+| + |\mathcal{J}_0^-| = r$$

for any $\beta$. We also define $\mathcal{J}_0$ as follows:

$$\mathcal{J}_0 = \mathcal{J}_0(\tilde{\theta}) := \{ \ell \in \Lambda_E(\tilde{\theta}) : \nabla e_{\ell}(\tilde{\theta}) = 0 \}.$$
If $E \neq 0$, Lemma 5.3 directly implies $\mathcal{J}_0 = \emptyset$. If $E = 0$, $\mathcal{J}_0$ is also empty. To see this, suppose on the contrary that $\ell = (\ell_1, \ell_2) \in \mathcal{J}_0$. Lemma 5.3 implies that

\begin{equation}
\frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} = \pi,
\end{equation}

and (5.5) forces

\begin{equation}
\frac{\tilde{\theta}_2 + 2\pi \ell_2^{(1)}}{p_2} = \arccos \left( -\frac{1}{3} \right).
\end{equation}

Subtracting (5.12) from (5.13) yields

\[\frac{\ell_2^{(1)} - \ell_2}{p_2} = \frac{1}{2\pi} \arccos \left( -\frac{1}{3} \right) - \frac{1}{2}.\]

However, this implies that $(2\pi)^{-1} \arccos(-1/3)$ is a rational number, which contradicts the following well-known fact, whose proof we supply at the end of the present section.

**Lemma 5.4.**

\[\frac{1}{2\pi} \arccos \left( -\frac{1}{3} \right) \in \mathbb{R} \setminus \mathbb{Q}.\]

Therefore $\mathcal{J}_0 = \emptyset$ for any $E \neq -1$.

We choose $\beta_1 = (1, 0)$. Then (5.6) implies $\ell^{(1)} \in \mathcal{J}^0_{\beta_1}$, and hence

\begin{equation}
\mathcal{J}^0_{\beta_1} \neq \emptyset.
\end{equation}

Next we are going to perturb the point $\tilde{\theta}$ and count the eigenvalues. Since $\mathcal{J}_0 = \emptyset$, we can choose a unit vector $\beta_2$ such that

\begin{equation}
\beta_2 \cdot \nabla e_\ell(\tilde{\theta}) \neq 0
\end{equation}

holds for any $\ell \in \Lambda_E(\tilde{\theta})$. Thus $\mathcal{J}^0_{\beta_2} = \emptyset$ and

\begin{equation}
|\mathcal{J}^+_{\beta_2}| + |\mathcal{J}^-_{\beta_2}| = r.
\end{equation}

Arguing as in the proof of Theorem 3.1.1, we deduce that

\begin{equation}
r = 2s.
\end{equation}

**Perturbation along $\beta_1$.** Now we perturb the eigenvalues along $\beta_1 = (1, 0)$. The case when $\ell \in \mathcal{J}^0_{\beta_1}$ is similar to that of $\beta_2$. The difference here is that, according to (5.14), $\mathcal{J}^0_{\beta_1} \neq \emptyset$. 

By Lemma 5.2, we have that for \((\ell_1, \ell_2) \in \mathcal{J}_{\beta_1}^0\),

\[
(E + 1) \left[ \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) + \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} - \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right. \\
+ \left. \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} + \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right] > 0.
\]

(5.18)

Indeed, if \((\ell_1, \ell_2) \in \mathcal{J}_{\beta_1}^0\), \((x, y) = (p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1), p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2))\) is a solution to (5.1). Hence Lemma 5.2 implies that we have either

\[
\frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} = 0, \quad 1 + 2 \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) = \frac{E + 1}{3},
\]

or

\[
\frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} = \pi, \quad 1 + 2 \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) = -(E + 1).
\]

Clearly, both cases lead to (5.18).

By employing (5.8), we obtain

\[
e^\ell(\tilde{\theta} + t\beta_1) = E - \frac{t^2}{2p_1^2} \left[ \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) + \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} - \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right. \\
+ \left. \cos \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} + \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right] + O(t^3)
\]

for \(\ell \in \mathcal{J}_{\beta_1}^0\). Combining this with (5.18), we obtain that for \(|t| > 0\) small enough

\[
e^\ell(\tilde{\theta} + t\beta_1) \begin{cases} < E, & \text{if } E + 1 > 0, \\ > E, & \text{if } E + 1 < 0. \end{cases}
\]

(5.20)

Notice that the choice of \(\beta_1\) causes the second \(t^2\) term of (5.8) to drop out.

Without loss of generality, we assume \(E \in (-1, 8)\). The complementary case when \(E \in (-4, -1)\) can be handled similarly. For \(E \in (-1, 8)\), (5.20) implies that

\[
e^\ell(\tilde{\theta} + t\beta_1) < E = \min F_{k+1}
\]

(5.21)

holds for \(|t| > 0\) small enough and for any \(\ell \in \mathcal{J}_{\beta_1}^0\).

Combining (5.21) with (5.7), we have the following:

For \(t > 0\) small enough:

- If \(\ell \in \mathcal{J}_{\beta_1}^0\), we have

\[
E_{k+r-s+1}(\tilde{\theta} + t\beta_1) > e^\ell(\tilde{\theta} + t\beta_1) > E = \max F_k,
\]
which implies

\[(5.22) \quad |J^\pm_i| \leq r - s = s,\]

where the equality follows from (5.17).

- If \( \ell \in J^0_\beta \cup J^\mp_\beta \), we have

\[E_{k-s-1}(\tilde{\theta} + t\beta_1) < e_{\ell}(\tilde{\theta} + t\beta_1) < E = \min F_{k+1},\]

which implies

\[(5.23) \quad |J^0_\beta| + |J^\mp_\beta| \leq s.\]

In view of (5.10) and (5.17), equations (5.22) and (5.23) yield

\[(5.24) \quad |J^+_\beta| = |J^0_\beta| + |J^-_\beta| = s.\]

As before, we may observe that \( J^0_\beta = J^\mp_\beta \) and \( J^\pm_\beta = J^\mp_\beta \). Then, the analysis above applied with \( \beta_1 \) replaced by \( -\beta_1 \) forces

\[(5.25) \quad |J^-_\beta| = |J^0_\beta| + |J^+_\beta| = s.\]

Taken together, (5.24) and (5.25) imply \( |J^0_\beta| = 0 \), which contradicts (5.14). \( \square \)

### 5.1.2 \( E = -1 \). First, we would like to make a remark on our strategy of the proof of the \( E = -1 \) case, and on the importance of one of the periods being not divisible by 3.

**Remark 5.5.** For the exceptional energy \( E = -1 \) of the EHM lattice, we cannot use eigenvalues with vanishing gradients to create un-even eigenvalue counts unless neither \( p_1 \) nor \( p_2 \) is divisible by 3. The reason is the following: suppose only \( p_1 \) is not divisible by 3 and we choose \( \tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2) \) and \( \ell^{(1)} = (\ell_1^{(1)}, \ell_2^{(1)}) \) such that \( e_{\ell_1^{(1)}}(\tilde{\theta}) = -1 \) and \( \nabla e_{\ell_1^{(1)}}(\tilde{\theta}) = 0 \). Lemma 5.3 yields four possibilities: \( (p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1^{(1)}), p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2^{(1)})) = (2\pi/3, 2\pi/3), (2\pi/3, 4\pi/3), (4\pi/3, 2\pi/3) \) or \( (4\pi/3, 4\pi/3) \). Without loss of generality, we choose \( (2\pi/3, 2\pi/3) \); the other three choices are essentially the same. Since \( p_2 \) is divisible by 3, there exists \( \ell^{(2)} \), such that \( (p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1^{(2)}), p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2^{(2)})) = (2\pi/3, 4\pi/3) \). Hence \( e_{\ell_2^{(1)}}(\tilde{\theta}) \) is also located at \( -1 \) with vanishing gradient. Perturbing \( e_{\ell_1^{(1)}}(\tilde{\theta}) \) and \( e_{\ell_2^{(1)}}(\tilde{\theta}) \) along a given direction \( (\beta_1, \beta_2) \) is equivalent to controlling the signs of the following two expressions:

\[\beta_1 \beta_2 \quad \text{and} \quad -\beta_1 \beta_2.\]
This means we can never choose two different directions that lead to un-even counts. Therefore we need to develop a new argument for this case.

Indeed, when \( p_1 \) is not divisible by 3, we choose \( p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1^{(1)}) = 2\pi/3 \) and \( \tilde{\theta}_2 \) such that \( p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2) \notin \{2\pi/3, 4\pi/3\} \) regardless of the choice of \( \ell_2 \). Such choices guarantee that there are in total \( p_2 \) eigenvalues located at \(-1\), which are \( \{e_\ell(\tilde{\theta}), \ell_1 = \ell_1^{(1)}\} \). It then suffices to control the movements of these eigenvalues along any given direction. A key observation is that along any direction, approximately \( 2p_2/3 \) eigenvalues will move up (down) while the other \( p_2/3 \) eigenvalues move down (up); see (5.33). This leads to un-even counting that we need. Let us point out that if both \( p_1, p_2 \) are divisible by 3, this argument does not work (as it must, given the example constructed in Theorem 1.6): there will be \( 2p_2 \) eigenvalues located at \(-1\), and \( p_2 \) of them move up while the other \( p_2 \) of them move down along any given direction.

**Proof of Theorem 5.1.2.** Without loss of generality, we assume \( p_1 \) is not divisible by 3. Let \( p_j = 3p'_j + k_j \), where \( p'_j, k_j \in \mathbb{Z} \) with \( 0 \leq k_j < 3 \) and then define \( \tilde{\theta} \) by

\[
\tilde{\theta}_1 = \frac{2\pi k_1}{3}, \quad \tilde{\theta}_2 = \frac{k_2 + 1}{4} \pi.
\]

As usual, denote \( \Lambda_E(\tilde{\theta}) = \Lambda_{-1}(\tilde{\theta}) = \{\ell \in \Lambda : e_\ell(\tilde{\theta}) = -1\} \). We first claim that

\[
\Lambda_{-1}(\tilde{\theta}) = \{(p'_1, \ell_2) : 0 \leq \ell_2 < p_2 \text{ and } \ell_2 \in \mathbb{Z}\}.
\]

Let us consider the trigonometric equation

\[
\cos(x) + \cos(y) + \cos(x - y) + \cos(x + y) = -\frac{1}{2} = E_2.
\]

Using the identity \( \cos(x - y) + \cos(x + y) = 2 \cos(x) \cos(y) \), we see that \( (5.27) \) is equivalent to

\[
(2 \cos(x) + 1)(2 \cos(y) + 1) = 0,
\]

whose solutions are \( \cos(x) = -1/2 \) or \( \cos(y) = -1/2 \). With our choice of \( \tilde{\theta} \), it is clear that

\[
\frac{\tilde{\theta}_1 + 2\pi p'_1}{p_1} = \frac{2\pi}{3}, \quad \cos\left(\frac{\tilde{\theta}_1 + 2\pi p'_1}{p_1}\right) = -\frac{1}{2}.
\]

Consequently,

\[
(5.29) \quad e_{(p'_1, \ell_2)}(\tilde{\theta}) = -1 \quad \text{for every } 0 \leq \ell_2 < p_2.
\]

Due to our choice of \( \tilde{\theta}_2 \), we get

\[
(5.30) \quad \cos(p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2)) \neq -\frac{1}{2} \quad \text{for any } \ell_2 \in [0, p_2) \cap \mathbb{Z}.
\]
Indeed, since \( p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2) \in [0, 2\pi) \), \( \cos(p_2^{-1}(\tilde{\theta}_2 + 2\pi \ell_2)) = -1/2 \) would force
\[
\frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \in \left\{ \frac{2\pi}{3}, \frac{4\pi}{3} \right\},
\]
which, after doing some algebra, leads to
\[
3(8\ell_2 + k_2 + 1) \in \{8p_2, 16p_2\},
\]
which is plainly impossible, since \( \ell_2, p_2 \in \mathbb{Z} \) and \( k_2 \in \{0, 1, 2\} \). Additionally, due to our choice of \( \tilde{\theta}_1 \), we also have
\[
(5.31) \quad \cos(p_1^{-1}(\tilde{\theta}_1 + 2\pi \ell_1)) \neq -1/2 \quad \text{for any} \quad \ell_1 \in ((0, p_1) \cap \mathbb{Z}) \setminus \{p'_1\}.
\]
To see this, suppose on the contrary that (5.31) fails. This forces
\[
\frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} = \frac{4\pi}{3}
\]
for some \( 0 \leq \ell_1 < p_1 \) with \( \ell_1 \neq p'_1 \). Since
\[
\frac{\tilde{\theta}_1 + 2\pi p'_1}{p_1} = \frac{2\pi}{3},
\]
this implies
\[
\frac{2\pi(\ell_1 - p'_1)}{p_1} = \frac{2\pi}{3},
\]
which is impossible since \( p_1 \) is not divisible by 3. Combining (5.30) and (5.31) yields
\[
(5.32) \quad e_\ell(\tilde{\theta}) \neq -1 \quad \text{for any} \quad \ell = (\ell_1, \ell_2) \in \Lambda \text{ such that } \ell_1 \neq p'_1.
\]
Taken together, (5.29) and (5.32) imply (5.26).

Let us choose \( \beta = (\beta_1, \beta_2) = (1, 0) \). We have that, for any \( \ell \in \Lambda \),
\[
p_1 \beta \cdot \nabla e_\ell(\tilde{\theta}) = -\sin \left( \frac{\hat{\theta}_1 + 2\pi \ell_1}{p_1} \right) - \sin \left( \frac{\hat{\theta}_1 + 2\pi \ell_1}{p_1} - \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) - \sin \left( \frac{\hat{\theta}_1 + 2\pi \ell_1}{p_1} + \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) = -\sin \left( \frac{\hat{\theta}_1 + 2\pi \ell_1}{p_1} \right) \left[ 1 + 2 \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \right].
\]
By (5.26), (5.28), and (5.30), we have the following for any \( \ell = (\ell_1, \ell_2) \in \Lambda_{-1}(\tilde{\theta}) \):
\[
\sin \left( \frac{\tilde{\theta}_1 + 2\pi \ell_1}{p_1} \right) = \frac{\sqrt{3}}{2}, \quad \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) \neq -\frac{1}{2}.
\]
This implies
\[
(5.33) \quad \mathcal{J}^0_\beta = \emptyset \quad \text{and} \quad \mathcal{J}^\pm_\beta = \left\{ \ell \in \Lambda_{-1}(\tilde{\theta}) : \mp \frac{1}{2} \mp \cos \left( \frac{\tilde{\theta}_2 + 2\pi \ell_2}{p_2} \right) > 0 \right\}.
\]
Hence we expect that \( |\mathcal{J}^+_\beta| \sim p_2/3 \) and \( |\mathcal{J}^-_\beta| \sim 2p_2/3 \). More precisely, we note that
\[
\mathcal{J}^+_\beta = \left\{ (p'_1, \ell_2) : \frac{2\pi}{3} < \frac{(k_2 + 1)\pi/4 + 2\pi \ell_2}{p_2} < \frac{4\pi}{3} \right\}.
\]
Using \( p_2 = 3p'_2 + k_2 \), we obtain
\[
\mathcal{J}^+_\beta = \left\{ (p'_1, \ell_2) : p'_2 + \frac{5k_2 - 3}{24} < \ell_2 < 2p'_2 + \frac{13k_2 - 3}{24} \right\}.
\]
Consequently,
\[
\mathcal{J}^-_\beta = \begin{cases} 
(p'_1, \ell_2) : p'_2 \leq \ell_2 \leq 2p'_2 - 1, & \text{if } k_2 = 0, \\
(p'_1, \ell_2) : p'_2 + 1 \leq \ell_2 \leq 2p'_2, & \text{if } k_2 = 1, 2.
\end{cases}
\]
Therefore
\[
(5.34) \quad (|\mathcal{J}^+_\beta|, |\mathcal{J}^-_\beta|) = \begin{cases} 
(p'_2, 2p'_2), & \text{if } k_2 = 0, \\
(p'_2, 2p'_2 + 1), & \text{if } k_2 = 1, \\
(p'_2, 2p'_2 + 2), & \text{if } k_2 = 2.
\end{cases}
\]
Note that \( p'_2 \geq 1 \) whenever \( k_2 = 0 \). Thus, a direct consequence of (5.34) is
\[
(5.35) \quad |\mathcal{J}^+_\beta| \neq |\mathcal{J}^-_\beta|.
\]
On the other hand, since \( \mathcal{J}^0_\beta = \emptyset \), following the same argument as in the proof of Theorem 3.1.1 yields \( |\mathcal{J}^+_\beta| = |\mathcal{J}^-_\beta| \), which contradicts (5.35).

\section*{5.2 Proofs of Lemmas 5.2, 5.3, and 5.4.}

\textbf{Proof of Lemma 5.2.} Let \( x \) and \( y \) solve (5.1) with \( E \neq -1 \). The second condition therein yields
\[
\sin(x) + 2\sin(x)\cos(y) = 0,
\]
leading to two possibilities: \( \sin(x) = 0 \) or \( \cos(y) = -1/2 \). If \( \sin(x) = 0 \), we get \( x = 0 \) or \( x = \pi \), which yields (5.2) and (5.3) upon plugging into the first condition in (5.1). In the event that \( \cos(y) = -1/2 \), we arrive at

\[
\cos(x) + \cos(y) + \cos(x - y) + \cos(x + y) = \cos(x) + \cos(y) + 2 \cos(x) \cos(y)
\]

\[
= \cos(x) - \frac{1}{2} - \cos(x) = -\frac{1}{2},
\]

in contradiction with \( E \neq -1 \). □

**Proof of Lemma 5.3.** Suppose \( x \) and \( y \) satisfy (5.4). From the proof of Lemma 5.2, the second condition of (5.4) implies

\[
\sin(x) = 0 \quad \text{or} \quad \cos(y) = -1/2.
\]

Thus, \( x = 0, x = \pi, y = 2\pi/3, \) or \( y = 4\pi/3 \). When \( \sin(x) = 0 \), the third condition of (5.4) forces \( \sin(y) = 0 \). The four points so obtained yield \( E = 8 \) when \( (x, y) = (0, 0) \), \( E = -4 \) when \( (x, y) = (0, \pi), (\pi, 0) \) and \( E = 0 \) when \( (x, y) = (\pi, \pi) \). Alternatively, when \( \cos(y) = -1/2 \), the third condition of (5.4) yields \( \cos(x) = -1/2 \), which implies \( x = 2\pi/3 \) or \( x = 4\pi/3 \). As in the proof of Lemma 5.2, the four points corresponding to

\[
x, y \in \left\{ \frac{2\pi}{3}, \frac{4\pi}{3} \right\}
\]

all yield \( E = -1 \).

**Proof of Lemma 5.4.** Suppose

(5.36) \[ \cos \left( \frac{2\pi m}{n} \right) = -\frac{1}{3}, \]

for \( m/n \in \mathbb{Q} \). Let \( T_n(\cdot) \) denote the \( n \)-th degree Chebyshev polynomial so that

(5.37) \[ T_n \left( \cos \left( \frac{2\pi m}{n} \right) \right) = \cos(2\pi m) = 1. \]

It is well-known that

\[
T_n(x) = \sum_{k=0}^{n} a_k x^k,
\]

where \( a_n = 2^{n-1} \) and \( a_k \in \mathbb{Z} \) for any \( k \). Hence (5.36) and (5.37) imply

\[
2^{n-1} \left( -\frac{1}{3} \right)^n + \sum_{k=0}^{n-1} a_k \left( -\frac{1}{3} \right)^k = 1.
\]

Multiplying by \( (-3)^n \) on both sides of the equation, we obtain

\[
2^{n-1} - 3 \sum_{k=0}^{n-1} a_k (-3)^{n-k-1} = (-3)^n,
\]

which implies \( 2^{n-1} \) is divisible by 3. Contradiction. □
5.3 Opening a gap at $-1$.

**Theorem 5.6.** Enumerate the vertices of a $3 \times 3$ fundamental cell of the square lattice as in Figure 8, denote $r = \sqrt{4 - \sqrt{15}}$, define a $(3, 3)$-periodic potential $Q$ on $\mathbb{Z}^2$ via

$$(q_1, \ldots, q_9) = \left(-r - \frac{1}{r} + 2, -r, -r + \frac{1}{r} - 2, -\frac{1}{r}, 0, +\frac{1}{r}, r - \frac{1}{r} - 2, r, r + \frac{1}{r} + 2 \right),$$

and denote $H_\lambda = \Delta_{\text{sqn}} + \lambda Q$. Then, for all $\lambda > 0$ sufficiently small, $\sigma(H_\lambda)$ consists of two connected components. Moreover, if $g_\lambda$ denotes the gap that opens at energy $-1$, one has

$$\left(-1 - \frac{\lambda}{10}, -1 + \frac{\lambda}{10}\right) \subseteq g_\lambda \subseteq \left(-1 - \frac{\lambda}{4}, -1 + \frac{\lambda}{4}\right).$$

In particular, the gap opens linearly.

Let us observe that the proof below can be refined a bit to yield sharper constants than $1/10$ and $1/4$. 

Figure 8. A $3 \times 3$ potential on the square lattice that opens a gap at $E = -1$ with small positive coupling (color figure available online).
Proof. For $\theta = (\theta_1, \theta_2) \in \mathbb{T}^2$, let $H_{\lambda}(\theta)$ denote the Floquet matrix corresponding to $H$. Ordering the vertices of the fundamental domain as in Figure 8, we obtain:

$$H_{\lambda}(\theta) = \begin{bmatrix}
\lambda q_1 & 1 & e^{-i\theta_1} & 1 & 1 & e^{-i\theta_2} & e^{-i\theta_2} & e^{-i(\theta_1+\theta_2)} \\
1 & \lambda q_2 & 1 & 1 & 1 & e^{-i\theta_1} & e^{-i\theta_2} & e^{-i\theta_2} \\
e^{i\theta_1} & 1 & \lambda q_3 & e^{i\theta_1} & 1 & e^{i(\theta_1-\theta_2)} & e^{-i\theta_2} & e^{-i\theta_2} \\
1 & 1 & e^{-i\theta_1} & \lambda q_4 & 1 & e^{-i\theta_1} & 1 & 1 & e^{-i\theta_1} \\
1 & 1 & 1 & 1 & \lambda q_5 & 1 & 1 & 1 & 1 \\
e^{i\theta_2} & e^{i\theta_2} & e^{-i(\theta_1-\theta_2)} & 1 & 1 & e^{-i\theta_1} & \lambda q_7 & 1 & e^{-i\theta_1} \\
e^{i\theta_2} & e^{i\theta_2} & e^{i\theta_2} & 1 & 1 & 1 & \lambda q_8 & 1 & \\
[e^{i(\theta_1+\theta_2)}] & e^{i\theta_2} & e^{i\theta_2} & e^{i\theta_1} & 1 & 1 & e^{i\theta_1} & 1 & \lambda q_9 
\end{bmatrix}.$$

For $s \in (-1, 1)$, let us consider

$$\det(H_{\lambda}(\theta) + (1 + s\lambda)I) = \sum_{k=0}^{9} X_k(\theta, s)\lambda^k.$$ 

Our goal is to show $\det(H_{\lambda}(\theta) + (1 + s\lambda)I)$ never vanishes for sufficiently small $\lambda > 0$ and for $|s| < 0.1$. Direct computations yield

$$X_0(\theta, s) = 4096 \sin^6\left(\frac{\theta_1}{2}\right) \sin^6\left(\frac{\theta_2}{2}\right),$$

$$X_1(\theta, s) = 0,$$

$$X_2(\theta, s) = Y_2(s) \sin^4\left(\frac{\theta_1}{2}\right) \sin^4\left(\frac{\theta_2}{2}\right),$$

$$X_3(\theta, s) = Y_3(s) \sin^4\left(\frac{\theta_1}{2}\right) \sin^4\left(\frac{\theta_2}{2}\right),$$

$$X_4(\theta, s) = Y_4(s) \sin^2\left(\frac{\theta_1}{2}\right) \sin^2\left(\frac{\theta_2}{2}\right),$$

$$X_5(\theta, s) = Y_5(s) \sin^2\left(\frac{\theta_1}{2}\right) \sin^2\left(\frac{\theta_2}{2}\right),$$

$$X_6(\theta, s) = Y_{6,1}(s) + Y_{6,2}(s) \cos(\theta_1) + Y_{6,3}(s) \cos(\theta_2),$$

$$+ Y_{6,4}(s) \cos(\theta_1) \cos(\theta_2) + Y_{6,5}(s) \sin(\theta_1) \sin(\theta_2),$$

$$X_7(\theta, s) = 0,$$

$$X_8(\theta, s) = Y_8(s),$$

$$X_9(\theta, s) = Y_9(s).$$
in which

\[ Y_2(s) = 512(20 - 9s^2), \]
\[ Y_3(s) = 256(4 - 20s + 3s^2), \]
\[ Y_4(s) = 16(364 + 144s - 504s^2 + 81s^4), \]
\[ Y_5(s) = 16(64 - 196s - 48s^2 + 104s^3 - 9s^5), \]
\[ Y_{6,1}(s) = 176 + 704s - 3132s^2 - 496s^3 + 1376s^4 - 96s^6, \]
\[ Y_{6,2}(s) = -80 + (96\sqrt{15} - 320)s + (1380 + 144\sqrt{15})s^2 + 208s^3, \]
\[ - (584 + 54\sqrt{15})s^4 + 42s^6, \]
\[ Y_{6,3}(s) = -80 - (320 + 96\sqrt{15})s + (1380 - 144\sqrt{15})s^2 + 208s^3, \]
\[ - (584 - 54\sqrt{15})s^4 + 42s^6, \]
\[ Y_{6,4}(s) = -16 - 64s + 372s^2 + 80s^3 - 208s^4 + 12s^6, \]
\[ Y_{6,5}(s) = 8(2s - 1)^3, \]
\[ Y_8(s) = 12 + 32s - 360s^2 - 512s^3 + 1025s^4 + 96s^5 - 224s^6 + 9s^8, \]
\[ Y_9(s) = 12s + 16s^2 - 120s^3 - 128s^4 + 205s^5 + 16s^6 - 32s^7 + s^9. \]

One simple observation is that

\[ (5.38) \quad Y_{6,1}(s) + Y_{6,2}(s) + Y_{6,3}(s) + Y_{6,4}(s) = 0. \]

It is easy to see that for \( |s| < 0.1, \)

\[ Y_2(s), Y_3(s), Y_5(s) > 0. \]

It is easy to compute that

\[ Y'_9(s) = 12 + 32s - 360s^2 - 512s^3 + 1025s^4 + 96s^5 - 224s^6 + 9s^8 = Y_8(s). \]

Thus,

\[ Y'_9(s) > 12 - 32 \times 0.1 - 360 \times (0.1)^2 - 512 \times (0.1)^3 - 96 \times (0.1)^5 \]
\[ - 224 \times (0.1)^6 \]
\[ > 4.5 > 0 \]

for \( |s| < 0.1, \) which implies

\[ (5.40) \quad Y_9(s) \geq Y_9(-0.1) > -1 \]
for all $|s| < 0.1$. Carefully estimating $Y_4(s)$ and $Y_5(s)$ will help us bound the $\lambda^6$ order term from below using the AM-GM inequality:

\[
Y_4(s) \geq 16(364 - 144 \times 0.1 - 504 \times (0.1)^2 - 81 \times (0.1)^4) > 5500, \\
Y_5(s) \geq 12 - 32 \times 0.1 - 360 \times (0.1)^2 - 512 \times (0.1)^3 - 96 \times (0.1)^5 - 224 \times (0.1)^6 \\
> 4.5.
\]

(5.41)

In fact, since $Y_5 = Y'_5$, the second inequality already follows from (5.39). For the $Y_{6,j}$ terms, we have

\[
Y_{6,1}(s) \geq 176 - 704 \times 0.1 - 3132 \times (0.1)^2 \\
- 496 \times (0.1)^3 - 96 \times (0.1)^6 > 0, \\
Y_{6,2}(s) \leq -80 + (96\sqrt{15} - 320) \times 0.1 + (1380 + 144\sqrt{15}) \times (0.1)^2, \\
+ 208 \times (0.1)^3 + 42 \times (0.1)^6 < 0 \\
(5.42) \\
Y_{6,3}(s) \leq -80 + (320 + 96\sqrt{15}) \times 0.1 + (1380 - 144\sqrt{15}) \times (0.1)^2, \\
+ 208 \times (0.1)^3 + 42 \times (0.1)^6 < 0, \\
Y_{6,4}(s) \leq -16 + 64 \times 0.1 + 372 \times (0.1)^2 + 80 \times (0.1)^3, \\
+ 12 \times (0.1)^6 < 0, \\
-14 \leq Y_{6,5}(s) < 0.
\]

Using (5.38) and (5.42), we obtain

\[
X_6(\theta) \geq Y_{6,1}(s) + Y_{6,2}(s) + Y_{6,3}(s) + Y_{6,4}(s) + Y_{6,5}(s) \sin(\theta_1) \sin(\theta_2)
\]

\[
= Y_{6,5}(s) \sin(\theta_1) \sin(\theta_2) \\
\geq -14|\sin(\theta_1) \sin(\theta_2)|.
\]

(5.43)

In particular, the first line uses $Y_{6,2}$, $Y_{6,3}$, $Y_{6,4} < 0$, the second line uses (5.38), and the final line uses $-14 \leq Y_{6,5} < 0$.

Now we combine our estimates together. Note that

\[
(5.44) \quad X_0(\theta, s) + X_2(\theta, s)\lambda^2 + X_3(\theta, s)\lambda^3 + X_5(\theta, s)\lambda^5 \geq 0.
\]

Using $a^2 + b^2 \geq 2|ab|$, we obtain the following from (5.41):

\[
X_4(\theta, s)\lambda^4 + \frac{1}{2}X_8(\theta, s)\lambda^8 \geq 2\sqrt{2.25 \times 5500} \sin \left(\frac{\theta_1}{2}\right) \sin \left(\frac{\theta_2}{2}\right) \lambda^6.
\]

Using $2|\sin(x/2)| \geq 2|\sin(x/2)\cos(x/2)| = |\sin(x)|$, we obtain from above that

\[
X_4(\theta, s)\lambda^4 + \frac{1}{2}X_8(\theta, s)\lambda^8 \geq 55|\sin(\theta_1) \sin(\theta_2)|\lambda^6.
\]
Combining this with (5.43), we have

\[(5.45) \quad X_4(\theta, s) \lambda^4 + \frac{1}{2} X_8(\theta, s) \lambda^8 + X_6(\theta, s) \lambda^6 \geq 41 |\sin(\theta_1) \sin(\theta_2)| \lambda^6 \geq 0.\]

Finally using (5.40) and (5.41), we have

\[(5.46) \quad \frac{1}{2} X_8(\theta, s) \lambda^8 + X_9(\theta, s) \lambda^9 = \frac{1}{2} Y_8(s) \lambda^8 + Y_9(s) \lambda^9 \geq 2.25 \lambda^8 - \lambda^9 > 0.25 \lambda^8,\]

provided that \(\lambda < 2\). Combining (5.44)–(5.46), we have

\[\det(H_\lambda(\theta)) + (1 + s \lambda) I \geq 0.25 \lambda^8 > 0,\]

for any \(\theta \in T^2\) and \(|s| < 0.1\). This proves the lower bound on the gap.

For the upper bound, observe that \(X_j((\pi, 0), s) = 0\) for all \(s\) and for every \(0 \leq j \leq 5\) and

\[X_6((\pi, 0), \pm 1/4) < -85.\]

Thus, for small \(\lambda > 0\),

\[\det(H_\lambda(\pi, 0) + (1 \pm \lambda/4) I) < -85 \lambda^6 + O(\lambda^8) < 0.\]

It is also clear that \(X_0((0, 0), s) = 4096\), which implies

\[\det(H_\lambda(0, 0) + (1 \pm \lambda/4) I) = 4096 + O(\lambda) > 0.\]

Thus we conclude that

\[1 \pm \frac{\lambda}{4} \in \sigma(H_\lambda),\]

which concludes the proof of the upper bound on the length of the gap. \(\square\)

**Acknowledgements.** We would like to thank Svetlana Jitomirskaya for comments on an earlier version of the manuscript, and Tom Spencer for useful discussions. R. H. would like to thank IAS, Princeton, for its hospitality during the 2017-18 academic year, and Virginia Tech for its hospitality during which part of the work was done. R. H. is supported in part by the National Science Foundation under Grant No. DMS-1638352 and DMS-1800689. J. F. was supported in part by an AMS Simons Travel Grant 2016–2018 and Simons Collaboration Grant #711663.
REFERENCES

[1] A. Avila, S. Jitomirskaya and C. Marx, Spectral theory of extended Harper’s model and a question by Erdős and Szekeres, Invent. Math. 210 (2017), 283–339.

[2] S. Becker and M. Zworski, Magnetic oscillations in a model of graphene, Comm. Math. Phys. 367 (2019), 941–989.

[3] S. Becker, R. Han and S. Jitomirskaya, Cantor spectrum of graphene in magnetic fields, Invent. Math. 218 (2019), 979–1041.

[4] J. Bellissard and B. Simon, Cantor spectrum for the almost Mathieu equation, J. Funct. Anal. 48 (1982), 408–419.

[5] G. Berkolaiko and A. Comech, Symmetry and Dirac points in graphene spectrum, J. Spectr. Theory 8 (2018), 1099–1148.

[6] A. Brouwer and W. Haemers, Spectra of Graphs, Springer, New York, 2012.

[7] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov and A. Geim, The electronic properties of graphene, Rev. Mod. Phys. 81 (2009), 109–162.

[8] F. Chung, Spectral Graph Theory, American Mathematical Society, Providence, RI, 1997.

[9] D. Cvetković, M. Doob, I. Gutman and A. Š托rgasev, Recent Results in the Theory of Graph Spectra, Elsevier, Amsterdam, 1988.

[10] D. Cvetković, M. Doob and H. Sachs, Spectra of Graphs, J. A. Barth, Heidelberg, 1995.

[11] P. Delplace and G. Montambaux, WKB analysis of edge states in graphene in a strong magnetic field, Phys. Rev. B 82 (2010), 205412.

[12] M. Embree and J. Fillman, Spectra of discrete two-dimensional periodic Schrödinger operators with small potentials, J. Spectr. Theory, 9 (2019), 1063–1087.

[13] C. Fefferman and M. Weinstein, Honeycomb lattice potentials and Dirac points, J. Amer. Math. Soc. 25 (2012), 1169–1220.

[14] C. Fefferman and M. Weinstein, Edge States of continuum Schroedinger operators for sharply terminated honeycomb structures, arXiv:1810.03497.

[15] C. Fefferman, J. P. Lee-Thorp and M. Weinstein, Honeycomb Schroedinger operators in the strong binding regime, Commun. Pure Appl. Math. 71 (2018), 1178–1270.

[16] J. Fillman and R. Han, preprint in preparation.

[17] D. Gieseker, H. Knörrer and E. Trubowitz, The Geometry of Algebraic Fermi Curves, Academic Press, Boston, MA, 1993.

[18] R. Han, Absence of point spectrum for the self-dual extended Harper’s model, Int. Math. Res. Not. IMRN 9 (2018), 2801–2809.

[19] R. Han, Dry Ten Martini problem for the non-self-dual extended Harper’s model, Trans. Amer. Math. Soc. 370 (2018), 197–217.

[20] R. Han and S. Jitomirskaya, Full measure reducibility and localization for quasiperiodic Jacobi operators: A topological criterion, Adv. Math. 319 (2017), 224–250.

[21] R. Han and S. Jitomirskaya, Discrete Bethe–Sommerfeld Conjecture, Commun. Math. Phys. 361 (2018), 205–216.

[22] J. H. Han, D. J. Thouless, H. Hiramoto and M. Kohmoto, Critical and bicritical properties of Harper’s equation with next-nearest-neighbor coupling, Phys. Rev. B 50 (1994), 11365.

[23] B. Helffer, P. Kerdelhué and J. Royo-Letelier, Chambers’s formula for the graphene and the Hou model with kagome periodicity and applications, Ann. Henri Poincaré 17 (2016), 795–818.

[24] B. Helffer and A. Mohamed, Asymptotics of the density of states for the Schrödinger operator with periodic electric potential, Duke Math. J. 92 (1998), 1–60.

[25] S. Jitomirskaya and C. A. Marx, Analytic quasi-periodic cocycles with singularities and the Lyapunov exponent of extended Harper’s model, Comm. Math. Phys. 316 (2012), 237–267.
[26] Y. E. Karpeshina, *Perturbation Theory for the Schrödinger Operator with a Periodic Potential*, Springer, Berlin, 1997.

[27] E. Korotyaev and N. Saburova, *Schrödinger operators on periodic discrete graphs*, J. Math. Anal. Appl. **420** (2014), 576–611.

[28] H. Krüger, *Periodic and limit-periodic discrete Schrödinger operators*, preprint, arXiv:1108.1584

[29] P. Kuchment, O. Post, *On the spectra of carbon nano-structures*, Comm. Math. Phys. **275** (2007), 805–882.

[30] P. Van Mouche, *The coexistence problem for the discrete Mathieu operator*, Comm. Math. Phys. **122** (1989), 23–33.

[31] K. Novoselov, *Nobel lecture: Graphene: Materials in the flatland*, Rev. Modern Phys. **83** (2011), 837–849.

[32] L. Parnovski, *Bethe–Sommerfeld conjecture*, Ann. Henri Poincaré **9** (2008), 457–508.

[33] L. Parnovski and A. V. Sobolev, *On the Bethe-Sommerfeld conjecture for the polyharmonic operator*, Duke Math. J. **107** (2001), 209–238.

[34] L. Parnovski and A. V. Sobolev, *Perturbation theory and the Bethe–Sommerfeld conjecture*, Ann. Henri Poincaré **2** (2001), 573–581.

[35] V. N. Popov and M. Skriganov, *A remark on the spectral structure of the two dimensional Schrödinger operator with a periodic potential*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **109** (1981), 131–133.

[36] O. Post, *Spectral Analysis on Graph-Like Spaces*, Springer, Heidelberg, 2012.

[37] M. Skriganov, *Proof of the Bethe–Sommerfeld conjecture in dimension two*, Soviet Math. Dokl. **20** (1979), 89–90.

[38] M. Skriganov, *Geometric and arithmetic methods in the spectral theory of multidimensional periodic operators*, Proc. Steklov Math. Inst. **171** (1984), 3–122.

[39] M. Skriganov, *The spectrum band structure of the three-dimensional Schrödinger operator with periodic potential*, Inv. Math. **80** (1985), 107–121.

[40] D. J. Thouless, *Bandwidth for a quasiperiodic tight binding model*, Phys. Rev. B **28** (1983), 4272–4276.

[41] O. A. Veliev, *Spectrum of multidimensional periodic operators*, Teor. Funktsii Funktsional. Anal. i Prilozhen **49** (1988), 17–34.

*Jake Fillman*

DEPARTMENT OF MATHEMATICS, MCS 470
TEXAS STATE UNIVERSITY
601 UNIVERSITY DR.
SAN MARCOS, TX 78666, USA
email: fillman@txstate.edu

*Rui Han*

DEPARTMENT OF MATHEMATICS
LOUISIANA STATE UNIVERSITY
LOCKETT HALL
BATON ROUGE, LA 70802, USA
email: rhan@lsu.edu

(Received October 23, 2018 and in revised form March 17, 2019)