MATHEMATICAL RESULTS ON THE CHIRAL MODEL OF TWISTED BILAYER GRAPHENE.

MACIEJ ZWORSKI

With an appendix by Mengxuan Yang and Zhongkai Tao

Abstract. The study of twisted bilayer graphene (TBG) is a hot topic in condensed matter physics with special focus on *magic angles* of twisting at which TBG acquires unusual properties. Mathematically, topologically non-trivial flat bands appear at those special angles. The chiral model of TBG pioneered by Tarnopolsky–Kruchkov–Vishwanath [TKV19] has particularly nice mathematical properties and we survey, and in some cases, clarify, recent rigorous results which exploit them.

Contents

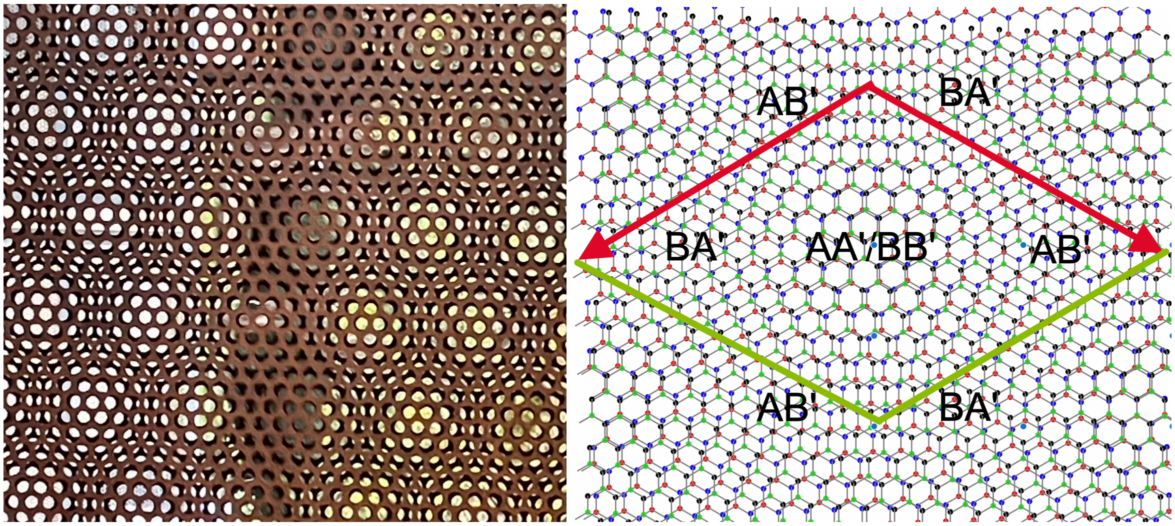


FIGURE 1. Left: a moiré pattern at CIRM in Luminy; right: a moiré fundamental cell with regions of different (AA', BB', AB'...) particle-type overlaps. Tunneling potential $|V(\mathbf{r})|$ concentrates in AA'/BB' regions and $|U(\mathbf{r})|$ concentrates at AB' regions.

1. Introduction

Investigation of physical properties of twisted bilayer graphene, and of similar structures, is a hot topics in condensed matter physics. One feature which is present when periodic structures are twisted is the emergence of *moiré patterns* – see Figure 1. These patterns create new periodic (or quasi-periodic) structures which now have much larger fundamental cells. That is very useful and, for instance, has led to experimental observation of the *Hofstadter butterfly* [H*13] – see [AvJi09] for the mathematical derivation and history.

The property on which we focus in this mathematical survey is existence of *flat bands* at certain angles of twisting (see §3.1 below for a review of the Bloch–Floquet theory and for definition of band spectrum). Flat bands correspond to eigenvalues of infinite multiplicity for the periodic Hamiltonian modeling the system. The first thought would then suggest existence of highly localized eigenstates which would prevent conductivity. If however the band topology is non-trivial (see §8 below) the localization is weak and can lead to superconductivity, in a somewhat mysterious mechanism, certainly not understood mathematically.

The Bistritzer–MacDonald Hamiltonian (BMH) [BiMa11] is widely considered to be a good model for the study of twisted bilayer graphene (TBG) and it achieved celebrity for an accurate prediction of the twisting angle at which superconductivity occurs [Ca*18]. The chiral limit of BMH is obtained by neglecting AA'/BB' tunneling (see Figure 1 and §2.2). It has many advantageous properties and was studied with great success by Tarnopolsky–Kruchkov–Vishwanath [TKV19] and their collaborators, see for instance Ledwith et al [Le*20]. One striking feature of the chiral limit, one which is not present in the BMH model, is the existence of *exact* flat bands. The Hamiltonian is of the form

$$H(\alpha) = \begin{pmatrix} 0 & D(\alpha)^* \\ D(\alpha) & 0 \end{pmatrix}, \quad D(\alpha) : H^1(\mathbb{C}; \mathbb{C}^2) \to L^2(\mathbb{C}; \mathbb{C}^2),$$

where $D(\alpha)$ is a first order (non-self-adjoint) matrix valued operator and α is dimensionless constant (a much appreciated feature for mathematicians) with $1/\alpha$ corresponding to the angle of twisting. The bands are the eigenvalues of $H_k(\alpha)$ which is obtained by replacing $D(\alpha)$ by $D(\alpha)+k$ in the definition of $H(\alpha)$ and by taking periodic boundary condition with respect to the lattice of periodicity of $H(\alpha)$, Γ. Hence,

$$H(\alpha) \text{ has a flat band at zero energy} \iff \operatorname{Spec}_{L^2(\mathbb{C}/\Gamma)} D(\alpha) = \mathbb{C}.$$

It turns out (see §§2.2,5) that the set of α's for which this happens is discrete – at other α's the spectrum is given by Γ^*, the reciprocal lattice of Γ (in the notation of §2.1, $\Gamma = 3\Lambda$ and $\Gamma^* = \frac{1}{3}\Lambda^*$).

In this survey we discuss distribution of α for which $H(\alpha)$ has a flat band at zero energy and properties of the corresponding eigenfunctions. We concentrate on presenting rigorous mathematical results familiar to the author with precise pointers to specific papers. In particular, we do not attempt to survey the vast physics literature on TBG. The motivation comes from beautiful and mysterious properties of the differential operator appearing in the chiral model (see Figure 2 for an illustration). We also highlight some open mathematical problems. The most interesting are perhaps Problems 1 and 9 as they still attract attention in the physics literature. Other problems concern finer aspects of the model and most are of purely mathematical interest – I find Problems 2,3,8,15,18 and 20 particularly appealing.

Mathematical study of the chiral model of TBG started with the work of Watson–Luskin [WaLu21] who showed existence of the first magic angle, and of Becker–Embree–Wittsten–Zworski [Be*21, Be*22] who gave a spectral characterization of magic angles and explained exponential squeezing of bands. It has been developed in several directions by Becker–Humbert–Zworski [BHZ22a, BHZ22b, BHZ22b] (trace formulas, existence of generalized magic angles, existence and properties of degenerate magic angles, topological properties), Becker–Humbert–Wittsten–Yang [Be*23] (magic angles for trilayer graphene), Becker–Oltman–Vogel [BOV23] (random perturbation of

TBG), Becker–Zworski [BeZw23a, BeZw23b] (TBG in a magnetic field parallel to the graphene sheets, deformation to the full Bistritzer–MacDonald model), Galkowski–Zworski [GaZw23] (an abstract formulation of the spectral characterization, a scalar model for magic angles), Hitrik–Zworski [HiZw23], Tao–Zworski [TaZw23b] (classically forbidden regions for eigenstates), and Yang [Ya23] (twisted multiple layer graphene). Many of these results are described in this survey.

During the writing of this survey it became apparent that we did not have a reference to the fact that the chiral model of TBG exhibits Dirac cones away from α's at which flat bands appear – see Open Problem 2. Mengxuan Yang and Zhongkai Tao immediately provided an argument for that and it is included here as an appendix.

Notation In this paper we use the physics notation: for an operator A on $L^2(M, dm)$, $\langle u|A|v\rangle := \int_M Av\,\bar{u}\,dm$. Also, $|u\rangle$ denotes the operator $\mathbb{C} \ni \mu \to \mu u \in L^2$ and $\langle u|$, its adjoint $L^2 \ni v \to \langle u|v\rangle \in \mathbb{C}$. For $z, w \in \mathbb{C} \simeq \mathbb{R}^2$, we use the real inner product, $\langle z, w\rangle := \operatorname{Re} z\bar{w}$. If H is a function space (such as L^2, Sobolev space H^s or spaces with given periodicity conditions) then $H(M; \mathbb{C}^n)$ denotes functions in H on M with values in $\mathbb{C}^n$. When the context is clear we may drop M and $\mathbb{C}^n$.

Acknowledgements. I would like to thank Mike Zaletel for introducing me to TBG and pointing out the semiclassical nature of small angle asymptotics. I am grateful to my many collaborators on projects related to TBG, on whose work this survey is based, especially to Simon Becker who produced most of the figures (and movies) in our recent joint papers, some of which are re-used here. I would also like to thank Patrick Ledwith, Lin Lin, Mitch Luskin, Allan MacDonald, Ashvin Vishwanath, and Alex Watson for valuable physics perspectives (many of which, alas, remain a mystery to this author). Simon Becker, Jens Wittsten and Mengxuan Yang provided many insightful comments on earlier versions of this survey and I am very grateful for that great help. Thanks go also to Zhongkai Tao for pointing out and clarifying a mistake in §8. Partial support by the NSF grant DMS-1901462 and by the Simons Foundation under a "Moiré Materials Magic" grant is also most gratefully acknowledged.

2. The Bistritzer–MacDonald Hamiltonian and its chiral limit

In this section we consider the Bistritzer–MacDonald Hamiltonian (BMH) [BiMa11] from the PDE point of view without addressing its physical motivation. It has been mathematically derived by Cancès–Garrigue–Gontier [CGG22] and Watson–Kong–MacDonald–Luskin [Wa*22] and we refer to these papers above and [TKV19] for physics background. As we will stress, its chiral limit exhibits beautiful and unusual mathematical properties which have been our main motivation.

The representation of BMH in the physics literature [BiMa11], [TKV19] is given as follows: for two parameters α and λ we define

$$H_{\rm BM}(\alpha,\lambda) = \begin{pmatrix} -i(\sigma_1\partial_{x_1}+\sigma_2\partial_{x_2}) & T(\alpha,\lambda) \\ T(\alpha,\lambda)^* & -i(\sigma_1\partial_{x_1}+\sigma_2\partial_{x_2}) \end{pmatrix} : H^1(\mathbb{R}^2;\mathbb{C}^4)\to L^2(\mathbb{R}^2;\mathbb{C}^4),$$

where we use Pauli matrices,

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbf{r} = (x_1,x_2)\in\mathbb{R}^2,$$

and $\bullet^*$ denotes the hermitian conjugate.

The interlayer tunnelling matrix is defined as follows

$$T(\alpha,\lambda) = \begin{pmatrix} \lambda V(\mathbf{r}) & \alpha\overline{U(-\mathbf{r})} \\ \alpha U(\mathbf{r}) & \lambda V(\mathbf{r}) \end{pmatrix}.$$

The non-equivalent pairs of atoms in a fundamental cell of the honeycomb lattice of graphene are labelled by A, B, with the labeling A', B' for the second sheet in TBG. In the matrix potential T, $U(\pm\bullet)$ and V model AB'/BA' and AA'/BB' tunnelling respectively, see Figure 1. They are defined as follows: with $\omega := \exp(2\pi i/3)$,

$$U(\mathbf{r}) = \sum_{i=0}^{2}\omega^\ell e^{-iq_\ell\cdot\mathbf{r}}, \quad V(\mathbf{r}) = \sum_{\ell=0}^{2} e^{-iq_\ell\cdot\mathbf{r}}, \quad q_\ell := R^\ell(0,-1), \quad R := \tfrac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

(We note that R is the $2\pi/3$ rotation matrix.) A useful equivalent representation of $H_{\rm BM}$ is given as follows:

$$AH_{\rm BM}(\alpha,\lambda)A = \begin{pmatrix} \lambda C & D(\alpha)^* \\ D(\alpha) & \lambda C \end{pmatrix}, \quad A := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : \mathbb{C}^4\to\mathbb{C}^4,$$

where (with $D_{x_j} := (1/i)\partial_{x_j}$)

$$D(\alpha) = \begin{pmatrix} D_{x_1}+iD_{x_2} & \alpha U(\mathbf{r}) \\ \alpha U(-\mathbf{r}) & D_{x_1}+iD_{x_2} \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & V(\mathbf{r}) \\ V(-\mathbf{r}) & 0 \end{pmatrix}.$$

In most of the figures we use the coordinates (x_1,x_2) and corresponding dual coordinates k.

2.1. **Change to the standard lattice $\mathbb{Z}+\omega\mathbb{Z}$.** The potentials U and V are periodic with respect to the lattice $\Gamma = 4\pi i(\mathbb{Z}+\omega\mathbb{Z})$ with finer twisted periodicity with respect to the moiré lattice $\frac{1}{3}\Gamma$. It is mathematically nicer, especially when dealing with theta functions, to use coordinates in which the moiré lattice is given by $\Lambda := \mathbb{Z}+\omega\mathbb{Z}$. This corresponds to changing the physics coordinates $\mathbf{r} = (x_1,x_2)\in\mathbb{R}^2$ to $z\in\mathbb{C}\simeq\mathbb{R}^2$ defined by

$$x_1 + ix_2 = \tfrac{4}{3}\pi i z.$$

This leads to an equivalent Hamiltonian,

$$H(\alpha,\lambda) := \begin{pmatrix} \lambda C & D(\alpha)^* \\ D(\alpha) & \lambda C \end{pmatrix} : H^1(\mathbb{C};\mathbb{C}^4) \to L^2(\mathbb{C};\mathbb{C}^4), \quad \alpha \in \mathbb{C}, \quad \lambda \in \mathbb{R}, \tag{2.1}$$

where (with $D_{\bar z} = (1/i)\partial_{\bar z} = (1/2i)(\partial_{x_1} + i\partial_{x_2})$)

$$D(\alpha) = \begin{pmatrix} 2D_{\bar z} & \alpha U(z) \\ \alpha U(-z) & 2D_{\bar z} \end{pmatrix}, \qquad C := \begin{pmatrix} 0 & V(z) \\ V(-z) & 0 \end{pmatrix}, \tag{2.2}$$

where the parameter α is proportional to the inverse relative twisting angle. With $\omega = e^{2\pi i/3}$ and $K := \frac{4}{3}\pi$, we assume that

$$U(z+\gamma) = e^{i\langle \gamma, K\rangle} U(z), \ \gamma \in \Lambda, \quad U(\omega z) = \omega U(z), \quad \overline{U(\bar z)} = -U(-z), \tag{2.3}$$

$\Lambda := \mathbb{Z} \oplus \omega\mathbb{Z}$, and

$$V(z) = V(\bar z) = \overline{V(-z)}, \quad V(\omega z) = V(z), \quad V(z+\gamma) = e^{i\langle\gamma,K\rangle} V(z). \tag{2.4}$$

The specific potentials in H_{BM} are, with $K = \frac{4}{3}\pi$,

$$U(z) = U_{\mathrm{BM}}(z) := -\tfrac{4}{3}\pi i \sum_{\ell=0}^{2} \omega^\ell e^{i\langle z, \omega^\ell K\rangle}, \quad V(z) = V_{\mathrm{BM}}(z) := \sum_{\ell=0}^{2} e^{i\langle z,\omega^\ell K\rangle}, \tag{2.5}$$

and these are the potentials used in (most) numerical experiments in the papers cited in the abstract.

2.2. **The chiral limit.** When we put $\lambda = 0$ in (2.1) (or equivalently in H_{BM}) we obtain an operator build from $D(\alpha)$ only and satisfying a chiral symmetry:

$$\begin{gathered} H(\alpha) := H(\alpha, 0) = \begin{pmatrix} 0 & D(\alpha)^* \\ D(\alpha) & 0 \end{pmatrix}, \\ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} H(\alpha) \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -H(\alpha). \end{gathered} \tag{2.6}$$

In particular, the spectrum of $H(\alpha)$ is symmetric with respect to 0. The great advantage comes from reducing some properties of $H(\alpha)$ to those of the operator $D(\alpha)$. We will see in §3 that $H(\alpha)$ has a perfect *flat band* at energy zero if and only if

$$\mathrm{Spec}_{L^2(\mathbb{C}/3\Lambda;\mathbb{C}^2)} D(\alpha) = \mathbb{C}, \tag{2.7}$$

and that the set of α's for which this happens is discrete. Outside of that discrete set the spectrum on $L^2(\mathbb{C}/3\Lambda)$ is given by $\frac{1}{3}\Lambda^*$. The domain of $D(\alpha)$ is given by $H^1(\mathbb{C}/3\Lambda)$ and it is a Fredholm operator of index 0. (In §3 we will consider a finer space $L^2_0(\mathbb{C};\mathbb{C}^2)$ which is more suitable for Floquet theory and the study of flat bands; the reason for $\mathbb{C}/3\Lambda$ is periodicity of potentials with respect to the lattice 3Λ.) The set, $\mathcal{A}$, of α's for which (2.7) holds satisfies the following symmetries (see [Be*22], [BHZ22b, §2.3]):

$$\mathcal{A} = -\mathcal{A} = \bar{\mathcal{A}}. \tag{2.8}$$

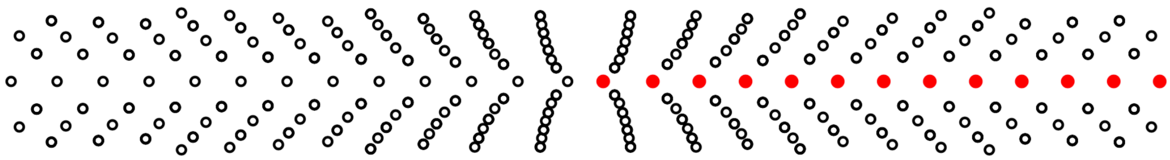


FIGURE 2. The set of α's for which (2.7) holds (with the potential given by (2.5)), that is for which the chiral Hamiltonian has a perfectly flat band at 0 energy. The regular distribution becomes less apparent when the potential is relaxed while all the properties (2.3) are maintained.

Another advantage of the operator $D(\alpha)$ is that scalar valued holomorphic functions act as scalars:

$$D(\alpha)(fu) = fD(\alpha)u, \quad u \in H^1_{\rm loc}(\mathbb{C}; \mathbb{C}^2), \quad f \in \mathcal{O}(\mathbb{C}; \mathbb{C}).$$

This was emphasized in [TKV19] and was a basis of the argument recalled in §6 below.

A crucial feature of $D(\alpha)$ is its non normality, $[D(\alpha), D(\alpha)^*] \neq 0$. This allows for exotic phenomena such as (2.7), which in turn produce exactly flat bands appreciated by physicists. As indicated in [Be*22] it also results in less desirable features such as exponential squeezing of bands (see §10.1) and spectral instability (see Figure 5). Those effects are exploited in [BOV23] where small random perturbations produce dramatic changes in spectral behaviour, suggesting high instability of all but the first magic angle.

The set of (complex) α's for which (2.7) holds for the potential (2.5) is shown in Figure 2. Its structure remains a mystery. One striking observation made in [TKV19] is the even spacing of real α's (shown in red and labeled $0 < \alpha_1 < \alpha_2 < \cdots$) roughly given by

$$\alpha_{j+1} - \alpha_j \simeq \tfrac{3}{2}. \tag{2.9}$$

(A more accurate computation based on the spectral characterization [Be*22] – see Theorem 5 – suggests the spacing $\simeq 1.515$).

Open Problem 1. *For U given in (2.5) establish an asymptotic quantization rule (2.9). At the moment, there are no convincing arguments. A more general question is obtaining asymptotics of real α's for more general potentials satisfying (2.3). In that case, a simple law similar to (2.9) is harder to observe – see the movie linked to Figure 3. See also §5.2 and §10 for discussions of related issues.*

3. Basic symmetries and band theory of TBG

The translation symmetry of BMH are given as follows: for $u \in L^2_{\rm loc}(\mathbb{C}; \mathbb{C}^2)$ we define

$$L_\gamma u(z) := \begin{pmatrix} e^{i\langle \gamma, K\rangle} & 0 \\ 0 & e^{-i\langle \gamma, K\rangle} \end{pmatrix} u(z+\gamma), \quad \gamma \in \Lambda, \quad K = \tfrac{4}{3}\pi. \tag{3.1}$$

We extend this action diagonally for $w \in L^2_{\rm loc}(\mathbb{C};\mathbb{C}^4)$:

$$\mathscr{L}_\gamma w = \begin{pmatrix} L_\gamma w_1 \\ L_\gamma w_2 \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \quad w_j \in L^2_{\rm loc}(\mathbb{C};\mathbb{C}^2).$$

We then have, in the notation of (2.1), (2.2) with U, V satisfying (2.3) and (5.9),

$$L_\gamma D(\alpha) = D(\alpha) L_\gamma, \quad \mathscr{L}_\gamma H(\alpha,\lambda) = H(\alpha,\lambda)\mathscr{L}_\gamma. \tag{3.2}$$

We also define the pull back of the rotation by $2\pi/3$:

$$\begin{gathered} \Omega : L^2_{\rm loc}(\mathbb{C};\mathbb{C}^2) \to L^2_{\rm loc}(\mathbb{C};\mathbb{C}^2), \quad \mathscr{C} : L^2_{\rm loc}(\mathbb{C};\mathbb{C}^4) \to L^2_{\rm loc}(\mathbb{C};\mathbb{C}^4), \\ \Omega u(z) := u(\omega z), \ \ \mathscr{C}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} := \begin{pmatrix} \Omega w_1 \\ \bar\omega \Omega w_2 \end{pmatrix}. \end{gathered} \tag{3.3}$$

This gives

$$\Omega D(\alpha) = \omega D(\alpha), \quad \mathscr{C} H(\alpha,\lambda) = H(\alpha,\lambda)\mathscr{C}. \tag{3.4}$$

The natural subspaces of $L^2_{\rm loc}(\mathbb{C};\mathbb{C}^p)$, $p = 2,4$, are given by

$$L^2_k(\mathbb{C};\mathbb{C}^2) := \{u \in L^2_{\rm loc}(\mathbb{C},\mathbb{C}^2) : L_\gamma u = e^{i\langle k,\gamma\rangle} u\}, \quad \|u\|_{L^2_k} = \int_{\mathbb{C}/\Lambda} |u(z)|^2 dm(z), \tag{3.5}$$

and similarly for $p = 4$ with L_γ replaced by $\mathscr{L}_\gamma$. We also define Sobolev spaces $H^s_k := L^2_k \cap H^s_{\rm loc}$. With $s = 1$ they can be used as domains of our operators.

These spaces depend only on the congruence class of k in $\mathbb{C}/\Lambda^*$,

$$\Lambda^* := \frac{4\pi i}{\sqrt{3}}\Lambda, \quad k \mapsto z(k) := \frac{\sqrt{3}k}{4\pi i}, \quad \Lambda^* \to \Lambda, \quad \langle p,\gamma\rangle \in 2\pi\mathbb{Z}, \quad p \in \Lambda^*, \ \gamma \in \Lambda. \tag{3.6}$$

The points of high symmetry, $\mathcal{K}$, are defined by demanding that

$$p \in \mathcal{K} \implies \omega p \equiv p \mod \Lambda^*.$$

They are given by

$$\mathcal{K} = \{K, -K, 0\} + \Lambda^*, \quad K = \tfrac{4}{3}\pi. \tag{3.7}$$

Mathematically, these are the fixed points of the action of $z \mapsto \omega z$ on $\mathbb{C}/\Lambda^*$. Physically, $\pm K$ are called the K-points at which Dirac points are present (see §6) and 0 is called a Γ-point – see Figure 4. (A different choice of L_γ in (3.1) can result in different sets of K-points – see [BeZw23a, §2].)

For $k \in \mathcal{K}/\Lambda^*$ and $p \in \mathbb{Z}_3$ we also define

$$L^2_{k,p}(\mathbb{C};\mathbb{C}^4) := \{u \in L^2_k(\mathbb{C};\mathbb{C}^4) : \mathscr{C} u = \bar\omega^p u\}, \tag{3.8}$$

with the definition of $L^2_{k,p}(\mathbb{C};\mathbb{C}^2)$ obtained by replacing $\mathscr{C}$ by Ω. We have orthogonal decompositions $L^2_k = \bigoplus_{p\in\mathbb{Z}_3} L_{k,p}$, $k \in \mathcal{K}/\Lambda^*$. Also, the actions of $\mathscr{L}_\gamma$ and $\mathscr{C}$ on $L^2_{p,k}$ commute. In general, $\mathscr{L}_\gamma \mathscr{C} = \mathscr{C}\mathscr{L}_{\omega\gamma}$ and the group generated by the action $\mathscr{L}_\gamma$ and $\mathscr{C}$ (or the actions of L_γ and $\mathscr{C}$) is a discrete Heisenberg group – see §[Be*22, §2.1]. These spaces play an important role in the study of protected states, multiplicities and trace formulas for magic angles.

3.1. **Bloch–Floquet theory.** The "twisted" translations $\mathscr{L}_\gamma$ can be used to define a Bloch transform

$$\mathcal{B}u(k,z) := |\mathbb{C}/\Lambda^*|^{-\frac{1}{2}} \sum_{\gamma\in\Lambda} e^{-i\langle z+\gamma,k\rangle} \mathscr{L}_\gamma u(z), \quad u \in \mathscr{S}(\mathbb{C}).$$

We then easily check that

$$\begin{gathered} \mathcal{B}u(k+p,z) = e^{-i\langle z,p\rangle}\mathcal{B}u(k,z), \quad p \in \Lambda^*, \\ \mathscr{L}_\alpha \mathcal{B}u(k,\bullet) = |\mathbb{C}/\Lambda^*|^{-\frac{1}{2}} \sum_{\gamma} e^{-i\langle z+\alpha+\gamma,k\rangle} \mathscr{L}_{\alpha+\gamma} u(z) = \mathcal{B}u(k,\bullet), \quad \alpha\in\Lambda. \end{gathered}$$

We can check that for $u \in \mathscr{S}(\mathbb{C})$,

$$\int_{\mathbb{C}/\Lambda}\int_{\mathbb{C}/\Lambda^*} |\mathcal{B}u(k,z)|^2 dm(z)dm(k) = \int_{\mathbb{C}} |u(z)|^2 dm(z),$$

and that

$$\mathcal{C}\mathcal{B}u(z) = u(z), \quad \mathcal{C}u(z) := |\mathbb{C}/\Lambda^*|^{-\frac{1}{2}} \int_{\mathbb{C}/\Lambda^*} v(z,k) e^{i\langle z,k\rangle} dm(k).$$

This shows that $\mathcal{B}$ extends to a unitary map $\mathcal{B} : L^2(\mathbb{C};\mathbb{C}^4) \to \mathscr{H}$, where

$$\mathscr{H} := \{v(k,z) \in L^2_{\mathrm{loc}}(\mathbb{C}; L^2_0(\mathbb{C};\mathbb{C}^4)), \quad v(k+p,z) = e^{-i\langle z,p\rangle}v(k,z),\ p \in \Lambda^*\}.$$

We then define

$$\begin{gathered} H_k(\alpha,\lambda) : \mathscr{D} \to \mathscr{H}, \quad \mathscr{D} := \mathscr{H} \cap L^2_{\mathrm{loc}}(\mathbb{C}_k; H^1_0(\mathbb{C},\mathbb{C}^4)), \\ H_k(\alpha,\lambda) := e^{-i\langle z,k\rangle} H(\alpha,\lambda) e^{i\langle z,k\rangle} = \begin{pmatrix} \lambda C & D(\alpha)^* + \bar{k} \\ D(\alpha)+k & \lambda C \end{pmatrix}, \\ [H_k(\alpha,\lambda)\mathcal{B}u](k,z) = [\mathcal{B}H(\alpha,\lambda)u](k,z). \end{gathered} \tag{3.9}$$

We see that $\mathrm{Spec}_{L^2_0}(H_k(\alpha,\lambda))$ (with the domain given by H^1_0) is discrete and

$$\mathrm{Spec}_{L^2(\mathbb{C};\mathbb{C}^4)}(H(\alpha,\lambda)) = \bigcup_{k\in\mathbb{C}/\Lambda^*} \mathrm{Spec}_{L^2_0} H_k(\alpha,\lambda).$$

The Hamiltonian (2.1) possesses other important symmetries called the *parity-inversion/time-reversal* symmetry, the *particle-hole* symmetry and the *mirror* symmetry – see [BeZw23b, §2.2] for a concise review. One consequence of the symmetries is the existence and properties of protected states:

Theorem 1 ([Be*21, Be*22])**.** *For the Hamiltonian* (3.9) *with U and V satisfying* (2.3),(5.9) *and $\alpha,\lambda \in \mathbb{R}$,*

$$\dim\ker_{H^1_0} H_{\pm K}(\alpha,\lambda) \geq 2, \quad K = \tfrac{4}{3}\pi, \tag{3.10}$$

In addition, for $\alpha \in \mathbb{C}$

$$\dim \ker_{H^1_0}(D(\alpha) \pm K) \geq 1. \tag{3.11}$$

Moreover we can find a holomorphic function

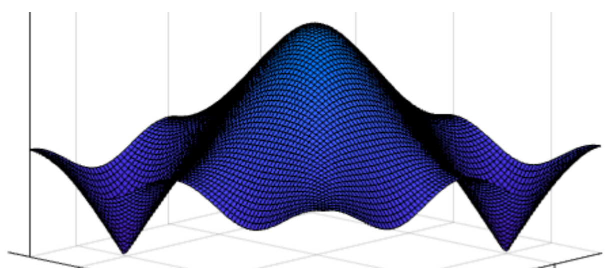
$$\mathbb{C} \ni \alpha \to u_{\pm K}(\alpha) \in (C^\infty \cap L^2_0)(\mathbb{C}; \mathbb{C}^2) \setminus \{0\},$$

such that

$$\begin{gathered} (D(\alpha) \pm K) u_{\pm K}(\alpha) = 0, \qquad u_{-K}(\alpha) = \tau(K) \mathscr{E} \tau(K) u_K(\alpha), \\ \tau(K) u_K(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \tau(\pm K) u_{\pm K}(\alpha) \in L^2_{\pm K, 0}, \\ \mathscr{E} \begin{pmatrix} u_1(z) \\ u_2(z) \end{pmatrix} := \begin{pmatrix} u_2(-z) \\ -u_1(-z) \end{pmatrix}, \qquad \tau(k) u(z) := e^{i\langle z, k \rangle} u(z). \end{gathered} \tag{3.12}$$

This was essentially established in [Be*21, Be*22] but for a streamlined proof of (3.10) see [BeZw23b, Proposition 2], and for the proofs of (3.11) and (3.12), [BHZ22b, Propositions 2.2, 2.3], respectively. An alternative proof of (3.11) which does not involve $H(\alpha, 0)$ is presented in [Be*23].

Open Problem 2. *Do upper and lower bands for the Bistritzer–MacDonald Hamiltonian have conic singularities at* $\pm K$ *for all real values of* α *and* λ*? That would mean that* $\pm K$ *are Dirac points:*



3.2. **Flat bands in the chiral limit.** The first advantage of the chiral model (2.6) is that the spectrum of $H_k(\alpha) := H_k(\alpha, 0)$ is symmetric with respect to 0 (that is not true in the case of BMH – see §4). In view of (3.11) we know that that two bands always touch at 0. Hence it is natural to label the spectrum of $H_k(\alpha)$ as follows:

$$\begin{gathered} \operatorname{Spec}_{L^2_0} H_k(\alpha) = \{E_\ell(\alpha, k)\}_{\ell \in \mathbb{Z} \setminus 0}, \quad E_{\ell+1}(\alpha, k) \geq E_\ell(\alpha, k), \\ E_\ell(\alpha, k) = -E_{-\ell}(\alpha, k), \quad E_{\pm 1}(\alpha, \pm K) = 0, \text{ for all } \alpha \in \mathbb{C}. \end{gathered} \tag{3.13}$$

We note that $E_\ell(\alpha, k)$, $\ell \geq 1$, are the ordered sequence of the singular values of the non-self-adjoint operator $D(\alpha) + k$.

A flat band at zero energy occurs at a given value of the parameter α if $E_1(\alpha, k) = 0$ for all $k \in \mathbb{C}$. We recall that in the BMH, $1/\alpha$ is proportional to the angle of twisting of the two sheets of graphene. For a specific potential U satisfying (2.3) the magic α (that is *magic angles*) and their multiplicities were defined as follows in [BHZ23]:

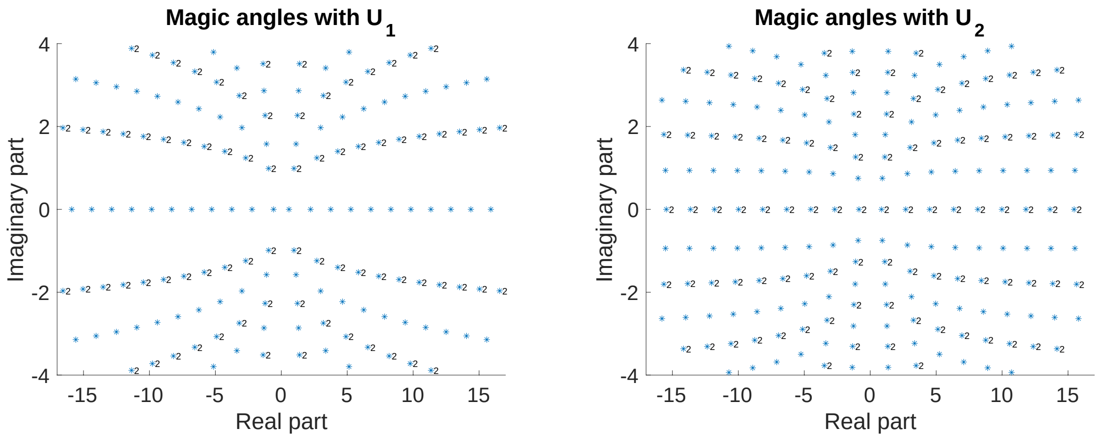


FIGURE 3. Magic angles α for $U_1(z) = U_{\rm BM}(z)$ given in (2.5) (left) and $U_2(z) = (U_{\rm BM}(z) - U_{\rm BM}(-2z))/\sqrt{2}$ (right). Multiplicity of the flat bands (no number $\to$ simple magic angle, 2 $\to$ two-fold degenerate magic angle) in the figure. The movie https://math.berkeley.edu/~zworski/Interpolation.mp4 shows the magic angles for interpolation between these potentials: $U(z) = (\cos\theta - \sin\theta)U_1(z) + \sin\theta U_2$; multiplicity one magic angles are coded by $*$ and multiplicity two by $*$.

Definition (Magic angles and their multiplicities). *A value of α in (2.2) is called magical if $H(\alpha)$ has a flat band at zero*

$$E_1(\alpha, k) \equiv 0, \quad k \in \mathbb{C}. \tag{3.14}$$

The set of magic α's is denoted by $\mathcal{A}$ or $\mathcal{A}(U)$ if we specify the dependence on the potential. The multiplicity of a magic α is defined as

$$m(\alpha) = m_U(\alpha) := \min\{j > 0 : \max_k E_{j+1}(\alpha, k) > 0\}. \tag{3.15}$$

Magic angles are (up to physical constants) reciprocals of $\alpha \in \mathcal{A}$.

A numerical illustration of the sets $\mathcal{A}$ for different potentials satisfying (2.3) is shown in Figure 3. Multiplicities are indicated there and in the linked animation. The computation was done based on the spectral characterization described in the next section. The protected nature of multiplicities one and two will be reviewed in §7.

Although the proof relies on the material presented in §5 we recall here a result stating that if $E_1(\alpha, k)$ touches 0 at some k away from the K-points then the band has to be perfectly flat:

Theorem 2 ([Be*22], [BHZ22b]). *For any U satisfying (2.3) and $\alpha \in \mathbb{C}$,*

$$\exists\ k \notin \{-K, K\} + \Lambda^* \quad E_1(\alpha, k) = 0 \implies \forall\ k \in \mathbb{C} \quad E_1(\alpha, k) = 0. \tag{3.16}$$

For the Bistritzer–MacDonald Hamiltonian (2.1) perfectly flat bands are not expected. That the antichiral model $H(0,\lambda)$ cannot have flat bands was shown in [Be*21].

A perfectly flat band at 0 energy for a periodic Hamiltonian corresponds to an eigenvalue of infinite multiplicity at 0 for the Hamiltonian acting on L^2 (in our case $L^2(\mathbb{C};\mathbb{C}^4)$ with the domain given by $H^1(\mathbb{C};\mathbb{C}^4)$). Physical properties, such as superconductivity, are then related to the decay of the corresponding eigenfunctions. That in turn is related to the *topology of the flat band* – see [TaZw23a, §8.5] and references given there. Trivial topology gives exponential decay while nontrivial topology forces the blow up of moments of the probability distribution of the Wannier functions [TaZw23a, Theorem 9]. We will discuss the topology of flat bands for TBG in §8.

Open Problem 3. *Show that the Hamiltonian* (2.1), $H(\alpha,\lambda)$, *with* U *and* $V \not\equiv 0$ *satisfying* (2.3),(5.9), *cannot have flat bands when* $\lambda \neq 0$. *(Or give a counterexample to this claim.)*

Open Problem 4. *Numerics indicate (see* [BHZ22b, Figure 2]*) that for* $U = U_{\rm BM}$, $k \mapsto E_1(\alpha,k)/[\max_{p\in\mathbb{C}} E_1(\alpha,p)]$ *does not vary much with* α*, in particular in neighbourhoods of* $\alpha \in \mathcal{A}$*, and its graph is close to that of* $k \mapsto |U_{\rm BM}(z(k))|$*, where* $z : \Lambda^* \to \Lambda$*, see* (3.6)*. What is the explanation of this phenomenon? For an animation of rescaled bands see* `https://math.berkeley.edu/~zworski/KKmovie.mp4`.

4. BMH as a perturbation of the chiral model

The Bistritzer–MacDonald Hamiltonian (BMH) (2.1) could, for small values of the coupling constant λ, be considered as a perturbation of the chiral model. The actual physical value of λ (see [BiMa11, TKV19]) is approximately given by $\lambda = 0.7\alpha$.

The simplest case to consider is of $\alpha \in \mathcal{A}$ which is positive and simple (which, in the case of the potential in (2.5) we know rigorously for the smallest magic α and numerically for other real α's – see §7). Then, in the notation of (3.13),

$$E_{-2}(\alpha,k) < E_{-1}(\alpha,k) = 0 = E_1(\alpha,k) < E_2(\alpha,k), \quad \text{for all } k.$$

This means that for $|\lambda| \ll 1$ in (3.13), the bands $E_{\pm 1}(\alpha,\lambda,k)$ are well defined.

A standard application of perturbation theory (see §9), the symmetries of $D(\alpha)$ and $H(\alpha,\lambda)$, and of some basic properties of theta functions (see (6.11),(6.9),(6.10) below) gives the following simple, but to us, surprising result:

Theorem 3 ([BeZw23b])**.** *Suppose that* $\alpha \in \mathcal{A} \cap \mathbb{R}$ *is simple and that* $k \mapsto E_{\pm 1}(\alpha,\lambda,k)$ *are the two lowest bands (in absolute value) of BMH in* (2.1)*. Then there exist* $e(\alpha,\bullet), f(\alpha,\bullet) \in C^\infty(\mathbb{C}/\Lambda^*)$ *such that*

$$E_{\pm 1}(\alpha,\lambda,k) = e(\alpha,k)\lambda \pm |f(\alpha,k)|\lambda^2 + \mathcal{O}(\lambda^3), \quad \lambda \to 0, \tag{4.1}$$

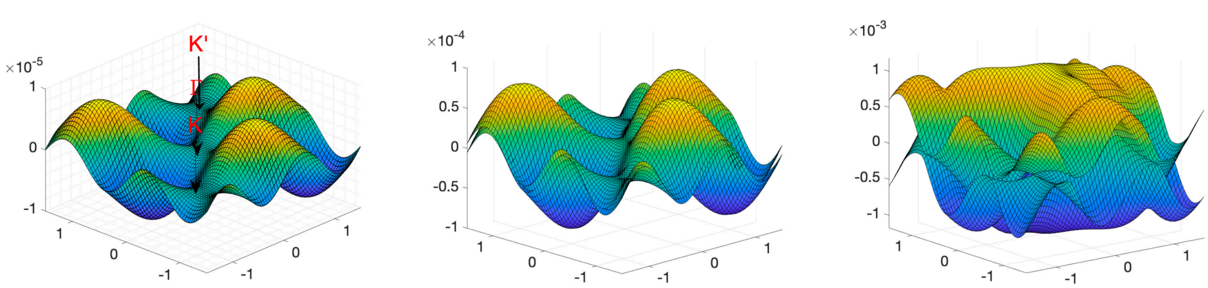


FIGURE 4. Plots of $k \mapsto E_{\pm 1}(\alpha, \lambda, k)$ for α the first real magic element of $\mathcal{A}$ and $\lambda = 10^{-3}, 10^{-2}, 10^{-1}$. We see that for very small coupling the flat bands "move together" and split only when the coupling gets larger; the quadratic term controls the splitting of the bands, see Figure 1. For an animated version see https://math.berkeley.edu/~zworski/Chiral2BM.mp4.

$f(\pm K) = 0$, *(*$\omega K \equiv K \bmod \Lambda^*$, $K \neq 0$*), and*

$$e(\alpha, k) = -e(\alpha, -k) = -e(\alpha, \bar{k}) = e(\alpha, \omega k), \quad \omega = e^{2\pi i/3}. \tag{4.2}$$

The surprising fact is that the leading linear term (for very small λ) does not depend on the band: when λ is switched on the two bands initially move together – see Figure 4. However, $|e(\alpha, k)| \ll |f(\alpha, k)|$ (except at the crossing points $k = \pm K$) and hence the quadratic term quickly dominates and is responsible for the splitting of the bands – see [BeZw23b, Figure 2]. For the first magic α (and the potential in (2.5)), the quadratic approximation provides an accurate description of the bands when $\lambda = 0.7\alpha$ (the physical λ). For a discussion of the splitting of bands in the case of double α's see [BeZw23b, §5].

Open Problem 5. *Show that* $|f(\alpha, \pm K + \zeta)| \sim |\zeta|$ *which is equivalent to showing that the Jacobian does not vanish:* $|\partial_k f(\alpha, \pm K)|^2 - |\partial_{\bar{k}} f(\alpha, \pm K)|^2 \neq 0$. *This is a simpler (infinitesimal) version of Problem 2 at a magic angle.*

5. Spectral characterization of magic angles

In §3 we gave the definition of $\mathcal{A} \subset \mathbb{C}$, the set of magic parameters α (corresponding to the reciprocals of magic angles). The purpose of this section is to give a general argument [GaZw23] for the discreteness of $\mathcal{A}$ which relies only on holomorphy of $\alpha \mapsto D(\alpha)$, Fredholm properties, and existence of protected states. In the case of operators appearing in [TKV19, Be*22, Be*23, Ya23] it also characterizes magic angles as eigenvalues of a compact operator, which in turn allows their accurate numerical computation (see Figure 3).

We replace the operator $D(\alpha)+k$ by a family of operators acting between Banach spaces X and Y. We let $\Omega \subset \mathbb{C}$ be an open set and assume that for $(\alpha,k)\in\Omega\times\mathbb{C}$

$$\begin{gathered} Q(\alpha,k) : X \to Y, \quad \text{is a holomorphic family of Fredholm operators of index 0,} \\ \tau_Y(p) Q(\alpha,k) \tau_X(p)^{-1} = Q(\alpha, k+p), \quad k \in \mathbb{C}, \quad p \in \Lambda^*, \end{gathered} \tag{5.1}$$

where the maps $\tau_\bullet(p) : \bullet \to \bullet$, $\bullet = X, Y$, are invertible bounded linear maps, and Λ^* is a lattice in $\mathbb{C}$. (The last condition can be significantly weakened but we leave in the form relevant to periodic problems.)

We have the following for dichotomy: for a fixed $\alpha \in \Omega$

$$k \mapsto Q(\alpha,k)^{-1} \text{ is a meromorphic for } k \in \mathbb{C} \text{ with poles of finite rank} \tag{5.2}$$

or

$$\ker_X Q(\alpha,k) \neq \{0\} \text{ for all } k \in \mathbb{C}. \tag{5.3}$$

(See [GaZw23] and also [DyZw19, Appendix C] for a brief introduction to Fredholm theory and families of meromorphic operators.)

We now define multiplicity as follows: if (5.2) holds then

$$m(\alpha,k) := \frac{1}{2\pi i} \operatorname{tr} \oint_{\partial D} Q(\alpha,\zeta)^{-1} \partial_\zeta Q(\alpha,\zeta) d\zeta, \tag{5.4}$$

where the integral is over the positively oriented boundary of a disc D which contains k as the only possible pole of $\zeta \mapsto Q(\alpha,\zeta)$. Otherwise, that is when (5.3) holds, we put $m(\alpha,k) = \infty$ for all $k \in \mathbb{C}$.

Although seemingly very general and abstract, this definition is necessary in natural examples as will be indicated in §§5.1,5.2.

Theorem 4 ([GaZw23])**.** *Suppose that* (5.1) *holds and that for some $\alpha_0 \in \Omega$ and every $k \in \mathbb{C}$, we have,*

$$m(\alpha,k) \geq m(\alpha_0,k) \neq \infty. \tag{5.5}$$

Then there exists a discrete set $\mathcal{A} \subset \Omega$ such that for all $k \in \mathbb{C}$

$$m(\alpha,k) = \begin{cases} \infty & \alpha \in \mathcal{A}, \\ m(\alpha_0,k) & \alpha \notin \mathcal{A}. \end{cases} \tag{5.6}$$

We illustrate the theorem with some simple examples.

Examples. 1. Consider

$$Q(\alpha,k) = e^{ix} D_x + (\alpha - \tfrac{1}{2}) e^{ix} + k, \quad x \in \mathbb{R}/2\pi\mathbb{Z}, \quad D_x := (1/i)\partial_x.$$

Then, in the notation of Theorem 4, $X = L^2(\mathbb{R}/2\pi\mathbb{Z})$, $Y = H^1(\mathbb{R}/2\pi\mathbb{Z})$ and

$$m(k,0) \equiv 0, \quad \Lambda^* = 2\pi\mathbb{Z}, \quad \mathcal{A} = \mathbb{Z} + \tfrac{1}{2}. \tag{5.7}$$

In this case we do not have the second condition in (5.1) but the proof in [GaZw23] still applies as $m(k,0) \equiv 0$. A direct elementary verification is of course much simpler.

This is a special case of the class of one dimensional examples constructed by Seeley [Se86] to show pathological properties of non-normal operators.

2. We can consider $Q(\alpha, k) = D(\alpha) + k$ given in (2.2) with U satisfying (2.3). In [Be*22] we took

$$X = L^2(\mathbb{C}/3\Lambda; \mathbb{C}^2), \quad Y = H^1(\mathbb{C}/3\Lambda; \mathbb{C}^2).$$

In that case the assumptions were satisfied by

$$m(0,k) = 2\, \mathbb{1}_{\frac{1}{3}\Lambda^*}(k), \quad \tau(p)u(z) := e^{i\langle p,z\rangle}u(z).$$

3. In [BHZ22b] we took the point of view closer to the physics literature and had $D(\alpha)$ act on

$$X = L^2_0(\mathbb{C}; \mathbb{C}^2), \quad Y = H^1_0(\mathbb{C}; \mathbb{C}^2),$$

where the spaces were defined in (3.5), so that

$$m(0,k) = \mathbb{1}_{\mathcal{K}_0}(k), \quad \mathcal{K}_0 := \{K, -K\} + \Lambda^*, \quad \tau(p)u(z) := e^{i\langle p,z\rangle}u(z).$$

(The protected states were reviewed in Theorem 1.) The sets $\mathcal{A}$ are the same in both cases. However, there are multiplicity issues illustrated in [BHZ22b, Figure 4].

More interesting examples, in which $m(\alpha_0, k) > \dim\ker Q(\alpha_0, k)$, will be given the next two sections.

5.1. **Spectral characterization.** For operators appearing in TBG (see the examples above) but also in the study of multilayer graphene – see [Be*23],[Ya23] and references given there – the structure of operators $Q(\alpha, k)$ in Theorem 4 is more special.

A natural generalization of $D(\alpha)$ in (2.2) is given as follows

$$\begin{gathered} D(\alpha) := 2D_{\bar z} \otimes I_{\mathbb{C}^n} + W(z) + \alpha V(z) : H^1_{\rm loc}(\mathbb{C}; \mathbb{C}^n) \to L^2_{\rm loc}(\mathbb{C}; \mathbb{C}^n), \\ H(\alpha) := \begin{pmatrix} 0 & D(\alpha)^* \\ D(\alpha) & 0 \end{pmatrix}, \end{gathered} \tag{5.8}$$

where $V(z), W(z) \in C^\infty(\mathbb{C}; \mathbb{C}^n \otimes \mathbb{C}^n)$. Here $2D_{\bar z} := \partial_{x_1} + i\partial_{x_2}$, $z = x_1 + ix_2$, and we will write $2D_{\bar z}$ for the diagonal action on $\mathbb{C}^n$-valued functions.

In (2.2) we had $n = 2$ and $W = 0$ but the presence of W is needed for other models. Mathematically, having that term seems essential when $n > 3$ is considered as it helps in controlling the number of protected states, see (5.10) below. We could consider an even more general case of $W(z) + \alpha V(z)$ replaced by $V(\alpha, z)$.

Let

$$\Lambda = c_\Lambda(\mathbb{Z} + \omega\mathbb{Z}), \quad c_\Lambda \in \mathbb{C}^*, \quad \omega = e^{2\pi i/3}.$$

One nice choice is $c_\Lambda = 1$ (used in [BeZw23a] and later papers and in §2.1 above) but the lattices in the physics literature have different c_Λ. Let $\Lambda^* := c_\Lambda^{-1}(4\pi i/\sqrt{3})\Lambda$, be the dual (reciprocal) lattice.

The class of very general periodicity conditions is given as follows:

$$\begin{gathered} V(z+\gamma) = \rho(\gamma)^{-1} V(z) \rho(\gamma), \quad W(z+\gamma) = \rho(\gamma)^{-1} W(z) \rho(\gamma), \\ \rho(\gamma) := \operatorname{diag} \left[(\exp(i \langle \gamma, k_j \rangle))_{j=1}^n \right], \quad k_j \in \mathbb{C}/\Lambda^*. \end{gathered} \tag{5.9}$$

We remark that $\rho(\gamma)$ is, up to a change of coordinates on $\mathbb{C}^n$ a general unitary representation of the group Λ on $\mathbb{C}^n$.

We then have

$$L_\gamma D(\alpha) = L_\gamma D(\alpha), \quad L_\gamma u(z) := \rho(\gamma) u(z+\gamma),$$

and Bloch–Floquet theory follows the same path as in §3.1 by considering the spectrum of

$$\begin{gathered} H_k(\alpha) := \begin{pmatrix} 0 & D(\alpha)^* + \bar{k} \\ D(\alpha) + k & 0 \end{pmatrix} : H^1_\rho \to L^2_\rho, \\ L^2_\rho := \{ u \in L^2_{\rm loc}(\mathbb{C}; \mathbb{C}^n), \quad L_\gamma u = u \}, \quad H^1_\rho := H^1_{\rm loc} \cap L^2_\rho. \end{gathered}$$

Equivalently we can consider

$$\begin{gathered} D_\rho(\alpha) := \rho(z) D(\alpha) \rho(z)^{-1} = \operatorname{diag} \left[(2D_{\bar z} - k_j)_{j=1}^n \right] + W_\rho(z) + \alpha V_\rho(z), \\ \bullet_\rho(z+\gamma) = \bullet_\rho(z), \quad \bullet_\rho(z) := \rho(z) \bullet (z) \rho(z)^{-1}, \quad \bullet = V, W, \end{gathered} \tag{5.10}$$

which is a periodic operator with respect to Λ and look at the corresponding $H_{\rho,k}(\alpha)$ on Λ-periodic functions.

By putting

$$Q(\alpha, k) := D(\alpha) + k, \quad X = L^2_\rho, \quad Y = H^1_\rho,$$

we can apply Theorem 4 to this case provided that $D(0)$ (corresponding to $\alpha_0 = 0$) has discrete spectrum. If the eigenvalues of $D(0)$ are *semisimple* then

$$m(k, 0) = \dim \ker_{H^1_\rho} (2D_{\bar z} + W(z) + k). \tag{5.11}$$

This happens when $W(z) \equiv 0$ in which case

$$m(k, 0) = |\{ j \in [1, \cdots, n] : k \equiv k_j \bmod \Lambda^* \}|.$$

The advantage of the special form of $D(\alpha)$ is that for $k \notin \operatorname{Spec}_{H_\rho} D(0)$, $(D(0)+k)^{-1} : L^2_\rho \to L^2_\rho$ is a compact operator. Combined with Theorem 4 this gives

Theorem 5 ([Be*22],[BHZ22b],[GaZw23])**.** *Suppose that $Q(\alpha, k) := D(\alpha) + k$ where $D(\alpha)$ is given in (5.8) and that $D(0)$ has discrete spectrum. If for all k (see definition (5.4))*

$$m(\alpha, k) \geq m(0, k), \tag{5.12}$$

then the Birman–Schwinger *operator,*

$$T_z := (D(0) - z)^{-1} W(z) : L^2_\rho \to H^1_\rho \hookrightarrow L^2_\rho, \quad z \notin \operatorname{Spec}(P(0)), \tag{5.13}$$

has discrete spectrum independent of z and, in the notation of Theorem 4,

$$m(k,\alpha) = \begin{cases} \infty, & 1/\alpha \in \operatorname{Spec}(T_z), \\ m(k,0), & \textit{otherwise}. \end{cases} \tag{5.14}$$

In particular, $H(\alpha)$ in (5.8) has a flat band at 0 if and only if $1/\alpha \in \operatorname{Spec}(T_z)$.

Conversely, if the spectrum of T_z is independent of $z \notin \operatorname{Spec} D(0)$, then (5.12) and (5.14) hold.

As pointed out above, this spectral characterization, with magic angles as the spectrum of a compact operator, has been very useful in computing elements of $\mathcal{A}$. Since T_z is non-selfadjoint, pseudospectral issues (see [DSZ04] and references given there), that is the large size of the norm of the resolvent of T_z, enter for large values of α. An explanation of this is provided in §10.1 but a striking numerical illustration is given in Figure 5.

An example of an operator with $n = 3$ can be found in [Be*23] where trilayer graphene was studied (and (5.11) holds). A more interesting case is given by twisted m-sheets of graphene studied mathematically in [Ya23]:

Example. Let us rename the operator $D(\alpha)$ in (2.2) as $D_1(\alpha)$. Following [Ya23] and the physics papers cited there, we put, for $N > 1$,

$$D(\alpha) = D_N(\alpha, \mathbf{t}) := \begin{pmatrix} D_1(\alpha) & t_1 T_+ & & & & \\ t_1 T_- & D_1(0) & t_2 T_+ & & & \\ & t_2 T_- & D_1(0) & \ddots & & \\ & & \ddots & & & \\ & & & & & t_{N-1} T_+ \\ & & & & t_{N-1} T_- & D_1(0) \end{pmatrix}, \tag{5.15}$$

with $\mathbf{t} = (t_1, t_2, \cdots, t_{N-1})$ and

$$T_+ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad T_- = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

To find a suitable ρ in (5.9) we first choose k_1 and k_2 which work for $D_1(\alpha)$ (for instance, as in (3.1)) and then check that k_j for $3 \leq j \leq 2N$ can be chosen consistently so that (5.9) holds. Then $D(\alpha)$ is an example of an operator to which Theorem 5 applies with $n = 2N$. In this case $m(0,k) = N\, \mathbb{1}_{\mathcal{K}}(k) > \dim\ker(D_N(0) + k) = \mathbb{1}_{\mathcal{K}}(k)$ ($\mathcal{K} = \{K, -K\} + \Lambda$ in the case of (3.1)). A direct argument in [Ya23, §4.2] showed that set of α's for which the spectrum of $D(\alpha)$ is a discrete set and that implies that the spectrum T_z in (5.13) is independent of $z \notin \mathcal{K}$. Hence Theorem 5 implies that (5.12) holds but it would be interesting to have a direct argument for that.

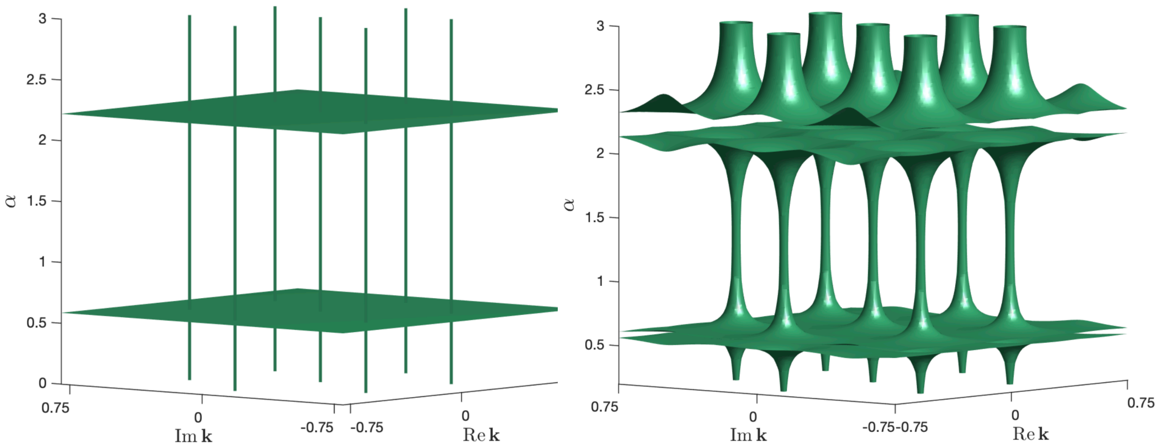


FIGURE 5. : Left: the spectrum of $D(\alpha)$ (in the k plane) as α varies (vertical axis). Flat surfaces indicate that $1/\alpha$ is a magic angle. Right: level surface of $\|(D(\alpha)-k)^{-1}\| = 10^2$ as a function of k and α: the norm blows up at magic angles for all k (α near the magic values 0.586 and 2.221). The thickening of the "trunks" reflects the exponential squeezing of the bands. This figure comes from [Be*22].

5.2. **A scalar model.** One of the difficulties of dealing with the operator $D(\alpha)$ given in (2.2) is that it acts on vector valued functions – some of that will be highlighted in §10. By increasing the order of the operator a scalar model non-equivalent to $D(\alpha)$ but exhibiting flat bands was proposed in [GaZw23].

We first observe that $D(-\alpha)$ is the co-adjoint matrix of $D(\alpha)$ and hence

$$\begin{gathered} D(-\alpha)D(\alpha) = Q(\alpha) \otimes I_{\mathbb{C}^2} + \begin{pmatrix} 0 & \alpha 2D_{\bar z} U(z) \\ -\alpha[2D_{\bar z}U](-z) & 0 \end{pmatrix}, \\ Q(\alpha) := (2D_{\bar z})^2 - \alpha^2 U(z)U(-z). \end{gathered} \tag{5.16}$$

From the semiclassical point (as $\alpha \to \infty$) of view the non-scalar term in (5.16) is of lower order (see §10) and is natural to consider the operator $Q(\alpha)$ on its own.

We can then consider a self-adjoint Hamiltonian, on $L^2(\mathbb{C};\mathbb{C}^2)$ with the domain given by $H^2(\mathbb{C};\mathbb{C}^2)$ (note that $D_{\bar z}^2$ is an elliptic operator),

$$H(\alpha) := \begin{pmatrix} 0 & Q(\alpha)^* \\ Q(\alpha) & 0 \end{pmatrix}. \tag{5.17}$$

This is a periodic operator with respect to the lattice Λ (note that for U satisfying (2.3), $U(z)U(-z)$ is Λ-periodic). And Floquet theory (as reviewed in §3.1) corresponds to studying the spectra of

$$H(\alpha,k) = \begin{pmatrix} 0 & Q(\alpha,k)^* \\ Q(\alpha,k) & 0 \end{pmatrix}, \quad Q(\alpha,k) := (2D_{\bar z}+k)^2 - \alpha^2 U(z)U(-z),$$

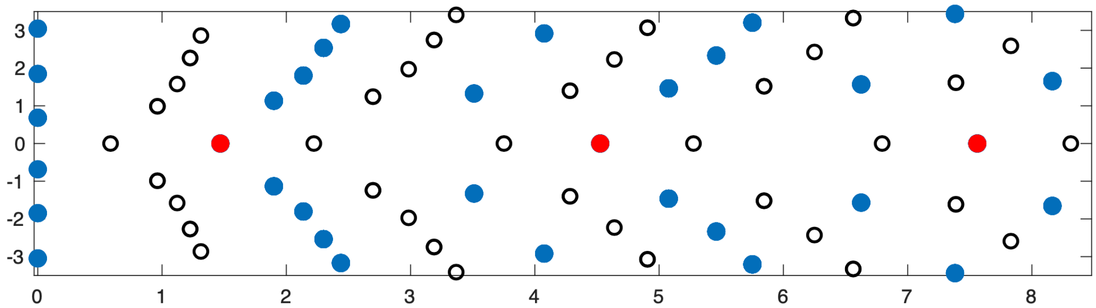


FIGURE 6. Comparison of the set of magic α's, $\mathcal{A}$ for the potential $U = U_{\rm BM}$ given in (2.5) (shown as $\circ$) and $\mathcal{A}_{\rm sc}$ the set for which (5.18) holds (with the same U; shown as $\bullet$). The real elements of $\mathcal{A}_{\rm sc}$ are shown as $\bullet$. They appear to have multiplicity two. When we interpolate between the chiral model and the scalar model, the multiplicity two real α's split and travel in opposite directions to become magic α's for the chiral model: see https://math.berkeley.edu/~zworski/Spec.mp4.

on $L^2(\mathbb{C}/\Lambda;\mathbb{C}^2)$ and with the domain $H^2(\mathbb{C}/\Lambda;\mathbb{C}^2)$.

A flat band of $H(\alpha)$ given in (5.17) corresponds to

$$\forall\, k \in \mathbb{C} \quad 0 \in \operatorname{Spec} H_k(\alpha) \iff \forall\, k \in \mathbb{C} \quad \ker_{H^1(\mathbb{C}/\Lambda)} Q(\alpha,k) \neq \{0\}. \tag{5.18}$$

To apply Theorem 4 we need to verify the first inequality in (using the definition (5.4))

$$m(\alpha,k) \geq m(0,k) = 2\,\mathbb{1}_{\Lambda^*}(k) > \dim\ker Q(0,k) = \mathbb{1}_{\Lambda^*}(k),$$

see [GaZw23, §3]. (Just as in the case of (5.15) it is important to consider the generalized multiplicities.) It then follows that there exists a discrete set $\mathcal{A}_{\rm sc}$ such that (5.18) holds if and only if $\alpha \in \mathcal{A}_{\rm sc}$ – see Figure 6.

The next two problems are probably the most doable on the list.

Open Problem 6. *Adapt the theta function argument recalled in §6 to the scalar model and show that the multiplicity of the flat bands is at least* 2.

Open Problem 7. *Adapt the trace argument is §7 to show that for the potential $U = U_{\rm BM}$ in (2.3), $|\mathcal{A}_{\rm sc}| = \infty$.*

The situation is less clear for the next

Open Problem 8. *Is the spectrum of $Q(\alpha,k)$ discrete for all α and k? Characterize the set for which $\operatorname{Spec} Q(\alpha,k) = \emptyset$. (We should stress that this is a mathematical curiosity: only the fact that $0 \in \operatorname{Spec} Q(\alpha,k)$ is relevant to the question of band theory, bands being given by characteristic values of $Q(\alpha,k)$ as k varies.*

The next problem is the analogue of Open Problem 1. One would like to hope that the scalar nature of the operator could be of some help in the semiclassical analysis (see §10).

Open Problem 9. *If $\beta_1 < \beta_2 < \cdots$ is the ordered sequence of elements of $\mathcal{A}_{\rm sc} \cap [0,\infty)$ then, for the potential* (2.5),

$$\beta_{j+1} - \beta_j = 2\gamma + o(1), \quad j \to \infty$$

where $\gamma \simeq \frac{3}{2}$ is the asymptotic spacing between the elements of $\mathcal{A} \cap [0,\infty)$ (see (2.9)*). What happens for more general potentials satisfying* (2.3)*?*

6. Theta function argument for magic angles

Magic angles for the chiral model were described in [TKV19] using a different approach than that recalled in §5 and coming from [Be*22]. It was based on an idea which appeared earlier, in a different but related context, in the work of Dubrovin–Novikov [DuNo80]. It was revisited in [BHZ22b] and here we present a slightly different variant.

The operator $2D_{\bar z} := (1/i)(\partial_{x_1} + i\partial_{x_2})$, $z = x_1 + ix_2$, acting on $L^2(\mathbb{C}/\Lambda;\mathbb{C})$ with the domain given by $H^1(\mathbb{C}/\Lambda;\mathbb{C})$ is a normal operator (a sum of two commuting sel-adjoint operators). Its spectrum is given by Λ^* with simple eigenvalues and normalized eigenfunctions given by $v(p) = \tau(p)v(0)$, $v(0) := |\mathbb{C}/\Lambda|^{-\frac{1}{2}}$, $[\tau(p)u](z) = e^{i\langle z,p\rangle}u(z)$, $p \in \Lambda^*$. Hence, its resolvent

$$(2D_{\bar z} + k)^{-1} : L^2(\mathbb{C}/\Lambda;\mathbb{C}) \to H^1(\mathbb{C}/\Lambda;\mathbb{C}),$$

is a *meromorphic family* of operators with simple poles at $p \in \Lambda^*$ and residues $|v(p)\rangle\langle v(p)|$. Since $2D_{\bar z}$ is translation invariant, we have

$$(2D_{\bar z} + k)^{-1} f(z) = \int_{\mathbb{C}/\Lambda} G_k(z-\zeta) f(\zeta) dm(\zeta), \quad (2D_{\bar z} + k) G_k(z) = \delta_0(z), \quad k \notin \Lambda^*.$$

If $a(k)$ is *any* entire function with the zero set given by simple zeros at Λ^*, then

$$\begin{gathered} k \mapsto F_k(z) := a(k) G_k(z), \quad \text{is a holomorphic family of distributions,} \\ (2D_{\bar z} + k) F_k(z) = a(k)\delta_0(z). \end{gathered} \tag{6.1}$$

If u_K is the protected state described in Theorem 1 and $z_0 \in \mathbb{C}/\Lambda$, then (6.1) gives (note that u_K is valued in $\mathbb{C}^2$ and F_k is scalar valued)

$$(2D_{\bar z} + k)(F_{k-K}(z - z_0) u_K(\alpha, z)) = a(k-K) u_K(\alpha, z_0) \delta(z - z_0). \tag{6.2}$$

Hence,

$$\exists\, z_0 \;\; u_K(\alpha, z_0) = 0 \implies \forall k\, \exists\, u(k) \in H^1_0 \;\; (D(\alpha) + k) u(k) = 0, \quad \|u(k)\|_{L^2_0} = 1. \tag{6.3}$$

The required vanishing condition is strong: we are looking for simultaneous vanishing of two complex valued functions of the complex variable (components of u_K).

Following [TKV19] we observe

$$\tau(K)u_K(\alpha,z)=\begin{pmatrix}\psi(z)\\ \varphi(z)\end{pmatrix}\implies\forall\,\alpha\quad\varphi(z(K))=0.$$

(Here we use the notation of (6.11) and recall from (3.8) and (3.12) that $L_\gamma\tau(K)u_K(\alpha)=e^{i\langle\gamma,K\rangle}\tau(K)u_K(\alpha)$ and that $\tau(K)u_K(\alpha,\omega z)=\tau(K)u_K(\alpha,z)$ which then implies, following the definitions, that $\varphi(z(K))=\bar\omega\varphi(z(K))$.) Using (3.12) we have

$$\tau(-K)u_K(\alpha,z)=\begin{pmatrix}\varphi(-z)\\ -\psi(-z)\end{pmatrix}.$$

Since $D(\alpha)(\tau(\pm K)u_{\pm K}(\alpha))=0$, the Wronskian of $\tau(\pm K)u_{\pm K}$ is a holomorphic Λ-periodic function. Hence it is a constant depending only on α:

$$v_F(\alpha):=\frac{\psi(z)\psi(-z)+\varphi(z)\varphi(-z)}{\|u_K(\alpha)\|^2}=\frac{\psi(z(K))\psi(-z(K))}{\|u_K(\alpha)\|^2}.\tag{6.4}$$

(For an interesting physical interpretation of $v_F(\alpha)$ as the Fermi velocity see [TKV19, (8),(21),(22)]. We lose holomorphy in α because of the normalization.) We conclude that

$$\exists\,z_0\quad u_K(\alpha,z_0)=0\iff v_F(\alpha)=0\iff\exists\,\varepsilon\in\{+,-\}\quad u_K(\alpha,\varepsilon z(K))=0.\tag{6.5}$$

This argument, essentially from [TKV19], establishes one implication in the first statement of

Theorem 6 ([TKV19],[Be*22],[BHZ22b])**.** *For any potential U satisfying (2.3), $\mathcal{A}$ defined in §3.2 and $v_F(\alpha)$ defined in (6.4), we have*

$$v_F(\alpha)=0\iff\alpha\in\mathcal{A}.\tag{6.6}$$

Moreover, if $\alpha\in\mathcal{A}$ is simple *then*

$$u_K(\alpha,z_0)=0\implies z_0=z(K),\tag{6.7}$$

and the zero is simple: $u_K(\alpha,z)=(z-z(K))w(z)$, $w\in C^\infty$, $w(z(K))\neq0$.

The implication $v_F(\alpha)\neq0\Rightarrow\alpha\notin\mathcal{A}$ follows easily from building a formula for $(D(\alpha)+k)^{-1}$ using $u_{\pm K}(\alpha)$ – see [Be*22, Proposition 3.3]. The implication (6.7) is a special case of [BHZ22b, Theorem 3]. The point $z(K)=:-z_S$ is called a stacking point – see Figure 9. The proof of (6.7) was simplified in [Be*23] in a way which allowed an adaptation to the trilayer case. For an animation showing the behaviour of $u_K(\alpha)$ as α increases along the real axis (for the potential (2.5)), see `https://math.berkeley.edu/~zworski/magic.mp4`.

We recall another characterization of simple $\alpha\in\mathcal{A}$:

Theorem 7 ([BHZ22b]). *We have the following equivalence (using definition (3.15) and denoting $\mathcal{K}_0 := \{K, -K\} + \Lambda^*$)*

$$\begin{aligned} m(\alpha) = 1 &\iff \forall\, k \in \mathbb{C} \quad \dim \ker_{L^2_0(\mathbb{C}/\Lambda)}(D(\alpha) + k) = 1 \\ &\iff \exists\, p \notin \mathcal{K}_0 \quad \dim \ker_{L^2_0(\mathbb{C}/\Lambda)}(D(\alpha) + p) = 1. \end{aligned} \tag{6.8}$$

Returning to (6.2) and (6.3) we see that for $\alpha \in \mathcal{A}$, simple, we can take (see [BeZw23a, (3.43)])

$$u(k, z) = c(k) F_k(z) u_0(z), \quad \ker_{H^1_0} D(\alpha) = \mathbb{C} u_0, \quad \ker_{H^1_0}(D(\alpha) + k) = \mathbb{C} u(k), \tag{6.9}$$

where $c(k)$ is the normalizing constant so that $\|u(k)\|_{L^2_0} = 1$. (We know that in this case u_0 has a simple zero at 0 – see [BeZw23a, Proposition 3.6]. Please note that $u_0 \in L^2_0$ exists only for $\alpha \in \mathcal{A}$, unlike $\tau(\pm K) u_{\pm K} \in L^2_{\pm K}$, $D(\alpha)\tau(\pm K) u_{\pm K} = 0$ which exist for all α.) Using symmetries of $D(\alpha)$ we can also describe the kernel of $(D(\alpha)+k)^*$ and that can be done in different ways. Following [BeZw23a, (3.44)] we can take (with the advantage that it works also for more general potentials (7.7))

$$u^*(k, z) = c(k) \overline{F_{-k}(z)} \begin{pmatrix} \overline{\varphi_0(z)} \\ -\overline{\psi_0(z)} \end{pmatrix}, \quad u_0 =: \begin{pmatrix} \psi_0 \\ \varphi_0 \end{pmatrix}, \quad \ker_{H^1_0}(D(\alpha) + k)^* = \mathbb{C} u^*(k), \tag{6.10}$$

and $\|u^*(k)\|_{L^2_0} = 1$. (For other choices of $u^*(k)$ when $D(\alpha)$ is given by (2.2) see [BeZw23b, (2.9)].)

There are many choices for F_k (that is, choices of entire functions with simple zeros precisely at Λ^*) and we can for instance follow [BeZw23a] and take

$$\begin{gathered} F_k(z) := e^{\frac{i}{2}(z-\bar z)k} \frac{\theta(z - z(k))}{\theta(z)}, \quad z(k) = \frac{\sqrt{3}}{4\pi i} k, \quad a(k) := \frac{2\pi i \theta(z(k))}{\theta'(0)}, \\ \theta(z) := \theta_1(z|\omega) := -\sum_{n \in \mathbb{Z}} \exp(\pi i (n + \tfrac12)^2 \omega + 2\pi i (n + \tfrac12)(z + \tfrac12)), \end{gathered} \tag{6.11}$$

that is θ is the first Jacobi theta function and its simple zeros coincide with Λ – see [Mu83] or [KhZa15]. Weierstrass σ function was used explicitly in [DuNo80] and the theta function in [TKV19], but in fact it is only the canonical nature of Green's function and the set Λ^* that matter (though of course constructing a function which vanishes precisely at Λ^* hides those special functions).

7. Existence and multiplicities of magic angles

So far we have not addressed the question of existence of magic α's, and in particular of existence of real simple α's (see the definition in §3.2). It is not clear if there exist more than one physical magic angle and the current experimental and theoretical evidence suggests that there may only be one. The work of Becker–Oltman–Vogel [BOV23] on random perturbations of TBG provides some mathematical evidence for that.

In the chiral model rigorous existence and simplicity of the first real magic angle has however been established:

Theorem 8 ([WaLu21],[BHZ22a])**.** *For the potential* (2.5) *and for the (discrete) set of magic α's, $\mathcal{A}$, defined in §3.2, we have*

$$\min \mathcal{A} \cap [0,\infty) = \alpha_1 \simeq 0.586. \tag{7.1}$$

In addition, in the sense of (3.15),

$$m(\alpha_1) = 1, \tag{7.2}$$

that is, α_1 is simple.

Watson and Luskin [WaLu21] followed the approach of [TKV19] and proved existence of a zero of $v_F(\alpha)$ given in (6.4) (see Theorem 6). That was done by a careful analysis of the Taylor series at 0, with precise estimates of the remainder, and floating point arithmetic.

The approach of [BHZ22a] was based on the spectral characterization from [Be*22] (see §5) and the evaluation, theoretical and numerical, of sums of powers of magic α's:

Theorem 9 ([Be*22],[BHZ22a])**.** *For the potential in* (2.5) *we have*

$$\sum_{\alpha\in\mathcal{A}} \alpha^{-4} = \frac{8\pi}{\sqrt{3}}, \tag{7.3}$$

and more generally, for $p \in \mathbb{N}+2$,

$$\sum_{\alpha\in\mathcal{A}} \alpha^{-2p} \in \frac{\pi}{\sqrt{3}}\mathbb{Q}. \tag{7.4}$$

In the above sums the multiplicity of $\alpha \in \mathcal{A}$ is given by the algebraic multiplicity of $1/\alpha$ as an eigenvalue of T_k, $k \notin \Lambda^*$, where T_k is the Birman–Schwinger operator (5.13).

These identities are based on writing $\sum_{\alpha\in\mathcal{A}} \alpha^{-2p} = \operatorname{tr} T_k^{2p}$, and (7.3) was proved in [Be*22, §3.3] (the sum in (7.4) with $p=4$ was also given as $80\pi/\sqrt{3}$; since there we considered action on $L^2(\mathbb{C}/3\Lambda)$ rather than on L^2_0, the multiplicities were nine fold higher; we note that for odd powers of T_k the traces are 0 in view of (2.8)). The far reaching generalization in (7.4) happened thanks to the expansion of the collaboration in [BHZ22a]. It holds for a greater class of potentials. The existence of algebraic multiplicities greater than geometric multiplicities (Jordan blocks) is suggested by numerical experiments – see [BHZ23, §10.1].

The method for proving (7.3) provides an algorithm for finding the rational number $(\sqrt{3}/\pi)\operatorname{tr} T_k^{2p}$. This allows a precise evaluation of regularized determinants of $I - T_k$ and that leads to an alternative proof of (7.1) and a proof of (7.2).

An immediate consequence of (7.3), (7.4), the transcendental nature of $\pi/\sqrt{3}$, and of Newton identities is (see [BHZ22a, Theorem 6] for a more general version):

Theorem 10 ([BHZ22a]). *For the potential* (2.3),

$$|\mathcal{A}| = \infty. \tag{7.5}$$

Before moving to the discussion of higher multiplities we present some open problems related to the above theorems. They all seem quite hard.

Open Problem 10. *Show that* (7.5) *holds for any non zero potential satisfying* (2.3).

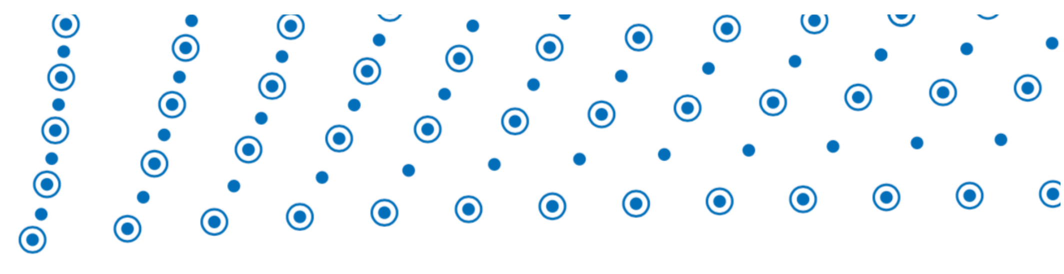
Open Problem 11. *Using Theorem 5 it is not difficult to see that* $|\{\alpha \in \mathcal{A} : |\alpha| \leq r\} \leq Cr^2$. *Do we have lower bounds? Is there a way to use methods of Christiansen* [Ch99] *("plurisubharmonic magic") to obtain results for generic potentials?*

Open Problem 12. *Show that for the potential* (2.5) *and* α_1 *given in* (7.1) *we have*

$$\tfrac{1}{2}\alpha_1^{2p} \sum_{\alpha \in \mathcal{A}} \alpha^{-2p} \to 1, \quad p \to \infty, \quad p \in \mathbb{N}.$$

This seems to be the case numerically as, $\min\{|\alpha| : \alpha \in \mathcal{A} \setminus \{\pm\alpha_1\}\} > 1$. *Any type of asymptotic result about* $\operatorname{tr} T_k^{2p}$ *would be interesting.*

We now turn attention to higher multiplicities. Figure 3 showed numerically computed multiplicities, including $\alpha \in \mathcal{A} \cap \mathbb{R}$ with $m(\alpha) > 1$ (see (3.15) for the definition of multiplicity). For the BM potential (2.5) "half" of the complex α's have multiplicity two (indicated by circles; we show α's is the first quadrant):



Using Theorem 12 below and analysis of traces of T_k^{2p} restricted to different spaces $L^2_{p,k}$ we obtain a partial mathematical confirmation of the above figure:

Theorem 11 ([BHZ23]). *For the Bistritzer–MacDonald potential* (2.5)

$$|\{\alpha \in \mathcal{A} : m(\alpha) > 1\}| = \infty,$$

that is, there exist infinitely many (complex) degenerate magic α's.

The double α's shown in the above figure are protected as we have a surprising rigidity result expressed using the spaces defined in (3.8):

Theorem 12 ([BHZ23])**.** *For any potential satisfying* (2.3) *we have, with the definition of multiplicity* (3.15)*,*

$$\begin{aligned} m(\alpha) = 1 &\implies \dim \ker_{L^2_{0,2}} D(\alpha) = 1, \\ m(\alpha) = 2 &\implies \dim \ker_{L^2_{0,0}} D(\alpha) = \dim \ker_{L^2_{0,1}} D(\alpha) = 1. \end{aligned} \tag{7.6}$$

In particular, a multiplicity two $\alpha \in \mathcal{A}$ cannot *be split into simple α's by deforming a potential within the class* (2.3)*.*

In [BHZ23, Theorem 4] we also have an analogue of Theorem 7 for the case of double α's.

Open Problem 14. *As suggested by* (6.3) *the multiplicity of α is closely related to the number of zeros (counted with multiplicity) of the eigenstate of $D(\alpha)$. For $1 \leq m(\alpha) \leq 2$, Theorem 12 can be used to obtain the precise description (see §8). What is the situation for higher multiplicities?*

It is natural to ask if generically we only have simple or double magic α's. We have established it by expanding the class of allowed potentials:

$$D(\alpha) := 2D_{\bar z} \otimes I_{\mathbb{C}^2} + W(z), \quad W(z) := \begin{pmatrix} 0 & \alpha U_+(z) \\ \alpha U_-(z) & 0 \end{pmatrix}, \tag{7.7}$$

where the potentials satisfy

$$U_\pm(z+\gamma) = e^{\pm i\langle \gamma, K\rangle} U_\pm(z), \quad \gamma \in \Lambda, \quad U_\pm(\omega z) = \omega U_\pm(z). \tag{7.8}$$

The self-adjoint Hamiltonian $H(\alpha)$ is defined by (2.6) and commutation relations (3.2),(3.4) still hold. We then have the same Bloch–Floquet theory as in §3.1 and the same definitions of $\mathcal{A}$ and $m(\alpha)$ (see §3.2).

As the space of allowed potentials W we use a Hilbert space of *real analytic* functions equipped with the following norm: for a fixed $\delta > 0$,

$$\|W\|_\delta^2 := \sum_{\pm} \sum_{k \in \Lambda^*/3} |a_k^\pm|^2 e^{2|k|\delta}, \qquad U_\pm(z) = \sum_{k \in K + \Lambda^*} a_k^\pm e^{\pm i \langle z, k\rangle}. \tag{7.9}$$

Then we define $\mathscr{V} = \mathscr{V}_\delta$ by

$$W \in \mathscr{V} \iff W \text{ satisfies (7.8)}, \ \|W\|_\delta < \infty. \tag{7.10}$$

With this in place we can state

Theorem 14 ([BHZ23])**.** *There exists a generic subset (an intersection of open dense sets), $\mathscr{V}_0 \subset \mathscr{V}$, such that if $W \in \mathscr{V}_0$ then for all $\alpha \in \mathcal{A}$ (defined using* (7.7)*)*

$$1 \leq m(\alpha) \leq 2.$$

A more precise formulation related to Theorem 12 is given in [BHZ23, Theorem 3].

Open Problem 15. *Does Theorem 14 hold for a generic set of potentials satisfying* (2.3)*?*

8. Topology of flat bands

Topology of flat bands refers to the topology of vector bundles over the k-space torus $\mathbb{C}/\Lambda^*$ obtained by considering eigenfunctions of $H_k(\alpha) = H_k(\alpha, 0)$ (see (3.9)) for $\alpha \in \mathcal{A}$, that is for α's at which we have perfectly flat bands. The eigenfunctions are given by

$$\Phi := \begin{pmatrix} u \\ v \end{pmatrix}, \quad u \in \ker_{H^1_0}(D(\alpha)+k), \quad v \in \ker_{H^1_0}(D(\alpha)^*+\bar{k}), \quad H_k(\alpha)\Phi = 0. \tag{8.1}$$

The two components u, v, are completely decoupled and hence we can consider them separately. Symmetries of $D(\alpha)$ (see [BeZw23b, §2.2] for a quick review) show that we only need to consider $\ker_{H^1_0}(D(\alpha)+k)$. As we already mentioned, the nontrivial topology implies blow up of moments of Wannier functions corresponding to lack of localization – see [TaZw23a, Theorem 9, §8.5] and references given there.

We now assume that $\alpha \in \mathcal{A}$ and that

$$1 \le m(\alpha) \le 2, \tag{8.2}$$

that is the band has multiplicity one or two in the sense of §3.2. In view of Theorem 7 and [BHZ23, Theorem 4] we have

$$V(k) := \ker_{H^1_0}(D(\alpha)+k) \subset L^2_0, \quad \dim V(k) = m(\alpha), \quad k \in \mathbb{C}, \tag{8.3}$$

and we can define a trivial vector bundle $\widetilde{E} \to \mathbb{C}$ of rank $m(\alpha)$:

$$\widetilde{E} := \{(k,v) : v \in V(k)\} \subset \mathbb{C} \times L^2_0(\mathbb{C}/\Lambda; \mathbb{C}^2).$$

To define a vector bundle over the torus $\mathbb{C}/\Lambda^*$ we need an equivalence relation on $\mathbb{C} \times L^2_0(\mathbb{C}/\Lambda; \mathbb{C}^2)$ based on

$$\begin{gathered} \tau(p)^* H_k(\alpha)\tau(p) = H_{k+p}(\alpha), \quad \tau(p)^*(D(\alpha)+k)\tau(p) = D(\alpha)+k+p, \\ \tau(p)^{-1}V(k) = V(k+p), \quad [\tau(p)u](z) := e^{i\langle z,p\rangle}v(z), \quad p \in \Lambda^*. \end{gathered} \tag{8.4}$$

It is given as follows:

$$\exists\, p \in \Lambda^* \ (k,u) \sim_\tau (k+p, \tau(p)^{-1}u) \tag{8.5}$$

Using this (see [TaZw23a, Lemma 8.4] or [BHZ22b, Lemma 5.1]),

$$E := \widetilde{E} \,/ \sim_\tau \to \mathbb{C}/\Lambda^* \tag{8.6}$$

is a holomorphic vector bundle over $\mathbb{C}/\Lambda^*$. In the case of $m(\alpha) = 1$ (and up to precise definitions) this observation was made by Ledwith et al [Le*20]. In view of (6.9) the line bundle can be identified with a theta bundle over the torus – see [BHZ22b, §5.3].

A natural connection on this vector bundle can be defined either as the Chern connection or the Berry connection, as they are equal in the holomorphic case – see [BHZ23, §9, Proposition 9.1] for a detailed presentation and definitions. The scalar curvature of this connection is a two form on $\mathbb{C}/\Lambda^*$,

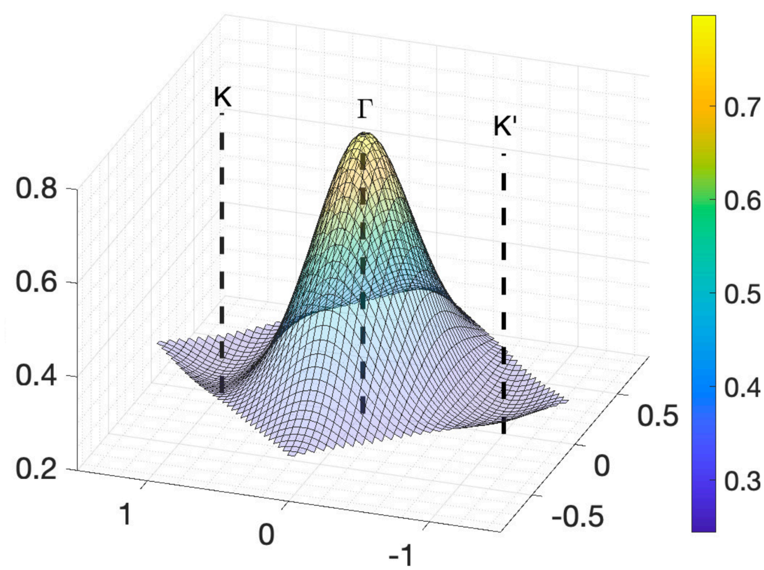
$$\operatorname{tr}\Theta = H(k)d\bar{k} \wedge dk, \tag{8.7}$$

see [BHZ23, §9]. Here Θ is the curvature form taking values in $\operatorname{Hom}(E, E)$. The following observations were made in [BHZ22b, §5.2] and [BHZ23, §9.3]:

$$H(k) \geq 0, \quad H(\omega k) = H(k), \quad H(k) = H(-k). \tag{8.8}$$

In particular, $\mathcal{K} = \{0, K, -K\}$ (see (3.7)) is contained in the set of critical points of H.

Open Problem 16. *Show that for the potential* (2.3) *(or for a more general class of potentials?) and $\alpha \in \mathcal{A} \cap \mathbb{R}$ (or simply for α_1 in* (7.1)*), $\mathcal{K}$ is the set of all critical points of $H(k)$, and that the maximum is attained at 0 (the Γ point) and the minimum at $\pm K$ (the K-points):*



For a discussion of analogous issues when multiplicity is equal to 2, see [BHZ23, §10.2].

The Chern number for complex vector bundles over a torus is defined using (8.7):

$$c_1(E) := \frac{i}{2\pi} \int_{\mathbb{C}/\Lambda^*} \operatorname{tr}\Theta = -\frac{1}{\pi} \int_F H(k) dm(k), \tag{8.9}$$

where F is a fundamental domain of Λ^* and $dm(k) = dxdy$, $k = x + iy$, the Lebesgue measure. We have $c_1(E) \in \mathbb{Z}$ (see [TaZw23a, Theorem 6] and references given there) and if $c_1(E) \neq 0$ then the vector bundle is *non-trivial*, that is it is not homeomorphic to $\mathbb{C}/\Lambda^* \times \mathbb{C}^n$. For complex vector bundles over tori $c_1(E)$ is the only topological invariant. (For instance, for a simple α we could consider the complex vector bundle defined using $\ker_{H_0^1(\mathbb{C},\mathbb{C}^4)} H_k(\alpha)$, see (8.1). Its Chern number vanishes and the bundle is trivial.)

For simple α's an evaluation of $c_1(E)$ follows easily from (6.9) – see [Le*20] for a direct calculation and [BHZ22b, (5.9),(B.8)] for an argument based on general principles. It turns out [BHZ23, Theorem 5] that the Chern number does not change if α is double:

Theorem 15 ([BHZ22b],[BHZ23])**.** *Suppose that (8.2) holds and that the complex vector bundle E is defined by (8.6). Then the Chern number defined in (8.9) is given by*

$$c_1(E) = -1. \tag{8.10}$$

Yang [Ya23] provided a mathematical justification of the Chern number calculation in [LVK22],[WaLi22] (and of other issues related to flat bands in their setting) for two twisted n-layer wafers of graphene. In that case, the analogue of the line bundle (8.6) satisfies $c_1(E) = -n$.

Open Problem 17. *Does (8.10) hold without the assumption (8.2)?*

9. Dynamics of Dirac points for in-plane magnetic field

Interesting mathematical phenomena arise when a constant magnetic field in the direction parallel to the two twisted layers of graphene is added. Following Kwan–Parameswaran–Sondhi [KPS20] and Qin–MacDonald [QiMa21] the new Hamiltonian for the chiral model is given by

$$H_B(\alpha) := \begin{pmatrix} 0 & D_B(\alpha)^* \\ D_B(\alpha) & 0 \end{pmatrix}, \quad D_B(\alpha) := D(\alpha) + \mathcal{B}, \quad \mathcal{B} := \begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix}, \tag{9.1}$$

where $B = |B|e^{2\pi i\theta}$ with $|B|$ corresponding to the strength of the magnetic field and $2\pi\theta$ is its in-plane direction; $D(\alpha)$ is the same as in (2.2).

For the BMH and the chiral model, the bands close to zero touch at 0 at $\pm K$ (see (3.7) – these are the K-points in our coordinates) and the intersection is expected to be conic (except for the perfectly flat bands), that is we see two Dirac points – see Open Problem 2 and the figure there. Theorem 2 shows that for the chiral model, $H(\alpha) = H(\alpha, 0)$ in the notation of (2.1), once the bands touch 0 away from $\pm K$, the bands are perfectly flat.

It was observed numerically in [KPS20] that for the chiral model with in-plane magnetic field (9.1) flat bands disappear when $B \neq 0$ and the two Dirac points move. Moreover for $\alpha \in \mathcal{A}$ the Dirac points seem to coalesce at the Γ point forming a quadratic band crossing point (QBCP) – see Figure 8. In [BeZw23a] we provided a

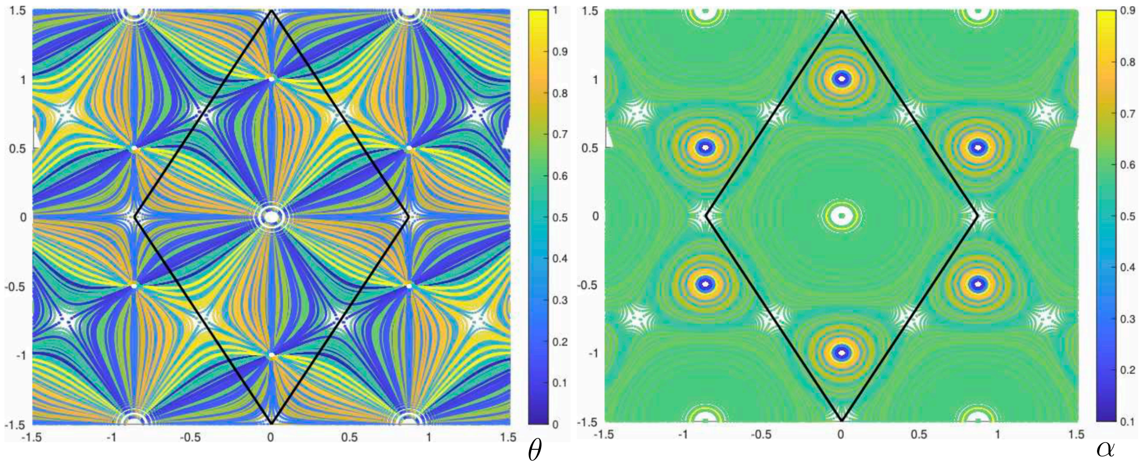


FIGURE 7. The dynamics of Dirac points for H_B in (9.1) with the BM potential (2.5). The magnetic field given by $B = B_0 e^{2\pi i\theta}$ with $B_0 = 0.1$. Colour coding (shown in colour bars) corresponds to different values of θ on the left, and different values of α on the right. In the left figure α varies between 0.1 and 0.9 and curves of different colour trace the corresponding Dirac points – see `https://math.berkeley.edu/~zworski/B01.mp4` for an animated version. When $3\theta \in \mathbb{N}$ we showed in [BeZw23a, Theorem 3] that the Dirac points move along straight lines – see `https://math.berkeley.edu/~zworski/Rectangle_1.mp4` where $\theta = \frac{1}{3}$. In the right figure, θ varies and curves of different colour trace the corresponding Dirac points. The predominance of green (corresponding to the range between 0.5 and 0.6) means that most of the motion happens near the (first) magic alpha – for the dance of Dirac points for fixed B and as α varies, see `https://math.berkeley.edu/~zworski/first_band.mp4` which shows $E_1(\alpha, k)/\max_k E_1(\alpha, k)$. (The boundary Brillouin zone is also shows; we take the image of the k plane by the map $k \mapsto z(k)$, see (3.6) so that Λ^* is mapped to $\mathbb{Z} + \omega\mathbb{Z}$.)

more precise description of the dynamics of Dirac points for small magnetic fields. In particular finer analysis and numerical evidence suggest that exact QBCP appear only when Dirac points move along straight lines which happens when $3\theta \in \mathbb{N}$ (the direction of the magnetic field is given by $2\pi\theta$) – see [BeZw23a, Theorem 3 and Figure 5].

The reason for the Dirac points appearing close to Γ when α is close to (simple) elements of $\mathcal{A}$, can be elegantly described using properties of theta functions. Since it is a simple consequence of (6.9) and (6.10) we recall it, referring to [BeZw23a, §4] for additional details. This also allows to present an approach to perturbation theory based on Schur's complement formula (via Grushin problems in the terminology of Sjöstrand who turned Schur's complement formula into a systematic tool) – see [TaZw23a, §2.6]. Same approach is used to obtain Theorem 3.

Suppose that $\alpha \in \mathcal{A}$ is simple, and in the notation of (6.9) and (6.10) the operator (see §1 for the review of notation)

$$\mathscr{D}(\alpha, k) := \begin{pmatrix} D(\alpha) + k & |u^*(k)\rangle \\ \langle u(k)| & 0 \end{pmatrix} : H_0^1 \times \mathbb{C} \to L_0^2 \times \mathbb{C},$$

is invertible with the inverse given by

$$\mathcal{E}(\alpha, k) = \begin{pmatrix} E(k) & |u(k)\rangle \\ \langle u^*(k)| & E_{-+}(k) \end{pmatrix} : L_0^2 \times \mathbb{C} \to H_0^1 \times \mathbb{C},$$

where $E_{-+}(k) \equiv 0$ is the effective Hamiltonian: from Schur's complement formula [TaZw23a, (2.15)] we see that $D(\alpha) + k$ is invertible if and only if $E_{-+}(k) = 0$. Since $\alpha \in \mathcal{A}$, $\operatorname{Spec}_{L_0^2} D(\alpha) = \mathbb{C}$ this is consistent with $E_{-+}(k) \equiv 0$. For $|B| \ll 1$, we can consider $D_B(\alpha)$ as a perturbation of $D(\alpha)$ and we still have invertibility

$$\begin{gathered} \begin{pmatrix} D_B(\alpha) + k & |u^*(k)\rangle \\ \langle u(k)| & 0 \end{pmatrix}^{-1} = \begin{pmatrix} E^B(k) & E_+^B(k) \\ E_-^B(k) & E_{-+}^B(k) \end{pmatrix}, \\ E_{-+}^B(k) = -\langle u^*(k)|\mathcal{B}|u(k)\rangle + \mathcal{O}(B^2), \end{gathered}$$

see [TaZw23a, Proposition 2.12]. From (6.9) and (6.10) we then obtain that

$$E_{-+}^B(k) = -c(k)^{-2} B(G(k) + \mathcal{O}(B)), \quad G(k) = 2 \int_{\mathbb{C}/\Lambda} F_k(z) F_{-k}(z) \varphi_0(z) \psi_0(z) dm(z),$$

where F_k is defined in (6.11). This definition combined with a theta function identity

$$\theta(z+u)\theta(z-u)\theta(\tfrac{1}{2})^2 = \theta^2(z)\theta^2(u+\tfrac{1}{2}) - \theta^2(z+\tfrac{1}{2})\theta^2(u),$$

and symmetries of ψ_0 and φ_0 (see §[BeZw23a, §4.1]) gives

$$G(k) = g_0 \frac{\theta(z(k))^2}{\theta(\frac{1}{2})^2}, \quad g_0 = g_0(\alpha) := 2 \int_{\mathbb{C}/\Lambda} \theta(z+\tfrac{1}{2})^2 \frac{\varphi_0(z)\psi_0(z)}{\theta(z)^2} dm(z). \tag{9.2}$$

For the Bistritzer–MacDonald potential and the first magic angle, α_1 (see Theorem 8) $|g_0| \simeq 0.07 \neq 0$. We now see that k is a Dirac point for (9.1) with $\alpha = \alpha_1$ if and only if $E_{-+}^B(k) = 0$, and in particular

$$k \in \operatorname{Spec}_{L_0^2} D_B(\alpha_1) \implies \theta(z(k))^2 + \mathcal{O}(B) = 0. \tag{9.3}$$

(For $g_0(\alpha)$ for other real magic α's see [BeZw23a, Table 1].)

Since $\theta(z(k))^2$ vanishes quadratically at 0 (the Γ point), equation (9.3) shows that at $\alpha = \alpha_1$ and for B small the Dirac points are near the Γ point. It also suggests QBCP – see Figure 8 and [BeZw23a, §5] for a discussion of the bifurcation at Γ and other points.

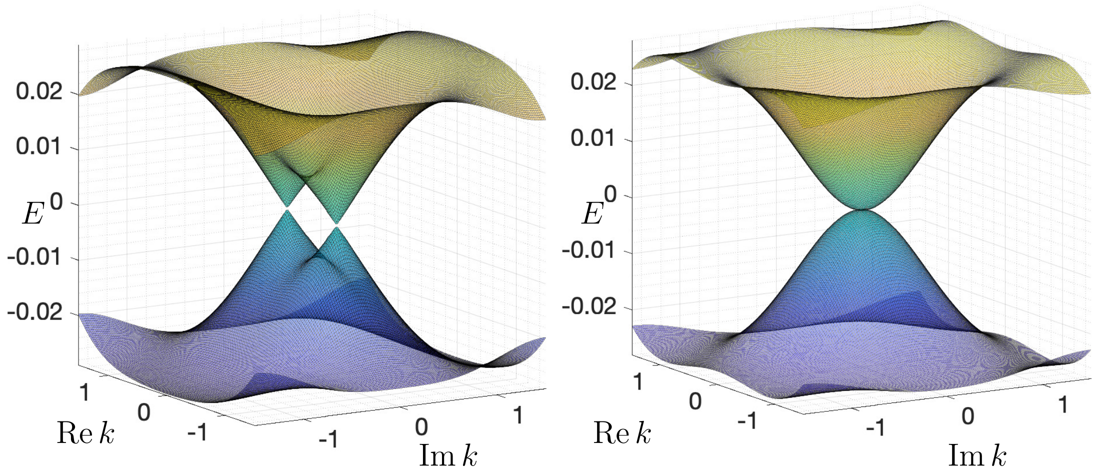


FIGURE 8. When B is real in (9.1) the two Dirac cones approach Γ point as $\alpha \to \alpha^* = \underline{\alpha} + \mathcal{O}(B^3)$ ($\underline{\alpha}$ a simple real magic α) on the line $\operatorname{Im} k = 0$ (left). For $\alpha = \alpha^*$, the quasi-momentum k at which the bifurcation happens are the boundary of the Brillouin zone and the Γ-point which is shown in the figure (right). The animation `https://math.berkeley.edu/~zworski/Rectangle_1.mp4` shows the motion of Dirac points in this case.

The study of the effective Hamiltonian $E_{-+}^B(k)$ (a scalar function in our case) and some additional arguments give the following result (see [BeZw23a, §2] for more detailed statements):

Theorem 16. *Suppose $\underline{\alpha} \in \mathcal{A}$ is simple and $g_0(\underline{\alpha}) \neq 0$ where g_0 is defined in (9.2). Then there exists $\delta_0 > 0$ such that for $0 < |B| < \delta_0$ and $|\alpha - \underline{\alpha}| < \delta_0$, the spectrum of $D_B(\alpha)$ on L_0^2 is discrete (that is the set of Dirac points) and*

$$|\operatorname{Spec}_{L_0^2}(D_B(\alpha)) \cap \mathbb{C}/\Gamma^*| = 2, \tag{9.4}$$

where the elements of the spectrum are included according to their (algebraic) multiplicity. In addition, for a fixed constant $a_0 > 0$ and for every ε there exists δ such that for $0 < |B| < \delta$, $|\alpha - \underline{\alpha}| < a_0 \delta |B|$,

$$\operatorname{Spec}_{L_0^2}(D_B(\alpha)) \subset \Lambda^* + D(0, \varepsilon), \tag{9.5}$$

where we recall that elements of Λ^*, in particular 0, correspond to the Γ point.

A more detailed description would be very desirable. Among things which were left open in [BeZw23a] is the behaviour near K points when $3\theta \in \mathbb{N}$ – see [BeZw23a, Figure 5]. We state one, somewhat vaguely formulated, problem:

Open Problem 18. *Is there a dynamical system which fully explains Figure 7? Basic symmetries of Dirac points are described in* [BeZw23a, (2.10)] *but the clean structure may be due to the special BM potential* (2.5)*. It becomes more complicated for other potentials – see* [BeZw23a, Figure 1].

The quantitative behaviour of Dirac points seems to remain similar for BMH and clarifying that would also be nice. The agreement is particularly striking for $3\theta \in \mathbb{N}$ – see `https://math.berkeley.edu/~zworski/Dirac_BMH.mp4` ($\alpha_0 = \lambda$, $\alpha_1 = \alpha$) where a comparison of the movement of Dirac points for chiral, weakly interacting, and BMH ($\lambda = 0.7\alpha$) is animated. It is harder to catch Dirac points when $\lambda \neq 0$ as we do not have a simple characterization as spectrum of $D_B(\alpha)$ on L^2_0. Hence the neighbourhoods of the Dirac points are shown.

10. Small angle limit as a semiclassical limit

The small angle limit corresponds to letting $\alpha \to \infty$. In that case it is natural to write $\alpha = \lambda/h$, $h \in (0,1]$, $\lambda \in K \Subset \mathbb{C} \setminus 0$ and consider asymptotic behaviour as $h \to 0$. When considering real and positive alpha we can simply take $\lambda = 1$.

The operator $D(\alpha)$ in (2.2) then becomes (up to an irrelevant factor of h^{-2})

$$P(x, hD) := \begin{pmatrix} 2hD_{\bar z} & \lambda U(z) \\ \lambda U(-z) & 2hD_{\bar z} \end{pmatrix}, \quad D_{\bar z} = (1/2i)(\partial_{x_1} + i\partial_{x_2}), \tag{10.1}$$

which is a semiclassical differential system in the sense of [DyZw19, Appendix E.1.1]. Its matrix valued symbol is given by

$$p(x, \xi) = \begin{pmatrix} 2\bar\zeta & U(z) \\ U(-z) & 2\bar\zeta \end{pmatrix}, \quad z = x_1 + ix_2, \quad \zeta = \tfrac{1}{2}(\xi_1 - i\xi_2). \tag{10.2}$$

Theorem 2 shows that (with $H^1_0 = H^1_{\rm loc} \cap L^2_0$ defined in (3.5))

$$\begin{aligned} h\lambda \in \mathcal{A} \quad &\Longleftrightarrow \quad \operatorname{Spec}_{L^2_0} P(x, hD) = \mathbb{C} \\ &\Longleftrightarrow \quad \exists\, u \in H^1_0, \ u \neq 0, \quad P(x, hD)u = 0. \end{aligned} \tag{10.3}$$

We note that $E_\ell(\lambda/h, k)^2$, defined in (3.13) (essentially the bands of $H(\alpha)$), are the eigenvalues of the self-adjoint operator

$$P_2(x, hD, hk) := (P(x, hD) + hk)^*(P(x, hD) + hk). \tag{10.4}$$

Since we only need to consider k in a fundamental domain of Λ^*, hk is a lower order terms when $h \to 0$.

In §10.1 we will see one reason for the difficulty of finding λ's with exactly 3Λ–periodic solutions to $P(x, hD)u = 0$ (or $u \in L^2_0$) when h is small, that is, the difficulty of using (10.3) to characterize magic $\alpha = \lambda/h$.

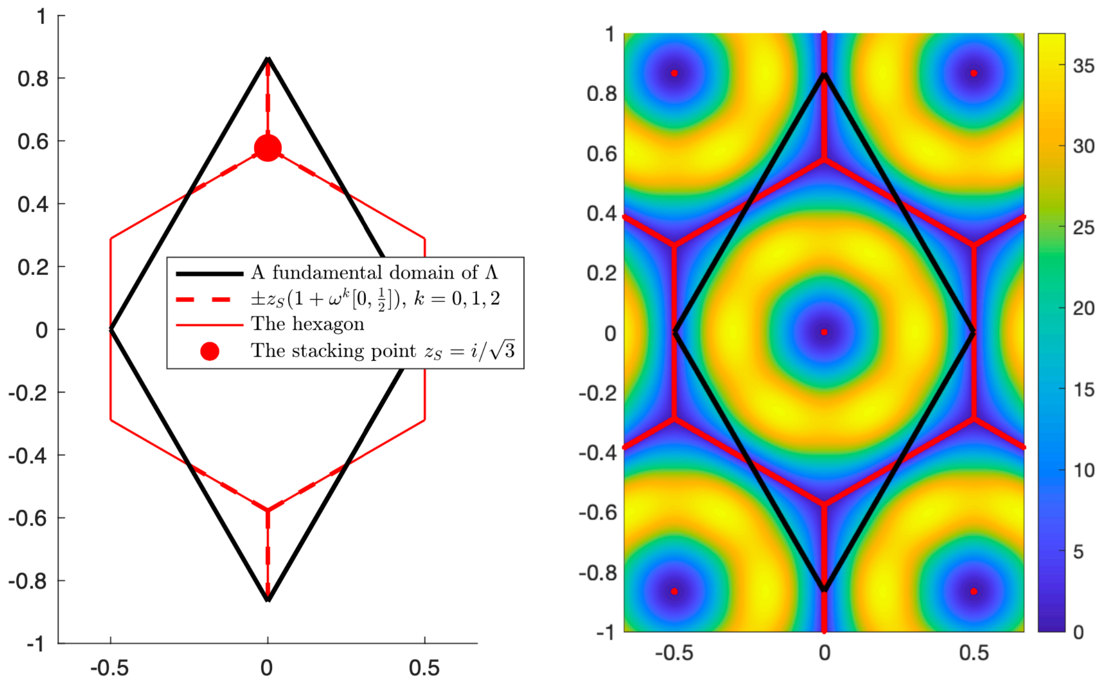


FIGURE 9. Left: the vertices of the hexagon in a fundamental domain of Λ are given by the *stacking points* $\pm z_S$, $z_S = i/\sqrt{3}$ (we use the coordinates of §2.1). They are non-zero points of high symmetry in the sense that $\pm\omega z_S \equiv \pm z_S \mod \Lambda$. Right: the contour plot of $|\{q,\bar{q}\}_{q^{-1}(0)}|$ for q given by the determinant of the semiclassical symbol of $D(\alpha)$ (see (10.7)), $\alpha = 1/h$; the set where $\{q,\bar{q}\}_{q^{-1}(0)} = 0$ is in red. We should stress that the structure of that set becomes more complicated for other potentials U satisfying the required symmetries – see [Be*22, Figure 6].

Instead of (10.3) one could attempt to analyse semiclassically the spectral characterization of Theorem 5: for $k \notin \mathcal{K}$ (see (3.7); we could take $k=0$),

$$\lambda h \in \mathcal{A} \iff \lambda^{-1} \in \mathrm{Spec}_{L^2_0}\left((2hD_{\bar{z}} - hk)^{-1}W(z)\right), \quad W(z) := \begin{pmatrix} 0 & U(z) \\ U(-z) & 0 \end{pmatrix},$$

which of course seems like a tautology. The problem here lies in the fact that $(2hD_{\bar{z}} - hk)^{-1}$, with the Schwartz kernel explicitely given in (6.11), is essentially independent of h and is *not* a semiclassical pseudodifferential operator: hk is a lower order term and the symbol of $2hD_{\bar{z}}$, $2\bar{\zeta}$ has all of $\mathbb{C}$ as its range.

We finally remark that Open Problem 1 (and also 9) is semiclassical in nature: it states a quantization rule

$$\lambda_{n+1} - \lambda_n = \gamma h + \mathcal{O}(h^2), \quad \gamma \simeq \tfrac{1}{2}.$$

10.1. **Exponential squeezing of bands.** In [Be*22] we observed that the results on the existence of localized quasimodes for non-normal semiclassical differential operators with analytic coefficients implies existence of many exponentially small (as $\alpha \to \infty$) Bloch eigenvalues for the chiral model. That means that as α gets large it is hard to distinguish an exactly flat band from many bands that seem flat. Since the phenomenon is semiclassical we use the notation of this section:

Theorem 17 ([Be*22])**.** *Suppose that U is given by (2.3) and $E_\ell(1/h,k)$ are defined in (3.13). Then, there exist constants $c_0, c_1, c_2 > 0$ such that*

$$|E_\ell(1/h,k)| \le c_1 e^{-c_0/h}, \quad |\ell| \le c_2/h. \tag{10.5}$$

The proof is based on a result of Dencker–Sjöstrand–Zworski [DSZ04, Theorem 1.2′] (see also [HiSj18, §II.2.8]) which in turn was based on works of Hörmander and of Kashiwara, Kawai, and Sato. Roughly, it states the following fact: suppose that $Q(x,hD,h)$ is a (scalar) semiclassical differential operator with analytic coefficients and $q(x,\xi)$ is its principal symbol. Then

$$q(x_0,\xi_0) = 0, \ \{\operatorname{Re} q, \operatorname{Im} q\}(x_0,\xi_0) < 0 \Longrightarrow \begin{cases} \exists \ u(h) \in C^\infty, \|u(h)\|_{L^2} = 1, \\ \|Q(x,hD,h)u(h)\|_{L^2} \le Ce^{-C/h}, \\ u(h) \text{ is } \textit{microlocalized} \text{ at } (x_0,\xi_0), \end{cases}$$

see [DSZ04] and references given there. Here $\{a,b\}$ denotes the *Poisson bracket* which in our $2D$ case and using the notation z and ζ in (10.2) is given by

$$\{a,b\} = \partial_\zeta a \, \partial_z b - \partial_\zeta b \, \partial_z a + \partial_{\bar\zeta} a \, \partial_{\bar z} b - \partial_{\bar\zeta} b \, \partial_{\bar z} a,$$

see [Zw12, §2.4] for an introduction to its geometric significance.

The type of microlocalization we have for $u(h)$, implies in particular that that $|u(h,x)| \le e^{-|x-x_0|^2/Ch}$, which means that $u(h)$ "lives" in $B(x_0, h^{\frac12 - \varepsilon})$, for any $\varepsilon > 0$. From such local approximate solutions we can built many approximate solutions with any periodicity properties. (A model to keep in mind is the annihilation operator $Q(x,hD) = hD_{x_1} - ix_1$ with $(x_0,\xi_0) = (0,0) \in \mathbb{R}^2 \times \mathbb{R}^2$; we can then take $u(h,x) = c(h)e^{-x_1^2/h - x_2^2/h}$.)

At points z_0 with $U(z_0) \ne 0$ an easy reduction (see [Be*22, Proof of Proposition 4.1]) shows that to construct $u(h) \in C^\infty(\mathbb{C};\mathbb{C}^2)$ localized at z_0 and satisfying

$$\|(P(x,hD,h) + hk)u(h)\| \le Ce^{-1/Ch}\|u(h)\|, \tag{10.6}$$

it is enough to find $v(h) \in C^\infty(\mathbb{C};\mathbb{C})$, localized to z_0 such that $\|Q(x,hD_x,h)v(h)\|_{L^2} \leq Ce^{-1/Ch}\|v(h)\|_{L^2}$ where Q is a *scalar* operator with the principal symbol given by the determinant of p in (10.2):

$$q(x,\xi) = (2\bar{\zeta})^2 - \lambda^2 U(z)U(-z), \quad z = x_1 + ix_2, \ \ \zeta = \tfrac{1}{2}(\xi_1 - i\xi_2). \tag{10.7}$$

Hence, in view of the discussion above, we need to look for (x_0,ξ_0) such that $q(x_0,\xi_0) = 0$ and $\{\operatorname{Re} q, \operatorname{Im} q\}(x_0,\xi_0) < 0$. Such points are indeed plentiful – see Figure 9 for the case of $\lambda = 1$ and [Be*22, §4] for more examples.

Once we have (10.6) we obtain an en exponentially accurate approximate solution to $P_2(x,hD,hk)u(h) = 0$ where P_2 was defined in (10.4). Self-adjointness of P_2 then implies existence of exponentially small eigenvalues. Using many localized approximate solutions we can bound their number from below by $1/h$, see [Be*22, §4].

Open Problem 19. *Relate the geometry of level sets of* $z \mapsto |\{q,\bar{q}\}|_{q^{-1}(z)=0}|$ *(see Figure 9) to the concentration of mass of the protected states* $u_K(\lambda/h)$ *(see Theorem 1) as* λ *varies in a compact set and* $h \to 0$. *For an animated example see* https://math.berkeley.edu/~zworski/bracket_dynamics.mp4 *where* $h = 1/8$ *and* λ *varies on a circle of radius* 1. *This problem is related to the issues discussed in §10.2 below.*

10.2. **Classically forbidden regions.** The contour plot of $z \mapsto \log|u_K(\alpha,z)|$ as α changes (and U is given in (2.5)) as well as the link in Open Problem 19, suggest that solutions to $(D(\alpha)+k)u = 0$, $u \in H^1_0$ (nontrivial only for $\alpha \in \mathcal{A}$ if $k \neq \pm K$) *decay exponentially* in α near the hexagon spanned by the stacking points (see Figure 9) and near the center of the hexagon. From the semiclassical point of view presented in this section this means decay $e^{-c/h}$ which typically corresponds to classically forbidden regions.

The standard notion of classically forbidden regions is based on ellipticity: if Q is a principally scalar semiclassical differential operator, elliptic in the classical sense (that is, for fixed h), with analytic coefficients and a scalar principal symbol $q(x,\xi)$ then (with neigh(x_0) denoting *some* neighbourhood of x_0)

$$q|_{\pi^{-1}(x_0)} \neq 0, \ \ Qu = 0 \text{ in neigh}(x_0), \ \|u\|_{L^2} = 1 \implies \|u\|_{L^2(\text{neigh}(x_0))} \leq Ce^{-c/h}, \tag{10.8}$$

see [Ma02, Theorem 4.1.5], [HiZw23, Proposition 6.4]. (A typical example is given by $Q = -h^2\Delta + V(x)$ where $V \in C^\infty$ is real valued – there is no need for analyticity in that case – see [Zw12, Theorem 7.3]; in that case the condition is simply that $V(x_0) > 0$ as then for all ξ, $q(x_0,\xi) = \xi^2 + V(x_0) > 0$.)

In the case of the operator $P(x,hD)$ given in (10.1) there are no classically forbidden regions: for every $x \in \mathbb{R}^2$ there exists $\xi \in \mathbb{R}^2$ at which the determinant of the principal symbol (see (10.7)) vanishes, $q(x,\xi) = 0$.

The remedy for this is to use analogues of results on (analytic) hypoellipticity due to Trépreau (with different proofs, following an approach due to Sjöstrand and reviewed

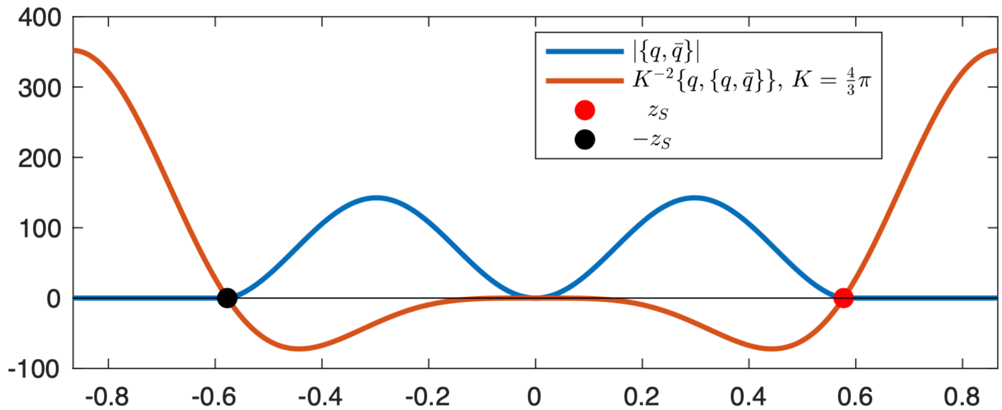


FIGURE 10. Plots of $|\{q,\bar q\}|$ and of (rescaled) $\{q,\{q,\bar q\}\}$ above the intersection of the imaginary axis and the fundamental domain in Figure 9. The edges of the hexagon emanate right of z_S and left of $-z_S$.

in [HiSj18], provided by Himonas), which followed ideas of Egorov, Hörmander, and Kashiwara (we defer to [HiZw23] for pointers to the literature). Hypoellipticity here refers to having the same conclusion $\|u\|_{L^2(\mathrm{neigh}(x_0))} \leq Ce^{-c/h}$ as in (10.8) but without the assumption that $q|_{q^{-1}(x_0)} \neq 0$.

A semiclassical version of a general hypoelliptic result we need is given as follows: let Q satisfy the same general assumptions as before (10.8);

$$\left.\begin{array}{l} \{q,\bar q\}|_{\pi^{-1}(x_0)\cap q^{-1}(0)} = 0, \\ \{q,\{q,\bar q\}\}|_{\pi^{-1}(x_0)\cap q^{-1}(0)} \neq 0, \\ Qu = 0 \text{ in } \mathrm{neigh}(x_0), \ \|u\|_{L^2} = 1 \end{array}\right\} \Longrightarrow \ \|u\|_{L^2(\mathrm{neigh}(x_0))} \leq Ce^{-c/h}, \tag{10.9}$$

see [HiZw23, Theorem 2].

To see why such a result could be true consider a simple one dimensional example: $q(x,\xi) = \xi + ix^2$, $(x,\xi) \in \mathbb{R}\times\mathbb{R}$, $x_0 = 0$. Then $\{q,\bar q\}(x_0,\xi) = -4ix_0 = 0$, $\{q,\{q,\bar q\}\}(x_0,\xi) = -4i$, so the condition holds. If $0 = q(x,hD)u = (h/i)(\partial_x - x^2/h)u$, then $u(x,h) = u(0,h)e^{\frac13 x^3/h}$. For this to be uniformly bounded near 0, we need $u(0,h) = e^{-c/h}$, $c > 0$. So $|u(x,h)| \leq e^{-c/2h}$ for $|x|$ small. We remark that similar bracket conditions in the semiclassical setting appeared recently in the work of Sjöstrand–Vogel [SjVo23] who provided fine tunneling estimates for a model operator. Any extension of their results to more general operators should have consequences in our setting as well.

As in (5.16) we can reduce the problem of looking at solutions to $h(D(\alpha)+k) = P(x,hD) + hk$ to a principally scalar problem, with the principal symbol given by $q(x,\xi)$ in (10.7). It then turns out that the condition in (10.9) holds at any x_0 on an *open* edge of the hexagon spanned by the stacking points – see Figure 10 for the case of $\lambda = 1$ and U given in (2.5). Remarkably, due to the special properties of the BM potential, the sign properties can be established analytically – see [HiZw23, §3].

At $\pm z_S$ the condition in (10.9) does not hold. However, $\pi^{-1}(\pm z_S) \cap q^{-1}(0) = \{(\pm z_S, 0)\}$ and

$$\{q, \bar{q}\}(\pm z_S, 0) = 0, \quad \{q, \{q, \{q, \{q, \bar{q}\}\}\}\}(\pm z_S, 0) \neq 0. \tag{10.10}$$

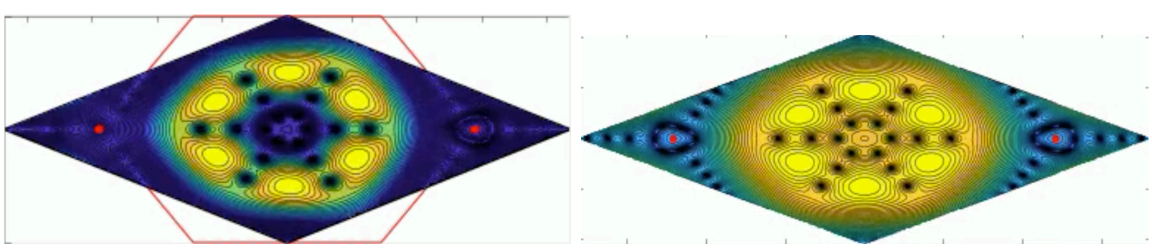
General hypoellipticity results of Trépreau do not apply to this case but a detailed analysis of our specific principal symbol [TaZw23b] allows an application of the same strategy as in the proof of (10.9) to obtain exponential decay near the stacking points.

Since the conditions in (10.9) and (10.10) are classical in the sense of involving the symbol (that is, the "classical observable", $q(x, \xi)$) and Poisson brackets (objects underlying classical dynamics), we obtain the following result about classically forbidden regions

Theorem 18 ([HiZw23],[TaZw23b])**.** *There exists a fixed open neighbourhood, Ω, of the hexagon spanned by the stacking points (see Figure 9) and $c > 0$ such that if $u(h) \in H^1_0$ satisfies $(P(x, hD) + hk)u = 0$ and $\|u(h)\|_{L^2_0} = 1$, then*

$$\|u(h)\|_{L^2(\Omega)} \leq c^{-1} e^{-c/h}. \tag{10.11}$$

The situation is more complicated at the center of the hexagon, $z_0 = 0$. In that case, the operator is not of principal type, that is, $q(0, 0) = 0$ ($\pi^{-1}(0) \cap q^{-1}(0) = \{(0, 0)\}$) and $dq(0, 0) = 0$. This means that lower order terms should matter. That is confirmed by comparing (5.16) with the scalar model $Q(\alpha)$ (with no lower order terms). For $Q(\alpha)$, unlike for the chiral model, we do not see exponential decay near 0 (the decay near the hexagon based on the properties of pricipal symbol q persists) : on the left $\log |u|$ for u a protected state for $D(\alpha)$ and on the right same for $Q(\alpha)$:



Open Problem 20. *Show that there exist a fixed neighbourhood Ω of 0 (see Figure 9) and $c > 0$ such that if $(P(x, hD) + hk)u = 0$, where P is given in (10.1), and $\|u(h)\|_{L^2_0} = 1$, then $\|u(h)\|_{L^2(\Omega)} \leq c^{-1} e^{-c/h}$.*

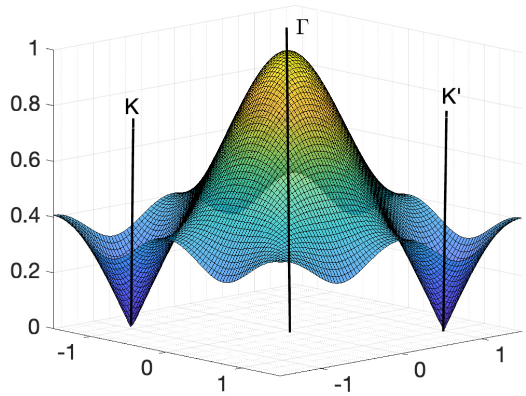


FIGURE 11. Two Dirac cones at K and K' points.

APPENDIX BY MENGXUAN YANG AND ZHONGKAI TAO

We prove the existence of conic singularities in the first band of the chiral limit [TKV19] of the Bistritzer–MacDonald Hamiltonian [BiMa11] of twisted bilayer graphene when $\alpha \notin \mathcal{A}$.

Theorem 19. *Near $\pm K$ points, the first band $E_1(\alpha, k)$ is given by*

$$E_1(\alpha, k) = c(\alpha) \cdot |k \pm K| + \mathcal{O}(|k \pm K|^2), \tag{A.1}$$

where $c(\alpha) \geq 0$ with the equality holds if and only if $\alpha \in \mathcal{A}$.

A key fact used in the proof is the *existence of protected eigenstates* [TKV19, Be*22] described in Theorem 1. We also remark that a dual result is the existence of protected states for the operator $D(\alpha)^*$: there exists $v_{\pm K}(\alpha) \in H_0^1(\mathbb{C}; \mathbb{C}^2)$ such that $\tau(K)v_K(0) = (1, 0)^T$, $\tau(-K)v_{-K}(0) = (0, 1)^T$,

$$v_{\pm K}(\alpha) \in \ker_{L_0^2(\mathbb{C};\mathbb{C}^2)}(D(\alpha)^* \pm \bar{K}). \tag{A.2}$$

It also follows from the proof that the generalized eigenspace also has dimension 1, i.e., the spectrum is simple.

Now we prove Theorem 19 by setting up a Grushin problem to compute the first band $E_1(\alpha, k)$ near $k = \pm K$ for $\alpha \notin \mathcal{A}$. We refer to [DyZw19, Appendix C] for a presentation of this method. The proof of Theorem 19 is based on the following general fact: suppose that $X_1 \subset X_2$ are two Banach spaces and $P : X_1 \to X_2$ be a Fredholm operator of index 0 such that

$$\ker P = \operatorname{span}\{\varphi\}, \quad \ker P^* = \operatorname{span}\{\varphi_*\}.$$

Then there is a dichotomy:

$$\begin{gathered} P-z \text{ is invertible in a punctured neighbourhood of } z=0; \\ \text{if moreover the eigenvalue } z=0 \text{ is simple, then } \langle\varphi,\varphi_*\rangle\neq 0 \\ \text{or} \\ P-z \text{ is not invertible for all } z, \text{ and } \langle\varphi,\varphi_*\rangle=0. \end{gathered} \tag{A.3}$$

Proof of (A.3). The first part of the dichotomy follows from the analytic Fredholm theory (see [DyZw19, Theorem C.8]), which says if $P-z$ is invertible at one point then $(P-z)^{-1}$ is a meromorphic family.

Now suppose $P-z$ is invertible in a neighbourhood of $z=0$ and 0 is a simple eigenvalue, then $(P-z)^{-1}$ has the following expansion near $z=0$:

$$(z-P)^{-1}=A_0(z)+\frac{\Pi}{z}$$

where $A_0(z)$ is holomorphic and Π is a rank one projector. From the expansion we see $P\Pi=\Pi P=0$. So

$$\operatorname{im}\Pi\subset\ker P=\operatorname{span}\{\varphi\},\quad \operatorname{im}P\subset\ker\Pi.$$

Thus Π is of the form $\Pi(y)=\langle y,v_*\rangle\varphi$ for some $v_*\in X_2^*$. Moreover, $\langle Px,v_*\rangle=0$ for any $x\in X_1$, which implies $P^*v_*=0$. Thus $v_*=c\varphi_*$ for some $c\in\mathbb{C}\setminus\{0\}$. Since $\Pi^2=\Pi$, we conclude $\langle\varphi,\varphi_*\rangle\neq 0$.

Suppose $P-z$ is not invertible for any z, then we consider the following Grushin problem:

$$\begin{pmatrix} P-z & R_- \\ R_+ & 0 \end{pmatrix} : X_1\times\mathbb{C}\to X_2\times\mathbb{C}$$

where $\varphi_*(R_-1)=1$ and $R_+\varphi=1$. One can compute from [DyZw19, Proposition C.3] that $E_{-+}(z)=z\langle\varphi,\varphi_*\rangle+\mathcal{O}(|z|^2)$. By assumption $E_{-+}(z)=0$, so we conclude $\langle\varphi,\varphi_*\rangle=0$. ☐

We can now give

Proof of Theorem 19. For the chiral Hamiltonian

$$H_k(\alpha):H_0^1(\mathbb{C};\mathbb{C}^4)\to L_0^2(\mathbb{C};\mathbb{C}^4),\quad \alpha\in\mathbb{C}, \tag{A.4}$$

we consider the existence of a Dirac cone at K point, as the point $-K$ is similar. By the existence of protected states, there exist two normalized protected states $\varphi(\alpha;z),\psi(\alpha;z)\in\ker_{L_0^2(\mathbb{C};\mathbb{C}^4)}H_K(\alpha)$ such that

$$\varphi(\alpha;z)=(u_K(\alpha),0_{\mathbb{C}^2})^T,\quad \psi(\alpha;z)=(0_{\mathbb{C}^2},v_K(\alpha))^T. \tag{A.5}$$

We consider the Grushin problem for the operator $H_k(\alpha) - z$ near $k = K$:

$$\mathcal{H}_k = \begin{pmatrix} H_k(\alpha) - z & R_- \\ R_+ & 0 \end{pmatrix} : H_0^1(\mathbb{C}; \mathbb{C}^4) \oplus \mathbb{C}^2 \longrightarrow L_0^2(\mathbb{C}; \mathbb{C}^4) \oplus \mathbb{C}^2 \tag{A.6}$$

with

$$R_- : (u_-^{(1)}, u_-^{(2)})^T \mapsto u_-^{(1)}\varphi + u_-^{(2)}\psi, \;\; R_+ : u \mapsto (\langle u, \varphi\rangle, \langle u, \psi\rangle)^T.$$

For $k = K$, the Grushin problem (A.6) is invertible with the inverse given by

$$\mathcal{E} = \begin{pmatrix} E & E_+ \\ E_- & E_{-+} \end{pmatrix} : L_0^2(\mathbb{C}; \mathbb{C}^4) \oplus \mathbb{C}^2 \longrightarrow H_0^1(\mathbb{C}; \mathbb{C}^4) \oplus \mathbb{C}^2 \tag{A.7}$$

with

$$Ev = \sum_{j \neq \pm 1} \frac{1}{E_j - z} \langle v, \varphi_j\rangle \varphi_j, \quad E_+ v_+ = R_- v_+, \quad E_- v = R_+ v, \quad E_{-+} = \begin{pmatrix} z & \\ & z \end{pmatrix},$$

where $\{\varphi_j\}$ is an orthonormal basis of eigenfunctions of $H_K(\alpha)$ with eigenvalue E_j such that $\varphi_1 = \varphi$ and $\varphi_{-1} = \psi$. By [DyZw19, Proposition C.3], the perturbed Grushin problem (A.6) is well-posed for $|k - K|$ sufficiently small and the eigenvalues of $H_k(\alpha)$ are given by zeros of the determinant of

$$F_{-+} = E_{-+} + \sum_{k=1}^{\infty} (-1)^k E_- A (EA)^{k-1} E_+, \quad A = \begin{pmatrix} & \overline{k - K} \\ k - K & \end{pmatrix}. \tag{A.8}$$

In particular, the leading order term is given by

$$E_- A E_+ = \begin{pmatrix} & \overline{(k - K)} \langle v_K, u_K\rangle \\ (k - K)\langle u_K, v_K\rangle & \end{pmatrix} \tag{A.9}$$

This yields that $E_{\pm 1}(\alpha, k) = \pm |\langle v_K, u_K\rangle| \cdot |k - K| + \mathcal{O}(|k - K|^2)$ near $k = 0$, where $\langle v_K, u_K\rangle = 0$ if and only if $\alpha \in \mathcal{A}$ by (A.3). $\square$

References

[AvJi09] A. Avila and S. Jitomirskaya, *The ten martini problem*, Ann. Math. **170**(2009), 303–342.

[Be*21] S. Becker, M. Embree, J. Wittsten and M. Zworski, *Spectral characterization of magic angles in twisted bilayer graphene,* Phys. Rev. B **103**, 165113, 2021.

[Be*22] S. Becker, M. Embree, J. Wittsten and M. Zworski, *Mathematics of magic angles in a model of twisted bilayer graphene*, Probab. Math. Phys. **3**(2022), 69–103.

[Be*23] S. Becker, T. Humbert, J. Wittsten and M. Yang, *Chiral limit of twisted trilayer graphene*, arXiv:2308.10859.

[BHZ22a] S. Becker, T. Humbert and M. Zworski, *Integrability in the chiral model of magic angles*, Comm. Math. Phys. **403**(2023), 1153–1169.

[BHZ22b] S. Becker, T. Humbert and M. Zworski, *Fine structure of flat bands in a chiral model of magic angles,* arXiv:2208.01628.

[BHZ23] S. Becker, T. Humbert, and M. Zworski, *Degenerate flat bands in twisted bilayer graphene,* arXiv:2306.02909.

[BOV23] S. Becker, I. Oltman, and M. Vogel, *Magic angle (in)stability and mobility edges in disordered Chern insulators,* arXiv:2309:02701.

[BeZw23a] S. Becker and M. Zworski, *Dirac points for twisted bilayer graphene with in-plane magnetic field*, arXiv:2303.00743.

[BeZw23b] S. Becker and M. Zworski, *From the chiral model of TBG to the Bistritzer–MacDonald model*, arXiv:2308.11555.

[BiMa11] R. Bistritzer and A. MacDonald, *Moiré bands in twisted double-layer graphene.* PNAS, **108**, 12233–12237, 2011.

[CGG22] E. Cancès, L. Garrigue, D. Gontier, *A simple derivation of moiré-scale continuous models for twisted bilayer graphene.* Phys. Rev. B **107**(2023), 155403.

[Ca*18] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, P. Jarillo-Herrero, *Unconventional superconductivity in magic-angle graphene superlattices.* Nature 556, 43-50, (2018).

[Ch99] T. Christiansen, *Some lower bounds on the number of resonances in Euclidean scattering,* Math. Res. Letters **6**(1999), 203–211.

[DSZ04] N. Dencker, J. Sjöstrand and M. Zworski, *Pseudospectra of semiclassical differential operators*, Comm. Pure Appl. Math. **57**(2004), 384—415.

[DuNo80] B.A. Dubrovin and S.P. Novikov, *Ground states in a periodic field. Magnetic Bloch functions and vector bundles.* Soviet Math. Dokl. **22**, 1, 240–244, 1980.

[DyZw19] S. Dyatlov and M. Zworski, *Mathematical Theory of Scattering Resonances,* AMS 2019, `http://math.mit.edu/~dyatlov/res/`

[GaZw23] J. Galkowski and M. Zworski, *An abstract formulation of the flat band condition*, arXiv:2307.04896.

[HiSj18] M. Hitrik and J. Sjöstrand, *Two minicourses on analytic microlocal analysis,* in "Algebraic and Analytic Microlocal Analysis", M. Hitrik, D. Tamarkin, B. Tsygan, and S. Zelditch, eds. Springer, 2018, arXiv:1508.00649.

[HiZw23] M. Hitrik and M. Zworski, *Classically forbidden regions in the chiral model of twisted bilayer graphene,* arXiv:2310.19140.

[H*13] B. Hunt, J.D. Sanchez-Yamagishi, A.F. Young, M. Yankowitz, B.J. LeRoy, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero, and R.C. Ashoori *Massive Dirac Fermions and Hofstadter Butterfly in a van der Waals Heterostructure*, Science **340**(2013), 1427–1430.

[KhZa15] S. Kharchev and A. Zabrodin, *Theta vocabulary I.* J. Geom. Phys. **94**(2015), 19–31.

[KPS20] Y.H. Kwan, S.A. Parameswaran, and S.L. Sondhi, *Twisted bilayer graphene in a parallel magnetic field* Phys. Rev. B **101**(2020), 205116.

[Le*20] P.J. Ledwith, G. Tarnopolsky, E. Khalaf, and A. Vishwanath, *Fractional Chern insulator states in twisted bilayer graphene: An analytical approach,* Phys. Rev. Research 2, 023237 (2020).

[LVK22] P. Ledwith, A. Vishwanath, and E. Khalaf. *Family of ideal Chern flatbands with arbitrary Chern number in chiral twisted graphene multilayers.* Phys. Rev. Lett. **128**(2022), 176404.

[Ma02] A. Martinez, *An Introduction to Semiclassical and Microlocal Analysis,* Springer, 2002.

[Mu83] D. Mumford, *Tata Lectures on Theta. I.* Progress in Mathematics, **28**, Birkhäuser, Boston, 1983.

[QiMa21] W. Qin and A.H. MacDonald, *In-Plane critical magnetic fields in magic-angle twisted trilayer graphene*, Phys. Rev. Lett. **127**(2021), 097001.

[Se86] R. Seeley, *A simple example of spectral pathology for differential operators,* Comm. PDE, **11**(1986), 595–598.

[SjVo23] J. Sjöstrand and M. Vogel, *Tunneling for the $\bar\partial$ operator*, arXiv:2303.06096.

[TaZw23a] Z. Tao and M. Zworski, *PDE methods in condensed matter physics,* Lecture Notes, 2023, https://math.berkeley.edu/~zworski/Notes_279.pdf.

[TaZw23b] Z. Tao and M. Zworski, Appendix to [HiZw23], arXiv:2310.19140.

[TKV19] G. Tarnopolsky, A.J. Kruchkov and A. Vishwanath, *Origin of magic angles in twisted bilayer graphene*, Phys. Rev. Lett. 122, 106405, 2019

[WaLi22] J. Wang and Z. Liu, *Hierarchy of ideal flatbands in chiral twisted multilayer graphene models.* Phys. Rev. Lett. **128**(2022), 176403.

[Wa*22] A. B. Watson, T. Kong, A. H. MacDonald, and M. Luskin *Bistritzer-MacDonald dynamics in twisted bilayer graphene*, J. Math. Phys. **64**(2023), 031502.

[WaLu21] A. Watson and M. Luskin, *Existence of the first magic angle for the chiral model of bilayer graphene*, J. Math. Phys. **62**, 091502 (2021).

[Ya23] M. Yang, *Flat bands and high Chern numbers in twisted multilayer graphene,* arXiv:2303.00103.

[Zw12] M. Zworski, *Semiclassical analysis,* Graduate Studies in Mathematics **138**, AMS, 2012.

Email address: zworski@math.berkeley.edu

Department of Mathematics, University of California, Berkeley, CA 94720, USA

Department of Mathematics, University of California, Berkeley, CA 94720, USA.

Email address: tzk320581@berkeley.edu

Department of Mathematics, University of California, Berkeley, CA 94720, USA.

Email address: mxyang@math.berkeley.edu