Complete asymptotic expansion of the integrated density of states of multidimensional almost-periodic pseudo-differential operators

Sergey Morozov, Leonid Parnovski and Roman Shterenberg

Abstract. We obtain a complete asymptotic expansion of the integrated density of states of operators of the form $H = (-\Delta)^w + B$ in $\mathbb{R}^d$. Here $w > 0$, and B belongs to a wide class of almost-periodic self-adjoint pseudo-differential operators of order less than $2w$. In particular, we obtain such an expansion for magnetic Schrödinger operators with either smooth periodic or generic almost-periodic coefficients.

Mathematics Subject Classification (2010). Primary 35P20, 47G30, 47A55; Secondary 81Q10.

Keywords. Periodic operators, almost-periodic pseudo-differential operators, integrated density of states.

1. Introduction

In [12], two of the authors of this paper have considered the following problem. Let

$$H = -\Delta + V \tag{1.1}$$

be a Schrödinger operator acting in $L_2(\mathbb{R}^d)$. The potential $V = V(\mathbf{x})$ is assumed to be real, smooth, and either periodic, or almost-periodic; in the almost-periodic case we assume that all the derivatives of V are almost-periodic as well. We are interested in the asymptotic behaviour of the (integrated) density of states $N(\lambda)$ as the spectral parameter λ tends to infinity.

SM and LP were partially supported by the EPSRC grant EP/ F029721/1. SM was also supported by the Lundbeck Foundation and the European Research Council under the European Community's Seventh Framework Program (FP7/2007–2013)/ERC grant agreement 202859. RS was partially supported by the NSF grant DMS-0901015. The authors would like to thank Gerassimos Barbatis for participation in preliminary discussions which led to this paper. SM would like to express his thanks for hospitality to the University of Athens and ESI Vienna, where part of this work was made.

The density of states of H can be defined by the formula

$$N(\lambda) = N(\lambda; H) := \lim_{L\to\infty} \frac{N(\lambda; H_D^{(L)})}{(2L)^d}. \tag{1.2}$$

Here, $H_D^{(L)}$ is the restriction of H to the cube $[-L, L]^d$ with the Dirichlet boundary conditions, and $N(\lambda; A)$ is the counting function of the discrete spectrum of A. Later, we will give several equivalent definitions of $N(\lambda)$ which are more convenient to work with. If we denote by $N_0(\lambda)$ the density of states of the unperturbed operator $H_0 = -\Delta$, one can easily see that for positive λ one has

$$N_0(\lambda) = C_d \lambda^{d/2}, \tag{1.3}$$

where

$$C_d = \frac{w_d}{(2\pi)^d} \text{ and } w_d = \frac{\pi^{d/2}}{\Gamma(1 + d/2)} \tag{1.4}$$

is the volume of the unit ball in $\mathbb{R}^d$. There was a long-standing conjecture that the density of states of H enjoys the following asymptotic behaviour as $\lambda \to \infty$:

$$N(\lambda) \sim \lambda^{d/2}\Big(C_d + \sum_{j=1}^{\infty} e_j \lambda^{-j}\Big), \tag{1.5}$$

meaning that for each $K \in \mathbb{N}$ one has

$$N(\lambda) = \lambda^{d/2}\Big(C_d + \sum_{j=1}^{K} e_j \lambda^{-j}\Big) + R_K(\lambda) \tag{1.6}$$

with $R_K(\lambda) = o(\lambda^{\frac{d}{2}-K})$. In those formulas, e_j are real numbers which depend on the potential V. These numbers can be calculated relatively easily using the heat kernel invariants (computed in [2]); they are equal to certain integrals of the potential V and its derivatives. In paper [7], all these coefficients were computed; in particular, it turns out that, if d is even, then e_j vanish whenever $j > d/2$.

Until [12], formula (1.5) was proved only in the case $d = 1$ in [17] for periodic V and in [14] for almost-periodic V. In [11], this formula was proved in the case $d = 2$ and periodic potential. In the periodic case and $d \geqslant 3$, only partial results were known, see [1], [4], [5], [13], [18], [21]. In particular, in [4] it was shown that formula (1.6) is valid with $K = 1$ and $R(\lambda) = O(\lambda^{-\delta})$ with some small positive δ when $d = 3$ and $R(\lambda) = O(\lambda^{\frac{d-3}{2}} \ln \lambda)$ when $d > 3$. In the multidimensional almost-periodic case, formula (1.6) was known only with $K = 0$ and $R(\lambda) = O(\lambda^{\frac{d-2}{2}})$, see [16].

In [12], formula (1.5) was obtained for operators (2.1) assuming that the real-valued potential V is either smooth periodic, or generic quasi-periodic, or belongs to a reasonably wide class of almost-periodic functions (see [12] for a complete set of conditions on V).

In the case of magnetic Schrödinger operator

$$H = (-i\nabla + \mathbf{A})^2 + V,$$

one expects the asymptotic formula (1.5) to be still valid (assuming similar restrictions on the magnetic potential $A = A(\mathbf{x})$ and the electric potential $V = V(\mathbf{x})$). However, until now only the partial asymptotic formula (1.6) with $K = 0$ and $R(\lambda) = O(\lambda^{\frac{d-2+\varepsilon}{2}})$ was known ([8]; see also [6]).

The main aim of the current paper is to obtain the complete asymptotic expansion of the integrated density of states for a more general class of operators than was considered in [12]. This class, in particular, will contain magnetic Schrödinger operators. We give a detailed description of this new class in the next section; here, we list the main properties of the operators belonging to it.

(i) We consider perturbations of the Laplacian, or any positive power of the Laplacian. More precisely, we work with operators of the form

$$H = (-\Delta)^w + B,$$

where B is a differential or pseudo-differential operator of order $\kappa < 2w$. Here H is self-adjoint and belongs to the standard algebra of almost-periodic pseudo-differential operators, see e.g. [15] and [16].

(ii) If B is a differential operator, we assume that its coefficients satisfy the same conditions the potential V had to satisfy in [12] (for example, the coefficients can be smooth periodic, or generic quasi-periodic functions). In particular, periodic magnetic Schrödinger operators are covered by our results.

(iii) If B is pseudo-differential, we assume that it is a classical pseudo-differential operator, or, more generally, the operator of classical type. By the latter we mean that the symbol of B admits an asymptotic decomposition in powers of $|\boldsymbol{\xi}|$ when $|\boldsymbol{\xi}| \to \infty$; however, these powers do not have to be integer. Note that operators with the relativistic kinetic energy $\sqrt{(-i\nabla + \mathbf{A})^2 + m^2}$ are admissible for (almost-)periodic smooth $\mathbf{A}$ and $m \geqslant 0$.

Under these assumptions we prove that the integrated density of states $N(\lambda)$ has the complete asymptotic expansion (2.17). This expansion contains powers of λ and powers of $\ln \lambda$; the values of the exponents in the powers of λ depend on the form of B, whereas logarithms are raised to integer powers smaller than d.

Our method is not efficient for explicitly calculating the coefficients in (2.17). However, as soon as the general form of the expansion is established, the coefficients can sometimes be calculated by, say, comparison with the expansion for the heat kernel (see [7, 12]). This applies, in particular, to the case of magnetic Schrödinger operators, for which it turns out that the logarithmic terms are absent (i.e., the corresponding coefficients are zero).

One immediate and slightly unexpected corollary of (2.17) is as follows:

Corollary 1.1. *Suppose, $H = (-\Delta)^w + B$ with B being periodic and either differential, or pseudo-differential operator of classical type. Then for sufficiently large λ the spectrum of H is purely absolutely continuous.*

Proof. Since H is periodic, the general Floquet-Bloch theory implies that the spectrum of H is absolutely continuous with the possible exception of eigenvalues of infinite multiplicity. If λ is such an eigenvalue, the integrated density of states has a jump at least $|\Gamma^\dagger|$ at λ, where $\Gamma^\dagger$ is the lattice dual to the lattice of periods of H. Due to (2.17), this cannot happen for large λ. $\square$

The approach of our paper is similar to the one of [12]. In particular, we use the method of gauge transform developed in [19], [20], and [13]. Nevertheless, there are plenty of new (mostly technical, but sometimes ideological) difficulties arising because the operator B is no longer bounded and no longer local. One example of the new methods employed in this paper is the proof of Lemma 10.5: not only this proof works for unbounded B, it also makes Condition D from [12] redundant. The biggest increase in technical difficulties comes in Section 10 where we express the contribution to the density of states from various regions in the momentum space as certain complicated integrals and then try to compute these integrals. As a result, our paper is technically more complicated than [12] (which already was quite difficult to read). Thus, we have reluctantly abandoned the idea of making our paper completely self-contained; we will skip all parts of the argument which are identical (or close) to corresponding parts of [12] and refer the reader to that paper. Nevertheless, we will present all the definitions and properties of the important objects.

Remark 1.2. Throughout the article we employ the convention that, if some statement is given without a proof, then an analogous statement can be found in [12], and the proof is the same up to obvious modifications. It comes without saying that the reader is strongly encouraged to read the article [12] first, before attempting to read this paper.

2. Preliminaries

For $w > 0$ we consider the operator

$$H = (-\Delta)^w + B \tag{2.1}$$

acting in $\mathsf{L}_2(\mathbb{R}^d)$. The action of the pseudo-differential operator B on functions from the Schwarz class $\mathsf{S}(\mathbb{R}^d)$ is defined by the formula

$$(Bf)(\mathbf{x}) := (2\pi)^{-d/2} \int b(\mathbf{x}, \boldsymbol{\xi}) e^{i\boldsymbol{\xi}\mathbf{x}} (\mathcal{F}f)(\boldsymbol{\xi}) d\boldsymbol{\xi}.$$

Here $\mathcal{F}$ is the Fourier transform

$$(\mathcal{F}f)(\boldsymbol{\xi}) := (2\pi)^{-d/2} \int e^{-i\boldsymbol{\xi}\mathbf{x}} f(\mathbf{x}) d\mathbf{x}, \qquad \boldsymbol{\xi} \in \mathbb{R}^d,$$

the integration is over $\mathbb{R}^d$, and b is the symbol of B. We assume that $b(\mathbf{x}, \boldsymbol{\xi})$, $\mathbf{x}, \boldsymbol{\xi} \in \mathbb{R}^d$, is a smooth almost-periodic in $\mathbf{x}$ complex-valued function and, moreover, that for some countable set $\boldsymbol{\Theta}$ of frequencies we have

$$b(\mathbf{x}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \mathsf{e}_{\boldsymbol{\theta}}(\mathbf{x}) \tag{2.2}$$

where

$$\mathsf{e}_{\boldsymbol{\theta}}(\mathbf{x}) := e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.3}$$

and

$$\hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) := \mathrm{M}_{\mathbf{x}}\big(b(\mathbf{x}, \boldsymbol{\xi})\mathsf{e}_{-\boldsymbol{\theta}}(\mathbf{x})\big)$$

are the Fourier coefficients of b. (For any almost–periodic function f in $\mathbb{R}^d$ its mean $\mathrm{M}_{\mathbf{x}}\big(f(\mathbf{x})\big)$ is defined by

$$\mathrm{M}_{\mathbf{x}}\big(f(\mathbf{x})\big) := \lim_{L\to\infty} L^{-d} \int_{[-L/2,L/2)^d} f(\mathbf{x})d\mathbf{x}.)$$

We assume that the series (2.2) converges absolutely, and that b satisfies the symmetry condition

$$\hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \overline{\hat{b}(-\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\theta})},$$

so that the operator B is formally self-adjoint. For $R > 0$ let $\mathbb{1}_{\mathcal{B}_R}$ be the indicator function of the ball $\mathcal{B}_R := \big\{\boldsymbol{\xi} : |\boldsymbol{\xi}| < R\big\}$. We assume that there exists a constant C_0 such that

$$\|b\, \mathbb{1}_{\mathcal{B}_{C_0}}\|_{L_\infty(\mathbb{R}^d\times\mathbb{R}^d)} < \infty,$$

and that

$$\big(1 - \mathbb{1}_{\mathcal{B}_{C_0}}(\boldsymbol{\xi})\big)b(\mathbf{x}, \boldsymbol{\xi}) = \sum_{\iota\in J} |\boldsymbol{\xi}|^\iota b_\iota\big(\mathbf{x}, \boldsymbol{\xi}/|\boldsymbol{\xi}|\big), \tag{2.4}$$

where J is a discrete subset of $(-\infty, \varkappa]$ with

$$0 \leqslant \varkappa < 2w \tag{2.5}$$

(the first inequality here is assumed for convenience without loss of generality), and $b_\iota(\mathbf{x}, \boldsymbol{\eta})$ are smooth functions on $\mathbb{R}^d \times \mathbb{S}^{d-1}$ almost-periodic with respect to $\mathbf{x}$.

Let

$$\tilde{w} := (w + \varkappa)/2. \tag{2.6}$$

We introduce $\chi \in \mathsf{C}^\infty(\mathbb{R}_+)$ so that

$$\chi(r) = \begin{cases} r, & r \geqslant C_0, \\ 0, & r \leqslant C_0/2. \end{cases} \tag{2.7}$$

Remark 2.1. Increasing C_0 if necessary, we can guarantee that for any $\widetilde{J} \subset J$ and any $\widetilde{\boldsymbol{\Theta}} \subset \boldsymbol{\Theta}$ the operator $\widetilde{B}$ with the symbol $\tilde{b}$ given by

$$\tilde{b}(\mathbf{x}, \boldsymbol{\xi}) := \sum_{\iota\in\widetilde{J}} \Big(\chi\big(|\boldsymbol{\xi}|\big)\Big)^\iota \sum_{\boldsymbol{\theta}\in\widetilde{\boldsymbol{\Theta}}} \hat{b}_\iota\big(\boldsymbol{\theta}, \boldsymbol{\xi}/|\boldsymbol{\xi}|\big)\mathsf{e}_{\boldsymbol{\theta}}(\mathbf{x}) \tag{2.8}$$

satisfies

$$(-\Delta)^{\tilde{w}} - |\widetilde{B}| \geqslant 0. \tag{2.9}$$

We also assume that the coefficients in the expansion

$$b_\iota(\mathbf{x},\boldsymbol{\eta}) = \sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \hat{b}_\iota(\boldsymbol{\theta},\boldsymbol{\eta})\mathsf{e}_{\boldsymbol{\theta}}(\mathbf{x}), \qquad \mathbf{x}\in\mathbb{R}^d, \quad \boldsymbol{\eta}\in\mathbb{S}^{d-1}, \quad \iota\in J \tag{2.10}$$

can be represented by a series

$$\hat{b}_\iota(\boldsymbol{\theta},\eta_1,\dots,\eta_d) = \sum_{\tau\in\mathbb{N}_0^d} \hat{b}_\iota^{(\tau)}(\boldsymbol{\theta})\eta_1^{\tau_1}\cdots\eta_d^{\tau_d} \tag{2.11}$$

which converges absolutely in a ball of radius greater than one of $\mathbb{R}^d$.

Under the above assumptions H is a selfadjoint operator on the Sobolev space $\mathsf{H}^{2w}(\mathbb{R}^d)$. We are interested in the asymptotic behaviour of its integrated density of states $N(\lambda)$ as the spectral parameter λ tends to infinity.

Definition 2.2. Let $e(\lambda;\mathbf{x},\mathbf{y})$ be the kernel of the spectral projection of H. We define the integrated density of states as

$$N(\lambda) := \mathrm{M}_{\mathbf{x}}\big(e(\lambda;\mathbf{x},\mathbf{x})\big).$$

It was proved in Theorem 4.1 of [16] that for differential operators this definition agrees with the traditional one (at least at its continuity points). The following lemma is proved at the end of Section 4 of [12].

Lemma 2.3. a. *If $A \geqslant B$, then $N(\lambda;A)\leqslant N(\lambda;B)$.*

b. *Suppose, $A = a(\mathbf{x},D)$ and $U = u(\mathbf{x},D)$ are two pseudo-differential operators with almost-periodic coefficients. Let operator A be elliptic self-adjoint and operator U be unitary. Then $N(\lambda;A) = N(\lambda;U^{-1}AU)$.*

Without loss of generality we assume that $\boldsymbol{\Theta}$ (recall (2.2)) spans $\mathbb{R}^d$, contains $\mathbf{0}$ and is symmetric about $\mathbf{0}$; we also put

$$\boldsymbol{\Theta}_k := \boldsymbol{\Theta}+\boldsymbol{\Theta}+\cdots+\boldsymbol{\Theta} \tag{2.12}$$

(algebraic sum taken k times) and $\boldsymbol{\Theta}_\infty := \cup_k\boldsymbol{\Theta}_k = Z(\boldsymbol{\Theta})$, where for a set $S\subset\mathbb{R}^d$ by $Z(S)$ we denote the set of all finite linear combinations of elements in S with integer coefficients. The set $\boldsymbol{\Theta}_\infty$ is countable and non-discrete (unless B is periodic). We will need

Condition A. *Suppose that $\boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_d\in\boldsymbol{\Theta}_\infty$. Then $Z(\boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_d)$ is discrete.*

It is easy to see that this condition can be reformulated like this: suppose that $\boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_d\in\boldsymbol{\Theta}_\infty$. Then either $\{\boldsymbol{\theta}_j\}$ are linearly independent, or $\sum_{j=1}^d n_j\boldsymbol{\theta}_j = 0$, where $n_j\in\mathbb{Z}$ and not all n_j are zeros. This reformulation shows that Condition A is generic: indeed, if we are choosing frequencies of b one after the other, then on each step we have to avoid choosing a new frequency from a countable set of hyperplanes, and this is obviously a generic restriction. Condition A is obviously satisfied for periodic B, but it becomes meaningful if B is quasi-periodic (i.e., if it is a linear combination of finitely many exponentials).

If $\boldsymbol{\Theta}$ and J are finite, Condition A is all we need. If, however, any (or both) of these sets is infinite, we need other conditions which describe, how well B can be approximated by operators with quasi-periodic symbols. In the proof we are going to work with quasi-periodic approximations of B, and we need these conditions to make sure that all estimates in the proof are uniform with respect to these approximations.

We introduce

$$\mathsf{b}_\iota(\boldsymbol{\theta}) := \sup_{|\boldsymbol{\eta}|=1} \left|\hat{b}_\iota(\boldsymbol{\theta}, \boldsymbol{\eta})\right|, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}.$$

Condition B. *Let k be a positive integer. Then there exists $R_0 \geqslant C_0$ such that for each $\rho > R_0$ there exist a finite symmetric set $\widetilde{\boldsymbol{\Theta}} \subset \left(\boldsymbol{\Theta} \cap \mathcal{B}(\rho^{1/k})\right)$ (where $\mathcal{B}(r)$ is the ball of radius r centered at 0) and a finite subset $\widetilde{J} \subset J$ with*

$$\operatorname{card} \widetilde{J} \leqslant \rho^{1/k} \tag{2.13}$$

such that

$$\sum_{(\boldsymbol{\theta},\iota) \in (\boldsymbol{\Theta} \times J) \setminus (\widetilde{\boldsymbol{\Theta}} \times \widetilde{J})} \left(1 + |\boldsymbol{\theta}|^2\right)^{\varkappa/4} |R_0|^{\iota - \varkappa} \mathsf{b}_\iota(\boldsymbol{\theta}) \leqslant \rho^{-k}. \tag{2.14}$$

The last condition we need is a version of the Diophantine condition on the frequencies of b. First, we need some definitions. We fix a natural number $\tilde{k}$ (the choice of $\tilde{k}$ will be determined later by the order of the remainder in the asymptotic expansion) and denote $\widetilde{\boldsymbol{\Theta}}'_{\tilde{k}} := \widetilde{\boldsymbol{\Theta}}_{\tilde{k}} \setminus \{0\}$ (see (2.12) for the notation). We say that $\mathfrak{V}$ is a quasi-lattice subspace of dimension m, if $\mathfrak{V}$ is a linear span of m linearly independent vectors $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_m$ from $\widetilde{\boldsymbol{\Theta}}_{\tilde{k}}$. Obviously, the zero space (which we will denote by $\mathfrak{X}$) is a quasi-lattice subspace of dimension 0, and $\mathbb{R}^d$ is a quasi-lattice subspace of dimension d. We denote by $\mathcal{V}_m$ the collection of all quasi-lattice subspaces of dimension m and put $\mathcal{V} := \cup_m \mathcal{V}_m$. If $\boldsymbol{\xi} \in \mathbb{R}^d$ and $\mathfrak{V}$ is a linear subspace of $\mathbb{R}^d$, we denote by $\boldsymbol{\xi}_{\mathfrak{V}}$ the orthogonal projection of $\boldsymbol{\xi}$ onto $\mathfrak{V}$, and put $\mathfrak{V}^\perp$ to be an orthogonal complement of $\mathfrak{V}$, so that $\boldsymbol{\xi} = \boldsymbol{\xi}_{\mathfrak{V}} + \boldsymbol{\xi}_{\mathfrak{V}^\perp}$. Let $\mathfrak{V}, \mathfrak{U} \in \mathcal{V}$. We say that these subspaces are *strongly distinct*, if neither of them is a subspace of the other one. This condition is equivalent to stating that if we put $\mathfrak{W} := \mathfrak{V} \cap \mathfrak{U}$, then $\dim \mathfrak{W}$ is strictly less than dimensions of $\mathfrak{V}$ and $\mathfrak{U}$. We put $\phi = \phi(\mathfrak{V}, \mathfrak{U}) \in [0, \pi/2]$ to be the angle between them, i.e. the angle between $\mathfrak{V} \ominus \mathfrak{W}$ and $\mathfrak{U} \ominus \mathfrak{W}$, where $\mathfrak{V} \ominus \mathfrak{W}$ is the orthogonal complement of $\mathfrak{W}$ in $\mathfrak{V}$. This angle is non-zero iff $\mathfrak{V}$ and $\mathfrak{W}$ are strongly distinct. We put $s = s(\rho) = s(\widetilde{\boldsymbol{\Theta}}_{\tilde{k}}) := \inf \sin\left(\phi(\mathfrak{V}, \mathfrak{U})\right)$, where the infimum is over all strongly distinct pairs of subspaces from $\mathcal{V}$, $R = R(\rho) := \sup_{\boldsymbol{\theta} \in \widetilde{\boldsymbol{\Theta}}_{\tilde{k}}} |\boldsymbol{\theta}|$, and $r = r(\rho) := \inf_{\boldsymbol{\theta} \in \widetilde{\boldsymbol{\Theta}}'_{\tilde{k}}} |\boldsymbol{\theta}|$. Obviously,

$$R(\rho) = O(\rho^{1/k}), \tag{2.15}$$

where the implied constant can depend on k and $\tilde{k}$.

Condition C. *For each fixed k and $\tilde{k}$ the sets $\widetilde{\mathbf{\Theta}}_{\tilde{k}}$ can be chosen in such a way that for sufficiently large ρ the number of elements in $\widetilde{\mathbf{\Theta}}_{\tilde{k}}$ satisfies* $\operatorname{card}\widetilde{\mathbf{\Theta}}_{\tilde{k}} \leqslant \rho^{1/k}$ *and we have*

$$s(\rho) \geqslant \rho^{-1/k} \tag{2.16}$$

and

$$r(\rho) \geqslant \rho^{-1/k},$$

where the implied constant (i.e. how large should ρ be) can depend on k and $\tilde{k}$.

Remark 2.4. Note that Condition C is automatically satisfied for quasi-periodic and smooth periodic B; see [12] for further discussion of this condition.

Condition A implies the following statement, which will be used crucially in our constructions.

Corollary 2.5. *Suppose, $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_l \in \widetilde{\mathbf{\Theta}}_{\tilde{k}}$, $l \leqslant d-1$. Let $\mathfrak{V}$ be the span of $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_l$. Then each element of the set $\widetilde{\mathbf{\Theta}}_{\tilde{k}} \cap \mathfrak{V}$ is a linear combination of $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_l$ with rational coefficients. Since the set $\widetilde{\mathbf{\Theta}}_{\tilde{k}} \cap \mathfrak{V}$ is finite, this implies that the set $Z(\widetilde{\mathbf{\Theta}}_{\tilde{k}} \cap \mathfrak{V})$ is discrete and is, therefore, a lattice in $\mathfrak{V}$.*

From now on, we always assume that B satisfies all the conditions from this section; we will also denote

$$\rho := \lambda^{1/2w}.$$

Now we can formulate our main theorem.

Theorem 2.6. *Let H be an operator* (2.1) *satisfying Conditions* A, B *and* C. *Then for each $K \in \mathbb{R}$ there exists a finite positive integer L and a finite subset $J_0 \subset J$ such that*

$$\begin{aligned} &N(\rho^{2w}) \\ &= \sum_{q=0}^{d-1} \sum_{h=0}^{L} \sum_{\iota_1,\dots,\iota_h \in J_0} \sum_{j=0}^{[K+d+(2-2w)h+\iota_1+\dots+\iota_h]} C_{q\,h\,j}^{\iota_1\cdots\iota_h} \rho^{d+(2-2w)h+\iota_1+\dots+\iota_h-j} \ln^q \rho \\ &\quad + O(\rho^{-K}). \end{aligned} \tag{2.17}$$

as $\rho \to \infty$.

Remark 2.7. The powers of ρ present in (2.17) are equal to $d + (2-2w)h + \iota_1 + \dots + \iota_h - j$, and the first impression is that there are far too many of them (indeed, a priori the set of all such powers can be dense in $\mathbb{R}$, for instance). However, many of these powers are, in fact, spurious (i.e. the corresponding coefficients $C_{q\,h\,j}^{\iota_1\cdots\iota_h}$ are zero). This happens, for example, when $d + (2-2w)h + \iota_1 + \dots + \iota_h - j > d$ (for obvious reasons). Equally obviously, these powers do not 'multiply' when we increase K. This means that if $K_1 < K_2$, then expansion (2.17) with $K = K_2$ does not contain extra terms with $d + (2-2w)h + \iota_1 + \dots + \iota_h - j > -K_1$, compared to this expansion for $K = K_1$.

In the case of magnetic Schrödinger operators, Theorem 2.6 and calculations similar to those of [2] and [12] imply that most of the terms in (2.17) will indeed disappear:

Corollary 2.8. *Suppose that smooth almost-periodic functions* $\mathbf{A}$ *and* V *are such that*

$$H = (-i\nabla + \mathbf{A})^2 + V =: -\Delta + B$$

satisfies the hypothesis of Theorem 2.6. Then for each $K \in \mathbb{N}$ *we have:*

$$N(\lambda) = \lambda^{d/2}\Big(C_d + \sum_{j=1}^{K} e_j \lambda^{-j} + o(\lambda^{-K})\Big) \tag{2.18}$$

as $\lambda \to \infty$.

Remark 2.9. By taking the Laplace transform of (2.17), one can obtain an asymptotic expansion of the (regularised) heat trace as $t \to 0$. However, it seems that using the approach of [2] and [3], it is possible to obtain even stronger results (the pointwise asymptotic expansion of the heat kernel).

Remark 2.10. Of course, formula (2.17) cannot be differentiated; moreover, we do not even know if in the almost periodic case $N(\lambda)$ is strictly increasing. However, in the periodic Schrödinger case there are some results on the high-energy behaviour of the (non-integrated) density of states, see e. g. [9].

Given Conditions B and C, we want to introduce the following definition. We say that a non-negative function $f = f(\rho) = f(\rho; k, \tilde{k})$ satisfies the estimate $f(\rho) \leqslant \rho^{0+}$ (resp. $f(\rho) \geqslant \rho^{0-}$), if for each positive ε and for each $\tilde{k}$ we can achieve $f(\rho) \leqslant \rho^{\varepsilon}$ (resp. $f(\rho) \geqslant \rho^{-\varepsilon}$) for sufficiently large ρ by choosing parameter k from Conditions B and C sufficiently large. For example, we have

$$R(\rho) \leqslant \rho^{0+}, \tag{2.19}$$

card $\widetilde{\mathbf{\Theta}} \leqslant \rho^{0+}$, $s(\rho) \geqslant \rho^{0-}$, and $r(\rho) \geqslant \rho^{0-}$.

Throughout the paper, we always assume that the value of k is chosen sufficiently large so that all inequalities of the form $\rho^{0+} \leqslant \rho^{\varepsilon}$ or $\rho^{0-} \geqslant \rho^{-\varepsilon}$ we encounter in the proof are satisfied.

The next statement proved in [12] is an example of how this new notation is used.

Lemma 2.11. *Suppose,* $\boldsymbol{\theta}, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_d \in \widetilde{\mathbf{\Theta}}'_{\tilde{k}}$*, the set* $\{\boldsymbol{\mu}_j\}$ *is linearly independent, and* $\boldsymbol{\theta} = \sum_{j=1}^d c_j \boldsymbol{\mu}_j$*. Then each non-zero coefficient* c_j *satisfies*

$$\rho^{0-} \leqslant |c_j| \leqslant \rho^{0+}.$$

In this paper, by C we denote positive constants, the exact value of which can be different each time they occur in the text, possibly even in the same formula. On the other hand, the constants which are labeled (like C_1, C_{21}, etc) have their values being fixed throughout the text. Given two positive functions f and g, we say that $f \gtrsim g$, or $g \lesssim f$, or $g = O(f)$ if the ratio g/f is bounded. We say $f \asymp g$ if $f \gtrsim g$ and $f \lesssim g$.

We will also need a number of auxiliary constants. Let us choose numbers $\{\alpha_j\}_{j=1}^d$, β, ϑ, and ς satisfying

$$\max\{1-w+\varkappa/2, 1/2\} < \beta < \alpha_1 < \alpha_2 < \dots < \alpha_d < \vartheta < \varsigma < 1 \tag{2.20}$$

(recall (2.5)), and set

$$\alpha := \varkappa/\beta. \tag{2.21}$$

3. Reduction to a finite interval of spectral parameter

To begin with, we choose sufficiently large $\rho_0 > C_0$ (to be fixed later on) and for $n \in \mathbb{N}$ put $\rho_n := 2\rho_{n-1} = 2^n\rho_0$. We also define the intervals $I_n := [\rho_n, 4\rho_n]$. The proof of Theorem 2.6 will be based on the following lemma:

Lemma 3.1. *For each $M \in \mathbb{R}$ there exist $L > 0$ and a finite subset $J_0 \subset J$ such that for every $n \in \mathbb{N}$ and $\rho \in I_n$*

$$\begin{aligned} &N(\rho^{2w}) \\ &= \sum_{q=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\dots,\iota_h\in J_0}\sum_{j=0}^{[\frac{d+M}{1-\varsigma}]} C_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M)\rho^{d+(2-2w)h+\iota_1+\dots+\iota_h-j}\ln^q\rho \\ &\quad + O(\rho_n^{-M}). \end{aligned} \tag{3.1}$$

Here, $C_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M)$ are some real numbers satisfying

$$C_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M) = O(\rho_n^{-2\beta h+\varsigma j}). \tag{3.2}$$

The constants in the O-terms do not depend on n (but they may depend on M).

Remark 3.2. Note that (3.1) is not a 'proper' asymptotic formula, since the coefficients are allowed to grow with n (and, therefore, with ρ).

Some of the powers of ρ on the right hand side of (3.1) may coincide. In order to avoid the ambiguity let us redefine coefficients $C_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M)$ in such a way that, for any given values of q and $d+(2-2w)h+\iota_1+\dots+\iota_h-j$, only the coefficient with the minimal possible value of h and maximal possible values of $j, \iota_1, \dots, \iota_h$ (in this order) is nonzero. Note that these new coefficients still satisfy (3.2).

Let us prove Theorem 2.6 assuming that we have proved Lemma 3.1. Let M be fixed. Denote the sum on the right hand side of (3.1) by $N_n(\rho^{2w})$. Then, for $n \geqslant 1$, whenever $\rho \in I_{n-1} \cap I_n = [\rho_n, 2\rho_n]$, we have:

$$\begin{aligned} &N_n(\rho^{2w}) - N_{n-1}(\rho^{2w}) \\ &= \sum_{q=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\dots,\iota_h\in J_0}\sum_{j=0}^{[\frac{d+M}{1-\varsigma}]} t_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M)\rho^{d+(2-2w)h+\iota_1+\dots+\iota_h-j}\ln^q\rho \\ &\quad + O(\rho_n^{-M}), \end{aligned} \tag{3.3}$$

where

$$t^{\iota_1\cdots\iota_h}_{q\,h\,j}(n,M) := C^{\iota_1\cdots\iota_h}_{q\,h\,j}(n,M) - C^{\iota_1\cdots\iota_h}_{q\,h\,j}(n-1,M).$$

On the other hand, since for $\rho \in I_{n-1} \cap I_n$ we have both $N(\rho^{2w}) = N_n(\rho^{2w}) + O(\rho_n^{-M})$ and $N(\rho^{2w}) = N_{n-1}(\rho^{2w}) + O(\rho_n^{-M})$, this implies that

$$\sum_{q=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\ldots,\iota_h\in J_0}\sum_{j=0}^{[\frac{d+M}{1-\varsigma}]} t^{\iota_1\cdots\iota_h}_{q\,h\,j}(n,M)\rho^{d+(2-2w)h+\iota_1+\cdots+\iota_h-j}\ln^q\rho = O(\rho_n^{-M}). \tag{3.4}$$

Claim 3.3. *For each combination of indices present on the right hand side of* (3.3) *we have:*

$$t^{\iota_1\cdots\iota_h}_{q\,h\,j}(n,M) = O(\rho_n^{j-M-d+(2w-2)h-\iota_1-\cdots-\iota_h}\ln^{d-1-q}\rho_n). \tag{3.5}$$

Proof. Put $y := \rho_n/\rho$ and let

$$\tau^{\iota_1\cdots\iota_h}_{p\,h\,j}(n,M) := \rho_n^{M+d+(2-2w)h+\iota_1+\cdots+\iota_h-j}\sum_{q=p}^{d-1}\binom{q}{p}(-1)^p t^{\iota_1\cdots\iota_h}_{q\,h\,j}(n,M)\ln^{q-p}\rho_n. \tag{3.6}$$

Then by (3.4) for $y \in [1/2,1]$

$$P(y) := \sum_{p=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\ldots,\iota_h\in J_0}\sum_{j=0}^{[\frac{d+M}{1-\varsigma}]} \tau^{\iota_1\cdots\iota_h}_{p\,h\,j}(n,M)y^{j-d+(2w-2)h-\iota_1-\cdots-\iota_h}\ln^p y = O(1) \quad \text{as } n\to\infty. \tag{3.7}$$

Let us denote by $f_1,\ldots,f_T$ the functions $y^{j-d+(2w-2)h-\iota_1-\cdots-\iota_h}\ln^p y$ entering the sum in (3.7) with non-zero coefficients; these functions are linearly independent on the interval $[1/2,1]$. Therefore, there exist points $y_1,\ldots,y_T \in [1/2,1]$ such that the determinant of the matrix $\big(f_j(y_l)\big)_{j,l=1}^T$ is non-zero. Now (3.7) and the Cramer's Rule imply that the values of $\tau^{\iota_1\cdots\iota_h}_{p\,h\,j}(n,M)$ are fractions with a bounded expression in the numerator and a fixed non-zero number in the denominator. Therefore,

$$\tau^{\iota_1\cdots\iota_h}_{p\,h\,j}(n,M) = O(1) \quad \text{as } n\to\infty. \tag{3.8}$$

Thus, choosing $p = d-1$ in (3.6), we obtain

$$t^{\iota_1\cdots\iota_h}_{d-1\,h\,j}(n,M) = O(\rho_n^{j-M-d+(2w-2)h-\iota_1-\cdots-\iota_h}).$$

Now we can put $p = d-2$ into (3.8) and obtain

$$t^{\iota_1\cdots\iota_h}_{d-1\,h\,j}(n,M) = O(\rho_n^{j-M-d+(2w-2)h-\iota_1-\cdots-\iota_h}\ln\rho_n).$$

Continuing this process until $p = 0$, we obtain (3.5). □

Therefore, the series $\sum_{m=0}^{\infty} t_{q\,h\,j}^{\iota_1\cdots\iota_h}(m,M)$ is absolutely convergent for $j < M + d + (2-2w)h + \iota_1 + \cdots + \iota_h$; moreover, for such j we have:

$$
\begin{aligned}
C_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M) &= C_{q\,h\,j}^{\iota_1\cdots\iota_h}(0,M) + \sum_{m=1}^{n} t_{q\,h\,j}^{\iota_1\cdots\iota_h}(m,M) \\
&= C_{q\,h\,j}^{\iota_1\cdots\iota_h}(0,M) + \sum_{m=1}^{\infty} t_{q\,h\,j}^{\iota_1\cdots\iota_h}(m,M) \\
&\quad + O(\rho_n^{j-M-d+(2w-2)h-\iota_1-\cdots-\iota_h} \ln^{d-1-q}\rho_n) \\
&=: C_{q\,h\,j}^{\iota_1\cdots\iota_h}(M) + O(\rho_n^{j-M-d+(2w-2)h-\iota_1-\cdots-\iota_h} \ln^{d-1-q}\rho_n),
\end{aligned}
$$

where we have denoted $C_{q\,h\,j}^{\iota_1\cdots\iota_h}(M) := C_{q\,h\,j}^{\iota_1\cdots\iota_h}(0,M) + \sum_{m=1}^{\infty} t_{q\,h\,j}^{\iota_1\cdots\iota_h}(m,M)$. For bigger values of j we use (3.2) and (2.20) to obtain

$$
\begin{aligned}
&\sum_{q=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\ldots,\iota_h\in J_0}\ \sum_{j\geqslant M+d+(2-2w)h+\iota_1+\cdots+\iota_h}^{[\frac{d+M}{1-\varsigma}]} \left|C_{q\,h\,j}^{\iota_1\cdots\iota_h}(n,M)\right| \\
&\quad\times \rho^{d+(2-2w)h+\iota_1+\cdots+\iota_h-j}\ln^q\rho \\
&\lesssim \sum_{q=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\ldots,\iota_h\in J_0} \rho_n^{\varsigma d+(2\varsigma-2\beta-2\varsigma w+\varsigma\varkappa)h-(1-\varsigma)M}\ln^q\rho_n \\
&\lesssim \rho_n^{\varsigma d-(1-\varsigma)M}\ln^{d-1}\rho_n.
\end{aligned}
$$

Thus, when $\rho\in I_n$, we have:

$$
\begin{aligned}
N(\rho^{2w}) = &\sum_{q=0}^{d-1}\sum_{h=0}^{L}\sum_{\iota_1,\ldots,\iota_h\in J_0}\ \sum_{j=0}^{[M+d+(2-2w)h+\iota_1+\cdots+\iota_h]} C_{q\,h\,j}^{\iota_1\cdots\iota_h}(M) \\
&\times \rho^{d+(2-2w)h+\iota_1+\cdots+\iota_h-j}\ln^q\rho \\
&+ O(\rho^{-M}\ln^{d-1}\rho) + O(\rho^{\varsigma d-(1-\varsigma)M}\ln^{d-1}\rho).
\end{aligned}
$$

Since the constants in O terms do not depend on n, it is sufficient to choose

$$M := \big[(\varsigma d + K)/(1-\varsigma)\big] + 1$$

to get (2.17) for all $\rho \geqslant \rho_0$.

Remark 3.4. The main reason why we need the representation (2.4) is to match the asymptotic expansions in different intervals I_n. If we did not have (2.4), we would have obtained the asymptotic expansions containing the general 'phase volumes' (like in [19]), and it is not clear how to merge the results for different n.

The rest of the paper is devoted to proving Lemma 3.1. The first step of the proof is fixing n and fixing large $\tilde{k}$ and k. The precise value of $\tilde{k}$ will be chosen later; the only restriction on it will be to satisfy inequality (9.9) (it says that the more asymptotic terms we want to have in (3.1), the bigger $\tilde{k}$ we need to choose; note that the choice of $\tilde{k}$ does not depend on k). We will have several requirements on how large k should be (most of them will be of the form $\rho_n^{0+} < \rho_n^{\varepsilon}$ or $\rho_n^{0-} > \rho_n^{-\varepsilon}$); each time we have such an inequality, we assume that k is chosen sufficiently large to satisfy it.

Remark 3.5. Our choice of k will only depend on M, w, $\varkappa$, and the constants introduced in (2.20). The set J_0 in Lemma 3.1 can be chosen to be

$$J_0 := J \cap [\varkappa - d - M - 1, \varkappa]. \tag{3.9}$$

The first requirement on k we have is that

$$k > d + M + \varkappa(d + M)/(w - \varkappa) - 2w. \tag{3.10}$$

After fixing $\tilde{k}$ and k we get R_0 from Condition B. Then, taking

$$\rho_0 \geqslant R_0 \tag{3.11}$$

and fixing n, we choose $\widetilde{\boldsymbol{\Theta}}$ and $\widetilde{J}$ so that Conditions B and C are satisfied for $\rho := 4\rho_n$. Without loss of generality we may assume that $\widetilde{J} \supset J_0$. Then we introduce an auxiliary pseudo–differential operator $\widetilde{B}$ with the symbol $\tilde{b}$ given by (2.8).

From now on we prove Lemma 3.1 for $B = \widetilde{B}$ and with J_0 replaced by $\widetilde{J}$. However, in view of (2.13) and (2.20), the results with $\widetilde{J}$ and J_0 are equivalent. Afterwards, in Section 11 we will prove that the asymptotics (3.1) for the original B follows from Condition B and (3.1) for $\widetilde{B}$.

4. Pseudo-differential operators

Most of the material in this and several subsequent sections is very similar to the corresponding sections of [12] and [13], as are the proofs of most of the statements. Therefore, we will often omit the proofs, instead referring the reader to [12], [19], and [13].

4.1. Classes of PDO's

Before we define the pseudo-differential operators (PDO's), we introduce the relevant classes of symbols. Let $b = b(\mathbf{x}, \boldsymbol{\xi})$, $\mathbf{x}, \boldsymbol{\xi} \in \mathbb{R}^d$, be an almost-periodic (in $\mathbf{x}$) complex-valued function and, moreover, for some countable set $\hat{\boldsymbol{\Theta}}$ of frequencies (we always assume that $\hat{\boldsymbol{\Theta}}$ is symmetric and contains 0; starting from the middle of this section, $\hat{\boldsymbol{\Theta}}$ will be assumed to be finite)

$$b(\mathbf{x}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\theta} \in \hat{\boldsymbol{\Theta}}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \mathsf{e}_{\boldsymbol{\theta}}(\mathbf{x}), \tag{4.1}$$

where

$$\hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) := \mathrm{M}_{\mathbf{x}}\big(b(\mathbf{x}, \boldsymbol{\xi}) \mathsf{e}_{-\boldsymbol{\theta}}(\mathbf{x})\big)$$

are the Fourier coefficients of $b(\cdot, \boldsymbol{\xi})$ (recall that M is the mean of an almost-periodic function). We always assume that (4.1) converges absolutely. Let us now define the classes of symbols we will consider and operators associated with them. For $\boldsymbol{\xi} \in \mathbb{R}^d$ let $\langle \boldsymbol{\xi} \rangle := \sqrt{1 + |\boldsymbol{\xi}|^2}$. We notice that

$$\langle \boldsymbol{\xi} + \boldsymbol{\eta} \rangle \leqslant 2 \langle \boldsymbol{\xi} \rangle \langle \boldsymbol{\eta} \rangle, \ \forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^d. \tag{4.2}$$

We say that a symbol b belongs to the class $\mathbf{S}_\alpha = \mathbf{S}_\alpha(\beta) = \mathbf{S}_\alpha(\beta, \hat{\mathbf{\Theta}})$, if for any $l \geqslant 0$ and any non-negative $s \in \mathbb{Z}$ the conditions

$$\big|b\big|_{l,s}^{(\alpha)} := \max_{|\mathbf{s}| \leqslant s} \sum_{\boldsymbol{\theta} \in \hat{\mathbf{\Theta}}} \langle \boldsymbol{\theta} \rangle^l \sup_{\boldsymbol{\xi}} \langle \boldsymbol{\xi} \rangle^{(-\alpha + |\mathbf{s}|)\beta} \big|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})\big| < \infty, \ |\mathbf{s}| = s_1 + s_2 + \cdots + s_d, \tag{4.3}$$

are fulfilled. The quantities (4.3) define norms on the class $\mathbf{S}_\alpha$. Note that $\mathbf{S}_\alpha$ is an increasing function of α, i.e. $\mathbf{S}_\alpha \subset \mathbf{S}_\gamma$ for $\alpha < \gamma$.

Given $\boldsymbol{\theta} \in \mathbb{R}^d$, let us introduce a linear map $\nabla_{\boldsymbol{\theta}}$ on symbols which acts according to the rule

$$\widehat{(\nabla_{\boldsymbol{\theta}} a)}(\boldsymbol{\phi}, \boldsymbol{\xi}) := \hat{a}(\boldsymbol{\phi}, \boldsymbol{\xi} + \boldsymbol{\theta}) - \hat{a}(\boldsymbol{\phi}, \boldsymbol{\xi}). \tag{4.4}$$

If the Fourier transform of the symbol is factorized, i.e.

$$\hat{a}(\boldsymbol{\phi}, \boldsymbol{\xi}) = \prod_{q=1}^{Q} \hat{a}_q(\boldsymbol{\phi}, \boldsymbol{\xi}),$$

then the action of $\nabla_{\boldsymbol{\theta}}$ can be written as a sum of actions on each factor separately:

$$\widehat{(\nabla_{\boldsymbol{\theta}} a)}(\boldsymbol{\phi}, \boldsymbol{\xi}) = \sum_{q=1}^{Q} \prod_{l=1}^{q-1} \hat{a}_l(\boldsymbol{\phi}, \boldsymbol{\xi} + \boldsymbol{\theta}) \widehat{(\nabla_{\boldsymbol{\theta}} a_q)}(\boldsymbol{\phi}, \boldsymbol{\xi}) \prod_{s=q-1}^{Q} \hat{a}_s(\boldsymbol{\phi}, \boldsymbol{\xi}). \tag{4.5}$$

For later reference we mention here the following convenient bound that follows from definition (4.3) and property (4.2):

$$\begin{aligned} &\sum_{\boldsymbol{\theta} \in \hat{\mathbf{\Theta}}} \langle \boldsymbol{\theta} \rangle^l \sup_{\boldsymbol{\xi}} \langle \boldsymbol{\xi} \rangle^{(-\alpha + s + 1)\beta} \Big(\big|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}} \widehat{(\nabla_{\boldsymbol{\eta}} b)}(\boldsymbol{\theta}, \boldsymbol{\xi})\big| \Big) \\ &\leqslant C \big|b\big|_{l,s+1}^{(\alpha)} \langle \boldsymbol{\eta} \rangle^{|\alpha - s - 1|\beta} |\boldsymbol{\eta}|, \qquad s = |\mathbf{s}|, \end{aligned} \tag{4.6}$$

with a constant C depending only on α, s, and β. The estimate (4.6) implies that for all $\boldsymbol{\eta}$ with $|\boldsymbol{\eta}| \leqslant C$ we have a uniform bound

$$\big|\nabla_{\boldsymbol{\eta}} b\big|_{l,s}^{(\alpha-1)} \leqslant C \big|b\big|_{l,s+1}^{(\alpha)} |\boldsymbol{\eta}|.$$

Now we define the PDO $\mathrm{Op}(b)$ in the usual way:

$$\mathrm{Op}(b) u(\mathbf{x}) = (2\pi)^{-d/2} \int b(\mathbf{x}, \boldsymbol{\xi}) e^{i\boldsymbol{\xi}\mathbf{x}} (\mathcal{F}u)(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{4.7}$$

the integral being over $\mathbb{R}^d$. Under the condition $b \in \mathbf{S}_\alpha$ the integral on the r.h.s. is clearly finite for any u from the Schwarz class $\mathsf{S}(\mathbb{R}^d)$. Moreover, the property $b \in \mathbf{S}_0$ guarantees the boundedness of $\mathrm{Op}(b)$ in $\mathsf{L}_2(\mathbb{R}^d)$, see Proposition 4.1. Unless otherwise stated, from now on $\mathsf{S}(\mathbb{R}^d)$ is taken as a natural domain for all PDO's when they act in $\mathsf{L}_2(\mathbb{R}^d)$.

Applying the standard regularization procedures to definition (4.7) (see, e.g., [15]), we can also consider the action of $\mathrm{Op}(b)$ on the exponentials $\mathsf{e}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} \in \mathbb{R}^d$. Namely, we have

$$\mathrm{Op}(b)\mathsf{e}_{\boldsymbol{\nu}} = \sum_{\boldsymbol{\theta} \in \hat{\mathbf{\Theta}}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\nu}) \mathsf{e}_{\boldsymbol{\nu} + \boldsymbol{\theta}}.$$

This action can be extended by linearity to all quasi-periodic functions (i.e. finite linear combinations of $\mathsf{e}_{\boldsymbol{\nu}}$ with different $\boldsymbol{\nu}$). By taking the closure, we can extend this action of $\mathrm{Op}(b)$ to the Besicovitch space $\mathsf{B}_2(\mathbb{R}^d)$. This is the space of all formal sums

$$\sum_{j=1}^{\infty} a_j \mathsf{e}_{\boldsymbol{\theta}_j}(\mathbf{x}), \quad \text{with} \quad \sum_{j=1}^{\infty} |a_j|^2 < +\infty.$$

It is known (see [15]) that the spectra of $\mathrm{Op}(b)$ acting in $\mathsf{L}_2(\mathbb{R}^d)$ and $\mathsf{B}_2(\mathbb{R}^d)$ are the same, although the types of the spectra can be entirely different. It is very convenient, when working with the gauge transform constructions, to assume that all the operators involved act in $\mathsf{B}_2(\mathbb{R}^d)$, although in the end we will return to operators acting in $\mathsf{L}_2(\mathbb{R}^d)$. This trick (working with operators acting in $\mathsf{B}_2(\mathbb{R}^d)$) is similar to working with fibre operators in the periodic case in the sense that we can freely consider the action of an operator on one, or finitely many, exponentials (2.3), despite the fact that these exponentials do not belong to our original function space.

Moreover, if the order $\alpha = 0$ then by continuity this action can be extended to all of $\mathsf{B}_2(\mathbb{R}^d)$, and the extension has the same norm as $\mathrm{Op}(b)$ acting in L_2 (see [15]). Thus, in what follows, when we speak about a pseudo-differential operator with almost-periodic symbol acting in B_2, we mean that its domain is either whole B_2 (when the order is non-positive), or the space of all quasi-periodic functions (for operators with positive order). And, when we make a statement about the norm of a pseudo-differential operator with almost-periodic symbol, we will not specify whether the operator acts in $\mathsf{L}_2(\mathbb{R}^d)$ or $\mathsf{B}_2(\mathbb{R}^d)$, since these norms are the same.

4.2. Some basic results on the calculus of almost-periodic PDO's

We begin by listing some elementary results for almost-periodic PDO's. The proofs are very similar (with obvious changes) to the proof of analogous statements in [19].

Proposition 4.1. *Suppose that b satisfies (4.1) and that $\left|b\right|^{(0)}_{0,0} < \infty$. Then $\mathrm{Op}(b)$ is bounded in both $\mathsf{L}_2(\mathbb{R}^d)$ and $\mathsf{B}_2(\mathbb{R}^d)$ and $\|\mathrm{Op}(b)\| \leqslant \left|b\right|^{(0)}_{0,0}$.*

In what follows, *if we need to calculate a product of two (or more) operators with some symbols $b_j \in \mathbf{S}_{\alpha_j}(\hat{\boldsymbol{\Theta}}_j)$ we will always consider that $b_j \in \mathbf{S}_{\alpha_j}(\sum_j \hat{\boldsymbol{\Theta}}_j)$ where, of course, all extra terms are assumed to have zero coefficients in front of them.*

Since $\mathrm{Op}(b)u \in \mathsf{S}(\mathbb{R}^d)$ for any $b \in \mathbf{S}_\alpha$ and $u \in \mathsf{S}(\mathbb{R}^d)$, the product $\mathrm{Op}(b)\,\mathrm{Op}(g)$, $b \in \mathbf{S}_\alpha(\hat{\boldsymbol{\Theta}}_1), g \in \mathbf{S}_\gamma(\hat{\boldsymbol{\Theta}}_2)$, is well defined on $\mathsf{S}(\mathbb{R}^d)$. A straightforward calculation leads to the following formula for the symbol $b \circ g$ of the product $\mathrm{Op}(b)\,\mathrm{Op}(g)$:

$$(b \circ g)(\mathbf{x}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\theta} \in \hat{\boldsymbol{\Theta}}_1,\, \boldsymbol{\phi} \in \hat{\boldsymbol{\Theta}}_2} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\phi}) \hat{g}(\boldsymbol{\phi}, \boldsymbol{\xi}) e^{i(\boldsymbol{\theta}+\boldsymbol{\phi})\mathbf{x}},$$

and hence

$$\widehat{(b\circ g)}(\boldsymbol\chi, \boldsymbol\xi) = \sum_{\boldsymbol\theta+\boldsymbol\phi=\boldsymbol\chi} \hat b(\boldsymbol\theta, \boldsymbol\xi+\boldsymbol\phi)\hat g(\boldsymbol\phi, \boldsymbol\xi),\ \boldsymbol\chi\in\hat{\boldsymbol\Theta}_1+\hat{\boldsymbol\Theta}_2,\ \boldsymbol\xi\in\mathbb R^d. \tag{4.8}$$

We have

Proposition 4.2. *Let* $b\in \mathbf S_\alpha(\hat{\boldsymbol\Theta}_1)$, $g\in\mathbf S_\gamma(\hat{\boldsymbol\Theta}_2)$. *Then* $b\circ g\in \mathbf S_{\alpha+\gamma}(\hat{\boldsymbol\Theta}_1+\hat{\boldsymbol\Theta}_2)$ *and*

$$\big|b\circ g\big|^{(\alpha+\gamma)}_{l,s}\leqslant C\big|b\big|^{(\alpha)}_{l,s}\big|g\big|^{(\gamma)}_{l+(|\alpha|+s)\beta,s},$$

with the constant C *depending only on* l, α, *and* s.

We are also interested in the estimates for symbols of commutators. For PDO's $A, \Psi_l,\ l=1,2,\dots,N$, denote

$$\mathrm{ad}(A;\Psi_1,\Psi_2,\dots,\Psi_N) := i\big[\mathrm{ad}(A;\Psi_1,\Psi_2,\dots,\Psi_{N-1}),\Psi_N\big],$$

$$\mathrm{ad}(A;\Psi):=i[A,\Psi],\quad \mathrm{ad}^N(A;\Psi):=\mathrm{ad}(A;\Psi,\Psi,\dots,\Psi),\quad \mathrm{ad}^0(A;\Psi):=A.$$

For the sake of convenience, we use the notation $\mathrm{ad}(a;\psi_1,\psi_2,\dots,\psi_N)$ and $\mathrm{ad}^N(a,\psi)$ for the symbols of multiple commutators.

Let

$$\mathrm{supp}\,\hat b:=\big\{\boldsymbol\theta\in\mathbb R^d:\hat b(\boldsymbol\theta,\cdot)\not\equiv 0\big\}.$$

It follows from (4.8) that the Fourier coefficients of the symbol $\mathrm{ad}(b,g)$ are given by

$$\begin{aligned}&\widehat{\mathrm{ad}(b,g)}(\boldsymbol\chi,\boldsymbol\xi)\\ &= i\sum_{\boldsymbol\theta\in(\mathrm{supp}\,\hat b)\cup(\boldsymbol\chi-\mathrm{supp}\,\hat g)}\big[(\widehat{\nabla_{\boldsymbol\chi-\boldsymbol\theta}b})(\boldsymbol\theta,\boldsymbol\xi)\hat g(\boldsymbol\chi-\boldsymbol\theta,\boldsymbol\xi)-\hat b(\boldsymbol\theta,\boldsymbol\xi)\widehat{(\nabla_{\boldsymbol\theta}g)}(\boldsymbol\chi-\boldsymbol\theta,\boldsymbol\xi)\big].\end{aligned} \tag{4.9}$$

Proposition 4.3. *Let* $b\in\mathbf S_\alpha(\hat{\boldsymbol\Theta})$ *and* $g_j\in\mathbf S_{\gamma_j}(\hat{\boldsymbol\Theta}_j)$, $j=1,2,\dots,N$. *Then* $\mathrm{ad}(b;g_1,\dots,g_N)\in\mathbf S_\gamma(\hat{\boldsymbol\Theta}+\sum_j\hat{\boldsymbol\Theta}_j)$ *with*

$$\gamma=\alpha+\sum_{j=1}^N(\gamma_j-1),$$

and

$$\big|\,\mathrm{ad}(b;g_1,\dots,g_N)\big|^{(\gamma)}_{l,s}\leqslant C\big|b\big|^{(\alpha)}_{p,s+N}\prod_{j=1}^N\big|g_j\big|^{(\gamma_j)}_{p,s+N-j+1},$$

where C *and* p *depend on* l,s,N,α *and* γ_j.

5. Resonant regions

We now define resonant regions and mention some of their properties. This material is essentially identical to Section 5 of [12], where the reader can find the proofs of all the statements of this section.

Recall the definition of the set $\boldsymbol\Theta=\widetilde{\boldsymbol\Theta}$ as well as of the quasi-lattice subspaces from Section 2. As before, by $\boldsymbol\Theta_{\tilde k}$ we denote the algebraic sum of $\tilde k$ copies of $\boldsymbol\Theta$; remember that we consider $\tilde k$ fixed. We also put $\boldsymbol\Theta'_{\tilde k}:=\boldsymbol\Theta_{\tilde k}\setminus\{0\}$.

For each $\mathfrak{V} \in \mathcal{V}$ we put $S_{\mathfrak{V}} := \{\boldsymbol{\xi} \in \mathfrak{V},\ |\boldsymbol{\xi}| = 1\}$. For each non-zero $\boldsymbol{\theta} \in \mathbb{R}^d$ we put $\mathbf{n}(\boldsymbol{\theta}) := \boldsymbol{\theta}|\boldsymbol{\theta}|^{-1}$.

Let $\mathfrak{V} \in \mathcal{V}_m$. We say that $\mathfrak{F}$ is a *flag* generated by $\mathfrak{V}$, if $\mathfrak{F}$ is a sequence $\mathfrak{V}_j \in \mathcal{V}_j$ $(j = 0, 1, \dots, m)$ such that $\mathfrak{V}_{j-1} \subset \mathfrak{V}_j$ and $\mathfrak{V}_m = \mathfrak{V}$. We say that $\{\boldsymbol{\nu}_j\}_{j=1}^m$ is a sequence generated by $\mathfrak{F}$ if $\boldsymbol{\nu}_j \in \mathfrak{V}_j \ominus \mathfrak{V}_{j-1}$ and $\|\boldsymbol{\nu}_j\| = 1$ (obviously, this condition determines each $\boldsymbol{\nu}_j$ up to multiplication by -1). We denote by $\mathcal{F}(\mathfrak{V})$ the collection of all flags generated by $\mathfrak{V}$. We put

$$L_j := \rho_n^{\alpha_j}, \tag{5.1}$$

recall (2.20).

Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}'_{\tilde{k}}$. The *resonant region* generated by $\boldsymbol{\theta}$ is defined as

$$\Lambda(\boldsymbol{\theta}) := \Big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \big|\langle \boldsymbol{\xi}, \mathbf{n}(\boldsymbol{\theta})\rangle\big| \leqslant L_1\Big\}. \tag{5.2}$$

Suppose, $\mathfrak{F} \in \mathcal{F}(\mathfrak{V})$ is a flag and $\{\boldsymbol{\nu}_j\}_{j=1}^m$ is a sequence generated by $\mathfrak{F}$. We define

$$\Lambda(\mathfrak{F}) := \Big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \big|\langle \boldsymbol{\xi}, \boldsymbol{\nu}_j\rangle\big| \leqslant L_j\Big\}. \tag{5.3}$$

If $\dim \mathfrak{V} = 1$, definition (5.3) is reduced to (5.2). Obviously, if $\mathfrak{F}_1 \subset \mathfrak{F}_2$, then $\Lambda(\mathfrak{F}_2) \subset \Lambda(\mathfrak{F}_1)$.

Suppose, $\mathfrak{V} \in \mathcal{V}_j$. We denote

$$\Xi_1(\mathfrak{V}) := \cup_{\mathfrak{F} \in \mathcal{F}(\mathfrak{V})} \Lambda(\mathfrak{F}).$$

Note that $\Xi_1(\mathfrak{X}) = \mathbb{R}^d$ and $\Xi_1(\mathfrak{V}) = \Lambda(\boldsymbol{\theta})$ if $\mathfrak{V} \in \mathcal{V}_1$ is spanned by $\boldsymbol{\theta}$. Finally, we put

$$\Xi(\mathfrak{V}) := \Xi_1(\mathfrak{V}) \setminus \big(\cup_{\mathfrak{U} \supsetneq \mathfrak{V}} \Xi_1(\mathfrak{U})\big) = \Xi_1(\mathfrak{V}) \setminus \big(\cup_{\mathfrak{U} \supsetneq \mathfrak{V}} \cup_{\mathfrak{F} \in \mathcal{F}(\mathfrak{U})} \Lambda(\mathfrak{F})\big). \tag{5.4}$$

We call $\Xi(\mathfrak{V})$ the resonance region generated by $\mathfrak{V}$. Very often, the region $\Xi(\mathfrak{X})$ is called the non-resonance region. We, however, will omit using this terminology since we will treat all regions $\Xi(\mathfrak{V})$ in the same way.

The first set of properties follows immediately from the definitions.

Lemma 5.1. *(i) We have*

$$\cup_{\mathfrak{V} \in \mathcal{V}} \Xi(\mathfrak{V}) = \mathbb{R}^d.$$

(ii) $\boldsymbol{\xi} \in \Xi_1(\mathfrak{V})$ *iff* $\boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})$, *where* $\Omega(\mathfrak{V}) \subset \mathfrak{V}$ *is a certain bounded set (more precisely,* $\Omega(\mathfrak{V}) = \Xi_1(\mathfrak{V}) \cap \mathfrak{V} \subset \mathcal{B}(mL_m)$ *if* $\dim \mathfrak{V} = m$*).*

(iii) $\Xi_1(\mathbb{R}^d) = \Xi(\mathbb{R}^d)$ *is a bounded set,* $\Xi(\mathbb{R}^d) \subset \mathcal{B}(dL_d)$*; all other sets* $\Xi_1(\mathfrak{V})$ *are unbounded.*

Now we move to slightly less obvious properties. From now on we always assume that ρ_0 (and thus ρ_n) is sufficiently large. We also assume, as we always do, that the value of k is sufficiently large so that, for example, $L_j \rho_n^{0+} < L_{j+1}$.

Lemma 5.2. *Let* $\mathfrak{V}, \mathfrak{U} \in \mathcal{V}$. *Then* $\big(\Xi_1(\mathfrak{V}) \cap \Xi_1(\mathfrak{U})\big) \subset \Xi_1(\mathfrak{W})$, *where* $\mathfrak{W} := \mathfrak{V} + \mathfrak{U}$ *(algebraic sum).*

Corollary 5.3. *(i) We can re-write definition* (5.4) *like this:*

$$\boldsymbol{\Xi}(\mathfrak{V}) := \boldsymbol{\Xi}_1(\mathfrak{V}) \setminus \big(\cup_{\mathfrak{U} \not\subset \mathfrak{V}} \boldsymbol{\Xi}_1(\mathfrak{U}) \big).$$

(ii) If $\mathfrak{V} \neq \mathfrak{U}$*, then* $\boldsymbol{\Xi}(\mathfrak{V}) \cap \boldsymbol{\Xi}(\mathfrak{U}) = \emptyset$.

(iii) We have $\mathbb{R}^d = \sqcup_{\mathfrak{V} \in \mathcal{V}} \boldsymbol{\Xi}(\mathfrak{V})$ *(the disjoint union).*

Lemma 5.4. *Let* $\mathfrak{V} \in \mathcal{V}_m$ *and* $\mathfrak{V} \subset \mathfrak{W} \in \mathcal{V}_{m+1}$*. Let* $\boldsymbol{\mu}$ *be (any) unit vector from* $\mathfrak{W} \ominus \mathfrak{V}$*. Then, for* $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{V})$*, we have* $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{W})$ *if and only if the estimate* $\big|\langle \boldsymbol{\xi}, \boldsymbol{\mu}\rangle\big| = \big|\langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\mu}\rangle\big| \leqslant L_{m+1}$ *holds.*

Lemma 5.5. *We have*

$$\boldsymbol{\Xi}_1(\mathfrak{V}) \cap \cup_{\mathfrak{U} \supsetneq \mathfrak{V}} \boldsymbol{\Xi}_1(\mathfrak{U}) = \boldsymbol{\Xi}_1(\mathfrak{V}) \cap \cup_{\mathfrak{W} \supsetneq \mathfrak{V},\ \dim \mathfrak{W} = 1 + \dim \mathfrak{V}} \boldsymbol{\Xi}_1(\mathfrak{W}).$$

Corollary 5.6. *We can re-write* (5.4) *as*

$$\boldsymbol{\Xi}(\mathfrak{V}) := \boldsymbol{\Xi}_1(\mathfrak{V}) \setminus \big(\cup_{\mathfrak{W} \supsetneq \mathfrak{V}, \dim \mathfrak{W} = 1 + \dim \mathfrak{V}} \boldsymbol{\Xi}_1(\mathfrak{W}) \big). \tag{5.5}$$

Lemma 5.7. *Let* $\mathfrak{V} \in \mathcal{V}$ *and* $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}}$*. Suppose that* $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})$ *and both points* $\boldsymbol{\xi}$ *and* $\boldsymbol{\xi} + \boldsymbol{\theta}$ *are inside* $\Lambda(\boldsymbol{\theta})$*. Then* $\boldsymbol{\theta} \in \mathfrak{V}$ *and* $\boldsymbol{\xi} + \boldsymbol{\theta} \in \boldsymbol{\Xi}(\mathfrak{V})$.

Definition 5.8. Let $\boldsymbol{\theta}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_l$ be some vectors from $\boldsymbol{\Theta}'_{\tilde{k}}$, which are not necessarily distinct.

a. We say that two vectors $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^d$ are $\boldsymbol{\theta}$*-resonant congruent* if both $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are inside $\Lambda(\boldsymbol{\theta})$ and $(\boldsymbol{\xi} - \boldsymbol{\eta}) = l\boldsymbol{\theta}$ with $l \in \mathbb{Z}$. In this case we write $\boldsymbol{\xi} \leftrightarrow \boldsymbol{\eta} \mod \boldsymbol{\theta}$.

b. For each $\boldsymbol{\xi} \in \mathbb{R}^d$ we denote by $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}(\boldsymbol{\xi})$ the set of all points which are $\boldsymbol{\theta}$-resonant congruent to $\boldsymbol{\xi}$. For $\boldsymbol{\theta} \neq \mathbf{0}$ we say that $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \varnothing$ if $\boldsymbol{\xi} \notin \Lambda(\boldsymbol{\theta})$.

c. We say that $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_l$*-resonant congruent*, if there exists a sequence $\boldsymbol{\xi}_j \in \mathbb{R}^d, j = 0, 1, \dots, l$ such that $\boldsymbol{\xi}_0 = \boldsymbol{\xi}$, $\boldsymbol{\xi}_l = \boldsymbol{\eta}$, and $\boldsymbol{\xi}_j \in \boldsymbol{\Upsilon}_{\boldsymbol{\theta}_j}(\boldsymbol{\xi}_{j-1})$ for $j = 1, 2, \dots, l$.

d. We say that $\boldsymbol{\eta} \in \mathbb{R}^d$ and $\boldsymbol{\xi} \in \mathbb{R}^d$ are *resonant congruent*, if either $\boldsymbol{\xi} = \boldsymbol{\eta}$ or $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_l$-resonant congruent with some $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_l \in \boldsymbol{\Theta}'_{\tilde{k}}$. The set of **all** points, resonant congruent to $\boldsymbol{\xi}$, is denoted by $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$. For points $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$ (note that this condition is equivalent to $\boldsymbol{\xi} \in \boldsymbol{\Upsilon}(\boldsymbol{\eta})$) we write $\boldsymbol{\eta} \leftrightarrow \boldsymbol{\xi}$.

Note that $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) = \{\boldsymbol{\xi}\}$ for any $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{X})$. Now Lemma 5.7 immediately implies

Corollary 5.9. *For each* $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})$ *we have* $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \boldsymbol{\Xi}(\mathfrak{V})$ *and thus*

$$\boldsymbol{\Xi}(\mathfrak{V}) = \sqcup_{\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})} \boldsymbol{\Upsilon}(\boldsymbol{\xi}).$$

Lemma 5.10. *The diameter of* $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$ *is bounded above by* mL_m*, if* $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})$*,* $\mathfrak{V} \in \mathcal{V}_m$*.*

Lemma 5.11. *For each* $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})$*,* $\mathfrak{V} \neq \mathbb{R}^d$*, the set* $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$ *is finite, and* $\operatorname{card} \boldsymbol{\Upsilon}(\boldsymbol{\xi})$ *is bounded uniformly in* $\boldsymbol{\xi} \in \mathbb{R}^d \setminus \boldsymbol{\Xi}(\mathbb{R}^d)$.

6. Description of the approach

We first prove (3.1) assuming that the symbol b of B is replaced by $\tilde{b}$ which satisfies (2.8). In particular, it belongs to the class $\mathbf{S}_\alpha$. At the end, in Section 11, we will use (2.14) to show that Theorem 2.6 holds as stated.

For any set $\mathcal{C} \subset \mathbb{R}^d$ by $\mathcal{P}(\mathcal{C})$ we denote the orthogonal projection onto $\mathrm{span}\{\mathsf{e}_{\boldsymbol{\xi}}\}_{\boldsymbol{\xi}\in\mathcal{C}}$ in $\mathsf{B}_2(\mathbb{R}^d)$ and by $\mathcal{P}^L(\mathcal{C})$ the same projection considered in $\mathsf{L}_2(\mathbb{R}^d)$, i.e.

$$\mathcal{P}^L(\mathcal{C}) = \mathcal{F}^* \, \mathbb{1}_{\mathcal{C}} \, \mathcal{F}, \tag{6.1}$$

where $\mathcal{F}$ is the Fourier transform and $\mathbb{1}_{\mathcal{C}}$ is the operator of multiplication by the indicator function of $\mathcal{C}$. Obviously, $\mathcal{P}^L(\mathcal{C})$ is a well-defined (respectively, non-zero) projection iff $\mathcal{C}$ is measurable (respectively, has non-zero measure). Let us fix sufficiently large n, and denote (recall that $\lambda_n = \rho_n^{2w}$)

$$\mathcal{X}_n := \Big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ |\boldsymbol{\xi}|^{2w} \in \big[(5/6)^{2w}\lambda_n, 5^{2w}\lambda_n\big]\Big\}. \tag{6.2}$$

We also put

$$\mathcal{A} = \mathcal{A}_n := \cup_{\boldsymbol{\xi}\in\mathcal{X}_n} \boldsymbol{\Upsilon}(\boldsymbol{\xi}).$$

Lemma 5.10 implies that, if ρ_0 is big enough,

$$\text{for each } \boldsymbol{\xi} \in \mathcal{A} \text{ we have } |\boldsymbol{\xi}|^{2w} \in \big[(2/3)^{2w}\lambda_n, 6^{2w}\lambda_n\big]. \tag{6.3}$$

In particular, we have

$$\mathcal{A} \cap \boldsymbol{\Xi}(\mathbb{R}^d) = \varnothing. \tag{6.4}$$

Let us define

$$\hat{\mathcal{A}} := \big\{\boldsymbol{\xi} \notin \mathcal{A},\ |\boldsymbol{\xi}|^{2w} < \lambda_n\big\}$$

and

$$\check{\mathcal{A}} := \big\{\boldsymbol{\xi} \notin \mathcal{A},\ |\boldsymbol{\xi}|^{2w} > \lambda_n\big\}. \tag{6.5}$$

We now plan to apply the gauge transform as in Sections 8 and 9 of [12] to the operator H. The details of this procedure will be explained in Sections 8 and 9; here, we just mention that we are going to introduce two operators: H_1 and H_2. The operator H_1 is unitary equivalent to H: $H_1 = U^{-1}HU$, where $U = e^{i\Psi}$ with a bounded pseudo-differential operator Ψ with almost-periodic coefficients (then Lemma 2.3 implies that the densities of states of H and H_1 are the same). Moreover, $H_1 = H_2 + R_{\tilde{k}}$, where

$$\|R_{\tilde{k}}\| \lesssim \rho_n^{-M+2w-d} \tag{6.6}$$

and $H_2 = (-\Delta)^w + W_{\tilde{k}}$ is a self-adjoint pseudo-differential operator with symbol $|\boldsymbol{\xi}|^{2w} + w_{\tilde{k}}(\mathbf{x}, \boldsymbol{\xi})$ which satisfies the following property:

$$\hat{w}_{\tilde{k}}(\boldsymbol{\theta}, \boldsymbol{\xi}) = 0, \text{ if } \big(\boldsymbol{\xi} \notin \Lambda(\boldsymbol{\theta}) \ \&\ \boldsymbol{\xi} \in \mathcal{A}\big), \text{ or } \big(\boldsymbol{\xi}+\boldsymbol{\theta} \notin \Lambda(\boldsymbol{\theta}) \ \&\ \boldsymbol{\xi} \in \mathcal{A}\big), \text{ or } (\boldsymbol{\theta} \notin \boldsymbol{\Theta}_{\tilde{k}}). \tag{6.7}$$

We can now use a simple statement which follows from Lemma 2.3 and Remark 2.7:

Lemma 6.1. *Suppose, H_1 and H_2 are two elliptic self-adjoint pseudo-differential operators with almost-periodic coefficients such that $\|H_1 - H_2\| \lesssim \rho_n^{-M+2w-d}$. Suppose that $N(H_2; \rho^{2w})$ satisfies asymptotic expansion* (3.1). *Then $N(H_1; \rho^{2w})$ also satisfies* (3.1) *with the same coefficients.*

This means that it is enough to establish the asymptotic expansion (3.1) for the operator H_2 instead of H. Condition (6.7) implies that for each $\boldsymbol{\xi} \in \mathcal{A}$ the subspace $\mathcal{P}\big(\boldsymbol{\Upsilon}(\boldsymbol{\xi})\big)\mathsf{B}_2(\mathbb{R}^d)$ is an invariant subspace of H_2; its dimension is finite by Lemma 5.11. We put

$$H_2(\boldsymbol{\xi}) := H_2|_{\mathcal{P}(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))\mathsf{B}_2(\mathbb{R}^d)}.$$

Note that the subspaces $\mathcal{P}(\hat{\mathcal{A}})\mathsf{B}_2(\mathbb{R}^d)$ and $\mathcal{P}(\check{\mathcal{A}})\mathsf{B}_2(\mathbb{R}^d)$ are invariant as well; by $H_2(\hat{\mathcal{A}})$ and $H_2(\check{\mathcal{A}})$ we denote the restrictions of H_2 to these subspaces; we also denote by $H_2(\mathcal{A})$ the restriction of H_2 to $\mathcal{P}(\mathcal{A})\mathsf{B}_2(\mathbb{R}^d)$. If we consider the operator H_2 acting in $\mathsf{L}_2(\mathbb{R}^d)$, then $\mathcal{P}^L(\hat{\mathcal{A}})\mathsf{L}_2(\mathbb{R}^d)$, $\mathcal{P}^L(\check{\mathcal{A}})\mathsf{L}_2(\mathbb{R}^d)$, and $\mathcal{P}^L(\mathcal{A})\mathsf{L}_2(\mathbb{R}^d)$ are still invariant subspaces. It follows from (6.2) – (6.5) that $UH_2(\hat{\mathcal{A}})U^* < (5/6)^{2w}\lambda_n I$ and $UH_2(\check{\mathcal{A}})U^* > 5^{2w}\lambda_n I$.

For each $\boldsymbol{\xi} \in \mathcal{A}$ the operator $H_2(\boldsymbol{\xi})$ is a finite-dimensional self-adjoint operator, so its spectrum is purely discrete; we denote its eigenvalues (counting multiplicities) by $\lambda_1(\boldsymbol{\xi}) \leqslant \lambda_2(\boldsymbol{\xi}) \leqslant \ldots \leqslant \lambda_{\mathrm{card}\,\boldsymbol{\Upsilon}(\boldsymbol{\xi})}(\boldsymbol{\xi})$. Next, we list all points $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$ in increasing order of their absolute values; thus, we have put into correspondence to each point $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$ a natural number $t = t(\boldsymbol{\eta})$ so that $t(\boldsymbol{\eta}) < t(\boldsymbol{\eta}')$ if $|\boldsymbol{\eta}| < |\boldsymbol{\eta}'|$. If two points $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_d)$ and $\boldsymbol{\eta}' = (\eta_1', \ldots, \eta_d')$ have the same absolute values, we put them in the lexicographic order of their coordinates, i.e. we say that $t(\boldsymbol{\eta}) < t(\boldsymbol{\eta}')$ if $\eta_1 < \eta_1'$, or $\eta_1 = \eta_1'$ and $\eta_2 < \eta_2'$, etc. Now we define the map $g : \mathcal{A} \to \mathbb{R}$ which to each point $\boldsymbol{\eta} \in \mathcal{A}$ brings into correspondence the number $\lambda_{t(\boldsymbol{\eta})}\big(\boldsymbol{\Upsilon}(\boldsymbol{\eta})\big)$. This map is an injection from $\mathcal{A}$ onto the set of eigenvalues of H_2, counting multiplicities (recall that we consider the operator H_2 acting in $\mathsf{B}_2(\mathbb{R}^d)$, so there is nothing miraculous about its spectrum consisting of eigenvalues and their limit points). Moreover, all eigenvalues of H_2 inside the interval $\big[(7/8)^{2w}\lambda_n, (9/2)^{2w}\lambda_n\big]$ have a pre-image under g. We define

$$g(\boldsymbol{\xi}) := |\boldsymbol{\xi}|^{2w}, \quad \text{for} \quad \boldsymbol{\xi} \in \mathbb{R}^d \setminus \mathcal{A}. \tag{6.8}$$

Arguments similar to the ones used in [13] show that g is a measurable function.

We introduce

$$\mathcal{G}_\lambda := \big\{\boldsymbol{\xi} \in \mathbb{R}^d,\, g(\boldsymbol{\xi}) \leqslant \lambda\big\}.$$

Lemma 6.2. *For $\lambda \in [\lambda_n, 4^{2w}\lambda_n]$ being a continuity point of $N(\lambda; H_2)$ we have:*

$$N(\lambda; H_2) = (2\pi)^{-d} \operatorname{vol} \mathcal{G}_\lambda. \tag{6.9}$$

Since points of continuity of $N(\lambda)$ are dense, *the asymptotic expansion proven for such λ can be extended to all $\lambda \in [\lambda_n, 4^{2w}\lambda_n]$ by taking the limit.* Thus, our next task is to compute $\operatorname{vol} \mathcal{G}_\lambda$. Let us put

$$\mathcal{A}^+(\rho) := \big\{\boldsymbol{\xi} \in \mathbb{R}^d,\, g(\boldsymbol{\xi}) < \rho^{2w} < |\boldsymbol{\xi}|^{2w}\big\}$$

and

$$\mathcal{A}^-(\rho) := \big\{\boldsymbol{\xi} \in \mathbb{R}^d,\, |\boldsymbol{\xi}|^{2w} < \rho^{2w} < g(\boldsymbol{\xi})\big\}.$$

Lemma 6.3.

$$\operatorname{vol}(\mathcal{G}_\lambda) = \omega_d\rho^d + \operatorname{vol}\mathcal{A}^+(\rho) - \operatorname{vol}\mathcal{A}^-(\rho), \tag{6.10}$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$.

Proof. We obviously have $\mathcal{G}_\lambda = \mathcal{B}(\rho) \cup \mathcal{A}^+(\rho) \setminus \mathcal{A}^-(\rho)$. Since $\mathcal{A}^-(\rho) \subset \mathcal{B}(\rho)$ and $\mathcal{A}^+(\rho) \cap \mathcal{B}(\rho) = \emptyset$, this implies (6.10). □

Remark 6.4. Properties of the mapping g imply that $\mathcal{A}^+(\rho) \cup \mathcal{A}^-(\rho) \subset \mathcal{A}$. Thus, in order to compute $N(\lambda)$, we need to analyze the behavior of g only inside $\mathcal{A}$.

We will compute volumes of $\mathcal{A}^\pm(\rho)$ by means of integrating their characteristic functions in a specially chosen set of coordinates. The next section is devoted to introducing these coordinates.

7. Coordinates

In this section, we do some preparatory work before computing $\operatorname{vol}\mathcal{A}^\pm(\rho)$. Namely, we are going to introduce a convenient set of coordinates in $\boldsymbol{\Xi}(\mathfrak{V})$. Let $\mathfrak{V} \in \mathcal{V}_m$ be fixed; since $\mathcal{A}^\pm(\rho) \cap \boldsymbol{\Xi}(\mathbb{R}^d) = \emptyset$, we will assume that $m < d$. Then, as we have seen, $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{V})$ if and only if $\boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})$. Let $\{\mathfrak{U}_j\}$ be a collection of all subspaces $\mathfrak{U}_j \in \mathcal{V}_{m+1}$ such that each $\mathfrak{U}_j$ contains $\mathfrak{V}$. Let $\boldsymbol{\mu}_j = \boldsymbol{\mu}_j(\mathfrak{V})$ be (any) unit vector from $\mathfrak{U}_j \ominus \mathfrak{V}$. Then it follows from Lemma 5.4 that for $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{V})$, we have $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{U}_j)$ if and only if the estimate $\big|\langle \boldsymbol{\xi}, \boldsymbol{\mu}_j\rangle\big| = \big|\langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\mu}_j\rangle\big| \leqslant L_{m+1}$ holds. Thus, formula (5.5) implies that

$$\boldsymbol{\Xi}(\mathfrak{V}) = \Big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})\ \&\ \forall j\ \big|\langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\mu}_j(\mathfrak{V})\rangle\big| > L_{m+1}\Big\}.$$

The collection $\big\{\boldsymbol{\mu}_j(\mathfrak{V})\big\}$ obviously coincides with

$$\big\{\mathbf{n}(\boldsymbol{\theta}_{\mathfrak{V}^\perp}),\ \boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}} \setminus \mathfrak{V}\big\}.$$

The set $\boldsymbol{\Xi}(\mathfrak{V})$ is, in general, disconnected; it consists of several connected components which we will denote by $\big\{\boldsymbol{\Xi}(\mathfrak{V})_p\big\}_{p=1}^P$. Let us fix a connected component $\boldsymbol{\Xi}(\mathfrak{V})_p$. Then for some vectors $\big\{\tilde{\boldsymbol{\mu}}_j(p)\big\}_{j=1}^{J_p} \subset \{\pm\boldsymbol{\mu}_j\}$ we have

$$\boldsymbol{\Xi}(\mathfrak{V})_p = \big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})\ \&\ \forall j\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > L_{m+1}\big\};$$

we assume that $\big\{\tilde{\boldsymbol{\mu}}_j(p)\big\}_{j=1}^{J_p}$ is the minimal set with this property, so that each hyperplane

$$\big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})\ \ \&\ \ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle = L_{m+1}\big\},\ j = 1, \dots, J_p$$

has a non-empty intersection with the boundary of $\boldsymbol{\Xi}(\mathfrak{V})_p$. It is not hard to see that $J_p \geqslant d - m$. Indeed, otherwise $\boldsymbol{\Xi}(\mathfrak{V})_p$ would have non-empty intersection with $\boldsymbol{\Xi}_1(\mathfrak{V}')$ for some $\mathfrak{V}'$, $\mathfrak{V} \subsetneq \mathfrak{V}'$. We also introduce

$$\tilde{\boldsymbol{\Xi}}(\mathfrak{V})_p := \big\{\boldsymbol{\xi} \in \mathfrak{V}^\perp,\ \forall j\ \langle \boldsymbol{\xi}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > 0\big\}.$$

Note that our assumption that $\Xi(\mathfrak{V})_p$ is a connected component of $\Xi(\mathfrak{V})$ implies that for any $\boldsymbol{\xi} \in \tilde{\Xi}(\mathfrak{V})_p$ and any $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}} \setminus \mathfrak{V}$ we have

$$\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle = \langle \boldsymbol{\xi}, \boldsymbol{\theta}_{\mathfrak{V}^\perp} \rangle \neq 0.$$

We also put

$$K := d - m - 1.$$

Without loss of generality we may (and will) assume that the number J_p of 'defining planes' is the minimal possible, i.e. $J_p = K+1$. Indeed, the argument presented in Section 11 of [12] explains how to derive the result for arbitrary $\Xi(\mathfrak{V})_p$, assuming we have proved it in the case $J_p = K+1$.

If $J_p = K+1$, then the set $\{\tilde{\boldsymbol{\mu}}_j(p)\}_{j=1}^{K+1}$ is linearly independent. Let $\mathbf{a} = \mathbf{a}(p)$ be a unique point from $\mathfrak{V}^\perp$ satisfying the following conditions: $\langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_j(p) \rangle = L_{m+1}$, $j = 1, \dots, K+1$. Then, since the determinant of the Gram matrix of vectors $\tilde{\boldsymbol{\mu}}_j(p)$ is $\gtrsim \rho_n^{0-}$ by (2.16), we have

$$|\mathbf{a}| \lesssim L_{m+1}\rho_n^{0+} = \rho_n^{\alpha_{m+1}+0+}. \tag{7.1}$$

We introduce shifted cylindrical coordinates in $\Xi(\mathfrak{V})_p$. These coordinates will be denoted by $\boldsymbol{\xi} = (r; \boldsymbol{\Phi}; \mathbf{X})$. Here, $\mathbf{X} = (X_1, \dots, X_m)$ is an arbitrary set of cartesian coordinates in $\Omega(\mathfrak{V})$. These coordinates do not depend on the choice of the connected component $\Xi(\mathfrak{V})_p$. The rest of the coordinates $(r, \boldsymbol{\Phi})$ are shifted spherical coordinates in $\mathfrak{V}^\perp$, centered at $\mathbf{a}$. This means that

$$r(\boldsymbol{\xi}) = |\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}|$$

and

$$\boldsymbol{\Phi} = \mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}) \in S_{\mathfrak{V}^\perp}.$$

More precisely, $\boldsymbol{\Phi} \in \mathcal{M}_p$, where $\mathcal{M}_p := \{\mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}),\ \boldsymbol{\xi} \in \Xi(\mathfrak{V})_p\} \subset S_{\mathfrak{V}^\perp}$ is a K-dimensional spherical simplex with $K+1$ sides. Note that

$$\begin{aligned} \mathcal{M}_p &= \{\mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}),\ \boldsymbol{\xi} \in \Xi(\mathfrak{V})_p\} = \{\mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}),\ \forall j\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p) \rangle > L_{m+1}\} \\ &= \{\mathbf{n}(\boldsymbol{\eta}),\ \boldsymbol{\eta} := \boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a} \in \mathfrak{V}^\perp,\ \forall j\ \langle \boldsymbol{\eta}, \tilde{\boldsymbol{\mu}}_j(p) \rangle > 0\} = S_{\mathfrak{V}^\perp} \cap \tilde{\Xi}(\mathfrak{V})_p. \end{aligned}$$

We will denote by $d\boldsymbol{\Phi}$ the spherical Lebesgue measure on $\mathcal{M}_p$. For each non-zero vector $\boldsymbol{\mu} \in \mathfrak{V}^\perp$, we denote

$$\mathcal{W}(\boldsymbol{\mu}) := \{\boldsymbol{\eta} \in \mathfrak{V}^\perp,\ \langle \boldsymbol{\eta}, \boldsymbol{\mu} \rangle = 0\}.$$

Thus, the sides of the simplex $\mathcal{M}_p$ are intersections of $\mathcal{W}(\tilde{\boldsymbol{\mu}}_j(p))$ with the sphere $S_{\mathfrak{V}^\perp}$. Each vertex $\mathbf{v} = \mathbf{v}_t$, $t = 1, \dots, K+1$ of $\mathcal{M}_p$ is an intersection of $S_{\mathfrak{V}^\perp}$ with K hyperplanes $\mathcal{W}(\tilde{\boldsymbol{\mu}}_j(p))$, $j = 1, \dots, K+1$, $j \neq t$. This means that $\mathbf{v}_t$ is a unit vector from $\mathfrak{V}^\perp$ which is orthogonal to $\{\tilde{\boldsymbol{\mu}}_j(p)\}$, $j = 1, \dots, K+1$, $j \neq t$; this defines $\mathbf{v}$ up to a multiplication by -1.

Lemma 7.1. *Let $\mathfrak{U}_1$ and $\mathfrak{U}_2$ be two strongly distinct subspaces each of which is a linear combination of some of the vectors from $\{\tilde{\boldsymbol{\mu}}_j(p)\}$. Then the angle between them is not smaller than $s(\rho_n)$. In particular, all non-zero angles between two sides of any dimensions of $\mathcal{M}_p$ as well as all the distances between two vertexes $\mathbf{v}_t$ and $\mathbf{v}_\tau$, $t \neq \tau$, are bounded below by $s(\rho_n)$.*

Lemma 7.2. *Let p be fixed. Suppose, $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}} \setminus \mathfrak{V}$ and $\boldsymbol{\theta}_{\mathfrak{V}^\perp} = \sum_{j=1}^{K+1} c_j \tilde{\boldsymbol{\mu}}_j(p)$. Then either all coefficients c_j are non-positive, or all of them are non-negative.*

By taking sufficiently large $\tilde{k}$ we can assure that the diameter of $\mathcal{M}_p$ does not exceed $(100d^2)^{-1}$. We put $\Phi_q := \frac{\pi}{2} - \phi\big(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}, \tilde{\boldsymbol{\mu}}_q(p)\big)$, $q = 1, \dots, K+1$. The geometrical meaning of these coordinates is simple: Φ_q is the spherical distance between $\boldsymbol{\Phi} = \mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a})$ and $\mathcal{W}\big(\tilde{\boldsymbol{\mu}}_q(p)\big)$. The reason why we have introduced Φ_q is that in these coordinates some important objects will be especially simple (see e.g. Lemma 7.5 below) which is very convenient for integration. At the same time, the set of coordinates $\big(r, \{\Phi_q\}\big)$ contains $K+2$ variables, whereas we only need $K+1$ coordinates in $\mathfrak{V}^\perp$. Thus, we have one constraint for variables Φ_j. Namely, let $\{\mathbf{h}_j\}$, $j = 1, \dots, K+1$ be a fixed orthonormal basis in $\mathfrak{V}^\perp$ chosen in such a way that the $K+1$-st axis is directed along $\mathbf{a}$, and thus passes through $\mathcal{M}_p$. Then we have $\mathbf{h}_j = \sum_{l=1}^{K+1} a_{jl} \tilde{\boldsymbol{\mu}}_l$ with some matrix $\{a_{jl}\}$, $j, l = 1, \dots, K+1$, and $\tilde{\boldsymbol{\mu}}_l = \tilde{\boldsymbol{\mu}}_l(p)$. Therefore (recall that we denote $\boldsymbol{\eta} := \boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}$),

$$\eta_j = \langle \boldsymbol{\eta}, \mathbf{h}_j \rangle = r \sum_{q=1}^{K+1} a_{jq} \sin \Phi_q$$

and, since $r^2(\boldsymbol{\xi}) = |\boldsymbol{\eta}|^2 = \sum_{j=1}^{K+1} \eta_j^2$, this implies that

$$\sum_{j=1}^{K+1} \Big(\sum_{q=1}^{K+1} a_{jq} \sin \Phi_q \Big)^2 = 1,$$

which is our constraint.

Let us also put

$$\eta_j' := \frac{\eta_j}{|\boldsymbol{\eta}|} = \sum_{q=1}^{K+1} a_{jq} \sin \Phi_q. \tag{7.2}$$

Then we can write the surface element $d\boldsymbol{\Phi}$ in the coordinates $\{\eta_j'\}$ as

$$d\boldsymbol{\Phi} = \frac{d\eta_1' \dots d\eta_K'}{\eta_{K+1}} = \frac{d\eta_1' \dots d\eta_K'}{\big(1 - \sum_{j=1}^{K} (\eta_j')^2\big)^{1/2}},$$

where the denominator is bounded below by $1/2$ by our choice of the basis $\{\mathbf{h}_j\}$. It follows from our choice of the coordinates and (7.2) that

$$\langle \mathbf{a}, \boldsymbol{\Phi} \rangle = \langle \mathbf{a}, \mathbf{n}(\boldsymbol{\eta}) \rangle = |\mathbf{a}| \eta_{K+1}' = |\mathbf{a}| \sum_{q=1}^{K+1} a_{K+1\,q} \sin \Phi_q. \tag{7.3}$$

Lemma 7.3. *For each p, l we have $|a_{pl}| \leqslant s(\rho_n)^{-1}$.*

Lemma 7.4. *We have $\max_j \sin \Phi_j(\boldsymbol{\eta}) \geqslant s(\rho_n) d^{-3/2}$.*

The next lemma describes the dependence on r of all possible inner products $\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}}$, $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})_p$.

Lemma 7.5. *Let* $\boldsymbol{\xi} \in \Xi(\mathfrak{V})_p$, $\mathfrak{V} \in \mathcal{V}_m$, *and* $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}}$.

(i) If $\boldsymbol{\theta} \in \mathfrak{V}$, *then* $\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle$ *does not depend on* r.

(ii) If $\boldsymbol{\theta} \notin \mathfrak{V}$ *and* $\boldsymbol{\theta}_{\mathfrak{V}^\perp} = \sum_q c_q \tilde{\boldsymbol{\mu}}_q(p)$, *then*

$$\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle = \langle \mathbf{X}, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + L_{m+1} \sum_q c_q + r(\boldsymbol{\xi}) \sum_q c_q \sin \Phi_q.$$

In the case (ii) all the coefficients c_q *are either non-positive or non-negative and each non-zero coefficient* c_q *satisfies*

$$\rho_n^{0-} \lesssim |c_q| \lesssim \rho_n^{0+}. \tag{7.4}$$

8. Partition of the perturbation

The symbols we are going to construct in this section will depend on ρ_n; this dependence will usually be omitted from the notation.

Let $\varpi \in \mathsf{C}^\infty(\mathbb{R})$ be such that

$$0 \leqslant \varpi \leqslant 1, \quad \varpi(z) = \begin{cases} 1, & z \leqslant 1; \\ 0, & z \geqslant 21/20. \end{cases} \tag{8.1}$$

For $\boldsymbol{\theta} \in \boldsymbol{\Theta}'$ we define several C^∞-cut-off functions:

$$\begin{cases} e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) & := \varpi\Big(\big|2|2\boldsymbol{\xi} + \boldsymbol{\theta}|/\rho_n - 15\big|/13\Big), \\ \ell^{>}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) & := 1 - \varpi\Big(\big(2|2\boldsymbol{\xi} + \boldsymbol{\theta}|/\rho_n - 15\big)/13\Big), \\ \ell^{<}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) & := 1 - \varpi\Big(\big(15 - 2|2\boldsymbol{\xi} + \boldsymbol{\theta}|/\rho_n\big)/13\Big), \end{cases} \tag{8.2}$$

and

$$\begin{cases} \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}) & := \varpi\left(\dfrac{\big|\langle \boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\theta}/2 \rangle\big|}{\rho_n^{\beta} |\boldsymbol{\theta}|}\right), \\ \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) & := 1 - \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}). \end{cases} \tag{8.3}$$

Remark 8.1. Note that $e_{\boldsymbol{\theta}} + \ell^{>}_{\boldsymbol{\theta}} + \ell^{<}_{\boldsymbol{\theta}} = 1$. The function $\ell^{>}_{\boldsymbol{\theta}}$ is supported on the set $|\boldsymbol{\xi} + \boldsymbol{\theta}/2| \geqslant 7\rho_n$, and $\ell^{<}_{\boldsymbol{\theta}}$ is supported on the set $|\boldsymbol{\xi} + \boldsymbol{\theta}/2| \leqslant \rho_n/2$. The function $e_{\boldsymbol{\theta}}$ is supported in the shell $\rho_n/3 \leqslant |\boldsymbol{\xi} + \boldsymbol{\theta}/2| \leqslant 8\rho_n$.

Using the notation $\ell_{\boldsymbol{\theta}}$ for any of the functions $\ell^{>}_{\boldsymbol{\theta}}$ or $\ell^{<}_{\boldsymbol{\theta}}$, we point out that

$$\begin{cases} e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = e_{-\boldsymbol{\theta}}(\boldsymbol{\xi} + \boldsymbol{\theta}), & \ell_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \ell_{-\boldsymbol{\theta}}(\boldsymbol{\xi} + \boldsymbol{\theta}), \\ \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \varphi_{-\boldsymbol{\theta}}(\boldsymbol{\xi} + \boldsymbol{\theta}), & \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \zeta_{-\boldsymbol{\theta}}(\boldsymbol{\xi} + \boldsymbol{\theta}). \end{cases}$$

Note that the above functions satisfy the estimates

$$\begin{cases} \big|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} e_{\boldsymbol{\theta}}(\boldsymbol{\xi})\big| + \big|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \ell_{\boldsymbol{\theta}}(\boldsymbol{\xi})\big| \lesssim \rho_n^{-|\mathbf{s}|}, \\ \big|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})\big| + \big|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi})\big| \lesssim \rho_n^{-\beta|\mathbf{s}|}. \end{cases} \tag{8.4}$$

Now for any symbol $b \in \mathbf{S}_\alpha(\beta)$ we introduce five new symbols:

$$b^{\mathcal{LE}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) := \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \ell^{>}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}},$$

$$b^{\mathcal{NR}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) := \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}},$$

$$b^{\mathcal{R}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) := \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}},$$

$$b^{\mathcal{SE}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) := \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \ell^{<}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}},$$

$$b^{o}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) = b^{o}(\boldsymbol{\xi}; \rho_n) := \hat{b}(0, \boldsymbol{\xi}).$$

The superscripts here are chosen to mean, respectively: 'large energy', 'non-resonant', 'resonant', 'small energy' and 0-th Fourier coefficient. The corresponding operators are denoted by $B^{\mathcal{LE}}$, $B^{\mathcal{NR}}$, $B^{\mathcal{R}}$, $B^{\mathcal{SE}}$, and B^{o}. By definitions (8.1), (8.2) and (8.3)

$$b = b^{o} + b^{\mathcal{SE}} + b^{\mathcal{R}} + b^{\mathcal{NR}} + b^{\mathcal{LE}}. \tag{8.5}$$

The role of each of these operators is easy to explain. Note that on the support of the functions $\hat{b}^{\mathcal{NR}}(\boldsymbol{\theta}, \cdot; \rho_n)$ and $\hat{b}^{\mathcal{R}}(\boldsymbol{\theta}, \cdot; \rho_n)$ we have (using (2.19))

$$\rho_n/3 - O(\rho_n^{0+}) \leqslant |\boldsymbol{\xi}| \leqslant 8\rho_n + O(\rho_n^{0+}).$$

On the support of $b^{\mathcal{SE}}(\boldsymbol{\theta}, \cdot; \rho_n)$ we have

$$|\boldsymbol{\xi}| \leqslant \rho_n/2 + O(\rho_n^{0+}). \tag{8.6}$$

On the support of $b^{\mathcal{LE}}(\boldsymbol{\theta}, \cdot\, ; \rho_n)$ we have

$$|\boldsymbol{\xi}| \geqslant 7\rho_n - O(\rho_n^{0+}). \tag{8.7}$$

The introduced symbols play a central role in the proof of Lemma 3.1. As we have seen in Section 6, due to (8.6) and (8.7) the symbols $b^{\mathcal{SE}}$ and $b^{\mathcal{LE}}$ make only a negligible contribution to the spectrum of the operator H near $\lambda = \rho^{2w}$ for $\rho \in I_n$. The only significant components of b are the symbols $b^{\mathcal{NR}}, b^{\mathcal{R}}$ and b^{o}. The symbol b^{o} will remain as it is, and the symbol $b^{\mathcal{NR}}$ will be transformed in the next section to another symbol, independent of $\mathbf{x}$.

Under the condition $b \in \mathbf{S}_\alpha(\beta)$ the above symbols belong to the same class $\mathbf{S}_\alpha(\beta)$ and the following bounds hold:

$$\big|b^{\mathcal{R}}\big|^{(\alpha)}_{l,s} + \big|b^{\mathcal{NR}}\big|^{(\alpha)}_{l,s} + \big|b^{\mathcal{LE}}\big|^{(\alpha)}_{l,s} + \big|b^{o}\big|^{(\alpha)}_{l,s} + \big|b^{\mathcal{SE}}\big|^{(\alpha)}_{l,s} \lesssim \big|b\big|^{(\alpha)}_{l,s}.$$

If b symmetric, then so are the symbols on the right hand side of (8.5).

Let us mention some other elementary properties of the introduced operators. In the lemma below we use the projection $\mathcal{P}(\mathcal{C})$, $\mathcal{C} \subset \mathbb{R}^d$ which was defined in Section 6.

Lemma 8.2. *Let $b \in \mathbf{S}_\alpha(\beta)$ with some $\alpha \in \mathbb{R}$. Then:*

(i) *The operator* $B^{\mathcal{S}\mathcal{E}}$ *is bounded and*

$$\|B^{\mathcal{S}\mathcal{E}}\| \lesssim \big|b\big|^{(\alpha)}_{0,0}\rho_n^{\beta\max(\alpha,0)}.$$

Moreover,

$$\Big(I - \mathcal{P}\big(\mathcal{B}(2\rho_n/3)\big)\Big)B^{\mathcal{S}\mathcal{E}} = B^{\mathcal{S}\mathcal{E}}\Big(I - \mathcal{P}\big(\mathcal{B}(2\rho_n/3)\big)\Big) = 0.$$

(ii) *The operator* $B^{\mathcal{R}}$ *satisfies the relations*

$$\begin{aligned}&\mathcal{P}\big(\mathcal{B}(\rho_n/6)\big)B^{\mathcal{R}} = B^{\mathcal{R}}\mathcal{P}\big(\mathcal{B}(\rho_n/6)\big)\\ &= \Big(I - \mathcal{P}\big(\mathcal{B}(9\rho_n)\big)\Big)B^{\mathcal{R}} = B^{\mathcal{R}}\Big(I - \mathcal{P}\big(9\rho_n)\big)\Big) = 0,\end{aligned}$$

and similar relations hold for the operator $B^{\mathcal{N}\mathcal{R}}$ *as well. Moreover,* $b^{\mathcal{N}\mathcal{R}}, b^{\mathcal{R}} \in \mathbf{S}_\gamma$ *for any* $\gamma \in \mathbb{R}$, *and for all* l *and* s

$$\big|b^{\mathcal{N}\mathcal{R}}\big|^{(\gamma)}_{l,s} + \big|b^{\mathcal{R}}\big|^{(\gamma)}_{l,s} \lesssim \rho_n^{\beta(\alpha-\gamma)}\big|b\big|^{(\alpha)}_{l,s},$$

with the implied constant independent of b *and* $n \geqslant 1$. *In particular, the operators* $B^{\mathcal{N}\mathcal{R}}, B^{\mathcal{R}}$ *are bounded and*

$$\|B^{\mathcal{N}\mathcal{R}}\| + \|B^{\mathcal{R}}\| \lesssim \rho_n^{\beta\alpha}\big|b\big|^{(\alpha)}_{0,0}.$$

(iii)

$$\mathcal{P}\big(\mathcal{B}(6\rho_n)\big)B^{\mathcal{L}\mathcal{E}} = B^{\mathcal{L}\mathcal{E}}\mathcal{P}\big(\mathcal{B}(6\rho_n)\big) = 0.$$

9. Operators H_1 and H_2

9.1. Preparation

As mentioned at the end of Section 3, we assume that the symbol b of B satisfies (2.8), and thus belongs to the class $\mathbf{S}_\alpha(\beta)$ with α defined in (2.21). Our strategy is to find a unitary operator which reduces $H = H_0 + B$, $H_0 := (-\Delta)^w$, to another PDO, whose symbol, essentially, depends only on $\boldsymbol{\xi}$. More precisely, we want to find operators H_1 and H_2 with the properties discussed in Section 6.

Repeating the calculations of Subsection 9.1 of [12] we find that H is unitarily equivalent to

$$H_1 = H_0 + Y^{(o)}_{\tilde{k}} + Y^{\mathcal{R}}_{\tilde{k}} + Y^{\mathcal{S}\mathcal{E},\mathcal{L}\mathcal{E}}_{\tilde{k}} + R_{\tilde{k}}, \tag{9.1}$$

where

$$Y_{\tilde{k}} := \sum_{l=1}^{\tilde{k}} B_l + \sum_{l=2}^{\tilde{k}} T_l, \tag{9.2}$$

$$B_1 := \mathrm{Op}(b),$$

$$B_l := \sum_{j=1}^{l-1} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l-1} \mathrm{ad}\big(\mathrm{Op}(b); \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j}\big),\ l \geqslant 2, \tag{9.3}$$

$$T_l := \sum_{j=2}^{l} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l} \mathrm{ad}(H_0; \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j}),\ l \geqslant 2, \tag{9.4}$$

$$R_{\tilde{k}} := \int_0^1 dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{\tilde{k}}} \exp(-it\Psi)\,\mathrm{ad}^{\tilde{k}+1}(H;\Psi)\exp(it\Psi)dt$$
$$+ \sum_{j=1}^{\tilde{k}} \frac{1}{j!} \sum_{\substack{k_1+k_2+\cdots+k_j \geqslant \tilde{k}+1, \\ k_q \leqslant \tilde{k},\ q=1,\ldots,j}} \mathrm{ad}(H; \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j}),$$

$$\Psi := \sum_{p=1}^{\tilde{k}} \Psi_p.$$

The symbols ψ_j of PDO Ψ_j are found from the following system of commutator equations:

$$\mathrm{ad}(H_0; \Psi_1) + B_1^{\mathcal{NR}} = 0, \tag{9.5}$$

$$\mathrm{ad}(H_0; \Psi_l) + B_l^{\mathcal{NR}} + T_l^{\mathcal{NR}} = 0,\ l \geqslant 2. \tag{9.6}$$

By Lemma 8.2(ii), the operators $B_l^{\mathcal{NR}}$, $T_l^{\mathcal{NR}}$ are bounded. This, in view of (9.5) and (9.6), implies boudedness of the commutators $\mathrm{ad}(H_0;\Psi_l)$, $l \geqslant 1$. Below we denote by $y_{\tilde{k}}$ the symbol of the PDO $Y_{\tilde{k}}$.

9.2. Commutator equations

Put

$$\tilde{\chi}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) := e_{\boldsymbol{\theta}}(\boldsymbol{\xi})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})\big(|\boldsymbol{\xi}+\boldsymbol{\theta}|^{2w} - |\boldsymbol{\xi}|^{2w}\big)^{-1} \tag{9.7}$$

when $\boldsymbol{\theta} \neq \mathbf{0}$, and $\tilde{\chi}_{\mathbf{0}}(\boldsymbol{\xi}) := 0$.

We have

Lemma 9.1. *Let $A = \mathrm{Op}(a)$ be a symmetric PDO with $a \in \mathbf{S}_\omega$. Then the PDO Ψ with the Fourier coefficients of the symbol $\psi(\mathbf{x},\boldsymbol{\xi})$ given by*

$$\hat{\psi}(\boldsymbol{\theta},\boldsymbol{\xi}) := i\,\hat{a}(\boldsymbol{\theta},\boldsymbol{\xi})\tilde{\chi}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) \tag{9.8}$$

solves the equation

$$\mathrm{ad}(H_0;\Psi) + \mathrm{Op}(a^{\mathcal{NR}}) = 0.$$

Moreover, the operator Ψ is bounded and self-adjoint, its symbol ψ belongs to $\mathbf{S}_\gamma$ with any $\gamma \in \mathbb{R}$ and the following bound holds:

$$\big|\psi\big|_{l,s}^{(\gamma)} \lesssim \rho_n^{\beta(\omega-\gamma-1)-2w+2} r(\rho_n)^{-1}\big|a\big|_{l-1,s}^{(\omega)} \lesssim \rho_n^{\beta(\omega-\gamma-1)-2w+2+0+}\big|a\big|_{l-1,s}^{(\omega)}.$$

The proof of this lemma is analogous to that of Lemma 4.1 of [13] and is based on the estimate

$$|\boldsymbol{\xi}+\boldsymbol{\theta}|^{2w}-|\boldsymbol{\xi}|^{2w}=|\boldsymbol{\xi}|^{2w}\Big(\big(1+|\boldsymbol{\xi}|^{-2}(2\boldsymbol{\xi}+\boldsymbol{\theta})\cdot\boldsymbol{\theta}\big)^{w}-1\Big)\asymp\rho^{2w-2}\big|\boldsymbol{\theta}\cdot(\boldsymbol{\xi}+\boldsymbol{\theta}/2)\big|$$

which holds for $\boldsymbol{\xi}$ in the support of $e_{\boldsymbol{\theta}}\varphi_{\boldsymbol{\theta}}$.

Using Propositions 4.1, 4.2, 4.3, Lemma 9.1, and repeating arguments from the proof of Lemma 4.2 from [13] (with $\sigma_j:=j\big(\alpha-2-(2w-2)\beta^{-1}\big)+1$), we obtain the following

Lemma 9.2. *Let $b\in\mathbf{S}_\alpha(\beta)$ be a symmetric symbol. Suppose that k is large enough so that $r(\rho_n)^{-1}\lesssim\rho_n^{0+}\lesssim\rho_n^{w+\beta-\frac{\alpha\beta}{2}-1}$ and $\tilde{k}$ satisfies*

$$\tilde{k}>2(M+\alpha\beta+d-2w)/(2w+2\beta-\alpha\beta-2). \tag{9.9}$$

Then ψ_j, b_j, $t_j\in\mathbf{S}_\gamma(\beta)$ for any $\gamma\in\mathbb{R}$ and there exists sufficiently large ρ_0, such that

$$\|R_{\tilde{k}}\|\lesssim\rho_n^{-M+2w-d}. \tag{9.10}$$

Remark 9.3. Note that the expression in the denominator of (9.9) is positive by (2.20) and (2.21).

Now Lemmas 6.1 and 9.2 imply that the contribution of $R_{\tilde{k}}$ to the integrated density of states can be neglected. More precisely, let $W_{\tilde{k}}$ be the operator with symbol

$$\begin{aligned} w_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi})&:=y_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi})-y_{\tilde{k}}^{\mathcal{NR}}(\mathbf{x},\boldsymbol{\xi}),\\ \text{i.e.}\quad \hat{w}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi})&=\hat{y}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi})\big(1-e_{\boldsymbol{\theta}}(\boldsymbol{\xi})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})\big). \end{aligned} \tag{9.11}$$

We introduce $H_2:=(-\Delta)^w+W_{\tilde{k}}$. Then, by (9.1) and (9.10), $\|H_1-H_2\|\lesssim\rho_n^{-M+2w-d}$ and, moreover, the symbol $w_{\tilde{k}}$ satisfies (6.7). This means that all the constructions of Section 6 are valid, and all we need to do is to compute $\operatorname{vol}\mathcal{G}_\lambda$.

Until this point, the material in our paper was quite similar to the corresponding parts of [12]. From now on, the differences will be substantial.

9.3. Computing the symbol of the operator after gauge transform

The following lemma provides us with a more explicit form of the symbol $\hat{y}_{\tilde{k}}$.

Lemma 9.4. *We have $\hat{y}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi})=0$ for $\boldsymbol{\theta}\notin\boldsymbol{\Theta}_{\tilde{k}}$. Otherwise,*

$$\begin{aligned} &\hat{y}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi})=\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\\ &+\sum_{s=1}^{\tilde{k}-1}\sum_{\substack{\boldsymbol{\theta}_j,\boldsymbol{\theta}_{s+1}\in\boldsymbol{\Theta}\\ \boldsymbol{\phi}_j,\boldsymbol{\phi}_{s+1},\boldsymbol{\phi}'_j\in\boldsymbol{\Theta}_{s+1}\\ \boldsymbol{\theta}'_j\in\boldsymbol{\Theta}'_{s+1}\\ 1\leqslant j\leqslant s}}\sum_{p=1}^{s}\sum_{\substack{\boldsymbol{\theta}''_q,\boldsymbol{\phi}''_q\in\boldsymbol{\Theta}'_{s+1}\\ 1\leqslant q\leqslant p-1}}\sum_{\substack{\nu_1,\dots,\nu_{2s+p}\geqslant0\\ \sum\nu_i=s}}\prod_{q=1}^{p-1}(\widehat{\nabla^{\nu_q}e_{\boldsymbol{\theta}''_q}\varphi_{\boldsymbol{\theta}''_q}})(\boldsymbol{\xi}+\boldsymbol{\phi}''_q)\\ &\times(\widehat{\nabla^{\nu_p}b})(\boldsymbol{\theta}_{s+1},\boldsymbol{\xi}+\boldsymbol{\phi}_{s+1})\prod_{j=1}^{s}(\widehat{\nabla^{\nu_{p+j}}b})(\boldsymbol{\theta}_j,\boldsymbol{\xi}+\boldsymbol{\phi}_j)(\widehat{\nabla^{\nu_{p+s+j}}\tilde{\chi}_{\boldsymbol{\theta}'_j}})(\boldsymbol{\xi}+\boldsymbol{\phi}'_j). \end{aligned} \tag{9.12}$$

Here for $\nu \in \mathbb{N}$

$$\begin{aligned}\nabla^{\nu} &:= \sum_{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_\nu \in \boldsymbol{\Theta}} C^{(s,p)}_{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_\nu}\big(\{\boldsymbol{\theta},\boldsymbol{\phi}\}\big)\nabla_{\boldsymbol{\eta}_1}\cdots\nabla_{\boldsymbol{\eta}_\nu};\\ \nabla^0 &:= C^{(s,p)}\big(\{\boldsymbol{\theta},\boldsymbol{\phi}\}\big),\end{aligned} \tag{9.13}$$

and, for $\boldsymbol{\theta} \in \mathbb{R}^d$*, the action of* $\nabla_{\boldsymbol{\theta}}$ *on symbols of PDO is defined in* (4.4)*, whereas for any function* f *on* $\mathbb{R}^d$

$$(\nabla_{\boldsymbol{\theta}} f)(\boldsymbol{\xi}) := f(\boldsymbol{\xi}+\boldsymbol{\theta}) - f(\boldsymbol{\xi}).$$

The coefficients $C^{(s,p)}\big(\{\boldsymbol{\theta},\boldsymbol{\phi}\}\big)$ *and* $C^{(s,p)}_{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_\nu}\big(\{\boldsymbol{\theta},\boldsymbol{\phi}\}\big)$ *depend on* s*,* p *and all vectors* $\boldsymbol{\theta}$*,* $\boldsymbol{\theta}_j$*,* $\boldsymbol{\theta}_{s+1}$*,* $\boldsymbol{\phi}_j$*,* $\boldsymbol{\phi}_{s+1}$*,* $\boldsymbol{\theta}'_j$*,* $\boldsymbol{\phi}'_j$*,* $\boldsymbol{\theta}''_q$*,* $\boldsymbol{\phi}''_q$ *(and on* $\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_\nu$ *if these subscripts are present). Moreover, these coefficients can differ for each particular* ∇^ν*,* $\nu \in \mathbb{N}_0$*. At the same time, they are uniformly bounded by a constant which depends on* $\tilde{k}$ *only.*
We apply the convention that $\prod_{q=1}^{0}(\nabla^{\nu_q}\widehat{e_{\boldsymbol{\theta}''_q}\varphi_{\boldsymbol{\theta}''_q}})(\boldsymbol{\xi}+\boldsymbol{\phi}''_q) = 1.$

Proof. We will prove the lemma by induction. Namely, let $\ell \geqslant 2$. We claim that:

1) For any $m = 1,\dots,\ell-1$, $\hat{\psi}_m(\boldsymbol{\theta},\boldsymbol{\xi}) = 0$ for $\boldsymbol{\theta} \notin \boldsymbol{\Theta}_m$. Otherwise,

$$\begin{aligned}&\hat{\psi}_m(\boldsymbol{\theta},\boldsymbol{\xi})\\ &= \sum_{\substack{\boldsymbol{\theta}_j\in\boldsymbol{\Theta}\\ \boldsymbol{\phi}_j,\boldsymbol{\phi}'_j\in\boldsymbol{\Theta}_m\\ \boldsymbol{\theta}'_j\in\boldsymbol{\Theta}'_m\\ 1\leqslant j\leqslant m}} \sum_{p=1}^{m} \sum_{\substack{\boldsymbol{\theta}''_q,\boldsymbol{\phi}''_q\in\boldsymbol{\Theta}'_m\\ 1\leqslant q\leqslant p-1}} \sum_{\substack{\nu_1,\dots,\nu_{2m+p-1}\geqslant 0\\ \sum\nu_i=m-1}} \prod_{q=1}^{p-1}(\nabla^{\nu_q}\widehat{e_{\boldsymbol{\theta}''_q}\varphi_{\boldsymbol{\theta}''_q}})(\boldsymbol{\xi}+\boldsymbol{\phi}''_q)\\ &\quad\times\prod_{j=1}^{m}(\nabla^{\nu_{p-1+j}}\widehat{b})(\boldsymbol{\theta}_j,\boldsymbol{\xi}+\boldsymbol{\phi}_j)(\nabla^{\nu_{p-1+m+j}}\widehat{\tilde{\chi}_{\boldsymbol{\theta}'_j}})(\boldsymbol{\xi}+\boldsymbol{\phi}'_j).\end{aligned} \tag{9.14}$$

2) For any $s = 1,\dots,\ell-1$ and any $k_1,\dots,k_p$ ($p \geqslant 1$) such that $k_1+\dots+k_p = s$, $\widehat{\mathrm{ad}\big(\mathrm{Op}(b);\Psi_{k_1},\dots,\Psi_{k_p}\big)}(\boldsymbol{\theta},\boldsymbol{\xi}) = 0$ for $\boldsymbol{\theta} \notin \boldsymbol{\Theta}_{s+1}$. Otherwise,

$$\begin{aligned}&\widehat{\mathrm{ad}\big(\mathrm{Op}(b);\Psi_{k_1},\dots,\Psi_{k_p}\big)}(\boldsymbol{\theta},\boldsymbol{\xi})\\ &= \sum_{\substack{\boldsymbol{\theta}_j,\boldsymbol{\theta}_{s+1}\in\boldsymbol{\Theta}\\ \boldsymbol{\phi}_j,\boldsymbol{\phi}_{s+1},\boldsymbol{\phi}'_j\in\boldsymbol{\Theta}_{s+1}\\ \boldsymbol{\theta}'_j\in\boldsymbol{\Theta}'_{s+1}\\ 1\leqslant j\leqslant s}} \sum_{p=1}^{s} \sum_{\substack{\boldsymbol{\theta}''_q,\boldsymbol{\phi}''_q\in\boldsymbol{\Theta}'_{s+1}\\ 1\leqslant q\leqslant p-1}} \sum_{\substack{\nu_1,\dots,\nu_{2s+p}\geqslant 0\\ \sum\nu_i=s}}\\ &\quad\prod_{q=1}^{p-1}(\nabla^{\nu_q}\widehat{e_{\boldsymbol{\theta}''_q}\varphi_{\boldsymbol{\theta}''_q}})(\boldsymbol{\xi}+\boldsymbol{\phi}''_q)(\widehat{\nabla^{\nu_p}b})(\boldsymbol{\theta}_{s+1},\boldsymbol{\xi}+\boldsymbol{\phi}_{s+1})\\ &\quad\times\prod_{j=1}^{s}(\widehat{\nabla^{\nu_{p+j}}b})(\boldsymbol{\theta}_j,\boldsymbol{\xi}+\boldsymbol{\phi}_j)(\nabla^{\nu_{p+s+j}}\widehat{\tilde{\chi}_{\boldsymbol{\theta}'_j}})(\boldsymbol{\xi}+\boldsymbol{\phi}'_j).\end{aligned} \tag{9.15}$$

3) For any $s=2,\dots,\ell$ and any $k_1,\dots,k_p$ $(p\geqslant 2)$ such that $k_1+\dots+k_p=s$, $\widehat{\mathrm{ad}(H_0;\Psi_{k_1},\dots,\Psi_{k_p})}(\boldsymbol\theta,\boldsymbol\xi)=0$ for $\boldsymbol\theta\notin\boldsymbol\Theta_s$. Otherwise,

$$\begin{aligned}
&\widehat{\mathrm{ad}(H_0;\Psi_{k_1},\dots,\Psi_{k_p})}(\boldsymbol\theta,\boldsymbol\xi)\\
&=\sum_{\substack{\boldsymbol\theta_j,\boldsymbol\theta_s\in\boldsymbol\Theta\\ \boldsymbol\phi_j,\boldsymbol\phi_s,\boldsymbol\phi'_j\in\boldsymbol\Theta_s\\ \boldsymbol\theta'_j\in\boldsymbol\Theta'_s\\ 1\leqslant j\leqslant s-1}}\sum_{p=1}^{s}\sum_{\substack{\boldsymbol\theta''_q,\boldsymbol\phi''_q\in\boldsymbol\Theta_s\\ 1\leqslant q\leqslant p-1}}\sum_{\substack{\nu_1,\dots,\nu_{2s+p-2}\geqslant 0\\ \sum\nu_i=s-1}}\\
&\prod_{q=1}^{p-1}(\nabla^{\nu_q}\widehat{e_{\boldsymbol\theta''_q}\varphi_{\boldsymbol\theta''_q}})(\boldsymbol\xi+\boldsymbol\phi''_q)(\widehat{\nabla^{\nu_p}b})(\boldsymbol\theta_s,\boldsymbol\xi+\boldsymbol\phi_s)\\
&\times\prod_{j=1}^{s-1}(\widehat{\nabla^{\nu_{p+j}}b})(\boldsymbol\theta_j,\boldsymbol\xi+\boldsymbol\phi_j)(\nabla^{\nu_{p+s-1+j}}\widehat{\tilde\chi_{\boldsymbol\theta'_j}})(\boldsymbol\xi+\boldsymbol\phi'_j).
\end{aligned}\tag{9.16}$$

For $\ell=2$ assumptions 1)–3) can be easily checked. Indeed, by (9.8), (4.9) and (4.5),

$$\begin{aligned}
&\hat\psi_1(\boldsymbol\theta,\boldsymbol\xi)=i\hat b(\boldsymbol\theta,\boldsymbol\xi)\tilde\chi_{\boldsymbol\theta}(\boldsymbol\xi),\\
&\widehat{\mathrm{ad}\big(\mathrm{Op}(b);\Psi_1\big)}(\boldsymbol\theta,\boldsymbol\xi)=\sum_{\boldsymbol\chi\in\boldsymbol\Theta\cup(\boldsymbol\theta-\boldsymbol\Theta)}\big(\hat b(\boldsymbol\theta,\boldsymbol\xi)\hat b(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi+\boldsymbol\chi)(\nabla_{\boldsymbol\chi}\widehat{\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}})(\boldsymbol\xi)\\
&\quad+\hat b(\boldsymbol\theta,\boldsymbol\xi)\widehat{(\nabla_{\boldsymbol\chi}b)}(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi)\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}(\boldsymbol\xi)-(\widehat{\nabla_{\boldsymbol\theta-\boldsymbol\chi}b})(\boldsymbol\chi,\boldsymbol\xi)\hat b(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi)\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}(\boldsymbol\xi)\big),\\
&\widehat{\mathrm{ad}(H_0;\Psi_1,\Psi_1)}(\boldsymbol\theta,\boldsymbol\xi)\\
&=\sum_{\boldsymbol\chi\in\boldsymbol\Theta\cup(\boldsymbol\theta-\boldsymbol\Theta)}\big(\hat b(\boldsymbol\chi,\boldsymbol\xi+\boldsymbol\theta-\boldsymbol\chi)(\nabla_{\boldsymbol\theta-\boldsymbol\chi}\widehat{\varphi_{\boldsymbol\chi}e_{\boldsymbol\chi}})(\boldsymbol\xi)\hat b(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi)\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}(\boldsymbol\xi)\\
&\quad+(\widehat{\nabla_{\boldsymbol\theta-\boldsymbol\chi}b})(\boldsymbol\chi,\boldsymbol\xi)\varphi_{\boldsymbol\chi}(\boldsymbol\xi)e_{\boldsymbol\chi}(\boldsymbol\xi)\hat b(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi)\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}(\boldsymbol\xi)\\
&\quad-\varphi_{\boldsymbol\theta}(\boldsymbol\xi)e_{\boldsymbol\theta}(\boldsymbol\xi)\hat b(\boldsymbol\theta,\boldsymbol\xi)\hat b(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi+\boldsymbol\chi)(\nabla_{\boldsymbol\chi}\widehat{\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}})(\boldsymbol\xi)\\
&\quad-\varphi_{\boldsymbol\theta}(\boldsymbol\xi)e_{\boldsymbol\theta}(\boldsymbol\xi)\hat b(\boldsymbol\theta,\boldsymbol\xi)\widehat{(\nabla_{\boldsymbol\chi}b)}(\boldsymbol\theta-\boldsymbol\chi,\boldsymbol\xi)\tilde\chi_{\boldsymbol\theta-\boldsymbol\chi}(\boldsymbol\xi)\big).
\end{aligned}$$

Now, we complete the induction in several steps.

Step 1. First of all, notice that due to (9.3), (9.4), for any $m=2,\dots,\ell$ the symbol of B_m admits a representation of the form (9.15) with $s=m-1$, and symbol of T_m admits a representation of the form (9.16) with $s=m$. Then it follows from Lemma 9.1 and (9.6) that Ψ_ℓ admits a representation of the form (9.14).

Step 2. Proof of (9.15) with $s=\ell$. Let $k_1+\dots+k_p=\ell$. If $p\geqslant 2$. Then

$$\mathrm{ad}\big(\mathrm{Op}(b);\Psi_{k_1},\dots,\Psi_{k_p}\big)=\mathrm{ad}\Big(\mathrm{ad}\big(\mathrm{Op}(b);\Psi_{k_1},\dots,\Psi_{k_{p-1}}\big);\Psi_{k_p}\Big).$$

Since $k_1+\dots+k_{p-1}\leqslant\ell-1$ and $k_p\leqslant\ell-1$ we can apply (9.14) and (9.15). Combined with (4.9) it gives a representation of the form (9.15). If $p=1$ then $\mathrm{ad}\big(\mathrm{Op}(b);\Psi_\ell\big)$ satisfies (9.15) because of (4.9) and step 1.

Step 3. Proof of (9.16) with $s = \ell + 1$. Let $k_1 + \cdots + k_p = \ell + 1$, $p \geqslant 2$. If $p \geqslant 3$, then (cf. step 2)

$$\operatorname{ad}(H_0; \Psi_{k_1}, \ldots, \Psi_{k_p}) = \operatorname{ad}\big(\operatorname{ad}(H_0; \Psi_{k_1}, \ldots, \Psi_{k_{p-1}}); \Psi_{k_p}\big).$$

Since $k_1 + \cdots + k_{p-1} \leqslant \ell$, $p - 1 \geqslant 2$ and $k_p \leqslant \ell - 1$ we can apply (9.14) and (9.16). Together with (4.9) it gives a representation of the form (9.16). If $p = 2$ then (see (9.6))

$$\operatorname{ad}(H_0; \Psi_{k_1}, \Psi_{k_2}) = \operatorname{ad}\big(\operatorname{ad}(H_0; \Psi_{k_1}); \Psi_{k_2}\big) = -\operatorname{ad}(B_{k_1}^{\mathcal{NR}} + T_{k_1}^{\mathcal{NR}}; \Psi_{k_2}).$$

Since $k_1 \leqslant \ell$ and $k_2 \leqslant \ell$, the representation of the form (9.16) follows from (4.9) and step 1. (Formally exceptional case $k_1 = 1$, $k_2 = \ell$ can be treated separately in the same way using (9.5) instead of (9.6).)

Induction is complete.

Now, (9.15), (9.16) and (9.2), (9.3), (9.4) prove the lemma. □

10. Contribution from various resonant regions

Let us fix a subspace $\mathfrak{V} \in \mathcal{V}_m$, $m < d$, and a component Ξ_p of the resonant region $\Xi(\mathfrak{V})$. Our aim is to compute the contribution to the density of states from each component Ξ_p. Therefore, we define

$$\mathcal{A}_p^+(\rho) := \mathcal{A}^+(\rho) \cap \Xi_p \quad \text{and} \quad \mathcal{A}_p^-(\rho) := \mathcal{A}^-(\rho) \cap \Xi_p \tag{10.1}$$

and try to compute

$$\operatorname{vol} \mathcal{A}_p^+(\rho) - \operatorname{vol} \mathcal{A}_p^-(\rho). \tag{10.2}$$

Since formulas (6.10) and (6.4) obviously imply that

$$\operatorname{vol}(\mathcal{G}_\lambda) = \omega_d \rho^d + \sum_{m=0}^{d-1} \sum_{\mathfrak{V} \in \mathcal{V}_m} \sum_p \big(\operatorname{vol} \mathcal{A}_p^+(\rho) - \operatorname{vol} \mathcal{A}_p^-(\rho)\big), \tag{10.3}$$

Lemma 3.1 would be proved if we manage to compute (10.2) (or at least prove that this expression admits a complete asymptotic expansion in ρ).

Note that if $\boldsymbol{\xi} \in \Xi_p$, then we also have that $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \Xi_p$. We denote

$$H_2(\boldsymbol{\xi}) := H_2|_{\mathsf{H}_{\boldsymbol{\xi}}}, \quad \mathsf{H}_{\boldsymbol{\xi}} := \mathcal{P}\big(\boldsymbol{\Upsilon}(\boldsymbol{\xi})\big)\mathsf{B}_2(\mathbb{R}^d)$$

(recall that $\mathsf{H}_{\boldsymbol{\xi}}$ is an invariant subspace of H_2 acting in $\mathsf{B}_2(\mathbb{R}^d)$). Suppose now that two points $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ have the same coordinates $\mathbf{X}$ and $\boldsymbol{\Phi}$ and different coordinates r. Then $\boldsymbol{\xi} \in \Xi_p$ implies $\boldsymbol{\eta} \in \Xi_p$ and $\boldsymbol{\Upsilon}(\boldsymbol{\eta}) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}) + (\boldsymbol{\eta} - \boldsymbol{\xi})$. This shows that two spaces $\mathsf{H}_{\boldsymbol{\xi}}$ and $\mathsf{H}_{\boldsymbol{\eta}}$ have the same dimension and, moreover, there is a natural isometry $F_{\boldsymbol{\xi},\boldsymbol{\eta}} : \mathsf{H}_{\boldsymbol{\xi}} \to \mathsf{H}_{\boldsymbol{\eta}}$ given by $F : \mathsf{e}_{\boldsymbol{\nu}} \mapsto \mathsf{e}_{\boldsymbol{\nu}+(\boldsymbol{\eta}-\boldsymbol{\xi})}$, $\boldsymbol{\nu} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$. This isometry allows us to 'compare' operators acting in $\mathsf{H}_{\boldsymbol{\xi}}$ and $\mathsf{H}_{\boldsymbol{\eta}}$. Thus, abusing slightly our notation, we can assume that $H_2(\boldsymbol{\xi})$ and $H_2(\boldsymbol{\eta})$ act in the same (finite dimensional) Hilbert space $\mathsf{H}(\mathbf{X}, \boldsymbol{\Phi})$. We will fix the values $(\mathbf{X}, \boldsymbol{\Phi})$ and study how these operators depend on r. Thus, we denote by $H_2(r) = H_2(r; \mathbf{X}, \boldsymbol{\Phi})$ the operator $H_2(\boldsymbol{\xi})$ with $\boldsymbol{\xi} = (\mathbf{X}, r, \boldsymbol{\Phi})$, acting in $\mathsf{H}(\mathbf{X}, \boldsymbol{\Phi})$.

Let $W_{\tilde{k}}(r)$ be the operator in $\mathsf{H}(\mathbf{X},\boldsymbol{\Phi})$ with the symbol $w_{\tilde{k}}\big(\mathbf{x},\boldsymbol{\xi}(\mathbf{X},r,\boldsymbol{\Phi})\big)$. According to formula (7.3), for any $s\leqslant\tilde{k}-1$ and $\boldsymbol{\theta}\in\boldsymbol{\Theta}_{s+1}$

$$|\boldsymbol{\xi}+\boldsymbol{\phi}|^2=r^2+2r|\mathbf{a}|\sum_{q=1}^{K+1}a_{K+1\,q}\sin\Phi_q+2\langle\boldsymbol{\xi},\boldsymbol{\phi}\rangle+|\mathbf{X}|^2+|\mathbf{a}|^2+|\boldsymbol{\phi}|^2. \tag{10.4}$$

This, together with (2.4), (2.10) and (2.11), implies that for $|\boldsymbol{\xi}+\boldsymbol{\phi}|>C_0$ the coefficients $\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\phi})$ can be represented as the absolutely convergent series

$$\begin{aligned}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\phi})=\sum_{\iota\in\tilde{J}}\sum_{l=0}^{\infty}\sum_{\substack{n_1,\dots,n_{K+1}\geqslant0\\ n_1+\dots+n_{K+1}\leqslant l\\ j_1,\dots,j_d\geqslant0\\ j_1+\dots+j_d\leqslant l}}\\ C^{\iota\,j_1\cdots j_d}_{l\,n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta})r^{\iota-l}\phi_1^{j_1}\cdots\phi_d^{j_d}\prod_{a=1}^{K+1}(\sin\Phi_a)^{n_a},\end{aligned} \tag{10.5}$$

where the coefficients satisfy

$$\big|C^{\iota\,j_1\cdots j_d}_{l\,n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta})\big|\lesssim\rho_n^{(l-j_1-\dots-j_d)(\alpha_{m+1}+0+)}$$

In the next lemma, to facilitate the expansion of the RHS of (9.12) in a suitable form, we transform the denominator of $\tilde{\chi}_{\boldsymbol{\theta}'}$ (recall (9.7)).

In the subsequent calculations we will use the generalized binomial coefficeints:

$$\binom{p}{j}:=\begin{cases}1, & j=0;\\ \dfrac{1}{j!}\displaystyle\prod_{k=0}^{j-1}(p-k), & j\in\mathbb{N}.\end{cases} \tag{10.6}$$

Lemma 10.1. *For $s\leqslant\tilde{k}-1$, $\boldsymbol{\phi}'\in\boldsymbol{\Theta}_{s+1}$, $\boldsymbol{\theta}'\in\boldsymbol{\Theta}'_{s+1}$, and $\boldsymbol{\xi}$ in the support of $e_{\boldsymbol{\theta}'}\varphi_{\boldsymbol{\theta}'}$ let*

$$\begin{aligned}D:=\frac{1}{w}\sum_{j=2}^{\infty}\binom{w}{j}r^{2-2j}\sum_{k=0}^{j-1}\binom{j}{k}\\ \times\Big(2r|\mathbf{a}|\sum_{q=1}^{K+1}a_{K+1\,q}\sin\Phi_q+2\langle\boldsymbol{\xi},\boldsymbol{\phi}'\rangle+|\mathbf{X}|^2+|\mathbf{a}|^2+|\boldsymbol{\phi}'|^2\Big)^k\\ \times\big(2\langle\boldsymbol{\xi},\boldsymbol{\theta}'\rangle+2\langle\boldsymbol{\phi}',\boldsymbol{\theta}'\rangle+|\boldsymbol{\theta}'|^2\big)^{j-k-1}.\end{aligned}$$

Then $|D|\lesssim\rho_n^{-1+\alpha_{m+1}+0+}$ and

$$\begin{aligned}&\big(|\boldsymbol{\xi}+\boldsymbol{\phi}'+\boldsymbol{\theta}'|^{2w}-|\boldsymbol{\xi}+\boldsymbol{\phi}'|^{2w}\big)^{-1}\\ &=w^{-1}r^{2-2w}\big(2\langle\boldsymbol{\xi},\boldsymbol{\theta}'\rangle+2\langle\boldsymbol{\phi}',\boldsymbol{\theta}'\rangle+|\boldsymbol{\theta}'|^2\big)^{-1}\sum_{a=0}^{\infty}(-D)^a.\end{aligned} \tag{10.7}$$

Proof. We introduce a shorthand

$$N := 2r|\mathbf{a}| \sum_{q=1}^{K+1} a_{K+1\, q} \sin \Phi_q + 2\langle \boldsymbol{\xi}, \boldsymbol{\phi}' \rangle + |\mathbf{X}|^2 + |\mathbf{a}|^2 + |\boldsymbol{\phi}'|^2.$$

Then by (generalized) binomial formula and (10.4) we obtain

$$\begin{aligned} &|\boldsymbol{\xi} + \boldsymbol{\phi}' + \boldsymbol{\theta}'|^{2w} - |\boldsymbol{\xi} + \boldsymbol{\phi}'|^{2w} \\ &= \big(|\boldsymbol{\xi}|^2 + 2\langle \boldsymbol{\xi}, \boldsymbol{\phi}' + \boldsymbol{\theta}' \rangle + |\boldsymbol{\phi}' + \boldsymbol{\theta}'|^2\big)^w - \big(|\boldsymbol{\xi}|^2 + 2\langle \boldsymbol{\xi}, \boldsymbol{\phi}' \rangle + |\boldsymbol{\phi}'|^2\big)^w \\ &= \big(r^2 + N + 2\langle \boldsymbol{\xi}, \boldsymbol{\theta}' \rangle + 2\langle \boldsymbol{\phi}', \boldsymbol{\theta}' \rangle + |\boldsymbol{\theta}'|^2\big)^w - (r^2 + N)^w \\ &= r^{2w} \sum_{j=1}^{\infty} \binom{w}{j} r^{-2j} \Big(\big(N + 2\langle \boldsymbol{\xi}, \boldsymbol{\theta}' \rangle + 2\langle \boldsymbol{\phi}', \boldsymbol{\theta}' \rangle + |\boldsymbol{\theta}'|^2\big)^j - N^j \Big) \\ &= w r^{2w-2} \big(2\langle \boldsymbol{\xi}, \boldsymbol{\theta}' \rangle + 2\langle \boldsymbol{\phi}', \boldsymbol{\theta}' \rangle + |\boldsymbol{\theta}'|^2\big)(1 + D). \end{aligned} \tag{10.8}$$

The estimate on $|D|$ follows from estimates (2.19) and (7.1), and Lemmas 7.3 and 5.1. Now (10.7) follows from (10.8). □

As we have seen from the previous sections, the symbol of the operator H_2 satisfies

$$h_2(\mathbf{x}, \boldsymbol{\xi}) = |\boldsymbol{\xi}|^{2w} + w_{\tilde{k}}(\mathbf{x}, \boldsymbol{\xi}) = \big(r^2 + 2r\langle \mathbf{a}, \boldsymbol{\Phi} \rangle + |\mathbf{a}|^2 + |\mathbf{X}|^2\big)^w + w_{\tilde{k}}(\mathbf{x}, \boldsymbol{\xi}), \tag{10.9}$$

where $w_{\tilde{k}}$ are given by (9.11) and (9.12).

Remark 10.2. In this section we assume that $\boldsymbol{\xi} \in \mathcal{A}$, so by (6.3) we have $2\rho_n/3 \leqslant |\boldsymbol{\xi}| \leqslant 6\rho_n$, and by Remark 8.1 all functions $e_{\boldsymbol{\theta}}(\boldsymbol{\xi} + \cdot)$ from (9.11) and (9.12) are equal to 1. Note that if $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}}$, $\boldsymbol{\phi} \in \boldsymbol{\Theta}_{\tilde{k}}$, and $\boldsymbol{\theta} \notin \mathfrak{V}$, then (see Lemma 7.5 and (8.3)) $\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi} + \boldsymbol{\phi}) = 1$. This means that all cut-off functions from (9.11) and (9.12) are equal to 1 unless $\boldsymbol{\theta} \in \mathfrak{V}$. If, on the other hand, $\boldsymbol{\theta} \in \mathfrak{V}$, then $\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi} + \boldsymbol{\phi})$ depends only on the projection $\boldsymbol{\xi}_{\mathfrak{V}}$ and thus is a function only of the coordinates $\mathbf{X}$.

By Proposition 4.1, (9.11), Lemma 9.4, formulas (10.5) and (10.7), Lemma 7.5, and Remark 10.2, for $r \asymp \rho_n$

$$\Big\| \frac{d^l}{dr^l} W_{\tilde{k}}(r) \Big\| \lesssim \rho_n^{\varkappa - l + 0+}, \qquad l \geqslant 0. \tag{10.10}$$

This, together with (10.9), implies

Lemma 10.3. *The operator $H_2(r)$ is monotonically increasing in r; in particular, all its eigenvalues $\lambda_j\big(H_2(r)\big)$ are increasing in r.*

Thus the function $g\big(\boldsymbol{\xi}(\mathbf{X}, r, \boldsymbol{\Phi})\big)$ (defined in Section 6) is an increasing function of r if we fix the other coordinates of $\boldsymbol{\xi}$, so the equation

$$g(\boldsymbol{\xi}) = \rho^{2w}$$

has a unique solution for fixed values of $\mathbf{X}$ and $\boldsymbol{\Phi}$; we denote the r-coordinate of this solution by $\tau = \tau(\rho) = \tau(\rho; \mathbf{X}, \boldsymbol{\Phi})$, so that

$$g\big(\boldsymbol{\xi}(\mathbf{X}, \tau, \boldsymbol{\Phi})\big) = \rho^{2w}. \tag{10.11}$$

By $\tau_0 = \tau_0(\rho) = \tau_0(\rho; \mathbf{X}, \mathbf{\Phi})$ we denote the value of τ for $(-\Delta)^w$, i.e. τ_0 is a unique solution of the equation

$$\big|\boldsymbol{\xi}(\mathbf{X}, \tau_0, \mathbf{\Phi})\big| = \rho.$$

Obviously, we can write down a precise analytic expression for τ_0 (and we have done this in [11] in the two-dimensional case) and show that it allows an expansion in powers of ρ and $\ln\rho$, but we will not need it. The definition (10.1) of the sets $\mathcal{A}_p^{\pm}(\rho)$ implies that the intersection

$$\mathcal{A}_p^{+}(\rho) \cap \big\{\boldsymbol{\xi}(\mathbf{X}, r, \mathbf{\Phi}),\ r \in \mathbb{R}_+\big\}$$

consists of points with r-coordinate belonging to the interval $\big[\tau_0(\rho), \tau(\rho)\big]$ (where we assume the interval to be empty if $\tau_0 > \tau$). Similarly, the intersection

$$\mathcal{A}_p^{-}(\rho) \cap \big\{\boldsymbol{\xi}(\mathbf{X}, r, \mathbf{\Phi}),\ r \in \mathbb{R}_+\big\}$$

consists of points with r-coordinate belonging to the interval $\big[\tau(\rho), \tau_0(\rho)\big]$. Therefore,

$$\mathcal{A}_p^{+}(\rho) = \Big\{\boldsymbol{\xi} = \boldsymbol{\xi}(\mathbf{X}, r, \mathbf{\Phi}), \mathbf{X} \in \Omega(\mathfrak{V}), \mathbf{\Phi} \in \mathcal{M}_p, r \in \big[\tau_0(\rho; \mathbf{X}, \mathbf{\Phi}), \tau(\rho; \mathbf{X}, \mathbf{\Phi})\big]\Big\}$$

and

$$\mathcal{A}_p^{-}(\rho) = \Big\{\boldsymbol{\xi} = \boldsymbol{\xi}(\mathbf{X}, r, \mathbf{\Phi}), \mathbf{X} \in \Omega(\mathfrak{V}), \mathbf{\Phi} \in \mathcal{M}_p, r \in \big[\tau(\rho; \mathbf{X}, \mathbf{\Phi}), \tau_0(\rho; \mathbf{X}, \mathbf{\Phi})\big]\Big\}.$$

This implies that (recall that $K = d - m - 1$)

$$\begin{aligned} \operatorname{vol}\mathcal{A}_p^{+}(\rho) - \operatorname{vol}\mathcal{A}_p^{-}(\rho) &= \int_{\Omega(\mathfrak{V})} d\mathbf{X} \int_{\mathcal{M}_p} d\mathbf{\Phi} \int_{\tau_0(\rho; \mathbf{X}, \mathbf{\Phi})}^{\tau(\rho; \mathbf{X}, \mathbf{\Phi})} r^K dr \\ &= (K+1)^{-1} \int_{\mathcal{M}_p} d\mathbf{\Phi} \int_{\Omega(\mathfrak{V})} d\mathbf{X} \big(\tau(\rho; \mathbf{X}, \mathbf{\Phi})^{K+1} - \tau_0(\rho; \mathbf{X}, \mathbf{\Phi})^{K+1}\big). \end{aligned} \tag{10.12}$$

Remark 10.4. Note that in the case $K = 0$ the simplex $\mathcal{M}_p$ is degenerate and there is no integration in $d\mathbf{\Phi}$.

Obviously, it is enough to compute the part of (10.12) containing τ, since the second part (containing τ_0) can be computed analogously. We start by considering

$$\int_{\Omega(\mathfrak{V})} \tau(\rho; \mathbf{X}, \mathbf{\Phi})^{K+1} d\mathbf{X}. \tag{10.13}$$

First of all, we notice that if $\boldsymbol{\xi}, \boldsymbol{\eta} \in \Xi(\mathfrak{V})$ are resonant congruent points then, according to Lemma 5.7, all vectors $\boldsymbol{\theta}_j$ from Definition 5.8 of equivalence belong to $\mathfrak{V}$. This naturally leads to the definition of equivalence for projections $\boldsymbol{\xi}_{\mathfrak{V}}$ and $\boldsymbol{\eta}_{\mathfrak{V}}$. Namely, we say that two points $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ from $\Omega(\mathfrak{V})$ are $\mathfrak{V}$-equivalent (and write $\boldsymbol{\nu} \leftrightarrow_{\mathfrak{V}} \boldsymbol{\mu}$) if $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ are equivalent in the sense of Definition 5.8 with an additional requirement that all $\boldsymbol{\theta}_j \in \mathfrak{V}$. Then $\boldsymbol{\xi} \leftrightarrow \boldsymbol{\eta}$ implies $\boldsymbol{\xi}_{\mathfrak{V}} \leftrightarrow_{\mathfrak{V}} \boldsymbol{\eta}_{\mathfrak{V}}$. For $\boldsymbol{\nu} \in \Omega(\mathfrak{V})$ we denote by $\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})$ the class of equivalence of $\boldsymbol{\nu}$ generated by $\leftrightarrow_{\mathfrak{V}}$. Then $\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\xi}_{\mathfrak{V}})$ is a projection of $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$ to $\mathfrak{V}$ and is, therefore, finite.

Since $\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})$ is a finite set for each $\boldsymbol{\nu} \in \Omega(\mathfrak{V})$, we can re-write (10.13) as

$$\begin{aligned}&\int_{\Omega(\mathfrak{V})} \tau(\rho;\mathbf{X},\boldsymbol{\Phi})^{K+1} d\mathbf{X} \\ &= \int_{\Omega(\mathfrak{V})} \big(\operatorname{card} \boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})\big)^{-1} \sum_{\mathbf{X}\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \tau(\rho;\mathbf{X},\boldsymbol{\Phi})^{K+1} d\boldsymbol{\nu}\end{aligned} \tag{10.14}$$

and try to compute

$$\sum_{\mathbf{X}\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \tau(\rho;\mathbf{X},\boldsymbol{\Phi})^{K+1}.$$

Remark 10.2, together with equations (10.9), (9.11), and (9.12), shows that $H_2(r)$ depends on r analytically, so we can and will consider the family $H_2(z)$ with complex values of the parameter z with $\Re e\ z \asymp \rho$. Likewise, we analytically continue the function $\boldsymbol{\xi}(\mathbf{X}, r, \boldsymbol{\Phi})$ to

$$\boldsymbol{\xi}(\mathbf{X}, z, \boldsymbol{\Phi}) := \mathbf{X} + \mathbf{a} + z\boldsymbol{\Phi}. \tag{10.15}$$

We also introduce the analytic continuation $|\cdot|_{\mathbb{C}}$ of the modulus of vectors, so that

$$|\boldsymbol{\xi}|_{\mathbb{C}}^2 := z^2 + 2z\langle \mathbf{a}, \boldsymbol{\Phi}\rangle + |\mathbf{a}|^2 + |\mathbf{X}|^2. \tag{10.16}$$

Formulas (9.11) and (9.12) give matrix elements of $H_2(z)$ in an orthonormal basis even for complex z.

We choose a contour

$$\gamma := \Big\{ z \in \mathbb{C} :\ |z-\rho| = t\rho_n := \Big(8 \max\big\{(2w-2)/3, 1\big\}\Big)^{-1} \rho_n \Big\} \tag{10.17}$$

to be a circle in the complex plane going in the positive direction.

Estimates (10.10) remain valid after the analytic continuation: for all z inside and on γ

$$\Big\| \frac{d^l}{dz^l} W_{\tilde{k}}(z) \Big\| \lesssim \rho_n^{\varkappa - l + 0+}, \qquad l \geqslant 0. \tag{10.18}$$

Lemma 10.5. *For $\rho \in I_n = [\rho_n, 4\rho_n]$ all $\tau(\rho;\mathbf{X},\boldsymbol{\Phi})$ lie inside γ. These are the only zeros of the function* $\det\big(H_2(z) - \rho^{2w} I\big)$ *inside the contour.*

Proof. Let $r := \Re e\ z$, $y := \Im m\ z$. For $y = 0$ the operator $H_2(r)$ is self-adjoint. Thus it has $\operatorname{card} \boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})$ real eigenvalues.

Now for $r \geqslant \rho + t\rho_n \geqslant (1 + t/4)\rho$ relations (10.9), (7.1), Lemma 5.1(ii), and (10.18) imply

$$H_2(r) \geqslant \Big(\big((1+t/4)\rho\big)^{2w} \big(1 - O(\rho^{\alpha_{m+1}-1+0+})\big) - O(\rho^{\varkappa+0+}) \Big) I.$$

Thus by (2.20) and (2.21) for big ρ no eigenvalue of $H_2(r)$ can coincide with ρ^{2w}.

Likewise for $r \leqslant \rho - t\rho_n \leqslant (1 - t/4)\rho$ for big ρ we have

$$H_2(r) \leqslant \Big(\big((1-t/4)\rho\big)^{2w} \big(1 + O(\rho^{\alpha_{m+1}-1+0+})\big) + O(\rho^{\varkappa+0+}) \Big) I,$$

and no eigenvalue of $H_2(r)$ can coincide with ρ^{2w}. This implies that all the eigenvalues of $H_2(r)$ lie in the real interval $(\rho - t\rho_n, \rho + t\rho_n)$. By (10.11) and Lemma 10.3 these eigenvalues coincide with $\{\tau(\rho; \mathbf{X}, \mathbf{\Phi}) : \mathbf{X} \in \mathbf{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})\}$.

It remains to show that $H_2(r + iy)$ is invertible for any nonzero y such that $r + iy$ is inside or on γ. Relation (10.15), Lemma 5.1(ii), definition (5.1), and bound (7.1) imply that inside and on the contour

$$\boldsymbol{\xi} = (r + iy)\big(1 + O(\rho_n^{-1+\alpha_{m+1}+0+})\big)$$

and

$$\arg|\boldsymbol{\xi}|_{\mathbb{C}} \leqslant \big(1 + o(1)\big)\arcsin(t\rho_n/\rho) \leqslant \big(1 + o(1)\big)\arcsin t \leqslant t\big(1 + o(1)\big).$$

Hence

$$\big||\boldsymbol{\xi}|_{\mathbb{C}}^{2w}\big| = \big||\boldsymbol{\xi}|_{\mathbb{C}}^{2}\big|^{w} \asymp \rho^{2w} \quad \text{and} \quad \arg|\boldsymbol{\xi}|_{\mathbb{C}}^{2w} = w\arcsin\frac{2y\big(r + \langle \mathbf{a}, \mathbf{\Phi}\rangle\big)}{\big||\boldsymbol{\xi}|_{\mathbb{C}}^{2}\big|} \asymp y\rho^{-1},$$

which implies that

$$\Big|\operatorname{Im}\big(|\boldsymbol{\xi}|_{\mathbb{C}}^{2w}\big)\Big| \gtrsim |y|\rho^{2w-1}. \tag{10.19}$$

Now for any $\Psi \in \mathsf{H}(\mathbf{X}, \mathbf{\Phi})$ with $\|\Psi\| = 1$ we have by (10.19) and (10.18)

$$\begin{aligned}\Big\|\big(H_2(z) - \rho^{2w}I\big)\Psi\Big\| &\geqslant \Big|\operatorname{Im}\langle\big(H_2(z) - \rho^{2w}I\big)\Psi, \Psi\rangle\Big| \\ &\geqslant \Big|\operatorname{Im}\big(|\boldsymbol{\xi}|_{\mathbb{C}}^{2w}\big)\Big| - |y|\sup_{t\in[0,y]}\big\|W'(r + it)\big\| \gtrsim |y|\rho^{2w-1},\end{aligned}$$

where we have used that for $y = 0$ the quadratic form of $W(z)$ is real-valued. So the kernel of $H_2(r + iy) - \rho^{2w}$ is trivial for $y \neq 0$. □

Lemma 10.6. *For $z \in \gamma$ and $l \in \mathbb{N}$*

$$(z^{2w} - \rho^{2w})^{-l} = \rho^{-2wl}\sum_{j=0}^{\infty} A_{l\,j}\Big(\frac{z-\rho}{\rho}\Big)^{j-l}, \tag{10.20}$$

where $A_{l\,0} := (2w)^{-l}$ and for $j > 0$

$$A_{l\,j} := \frac{1}{(2w)^l}\sum_{p=1}^{j}\frac{1}{(2w)^p}\binom{-l}{p}\sum_{\substack{q_1,\dots,q_p\geqslant 1\\ q_1+\dots+q_p=j}}\binom{2w}{q_1+1}\binom{2w}{q_2+1}\cdots\binom{2w}{q_p+1}.$$

The series in (10.20) *converges absolutely.*

Proof. A striaghtforward calculation gives

$$\begin{aligned}
&(z^{2w}-\rho^{2w})^{-l}\\
&=\frac{1}{\rho^{2wl}}\Big(\Big(1+\frac{z-\rho}{\rho}\Big)^{2w}-1\Big)^{-l}\\
&=\frac{1}{\rho^{2wl}}\Big(\sum_{q=1}^{\infty}\binom{2w}{q}\Big(\frac{z-\rho}{\rho}\Big)^{q}\Big)^{-l}\\
&=\frac{\rho^{-2wl}}{(2w)^l}\Big(\frac{z-\rho}{\rho}\Big)^{-l}\Big(1+\frac{1}{2w}\sum_{q=1}^{\infty}\binom{2w}{q+1}\Big(\frac{z-\rho}{\rho}\Big)^{q}\Big)^{-l}.
\end{aligned}\tag{10.21}$$

If $2w\in\mathbb{N}$, then the series on the right hand side is finite. Otherwise, by (10.17) and (10.6), for $z\in\gamma$ the ratio of absolute values of any two sequential terms of the series satisfies

$$\Big|\frac{z-\rho}{\rho}\binom{2w}{q+2}\binom{2w}{q+1}^{-1}\Big|=\Big|\frac{z-\rho}{\rho}\Big|\frac{|2w-q-1|}{q+2}\leqslant\frac{1}{8},\qquad q\geqslant 1.$$

So, again by (10.17) and (10.6), we have

$$\Big|\sum_{q=1}^{\infty}\binom{2w}{q+1}\Big(\frac{z-\rho}{\rho}\Big)^{q}\Big|<\Big|\binom{2w}{2}\Big|\frac{|z-\rho|}{\rho}\sum_{q=0}^{\infty}\frac{1}{8^q}\leqslant\frac{4w}{7}.$$

Thus we can decompose the expression on the right hand side of (10.21) into an absolutely converging series obtaining

$$\begin{aligned}
(z^{2w}-\rho^{2w})^{-l}&=\frac{\rho^{-2wl}}{(2w)^l}\Big(\frac{z-\rho}{\rho}\Big)^{-l}\sum_{p=0}^{\infty}\binom{-l}{p}\frac{1}{(2w)^p}\Big(\sum_{q=1}^{\infty}\binom{2w}{q+1}\Big(\frac{z-\rho}{\rho}\Big)^{q}\Big)^{p}\\
&=\frac{\rho^{-2wl}}{(2w)^l}\Big(\frac{z-\rho}{\rho}\Big)^{-l}\\
&\times\Big(1+\sum_{j=1}^{\infty}\Big(\frac{z-\rho}{\rho}\Big)^{j}\sum_{p=1}^{j}\frac{1}{(2w)^p}\binom{-l}{p}\sum_{\substack{q_1,\dots,q_p\geqslant 1\\ q_1+\dots+q_p=j}}\binom{2w}{q_1+1}\cdots\binom{2w}{q_p+1}\Big),
\end{aligned}$$

which finishes the proof. □

Let $S(z):=H_2(z)-z^{2w}I$ in $\mathsf{H}(\mathbf{X},\mathbf{\Phi})$. Then by (10.9) on γ the symbol of $S(z)$ admits the representaion

$$s(z)=\sum_{v=1}^{\infty}\binom{w}{v}z^{2w-v}\Big(2\langle\mathbf{a},\mathbf{\Phi}\rangle+z^{-1}\big(|\mathbf{a}|^2+|\mathbf{X}|^2\big)\Big)^{v}+w_{\tilde{k}}(z).\tag{10.22}$$

Relations (10.22), (10.18), (7.1), Lemma 5.1(ii), and (2.20) imply that everywhere inside and on γ

$$\Big\|\frac{d^l}{dz^l}S(z)\Big\|\lesssim\rho_n^{2w-1+\alpha_{m+1}-l+0+},\quad l\geqslant 0.\tag{10.23}$$

A version of the Jacobi's formula states that for any differentiable invertible matrix-valued function $F(z)$ we have

$$\operatorname{tr}\left[F'(z)F^{-1}(z)\right] = \Big(\det\left[F(z)\right]\Big)'\Big(\det\left[F(z)\right]\Big)^{-1}$$

(it can be proved, for example, using the expansion of the determinant along rows and the induction in the size of F).

Then by Lemma 10.6 and the residue theorem

$$\begin{aligned}
&\sum_{\mathbf{X}\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \tau(\rho;\mathbf{X},\mathbf{\Phi})^{K+1}\\
&= \frac{1}{2\pi i}\oint_\gamma z^{K+1}\Big(\det\left[H_2(z)-\rho^{2w}I\right]\Big)'\Big(\det\left[H_2(z)-\rho^{2w}I\right]\Big)^{-1}dz\\
&= \frac{1}{2\pi i}\oint_\gamma \operatorname{tr}\left[z^{K+1}H_2'(z)\big(H_2(z)-\rho^{2w}I\big)^{-1}\right]dz\\
&= \frac{1}{2\pi i}\oint_\gamma \operatorname{tr}\Big[\big(2wz^{2w+K}I+z^{K+1}S'(z)\big)\\
&\qquad\qquad\times\sum_{l=0}^{\infty}(-1)^l S^l(z)(z^{2w}-\rho^{2w})^{-1-l}\Big]dz \qquad (10.24)\\
&= \frac{1}{2\pi i}\oint_\gamma \operatorname{tr}\Big[\big(2wz^{2w+K}I+z^{K+1}S'(z)\big)\sum_{l=-\infty}^{\infty}(z-\rho)^{-1-l}\\
&\qquad\qquad\times\sum_{j=0}^{\infty}(-1)^{l+j}A_{1+l+j\,j}\,\rho^{1+l-2w(1+l+j)}S^{l+j}(z)\Big]dz\\
&= \sum_{l=0}^{\infty}\frac{1}{l!}\operatorname{tr}\frac{d^l}{dr^l}\Big[\big(2wr^{2w+K}I+r^{K+1}S'(r)\big)\\
&\qquad\qquad\times\sum_{j=0}^{\infty}(-1)^{l+j}A_{1+l+j\,j}\,\rho^{1+l-2w(1+l+j)}S^{l+j}(r)\Big]\Big|_{r=\rho}.
\end{aligned}$$

We can restrict the summation on the RHS of (10.24) to

$$l+j\leqslant l_0 := \big(M+K+d+1+(d-1)\alpha_{d-1}-2w\big)/(1-\alpha_{m+1}).$$

Indeed, using the trivial fact that for any linear operator A in the finite dimensional Hilbert space spanned by $\mathsf{e}_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta}\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})$

$$|\operatorname{tr}A|\leqslant\|A\|\operatorname{card}\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu}),$$

estimate (10.23), and relation (2.20) we can see that the sum of the terms in (10.24) with $l+j>l_0$ contributes only to the order $O(\rho_n^{-M+2w-d})$ in (10.14), and thus after integration in $\mathbf{\Phi}$ the corresponding term can be included into the remainder $R_{\tilde{k}}$ of Section 6.

Formula (10.12) shows that in order to compute the contribution to the density of states from $\boldsymbol{\Xi}(\mathfrak{V})_p$, we need to integrate the RHS of (10.24) against

$d\boldsymbol{\nu}$ and $d\boldsymbol{\Phi}$. We are going to integrate against $d\boldsymbol{\Phi}$ first:

$$\begin{aligned}
&\int_{\mathcal{M}_p} d\boldsymbol{\Phi} \int_{\Omega(\mathfrak{V})} \big(\operatorname{card} \boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})\big)^{-1} \sum_{\mathbf{X}\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \tau(\rho;\mathbf{X},\boldsymbol{\Phi})^{K+1} \\
&= \int_{\Omega(\mathfrak{V})} \frac{d\boldsymbol{\nu}}{\operatorname{card}\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \sum_{l=0}^{l_0}\sum_{j=0}^{l_0-l} \frac{(-1)^{l+j}}{l!} A_{1+l+j\,j}\rho^{1+l-2w(1+l+j)} \\
&\qquad \times \operatorname{tr} \frac{d^l}{dr^l}\Big[\int_{\mathcal{M}_p} d\boldsymbol{\Phi}\big(2wr^{2w+K}I + r^{K+1}S'(r)\big)S^{l+j}(r)\Big]\Big|_{r=\rho} \\
&\quad + O(\rho_n^{-M+2w-d}) \\
&= O(\rho_n^{-M+2w-d}) \\
&\quad + \int_{\Omega(\mathfrak{V})} \frac{d\boldsymbol{\nu}}{\operatorname{card}\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \sum_{l=0}^{l_0}\sum_{j=0}^{l_0-l} \frac{(-1)^{l+j}}{l!} A_{1+l+j\,j}\rho^{1+l-2w(1+l+j)} \\
&\qquad \times \operatorname{tr}\Bigg[\frac{d^l}{dr^l}\Big(2wr^{2w+K}\int_{\mathcal{M}_p} S^{l+j}(r)d\boldsymbol{\Phi} \\
&\qquad\qquad - \frac{(K+1)r^K}{l+j+1}\int_{\mathcal{M}_p} S^{l+j+1}(r)d\boldsymbol{\Phi}\Big) \\
&\qquad\qquad + \frac{d^{l+1}}{dr^{l+1}}\Big(\frac{r^{K+1}}{l+j+1}\int_{\mathcal{M}_p} S^{l+j+1}(r)d\boldsymbol{\Phi}\Big)\Bigg]\Bigg|_{r=\rho}.
\end{aligned} \tag{10.25}$$

We will prove that the integrand of the exterior integral in (10.25) is a convergent series of products of powers of ρ and $\ln\rho$. The coefficients in front of all terms will be bounded functions of $\mathbf{X}$, so afterwards we will just integrate these coefficients to obtain the desired asymptotic expansion.

Let us discuss, how $S(r)$ depends on ρ, $\mathbf{X}$ and $\boldsymbol{\Phi}$. In order to do this, we first look again at (9.12). As follows from Remark 10.2, the product $e_{\boldsymbol{\theta}''_q}\varphi_{\boldsymbol{\theta}''_q}$ does not depend on r and $\boldsymbol{\Phi}$, and by (8.4)

$$\|\nabla^{\nu}\widehat{e_{\boldsymbol{\theta}''_q}\varphi_{\boldsymbol{\theta}''_q}}\|_{\mathsf{L}_\infty(\mathbb{R}^d)} \lesssim \rho_n^{-\nu\beta}. \tag{10.26}$$

For any $\boldsymbol{\eta}\in\boldsymbol{\Theta}_{s+1}$ the application of the finite difference operator $\nabla_{\boldsymbol{\eta}}$ to a polynomial decreases its degree by 1. Hence formula (10.5) ensures that

$$\begin{aligned}
\widehat{(\nabla^\nu b)}(\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\phi}) = \sum_{\iota\in\widetilde{J}}\sum_{i=\nu}^{\infty} \sum_{\substack{n_1,\dots,n_{K+1}\geqslant 0\\ n_1+\dots+n_{K+1}\leqslant i\\ j_1,\dots,j_d\geqslant 0\\ j_1+\dots+j_d\leqslant i-\nu}} \\
\widetilde{C}^{\iota\, j_1\cdots j_d}_{i\, n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta}) r^{\iota-i}\phi_1^{j_1}\cdots\phi_d^{j_d}\prod_{a=1}^{K+1}(\sin\Phi_a)^{n_a}.
\end{aligned} \tag{10.27}$$

Here $\widetilde{C}^{\iota\, j_1\cdots j_d}_{i\, n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta})$ depend on the coefficients of (9.13) and satisfy a uniform estimate

$$\big|\widetilde{C}^{\iota\, j_1\cdots j_d}_{i\, n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta})\big| \lesssim \rho_n^{(i-\nu-j_1-\dots-j_d)(\alpha_{m+1}+0+)}.$$

Now

$$(\widehat{\nabla^\nu \tilde{\chi}_{\boldsymbol{\theta}}})(\boldsymbol{\xi}) = \sum_{\tilde{\nu}=0}^{\nu} (\widehat{\nabla^{\tilde{\nu}} e_{\boldsymbol{\theta}} \varphi_{\boldsymbol{\theta}}})\Big(\boldsymbol{\xi} + \sum_{p=1}^{\tilde{\nu}} \boldsymbol{\eta}_k\Big)\Big(\nabla^{\nu-\tilde{\nu}} \big(\widehat{|\cdot + \boldsymbol{\theta}|_{\mathbb{C}}^{2w} - |\cdot|_{\mathbb{C}}^{2w}}\big)^{-1}\Big)(\boldsymbol{\xi}).$$

The factors $(\widehat{\nabla^{\tilde{\nu}} e_{\boldsymbol{\theta}} \varphi_{\boldsymbol{\theta}}})$ satisfy the estimate (10.26). For $\boldsymbol{\eta} \in \boldsymbol{\Theta}_{s+1}$ we have

$$\begin{aligned}&\Big(\nabla_{\boldsymbol{\eta}} \big(\widehat{|\cdot + \boldsymbol{\theta}|_{\mathbb{C}}^{2w} - |\cdot|_{\mathbb{C}}^{2w}}\big)^{-1}\Big)(\boldsymbol{\xi}) \\ &= \big(|\boldsymbol{\xi} + \boldsymbol{\eta} + \boldsymbol{\theta}|_{\mathbb{C}}^{2w} - |\boldsymbol{\xi} + \boldsymbol{\eta}|_{\mathbb{C}}^{2w}\big)^{-1} \big(|\boldsymbol{\xi} + \boldsymbol{\theta}|_{\mathbb{C}}^{2w} - |\boldsymbol{\xi}|_{\mathbb{C}}^{2w}\big)^{-1} G(\boldsymbol{\xi}; \boldsymbol{\theta}, \boldsymbol{\eta}),\end{aligned}$$

where

$$\begin{aligned}G(\boldsymbol{\xi}; \boldsymbol{\theta}, \boldsymbol{\eta}) &:= |\boldsymbol{\xi} + \boldsymbol{\theta}|_{\mathbb{C}}^{2w} - |\boldsymbol{\xi}|_{\mathbb{C}}^{2w} - |\boldsymbol{\xi} + \boldsymbol{\eta} + \boldsymbol{\theta}|_{\mathbb{C}}^{2w} + |\boldsymbol{\xi} + \boldsymbol{\eta}|_{\mathbb{C}}^{2w} \\ &= -2w\langle \boldsymbol{\eta}, \boldsymbol{\theta}\rangle |\boldsymbol{\xi}|_{\mathbb{C}}^{2w-2} \\ &\quad + \sum_{j=2}^{\infty} \binom{w}{j} |\boldsymbol{\xi}|_{\mathbb{C}}^{2w-2j} \Big(\big(2\langle \boldsymbol{\xi}, \boldsymbol{\theta}\rangle + |\boldsymbol{\theta}|^2\big)^j \\ &\qquad - \big(2\langle \boldsymbol{\xi}, \boldsymbol{\eta} + \boldsymbol{\theta}\rangle + |\boldsymbol{\eta} + \boldsymbol{\theta}|^2\big)^j + \big(2\langle \boldsymbol{\xi}, \boldsymbol{\eta}\rangle + |\boldsymbol{\eta}|^2\big)^j\Big).\end{aligned}$$

In analogy to (10.27) we have

$$\begin{aligned}&\big(\nabla^\nu \widehat{G(\cdot; \boldsymbol{\theta}, \boldsymbol{\eta})}\big)(\boldsymbol{\xi}) \\ &= \sum_{i=\nu}^{\infty} \sum_{\substack{n_1, \dots, n_{K+1} \geqslant 0 \\ n_1 + \dots + n_{K+1} \leqslant i+2}} \widetilde{C}^i_{n_1 \cdots n_{K+1}}(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\eta}) r^{2w-2-i} \prod_{a=1}^{K+1} (\sin \Phi_a)^{n_a},\end{aligned}$$

with

$$\big|\widetilde{C}^i_{n_1 \cdots n_{K+1}}(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\eta})\big| \lesssim \rho_n^{(i-\nu)(\alpha_{m+1}+0+)+0+}. \tag{10.28}$$

Altogether, applying relations (10.26) – (10.28) to (9.11) and (9.12) we obtain

$$\begin{aligned}&w_{\tilde{k}}(\boldsymbol{\theta}, \boldsymbol{\xi}) \\ &= \sum_{s=0}^{\tilde{k}-1} \sum_{\iota_0, \dots, \iota_s \in \tilde{J}} \sum_{\mu=0}^{s} \sum_{\substack{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_{s+\mu} \in \boldsymbol{\Theta}_{s+1} \\ \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{s+\mu} \in \boldsymbol{\Theta}'_{s+1}}} \sum_{p=0}^{s-\mu} \sum_{i=0}^{\infty} \sum_{\substack{n_1, \dots, n_{K+1} \geqslant 0 \\ n_1 + \dots + n_{K+1} \leqslant 2\mu+p+i}} \\ &\qquad C^{\boldsymbol{\eta}_1 \cdots \boldsymbol{\eta}_{s+\mu}\, \boldsymbol{\theta}_1 \cdots \boldsymbol{\theta}_{s+\mu}}_{s\, \mu\, p\, i\, \iota_0 \cdots \iota_s\, n_1 \cdots n_{K+1}}(\mathbf{X}; \boldsymbol{\theta}) r^{(2w-2)\mu + \iota_0 + \dots + \iota_s - p - i} \\ &\qquad \times \prod_{a=1}^{K+1} (\sin \Phi_a)^{n_a} \prod_{v=1}^{s+\mu} \big(|\boldsymbol{\xi} + \boldsymbol{\eta}_v + \boldsymbol{\theta}_v|_{\mathbb{C}}^{2w} - |\boldsymbol{\xi} + \boldsymbol{\eta}_v|_{\mathbb{C}}^{2w}\big)^{-1},\end{aligned} \tag{10.29}$$

where

$$\big|C^{\boldsymbol{\eta}_1 \cdots \boldsymbol{\eta}_{s+\mu}\, \boldsymbol{\theta}_1 \cdots \boldsymbol{\theta}_{s+\mu}}_{s\, \mu\, p\, i\, \iota_0 \cdots \iota_s\, n_1 \cdots n_{K+1}}(\mathbf{X}; \boldsymbol{\theta})\big| \lesssim \rho_n^{i(\alpha_{m+1}+0+) - (s-\mu-p)\beta + 0+}.$$

According to Lemma 10.1,

$$\begin{aligned}&\big(|\boldsymbol{\xi}+\boldsymbol{\eta}_v+\boldsymbol{\theta}_v|_{\mathbb{C}}^{2w}-|\boldsymbol{\xi}+\boldsymbol{\eta}_v|_{\mathbb{C}}^{2w}\big)^{-1}\\&=r^{2-2w}\big(2\langle\boldsymbol{\xi},\boldsymbol{\theta}_v\rangle+2\langle\boldsymbol{\eta}_v,\boldsymbol{\theta}_v\rangle+|\boldsymbol{\theta}_v|^2\big)^{-1}\\&\qquad\times\sum_{i=0}^{\infty}\sum_{\substack{n_1,\dots,n_{K+1}\geqslant 0\\ n_1+\dots+n_{K+1}\leqslant i}}C^i_{n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\eta}_v,\boldsymbol{\theta}_v)r^{-i}\prod_{a=1}^{K+1}(\sin\Phi_a)^{n_a},\end{aligned}\tag{10.30}$$

and here

$$\big|C^i_{n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\eta}_v,\boldsymbol{\theta}_v)\big|\lesssim\rho_n^{i(\alpha_{m+1}+0+)}.$$

If now subsitute (10.30) to (10.29), we obtain

$$\begin{aligned}&w_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi})\\&=\sum_{s=0}^{\tilde{k}-1}\sum_{\iota_0,\dots,\iota_s\in\widetilde{J}}\sum_{\mu=0}^{s}\sum_{\substack{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_{s+\mu}\in\boldsymbol{\Theta}_{s+1}\\ \boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_{s+\mu}\in\boldsymbol{\Theta}'_{s+1}}}\sum_{p=0}^{s-\mu}\sum_{i=0}^{\infty}\sum_{\substack{n_1,\dots,n_{K+1}\geqslant 0\\ n_1+\dots+n_{K+1}\\ \leqslant 2\mu+p+i}}\\&\qquad C^{\boldsymbol{\eta}_1\cdots\boldsymbol{\eta}_{s+\mu}\,\boldsymbol{\theta}_1\cdots\boldsymbol{\theta}_{s+\mu}}_{s\,\mu\,p\,i\,\iota_0\cdots\iota_s\,n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta},\boldsymbol{\eta}_v,\boldsymbol{\theta}_v)r^{(2-2w)s+\iota_0+\dots+\iota_s-p-i}\\&\qquad\times\prod_{a=1}^{K+1}(\sin\Phi_a)^{n_a}\prod_{v=1}^{s+\mu}\big(2\langle\boldsymbol{\xi},\boldsymbol{\theta}_v\rangle+2\langle\boldsymbol{\eta}_v,\boldsymbol{\theta}_v\rangle+|\boldsymbol{\theta}_v|^2\big)^{-1},\end{aligned}\tag{10.31}$$

with

$$\big|C^{\boldsymbol{\eta}_1\cdots\boldsymbol{\eta}_{s+\mu}\,\boldsymbol{\theta}_1\cdots\boldsymbol{\theta}_{s+\mu}}_{s\,\mu\,p\,i\,\iota_0\cdots\iota_s\,n_1\cdots n_{K+1}}(\mathbf{X};\boldsymbol{\theta},\boldsymbol{\eta}_v,\boldsymbol{\theta}_v)\big|\lesssim\rho_n^{i(\alpha_{m+1}+0+)-(s-\mu-p)\beta+0+}.$$

The first sum in (10.22) can be written in the form

$$\begin{aligned}&\sum_{v=1}^{\infty}\binom{w}{v}z^{2w-v}\Big(2\langle\mathbf{a},\boldsymbol{\Phi}\rangle+z^{-1}\big(|\mathbf{a}|^2+|\mathbf{X}|^2\big)\Big)^v\\&=\sum_{i=0}^{\infty}\sum_{\substack{n_1,\dots,n_{K+1}\geqslant 0\\ n_1+\dots+n_{K+1}\leqslant i+1}}C^i_{n_1\cdots n_{K+1}}(\mathbf{X})z^{2w-1-i}\prod_{a=1}^{K+1}(\sin\Phi_a)^{n_a},\end{aligned}\tag{10.32}$$

where

$$\big|C^i_{n_1\cdots n_{K+1}}(\mathbf{X})\big|\lesssim\rho_n^{(i+1)(\alpha_{m+1}+0+)}.$$

Substituting (10.31) and (10.32) into (10.22) we can calculate the series for the symbol of the operator S^f for $f \in \mathbb{N}$:

$$\begin{aligned}
\widehat{s^f}(\boldsymbol{\theta},\boldsymbol{\xi}) &= \sum_{\substack{\boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_f \in \boldsymbol{\Theta}_{\tilde{k}} \\ \boldsymbol{\phi}_1,\dots,\boldsymbol{\phi}_f \in \boldsymbol{\Theta}_{\tilde{k}}}} C^{\boldsymbol{\theta}_1 \cdots \boldsymbol{\theta}_f}_{\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_f}(\boldsymbol{\theta}) \prod_{g=1}^{f} \hat{s}(\boldsymbol{\theta}_g, \boldsymbol{\xi} + \boldsymbol{\phi}_g) \\
&= \sum_{\nu=0}^{f} \sum_{h=\nu}^{\nu\tilde{k}} \sum_{\iota_1,\dots,\iota_h \in \tilde{J}} \sum_{\mu=0}^{h-\nu} \sum_{\substack{\boldsymbol{\phi}_1,\dots,\boldsymbol{\phi}_{h-\nu+\mu} \in \boldsymbol{\Theta}_{\tilde{k}} \\ \boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_{h-\nu+\mu} \in \boldsymbol{\Theta}'_{\tilde{k}}}} \sum_{p=0}^{h-\nu-\mu} \sum_{i=0}^{\infty} \\
&\quad \sum_{\substack{n_1,\dots,n_{K+1} \geqslant 0 \\ n_1+\dots+n_{K+1} \leqslant 2\mu+p+i+f-\nu}} C(\mathbf{X};\boldsymbol{\theta},\dots) r^{(2-2w)h+(2w-1)f-\nu+\iota_1+\dots+\iota_h-p-i} \\
&\quad \times \prod_{a=1}^{K+1} (\sin \Phi_a)^{n_a} \prod_{v=1}^{h-\nu+\mu} \left(2\langle \boldsymbol{\xi}, \boldsymbol{\theta}_v \rangle + 2\langle \boldsymbol{\phi}_v, \boldsymbol{\theta}_v \rangle + |\boldsymbol{\theta}_v|^2\right)^{-1},
\end{aligned} \tag{10.33}$$

with

$$\big|C(\mathbf{X};\boldsymbol{\theta},\dots)\big| \lesssim \rho_n^{(f-\nu+i)(\alpha_{m+1}+0+)-(h-\nu-\mu-p)\beta+0+}.$$

Note that the last product on the right hand side of (10.33) is of the form

$$\prod_{t=1}^{T} \Big(l_t + \rho \sum_q c_q^t \sin \Phi_q\Big)^{-k_t}.$$

Here we have expanded the inner products $\langle \boldsymbol{\xi}, \boldsymbol{\theta}_v \rangle$ using Lemma 7.5(ii). The coefficients $\{c_q^t\}$ in the decomposition $(\boldsymbol{\theta}_v)_{\mathfrak{V}^\perp} = \sum_q c_q^t \tilde{\boldsymbol{\mu}}_q$ are all of the same sign and satisfy (7.4). Without loss of generality we may assume that all c_q^t are non-negative. The numbers

$$l_t = l(c_1^t, \dots, c_{K+1}^t) := 2L_{m+1} \sum_q c_q^t + 2\langle \mathbf{X}, (\boldsymbol{\theta}_v)_{\mathfrak{V}} \rangle + 2\langle \boldsymbol{\phi}_v, \boldsymbol{\theta}_v \rangle + |\boldsymbol{\theta}_v|^2$$

satisfy $\rho_n^{\alpha_{m+1}} \rho_n^{0-} \lesssim l_t \lesssim \rho_n^{\alpha_{m+1}} \rho_n^{0+}$, since

$$\big|2\langle \mathbf{X}, (\boldsymbol{\theta}_v)_{\mathfrak{V}} \rangle + 2\langle \boldsymbol{\phi}_v, \boldsymbol{\theta}_v \rangle + |\boldsymbol{\theta}_v|^2\big| \lesssim \rho_n^{\alpha_m+0+}.$$

This numbers depend on $\mathbf{X}$, but not on $\boldsymbol{\Phi}$ or ρ. The numbers $k_t = k(c_1^t, \dots, c_{K+1}^t)$ are positive, integer, and independent of $\boldsymbol{\xi}$.

The following lemma is identical to Lemma 10.4 of [12], where for our purposes we have replaced the explicit constants 1/2 and 2/3 by ϑ and ς, respectively.

Lemma 10.7. *For* $1 \leqslant K \leqslant d-1$*;* $n_1, \dots, n_{K+1} \in \mathbb{N}_0$*;* $k_1, \dots, k_T \in \mathbb{N}$ *let* $Q := \sum_{t=1}^{T} k_t$*,*

$$\hat{J}_K := \int_{\mathcal{M}_p} \frac{(\sin \Phi_1)^{n_1} \dots (\sin \Phi_K)^{n_K} (\sin \Phi_{K+1})^{n_{K+1}}\, d\boldsymbol{\Phi}}{\prod_{t=1}^{T} \left(l_t + \rho \sum_{j=1}^{K+1} c_j^t \sin \Phi_j\right)^{k_t}}.$$

Then there exist positive numbers δ_0, p_K, and q_K depending only on the constants (2.20) *and K such that*

$$\hat{J}_K = \sum_{q=0}^{K} (\ln\rho)^q \sum_{p=0}^{\infty} e(p,q)\rho^{-p},$$

where

$$\big|e(p,q)\big| \lesssim \rho_n^{(\varsigma - p_K)p} \rho_n^{-Q\beta}.$$

These estimates are uniform in the following regions of variables:

$$\rho_n^{\beta} \lesssim l_t \lesssim \rho_n^{\vartheta}, \quad \rho_n^{-\delta_0} \lesssim c_j^t \lesssim \rho_n^{\delta_0}, \quad \rho_n^{\varsigma - q_K} < \rho.$$

Now using Lemma 10.7 we can compute the integrals of (10.33) over the domain $\{\mathbf{\Phi} \in \mathcal{M}_p\}$ (recall that this integration is not needed for $K = 0$ by Remark 10.4). Substituting the result into (10.25), integrating in $d\boldsymbol{\nu}$ over $\Omega(\mathfrak{V})$, and taking into account (10.12) and (10.14) we obtain in the region $2\rho_n/3 < \rho < 6\rho_n$

$$\begin{aligned}
&\operatorname{vol}\mathcal{A}_p^+(\rho) - \operatorname{vol}\mathcal{A}_p^-(\rho)\\
&= \sum_{q=0}^{K} \sum_{h=0}^{(l_0+1)\tilde{k}} \sum_{\iota_1,\dots,\iota_h \in \widetilde{J}} \sum_{j=0}^{\infty} C_{q\,h\,j}^{\iota_1\cdots\iota_h} \rho^{K+1+(2-2w)h+\iota_1+\cdots+\iota_h-j} (\ln\rho)^q\\
&\quad + O(\rho_n^{-M+2w-d}),
\end{aligned}$$

with the coefficients satisfying

$$|C_{q\,h\,j}^{\iota_1\cdots\iota_h}| \lesssim \rho_n^{-2\beta h+\varsigma j}.$$

This, together with equations (10.3), (6.9), Lemma 6.1, relation (2.20), Section 11 of [12], and the observation that the number of different quasi-lattice subspaces $\mathfrak{V}$ is $\lesssim \rho_n^{0+}$, completes the proof of Lemma 3.1 and, thus, of our main theorem in the case of $B = \widetilde{B}$ with the symbol satisfying (2.8). As explained at the end of Section 3, the summation over $\widetilde{J}$ may be replaced by summation over J_0.

It remains to relax the assumptions on B. This will be done in the subsequent section.

11. Approximation

In this section we prove Lemma 3.1 and thus Theorem 2.6 for general B using the fact that the proof is complete for $\widetilde{B}$ whose symbol fulfills the extra assumption (2.8).

1. Given B satisfying the hypothesis of Theorem 2.6 and the number M, we fix the values of k and $\tilde{k}$ in such a way that Lemma 3.1 holds true for $H = (-\Delta)^w + \widetilde{B}$, where the symbol $\tilde{b}$ of $\widetilde{B}$ satisfying (2.8) is constructed at the end of Section 3. For $R > 0$ let us define (recall (6.1))

$$\mathcal{P}_R := \mathcal{P}^L(\mathcal{B}_R), \quad \mathcal{P}_R^c := \mathcal{P}^L(\mathbb{R}^d \setminus \mathcal{B}_R)$$

We start by estimating the quadratic form of $B - \widetilde{B}$.
For any $\psi \in \mathsf{H}^{2w}(\mathbb{R}^d)$

$$\begin{aligned} \big|\langle\psi, (B-\widetilde{B})\psi\rangle\big| \leqslant \big|\langle\psi, \mathcal{P}_{R_0}(B-\widetilde{B})\mathcal{P}_{R_0}\psi\rangle\big| + \big|\langle\psi, \mathcal{P}_{R_0}(B-\widetilde{B})\mathcal{P}_{R_0}^c\psi\rangle\big| \\ + \big|\langle\psi, \mathcal{P}_{R_0}^c(B-\widetilde{B})\mathcal{P}_{R_0}\psi\rangle\big| + \big|\langle\psi, \mathcal{P}_{R_0}^c(B-\widetilde{B})\mathcal{P}_{R_0}^c\psi\rangle\big|. \end{aligned} \tag{11.1}$$

By Condition (2.14), the symbol of $(B - \widetilde{B})\mathcal{P}_{R_0}^c$ satisfies

$$\Big|(b - \tilde{b})\, \mathbb{1}_{\mathbb{R}^d \setminus \mathcal{B}_{R_0}}\Big|^{(\varkappa/\beta)}_{\varkappa/2,\,0} < \rho_n^{-k}.$$

Now Propositions 4.1 and 4.2 imply that

$$\big\|(-\Delta+1)^{-\varkappa/4}(B-\widetilde{B})(-\Delta+1)^{-\varkappa/4}\mathcal{P}_{R_0}^c\big\| \leqslant C\rho_n^{-k}. \tag{11.2}$$

Hence

$$\begin{aligned} &\big|\langle\psi, \mathcal{P}_{R_0}(B-\widetilde{B})\mathcal{P}_{R_0}^c\psi\rangle\big| \\ &= \big|\langle(-\Delta+1)^{\varkappa/4}\psi, \\ &\qquad \mathcal{P}_{R_0}(-\Delta+1)^{-\varkappa/4}(B-\widetilde{B})\mathcal{P}_{R_0}^c(-\Delta+1)^{-\varkappa/4}(-\Delta+1)^{\varkappa/4}\psi\rangle\big| \\ &\leqslant C\rho_n^{-k}\langle\psi, (-\Delta+1)^{\varkappa/2}\psi\rangle, \end{aligned} \tag{11.3}$$

and the analogous estimates hold for the last two terms in (11.1). Thus (11.1) implies

$$|B - \widetilde{B}| \leqslant B^{(k)}, \tag{11.4}$$

where $B^{(k)}$ is the operator of multiplication by the function

$$b^{(k)}(\boldsymbol{\xi}) := \begin{cases} \|b\|_{\mathsf{L}_\infty(\mathbb{R}^d\times\mathcal{B}_{R_0})} + \|\tilde{b}\|_{\mathsf{L}_\infty(\mathbb{R}^d\times\mathcal{B}_{R_0})}, & |\boldsymbol{\xi}| \leqslant R_0, \\ C\rho_n^{-k}\big(1+|\boldsymbol{\xi}|^2\big)^{\varkappa/2}, & |\boldsymbol{\xi}| > R_0 \end{cases} \tag{11.5}$$

in the momentum space.

In view of Lemma 2.3(a), we conclude that

$$N\big((-\Delta)^w + B, \lambda\big) \gtrless N\big((-\Delta)^w + \widetilde{B} \pm B^{(k)}, \lambda\big). \tag{11.6}$$

So to prove (3.1) it will be sufficient to show that for $\rho \in I_n$ (which we assume everywhere below) the right hand side of (11.6) does not differ from $N\big((-\Delta)^w + \widetilde{B}, \rho^{2w}\big)$ by more than $O(\rho_n^{-M})$. By (3.1) and Remark 2.7, it is enough to prove that

$$N\big((-\Delta)^w + \widetilde{B} \pm B^{(k)}, \lambda\big) = N\big((-\Delta)^w + \widetilde{B}, \lambda + O(\rho_n^{2w-d-M})\big). \tag{11.7}$$

2. We note that for

$$R_* := (4\rho_n^{d+M})^{1/(w-\varkappa)} \tag{11.8}$$

we have

$$\begin{aligned} &N\big((-\Delta)^w + \widetilde{B} \pm B^{(k)}, \lambda\big) \\ &= N\big((-\Delta)^w + \widetilde{B} \pm \mathcal{P}_{R_0} B^{(k)} \pm \mathcal{P}^c_{R_*} B^{(k)}, \lambda + O(\rho_n^{2w-d-M})\big). \end{aligned} \tag{11.9}$$

Indeed,

$$\big\|(\mathcal{P}_{R_*} - \mathcal{P}_{R_0})B^{(k)}\big\| = C\rho_n^{-k}(1+R_*^2)^{\varkappa/2} = O(\rho_n^{2w-d-M})$$

in view of (3.10).

3. Now we are going to prove that

$$\begin{aligned} &N\big((-\Delta)^w + \widetilde{B} \pm \mathcal{P}_{R_0} B^{(k)} \pm \mathcal{P}^c_{R_*} B^{(k)}, \lambda\big) \\ &= N\big((-\Delta)^w + \widetilde{B} \pm \mathcal{P}^c_{R_*} B^{(k)}, \lambda + O(\rho_n^{2w-d-M})\big). \end{aligned} \tag{11.10}$$

This will be done with the help of the following lemma, which is a development of Lemma 3.1 from [10].

Lemma 11.1. *Let H_0, V, A be pseudo–differential operators with almost–periodic coefficients. Suppose that $H := H_0 + V$ is elliptic, selfadjoint and bounded below, and there exists a collection of orthogonal projections $\{P_l\}_{l=0}^L$ commuting with H_0 such that*

$$\sum_{l=0}^{L} P_l = I \quad \textit{and} \quad V_{n\,l} := P_n V P_l = 0 \quad \textit{for} \quad |l-n| > 1. \tag{11.11}$$

Suppose that $A = P_0 A$ and that

$$a := \|A\| < \infty.$$

At last, suppose that for $\lambda \in \mathbb{R}$

$$D_l := \operatorname{dist}\big(\lambda, \sigma(P_l H P_l)\big) - (4 + 2^{5-L})a > 0, \quad l = 0, \dots, L-1 \tag{11.12}$$

and

$$\max_{0 \leqslant l \leqslant L-1} \big(a + \|V_{l\,l-1}\| + \|V_{l\,l+1}\|\big)/D_l \leqslant 1/4. \tag{11.13}$$

Then for

$$\varepsilon := 2^{4-L} a \tag{11.14}$$

we have

$$N(H, \lambda - \varepsilon) \leqslant N(H + A, \lambda) \leqslant N(H, \lambda + \varepsilon). \tag{11.15}$$

Proof. We will prove the first inequality; the second follows by interchanging the roles of H_0 and $H_0 + A$. Let E_λ be the spectral projection of $(-\infty, \lambda]$ for H. By Lemma 4.1 of [12] it is enough to prove that

$$\langle \phi, (H + A)\phi \rangle \leqslant \lambda \|\phi\|^2 \quad \text{for every} \quad \phi \in E_{\lambda-\varepsilon} \mathsf{L}_2(\mathbb{R}^d). \tag{11.16}$$

Let

$$\delta := \min\{a, 2^{-3-L} \min_{0 \leqslant l \leqslant L-1} D_l\}, \quad K := [2a/\delta] + 2, \tag{11.17}$$

so that

$$2a \leqslant (K-1)\delta \leqslant 3a \tag{11.18}$$

and by (11.13)

$$K-1 \leqslant 3a/\delta \leqslant 3\max\{1, 2^{L+3}a/\min_{0\leqslant l\leqslant L-1} D_l\} \leqslant 2^{L+3}. \tag{11.19}$$

For $\phi \in E_{\lambda-\varepsilon}$ introduce

$$\phi^k := (E_{\lambda-\varepsilon-(k-1)\delta} - E_{\lambda-\varepsilon-k\delta})\phi, \quad k = 1,\dots,K-1,$$

$$\phi^K := E_{\lambda-\varepsilon-(K-1)\delta}\phi, \quad \phi' := \phi - \phi^K = \sum_{k=1}^{K-1}\phi^k.$$

Then $\phi = \sum_{k=1}^{K}\oplus\phi^k$ and, letting

$$\eta^k := H\phi^k - \big(\lambda-\varepsilon-(k-1)\delta\big)\phi^k, \quad k=1,\dots,K-1, \tag{11.20}$$

we have

$$\|\eta^k\| \leqslant \delta\|\phi^k\|. \tag{11.21}$$

Let $P_{-1} := P_{L+1} := 0$. Projecting (11.20) with P_l we obtain

$$\eta_l^k = V_{l\,l-1}\phi_{l-1}^k + \Big(P_lHP_l - \big(\lambda-\varepsilon-(k-1)\delta\big)\Big)\phi_l^k + V_{l\,l+1}\phi_{l+1}^k, \quad l=0,\dots,L,$$

and thus by (11.13), (11.17) and (11.21)

$$\begin{aligned}\|\phi_l^k\| &\leqslant \big(\|\eta_l^k\| + \|V_{l\,l-1}\|\|\phi_{l-1}^k\| + \|V_{l\,l+1}\|\|\phi_{l+1}^k\|\big)/D_l \\ &\leqslant 2^{-3-L}\|\phi^k\| + \|\phi_{l-1}^k\|/4 + \|\phi_{l+1}^k\|/4, \quad l=0,\dots,L-1.\end{aligned}$$

By induction, starting from $l=0$ we obtain

$$\|\phi_l^k\| \leqslant 2^{-2-L}\|\phi^k\| + 3\|\phi_{l+1}^k\|/8, \quad l=0,\dots,L-1.$$

Again by induction, using that $\|\phi_L^k\| \leqslant \|\phi^k\|$, we get $\|\phi_l^k\| \leqslant 2^{l-L}\|\phi^k\|$, $l=1,\dots,L$ and thus $\|\phi_0^k\| \leqslant 2^{-L}\|\phi^k\|$. Therefore, for $k=1,\dots,K-1$,

$$\|A\phi^k\| = \|A\phi_0^k\| \leqslant 2^{-L}a\|\phi^k\|,$$

and thus

$$\|A\phi'\| \leqslant \sum_{k=1}^{K-1}\|A\phi^k\| \leqslant 2^{-L}\sqrt{K-1}\,a\|\phi'\|$$

and

$$\big|\langle\phi', A\phi'\rangle\big| = \Big|\sum_{k,m=1}^{K-1}\langle\phi_0^k, A\phi_0^m\rangle\Big| \leqslant 2^{-2L}(K-1)a\|\phi'\|^2.$$

Hence

$$\begin{aligned}&\langle\phi,(H+A)\phi\rangle \\ &= \langle\phi', H\phi'\rangle + \langle\phi', A\phi'\rangle + 2\Re\langle\phi^K, A\phi'\rangle + \langle\phi^K, H\phi^K\rangle + \langle\phi^K, A\phi^K\rangle \\ &\leqslant (\lambda-\varepsilon)\|\phi'\|^2 + 2^{-2L}(K-1)a\|\phi'\|^2 + 2^{1-L}\sqrt{K-1}\,a\|\phi'\|\|\phi^K\| \\ &\qquad + \big(\lambda-\varepsilon-(K-1)\delta\big)\|\phi^K\|^2 + a\|\phi^K\|^2 \\ &\leqslant \big(\lambda-\varepsilon+2^{1-2L}(K-1)a\big)\|\phi'\|^2 + \big(\lambda-\varepsilon-(K-1)\delta+2a\big)\|\phi^K\|^2 \\ &\leqslant \lambda\|\phi\|^2,\end{aligned}$$

where the last inequality follows from (11.18) and (11.19). □

We now want to apply Lemma 11.1 to

$$H_0^{\pm} := (-\Delta)^w \pm \mathcal{P}_{R_*}^c B^{(k)}, \quad V := \widetilde{B}, \quad A^{\pm} := \pm \mathcal{P}_{R_0} B^{(k)}.$$

Note that

$$a := \|b\|_{\mathsf{L}_\infty(\mathbb{R}^d \times \mathcal{B}_{R_0})} + \|\tilde{b}\|_{\mathsf{L}_\infty(\mathbb{R}^d \times \mathcal{B}_{R_0})} \tag{11.22}$$

does not depend on ρ_n. For

$$L := \big[4 + \log_2 a + (M + d - 2w) \log_2 \rho_n\big] + 1 \tag{11.23}$$

we let

$$R_l := R_0 + l\rho_n^{2/k}, \quad l = 0, \dots, L-1, \tag{11.24}$$

and introduce a family of projections

$$P_0 := \mathcal{P}_{R_0}, \quad P_l := \mathcal{P}_{R_l} - \mathcal{P}_{R_{l-1}}, \quad l = 1, \dots, L-1, \quad P_L := \mathcal{P}_{R_{L-1}}^c. \tag{11.25}$$

Let us check that the hypothesis of Lemma 11.1 is satisfied. Relation (11.11) follows from (2.15) and (11.24). It follows from (2.9) that for $l \leqslant L-1$

$$\begin{aligned} \|P_l H P_l\| &\leqslant \Big\| P_{L-1}\big((-\Delta)^w + \widetilde{B}\big) P_{L-1} \Big\| \\ &\leqslant 2\big\|P_{L-1}(-\Delta)^w P_{L-1}\big\| \leqslant 2(R_{L-1})^{2w}. \end{aligned} \tag{11.26}$$

Also, for $l \leqslant L-1$

$$\|V_{l\,l-1}\| + \|V_{l\,l+1}\| \leqslant 2(R_{L-1} + \rho_n^{2/k})^{2\tilde{w}}. \tag{11.27}$$

Since by (11.23) and (11.24) we have

$$R_{L-1} = R_0 + \big[4 + \log_2 a + (M + d - 2w) \log_2 \rho_n\big] \rho_n^{2/k} \lesssim \rho_n^{2/k} \log \rho_n, \tag{11.28}$$

relations (11.12) and (11.13) follow from (11.26) and (11.27) if ρ_n is big enough.

Applying Lemma 11.1, we get (11.15) with

$$\varepsilon = 2^{4-L} a \leqslant \rho_n^{-M+2w-d},$$

which implies (11.10).

4. It remains to prove that

$$N\big((-\Delta)^w + \widetilde{B} \pm \mathcal{P}_{R_*}^c B^{(k)}, \lambda\big) = N\big((-\Delta)^w + \widetilde{B}, \lambda + O(\rho_n^{2w-d-M})\big). \tag{11.29}$$

Choose

$$\varepsilon := \rho_n^{-d-M}. \tag{11.30}$$

In view of (2.9), we have

$$(-\Delta)^{\tilde{w}} + \widetilde{B} \lesseqgtr \mathcal{P}_{R_*}(1 \pm \varepsilon)\big((-\Delta)^{\tilde{w}} + \widetilde{B}\big)\mathcal{P}_{R_*} \oplus \mathcal{P}_{R_*}^c(1 \pm 1/\varepsilon)\big((-\Delta)^{\tilde{w}} + \widetilde{B}\big)\mathcal{P}_{R_*}^c.$$

Therefore,

$$\begin{aligned} &(-\Delta)^w + \widetilde{B} \pm \mathcal{P}_{R_*}^c B^{(k)} \\ &\lesseqgtr \mathcal{P}_{R_*}\big((-\Delta)^w \pm \varepsilon(-\Delta)^{\tilde{w}} + (1 \pm \varepsilon)\widetilde{B}\big)\mathcal{P}_{R_*} \\ &\qquad \oplus \mathcal{P}_{R_*}^c\big((-\Delta)^w \pm (-\Delta)^{\tilde{w}}/\varepsilon + (1 \pm 1/\varepsilon)\widetilde{B} \pm B^{(k)}\big)\mathcal{P}_{R_*}^c. \end{aligned} \tag{11.31}$$

Using (2.9) again and recalling the definitions (11.30), (11.5) and (11.8), we can estimate the last term on the right hand side of (11.31) from below:

$$\begin{aligned}&\mathcal{P}^c_{R_*}\big((-\Delta)^w \pm (-\Delta)^{\tilde w}/\varepsilon + (1\pm 1/\varepsilon)\widetilde{B} \pm B^{(k)}\big)\mathcal{P}^c_{R_*}\\ &\ > \big((-\Delta)^w - 2(-\Delta)^{\tilde w}/\varepsilon\big)\mathcal{P}^c_{R_*} \geqslant (R_*^{2w} - 2R_*^{2\tilde w}/\varepsilon)\mathcal{P}^c_{R_*} \geqslant (5\rho_n)^{2w}\mathcal{P}^c_{R_*},\end{aligned}$$

so it does not contribute to the density of states for $\rho \in I_n$. For the first term we have

$$\mathcal{P}_{R_*}\big((-\Delta)^w \pm \varepsilon(-\Delta)^{\tilde w} + (1\pm\varepsilon)\widetilde{B}\big)\mathcal{P}_{R_*} \lessgtr \mathcal{P}_{R_*}(1\pm\varepsilon)\big((-\Delta)^w + \widetilde{B}\big)\mathcal{P}_{R_*},$$

so

$$\begin{aligned}&N\Big((-\Delta)^w + \widetilde{B} \pm \mathcal{P}^c_{R_*}B^{(k)}, \lambda\Big)\\ &\gtrless N\Big(\mathcal{P}_{R_*}\big((-\Delta)^w \pm \varepsilon(-\Delta)^{\tilde w} + (1\pm\varepsilon)\widetilde{B}\big)\big|_{\mathcal{P}_{R_*}\mathsf{L}_2(\mathbb{R}^d)}, \lambda\Big)\\ &\gtrless N\Big(\mathcal{P}_{R_*}\big((-\Delta)^w + \widetilde{B}\big)\big|_{\mathcal{P}_{R_*}\mathsf{L}_2(\mathbb{R}^d)}, \lambda/(1\pm\varepsilon)\Big),\end{aligned} \tag{11.32}$$

and the same estimates hold true for $B^{(k)}$ replaced by 0. Combining these two versions of (11.32), we obtain

$$\begin{aligned}N\big((-\Delta)^w + \widetilde{B}, \lambda\big) &\lessgtr N\big((-\Delta)^w + \widetilde{B} \mp \mathcal{P}^c_{R_*}B^{(k)}, \lambda\big)\\ &\lessgtr N\Big(\mathcal{P}_{R_*}\big((-\Delta)^w + \widetilde{B}\big)\big|_{\mathcal{P}_{R_*}\mathsf{L}_2(\mathbb{R}^d)}, \lambda/(1\mp\varepsilon)\Big)\\ &\lessgtr N\big((-\Delta)^w + \widetilde{B}, (1\pm\varepsilon)\lambda/(1\mp\varepsilon)\big).\end{aligned} \tag{11.33}$$

Recalling that $\lambda = \rho^{2w} \leqslant (4\rho_n)^{2w}$ and (11.30), we arrive at (11.29).

Combining (11.9), (11.10) and (11.29), we get (11.7).

References

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

[2] M. Hitrik, I. Polterovich, *Regularized traces and Taylor expansions for the heat semigroup.* J. London Math. Soc. **68**(2) (2003), 402–418.

[3] M. Hitrik, I. Polterovich, *Resolvent expansions and trace regularizations for Schrödinger operators.* Adv. in Diff. Eq. and Math. Phys., Contemp. Math., AMS, 2003.

[4] Yu. Karpeshina, *On the density of states for the periodic Schrödinger operator*, Ark. Mat. **38** (2000), 111–137.

[5] Yu. Karpeshina, *Perturbation theory for the Schrödinger operator with a periodic potential*, Lecture Notes in Math., Vol. 1663, Springer Berlin 1997.

[6] Y. E. Karpeshina, *Spectral properties of periodic magnetic Schrödinger operator in the high-energy region. Two-dimensional case*, Comm. Math. Phys. **251**(3) (2004), 473 – 514.

[7] E. Korotyaev, A. Pushnitski, *On the High-Energy Asymptotics of the Integrated Density of States.* Bull. LMS **35** (2003), No. 6, 770–776.

[8] A. Mohamed, *Asymptotic of the density of states for the Schrödinger operator with periodic electromagnetic potential*, J. Math. Phys. **38** no. 8 (1997), 4023-4051.

[9] S. Morozov, L. Parnovski, I. Pchelintseva, *Lower bound on the density of states for periodic Schrödinger operators*, Operator Theory and Its Applications: In Memory of V. B. Lidskii (1924-2008), AMS Translations - Series 2 (231), Adv. in Math. Sciences, 161–172.

[10] L. Parnovski, *Bethe-Sommerfeld conjecture.* Ann. H. Poincaré **9**(3) (2008), 457–508.

[11] L. Parnovski, R. Shterenberg, *Asymptotic expansion of the integrated density of states of a two-dimensional periodic Schrödinger operator.* Invent. Math. **176**(2) (2009), 275–323.

[12] L. Parnovski, R. Shterenberg, *Complete asymptotic expenasion of the integrated density of states of multidimensional almost-periodic Schrödinger operators.* Ann. of Math. **176**(2) (2012), 1039–1096.

[13] L. Parnovski, A. Sobolev, *Bethe-Sommerfeld conjecture for periodic operators with strong perturbations.* Invent. Math. **181**(3) (2010), 467 - 540.

[14] A. V. Savin, *Asymptotic expansion of the density of states for one-dimensional Schrödinger and Dirac operators with almost periodic and random potentials*, Sb. Nauchn. Tr. , I.F.T.P., Moscow (Russian), 1988.

[15] M. Shubin, *Almost periodic functions and partial differential operators*, Russian Math. Surveys, **33**(2) (1978), 1–52.

[16] M. Shubin, *The spectral theory and the index of elliptic operators with almost periodic coefficients*, Russian Math. Surveys, **34**(2) (1979), 109–157.

[17] D. Shenk and M. Shubin, *Asymptotic expansion of the state density and the spectral function of a Hill operator*, Math. USSR Sbornik **56** (1987), No. 2, 473–490.

[18] M. Skriganov, *Geometrical and arithmetical methods in the spectral theory of the multi-dimensional periodic operators*, Proc. Steklov Math. Inst., Vol. 171, 1984.

[19] A. V. Sobolev, *Asymptotics of the integrated density of states for periodic elliptic pseudo-differential operators in dimension one*, Rev. Mat. Iberoam. **22** (2006), no. 1, 55–92.

[20] A. V. Sobolev, *Integrated Density of States for the Periodic Schrödinger Operator in Dimension Two*, Ann. H. Poincaré, **6** (2005), 31–84.

[21] O. A. Veliev, *On the spectrum of multidimensional periodic operators*, Theory of Functions, functional analysis and their applications, Kharkov University, **49** (1988), 17–34 (in Russian).

Sergey Morozov
Department of Mathematics
Aarhus University
Ny Munkegade 118
DK-8000 Aarhus C
Denmark
e-mail: `smorozov@imf.au.dk`

Leonid Parnovski
Department of Mathematics
University College London
Gower Street
London
WC1E 6BT
UK
e-mail: leonid@math.ucl.ac.uk

Roman Shterenberg
Department of Mathematics
University of Alabama at Birmingham
1300 University Blvd.
Birmingham AL 35294
USA
e-mail: shterenb@math.uab.edu