Annals of Mathematics **176** (2012), 1039–1096
http://dx.doi.org/10.4007/annals.2012.176.2.8

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

By Leonid Parnovski and Roman Shterenberg

Abstract

We prove the complete asymptotic expansion of the integrated density of states of a Schrödinger operator $H = -\Delta + b$ acting in $\mathbb{R}^d$ when the potential b is either smooth periodic, or generic quasi-periodic (finite linear combination of exponentials), or belongs to a wide class of almost-periodic functions.

1. Introduction

We consider the Schrödinger operator

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

acting in $L_2(\mathbb{R}^d)$. The potential $b = b(\mathbf{x})$ is assumed to be real, smooth, and either periodic, or almost-periodic; in the almost-periodic case we assume that all the derivatives of b 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. 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 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

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

is the volume of the unit ball in $\mathbb{R}^d$. There is a long-standing conjecture that, at least in the case of periodic b, 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 that depend on the potential b. They can be calculated relatively easily using the heat kernel invariants (computed in [3]); they are equal to certain integrals of the potential b and its derivatives. Indeed, in the paper [8], all these coefficients were computed; in particular, it turns out that if d is even, then e_j vanish whenever $j > d/2$.

Until recently, formula (1.5) was proved only in the case $d = 1$ in [15] for periodic b and in [14] for almost-periodic b. In the recent paper [11], this formula was proved in the case $d = 2$ and periodic potential. In the periodic case and $d \geq 3$, only partial results are known; see [2], [6], [5], [12], [18]. In particular, in [6] 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 $d > 3$. Finally, in the multidimensional almost-periodic case, formula (1.6) is known only with $K = 0$ and $R(\lambda) = O(\lambda^{\frac{d-2}{2}})$; see [17].

The aim of our paper is to prove formula (1.6) with arbitrary K for all dimensions d and for periodic or almost-periodic potentials. In the case of periodic potential, we do not impose any additional assumptions (besides infinite smoothness) on it. However, if the potential b is almost-periodic, we need it to satisfy certain extra conditions; since the formulation of them requires several definitions, we will list these conditions and formulate our main result in the next section.

Now we discuss the difference in the approaches of [11] and this paper. To begin with, let us assume that the potential b is periodic. Then we can perform the Floquet-Bloch decomposition (see, e.g., [13]) and express the operator H as a direct integral

$$H = \int_{\oplus} H(\mathbf{k}) d\mathbf{k}, \tag{1.7}$$

quasi-momentum $\mathbf{k}$ running over $\mathcal{O}^\dagger$ — the cell of the lattice $\Gamma^\dagger$, dual to the lattice of periods Γ. The very first thing we need to do is to replace definition (1.2) with a different one. There are two problems with definition (1.2). The first problem is that this definition is rather difficult to work with. The second problem is that this definition makes sense only for differential operators. And, although we are working with a differential operator (1.1) in the beginning, our methods require us to replace this operator with a pseudodifferential one, and so we need a definition that works for pseudodifferential operators as well. In the periodic case, the alternative definition is given by the formula

$$N(\lambda) := \frac{1}{(2\pi)^d} \int_{\mathcal{O}^\dagger} N(\lambda, H(\mathbf{k})) d\mathbf{k}, \tag{1.8}$$

where $N(\lambda, H(\mathbf{k}))$ is the eigenvalue counting function of $H(\mathbf{k})$. The first step of obtaining information on the density of states is to compute the precise asymptotics of the eigenvalues of $H(\mathbf{k})$. There are two different approaches to doing this. The first method, called the method of spectral projections, was developed in [10]. When using this method, we study, instead of H, the operator $\tilde{H} = \sum_j P_j H P_j$, where $\{P_j\}$ are spectral projections of the unperturbed operator $H_0 := -\Delta$. It was shown in [10] that if we carefully choose these projections, then the spectra of $H(\mathbf{k})$ and $\tilde{H}(\mathbf{k})$ are close to each other. Next, we can decompose the operator $\tilde{H}$ into invariant subspaces. There are two types of such subspaces. The first type (called stable, or nonresonant subspaces) corresponds to eigenvalues of H that are far away from other eigenvalues; in studying them we can use straightforward perturbation theory to compute their precise asymptotic behaviour. This was done in [10], but in certain cases such computations were performed earlier (see, for example, [5] and [21]). The second type of subspaces (called unstable, or resonant) corresponds to clusters of eigenvalues of H lying close to each other. In order to study these eigenvalues, we have to use perturbation theory of multiple eigenvalues, and this theory is much more difficult and less precise than in the stable case. The methods of [10] allow us to reduce the study of resonant eigenvalues to the study of a family of operators $A + \varepsilon B$ when $\varepsilon \to 0$. Here, A and B are finite-dimensional self-adjoint operators and $\varepsilon \sim \lambda^{-1/2}$ is a small parameter. We are interested in the eigenvalues of $A + \varepsilon B$ that are perturbations of zero eigenvalues of A. Of course, we can write the formula $\lambda(A + \varepsilon B) \sim \sum \lambda_j \varepsilon^j$ (see [7]) but the coefficients λ_j will, in general, be unbounded functions of the quasi-momentum and, therefore, we cannot integrate these asymptotic expansions against $d\mathbf{k}$. Paper [11] deals with this problem in the case $d = 2$. We study the operator PBP, where P is the orthogonal projection onto the kernel of A, and show that the cluster of eigenvalues of this operator has multiplicity at most two. This allows us, using the Weierstrass Preparation Theorem, to prove that the eigenvalues of $A+\varepsilon B$ enjoy the asymptotic formula $\lambda(A+\varepsilon B) \sim \sum \lambda_j \varepsilon^j \pm \sqrt{\sum \tilde{\lambda}_j \varepsilon^j}$, where the coefficients λ_j and $\tilde{\lambda}_j$ are bounded functions of the quasi-momentum $\mathbf{k}$ and so can be integrated against $d\mathbf{k}$. Unfortunately, this approach does not work if $d \geq 3$, since then the cluster multiplicity of PBP becomes unbounded.

The second method of obtaining asymptotic formulas for the eigenvalues of $H(\mathbf{k})$ was developed in [20] and [19] and was also used in [12]. This method, which we call the gauge transform method, consists of constructing two pseudodifferential operators, H_1 and H_2. Here, $H_1 = e^{i\Psi} H e^{-i\Psi}$, where Ψ is a bounded periodic self-adjoint pseudo-differential operator of order 0. Thus, the eigenvalues of $H_1(\mathbf{k})$ coincide with the eigenvalues of $H(\mathbf{k})$. The operator H_2 is close to H_1 in norm; also, operators $H_2(\mathbf{k})$ have a lot of invariant subspaces. As in the previous method, these invariant subspaces can be generated by stable and unstable eigenvalues, and the case of the stable eigenvalues can be treated completely (i.e., the complete asymptotic formula for such eigenvalues can be obtained). The difference with the previous approach lies in the form of the restriction of the operator H_2 to a subspace generated by a resonant eigenvalue. Let $\boldsymbol{\xi}$ be a point in the phase space lying in a resonant region generated by a lattice subspace $\mathfrak{V}$ (see Section 5 for the definitions and more details). Then H_2, restricted to the invariant subspace generated by $\boldsymbol{\xi}$, has a form $r^2 I + S(r)$, where r is, essentially, the distance from $\boldsymbol{\xi}$ to $\mathfrak{V}$ and $S(r)$ is a finite-dimensional self-adjoint operator that can grow in r, but slower than r^2. Using the method of spectral projections, we could achieve that $S(r)$ has a simple form, namely, $S(r) = rA + B = r(A + \varepsilon B)$, where $\varepsilon = r^{-1}$, which was the advantage of that method. The advantage of the method of gauge transform is that all eigenvalues of the reduced operator $r^2 I + S(r)$ contribute to the integrated density of states, whereas in the method of spectral projections only the eigenvalues coming from the zero eigenvalues of A (and not even all such eigenvalues) were of interest to us. This observation makes the method of gauge transform much more convenient to use, despite the operator $S(r)$ being more complicated than in the method of spectral projections. Indeed, the fact that all the eigenvalues contribute to the density of states allows us to use the residue theorem in order to compute the sum of contributions from all eigenvalues without computing the contributions from individual eigenvalues; see (10.11) and (10.18). In fact, formula (10.18) is the most crucial observation, which has enabled us to compute the contribution to the density of states from the resonance regions.

When we were working on the details of this approach, we realized that, as a matter of fact, the decomposition (1.7) is not required, and all the steps can be written for the 'global' operators H without any references to the 'fibre' operators $H(\mathbf{k})$. This led us to believe that this method is likely to be applicable in a range of other settings. In particular, it turned out that this method works for quasi-periodic and almost-periodic potentials. So, let us assume that the potential b is almost-periodic. What is the analogue of definition (1.8) in this case? The answer to this question can be found in [17]. One possible way of defining the density of states is via the von Neumann algebras; we discuss this approach later in our paper. However, the ultimate definition is as follows. Let $e(\lambda; \mathbf{x}, \mathbf{y})$ be the kernel of the spectral projection of an elliptic pseudo-differential operator of positive order with almost-periodic coefficients. Then, it was proved in Theorem 4.1 of [17] that the density of states of this operator satisfies (at least at its continuity points)

$$N(\lambda) = \mathbf{M}_{\mathbf{x}}(e(\lambda; \mathbf{x}, \mathbf{x})), \tag{1.9}$$

where $\mathbf{M}$ is the mean of an almost-periodic function. Our main tool in the proof will be formula (1.9), but we will use the operator-algebraic definition sometimes (for example, to show that density of states decreases when the operator increases, something not immediately obvious from (1.9)). Another useful observation that helped us in extending our results to the almost-periodic case is this. Let A be an elliptic pseudo-differential operator with almost-periodic coefficients. We are usually assuming that A acts in $L_2(\mathbb{R}^d)$. However, we can consider actions of A (via the same Fourier integral operator formula) in different vector spaces, for example in the Besicovitch space $B_2(\mathbb{R}^d)$. The space $B_2(\mathbb{R}^d)$ is the space of all formal sums

$$\sum_{j=1}^{\infty} a_j \mathbf{e}_{\boldsymbol{\theta}_j}(\mathbf{x}),$$

where

$$\mathbf{e}_{\boldsymbol{\theta}}(\mathbf{x}) := e^{i\boldsymbol{\theta}\mathbf{x}} \tag{1.10}$$

and $\sum_{j=1}^{\infty} |a_j|^2 < +\infty$. It is known (see [16]) that the spectra of A acting in $L_2(\mathbb{R}^d)$ and $B_2(\mathbb{R}^d)$ are the same, although the types of those 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 $B_2(\mathbb{R}^d)$, although in the end we will return to operators acting in $L_2(\mathbb{R}^d)$. This trick (working with operators acting in $B_2(\mathbb{R}^d)$) is similar to working with fibre operators $A(\mathbf{k})$ in the periodic case, in the sense that we can freely consider the action of an operator on one, or finitely many, exponentials, without caring that these exponentials do not belong to our function space.

It seems likely that the approach of this paper can be applied to a wider class of operators than (1.1). The operators should be of the type $H = H_0 + b$, where H_0 has constant coefficients and b has order smaller than H_0. We plan to consider such operators in a subsequent publication.

Now we describe the structure of this paper. The proof of our main theorem consists of several parts, which are not always immediately related to each other; in particular, there is no natural order in which these parts should be presented. As a result, it is possible to read different sections of our paper in almost arbitrary order. The main principles we were following in determining the actual order of the sections were: trying to postpone the most difficult and technical parts of the proof for as long as possible, and trying to minimize the amount of references to definitions/results stated after the reference. In particular, Section 6 of this paper can be considered as a further general discussion of our approach that we have decided to postpone until the definitions and results of Sections 2–5 have been introduced. In Section 2, we give some basic definitions, formulate the conditions we impose on the potential, and state the main result. In Section 3, we explain why, instead of proving (1.6), it would be sufficient to prove a more general asymptotic formula (3.1) (which includes more powers of λ as well as logarithms). We also explain why it is enough to prove this asymptotic formula not for all large λ, but only for λ inside a fixed interval. The proof of these statements (as well as reasons why we need them) is similar to the corresponding section in [11]. In Section 4, we describe the definition of the density of states based on the operator algebraic constructions and prove several useful properties of $N(\lambda)$ that immediately follow from these constructions. In Section 5, we define resonance regions and prove their properties. The reader who has read several of the papers [19], [10], [1], [11], [12] may have noticed that in each of these papers the construction of the resonance regions is slightly different. The reason is that each time we define these regions, we need to fine tune the definition taking care of the problem we are trying to solve. Our present paper is not an exception, and the construction of the resonance regions in Section 5 is different from the constructions in all papers mentioned above. This new construction will be extremely convenient when we are going to integrate the contribution from individual eigenvalues to the density of states. In Section 6, we describe this procedure of integrating the contribution from individual eigenvalues over the resonance zones in more detail. In Section 7, we introduce the coordinates in each resonance region (or rather we cut each resonance region into pieces and introduce coordinates in each piece). These coordinates are introduced so that the integration, described in Section 6, will be as painless as it possibly can. Each resonance region will have two types of coordinates. The first type is Cartesian coordinates in $\mathfrak{V}$, where $\mathfrak{V}$ is the quasi-lattice subspace generating the resonance region. The second set of coordinates is the shifted polar coordinates in $\mathfrak{V}^{\perp}$. These coordinates are ideologically similar to the shifted polar coordinates we have introduced in [11], but the details are much more complicated now. Starting from Section 7, until the end of Section 10, we will assume that all the regions where the integration takes place are of the simplest possible type (the simplex case). In Sections 8 and 9, we discuss the main tool of this paper, the gauge transform method. A large proportion of the material contained in these two sections is similar to the relevant parts of [12], the only difference being definition (8.3) (we need to change the norm to accommodate it to the case of almost-periodic coefficients) and Lemma 9.3 (this lemma was not required in [12]). In Section 10, we compute the contribution to the density of states from each resonance region and, first, reduce this contribution to the explicit integral (10.36) and then, in Lemma 10.4, prove that this integral admits a decomposition in the powers of λ and logarithms. Finally, in Section 11, we discuss how to reduce integration over the region of arbitrary shape to the simplex case.

Acknowledgement. We are very grateful to Michael Levitin for running computer experiments that eventually led us to the proper understanding of formula (10.18) and to Eugene Shargorodsky for useful discussions. Thanks also go to Gerassimos Barbatis and Sergey Morozov who have read the preliminary version of this paper and made several useful comments and suggestions. The work of the first author was partially supported by the EPSRC grant EP/F029721/1. The second author was partially supported by the NSF grant DMS-0901015.

2. **Preliminaries**

Since our potential b is almost-periodic, it has the Fourier series

$$b(\mathbf{x}) \sim \sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} a_{\boldsymbol{\theta}}\mathbf{e}_{\boldsymbol{\theta}}(\mathbf{x}), \tag{2.1}$$

where

$$\mathbf{e}_{\boldsymbol{\theta}}(\mathbf{x}) := e^{i\boldsymbol{\theta}\mathbf{x}} \tag{2.2}$$

and $\boldsymbol{\Theta}$ is a (countable) set of frequencies. Without loss of generality we assume that $\boldsymbol{\Theta}$ spans $\mathbb{R}^d$, and contains 0 and is symmetric about 0; we also put

$$\boldsymbol{\Theta}_k := \boldsymbol{\Theta} + \boldsymbol{\Theta} + \cdots + \boldsymbol{\Theta} \tag{2.3}$$

(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 nondiscrete (unless the potential b is periodic). The first condition we impose on the potential is

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 as follows. Suppose $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_d \in \boldsymbol{\Theta}_\infty$. 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 potentials, but it becomes meaningful for quasi-periodic potentials. (We call a function quasi-periodic if it is a linear combination of finitely many exponentials.)

The rest of the conditions we have to impose describe how well we can approximate the potential b by means of quasi-periodic functions. 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.

Condition B. Let k be an arbitrary fixed natural number. Then for each sufficiently large real number ρ, there is a finite set $\mathbf{\Theta}(k;\rho) \subset (\mathbf{\Theta} \cap B(\rho^{1/k}))$ (where $B(r)$ is a ball of radius r centered at 0) and a 'cut-off' potential

$$b_{(k;\rho)}(\mathbf{x}) := \sum_{\boldsymbol{\theta} \in \mathbf{\Theta}(k;\rho)} \tilde{a}_{\boldsymbol{\theta}} \mathbf{e}_{\boldsymbol{\theta}}(\mathbf{x}) \tag{2.4}$$

which satisfies

$$||b - b_{(k;\rho)}||_{\infty} < \rho^{-k}. \tag{2.5}$$

Remark 2.1. First of all, notice that we can reformulate this condition as follows. For each (small) $\alpha > 0$ and $\epsilon \in (0, \epsilon_0(\alpha)]$, there is a 'cut-off' potential $b_{(\alpha,\epsilon)}$ so that $||b - b_{(\alpha;\epsilon)}||_{\infty} < \epsilon$ and the frequencies of $b_{(\alpha;\epsilon)}$ lie inside the ball of radius $\epsilon^{-\alpha}$. However, it will be rather more convenient in what follows to have Condition B formulated in terms of k and ρ. This condition is obviously satisfied for quasi-periodic potentials; for periodic potentials, it is equivalent to the infinite smoothness. For almost-periodic potentials Condition B does not seem to follow from the infinite smoothness of b. Note that we do not require the coefficients $\tilde{a}_{\boldsymbol{\theta}}$ to be equal to the 'old' coefficients $a_{\boldsymbol{\theta}}$; indeed, sometimes one can find a better approximation by using procedures different from the trivial 'chopping off' of b, like, for example, the Bochner-Fejér summation.

The next condition we need to impose 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 how many terms in (1.6) we want to obtain) and denote $\tilde{\mathbf{\Theta}} := [\mathbf{\Theta}(k;\rho)]_{\tilde{k}}$ (see (2.3) for the notation) and $\tilde{\mathbf{\Theta}}' := \tilde{\mathbf{\Theta}} \setminus \{0\}$. We say that $\mathfrak{V}$ is a quasi-lattice subspace of dimension m if $\mathfrak{V}$ is a linear span of m linear independent vectors $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m$ with $\boldsymbol{\theta}_j \in \tilde{\mathbf{\Theta}}$ for all j. 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 $\boldsymbol{\xi}_{\mathfrak{V}^\perp} = \boldsymbol{\xi} - \boldsymbol{\xi}_{\mathfrak{V}}$. 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 positive if and only if $\mathfrak{V}$ and $\mathfrak{W}$ are strongly distinct. We put $s = s(\rho) = s(\tilde{\boldsymbol{\Theta}}) := \inf \sin(\phi(\mathfrak{V}, \mathfrak{U}))$, where infimum is over all strongly distinct pairs of subspaces from $\mathcal{V}$, $R = R(\rho) := \sup_{\boldsymbol{\theta} \in \tilde{\boldsymbol{\Theta}}} |\boldsymbol{\theta}|$, and $r = r(\rho) := \inf_{\boldsymbol{\theta} \in \tilde{\boldsymbol{\Theta}}'} |\boldsymbol{\theta}|$. Obviously, $R(\rho) \ll \rho^{1/k}$ (where the implied constant can depend on k and $\tilde{k}$; we say that $f \ll g$ if $f = O(g)$).

Condition C. For each fixed k and $\tilde{k}$, the sets $\boldsymbol{\Theta}(k; \rho)$ satisfying (2.4) and (2.5) can be chosen in such a way that for sufficiently large ρ, we have

$$s(\rho) \geq \rho^{-1/k} \tag{2.6}$$

and

$$r(\rho) \geq \rho^{-1/k}, \tag{2.7}$$

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

Remark 2.2. First of all, we remark that condition (2.7) for $\tilde{k}$ can be derived from condition (2.6) for $\tilde{k}+1$, but we prefer to postulate both conditions. We also note that Condition C is automatically satisfied for quasi-periodic potentials; for smooth periodic potentials, Condition C is also automatically satisfied (see, for example, [10]). Finally, notice that condition (2.6) is equivalent to $s(\rho) \geq \rho^{-\alpha/k}$ for any fixed positive α. (Indeed, this equivalence can be proved by considering sets $\boldsymbol{\Theta}([\alpha^{-1}k]; \rho)$ instead of $\boldsymbol{\Theta}(k; \rho)$ in Condition B, since Condition B holds for all k.) Thus, if we consider potentials of the form $b = b_{\text{per}} + b_{\text{qua-per}}$, where b_{per} is smooth periodic and $b_{\text{qua-per}}$ is quasi-periodic, Condition C amounts to the Diophantine condition on the frequencies of $b_{\text{qua-per}}$ and is generic.

Condition A implies the following statement. Suppose $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l \in \tilde{\boldsymbol{\Theta}}$, $l \leq d-1$. Let $\mathfrak{V}$ be the span of $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l$. Then each element of the set $\tilde{\boldsymbol{\Theta}} \cap \mathfrak{V}$ is a linear combination of $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l$ with rational coefficients. Since the set $\tilde{\boldsymbol{\Theta}} \cap \mathfrak{V}$ is finite, this implies that the set $Z(\tilde{\boldsymbol{\Theta}} \cap \mathfrak{V})$ is discrete and is, therefore, a lattice in $\mathfrak{V}$. We denote this lattice by $\Gamma(\rho; \mathfrak{V})$. Our final condition states that this lattice cannot be too dense.

Condition D. We can choose $\boldsymbol{\Theta}(k; \rho)$ satisfying conditions B and C in such a way that for sufficiently large ρ and for each $\mathfrak{V} \in \mathcal{V}$, $\mathfrak{V} \neq \mathbb{R}^d$, we have

$$\operatorname{vol}(\mathfrak{V}/\Gamma(\rho; \mathfrak{V})) \geq \rho^{-1/k}. \tag{2.8}$$

Remark 2.3. As with Condition C, Condition D is satisfied for quasi-periodic and smooth periodic potentials. Also, similarly to Remark 2.2, condition (2.8) is equivalent to $\mathrm{vol}(\mathfrak{V}/\Gamma(\rho;\mathfrak{V})) \geq \rho^{-\alpha/k}$. Condition D is not essential for our methods, and it is likely that this condition can be relaxed. Indeed, the only place we are using this condition is to get an upper bound on the number of elements in $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$, and this estimate, in turn, is used only to prove (10.18). However, it seems likely that there may be another way of establishing (10.18). This alternative proof is more direct and much more difficult technically. Given that our paper is quite technically involved the way it is now, we have decided to present a proof that is considerably simpler, paying the price of assuming a slightly stronger condition on the potential.

Remark 2.4. One final remark that concerns all Conditions B–D. Given any symmetric set $\boldsymbol{\Theta}$ of frequencies, we can construct a real smooth almost-periodic potential b such that (2.1) holds, all Fourier coefficients $a_{\boldsymbol{\theta}}$ are nonzero, and Conditions B–D are satisfied. (Of course, the Fourier coefficients will have to converge to zero really fast.) For example, if b is a limit-periodic function with Fourier coefficients going to zero exponentially, then all our Conditions A–D are satisfied.

Now we can formulate our main theorem.

THEOREM 2.5. *Let H be an operator* (1.1) *with real smooth almost-periodic potential b satisfying Conditions* A, B, C, *and* D. *Then for each $K \in \mathbb{N}$, we have*

$$N(\lambda) = \lambda^{d/2}\left(C_d + \sum_{j=1}^{K} e_j\lambda^{-j} + o(\lambda^{-K})\right) \tag{2.9}$$

as $\lambda \to \infty$.

Remark 2.6. Following [3], [4], and [8], it is straightforward to compute the coefficients e_j. For example, we have

$$e_1 = -\frac{dw_d}{2(2\pi)^d}\mathbf{M}(b)$$

and

$$e_2 = \frac{d(d-2)w_d}{8(2\pi)^d}\mathbf{M}(b^2),$$

where $\mathbf{M}$ is the mean of an almost-periodic function.

From now on, we always assume that our potential satisfies all the conditions from this section; we also will denote $\rho := \sqrt{\lambda}$. Given Conditions B–D, we want to introduce the following definition. We say that a positive function $f = f(\rho) = f(\rho; k, \tilde{k})$ satisfies the estimate $f(\rho) \leq \rho^{0+}$ (resp. $f(\rho) \geq \rho^{0-}$) if for each positive ε and for each $\tilde{k}$, we can achieve $f(\rho) \leq \rho^{\varepsilon}$ (resp. $f(\rho) \geq \rho^{-\varepsilon}$)

for sufficiently large ρ by choosing parameter k from Conditions B–D sufficiently large. For example, we have $R(\rho) \le \rho^{0+}$, $s(\rho) \ge \rho^{0-}$, $r(\rho) \ge \rho^{0-}$, and $\mathrm{vol}(\mathfrak{V}/\Gamma(\rho;\mathfrak{V})) \ge \rho^{0-}$. One can also use a standard covering argument to show that the number of elements in $\mathbf{\Theta}(k;\rho)$ satisfies $|\mathbf{\Theta}(k;\rho)| \le \rho^{0+}$. This and (2.5) imply that without loss of generality we can assume that the 'new' Fourier coefficients $\tilde{a}_{\boldsymbol{\theta}}$ can be chosen to be equal to the 'old' coefficients $a_{\boldsymbol{\theta}}$, and we will always assume this in what follows. We will also assume that the value of k is chosen sufficiently large so that all inequalities of the form $\rho^{0+} \le \rho^{\varepsilon}$ or $\rho^{0-} \ge \rho^{-\varepsilon}$ we encounter in the proof are satisfied.

Remark 2.7. As we mention several times in this paper, the enormous amount of notation one has to keep in mind while reading it represents the very big problem for both the authors and the readers. The above definition is the first step towards our aim of making a substantial part of the notation obsolete and eventually to stop using it.

The next statement shows a bit more how this new notation is used.

LEMMA 2.8. *Suppose $\boldsymbol{\theta}, \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_d \in \tilde{\mathbf{\Theta}}'$, the set $\{\boldsymbol{\mu}_j\}$ is linearly independent, and $\boldsymbol{\theta} = \sum_{j=1}^d b_j \boldsymbol{\mu}_j$. Then each nonzero coefficient b_j satisfies*

$$\rho^{0-} \le |b_j| \le \rho^{0+}. \tag{2.10}$$

Proof. Let $\mathfrak{V} \in \mathcal{V}_{d-1}$ be a subspace spanned by $\boldsymbol{\mu}_j$, $j = 2, \dots, d$, and let $\mathbf{e}$ be a unit vector orthogonal to $\mathfrak{V}$. Then the sine of the angle between $\boldsymbol{\theta}$ and $\mathfrak{V}$ is $\langle \boldsymbol{\theta}, \mathbf{e}\rangle |\boldsymbol{\theta}|^{-1}$. Thus, if this angle is nonzero, we have $|\langle \boldsymbol{\theta}, \mathbf{e}\rangle| \ge s(\rho) r(\rho)$ and, hence, if $b_1 = \langle \boldsymbol{\theta}, \mathbf{e}\rangle \langle \boldsymbol{\mu}_1, \mathbf{e}\rangle^{-1}$ is nonzero, it satisfies $|b_1| \ge r(\rho)s(\rho)R(\rho)^{-1} \ge \rho^{0-}$. Similarly, since $\langle \boldsymbol{\mu}_1, \mathbf{e}\rangle \ne 0$, we have $|\langle \boldsymbol{\mu}_1, \mathbf{e}\rangle| \ge s(\rho) r(\rho)$, and thus $|b_1| \le R(\rho)(r(\rho)s(\rho))^{-1} \le \rho^{0+}$. The proof for $j \ne 1$ is similar. □

In this paper, by C or c we denote positive constants, the exact value of which can be different each time they occur in the text, possibly even each time they occur in the same formula. On the other hand, the constants that are labeled (like C_1, c_3, etc) have their values being fixed throughout the text. Given two positive functions f and g, we say that $f \gg g$, or $g \ll f$, or $g = O(f)$ if the ratio $\frac{g}{f}$ is bounded. We say $f \asymp g$ if $f \gg g$ and $f \ll g$.

3. Reduction to a finite interval of spectral parameter

The main result of our paper, Theorem 2.5, will follow from the following theorem. (Recall that we put $\rho := \sqrt{\lambda}$.)

THEOREM 3.1. *For each $K \in \mathbb{N}$, we have*

$$N(\rho^2) = C_d \rho^d + \sum_{p=0}^{d-1} \sum_{j=-d+1}^{K} e_{j,p} \rho^{-j} (\ln \rho)^p + o(\rho^{-K}) \tag{3.1}$$

as $\rho \to \infty$.

Once the theorem is proved, it immediately implies

Corollary 3.2. *For each $K \in \mathbb{N}$, we have*

$$N(\lambda) = \lambda^{d/2}\left(C_d + \sum_{j=1}^{K} e_j \lambda^{-j} + o(\lambda^{-K})\right) \tag{3.2}$$

as $\lambda \to \infty$.

Proof. First of all, we notice that [3], [4], and formula (2.9) from [17] imply that

$$\int_{-\infty}^{\infty} e^{-t\lambda} N(\lambda) d\lambda \sim t^{-(d+2)/2} \sum_{j=0}^{\infty} q_j t^j \tag{3.3}$$

as $t \to 0+$, where q_j are constants depending on the potential. Now the corollary follows from Theorem 3.1 and calculations similar to those of [8]. Indeed, consider the following integrals:

$$I_1(t;k,p) := \int_1^{\infty} e^{-t\lambda} \lambda^{-k} (\ln \lambda)^p d\lambda, \quad p \in \mathbb{Z}_+, \ k \in \mathbb{Z}, \tag{3.4}$$

$$I_2(t;k,p) := \int_1^{\infty} e^{-t\lambda} \lambda^{-k-\frac{1}{2}} (\ln \lambda)^p d\lambda, \quad p \in \mathbb{Z}_+, \ k \in \mathbb{Z}. \tag{3.5}$$

Elementary calculations show that

$$I_1(t;k,p) = t^{k-1}\left(\Gamma(-k+1)\left(\ln\frac{1}{t}\right)^p + \sum_{j=0}^{p-1} a_j \left(\ln\frac{1}{t}\right)^j\right) + f_1(t) \quad \text{for } k \le 0, \ t > 0, \tag{3.6}$$

$$I_1(t;k,p) = t^{k-1}\left(\frac{1}{p+1}\frac{(-1)^{k-1}}{(k-1)!}\left(\ln\frac{1}{t}\right)^{p+1} + \sum_{j=0}^{p} a'_j \left(\ln\frac{1}{t}\right)^j\right) + f_2(t) \quad \text{for } k \ge 1, \ t > 0, \tag{3.7}$$

$$I_2(t;k,p) = t^{k-\frac{1}{2}}\left(\Gamma(-k+\frac{1}{2})\left(\ln\frac{1}{t}\right)^p + \sum_{j=0}^{p-1} a''_j \left(\ln\frac{1}{t}\right)^j\right) + f_3(t) \quad \text{for any } k \in \mathbb{Z}, \ t > 0. \tag{3.8}$$

Here, $a_j = a_j(k,p)$, $a'_j = a'_j(k,p)$, $a''_j = a''_j(k,p)$ are some constants and $f_j(t) = f_j(t;k,p)$ are *entire* functions in t. Obviously, $\int_{-\infty}^{1} e^{-t\lambda} N(\lambda) d\lambda$ is an entire function in t. Comparing (3.3) and (3.6), (3.7), (3.8), it is not difficult to see that

1. if d is even, then $e_{j,p}$ can be nonzero only if $p = 0$ and j is nonpositive and even;
2. if d is odd, then $e_{j,p}$ can be nonzero only if $p = 0$ and j is odd. $\square$

Thus, we can concentrate on proving Theorem 3.1. To begin with, we choose sufficiently large $\rho_0 > 1$ (to be fixed later on) and put $\rho_n = 2\rho_{n-1} = 2^n\rho_0$, $\lambda_n := \rho_n^2$; we also define the interval $I_n = [\rho_n, 4\rho_n]$. The proof of Theorem 3.1 will be based on the following lemma.

LEMMA 3.3. *For each $M \in \mathbb{N}$ and $\rho \in I_n$, we have*

$$N(\rho^2) = C_d\rho^d + \sum_{p=0}^{d-1}\sum_{j=-d+1}^{6M} e_{j,p}(n)\rho^{-j}(\ln\rho)^p + O(\rho_n^{-M}). \tag{3.9}$$

Here, $e_{j,p}(n)$ are some real numbers depending on j, p, and n (and M) satisfying

$$e_{j,p}(n) = O(\rho_n^{(2j/3)+a}). \tag{3.10}$$

The constants in the O-terms do not depend on n (but they may depend on M). The value of a does not depend on either n or M.

Remark 3.4. Note that (3.9) is not a 'proper' asymptotic formula, since the coefficients $e_{j,p}(n)$ are allowed to grow with n (and, therefore, with ρ).

Let us prove Theorem 3.1 assuming that we have proved Lemma 3.3. Let M be fixed. Denote

$$N_n(\rho^2) := C_d\rho^d + \sum_{p=0}^{d-1}\sum_{j=-d+1}^{6M} e_{j,p}(n)\rho^{-j}(\ln\rho)^p. \tag{3.11}$$

Then, whenever $\rho \in J_n := I_{n-1} \cap I_n = [\rho_n, 2\rho_n]$, we have

$$N_n(\rho^2) - N_{n-1}(\rho^2) = \sum_{p=0}^{d-1}\sum_{j=-d+1}^{6M} t_{j,p}(n)\rho^{-j}(\ln\rho)^p, \tag{3.12}$$

where

$$t_{j,p}(n) := e_{j,p}(n) - e_{j,p}(n-1). \tag{3.13}$$

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

$$\sum_{p=0}^{d-1}\sum_{j=-d+1}^{6M} t_{j,p}(n)\rho^{-j}(\ln\rho)^p = O(\rho_n^{-M}).$$

CLAIM 3.5. *For each $j = -d+1, \dots, 6M$, we have*

$$t_{j,p}(n) = O(\rho_n^{j-M}(\ln\rho_n)^{d-1-p}).$$

Proof. Put $x := \rho^{-1}$. Then $\sum_{p=0}^{d-1}\sum_{j=-d+1}^{6M} t_{j,p}(n)x^j(-1)^p(\ln x)^p = O(\rho_n^{-M})$ whenever $x \in [\frac{\rho_n^{-1}}{2}, \rho_n^{-1}]$. Put $y := x\rho_n$ and

$$\tau_{j,p}(n) := \rho_n^{M-j}\sum_{s=p}^{d-1}\binom{s}{p}(-1)^p t_{j,s}(n)(\ln\rho_n)^{s-p}.$$

Then

$$P(y) := \sum_{p=0}^{d-1} \sum_{j=-d+1}^{6M} \tau_{j,p}(n) y^j (\ln y)^p = O(1) \tag{3.14}$$

whenever $y \in [\frac{1}{2}, 1]$. Consider the following $d(6M+d)$ functions: $y^j(\ln y)^p$ ($j = -d+1, \dots, 6M$, $p = 0, \dots, d-1$), and label them $h_1(y), \dots, h_{d(6M+d)}(y)$. These functions are linearly independent on the interval $[\frac{1}{2}, 1]$. Therefore, there exist points $y_1, \dots, y_{d(6M+d)} \in [\frac{1}{2}, 1]$ such that the determinant of the matrix $(h_j(y_l))_{j,l=1}^{d(6M+d)}$ is nonzero. Now (3.14) and the Cramer's Rule imply that for each j, the values $\tau_{j,p}(n)$ are fractions with a bounded expression in the numerator and a fixed nonzero number in the denominator. Therefore, $\tau_{j,p}(n) = O(1)$. This shows first that $t_{j,d-1}(n) = O(\rho_n^{j-M})$ and then, subsequently reducing index p from $p = d-1$ to $p = 0$, we obtain $t_{j,p}(n) = O(\rho_n^{j-M}(\ln \rho_n)^{d-1-p})$ as claimed. □

Thus, for $j < M$, the series $\sum_{m=0}^{\infty} t_{j,p}(m)$ is absolutely convergent; moreover, for such j, we have

(3.15)

$$\begin{aligned} e_{j,p}(n) = e_{j,p}(0) + \sum_{m=1}^{n} t_{j,p}(m) &= e_{j,p}(0) + \sum_{m=1}^{\infty} t_{j,p}(m) + O(\rho_n^{j-M}(\ln \rho_n)^{d-1-p}) \\ &=: e_{j,p} + O(\rho_n^{j-M}(\ln \rho_n)^{d-1-p}), \end{aligned}$$

where we have denoted $e_{j,p} := e_{j,p}(0) + \sum_{m=1}^{\infty} t_{j,p}(m)$.

Since $e_{j,p}(n) = O(\rho_n^{(2j/3)+a})$ (it was one of the assumptions of lemma), we have

$$\sum_{j=M}^{6M} |e_{j,p}(n)| \rho_n^{-j} = O\left(\rho_n^{a-\frac{M}{3}}\right) = O\left(\rho_n^{-\frac{M}{4}}\right), \tag{3.16}$$

assuming as we can without loss of generality that M is sufficiently large. Thus, when $\rho \in I_n$, we have

(3.17)

$$N(\rho^2) = C_d\rho^d + \sum_{p=0}^{d-1} \sum_{j=-d+1}^{M-1} e_{j,p}\rho^{-j}(\ln \rho)^p + O(\rho^{-M}(\ln \rho)^{d-1}) + O(\rho^{-\frac{M}{4}}(\ln \rho)^{d-1}).$$

Since constants in O terms do not depend on n, for all $\rho \geq \rho_0$, we have

$$\begin{aligned} N(\rho^2) &= C_d\rho^d + \sum_{p=0}^{d-1} \sum_{j=-d+1}^{M-1} e_{j,p}\rho^{-j}(\ln \rho)^p + O(\rho^{-\frac{M}{6}}) \\ &= C_d\rho^d + \sum_{p=0}^{d-1} \sum_{j=-d+1}^{[M/6]} e_{j,p}\rho^{-j}(\ln \rho)^p + O(\rho^{-\frac{M}{6}}). \end{aligned} \tag{3.18}$$

Taking $M = 6K+1$, we obtain (3.1).

The rest of the paper is devoted to proving Lemma 3.3. We will mostly concentrate on obtaining formula (3.9), since estimate (3.10) will usually follow by trivial but tedious arguments (like estimating coefficients in the product of several geometric series). However, in the cases when estimating the coefficients in our infinite series would present difficulties, we will carry out these estimates as well. 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 in order to satisfy estimate (9.18). (This estimate says that the more asymptotic terms we want to have in (3.9), 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 $\rho_n^{0+} < \rho_n^{\epsilon}$ or $\rho_n^{0-} > \rho_n^{-\epsilon}$); each time we have such an inequality, we assume that k is chosen sufficiently large to satisfy it. The first requirement on k we have is that $k > M$. After fixing n and k, we choose the finite set $\mathbf{\Theta}(k;\rho_n)$ and the approximating potential $b_{(k;\rho_n)}$ that satisfy all the Conditions B–D. Then condition (2.5) and definition (1.2) imply that the difference between the densities of states of operators with potentials b and $b_{(k;\rho_n)}$ is smaller than ρ_n^{-M}. Thus, from now on we will consider the operator with a potential $b_{(k;\rho_n)}$ and try to establish (3.9) for this new operator. Following our policy of getting rid of all indexes as soon as possible (i.e., immediately after we have fixed them), we will denote $\mathbf{\Theta} := \mathbf{\Theta}(k;\rho_n)$ and $b := b_{(k;\rho_n)}$. This means that from now on we will assume that b is a quasi-periodic potential with $\mathbf{\Theta}$ its spectrum of frequencies so that $\mathbf{\Theta}$ satisfies Conditions A–D with $\rho = \rho_n$.

4. **Abstract results**

In this section, we establish several abstract results concerning density of states for operators with almost-periodic coefficients. In the periodic setting, these results become either trivial or already known, so the reader who is mostly interested in the periodic case can skip this section.

In this and further sections, we will work with pseudo-differential operators with almost-periodic coefficients (or symbols). These operators were studied in [16] and [17]. In Section 8, we will introduce the classes of such operators. We will also see that one can naturally consider the action of such operators in both $L_2(\mathbb{R}^d)$ and $B_2(\mathbb{R}^d)$. These actions have many similarities between them. In particular, the norms and (for elliptic or bounded operators) the spectra of operators acting in $L_2(\mathbb{R}^d)$ and $B_2(\mathbb{R}^d)$ are the same; see [16]. As a result, often when we discuss a pseudo-differential operator with almost-periodic coefficients, we do not specify in which space it acts. Sometimes, however, it becomes important to emphasize the space where the operator acts, in which case we will do this.

Following [17], we denote by $\mathfrak{A}_B$ a II_∞ factor acting in $\tilde{\mathfrak{H}} := B_2(\mathbb{R}^d) \otimes L_2(\mathbb{R}^d)$. We denote by $\mathbf{e}_{\boldsymbol{\xi}}$ both the function $e^{i\boldsymbol{\xi}\mathbf{x}}$ and the operator of multiplication by this function. $T_{\boldsymbol{\xi}}$ is the operator of translation by $\boldsymbol{\xi}$ in $L_2(\mathbb{R}^d)$, i.e., $T_{\boldsymbol{\xi}}u(\mathbf{x}) = u(\mathbf{x}-\boldsymbol{\xi})$. The factor $\mathfrak{A}_B$ is defined as the von Neumann algebra acting in $\tilde{\mathfrak{H}}$ generated by two families of operators:

$$\{\mathbf{e}_{\boldsymbol{\xi}} \otimes \mathbf{e}_{\boldsymbol{\xi}},\, \boldsymbol{\xi} \in \mathbb{R}^d\} \tag{4.1}$$

and

$$\{I \otimes T_{\boldsymbol{\xi}},\, \boldsymbol{\xi} \in \mathbb{R}^d\}. \tag{4.2}$$

Let $A = a(\mathbf{x}, D)$ be a self-adjoint pseudo-differential operator with almost-periodic coefficients such that $a(\mathbf{x}, \boldsymbol{\xi}) \gg |\boldsymbol{\xi}|^m$ for some $m > 0$. We introduce operator $A^\sharp := a(\mathbf{x}+\mathbf{y}, D_{\mathbf{y}})$ acting in $\tilde{\mathfrak{H}}$; here, $\mathbf{x}$ is a variable of functions in $B_2(\mathbb{R}^d)$ and $\mathbf{y}$ is a variable of functions in $L_2(\mathbb{R}^d)$. We denote by $E_\lambda(A)$ the spectral projection of A; by $\tilde{E}_\lambda(A)$ we denote the spectral projection of $A^\sharp$. By $\mathbf{D}$ and $\mathbf{T}$ we denote the relative dimension and the relative trace in $\mathfrak{A}_B$ (see [9]).

If A is actually a *differential* operator, then (see [17]) one can define the density of states of A (denoted by $N(\lambda; A)$) by formula (1.2). It was also proved in [17] that

$$N(\lambda; A) = \mathbf{T}(\tilde{E}_\lambda(A^\sharp)) = \mathbf{D}(\tilde{E}_\lambda(A^\sharp)\tilde{\mathfrak{H}}). \tag{4.3}$$

Note that the relative dimension in a II_∞ factor can take any nonnegative value. Now it is natural to define the density of states for a general elliptic self-adjoint *pseudo-differential* operator with almost-periodic coefficients by (4.3).

By $\mathfrak{L}$ we denote a closed linear subspace of $\tilde{\mathfrak{H}}$ adjoint to $\mathfrak{A}_B$ (see [9] for the explanation of the terminology). The following lemma gives a variational description of the density of states.

Lemma 4.1.

$$N(\lambda; A) = \sup\{\mathbf{D}(\mathfrak{L}),\ (A^\sharp\phi, \phi) \le \lambda(\phi, \phi), \forall \phi \in \mathfrak{L}\}. \tag{4.4}$$

Proof. By taking $\mathfrak{L} := \tilde{E}_\lambda(A^\sharp)\tilde{\mathfrak{H}}$ and using (4.3), we see that the left-hand side of (4.4) is at most the right-hand side. Suppose now that we have found a subspace $\mathfrak{L}$ such that $\mathbf{D}(\mathfrak{L}) > \mathbf{D}(\tilde{E}_\lambda(A^\sharp)\tilde{\mathfrak{H}})$. Then the lemma from Section VII.37 of [9] implies that $\mathfrak{L}$ contains a nonzero vector ϕ orthogonal to $\tilde{E}_\lambda(A^\sharp)\tilde{\mathfrak{H}}$. But then $(A^\sharp\phi, \phi) > \lambda(\phi, \phi)$, which contradicts our assumption on $\mathfrak{L}$. This proves (4.4). □

Corollary 4.2. *If $A \ge B$, then $N(\lambda; A) \le N(\lambda; B)$.*

Corollary 4.3. *Suppose H_1 and H_2 are two elliptic self-adjoint pseudo-differential operators with almost-periodic coefficients such that $||H_1 - H_2|| \ll \rho_n^{-M+(2-d)}$. Suppose $N(H_2; \rho^2)$ satisfies asymptotic expansion (3.9). Then $N(H_1; \rho^2)$ also satisfies (3.9).*

Proof. Our assumptions imply that $H_2 - \delta \leq H_1 \leq H_2 + \delta$, where $\delta \ll \rho_n^{-M+(2-d)}$. The previous corollary now implies that

$$N(H_2 + \delta; \lambda) \leq N(H_1; \lambda) \leq N(H_2 - \delta; \lambda). \tag{4.5}$$

It remains to notice that if $N(H_2; \rho^2)$ satisfies (3.9), then the difference between the right-hand side and the left-hand side of (4.5) is $O(\rho_n^{-M})$. □

LEMMA 4.4. *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 $N(\lambda; A) = N(\lambda; U^{-1}AU)$.*

Proof. Obviously, operator $U^\sharp$ is unitary and $(U^{-1}AU)^\sharp = (U^\sharp)^{-1}A^\sharp U^\sharp$. Thus,

$$\begin{aligned} N(\lambda; U^{-1}AU) &= \mathbf{T}(\tilde{E}_\lambda((U^\sharp)^{-1}A^\sharp U^\sharp)) \\ &= \mathbf{T}((U^\sharp)^{-1}\tilde{E}_\lambda(A^\sharp)U^\sharp) = \mathbf{T}(\tilde{E}_\lambda(A^\sharp)) = N(\lambda; A). \end{aligned} \tag{4.6}$$

Here, the third equality follows, for example, from Sections 36–37, Chapter 7 of [9]. □

5. Resonance zones

In this section, we define resonance regions and establish some of their properties. Recall the definition of the set $\mathbf{\Theta} = \mathbf{\Theta}(k; \rho_n)$ as well as of the quasi-lattice subspaces from Section 2. As before, by $\mathbf{\Theta}_{\tilde{k}}$ we denote the algebraic sum of $\tilde{k}$ copies of $\mathbf{\Theta}$; remember that we consider the index $\tilde{k}$ fixed. We also put $\mathbf{\Theta}'_{\tilde{k}} := \mathbf{\Theta}_{\tilde{k}} \setminus \{0\}$. For each $\mathfrak{V} \in \mathcal{V}$, we put $S_{\mathfrak{V}} := \{\boldsymbol{\xi} \in \mathfrak{V},\ |\boldsymbol{\xi}| = 1\}$. For each nonzero $\boldsymbol{\theta} \in \mathbb{R}^d$, 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 $\mathfrak{V}_j \in \mathcal{V}_j$ $(j = 0, 1, \ldots, m)$ 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 the multiplication by -1.) We denote by $\mathcal{F}(\mathfrak{V})$ the collection of all flags generated by $\mathfrak{V}$. We also fix an increasing sequence of positive numbers α_j $(j = 1, \ldots, d)$ with $\alpha_d < \frac{1}{2d}$ (these numbers depend only on d) and put $L_j := \rho_n^{\alpha_j}$.

Let $\boldsymbol{\theta} \in \mathbf{\Theta}'_{\tilde{k}}$. We call by *resonance zone* generated by $\boldsymbol{\theta}$

$$\Lambda(\boldsymbol{\theta}) := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ |\langle \boldsymbol{\xi}, \mathbf{n}(\boldsymbol{\theta})\rangle| \leq L_1\}. \tag{5.1}$$

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

$$\Lambda(\mathfrak{F}) := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ |\langle \boldsymbol{\xi}, \boldsymbol{\nu}_j\rangle| \leq L_j\}. \tag{5.2}$$

If $\dim \mathfrak{V} = 1$, definition (5.2) is reduced to (5.1). 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

$$\Xi_1(\mathfrak{V}) := \cup_{\mathfrak{F}\in\mathcal{F}(\mathfrak{V})}\Lambda(\mathfrak{F}). \tag{5.3}$$

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

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

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

Let us establish some basic properties of resonance regions. The first set of properties follows immediately from the definitions.

LEMMA 5.1. (i) *We have*

$$\cup_{\mathfrak{V}\in\mathcal{V}} \Xi(\mathfrak{V}) = \mathbb{R}^d. \tag{5.5}$$

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

(iii) $\Xi_1(\mathbb{R}^d) = \Xi(\mathbb{R}^d)$ *is a bounded set,* $\Xi(\mathbb{R}^d) \subset B(dL_d)$; *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* $(\Xi_1(\mathfrak{V}) \cap \Xi_1(\mathfrak{U})) \subset \Xi_1(\mathfrak{W})$, *where* $\mathfrak{W} := \mathfrak{V} + \mathfrak{U}$ (*algebraic sum*).

Proof. Assume, without loss of generality, that $m_1 := \dim \mathfrak{V} \geq \dim \mathfrak{U} =: m_2$. If $\mathfrak{U} \subset \mathfrak{V}$, then the statement of the lemma is obvious. Consider the case when $\mathfrak{V}$ and $\mathfrak{U}$ are strongly distinct. Suppose $\boldsymbol{\xi} \in (\Xi_1(\mathfrak{V}) \cap \Xi_1(\mathfrak{U}))$. Then there is a flag $\mathfrak{F} \in \mathcal{F}(\mathfrak{V})$ such that $\boldsymbol{\xi} \in \Lambda(\mathfrak{F})$. Let $\mathfrak{F}_1 \in \mathcal{F}(\mathfrak{W})$ be any flag such that the first m_1 elements of $\mathfrak{F}_1$ coincide with $\mathfrak{F}$. Let us prove that $\boldsymbol{\xi} \in \Lambda(\mathfrak{F}_1)$. Let $\{\boldsymbol{\nu}_j\}_{j=1}^m$ be a sequence generated by $\mathfrak{F}_1$ ($m = \dim \mathfrak{W}$). Then the inclusion $\boldsymbol{\xi} \in \Lambda(\mathfrak{F})$ implies that $|\langle \boldsymbol{\xi}, \boldsymbol{\nu}_j\rangle| \leq L_j$ for $j = 1, \ldots, m_1$. Moreover, a simple geometry implies $|\boldsymbol{\xi}_{\mathfrak{W}}| \leq (|\boldsymbol{\xi}_{\mathfrak{V}}| + |\boldsymbol{\xi}_{\mathfrak{U}}|)[\sin(\phi(\mathfrak{V}, \mathfrak{U}))]^{-1} \leq 2m_1 L_{m_1} s(\rho_n)^{-1} < L_{m_1}\rho_n^{0+} < L_{m_1+1}$. Therefore, for $j \geq m_1 + 1$, we have $|\langle \boldsymbol{\xi}, \boldsymbol{\nu}_j\rangle| = |\langle \boldsymbol{\xi}_{\mathfrak{W}}, \boldsymbol{\nu}_j\rangle| \leq |\boldsymbol{\xi}_{\mathfrak{W}}| \leq L_{m_1+1} \leq L_j$. This shows that indeed $\boldsymbol{\xi} \in \Lambda(\mathfrak{F}_1)$ and, therefore, $\boldsymbol{\xi} \in \Xi_1(\mathfrak{W})$, which proves our lemma. $\square$

The next statement follows immediately from Lemma 5.2.

Corollary 5.3. (i) *We can rewrite definition (5.4) as follows*:

$$\boldsymbol{\Xi}(\mathfrak{V}) := \boldsymbol{\Xi}_1(\mathfrak{V}) \setminus (\cup_{\mathfrak{U}\not\subset\mathfrak{V}}\boldsymbol{\Xi}_1(\mathfrak{U})). \tag{5.6}$$

(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* $|\langle \boldsymbol{\xi}, \boldsymbol{\mu}\rangle| = |\langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\mu}\rangle| \leq L_{m+1}$ *holds.*

Proof. In one direction the statement is obvious. Now, we assume that $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{V}) \cap \boldsymbol{\Xi}_1(\mathfrak{W})$. Let $\mathfrak{F} = \{\mathfrak{W}_0, \dots, \mathfrak{W}_m, \mathfrak{W}\}$ be a flag for which $\boldsymbol{\xi} \in \Lambda(\mathfrak{F})$. If $\mathfrak{W}_m = \mathfrak{V}$, then the statement of the lemma is straightforward. Otherwise, we can apply the construction from the proof of Lemma 5.2 with $\mathfrak{U} = \mathfrak{W}_m$. This completes the proof. □

Lemma 5.5. *We have*

$$\boldsymbol{\Xi}_1(\mathfrak{V}) \cap \cup_{\mathfrak{U}\supsetneq\mathfrak{V}}\boldsymbol{\Xi}_1(\mathfrak{U}) = \boldsymbol{\Xi}_1(\mathfrak{V}) \cap \cup_{\mathfrak{W}\supsetneq\mathfrak{V},\dim\mathfrak{W}=1+\dim\mathfrak{V}}\boldsymbol{\Xi}_1(\mathfrak{W}). \tag{5.7}$$

Proof. Indeed, obviously, the right-hand side of (5.7) is a subset of the left-hand side. On the other hand, suppose $\mathfrak{U} \supsetneq \mathfrak{V}$ and $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{V}) \cap \boldsymbol{\Xi}_1(\mathfrak{U})$. Then $\boldsymbol{\xi} \in \Lambda(\mathfrak{F})$ for some $\mathfrak{F} \in \mathcal{F}(\mathfrak{U})$. Suppose that $\mathfrak{V}_1 \in \mathfrak{F}$ is a subspace such that $\dim\mathfrak{V}_1 = \dim\mathfrak{V}$. If $\mathfrak{V}_1 = \mathfrak{V}$, it immediately follows that $\boldsymbol{\xi}$ is contained in the right-hand side of (5.7). Assume that $\mathfrak{V}_1 \neq \mathfrak{V}$; in particular, $\dim\mathfrak{V} \geq 1$. Then there exists $\mathfrak{V}_2 \in \mathfrak{F}$ such that $\dim(\mathfrak{V} + \mathfrak{V}_2) = \dim\mathfrak{V} + 1$. Put $\mathfrak{W} := \mathfrak{V} + \mathfrak{V}_2$. Since $\boldsymbol{\xi} \in \Lambda(\mathfrak{F}) \subset \boldsymbol{\Xi}_1(\mathfrak{V}_2)$, by Lemma 5.2 we have $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{W})$, and so $\boldsymbol{\xi}$ is contained in the right-hand side of (5.7). □

Corollary 5.6. *We can rewrite* (5.4) *as*

$$\boldsymbol{\Xi}(\mathfrak{V}) := \boldsymbol{\Xi}_1(\mathfrak{V}) \setminus (\cup_{\mathfrak{W}\supsetneq\mathfrak{V},\dim\mathfrak{W}=1+\dim\mathfrak{V}}\boldsymbol{\Xi}_1(\mathfrak{W})). \tag{5.8}$$

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* $\boldsymbol{\xi} + \boldsymbol{\theta} \in \boldsymbol{\Xi}(\mathfrak{V})$.

Proof. If $\boldsymbol{\theta} \notin \mathfrak{V}$, then Lemma 5.2 implies that $\boldsymbol{\xi} \in \boldsymbol{\Xi}_1(\mathfrak{W})$, where $\mathfrak{W} = \mathrm{span}(\mathfrak{V}, \boldsymbol{\theta})$, which contradicts our assumption $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})$.

Let us prove that $\boldsymbol{\xi} + \boldsymbol{\theta} \in \boldsymbol{\Xi}_1(\mathfrak{V})$. Since $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V}) \subset \boldsymbol{\Xi}_1(\mathfrak{V})$, this implies that $\boldsymbol{\xi} \in \Lambda(\mathfrak{F})$ with $\mathfrak{F} \in \mathcal{F}(\mathfrak{V})$, $\mathfrak{F} = \{\mathfrak{V}_0 = \mathfrak{X}, \mathfrak{V}_1, \dots, \mathfrak{V}_m = \mathfrak{V}\}$.

Let J be the biggest number such that $\boldsymbol{\theta} \notin \mathfrak{V}_{J-1}$. (Obviously, $J \leq m := \dim\mathfrak{V}$.) We construct a new flag $\mathfrak{F}_1 = \{\mathfrak{U}_0 = \mathfrak{X}, \mathfrak{U}_1, \dots, \mathfrak{U}_m = \mathfrak{V}\}$ such that

$$\mathfrak{U}_j = \begin{cases} \mathfrak{X}, & j = 0, \\ \mathrm{span}(\mathfrak{V}_{j-1}, \boldsymbol{\theta}), & 0 < j \leq J, \\ \mathfrak{V}_j, & j > J. \end{cases}$$

We are going to prove that $\boldsymbol{\xi}+\boldsymbol{\theta}\in\Lambda(\mathfrak{F}_1)$. Let $\{\boldsymbol{\nu}_j\}_{j=1}^m$ be a sequence generated by $\mathfrak{F}_1$. Obviously, for $j>J$, we have $\langle\boldsymbol{\xi}+\boldsymbol{\theta},\boldsymbol{\nu}_j\rangle=\langle\boldsymbol{\xi},\boldsymbol{\nu}_j\rangle$, so that $|\langle\boldsymbol{\xi}+\boldsymbol{\theta},\boldsymbol{\nu}_j\rangle|\le L_j$ if $|\langle\boldsymbol{\xi},\boldsymbol{\nu}_j\rangle|\le L_j$. So, assume that $j\le J$. If $j=1$, we have $\boldsymbol{\nu}_1=\mathbf{n}(\boldsymbol{\theta}_1)$, so the assumption $\boldsymbol{\xi}+\boldsymbol{\theta}\in\Lambda(\boldsymbol{\theta})$ implies that $|\langle\boldsymbol{\xi}+\boldsymbol{\theta},\boldsymbol{\nu}_1\rangle|\le L_1$. Assume now that $1<j\le J$. Then $|\boldsymbol{\xi}_{\mathfrak{U}_j}|\le(|\boldsymbol{\xi}_{\mathfrak{V}_{j-1}}|+|\boldsymbol{\xi}_{\mathfrak{U}_1}|)s(\rho_n)^{-1}\le 2(j-1)L_{j-1}s(\rho_n)^{-1}$. Therefore, $|(\boldsymbol{\xi}+\boldsymbol{\theta})_{\mathfrak{U}_j}|\le 2(j-1)L_{j-1}s(\rho_n)^{-1}+|\boldsymbol{\theta}|\le 2(j-1)L_{j-1}s(\rho_n)^{-1}+R(\rho_n)\le L_j$. Thus, $|\langle\boldsymbol{\xi}+\boldsymbol{\theta},\boldsymbol{\nu}_j\rangle|\le|(\boldsymbol{\xi}+\boldsymbol{\theta})_{\mathfrak{U}_j}|\le L_j$. This shows that, indeed, we have $\boldsymbol{\xi}+\boldsymbol{\theta}\in\Lambda(\mathfrak{F}_1)$ and, therefore, $\boldsymbol{\xi}+\boldsymbol{\theta}\in\boldsymbol{\Xi}_1(\mathfrak{V})$.

Suppose now that $\boldsymbol{\xi}+\boldsymbol{\theta}\notin\boldsymbol{\Xi}(\mathfrak{V})$. This could only happen if $\boldsymbol{\xi}+\boldsymbol{\theta}\in\boldsymbol{\Xi}_1(\mathfrak{W})$ for some $\mathfrak{W}\supsetneq\mathfrak{V}$. But then the previous part of the proof would imply that $\boldsymbol{\xi}\in\boldsymbol{\Xi}_1(\mathfrak{W})$, which contradicts our assumption $\boldsymbol{\xi}\in\boldsymbol{\Xi}(\mathfrak{V})$. Thus, $\boldsymbol{\xi}+\boldsymbol{\theta}\in\boldsymbol{\Xi}(\mathfrak{V})$, which finishes the proof. □

The next definition is almost identical to the corresponding definition from [12].

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.

(1) 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 $\boldsymbol{\xi}\leftrightarrow\boldsymbol{\eta} \mod \boldsymbol{\theta}$.

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

(3) 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 $j=1,2,\dots,l$.

(4) 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 $\boldsymbol{\xi}\in\boldsymbol{\Xi}(\mathfrak{V})$, $\mathfrak{V}\in\mathcal{V}_m$.*

Proof. This follows from Lemmas 5.7 and 5.1. □

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* $\operatorname{card} \boldsymbol{\Upsilon}(\boldsymbol{\xi}) \ll \rho_n^{(d-1)\alpha_{d-1}+0+}$ *uniformly in* $\boldsymbol{\xi} \in \mathbb{R}^d \setminus \boldsymbol{\Xi}(\mathbb{R}^d)$.

Proof. This immediately follows from Lemmas 5.1, 5.7, 5.10, Conditions A and D, and a standard covering argument. □

6. Description of the approach

For any set $\mathcal{C} \subset \mathbb{R}^d$ by $\mathcal{P}(\mathcal{C})$ we denote the orthogonal projection onto $\operatorname{span}\{\mathbf{e}_{\boldsymbol{\xi}}\}_{\boldsymbol{\xi}\in\mathcal{C}}$ in $B_2(\mathbb{R}^d)$ and by $\mathcal{P}^L(\mathcal{C})$ the same projection considered in $L_2(\mathbb{R}^d)$, i.e.,

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

where $\mathcal{F}$ is the Fourier transform and $\chi_{\mathcal{C}}$ is the operator of multiplication by the characteristic function of $\mathcal{C}$. Obviously, $\mathcal{P}^L(\mathcal{C})$ is a well defined (resp. nonzero) projection if and only if $\mathcal{C}$ is measurable (resp. has nonzero measure). We also denote $\mathfrak{H} := B_2(\mathbb{R}^d)$. Let us fix sufficiently large n and denote (recall that $\lambda_n = \rho_n^2$)

$$\mathcal{X}_n := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ |\boldsymbol{\xi}|^2 \in [0.7\lambda_n, 17.5\lambda_n]\}. \tag{6.2}$$

We also put

$$\mathcal{A} = \mathcal{A}_n := \cup_{\boldsymbol{\xi}\in\mathcal{X}_n} \boldsymbol{\Upsilon}(\boldsymbol{\xi}). \tag{6.3}$$

Lemma 5.10 implies that for each $\boldsymbol{\xi} \in \mathcal{A}$, we have $|\boldsymbol{\xi}|^2 \in [0.5\lambda_n, 18\lambda_n]$. In particular, we have

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

For each $\mathfrak{V} \in \mathcal{V}_m$, $m < d$, we put

$$\mathcal{A}(\mathfrak{V}) := \mathcal{A}_n \cap \boldsymbol{\Xi}(\mathfrak{V}). \tag{6.5}$$

We also denote

$$\hat{\mathcal{A}} := \{\boldsymbol{\xi} \notin \mathcal{A},\ |\boldsymbol{\xi}|^2 < \lambda_n\} \tag{6.6}$$

and

$$\check{\mathcal{A}} := \{\boldsymbol{\xi} \notin \mathcal{A},\ |\boldsymbol{\xi}|^2 > \lambda_n\}. \tag{6.7}$$

We plan to apply the gauge transform similar to the one used in [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 4.4 implies that the densities of states of H and H_1 are the same.) Moreover, $H_1 = H_2 + \mathcal{R}$, where $||\mathcal{R}|| \ll \rho_n^{-M+(2-d)}$ and $H_2 = -\Delta + W$ is a self-adjoint pseudo-differential operator with symbol

$|\boldsymbol{\xi}|^2+w$ that satisfies the following property (see Section 8 for more discussion about pseudo-differential operators and their symbols):
(6.8)
$$\hat{w}(\boldsymbol{\theta},\boldsymbol{\xi})=0,\ \text{if } (\boldsymbol{\xi}\not\in\Lambda(\boldsymbol{\theta})\ \&\ \boldsymbol{\xi}\in\mathcal{A}),\ \text{or } (\boldsymbol{\xi}+\boldsymbol{\theta}\not\in\Lambda(\boldsymbol{\theta})\ \&\ \boldsymbol{\xi}\in\mathcal{A}),\ \text{or } (\boldsymbol{\theta}\not\in\boldsymbol{\Theta}_{\tilde{k}}).$$

Now Corollary 4.3 implies that if we prove that $N(\rho^2;H_2)$ satisfies (3.9), then $N(\rho^2;H_1)$ (and therefore $N(\rho^2;H)$) satisfies the same asymptotic formula. This means that it is enough to establish the asymptotic expansion (3.9) for the operator H_2 instead of H. Condition (6.8) implies that for each $\boldsymbol{\xi}\in\mathcal{A}$, the subspace $\mathcal{P}(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))\mathfrak{H}$ is an invariant subspace of H_2 (acting, remember, in $B_2(\mathbb{R}^d)$); we denote its dimension by m (which is finite by Lemma 5.11). We put

$$H_2(\boldsymbol{\Upsilon}(\boldsymbol{\xi})):=H_2\Big|_{\boldsymbol{\Upsilon}(\boldsymbol{\xi})\mathfrak{H}}. \tag{6.9}$$

Note that the subspaces $\mathcal{P}(\hat{\mathcal{A}})\mathfrak{H}$ and $\mathcal{P}(\check{\mathcal{A}})\mathfrak{H}$ 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 $\mathcal{P}(\mathcal{A})\mathfrak{H}$. Also notice that if we consider the operator H_2 acting in $L_2(\mathbb{R}^d)$, then $\mathcal{P}^L(\hat{\mathcal{A}})L_2(\mathbb{R}^d)$, $\mathcal{P}^L(\check{\mathcal{A}})L_2(\mathbb{R}^d)$, and $\mathcal{P}^L(\mathcal{A})L_2(\mathbb{R}^d)$ would still be invariant subspaces. For each $\boldsymbol{\xi}\in\mathcal{A}$, the operator $H_2(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))$ is a finite-dimensional self-adjoint operator, so its spectrum is purely discrete; we denote its eigenvalues (counting multiplicities) by $\lambda_1(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))\le\lambda_2(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))\le\cdots\le\lambda_m(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))$ and the corresponding orthonormalized eigenfunctions by $\{h_{j,\boldsymbol{\Upsilon}(\boldsymbol{\xi})}(\mathbf{x})\}$. 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 $\boldsymbol{\eta}=(\eta_1,\dots,\eta_d)$ and $\boldsymbol{\eta}'=(\eta_1',\dots,\eta_d')$ 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 mapping $g:\mathcal{A}\to\mathbb{R}$ that puts into correspondence to each point $\boldsymbol{\eta}\in\mathcal{A}$ the number $\lambda_{t(\boldsymbol{\eta})}(\boldsymbol{\Upsilon}(\boldsymbol{\eta}))$. This mapping 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 $B_2(\mathbb{R}^d)$, so there is nothing miraculous about its spectrum consisting of eigenvalues and their limit points.) Moreover, all eigenvalues of H_2 inside the interval $[0.75\lambda_n,17\lambda_n]$ have a pre-image under g. Arguments, similar to the ones used in [12], show that g is a measurable function. Similarly, we define the mapping $h:\mathcal{A}\to B_2(\mathbb{R}^d)$ by the formula $h_{\boldsymbol{\xi}}:=h_{t(\boldsymbol{\xi}),\boldsymbol{\Upsilon}(\boldsymbol{\xi})}$. Then for each $\boldsymbol{\xi}\in\mathcal{A}$, the expression $(2\pi)^{-d}\sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})}h_{\boldsymbol{\eta}}(\mathbf{x})\overline{h_{\boldsymbol{\eta}}(\mathbf{y})}$ is the integral kernel of the projection $\mathcal{P}(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))$. Therefore, we have

$$\sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})}h_{\boldsymbol{\eta}}(\mathbf{x})\overline{h_{\boldsymbol{\eta}}(\mathbf{y})}=\sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})}\mathbf{e}_{\boldsymbol{\eta}}(\mathbf{x})\overline{\mathbf{e}_{\boldsymbol{\eta}}(\mathbf{y})}. \tag{6.10}$$

Another, perhaps slightly simpler, way of establishing (6.10) is just to notice that $\vec{\mathbf{h}} = F\vec{\mathbf{e}}$, where $\vec{\mathbf{h}}$ is a column-vector with entries $\{h_{\boldsymbol{\eta}}\}_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})}$, $\vec{\mathbf{e}}$ is a column-vector with entries $\{\mathbf{e}_{\boldsymbol{\eta}}\}_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})}$, and F is a unitary matrix. Then

$$\sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})} h_{\boldsymbol{\eta}}(\mathbf{x})\overline{h_{\boldsymbol{\eta}}(\mathbf{y})} = \vec{\mathbf{h}}(\mathbf{x})^T\overline{\vec{\mathbf{h}}(\mathbf{y})} = \vec{\mathbf{e}}(\mathbf{x})^T F^T \overline{F}\overline{\vec{\mathbf{e}}(\mathbf{y})} \tag{6.11}$$
$$= \vec{\mathbf{e}}(\mathbf{x})^T\overline{\vec{\mathbf{e}}(\mathbf{y})} = \sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})} \mathbf{e}_{\boldsymbol{\eta}}(\mathbf{x})\overline{\mathbf{e}_{\boldsymbol{\eta}}(\mathbf{y})}.$$

When $\boldsymbol{\xi} \notin \mathcal{A}$, we put $g(\boldsymbol{\xi}) := |\boldsymbol{\xi}|^2$ and $h_{\boldsymbol{\xi}} := \mathbf{e}_{\boldsymbol{\xi}}$, so that now the functions g and h are defined on all $\mathbb{R}^d$. It follows from the construction that $\{h_{\boldsymbol{\xi}}\}_{\boldsymbol{\xi}\in\mathbb{R}^d}$ is an orthonormal basis in $B_2(\mathbb{R}^d)$.

All this implies that for each $\lambda \in [0.75\lambda_n, 17\lambda_n]$, the function

$$e(\lambda;\mathbf{x},\mathbf{y}) := (2\pi)^{-d}\int_{G_\lambda} h_{\boldsymbol{\xi}}(\mathbf{x})\overline{h_{\boldsymbol{\xi}}(\mathbf{y})}d\boldsymbol{\xi},\ \mathbf{x},\mathbf{y} \in \mathbb{R}^d \tag{6.12}$$

is the integral kernel of the spectral projection $E_\lambda(H_2; B_2(\mathbb{R}^d))$ of the operator H_2 in $B_2(\mathbb{R}^d)$; here, we have denoted

$$G_\lambda := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ g(\boldsymbol{\xi}) \le \lambda\}. \tag{6.13}$$

Notice that $e(\lambda;\mathbf{x},\mathbf{y})$ also gives the kernel of the spectral projection of the operator H_2 considered in $L_2(\mathbb{R}^d)$. Since this is the statement we will use in our proof, let us give a little bit more detailed proof of it. We define a mapping

$$U:\ f \mapsto (2\pi)^{-d/2}\int_{\mathbb{R}^d}\overline{h_{\boldsymbol{\xi}}(\mathbf{x})}f(\mathbf{x})d\mathbf{x}$$

and

$$\mathcal{M} := \mathcal{A}/\leftrightarrow, \tag{6.14}$$

where $\leftrightarrow$ is the equivalence relation introduced in Definition 5.8; Lemma 5.11 and property (6.4) imply that $\mathcal{M}$ is measurable. It is not hard to see that U is a unitary operator in $L_2(\mathbb{R}^d)$ and

$$U^*:\ z \mapsto (2\pi)^{-d/2}\int_{\mathbb{R}^d} h_{\boldsymbol{\xi}}(\mathbf{x})z(\boldsymbol{\xi})d\boldsymbol{\xi}.$$

Indeed, we have (recall the notation (6.14) and identity (6.10))

$$U^*Uf(\mathbf{x}) = (2\pi)^{-d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_{\boldsymbol{\xi}}(\mathbf{x})\overline{h_{\boldsymbol{\xi}}(\mathbf{y})}f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi} \tag{6.15}$$
$$= (2\pi)^{-d}\Big(\int_{\hat{\mathcal{A}}} + \int_{\check{\mathcal{A}}} + \int_{\mathcal{A}}\Big)\int_{\mathbb{R}^d} h_{\boldsymbol{\xi}}(\mathbf{x})\overline{h_{\boldsymbol{\xi}}(\mathbf{y})}f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi}$$
$$= (2\pi)^{-d}\Big(\int_{\hat{\mathcal{A}}} + \int_{\check{\mathcal{A}}}\Big)\int_{\mathbb{R}^d} h_{\boldsymbol{\xi}}(\mathbf{x})\overline{h_{\boldsymbol{\xi}}(\mathbf{y})}f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi}$$
$$+ (2\pi)^{-d}\int_{\mathcal{M}}\int_{\mathbb{R}^d}\sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})} h_{\boldsymbol{\eta}}(\mathbf{x})\overline{h_{\boldsymbol{\eta}}(\mathbf{y})}f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi}$$

$$= (2\pi)^{-d}\Big(\int_{\hat{\mathcal{A}}} + \int_{\check{\mathcal{A}}}\Big) \int_{\mathbb{R}^d} \mathbf{e}_{\boldsymbol{\xi}}(\mathbf{x})\overline{\mathbf{e}_{\boldsymbol{\xi}}(\mathbf{y})} f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi}$$
$$+ (2\pi)^{-d} \int_{\mathcal{M}} \int_{\mathbb{R}^d} \sum_{\boldsymbol{\eta}\in\boldsymbol{\Upsilon}(\boldsymbol{\xi})} \mathbf{e}_{\boldsymbol{\eta}}(\mathbf{x})\overline{\mathbf{e}_{\boldsymbol{\eta}}(\mathbf{y})} f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi}$$
$$= (2\pi)^{-d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{e}_{\boldsymbol{\xi}}(\mathbf{x})\overline{\mathbf{e}_{\boldsymbol{\xi}}(\mathbf{y})} f(\mathbf{y})d\mathbf{y}d\boldsymbol{\xi} = f(\mathbf{x}).$$

Now we notice that for $\lambda \in [0.75\lambda_n, 17\lambda_n]$, the function $e(\lambda; \mathbf{x}, \mathbf{y})$ is the kernel of $U^*\chi_{G_\lambda}U$. Moreover, $UH_2(\mathcal{A})U^*$ is the operator of multiplication by g acting in $L_2(\mathcal{A})$. It remains to notice that the other two restrictions of H_2 satisfy $UH_2(\hat{\mathcal{A}})U^* < 0.75\lambda_n I$ and $UH_2(\check{\mathcal{A}})U^* > 17\lambda_n I$. Now it immediately follows that for $\lambda \in [\lambda_n, 16\lambda_n]$, $e(\lambda; \mathbf{x}, \mathbf{y})$ is the kernel of the spectral projection of the operator H_2 considered in $L_2(\mathbb{R}^d)$.

Now (4.3) and Theorem 4.1 from [17] (see (1.9)) imply the following result. (Note that since g is a measurable function, G_λ is a measurable set.)

LEMMA 6.1. *For $\lambda \in [\lambda_n, 16\lambda_n]$ being a continuity point of $N(\lambda; H_2)$, we have*

$$N(\lambda; H_2) = (2\pi)^{-d} \operatorname{vol}(G_\lambda). \tag{6.16}$$

Proof. For the proof, it is enough to notice that $|h_{\boldsymbol{\xi}}(\mathbf{x})| = |\mathbf{e}_{\boldsymbol{\xi}}(\mathbf{x})| = 1$ for $\boldsymbol{\xi} \notin \mathcal{A}$ and

$$|h_{\boldsymbol{\xi}}(\mathbf{x})|^2 \le \operatorname{card} \boldsymbol{\Upsilon}(\boldsymbol{\xi}) \ll \rho_n^{(d-1)\alpha_{d-1}+0+}$$

for $\boldsymbol{\xi} \in \mathcal{A}$ by (6.10) and Lemma 5.11 and apply Lebesgue's Limit Theorem. □

Since points of continuity of $N(\lambda)$ are dense, *the asymptotic expansion proven for such λ can be extended to all $\lambda \in [\lambda_n, 16\lambda_n]$ by taking the limit.* Thus, our next task is to compute $\operatorname{vol}(G_\lambda)$. Let us put

$$\hat{A}^+ := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ g(\boldsymbol{\xi}) < \rho^2 < |\boldsymbol{\xi}|^2\} \tag{6.17}$$

and

$$\hat{A}^- := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ |\boldsymbol{\xi}|^2 < \rho^2 < g(\boldsymbol{\xi})\}. \tag{6.18}$$

LEMMA 6.2. *The following identity holds*:

$$\operatorname{vol}(G_\lambda) = w_d\rho^d + \operatorname{vol}\hat{A}^+ - \operatorname{vol}\hat{A}^-. \tag{6.19}$$

Proof. We obviously have $G_\lambda = B(\rho) \cup \hat{A}^+ \setminus \hat{A}^-$. Since $\hat{A}^- \subset B(\rho)$ and $\hat{A}^+ \cap B(\rho) = \emptyset$, this implies (6.19). □

Remark 6.3. Properties of the mapping g imply that we have $\hat{A}^+,\ \hat{A}^- \subset \mathcal{A}$. Thus, in order to compute $N(\lambda)$, we need to analyze the behaviour of g only inside $\mathcal{A}$.

We will compute volumes of $\hat{A}^{\pm}$ 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 $\operatorname{vol}\hat{A}^{\pm}$. Namely, we are going to introduce a convenient set of coordinates in $\Xi(\mathfrak{V})$. Let $\mathfrak{V} \in \mathcal{V}_m$ be fixed; since $\hat{A}^{\pm} \cap \Xi(\mathbb{R}^d) = \emptyset$, we will assume that $m < d$. Then, as we have seen, $\boldsymbol{\xi} \in \Xi_1(\mathfrak{V})$ 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 $\boldsymbol{\xi} \in \Xi_1(\mathfrak{V})$, we have $\boldsymbol{\xi} \in \Xi_1(\mathfrak{U}_j)$ if and only if the estimate $|\langle \boldsymbol{\xi}, \boldsymbol{\mu}_j\rangle| = |\langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\mu}_j\rangle| \le L_{m+1}$ holds. Thus, formula (5.8) implies that

$$\Xi(\mathfrak{V}) = \{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})\ \&\ \forall j\ |\langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\mu}_j(\mathfrak{V})\rangle| > L_{m+1}\}. \tag{7.1}$$

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

$$\{\mathbf{n}(\boldsymbol{\theta}_{\mathfrak{V}^\perp}),\ \boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}} \setminus \mathfrak{V}\}. \tag{7.2}$$

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

$$\Xi(\mathfrak{V})_p = \{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V})\ \&\ \forall j\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > L_{m+1}\}; \tag{7.3}$$

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

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

has a nonempty intersection with the boundary of $\Xi(\mathfrak{V})_p$. It is not hard to see that $J_p \ge d - m$. Indeed, otherwise $\Xi(\mathfrak{V})_p$ would have nonempty intersection with $\Xi_1(\mathfrak{V}')$ for some $\mathfrak{V}'$, $\mathfrak{V} \subsetneq \mathfrak{V}'$. We also introduce

$$\tilde{\Xi}(\mathfrak{V})_p := \{\boldsymbol{\xi} \in \mathfrak{V}^\perp,\ \forall j\ \langle \boldsymbol{\xi}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > 0\}. \tag{7.4}$$

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

$$\langle \boldsymbol{\xi}, \boldsymbol{\theta}\rangle = \langle \boldsymbol{\xi}, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle \ne 0. \tag{7.5}$$

We also put $K := d - m - 1$.

Let us first assume that the number J_p of 'defining planes' is the minimal possible, i.e., $J_p = K + 1$. We will carry on this assumption throughout most of the paper, and only in Section 11 will we discuss how to deal with the more general case of arbitrary $\Xi(\mathfrak{V})_p$. 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: $\langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_j(p)\rangle = L_{m+1}$, $j = 1, \ldots, K+1$. Then, since the determinant of the Gram matrix of vectors $\tilde{\boldsymbol{\mu}}_j(p)$ is $\gg \rho_n^{0-}$, we have $|\mathbf{a}| \ll L_{m+1}\rho_n^{0+}$. We introduce the shifted cylindrical coordinates in $\boldsymbol{\Xi}(\mathfrak{V})_p$. These coordinates will be denoted by $\boldsymbol{\xi} = (r; \tilde{\Phi}; X)$. Here, $X = (X_1, \ldots, X_m)$ is an arbitrary set of cartesian coordinates in $\Omega(\mathfrak{V})$. These coordinates do not depend on the choice of the connected component $\boldsymbol{\Xi}(\mathfrak{V})_p$. The rest of the coordinates $(r, \tilde{\Phi})$ are shifted spherical coordinates in $\mathfrak{V}^\perp$, centered at $\mathbf{a}$. This means that

$$r(\boldsymbol{\xi}) = |\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}| \tag{7.6}$$

and

$$\tilde{\Phi} = \mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}) \in S_{\mathfrak{V}^\perp}. \tag{7.7}$$

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

(7.8)
$$\begin{aligned} M_p = \{\mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}),\ \boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})_p\} &= \{\mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}),\ \forall j\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > L_{m+1}\} \\ = \{\mathbf{n}(\boldsymbol{\eta}),\ \boldsymbol{\eta} := \boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a} \in \mathfrak{V}^\perp,\ \forall j\ \langle \boldsymbol{\eta}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > 0\} &= S_{\mathfrak{V}^\perp} \cap \tilde{\boldsymbol{\Xi}}(\mathfrak{V})_p. \end{aligned}$$

We will denote by $d\tilde{\Phi}$ the spherical Lebesgue measure on M_p. For each nonzero vector $\boldsymbol{\mu} \in \mathfrak{V}^\perp$, we denote

$$W(\boldsymbol{\mu}) := \{\boldsymbol{\eta} \in \mathfrak{V}^\perp,\ \langle \boldsymbol{\eta}, \boldsymbol{\mu}\rangle = 0\}. \tag{7.9}$$

Thus, the sides of the simplex M_p are intersections of $W(\tilde{\boldsymbol{\mu}}_j(p))$ with the sphere $S_{\mathfrak{V}^\perp}$. Each vertex $\mathbf{v} = \mathbf{v}_t$, $t = 1, \ldots, K+1$ of M_p is an intersection of $S_{\mathfrak{V}^\perp}$ with K hyperplanes $W(\tilde{\boldsymbol{\mu}}_j(p))$, $j = 1, \ldots, K+1$, $j \neq t$. This means that $\mathbf{v}_t$ is a unit vector from $\mathfrak{V}^\perp$ that is orthogonal to $\{\tilde{\boldsymbol{\mu}}_j(p)\}$, $j = 1, \ldots, K+1$, $j \neq t$; 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 nonzero angles between two sides of any dimensions of 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 $s(\rho_n)$.*

Proof. First of all, we remark that $\mathfrak{U}_j$ are not, in general, quasi-lattice subspaces. However, each algebraic sum $\mathfrak{W}_j := \mathfrak{V} + \mathfrak{U}_j$ is a quasi-lattice subspace. Moreover, the angle between $\mathfrak{W}_1$ and $\mathfrak{W}_2$ is equal to the angle between $\mathfrak{U}_1$ and $\mathfrak{U}_2$, so the first statement follows from Condition C. To prove the second statement, it is enough to notice that any nonzero angle between two sides (of arbitrary dimension) of M_p is equal to the angle between two subspaces $\mathfrak{U}_1$ and $\mathfrak{U}_2$ of the type considered in the first statement; the same can be said about the distance between $\mathbf{v}_t$ and $\mathbf{v}_\tau$. □

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

Proof. Suppose not. Then, without loss of generality, we can assume that $b_1 > 0$ and $b_2 < 0$. Let L be a spherical interval joining $\mathbf{v}_1$ and $\mathbf{v}_2$, i.e.,

$$L = \{\mathbf{u} \in \overline{M_p}, \langle \mathbf{u}, \tilde{\boldsymbol{\mu}}_j(p)\rangle = 0, \ j = 3, \dots, K+1\}. \tag{7.10}$$

Note that $\langle \mathbf{v}_1, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle = b_1\langle \mathbf{v}_1, \tilde{\boldsymbol{\mu}}_1(p)\rangle > 0$ and $\langle \mathbf{v}_2, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle = b_2\langle \mathbf{v}_2, \tilde{\boldsymbol{\mu}}_2(p)\rangle < 0$. Therefore, there is a point $\mathbf{u} \in L$ such that $\langle \mathbf{u}, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle = 0$. This means that $W(\boldsymbol{\theta}_{\mathfrak{V}^\perp})$ has a nonempty intersection with M_p, which contradicts (7.5). ☐

Assume that the diameter of M_p is $\leq (100d^2)^{-1}$, which we can always achieve by taking sufficiently large $\tilde{k}$. We put $\Phi_q := \frac{\pi}{2} - \phi(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}, \tilde{\boldsymbol{\mu}}_q(p))$, $q = 1, \dots, K+1$. The geometrical meaning of these coordinates is simple: Φ_q is the spherical distance between $\tilde{\Phi} = \mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a})$ and $W(\tilde{\boldsymbol{\mu}}_q(p))$. 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 in Section 10. At the same time, the set of coordinates $(r, \{\Phi_q\})$ 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 $\{\mathbf{e}_j\}$, $j = 1, \dots, K+1$ be a fixed orthonormal basis in $\mathfrak{V}^\perp$ chosen in such a way that the $K+1$-st axis passes through M_p. Then we have $\mathbf{e}_j = \sum_{l=1}^{K+1} a_{jl}\tilde{\boldsymbol{\mu}}_l$ 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 $\boldsymbol{\eta} := \boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}$),

$$\eta_j = \langle \boldsymbol{\eta}, \mathbf{e}_j\rangle = r \sum_{q=1}^{K+1} a_{jq} \sin \Phi_q \tag{7.11}$$

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

$$\sum_j \Big(\sum_q a_{jq} \sin \Phi_q\Big)^2 = 1, \tag{7.12}$$

which is our constraint.

Let us also put

$$\eta_j' := \frac{\eta_j}{|\boldsymbol{\eta}|} = \sum_{q=1}^{K+1} a_{jq} \sin \Phi_q. \tag{7.13}$$

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

$$d\tilde{\Phi} = \frac{d\eta_1' \cdots d\eta_K'}{\eta_{K+1}} = \frac{d\eta_1' \cdots d\eta_K'}{(1 - \sum_{j=1}^{K} (\eta_j')^2)^{1/2}}, \tag{7.14}$$

where the denominator is bounded below by $1/2$ by our choice of the basis $\{\mathbf{e}_j\}$.

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

Proof. This follows from the fact that for each p, $|a_{pl}|$ is the length of the projection of $\mathbf{e}_p$ onto $\tilde{\boldsymbol{\mu}}_l$ parallel to the linear space spanned by all $\tilde{\boldsymbol{\mu}}_j$, $j \neq l$. Since the absolute value of the sine of the angle between $\tilde{\boldsymbol{\mu}}_l$ and the linear space spanned by all $\tilde{\boldsymbol{\mu}}_j$, $j \neq l$, is at least $s(\rho_n)$, this implies that for each l, p, we have $|a_{pl}| \leq s(\rho_n)^{-1}$, which finishes the proof. □

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

Proof. Suppose not. Then for each l, we have $\sin \Phi_l(\boldsymbol{\eta}) < s(\rho_n) d^{-3/2}$. Since all $\sin \Phi_l$ are positive, Lemma 7.3 implies

$$\sum_j \left(\sum_l a_{jl} \sin \Phi_l \right)^2 < d(ds(\rho_n)^{-1} s(\rho_n) d^{-3/2})^2 = 1,$$

which contradicts (7.12). □

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* $\boldsymbol{\xi} \in \boldsymbol{\Xi}(\mathfrak{V})_p$, $\mathfrak{V} \in \mathcal{V}_m$, *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* $\boldsymbol{\theta}_{\mathfrak{V}^\perp} = \sum_q b_q \tilde{\boldsymbol{\mu}}_q(p)$, *then*

$$\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle = \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + L_{m+1} \sum_q b_q + r(\boldsymbol{\xi}) \sum_q b_q \sin \Phi_q. \tag{7.15}$$

In case (ii), *all the coefficients* b_q *are either nonpositive or nonnegative and each nonzero coefficient* b_q *satisfies*

$$\rho_n^{0-} \leq |b_q| \leq \rho_n^{0+}. \tag{7.16}$$

Proof. We begin by noticing that

$$\langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle = \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \boldsymbol{\theta}_{\mathfrak{V}^\perp} \rangle, \tag{7.17}$$

from which part (i) immediately follows. Recalling that $\boldsymbol{\xi}_{\mathfrak{V}^\perp} = \mathbf{a} + \boldsymbol{\eta}$, we obtain

$$\begin{aligned} \langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle &= \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \langle \mathbf{a}, \boldsymbol{\theta}_{\mathfrak{V}^\perp} \rangle + \langle \boldsymbol{\eta}, \boldsymbol{\theta}_{\mathfrak{V}^\perp} \rangle \\ &= \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \sum_q b_q \langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_q \rangle + \sum_q b_q \langle \boldsymbol{\eta}, \tilde{\boldsymbol{\mu}}_q \rangle \\ &= \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \sum_q b_q L_{m+1} + r \sum_q b_q \sin \Phi_q. \end{aligned} \tag{7.18}$$

The last statement follows from Lemmas 7.2 and 2.8. The application of Lemma 7.2 is straightforward, so let us discuss the application of Lemma 2.8. Suppose $\boldsymbol{\theta}_{\mathfrak{V}^\perp} = \sum_{q=1}^{K+1} b_q \tilde{\boldsymbol{\mu}}_q(p)$, where $\boldsymbol{\theta}$ belongs to $\boldsymbol{\Theta}_{\tilde{k}}$ (but $\tilde{\boldsymbol{\mu}}_q(p)$, in general, is not in $\boldsymbol{\Theta}_{\tilde{k}}$). We know that the linear span of each $\tilde{\boldsymbol{\mu}}_q(p)$ and $\mathfrak{V}$ is an element of $\mathcal{V}_{m+1}$. Therefore, for each $q = 1, \ldots, K+1$, there is a vector $\boldsymbol{\nu}_q \in \boldsymbol{\Theta}_{\tilde{k}}$ such that $\tilde{\boldsymbol{\mu}}_q(p)$ is proportional to $(\boldsymbol{\nu}_q)_{\mathfrak{V}^\perp}$, $\tilde{\boldsymbol{\mu}}_q(p) = C(q)(\boldsymbol{\nu}_q)_{\mathfrak{V}^\perp}$, where $\rho_n^{0-} \leq |C(q)| \leq$

ρ_n^{0+}. Now we choose arbitrary linearly independent vectors $\boldsymbol{\nu}_{K+2}, \dots, \boldsymbol{\nu}_d \in (\mathfrak{V} \cap \boldsymbol{\Theta}_{\tilde{k}})$. Then we can write $\boldsymbol{\theta} = \boldsymbol{\theta}_{\mathfrak{V}^\perp} + \boldsymbol{\theta}_{\mathfrak{V}} = \sum_{q=1}^{K+1} C(q) b_q (\boldsymbol{\nu}_q)_{\mathfrak{V}^\perp} + \sum_{q=K+2}^{d} b_q \boldsymbol{\nu}_q$. Now we can apply Lemma 2.8 directly. □

8. Pseudo-differential operators

In this and the next sections, we construct operators H_1 and H_2 described in Section 6. Most of the material in these two sections is very similar to the corresponding sections of [12], as are the proofs of most of the statements. Therefore, we will often omit the proofs, instead referring the reader to [20], [19], and [12].

8.1. *Classes of* PDO*'s.* Before we define the pseudo-differential operators (PDO's), we introduce the relevant classes of symbols.

For any $f \in L_2(\mathbb{R}^d)$, we define the Fourier transform:

$$(\mathcal{F}f)(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\xi}\mathbf{x}} f(\mathbf{x}) d\mathbf{x}, \ \boldsymbol{\xi} \in \mathbb{R}^d.$$

Let us now define the symbols we will consider and operators associated with them. 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, i.e., for some countable set $\hat{\boldsymbol{\Theta}}$ of frequencies (we always assume $\hat{\boldsymbol{\Theta}}$ to be symmetric and to contain 0; starting from the middle of this section, the set $\hat{\boldsymbol{\Theta}}$ will be assumed to be finite),

$$b(\mathbf{x}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\theta} \in \hat{\boldsymbol{\Theta}}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \mathbf{e}_{\boldsymbol{\theta}}(\mathbf{x}) \tag{8.1}$$

where

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

are Fourier coefficients of b. (Recall that $\mathbf{M}$ is the mean of an almost-periodic function.) We always assume that (8.1) converges absolutely. Put $\langle \mathbf{t} \rangle := \sqrt{1 + |\mathbf{t}|^2}$ for all $\mathbf{t} \in \mathbb{R}^d$. We notice that

$$\langle \boldsymbol{\xi} + \boldsymbol{\eta} \rangle \le 2 \langle \boldsymbol{\xi} \rangle \langle \boldsymbol{\eta} \rangle, \ \forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^d. \tag{8.2}$$

We say that the symbol b belongs to the class $\mathbf{S}_\alpha = \mathbf{S}_\alpha(\beta) = \mathbf{S}_\alpha(\beta, \hat{\boldsymbol{\Theta}})$, $\alpha \in \mathbb{R}$, $0 < \beta \le 1$, if for any $l \ge 0$ and any nonnegative $s \in \mathbb{Z}$, the condition

$$|b|_{l,s}^{(\alpha)} := \max_{|\mathbf{s}| \le s} \sum_{\boldsymbol{\theta} \in \hat{\boldsymbol{\Theta}}} \langle \boldsymbol{\theta} \rangle^l \sup_{\boldsymbol{\xi}} \langle \boldsymbol{\xi} \rangle^{(-\alpha + |\mathbf{s}|)\beta} |\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})| < \infty, \quad |\mathbf{s}| = s_1 + s_2 + \dots + s_d \tag{8.3}$$

is fulfilled. The quantities (8.3) 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 $\alpha < \gamma$. For later reference we write here the following convenient bound that follows from definition (8.3) and property (8.2):

$$\sum_{\boldsymbol{\theta}\in\hat{\boldsymbol{\Theta}}} \langle\boldsymbol{\theta}\rangle^l \sup_{\boldsymbol{\xi}} \langle\boldsymbol{\xi}\rangle^{(-\alpha+s+1)\beta}(|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\eta}) - \mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})|)| \\ \le C \mathbf{|} b \mathbf{|}_{l,s+1}^{(\alpha)} \langle\boldsymbol{\eta}\rangle^{|\alpha-s-1|\beta}|\boldsymbol{\eta}|, \ s=|\mathbf{s}|, \tag{8.4}$$

with a constant C depending only on α, s, and β. For a vector $\boldsymbol{\eta} \in \mathbb{R}^d$, introduce the symbol

$$b_{\boldsymbol{\eta}}(\mathbf{x},\boldsymbol{\xi}) = b(\mathbf{x},\boldsymbol{\xi}+\boldsymbol{\eta}), \boldsymbol{\eta} \in \mathbb{R}^d, \tag{8.5}$$

so that $\hat{b}_{\boldsymbol{\eta}}(\boldsymbol{\theta},\boldsymbol{\xi}) = \hat{b}(\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\eta})$. The bound (8.4) implies that for all $|\boldsymbol{\eta}| \le C$, we have

$$\mathbf{|} b - b_{\boldsymbol{\eta}} \mathbf{|}_{l,s}^{(\alpha-1)} \le C \mathbf{|} b \mathbf{|}_{l,s+1}^{(\alpha)} |\boldsymbol{\eta}|, \tag{8.6}$$

uniformly in $\boldsymbol{\eta}$: $|\boldsymbol{\eta}| \le C$.

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

$$\mathrm{Op}(b)u(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int b(\mathbf{x},\boldsymbol{\xi}) e^{i\boldsymbol{\xi}\mathbf{x}}(\mathcal{F}u)(\boldsymbol{\xi})d\boldsymbol{\xi}, \tag{8.7}$$

the integral being over $\mathbb{R}^d$. Under the condition $b \in \mathbf{S}_\alpha$, the integral in the right-hand side is clearly finite for any u from the Schwarz class $\mathsf{S}(\mathbb{R}^d)$. Moreover, the condition $b \in \mathbf{S}_0$ guarantees the boundedness of $\mathrm{Op}(b)$ in $L_2(\mathbb{R}^d)$; see Proposition 8.1. Unless otherwise stated, from now on $\mathsf{S}(\mathbb{R}^d)$ is taken as a natural domain for all PDO's at hand, when they act in L_2. Applying the standard regularization procedures to definition (8.7) (see, e.g., [16]), we can also consider the action of $\mathrm{Op}(b)$ when we apply it to an exponential $\mathbf{e}_{\boldsymbol{\nu}}$. Then, we have

$$\mathrm{Op}(b)\mathbf{e}_{\boldsymbol{\nu}} = \sum_{\boldsymbol{\theta}\in\hat{\boldsymbol{\Theta}}} \hat{b}(\boldsymbol{\theta},\boldsymbol{\nu})\mathbf{e}_{\boldsymbol{\nu}+\boldsymbol{\theta}}. \tag{8.8}$$

This action can be extended by linearity to all quasi-periodic functions (i.e., finite linear combinations of $\mathbf{e}_{\boldsymbol{\nu}}$). Moreover, if the order $\alpha = 0$, by continuity this action can be extended to all of $B_2(\mathbb{R}^d)$; this extension has the same norm as $\mathrm{Op}(b)$ acting in L_2 (see [16]). 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 whole B_2 (when the order is nonpositive), 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 $L_2(\mathbb{R}^d)$ or $B_2(\mathbb{R}^d)$, since these norms are the same. Notice that the operator $\mathrm{Op}(b)$ is symmetric if its symbol satisfies the condition

$$\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi}) = \overline{\hat{b}(-\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\theta})}. \tag{8.9}$$

We shall call such symbols *symmetric*.

We note that in the very beginning when we consider (1.1), our operator $\mathrm{Op}(b)$ is a multiplication by a function b (in particular, $b \in \mathbf{S}_0$). However, during modifications and transformations below, our perturbation will eventually become a pseudo-differential operator. Thus, it is convenient in abstract statements to consider b a pseudo-differential symbol from some $\mathbf{S}_\alpha$ class.

8.2. *Some basic results on the calculus of almost-periodic* PDO's. We begin by listing some elementary results for almost-periodic PDO's. The proof is very similar (with obvious changes) to the proof of analogous statements in [20]. 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 added terms are assumed to have zero coefficients in front of them.*

PROPOSITION 8.1. *Suppose that $\left|b\right|_{0,0}^{(0)} < \infty$. Then $\mathrm{Op}(b)$ is bounded in both $L_2(\mathbb{R}^d)$ and $B_2(\mathbb{R}^d)$ and $\|\mathrm{Op}(b)\| \le \left|b\right|_{0,0}^{(0)}$.*

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,

$$\widehat{(b \circ g)}(\boldsymbol{\chi}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\theta}+\boldsymbol{\phi}=\boldsymbol{\chi}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\phi}) \hat{g}(\boldsymbol{\phi}, \boldsymbol{\xi}),\ \boldsymbol{\chi} \in \hat{\boldsymbol{\Theta}}_1 + \hat{\boldsymbol{\Theta}}_2,\ \boldsymbol{\xi} \in \mathbb{R}^d. \tag{8.10}$$

We have

PROPOSITION 8.2. *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*

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

with a constant C depending only on l, α, s.

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

$$\mathrm{ad}(A; \Psi_1, \Psi_2, \dots, \Psi_N) = i\Big[\mathrm{ad}(A; \Psi_1, \Psi_2, \dots, \Psi_{N-1}), \Psi_N\Big],$$
$$\mathrm{ad}(A; \Psi) = i[A, \Psi],\quad \mathrm{ad}^N(A; \Psi) = \mathrm{ad}(A; \Psi, \Psi, \dots, \Psi),\ \mathrm{ad}^0(A; \Psi) = A.$$

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. It follows from (8.10) that the Fourier coefficients of the symbol $\mathrm{ad}(b,g)$ are given by

$$\widehat{\mathrm{ad}(b,g)}(\boldsymbol{\chi},\boldsymbol{\xi}) = i \sum_{\boldsymbol{\theta}+\boldsymbol{\phi}=\boldsymbol{\chi}} \Big[\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\phi})\hat{g}(\boldsymbol{\phi},\boldsymbol{\xi}) - \hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\hat{g}(\boldsymbol{\phi},\boldsymbol{\xi}+\boldsymbol{\theta})\Big], \ \ \boldsymbol{\xi}\in\mathbb{R}^d. \tag{8.11}$$

PROPOSITION 8.3. *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*

$$\mathrm{ad}(b; g_1,\dots,g_N) \in \mathbf{S}_\gamma\left(\hat{\boldsymbol{\Theta}} + \sum_j \hat{\boldsymbol{\Theta}}_j\right)$$

with

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

and

$$\big|\mathrm{ad}(b; g_1,\dots,g_N)\big|^{(\gamma)}_{l,s} \le C \big|b\big|^{(\alpha)}_{p,s+N} \prod_{j=1}^N \big|g_j\big|^{(\gamma_j)}_{p,s+N-j+1}, \tag{8.12}$$

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

8.3. *Partition of the perturbation.* From now on we fix $\beta :\ 0<\beta<\alpha_1$, and put $\hat{\boldsymbol{\Theta}} := \boldsymbol{\Theta}$, which is finite. The symbols we are going to construct will depend on ρ_n; this dependence will usually be omitted from the notation.

Let $\iota \in \mathsf{C}^\infty(\mathbb{R})$ be a nonnegative function such that

$$0 \le \iota \le 1, \quad \iota(z) = \begin{cases} 1, & z \le \frac{1}{4}, \\ 0, & z \ge \frac{1.1}{4}. \end{cases} \tag{8.13}$$

For $\boldsymbol{\theta}\in\boldsymbol{\Theta}, \boldsymbol{\theta}\ne\mathbf{0}$, define the following C^∞-cut-off functions:

$$\begin{cases} e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \iota\left(\left|\dfrac{|\boldsymbol{\xi}+\boldsymbol{\theta}/2| - 3\rho_n}{10\rho_n}\right|\right), \\ \ell^>_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = 1 - \iota\left(\dfrac{|\boldsymbol{\xi}+\boldsymbol{\theta}/2| - 3\rho_n}{10\rho_n}\right), \\ \ell^<_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = 1 - \iota\left(\dfrac{3\rho_n - |\boldsymbol{\xi}+\boldsymbol{\theta}/2|}{10\rho_n}\right) \end{cases} \tag{8.14}$$

and

$$\begin{cases} \zeta_\theta(\boldsymbol{\xi}) = \iota\left(\dfrac{|\langle\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\theta}/2\rangle|}{\rho_n^\beta|\boldsymbol{\theta}|}\right), \\ \varphi_\theta(\boldsymbol{\xi}) = 1 - \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}). \end{cases} \tag{8.15}$$

Note that $e_{\boldsymbol{\theta}} + \ell^>_{\boldsymbol{\theta}} + \ell^<_{\boldsymbol{\theta}} = 1$. The function $\ell^>_{\boldsymbol{\theta}}$ is supported on the set $|\boldsymbol{\xi}+\boldsymbol{\theta}/2| \ge 11\rho_n/2$, and $\ell^<_{\boldsymbol{\theta}}$ is supported on the set $|\boldsymbol{\xi}+\boldsymbol{\theta}/2| \le \rho_n/2$. The function $e_{\boldsymbol{\theta}}$ is supported in the shell $\rho_n/4 \le |\boldsymbol{\xi} + \boldsymbol{\theta}/2| \le 23\rho_n/4$. Using the notation $\ell_{\boldsymbol{\theta}}$ for any of the functions $\ell^{>}_{\boldsymbol{\theta}}$ or $\ell^{<}_{\boldsymbol{\theta}}$, we point out that

$$\begin{cases} e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = e_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}),\ \ \ell_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \ell_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}), \\ \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \varphi_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}),\ \ \ \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \zeta_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}). \end{cases} \tag{8.16}$$

Note that the above functions satisfy the estimates

$$\begin{cases} |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} e_{\boldsymbol{\theta}}(\boldsymbol{\xi})| + |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \ell_{\boldsymbol{\theta}}(\boldsymbol{\xi})| \ll \rho_n^{-|\mathbf{s}|}, \\ |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})| + |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi})| \ll \rho_n^{-\beta|\mathbf{s}|}. \end{cases} \tag{8.17}$$

Using the above cut-off functions, for any symbol $b \in \mathbf{S}_\alpha(\beta)$, we introduce five new symbols $b^{\mathcal{SE}}, b^o, b^{\mathcal{LE}}, b^{\mathcal{NR}}, b^{\mathcal{R}}$ in the following way:

$$b^{\mathcal{LE}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) = \sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \ell^{>}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{8.18}$$

$$b^{\mathcal{NR}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) = \sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{8.19}$$

$$b^{\mathcal{R}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) = \sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{8.20}$$

$$b^{\mathcal{SE}}(\mathbf{x}, \boldsymbol{\xi}; \rho_n) = \sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}'} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) \ell^{<}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{8.21}$$

$$b^o(\mathbf{x}, \boldsymbol{\xi}; \rho_n) = b^o(\boldsymbol{\xi}; \rho_n) = \hat{b}(0, \boldsymbol{\xi}). \tag{8.22}$$

The superscripts here are chosen to mean correspondingly 'large energy,' 'non-resonant,' 'resonant,' 'small energy,' and 0-th Fourier coefficient. The corresponding operators are denoted by

$$B^{\mathcal{LE}} = \mathrm{Op}(b^{\mathcal{LE}}),\ B^{\mathcal{NR}} = \mathrm{Op}(b^{\mathcal{NR}}),$$
$$B^{\mathcal{R}} = \mathrm{Op}(b^{\mathcal{R}}),\ B^{\mathcal{SE}} = \mathrm{Op}(b^{\mathcal{SE}}),\ B^o = \mathrm{Op}(b^o).$$

By definitions (8.13), (8.14), and (8.15), we have

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

The role of each of these operator 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

$$|\boldsymbol{\theta}| \le \rho_n^{0+},\ \frac{1}{4}\rho_n \le |\boldsymbol{\xi}+\boldsymbol{\theta}/2| \le \frac{23}{4}\rho_n,\ \frac{1}{4}\rho_n - \frac{1}{2}\rho_n^{0+} \le |\boldsymbol{\xi}| \le \frac{23}{4}\rho_n + \frac{1}{2}\rho_n^{0+}. \tag{8.23}$$

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

$$\left|\boldsymbol{\xi} + \frac{\boldsymbol{\theta}}{2}\right| \le \frac{1}{2}\rho_n,\ |\boldsymbol{\xi}| \le \frac{1}{2}\rho_n + \frac{1}{2}\rho_n^{0+}. \tag{8.24}$$

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

$$\left|\boldsymbol{\xi}+\frac{\boldsymbol{\theta}}{2}\right| \geq \frac{11}{2}\rho_n,\ |\boldsymbol{\xi}| \geq \frac{11}{2}\rho_n - \frac{1}{2}\rho_n^{0+}. \tag{8.25}$$

The introduced symbols play a central role in the proof of Lemma 3.3. As we have seen in Section 6, due to (8.24) and (8.25), the symbols $b^{\mathcal{SE}}$ and $b^{\mathcal{LE}}$ make only a negligible contribution to the spectrum of the operator H near the point $\lambda=\rho^2$, $\rho\in I_n$. 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 $\mathbf{x}$.

We will often combine $B^{\mathcal{R}}$, $B^{\mathcal{LE}}$, and $B^{\mathcal{SE}}$. For instance, $B^{\mathcal{R},\mathcal{LE}}=B^{\mathcal{R}}+B^{\mathcal{LE}}$, $B^{\mathcal{R},\mathcal{LE},\mathcal{SE}}=B^{\mathcal{R},\mathcal{LE}}+B^{\mathcal{SE}}$. A similar convention applies to the symbols. 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:

$$|b^{\mathcal{R}}|^{(\alpha)}_{l,s}+|b^{\mathcal{NR}}|^{(\alpha)}_{l,s}+|b^{\mathcal{LE}}|^{(\alpha)}_{l,s}+|b^{o}|^{(\alpha)}_{l,s}+|b^{\mathcal{SE}}|^{(\alpha)}_{l,s}\ll|b|^{(\alpha)}_{l,s}. \tag{8.26}$$

Indeed, let us check this for the symbol $b^{\mathcal{NR}}$, for instance. According to (8.23) and (8.17), on the support of the function $\hat{b}^{\mathcal{NR}}(\boldsymbol{\theta},\ \cdot\ ;\rho_n)$, we have

$$|\mathbf{D}^{\mathbf{s}}\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})|\ll\rho_n^{-\beta|\mathbf{s}|}\ll\langle\boldsymbol{\xi}\rangle^{-|\mathbf{s}|\beta},$$

$$|\mathbf{D}^{\mathbf{s}}\ell^{>}_{\boldsymbol{\theta}}(\boldsymbol{\xi})|+|\mathbf{D}^{\mathbf{s}}\ell^{<}_{\boldsymbol{\theta}}(\boldsymbol{\xi})|+|\mathbf{D}^{\mathbf{s}}e_{\boldsymbol{\theta}}(\boldsymbol{\xi})|\ll\rho_n^{-|\mathbf{s}|}\ll\langle\boldsymbol{\xi}\rangle^{-|\mathbf{s}|\beta}.$$

This immediately leads to the bound of the form (8.26) for the symbol $b^{\mathcal{NR}}$.

The introduced operations also preserve symmetry. Indeed, let us calculate using (8.16):

$$\begin{aligned}\overline{\hat{b}^{\mathcal{R}}(-\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\theta})}&=\overline{\hat{b}(-\boldsymbol{\theta},\boldsymbol{\xi}+\boldsymbol{\theta})}\zeta_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta})e_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta})\\&=\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi})e_{\boldsymbol{\theta}}(\boldsymbol{\xi})=\hat{b}^{\mathcal{R}}(\boldsymbol{\theta},\boldsymbol{\xi}).\end{aligned}$$

Therefore, by (8.9) the operator $B^{\mathcal{R}}$ is symmetric if $\mathrm{Op}(b)$ is symmetric. The proof is similar for the rest of the operators introduced above.

Let us list some other elementary properties of the introduced operators. In the lemma below, we use the projection $\mathcal{P}(\mathcal{C}), \mathcal{C}\subset\mathbb{R}$, whose definition was given in Section 6.

LEMMA 8.4. *Let $b\in\mathbf{S}_\alpha(\beta)$ with some $\alpha\in\mathbb{R}$. Then the following hold:*

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

$$\|B^{\mathcal{SE}}\|\ll|b|^{(\alpha)}_{0,0}\rho_n^{\beta\max(\alpha,0)}.$$

Moreover,

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

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

$$\mathcal{P}(B(\rho_n/8))B^{\mathcal{R}} = B^{\mathcal{R}}\mathcal{P}(B(\rho_n/8)) = \big(I - \mathcal{P}(B(6\rho_n)\big)B^{\mathcal{R}} = B^{\mathcal{R}}\big(I - \mathcal{P}(B(6\rho_n))\big) = 0, \tag{8.27}$$

and similar relations hold for the operator $B^{\mathcal{NR}}$ *as well.*

Moreover, for any $\gamma \in \mathbb{R}$, *one has* $b^{\mathcal{NR}}, b^{\mathcal{R}} \in \mathbf{S}_\gamma$ *and*

$$\big|b^{\mathcal{NR}}\big|^{(\gamma)}_{l,s} + \big|b^{\mathcal{R}}\big|^{(\gamma)}_{l,s} \ll \rho_n^{\beta(\alpha-\gamma)} \big|b\big|^{(\alpha)}_{l,s}, \tag{8.28}$$

for all l *and* s, *with an implied constant independent of* b *and* $n \geq 1$. *In particular, the operators* $B^{\mathcal{NR}}, B^{\mathcal{R}}$ *are bounded and*

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

(iii) $\mathcal{P}\big(B(5\rho_n)\big)B^{\mathcal{LE}} = B^{\mathcal{LE}}\mathcal{P}\big(B(5\rho_n)\big) = 0.$

Proof of (i). It follows from (8.3) and (8.24) that

$$\sum_{\boldsymbol\theta} \sup_{\boldsymbol\xi} |\hat b^{\mathcal{SE}}(\boldsymbol\theta, \boldsymbol\xi; \rho_n)| \ll \rho_n^{\beta\max(\alpha,0)} \sum_{\boldsymbol\theta} \sup_{\boldsymbol\xi} \langle\boldsymbol\xi\rangle^{-\beta\alpha} |\hat b(\boldsymbol\theta, \boldsymbol\xi; \rho_n)| = \big|b\big|^{(\alpha)}_{0,0} \rho_n^{\beta\max(\alpha,0)}. \tag{8.29}$$

By Proposition 8.1 this implies the sought bound for the norm $\|B^{\mathcal{SE}}\|$.

In view of (8.24), the second part of statement (i) follows from (8.21).

Proof of (ii). Relations (8.27) follow from definitions (8.20) and (8.19) in view of (8.23).

Furthermore, by (8.23) and (8.26),

$$\begin{aligned}\sum_{\boldsymbol\theta} \langle\boldsymbol\theta\rangle^l \sup_{\boldsymbol\xi} \langle\boldsymbol\xi\rangle^{(-\gamma+s)\beta} &|\mathbf{D}^{\mathbf{s}}_{\boldsymbol\xi} \hat b^{\mathcal{NR}}(\boldsymbol\theta, \boldsymbol\xi; \rho_n)| \\ &\ll \rho_n^{\beta(\alpha-\gamma)} \sum_{\boldsymbol\theta} \langle\boldsymbol\theta\rangle^l \sup_{\boldsymbol\xi} \langle\boldsymbol\xi\rangle^{(-\alpha+s)\beta} |\mathbf{D}^{\mathbf{s}}_{\boldsymbol\xi} \hat b^{\mathcal{NR}}(\boldsymbol\theta, \boldsymbol\xi; \rho_n)| \\ &\leq \big|b^{\mathcal{NR}}\big|^{(\alpha)}_{l,s} \rho_n^{\beta(\alpha-\gamma)} \ll \big|b\big|^{(\alpha)}_{l,s} \rho_n^{\beta(\alpha-\gamma)}.\end{aligned}$$

This means that $b^{\mathcal{NR}} \in \mathbf{S}_\gamma$ for any $\gamma \in \mathbb{R}$ and (8.28) holds for $b^{\mathcal{NR}}$. The bound for the norm follow from (8.28), with $\gamma = 0$, and Proposition 8.1. The proof for $b^{\mathcal{R}}$ is analogous.

Proof of (iii). Similar to (i). The required result follows from (8.25). ☐

9. Gauge transform and the symbol of the resulting operator

9.1. *Preparation.* Our strategy will be to find a unitary operator that reduces $H = H_0 + \mathrm{Op}(b)$, $H_0 := -\Delta$, to another PDO, whose symbol, essentially, depends only on $\boldsymbol\xi$. (Notice that now we have started to distinguish between the potential b and the operator of multiplication by it $\mathrm{Op}(b)$.) More precisely, we want to find operators H_1 and H_2 with the properties discussed in Section 6. The unitary operator will be constructed in the form $U = e^{i\Psi}$ with a suitable bounded self-adjoint quasi-periodic PDO Ψ. This is why we sometimes call it a 'gauge transform.' It is useful to consider $e^{i\Psi}$ as an element of the group

$$U(t) = \exp\{i\Psi t\}, \quad \forall t \in \mathbb{R}.$$

We assume that the operator $\mathrm{ad}(H_0, \Psi)$ *is bounded, so that* $U(t)D(H_0) = D(H_0)$. This assumption will be justified later on. Let us express the operator

$$A_t := U(-t)HU(t)$$

via its (weak) derivative with respect to t:

$$A_t = H + \int_0^t U(-t')\,\mathrm{ad}(H;\Psi)U(t')dt'.$$

By induction it is easy to show that

$$A_1 = H + \sum_{j=1}^{\tilde{k}} \frac{1}{j!}\,\mathrm{ad}^j(H;\Psi) + R^{(1)}_{\tilde{k}+1}, \tag{9.1}$$

$$R^{(1)}_{\tilde{k}+1} := \int_0^1 dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{\tilde{k}}} U(-t_{\tilde{k}+1})\,\mathrm{ad}^{\tilde{k}+1}(H;\Psi)U(t_{\tilde{k}+1})dt_{\tilde{k}+1}.$$

The operator Ψ is sought in the form

$$\Psi = \sum_{j=1}^{\tilde{k}} \Psi_j, \ \ \Psi_j = \mathrm{Op}(\psi_j), \tag{9.2}$$

with symbols ψ_j from some suitable class $\mathbf{S}_{\sigma_j}$ to be specified later on. Substitute this formula in (9.1) and rewrite, regrouping the terms:

$$\begin{aligned} A_1 &= H_0 + \mathrm{Op}(b) \\ &\quad + \sum_{j=1}^{\tilde{k}} \frac{1}{j!} \sum_{l=j}^{\tilde{k}} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(H;\Psi_{k_1},\Psi_{k_2},\dots,\Psi_{k_j}) + R^{(1)}_{\tilde{k}+1} + R^{(2)}_{\tilde{k}+1}, \end{aligned} \tag{9.3}$$

$$R^{(2)}_{\tilde{k}+1} := \sum_{j=1}^{\tilde{k}} \frac{1}{j!} \sum_{k_1+k_2+\dots+k_j \ge \tilde{k}+1} \mathrm{ad}(H;\Psi_{k_1},\Psi_{k_2},\dots,\Psi_{k_j}).$$

Changing this expression yet again produces

$$\begin{aligned} A_1 = H_0 + \mathrm{Op}(b) &+ \sum_{l=1}^{\tilde{k}} \mathrm{ad}(H_0;\Psi_l) \\ &+ \sum_{j=2}^{\tilde{k}} \frac{1}{j!} \sum_{l=j}^{\tilde{k}} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(H_0;\Psi_{k_1},\Psi_{k_2},\dots,\Psi_{k_j}) \\ &+ \sum_{j=1}^{\tilde{k}} \frac{1}{j!} \sum_{l=j}^{\tilde{k}} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(\mathrm{Op}(b);\Psi_{k_1},\Psi_{k_2},\dots,\Psi_{k_j}) + R^{(1)}_{\tilde{k}+1} + R^{(2)}_{\tilde{k}+1}. \end{aligned}$$

Next, we switch the summation signs and decrease l by one in the second summation:

$$A_1 = H_0 + \mathrm{Op}(b) + \sum_{l=1}^{\tilde{k}} \mathrm{ad}(H_0; \Psi_l)$$
$$+ \sum_{l=2}^{\tilde{k}} \sum_{j=2}^{l} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l} \mathrm{ad}(H_0; \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j})$$
$$+ \sum_{l=2}^{\tilde{k}+1} \sum_{j=1}^{l-1} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l-1} \mathrm{ad}(\mathrm{Op}(b); \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j}) + R^{(1)}_{\tilde{k}+1} + R^{(2)}_{\tilde{k}+1}.$$

Now we introduce the notation

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

$$B_l := \sum_{j=1}^{l-1} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l-1} \mathrm{ad}(\mathrm{Op}(b); \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j}),\ l \geq 2, \tag{9.4}$$

$$T_l := \sum_{j=2}^{l} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l} \mathrm{ad}(H_0; \Psi_{k_1}, \Psi_{k_2}, \ldots, \Psi_{k_j}),\ l \geq 2. \tag{9.5}$$

We emphasise that the operators B_l and T_l depend only on $\Psi_1, \Psi_2, \ldots, \Psi_{l-1}$. Let us make one more rearrangement:

$$A_1 = H_0 + \mathrm{Op}(b) + \sum_{l=1}^{\tilde{k}} \mathrm{ad}(H_0, \Psi_l) + \sum_{l=2}^{\tilde{k}} B_l + \sum_{l=2}^{\tilde{k}} T_l + R_{\tilde{k}+1}, \tag{9.6}$$
$$R_{\tilde{k}+1} = B_{\tilde{k}+1} + R^{(1)}_{\tilde{k}+1} + R^{(2)}_{\tilde{k}+1}.$$

Now we can specify our algorithm for finding Ψ_j's. The symbols ψ_j will be found from the following system of commutator equations:

$$\mathrm{ad}(H_0; \Psi_1) + B_1^{\mathcal{NR}} = 0, \tag{9.7}$$

$$\mathrm{ad}(H_0; \Psi_l) + B_l^{\mathcal{NR}} + T_l^{\mathcal{NR}} = 0,\ l \geq 2, \tag{9.8}$$

and hence

$$\begin{cases} A_1 = H_0 + Y^{(o)}_{\tilde{k}} + Y^{\mathcal{R}}_{\tilde{k}} + Y^{\mathcal{SE},\mathcal{LE}}_{\tilde{k}} + R_{\tilde{k}+1}, \\ Y_{\tilde{k}} = \sum_{l=1}^{\tilde{k}} B_l + \sum_{l=2}^{\tilde{k}} T_l. \end{cases} \tag{9.9}$$

Below we denote by $y_{\tilde{k}}$ the symbol of the PDO $Y_{\tilde{k}}$. Please recall that by Lemma 8.4(ii), the operators $B_l^{\mathcal{NR}}, T_l^{\mathcal{NR}}$ are bounded, and therefore, in view of (9.7), (9.8), so is the commutator $\mathrm{ad}(H_0; \Psi)$. This justifies the assumption made in the beginning of the formal calculations in this section.

9.2. *Commutator equations.* Put

$$\tilde{\chi}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) := e_{\boldsymbol{\theta}}(\boldsymbol{\xi})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})(|\boldsymbol{\xi}+\boldsymbol{\theta}|^2 - |\boldsymbol{\xi}|^2)^{-1} = \frac{e_{\boldsymbol{\theta}}(\boldsymbol{\xi})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})}{2\langle\boldsymbol{\theta}, \boldsymbol{\xi}+\frac{\boldsymbol{\theta}}{2}\rangle}$$

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*

$$\hat{\psi}(\boldsymbol{\theta}, \boldsymbol{\xi}) = i\, \hat{a}(\boldsymbol{\theta}, \boldsymbol{\xi})\tilde{\chi}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) \tag{9.10}$$

solves the equation

$$\mathrm{ad}(H_0; \Psi) + \mathrm{Op}(a^{\mathcal{NR}}) = 0. \tag{9.11}$$

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:

$$|\psi|^{(\gamma)}_{l,s} \ll \rho_n^{\beta(\omega-\gamma-1)} r(\rho_n)^{-1} |a|^{(\omega)}_{l-1,s} \ll \rho_n^{\beta(\omega-\gamma-1)+0+} |a|^{(\omega)}_{l-1,s}. \tag{9.12}$$

Using Propositions 8.1, 8.2, 8.3, Lemma 9.1, and repeating arguments from the proof of Lemma 4.2 from [12], we obtain the following estimates for the symbols introduced above.

LEMMA 9.2. *Let $b \in \mathbf{S}_0(\beta)$ be a symmetric symbol. Then ψ_j, b_j, $t_j \in \mathbf{S}_\gamma(\beta)$ for any $\gamma \in \mathbb{R}$, and*

$$|\psi_j|^{(\gamma)}_{l,s} \le C_j \rho_n^{\beta(1-\gamma-2j)} r(\rho_n)^{-j} \left(|b|^{(0)}_{l_j,s_j}\right)^j, \quad j \ge 1, \tag{9.13}$$

$$|b_j|^{(\gamma)}_{l,s} + |t_j|^{(\gamma)}_{l,s} \le C_j \rho_n^{\beta(2-\gamma-2j)} r(\rho_n)^{-j+1} \left(|b|^{(0)}_{l_j,s_j}\right)^j, \quad j \ge 2. \tag{9.14}$$

Here C_j, l_j, s_j depend only on j, l, s, and γ. Moreover, assuming ρ_0 is large enough (*depending on l, s, γ, b, and $\tilde{k}$*), *we get*

$$|\psi|^{(\gamma)}_{l,s} \ll \rho_n^{-\beta(1+\gamma)} r(\rho_n)^{-1} |b|^{(0)}_{l,s}, \tag{9.15}$$

$$|y_{\tilde{k}}|^{(0)}_{l,s} \le 2|b|^{(0)}_{l,s}, \tag{9.16}$$

$$\|R_{\tilde{k}+1}\| \ll \rho_n^{-2\beta\tilde{k}} r(\rho_n)^{-\tilde{k}} \left(|b|^{(0)}_{l_{\tilde{k}+1},s_{\tilde{k}+1}}\right)^{\tilde{k}+1}. \tag{9.17}$$

Now, we take

$$\tilde{k} > (M + (d-2))/\beta \tag{9.18}$$

and assume that k is large enough so that $r(\rho_n)^{-1} \ll \rho_n^{0+} \ll \rho_n^{\beta}$. Then

$$\|R_{\tilde{k}+1}\| \ll \rho_n^{-M+(2-d)},$$

and we can disregard $R_{\tilde{k}+1}$ due to Corollary 4.3. More precisely, let $W = W_{\tilde{k}}$ be the operator with symbol

(9.19)
$$w_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi}) := y_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi}) - y_{\tilde{k}}^{\mathcal{NR}}(\mathbf{x},\boldsymbol{\xi}), \quad \text{i.e.,} \quad \hat{w}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi}) = \hat{y}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi})(1 - e_{\boldsymbol{\theta}}(\boldsymbol{\xi})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi})).$$

We put $H_1 := A_1$ and $H_2 := -\Delta + W$. Then $||H_1 - H_2|| \ll \rho_n^{-M+(2-d)}$ and, moreover, the symbol w satisfies condition (6.8). This means that all the constructions of Section 6 are valid, and all we need to do is to compute $\mathrm{vol}(G_\lambda)$.

9.3. *Computing the symbol of the operator after gauge transform.* The following lemma provides us with more explicit form of the symbol $y_{\tilde{k}}$.

LEMMA 9.3. *We have $\hat{y}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi}) = 0$ for $\boldsymbol{\theta} \notin \boldsymbol{\Theta}_{\tilde{k}}$. Otherwise,*

(9.20)
$$\hat{y}_{\tilde{k}}(\boldsymbol{\theta},\boldsymbol{\xi}) = \hat{b}(\boldsymbol{\theta}) + \sum_{s=1}^{\tilde{k}-1}\sum C_s(\boldsymbol{\theta},\boldsymbol{\xi})\hat{b}(\boldsymbol{\theta}_{s+1})\prod_{j=1}^{s}\hat{b}(\boldsymbol{\theta}_j)\tilde{\chi}_{\boldsymbol{\theta}'_j}(\boldsymbol{\xi}+\boldsymbol{\phi}'_j)$$
$$= \hat{b}(\boldsymbol{\theta}) + \sum_{s=1}^{\tilde{k}-1}\sum C_s(\boldsymbol{\theta},\boldsymbol{\xi})\hat{b}(\boldsymbol{\theta}_{s+1})\prod_{j=1}^{s}\hat{b}(\boldsymbol{\theta}_j)\frac{e_{\boldsymbol{\theta}'_j}(\boldsymbol{\xi}+\boldsymbol{\phi}'_j)\varphi_{\boldsymbol{\theta}'_j}(\boldsymbol{\xi}+\boldsymbol{\phi}'_j)}{2\langle\boldsymbol{\theta}'_j,\boldsymbol{\xi}+\boldsymbol{\phi}'_j+\frac{\boldsymbol{\theta}'_j}{2}\rangle},$$

where the second sums are taken over all $\boldsymbol{\theta}_j \in \boldsymbol{\Theta}$, $\boldsymbol{\theta}'_j, \boldsymbol{\phi}'_j \in \boldsymbol{\Theta}_{s+1}$, and

$$C_s(\boldsymbol{\theta},\boldsymbol{\xi}) = \sum_{p=1}^{s}\sum_{\boldsymbol{\theta}''_j,\boldsymbol{\phi}''_j\in\boldsymbol{\Theta}_{s+1}\ (1\le j\le p)} C_s^{(p)}(\boldsymbol{\theta})\prod_{j=1}^{p} e_{\boldsymbol{\theta}''_j}(\boldsymbol{\xi}+\boldsymbol{\phi}''_j)\varphi_{\boldsymbol{\theta}''_j}(\boldsymbol{\xi}+\boldsymbol{\phi}''_j). \tag{9.21}$$

Here $C_s^{(p)}(\boldsymbol{\theta})$ depend on s, p, and all vectors $\boldsymbol{\theta}, \boldsymbol{\theta}_j, \boldsymbol{\theta}'_j, \boldsymbol{\phi}'_j, \boldsymbol{\theta}''_j, \boldsymbol{\phi}''_j$. At the same time, coefficients $C_s^{(p)}(\boldsymbol{\theta})$ can be bounded uniformly by a constant that depends on s only. We apply the convention that $0/0 = 0$.

Proof. We will prove the lemma by induction. Namely, let $\ell \ge 2$. 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,

$$\hat{\psi}_m(\boldsymbol{\theta},\boldsymbol{\xi}) = \sum C'_m(\boldsymbol{\theta},\boldsymbol{\xi})\prod_{j=1}^{m}\hat{b}(\boldsymbol{\theta}_j)\tilde{\chi}_{\boldsymbol{\theta}'_j}(\boldsymbol{\xi}+\boldsymbol{\phi}'_j), \tag{9.22}$$

where the sum is taken over all $\boldsymbol{\theta}_j \in \boldsymbol{\Theta}$, $\boldsymbol{\theta}'_j, \boldsymbol{\phi}'_j \in \boldsymbol{\Theta}_m$ and $C'_m(\boldsymbol{\theta},\boldsymbol{\xi})$ admit representation similar to (9.21).

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

$$\widehat{\mathrm{ad}(\mathrm{Op}(b); \Psi_{k_1}, \dots, \Psi_{k_p})}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \sum C''_s(\boldsymbol{\theta}, \boldsymbol{\xi}) \hat{b}(\boldsymbol{\theta}_{s+1}) \prod_{j=1}^{s} \hat{b}(\boldsymbol{\theta}_j) \tilde{\chi}_{\boldsymbol{\theta}'_j}(\boldsymbol{\xi} + \boldsymbol{\phi}'_j), \tag{9.23}$$

where the sum is taken over all $\boldsymbol{\theta}_j \in \boldsymbol{\Theta}$, $\boldsymbol{\theta}'_j, \boldsymbol{\phi}'_j \in \boldsymbol{\Theta}_{s+1}$ and $C''_s(\boldsymbol{\theta}, \boldsymbol{\xi})$ admit representation similar to (9.21).

(3) For any $s = 2, \dots, \ell$ and any $k_1, \dots, k_p$ $(p \geq 2)$ 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,

$$\widehat{\mathrm{ad}(H_0; \Psi_{k_1}, \dots, \Psi_{k_p})}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \sum C'''_s(\boldsymbol{\theta}, \boldsymbol{\xi}) \hat{b}(\boldsymbol{\theta}_s) \prod_{j=1}^{s-1} \hat{b}(\boldsymbol{\theta}_j) \tilde{\chi}_{\boldsymbol{\theta}'_j}(\boldsymbol{\xi} + \boldsymbol{\phi}'_j), \tag{9.24}$$

where the sum is taken over all $\boldsymbol{\theta}_j \in \boldsymbol{\Theta}$, $\boldsymbol{\theta}'_j, \boldsymbol{\phi}'_j \in \boldsymbol{\Theta}_s$, and $C'''_s(\boldsymbol{\theta}, \boldsymbol{\xi})$ admit representation similar to (9.21).

For $\ell = 2$ statements, (1)–(3) can be easily checked. Indeed,

$$\hat{\psi}_1(\boldsymbol{\theta}, \boldsymbol{\xi}) = i\hat{b}(\boldsymbol{\theta}) \tilde{\chi}_{\boldsymbol{\theta}}(\boldsymbol{\xi}), \tag{9.25}$$

$$\widehat{\mathrm{ad}(\mathrm{Op}(b); \Psi_1)}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\chi}+\boldsymbol{\phi}=\boldsymbol{\theta}} \left(\hat{b}(\boldsymbol{\chi}) \hat{b}(\boldsymbol{\phi}) \tilde{\chi}_{\boldsymbol{\phi}}(\boldsymbol{\xi} + \boldsymbol{\chi}) - \hat{b}(\boldsymbol{\chi}) \hat{b}(\boldsymbol{\phi}) \tilde{\chi}_{\boldsymbol{\phi}}(\boldsymbol{\xi}) \right), \tag{9.26}$$

$$\widehat{\mathrm{ad}(H_0; \Psi_1, \Psi_1)}(\boldsymbol{\theta}, \boldsymbol{\xi}) = - \sum_{\boldsymbol{\chi}+\boldsymbol{\phi}=\boldsymbol{\theta}} \left(\hat{b}^{\mathcal{NR}}(\boldsymbol{\chi}) \hat{b}(\boldsymbol{\phi}) \tilde{\chi}_{\boldsymbol{\phi}}(\boldsymbol{\xi} + \boldsymbol{\chi}) - \hat{b}^{\mathcal{NR}}(\boldsymbol{\chi}) \hat{b}(\boldsymbol{\phi}) \tilde{\chi}_{\boldsymbol{\phi}}(\boldsymbol{\xi}) \right). \tag{9.27}$$

Now, we complete the induction in several steps.

Step 1. First of all, notice that due to (9.4), (9.5), for any $m = 2, \dots, \ell$, the symbol of B_m admits representation of the form (9.23) with $s = m-1$ and the symbol of T_m admits representation of the form (9.24) with $s = m$. Then it follows from Lemma 9.1 and (9.8) that Ψ_ℓ admits representation of the form (9.22).

Step 2. *Proof of* (9.23) *with* $s = \ell$. Let $k_1 + \dots + k_p = \ell$. If $p \geq 2$, then

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

Since $k_1 + \dots + k_{p-1} \leq \ell - 1$ and $k_p \leq \ell - 1$, we can apply (9.22) and (9.23). Combined with (8.11) it gives representation of the form (9.23). If $p = 1$, then $\mathrm{ad}\,\mathrm{Op}(b); \Psi_\ell)$ satisfies (9.23) because of (8.11) and Step 1.

Step 3. *Proof of* (9.24) *with* $s = \ell + 1$. Let $k_1 + \cdots + k_p = \ell + 1$, $p \geq 2$. If $p \geq 3$, then (cf. Step 2)

$$\mathrm{ad}(H_0; \Psi_{k_1}, \dots, \Psi_{k_p}) = \mathrm{ad}(\mathrm{ad}(H_0; \Psi_{k_1}, \dots, \Psi_{k_{p-1}}); \Psi_{k_p}).$$

Since $k_1 + \cdots + k_{p-1} \leq \ell$, $p - 1 \geq 2$, and $k_p \leq \ell - 1$, we can apply (9.22) and (9.24). Together with (8.11), it gives representation of the form (9.24). If $p = 2$, then (see (9.8))

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

Since $k_1 \leq \ell$ and $k_2 \leq \ell$, the representation of the form (9.24) follows from (8.11) and Step 1. (Formally exceptional case $k_1 = 1$, $k_2 = \ell$ can be treated separately in the same way using (9.7) instead of (9.8).)

Induction is complete.

Now, (9.23), (9.24) and (9.4), (9.5), (9.9) prove the lemma. □

10. Contribution from various resonance regions

10.1. *Summing the contributions from individual eigenvalues.* Let us fix a subspace $\mathfrak{V} \in \mathcal{V}_m$, $m < d$, and a component $\boldsymbol{\Xi}_p$ of the resonance region $\boldsymbol{\Xi}(\mathfrak{V})$. Our aim is to compute the contribution to the density of states from each component $\boldsymbol{\Xi}_p$. This means that we define $\hat{A}^+(\boldsymbol{\Xi}_p) := \hat{A}^+ \cap \boldsymbol{\Xi}(\mathfrak{V})_p$ and $\hat{A}^-(\boldsymbol{\Xi}_p) := \hat{A}^- \cap \boldsymbol{\Xi}(\mathfrak{V})_p$ and try to compute

$$\mathrm{vol}\, \hat{A}^+(\boldsymbol{\Xi}_p) - \mathrm{vol}\, \hat{A}^-(\boldsymbol{\Xi}_p). \tag{10.1}$$

Since formulas (6.19) and (6.4) obviously imply that

$$\mathrm{vol}(G_\lambda) = w_d \rho^d + \sum_{m=0}^{d-1} \sum_{\mathfrak{V} \in \mathcal{V}_m} \sum_p \Big(\mathrm{vol}\, \hat{A}^+(\boldsymbol{\Xi}_p) - \mathrm{vol}\, \hat{A}^-(\boldsymbol{\Xi}_p)\Big), \tag{10.2}$$

if we manage to compute (10.1) (or at least prove that this expression admits a complete asymptotic expansion in ρ), Lemma 3.3 would be proved. Thus, we fix $\mathfrak{V}$ and, moreover, we fix a component $\boldsymbol{\Xi}(\mathfrak{V})_p$ of the resonance region. Recall that $K = d - m - 1$.

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

$$H_2(\boldsymbol{\xi}) := \mathcal{P}(\boldsymbol{\Upsilon}(\boldsymbol{\xi})) H_2 \mathcal{P}(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))$$

as an operator acting in $\mathfrak{H}_{\boldsymbol{\xi}} := \mathcal{P}(\boldsymbol{\Upsilon}(\boldsymbol{\xi}))\mathfrak{H}$. (Recall that $\mathfrak{H}_{\boldsymbol{\xi}}$ is an invariant subspace of H_2 acting in $B_2(\mathbb{R}^d)$.) Suppose now that two points $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ have the same coordinates X and $\tilde{\Phi}$ and different coordinates r. Then $\boldsymbol{\xi} \in \boldsymbol{\Xi}_p$ implies $\boldsymbol{\eta} \in \boldsymbol{\Xi}_p$ and $\boldsymbol{\Upsilon}(\boldsymbol{\eta}) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}) + (\boldsymbol{\eta} - \boldsymbol{\xi})$. This shows that two spaces $\mathfrak{H}_{\boldsymbol{\xi}}$ and $\mathfrak{H}_{\boldsymbol{\eta}}$ have the same dimension and, moreover, there is a natural isometry $F_{\boldsymbol{\xi},\boldsymbol{\eta}} : \mathfrak{H}_{\boldsymbol{\xi}} \to \mathfrak{H}_{\boldsymbol{\eta}}$ given by $F : \mathbf{e}_{\boldsymbol{\nu}} \mapsto \mathbf{e}_{\boldsymbol{\nu}+(\boldsymbol{\eta}-\boldsymbol{\xi})}$, $\boldsymbol{\nu} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$. This isometry allows us to 'compare' operators acting in $\mathfrak{H}_{\boldsymbol{\xi}}$ and $\mathfrak{H}_{\boldsymbol{\eta}}$. 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 $\mathfrak{H}(X,\tilde{\Phi})$. We will fix the values $(X,\tilde{\Phi})$ and study how these operators depend on r. Thus, we denote by $H_2(r) = H_2(r;X,\tilde{\Phi})$ the operator $H_2(\boldsymbol{\xi})$ with $\boldsymbol{\xi} = (X,r,\tilde{\Phi})$, acting in $\mathfrak{H}(X,\tilde{\Phi})$.

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

$$h_2(\mathbf{x},\boldsymbol{\xi}) = |\boldsymbol{\xi}|^2 + w_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi}) = r^2 + 2r\langle \mathbf{a},\mathbf{n}(\boldsymbol{\eta})\rangle + |\mathbf{a}|^2 + w_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi}) + |X|^2, \tag{10.3}$$

where the Fourier coefficients of $w_{\tilde{k}}$ satisfy (9.16), (9.19), (9.20), (9.21) and we denote, as usual, $\boldsymbol{\eta} = \boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}$. This immediately implies that the operator $H_2(r)$ is monotonically increasing in r; in particular, all its eigenvalues $\lambda_j(H_2(r))$ are increasing in r. Thus, the function $g(\boldsymbol{\xi})$ (defined in Section 6) is an increasing function of $r(\boldsymbol{\xi})$ if we fix other coordinates of $\boldsymbol{\xi}$, so the equation

$$g(\boldsymbol{\xi}) = \rho^2 \tag{10.4}$$

has a unique solution if we fix the values $(X,\tilde{\Phi})$; we denote the r-coordinate of this solution by $\tau = \tau(\rho) = \tau(\rho;X,\tilde{\Phi})$, so that

$$g(\boldsymbol{\xi}(X,\tau,\tilde{\Phi}) = \rho^2. \tag{10.5}$$

By $\tau_0 = \tau_0(\rho) = \tau_0(\rho;X,\tilde{\Phi})$ we denote the value of τ for the unperturbed operator; i.e., τ_0 is a unique solution of the equation

$$|\boldsymbol{\xi}(X,\tau_0,\tilde{\Phi})| = \rho. \tag{10.6}$$

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 of the sets $\hat{A}^\pm$ implies that the intersection

$$\hat{A}^+ \cap \{\boldsymbol{\xi}(X,r,\tilde{\Phi}),\ r\in\mathbb{R}_+\} \tag{10.7}$$

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

$$\hat{A}^- \cap \{\boldsymbol{\xi}(X,r,\tilde{\Phi}),\ r\in\mathbb{R}_+\} \tag{10.8}$$

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

$$\hat{A}^+(\Xi_p) = \{\boldsymbol{\xi} = \boldsymbol{\xi}(X,r,\tilde{\Phi}),\ X\in\Omega(\mathfrak{V}),\tilde{\Phi}\in M_p,\ r\in[\tau_0(\rho;X,\tilde{\Phi}),\tau(\rho;X,\tilde{\Phi})]\} \tag{10.9}$$

and

$$\hat{A}^-(\Xi_p) = \{\boldsymbol{\xi} = \boldsymbol{\xi}(X,r,\tilde{\Phi}),\ X\in\Omega(\mathfrak{V}),\tilde{\Phi}\in M_p,\ r\in[\tau(\rho;X,\tilde{\Phi}),\tau_0(\rho;X,\tilde{\Phi})]\}. \tag{10.10}$$

This implies that

$$\text{(10.11)}\qquad \begin{aligned}\operatorname{vol}\hat{A}^{+}(\boldsymbol{\Xi}_p)-\operatorname{vol}\hat{A}^{-}(\boldsymbol{\Xi}_p)&=\int_{\Omega(\mathfrak{V})}dX\int_{M_p}d\tilde{\Phi}\int_{\tau_0(\rho;X,\tilde{\Phi})}^{\tau(\rho;X,\tilde{\Phi})}r^K dr\\&=(K+1)^{-1}\int_{M_p}d\tilde{\Phi}\int_{\Omega(\mathfrak{V})}dX(\tau(\rho;X,\tilde{\Phi})^{K+1}-\tau_0(\rho;X,\tilde{\Phi})^{K+1}).\end{aligned}$$

Obviously, it is enough to compute the part of (10.11) containing τ, since the second part (containing τ_0) can be computed analogously. We start by considering

$$\text{(10.12)}\qquad \int_{\Omega(\mathfrak{V})}\tau(\rho;X,\tilde{\Phi})^{K+1}dX.$$

First of all, we notice that if $\boldsymbol{\xi},\boldsymbol{\eta}\in\boldsymbol{\Xi}(\mathfrak{V})$ are equivalent points then, according to Lemma 5.7, all vectors $\boldsymbol{\theta}_j$ from the 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 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.

Denote by $\mathcal{M}_{\mathfrak{V}}$ the quotient space $\mathcal{M}_{\mathfrak{V}}:=\Omega(\mathfrak{V})/\leftrightarrow_{\mathfrak{V}}$. Since $\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})$ is a finite set for each $\boldsymbol{\nu}\in\Omega(\mathfrak{V})$, there is a natural measure on $\mathcal{M}_{\mathfrak{V}}$ generated by the Lebesgue measure on $\Omega(\mathfrak{V})$. Therefore, we can rewrite (10.12) as

$$\text{(10.13)}\qquad \int_{\mathcal{M}_{\mathfrak{V}}}\sum_{X\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})}\tau(\rho;X,\tilde{\Phi})^{K+1}d\boldsymbol{\nu}$$

and try to compute

$$\text{(10.14)}\qquad \sum_{X\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})}\tau(\rho;X,\tilde{\Phi})^{K+1}.$$

Let us denote by $S=S(r)$ the operator with symbol $2r\langle\mathbf{a},\mathbf{n}(\boldsymbol{\eta})\rangle+|\mathbf{a}|^2+w_{\tilde{k}}(\mathbf{x},\boldsymbol{\xi})+|X|^2$ acting in $\mathfrak{H}(X,\tilde{\Phi})$, so that $H_2(r)=r^2I+S(r)$.

Remark 10.1. We always assume that $\boldsymbol{\xi}\in\mathcal{A}$, so that $0.7\rho_n\le|\boldsymbol{\xi}|\le 5\rho_n$ and all functions $e_{\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\phi})$ from (9.19)–(9.21) 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 (8.15)) $\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\phi})=1$. This means that all cut-off functions from (9.19)–(9.21) 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 X. Thus, equations (9.19)–(9.21) show 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.

Formulas (9.19) and (9.20) imply

$$||S(r)|| \ll \rho_n^{1+\alpha_d+0+}, \quad ||S'(r)|| \ll \rho_n^{\alpha_d+0+}, \tag{10.15}$$

and

$$\Big\|\frac{d^l}{dr^l}S(r)\Big\| \ll \rho_n^{-l}, \quad l \geq 2. \tag{10.16}$$

Let $\gamma: \{|z-\rho| = \rho_n/8\}$ be a circle in the complex plane going in the positive direction. Then for $\rho \in I_n$, all $\tau(\rho; X, \tilde{\Phi})$ lie inside γ. It is not hard to see that estimates (10.15) and (10.16) hold inside and on γ. (Indeed, formulas (9.19)–(9.21) give matrix elements of $S(z)$ in an orthonormal basis even for complex z.)

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

$$\mathrm{tr}[F'(z)F^{-1}(z)] = (\det[F(z)])'(\det[F(z)])^{-1}.$$

(It can be proved, for example, using the expansion of the determinant along rows and the induction in the size of F.) Let $\#(S,\gamma)$ be the total number of zeros (counting multiplicity) of $\det[S(z) + z^2I - \rho^2I]$ inside γ. We have

(10.17)

$$\#(S,\gamma) = \frac{1}{2\pi i}\oint_\gamma (\det[S(z) + z^2I - \rho^2I])'(\det[S(z) + z^2I - \rho^2I])^{-1}dz$$
$$= \frac{1}{2\pi i}\oint_\gamma \mathrm{tr}[(2zI + S'(z))(S(z) + z^2I - \rho^2I)^{-1}]dz = \mathrm{tr}[(1 + O(\rho_n^{\alpha_d-1+0+}))I],$$

where $I = I_{\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})}$. Since $\mathrm{card}\,\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu}) \ll \rho_n^{(d-1)\alpha_{d-1}+0+}$ by Lemma 5.11 and $d\alpha_d < 1$, we conclude that there are precisely $\mathrm{card}\,\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})$ zeros (counting multiplicities) of $\det[S(z) + z^2I - \rho^2I]$ inside γ, and thus the points $\tau(\rho; X, \tilde{\Phi})$ are the only zeros of $\det[S(z) + z^2I - \rho^2I]$ inside γ.

Then by the residue theorem, we have

(10.18)

$$\sum_{X\in\boldsymbol{\Upsilon}_{\mathfrak{V}}(\boldsymbol{\nu})} \tau(\rho; X, \tilde{\Phi})^{K+1}$$
$$= \frac{1}{2\pi i}\oint_\gamma z^{K+1}(\det[S(z) + z^2I - \rho^2I])'(\det[S(z) + z^2I - \rho^2I])^{-1}dz$$
$$= \frac{1}{2\pi i}\oint_\gamma \mathrm{tr}[z^{K+1}(2zI + S'(z))(S(z) + z^2I - \rho^2I)^{-1}]dz$$
$$= \frac{1}{2\pi i}\oint_\gamma \mathrm{tr}[(2z^{K+2}I + z^{K+1}S'(z))(z^2 - \rho^2)^{-1}\sum_{l=0}^{\infty}(-1)^lS^l(z)(z^2 - \rho^2)^{-l}]dz$$

$$= \frac{1}{2\pi i}\sum_{l=0}^{\infty}(-1)^l \oint_\gamma \operatorname{tr}[(2z^{K+2}I + z^{K+1}S'(z))S^l(z)(z-\rho)^{-(l+1)}(z+\rho)^{-(l+1)}]dz$$
$$= \sum_{l=0}^{\infty}\frac{(-1)^l}{l!}\operatorname{tr}\frac{d^l}{dr^l}[(2r^{K+2}I + r^{K+1}S'(r))S^l(r)(r+\rho)^{-(l+1)}]\Big|_{r=\rho}.$$

Formula (10.11) shows that in order to compute the contribution to the density of states from $\boldsymbol{\Xi}(\mathfrak{V})_p$, we need to integrate the right-hand side of (10.18) against dX (or rather $d\boldsymbol{\nu}$) and $d\tilde{\Phi}$. We are going to integrate against $d\tilde{\Phi}$ first. We will prove that this integral 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 X, so afterwards we will just integrate these coefficients to obtain the desired asymptotic expansion.

Let us discuss how the right-hand side of (10.18) depends on the coordinates X and $\tilde{\Phi}$ (or rather Φ). Equations (9.19)–(9.21), Lemma 7.5, and Remark 10.1 show that the right-hand side of (10.18) is a sum of terms of the following form:

$$C\rho^p f_1(X)f_2(\Phi)f_3(X;\rho;\Phi). \tag{10.19}$$

Here, f_1 is a uniformly bounded function of X coordinates only. It consists of contributions from the cut-off functions $\varphi_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta}\in\mathfrak{V}$ and from the terms in (9.20), (9.21) corresponding to $\boldsymbol{\theta}'_j, \boldsymbol{\theta}''_j \in \mathfrak{V}$. The function $f_2(\Phi)$ is a product of powers of $\{\sin\Phi_q\}$. This function comes from differentiating (7.15) with respect to r. Finally, f_3 is of the following form:

$$f_3(X;\rho;\Phi) = \prod_{t=1}^{T}\left(l_t + \rho\sum_q b_q^t\sin(\Phi_q)\right)^{-k_t}. \tag{10.20}$$

This function corresponds to the negative powers of inner products $\langle\boldsymbol{\xi},\boldsymbol{\theta}_t\rangle$ given by Lemma 7.5, part (ii). Here, $\{b_q^t\}$ are coefficients in the decomposition $(\boldsymbol{\theta}_t)_{\mathfrak{V}^\perp} = \sum_q b_q^t\tilde{\boldsymbol{\mu}}_q$; recall that these numbers are all of the same sign and satisfy (7.16). Without loss of generality we will assume that all b_q^t are nonnegative. The number $l_t = l(b_1^t,\dots,b_{K+1}^t) := \langle X,(\boldsymbol{\theta}_t)_{\mathfrak{V}}\rangle + L_{m+1}\sum_q b_q^t$ satisfies $\rho_n^{\alpha_{m+1}}\rho_n^{0-} \ll l_t \ll \rho_n^{\alpha_{m+1}}\rho_n^{0+}$, since our assumptions imply $|\langle X,\boldsymbol{\theta}_{\mathfrak{V}}\rangle| \ll \rho_n^{\alpha_m}$. This number depends on X, but not on Φ or ρ. The number $k_t = k(b_1^t,\dots,b_{K+1}^t)$ is positive, integer, and independent of $\boldsymbol{\xi}$. Our next objective is to compute the integrals of (10.19) over the domain $\{\tilde{\Phi}\in M_p\}$ and prove that these integrals enjoy asymptotic behaviour (3.9) with uniformly bounded coefficients (as functions of X). The calculations will be rather messy technically, although the main ideas of computing them are not too difficult.

10.2. *Computing the model integral.* Before computing the integral of (10.19), we will deal with a simpler integral,

$$J_K := \int_0^\gamma \int_0^{\hat\Phi_K} \dots \int_0^{\hat\Phi_2} \frac{\hat\Phi_1^{n_1} \dots \hat\Phi_K^{n_K}\, d\hat\Phi_1 \dots d\hat\Phi_K}{\prod_{t=1}^T (l_t + \rho \sum_{j=1}^K b_j^t \hat\Phi_j)^{k_t} (c_t + \sum_{j=1}^K \tilde b_j^t \hat\Phi_j)^{k_t'}}, \tag{10.21}$$

and then we will discuss how to reduce our initial integral to (10.21). Here, $n_j, k_t, k_t' \in (\mathbb{N} \cup \{0\})$, $\gamma \le 1$, $\rho_n^\beta \ll \rho_n^{\alpha_1+0-} \ll l_t \ll \rho_n^{\alpha_d+0+} \ll \rho_n^{1/2}$, $0 < c_t \ll \rho_n^{1/2}$, $\rho_n^{-\delta_0} \ll b_j^t \ll \rho_n^{\delta_0}$, and $\rho_n^{-\delta_0} \ll \tilde b_j^t \ll \rho_n^{\delta_0}$, where $\delta_0 > 0$ is sufficiently small. (For the sake of definiteness, we put $\delta_0 := \frac{1}{3^d 3000}$; obviously, we assume that these inequalities hold only for nonzero values of b_j^t and $\tilde b_j^t$.) We introduce the following notation:

$$P := \sum_j n_j, \qquad Q := \sum_t k_t, \qquad Q' := \sum_t k_t', \tag{10.22}$$

and we sometimes will denote the integral (10.21) as $J_K(P, Q, Q')$. We will also need the auxiliary positive numbers $p_j,\ q_j,\ j = 0, \dots, d$, defined by

$$q_d = \frac{1}{3^d 300}, \quad p_j = q_j + \frac{1}{3^d 300}, \quad q_{j-1} = q_j + p_j + \frac{1}{3^d 300} = 2p_j.$$

Obviously, $p_0 < 1/100$.

LEMMA 10.2. *Assume that $c_t \gg \rho_n^{-q_K}$. Then, we have*

$$J_K(P, Q, Q') = \sum_{q=0}^K (\ln(\rho))^q \sum_{p=0}^\infty e(p, q; P, Q, Q') \rho^{-p}, \tag{10.23}$$

where

$$|e(p, q; P, Q, Q')| \ll \rho_n^{(2/3 - p_K)p} \rho_n^{-Q\beta} 2^{Q'} \prod_{t=1}^T c_t^{-k_t'}. \tag{10.24}$$

These estimates are uniform in the following regions of variables:

$$\begin{aligned} &0 \le \gamma \le 1,\ \rho_n^\beta \ll l_t \ll \rho_n^{1/2},\ \rho_n^{-q_K} \ll c_t \ll \rho_n^{1/2}, \\ &\rho_n^{-\delta_0} \ll b_j^t \ll \rho_n^{\delta_0},\ \rho_n^{-\delta_0} \ll \tilde b_j^t \ll \rho_n^{\delta_0},\ \rho_n^{2/3 - q_K} < \rho. \end{aligned} \tag{10.25}$$

Remark 10.3. The estimates (10.24) are more natural than they may look. Indeed, each time we run the inductive argument in the proof (i.e., each time we increase K), we apply the geometric series expansion, which results in a slight worsening of the estimates. This accounts for the need to have p_K in the exponent of ρ_n.

Proof. The proof will go by induction in K. The base of induction ($K = 0$) is trivial (and the case $K = 1$ has been discussed in [11]). Suppose we have proved this statement for $K = S - 1$, and let us prove it for $K = S$.

Step I. First, we consider the area where $\hat{\Phi}_S \geq \rho_n^{-p_S}$. We do not change terms in the denominator of J_K where $b_S^t = 0$. If $b_S^t \neq 0$, then we proceed with the following transformations:

$$\begin{aligned}(10.26)\quad \Big(\rho\sum_j b_j^t\hat{\Phi}_j+l_t\Big)^{-k_t} &= \Big(\rho\sum_j b_j^t\hat{\Phi}_j\Big)^{-k_t}\Big(1+\frac{l_t}{\rho\sum_j b_j^t\hat{\Phi}_j}\Big)^{-k_t}\\ &= \Big(\rho\sum_j b_j^t\hat{\Phi}_j\Big)^{-k_t}\sum_{m=0}^{\infty}\binom{m+k_t-1}{m}\Big(\frac{-l_t}{\rho\sum_j b_j^t\hat{\Phi}_j}\Big)^{m}\\ &= \sum_{m=0}^{\infty} C_m(k_t)\Big(\frac{1}{\rho\sum_j b_j^t\hat{\Phi}_j}\Big)^{m+k_t},\end{aligned}$$

where constants $C_m(k_t)$ satisfy the estimate

$$(10.27)\qquad |C_m(k_t)| \leq \rho_n^{m/2}\binom{m+k_t-1}{m} \leq \rho_n^{m/2}2^{m+k_t-1}.$$

Now, we can move powers of ρ out of the integral and denote $\tilde{c}_t := b_S^t\hat{\Phi}_S$. Obviously, $\tilde{c}_t$ satisfies (10.25) with $K = S-1$. For terms that do not contain ρ, we just denote $\hat{c}_t := c_t + \tilde{b}_S^t\hat{\Phi}_S$. Then, we can apply the induction assumption for $K = S-1$. Corresponding coefficients will depend on $\hat{\Phi}_S$ uniformly. As a result, we obtain the following expression as the contribution to J_K from the region $\{\hat{\Phi}_S \geq \rho_n^{-p_S}\}$ (we denote $m := m_1 + \cdots + m_T$ and assume for simplicity that $b_K^t \neq 0$ for all t):

$$(10.28)\qquad \sum_{m_1=0}^{\infty}\cdots\sum_{m_T=0}^{\infty} C_{m_1}(k_1)\dots C_{m_T}(k_T)\rho^{-m-Q}J_{S-1}(P,0,m+Q+Q').$$

Of course, each time we write $J_{S-1}(P,0,m+Q+Q')$, it denotes a different integral of the form (10.21), but by the assumption of induction all of them satisfy

$$(10.29)\quad J_{S-1}(P,0,Q+m+Q') = \sum_{q=0}^{S-1}(\ln(\rho))^q\sum_{p=0}^{\infty}\breve{e}(p,q;P,0,Q+m+Q')\rho^{-p}$$

with

$$(10.30)\quad |\breve{e}(p,q;P,0,Q+m+Q')| \ll \rho_n^{(2/3-p_{S-1})p}\rho_n^{(Q+m)q_0}2^{Q+m+Q'}\prod_{t=1}^{T}c_t^{-k_t'}.$$

(When we use (10.24) as the induction hypothesis, we replace $\tilde{c}_t$ and $\hat{c}_t$ by the corresponding lower bounds $\rho_n^{-q_0}$ and c_t.)

Contribution from this region of integration into coefficients $e(p,q)$ of the integral J_S can therefore be estimated from above by the following expression:

(10.31)
$$\sum_{m=0}^{p-Q} \rho_n^{(2/3-p_{S-1})(p-m-Q)} \times \rho_n^{(Q+m)q_0} \rho_n^{m/2} 2^{m+Q-T} 2^{Q+m+Q'} \left(\prod_{t=1}^{T} c_t^{-k'_t}\right) \sum_{m_1+\dots+m_T=m,\ m_j\geq 0} 1$$
$$\leq \rho_n^{(2/3-p_S)p} \rho_n^{-Q\beta} 2^{Q'} \left(\prod_{t=1}^{T} c_t^{-k'_t}\right) \times \rho_n^{Q\beta+Qq_0-(2/3-p_{S-1})Q} 2^{2Q-T} \sum_{m=0}^{\infty} \rho_n^{mq_0+m/2-(2/3-p_{S-1})m} 2^{2m} 2^{m+T-1}$$
$$\ll \rho_n^{(2/3-p_S)p} \rho_n^{-Q\beta} 2^{Q'} \prod_{t=1}^{T} c_t^{-k'_t}.$$

Notice that at this step we have $S-1$ as the largest power of $\ln(\rho)$.

Step II. From now we are in the area $\hat{\Phi}_S \leq \rho_n^{-p_S}$. Then we can transform all terms not containing ρ in the denominator of J_K:

(10.32)
$$\left(\sum_j \tilde{b}_j^t \hat{\Phi}_j + c_t\right)^{-k'_t} = (c_t)^{-k'_t} \left(1 + \frac{\sum_j \tilde{b}_j^t \hat{\Phi}_j}{c_t}\right)^{-k'_t} = \sum_{m=0}^{\infty} C'_m(k'_t) \left(\sum_j \tilde{b}_j^t \hat{\Phi}_j\right)^m,$$

where

(10.33)
$$|C'_m(k'_t)| \leq c_t^{-k'_t-m} 2^{m+k'_t-1} \leq \rho_n^{q_S m} c_t^{-k'_t} 2^{m+k'_t-1}.$$

Then, only terms with ρ are left in the denominator. We change variables $x_j := \hat{\Phi}_j \rho$ and obtain the integral

(10.34)
$$\rho^{-P-S} \int_0^{\rho\rho_n^{-p_S}} \int_0^{x_S} \dots \int_0^{x_2} \frac{x_1^{n_1} \dots x_S^{n_S}\, dx_1 \dots dx_S}{\prod_t (l_t + \sum_{j=1}^S b_j^t x_j)^{k_t}} \prod_t \sum_{m_t=0}^{\infty} \rho^{-m_t} C'_{m_t}(k'_t) \left(\sum_{j=1}^{S} \tilde{b}_j^t x_j\right)^{m_t}.$$

First, we consider the integral along the region where $0 \leq x_S \leq \rho_n^{2/3-q_{S-1}}$. Obviously, the corresponding contribution to J_S can be computed as

$$\rho^{-P-S} \sum_{m=0}^{\infty} \rho^{-m} \tilde{C}_m,$$

where

$$|\tilde{C}_m| \le \rho_n^{(2/3-q_{S-1})(P+S+m)}\rho_n^{\delta_0 m}S^m\rho_n^{-Q\beta}2^{m+T-1}2^{m+Q'-T}\rho_n^{q_S m}\prod_{t=1}^{T}c_t^{-k_t'}.$$

Thus, if we put $m = p - P - S$, we obtain the following estimate for the coefficient in front of ρ^{-p} (notice that $p_S < q_{S-1} - q_S - 2\delta_0$ for $S \ge 1$):

$$\rho_n^{(2/3-p_S)p}\rho_n^{-Q\beta}\rho_n^{-(P+S)q_S}2^{Q'}\prod_{t=1}^{T}c_t^{-k_t'}.$$

Step III. *The case when* $x_S \in (\rho_n^{2/3-q_{S-1}}, \rho\rho_n^{-p_S})$. Once again, if $b_S^t = 0$, then we leave such terms unchanged. If $b_S^t \neq 0$ and thus $b_S^t \ge \rho_n^{-\delta_0}$, we perform the following transform to (10.34) (cf. (10.26)):

$$\left(l_t + \sum_{j=1}^{S} b_j^t x_j\right)^{-k_t} = \sum_{m=0}^{\infty} C_m(k_t)\left(\frac{1}{\sum_j b_j^t x_j}\right)^{m+k_t} \tag{10.35}$$

and introduce new variables $z_j := x_j/x_S$, $j = 1, \dots, S-1$. Thus, we reduce the problem to the integrals of the following form:

$$\int_{\rho_n^{2/3-q_{S-1}}}^{\rho\rho_n^{-p_S}}\int_0^1\int_0^{z_{S-1}}\cdots\int_0^{z_2}\frac{z_1^{\tilde{n}_1}\cdots z_{S-1}^{\tilde{n}_{S-1}}x_S^{\tilde{n}_S}\,dz_1\cdots dz_{S-1}dx_S}{\prod_t(l_t + x_S\sum_{j=1}^{S-1}b_j^t z_j)^{\tilde{k}_t}(b_S^t x_S + x_S\sum_{j=1}^{S-1}b_j^t z_j)^{\tilde{k}_t'}}.$$

Now we can remove x_S from the second bracket in the denominator and then apply the induction assumption for the internal $S-1$ integrals (i.e., the integrals against $dz_1 \dots dz_{S-1}$ with $c_t := b_S^t$ and $\rho := x_S$). This induction assumption guarantees that these internal integrals can be expressed as a series in powers of x_S and $\ln x_S$, with the biggest power of $\ln x_S$ being $S-1$. Then, we multiply this expansion by a (possibly negative) power of x_S and integrate the product against dx_S. As a result, we obtain a decomposition (10.23) of J_S, with the biggest power of $\ln \rho$ being equal to S. The estimate of the contribution of Step III to the coefficients $e(p,q)$ is similar (but rather more tedious) to the estimates in the first two steps, and we will skip it. □

10.3. *Reduction to the model integral.* Now we will discuss how to deal with our initial integral

$$\hat{J}_K := \int_{M_p}\frac{(\sin\Phi_1)^{n_1}\cdots(\sin\Phi_K)^{n_K}(\sin\Phi_{K+1})^{n_{K+1}}\,d\tilde{\Phi}}{\prod_{t=1}^{T}(l_t + \rho\sum_{j=1}^{K+1}b_j^t\sin\Phi_j)^{k_t}}. \tag{10.36}$$

The main problem with reducing the integral along M_p (or even along $M_p \cap \{\Phi_1 \le \cdots \le \Phi_K \le \Phi_{K+1}\}$) to the model integral (10.21) is the limits of integration: the upper limit of integration against $d\Phi_K$ is not a constant (since the collection of points where $\Phi_K = \Phi_{K+1}$ has variable coordinate Φ_K). In order to rectify this, we define

(10.37)
$$\hat{M} := \{\sin\Phi_1 \le \dots \le \sin\Phi_{s-1} \le \rho_n^{-p_d},\ \rho_n^{-p_d} \le \sin\Phi_q,\ q = s, \dots, K+1\}.$$

It is clear that M_p can be represented as a union of several domains of this type. Lemma 7.4 shows that we always have at least one 'large' variable in $\hat{M}$, i.e., $s \le K+1$. We also introduce the 'spherical' coordinates in the $(\Phi_s, \dots, \Phi_{K+1})$-subspace: we put

$$\begin{aligned} \sin\Phi_j &= \hat{\Phi}_j, \quad j = 1, \dots, s-1, \\ \sin\Phi_s - \rho_n^{-p_d} &= \hat{r}\cos\hat{\Phi}_s, \\ \sin\Phi_{s+1} - \rho_n^{-p_d} &= \hat{r}\sin\hat{\Phi}_s\cos\hat{\Phi}_{s+1}, \\ &\dots \\ \sin\Phi_K - \rho_n^{-p_d} &= \hat{r}\sin\hat{\Phi}_s\sin\hat{\Phi}_{s+1}\cdots\sin\hat{\Phi}_{K-1}\cos\hat{\Phi}_K, \\ \sin\Phi_{K+1} - \rho_n^{-p_d} &= \hat{r}\sin\hat{\Phi}_s\sin\hat{\Phi}_{s+1}\cdots\sin\hat{\Phi}_{K-1}\sin\hat{\Phi}_K, \end{aligned} \tag{10.38}$$

so that $(\hat{r}, \hat{\Phi}_1, \dots, \hat{\Phi}_K)$ are the new coordinates. Since only K of the coordinates Φ_j were independent, we can consider the new variables $(\hat{\Phi}_1, \dots, \hat{\Phi}_K)$ as independent (and $\hat{r}$ as a function of the independent variables). We remind the reader (see Lemma 7.4) that we always have $\max_j \sin\Phi_j \gg \rho_n^{0-}$, and thus

$$\rho_n^{0-} \ll \hat{r} \ll 1. \tag{10.39}$$

The point in introducing these variables is that the limits of integration over $\hat{M}$ become simple:

(10.40)
$$\int_0^{\pi/2} d\hat{\Phi}_K \int_0^{\pi/2} d\hat{\Phi}_{K-1} \int_0^{\pi/2} d\hat{\Phi}_s \int_0^{\rho_n^{-p_d}} d\hat{\Phi}_{s-1} \int_0^{\hat{\Phi}_{s-1}} d\hat{\Phi}_{s-2} \dots \int_0^{\hat{\Phi}_2} d\hat{\Phi}_1.$$

When we insert the values of $\sin\Phi_q$, $q = 1, \dots, K+1$ given by (10.38) into (7.12), we obtain the quadratic equation for finding $\hat{r}$:

$$\hat{A}_s\hat{r}^2 + 2\hat{B}_s\hat{r} + (\hat{C}_s - 1) = 0, \tag{10.41}$$

where

$$\begin{aligned} \hat{A}_s = \sum_j &(a_{js}\cos\hat{\Phi}_s + a_{j\,s+1}\sin\hat{\Phi}_s\cos\hat{\Phi}_{s+1} \\ &+ \dots + a_{j\,K+1}\sin\hat{\Phi}_s\sin\hat{\Phi}_{s+1}\cdots\sin\hat{\Phi}_{K-1}\sin\hat{\Phi}_K)^2 > 0, \end{aligned} \tag{10.42}$$

(10.43)

$$\hat{B}_s = \sum_j \left(\sum_{q=1}^{s-1} a_{jq}\hat{\Phi}_q + \rho_n^{-p_d} \sum_{q=s}^{K+1} a_{jq} \right) (a_{js}\cos\hat{\Phi}_s + a_{j\,s+1}\sin\hat{\Phi}_s\cos\hat{\Phi}_{s+1} + \cdots + a_{j\,K+1}\sin\hat{\Phi}_s\sin\hat{\Phi}_{s+1}\cdots\sin\hat{\Phi}_{K-1}\sin\hat{\Phi}_K),$$

and

$$\hat{C}_s = \sum_j \left(\sum_{q=1}^{s-1} a_{jq}\hat{\Phi}_q + \rho_n^{-p_d} \sum_{q=s}^{K+1} a_{jq} \right)^2 > 0. \tag{10.44}$$

Therefore, we have

$$\hat{r} = \frac{-\hat{B}_s \pm \sqrt{\hat{B}_s^2 - \hat{A}_s\hat{C}_s + \hat{A}_s}}{\hat{A}_s}. \tag{10.45}$$

Note that the Cauchy-Schwarz inequality implies $\hat{A}_s\hat{C}_s \geq \hat{B}_s^2$. Since equation (10.41) obviously has at least one real solution, we have $0 \leq \hat{A}_s + \hat{B}_s^2 - \hat{A}_s\hat{C}_s \leq \hat{A}_s$. Next, due to Lemma 7.3, we have

$$|\hat{B}_s| \ll \rho_n^{-p_d+0+}, \quad |\hat{C}_s| \ll \rho_n^{-2p_d+0+}. \tag{10.46}$$

Then using (10.39), (10.41), and (10.46), we get

$$\rho_n^{0+} \gg \frac{2}{\hat{r}^2} \geq \hat{A}_s \geq \frac{1}{2\hat{r}^2} \gg 1. \tag{10.47}$$

Since $\hat{r}$ is positive, we obviously have

$$\hat{r} = \frac{-\hat{B}_s + \sqrt{\hat{B}_s^2 - \hat{A}_s\hat{C}_s + \hat{A}_s}}{\hat{A}_s} = \frac{-\hat{B}_s + \sqrt{\hat{A}_s}\sqrt{1 - (\hat{C}_s - \hat{B}_s^2\hat{A}_s^{-1})}}{\hat{A}_s},$$

and thus $\hat{r}$ is *analytic* with respect to $\hat{B}_s$, $\hat{C}_s$, i.e., with respect to all $\hat{\Phi}_j$, $j = 1, \ldots, s-1$, *uniformly* in $\hat{\Phi}_l$, $l = s, \ldots, K$, inside $\hat{M}$. It is easy to see from (7.13), (7.14), and (10.38) that the same is true for the Jacobian $\frac{\partial(\tilde{\Phi})}{\partial(\hat{\Phi}_1,\ldots,\hat{\Phi}_K)}$.

We also notice that if we denote

$$\hat{r}_0 := \hat{r} - \hat{A}_s^{-1/2}, \tag{10.48}$$

then $\hat{r}_0$ satisfies the same analyticity properties as $\hat{r}$ and $\hat{r}_0 = O(\rho_n^{-p_d+0+})$.

Thus, we arrive at the integrals of the following form:

(10.49)

$$\int_0^{\pi/2} d\hat{\Phi}_K \int_0^{\pi/2} d\hat{\Phi}_{K-1} \cdots \int_0^{\pi/2} d\hat{\Phi}_s \times \int_0^{\rho_n^{-p_d}} d\hat{\Phi}_{s-1} \int_0^{\hat{\Phi}_{s-1}} d\hat{\Phi}_{s-2} \cdots \int_0^{\hat{\Phi}_2} d\hat{\Phi}_1 \frac{F(\hat{\Phi}_s,\ldots,\hat{\Phi}_K)\hat{\Phi}_1^{n_1}\cdots\hat{\Phi}_{s-1}^{n_{s-1}}}{\prod_{t=1}^T (l_t + \rho S(\hat{\Phi}_1,\ldots,\hat{\Phi}_K))^{k_t}}.$$

Here the function $F(\hat{\Phi}_s,\dots,\hat{\Phi}_K)$ is uniformly bounded with respect to $\hat{\Phi}_s,\dots,\hat{\Phi}_K$ in $\hat{M}$ and

(10.50)
$$S = S(\hat{\Phi}_1,\dots,\hat{\Phi}_K) := \sum_{j=1}^{s-1} b_j^t \hat{\Phi}_j + \sum_{j=s}^{K+1} b_j^t \rho_n^{-p_d} + (\hat{A}_s^{-1/2} + \hat{r}_0)\tilde{F}(\hat{\Phi}_s,\dots,\hat{\Phi}_K),$$

where

$$\tilde{F} := b_s^t \cos\hat{\Phi}_s + b_{s+1}^t \sin\hat{\Phi}_s \cos\hat{\Phi}_{s+1} + \dots + b_{K+1}^t \sin\hat{\Phi}_s \sin\hat{\Phi}_{s+1}\cdots\sin\hat{\Phi}_{K-1}\sin\hat{\Phi}_K.$$

Now we apply the construction from Step I of the proof of Lemma 10.2. We do not change terms $(l_t+\rho S)$ with $b_j^t = 0$ for all $j = s,\dots,K+1$ (such terms are equal to $(l_t + \rho\sum_{j=1}^{s-1} b_j^t\hat{\Phi}_j)$). Otherwise, we write (cf. (10.26))

$$(l_t+\rho S)^{-k_t} = \sum_{m=0}^{\infty} C_m(k_t)\left(\frac{1}{\rho S}\right)^{m+k_t}.$$

It remains to notice that

$$S^{-1} = \Big(\sum_{j=1}^{s-1} b_j^t\hat{\Phi}_j + \sum_{j=s}^{K+1} b_j^t\rho_n^{-p_d} + \hat{A}_s^{-1/2}\tilde{F}\Big)^{-1} \times \left(1 + \frac{\hat{r}_0\tilde{F}}{\sum_{j=1}^{s-1} b_j^t\hat{\Phi}_j + \sum_{j=s}^{K+1} b_j^t\rho_n^{-p_d} + \hat{A}_s^{-1/2}\tilde{F}}\right)^{-1}$$

and decompose the last expression using geometric progression. We remind the reader (see (10.47) and (10.48)) that $\hat{r}_0\hat{A}_s^{1/2} \ll \rho_n^{-p_d+0+}$. Since $\hat{r}_0$ is analytic in $\hat{\Phi}_1,\dots,\hat{\Phi}_{s-1}$, we end up with the model integrals J_{s-1} (with $c_t := \sum_{j=s}^{K+1} b_j^t\rho_n^{-p_d} + \hat{A}_s^{-1/2}\tilde{F}$) uniformly depending on the parameters $\hat{\Phi}_s,\dots,\hat{\Phi}_{K+1}$. Summing this and using Lemma 10.2, we have proved the following result.

Lemma 10.4. *We have*

(10.51)
$$\hat{J}_K = \sum_{q=0}^{K}(\ln(\rho))^q\sum_{p=0}^{\infty} e(p,q)\rho^{-p},$$

where

(10.52)
$$|e(p,q)| \ll \rho_n^{(2/3-p_K)p}\rho_n^{-Q\beta}.$$

These estimates are uniform in the following regions of variables:

(10.53)
$$\rho_n^{\beta} \ll l_t \ll \rho_n^{1/2}, \quad \rho_n^{-\delta_0} \ll b_j^t \ll \rho_n^{\delta_0}, \quad \rho_n^{2/3-q_K} < \rho.$$

Now Lemma 10.4, Remark 10.1, and equation (10.18) show that the integral (10.11) admits decomposition of the form (10.51) for $0.7\rho_n < \rho < 5\rho_n$. This, together with equations (10.11), (10.2), (6.16), Corollary 4.3, and the observation that the number of different quasi-lattice subspaces $\mathfrak{V}$ is $\le \rho_n^{0+}$, completes the proof of Lemma 3.3 and, thus, of our main theorem in the case of all domains M_p being simplexes. It remains to discuss how to reduce the case of general region $\Xi(\mathfrak{V})_p$ to the case of a simplex.

11. **Integration in nonsimplex domains**

Now let us consider the case when the number J_p of defining hyperplanes is bigger than $K+1$. Recall that we have

(11.1)
$$\boldsymbol{\Xi}(\mathfrak{V})_p = \{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V}) \ \text{ and } \ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > L_{m+1},\ j = 1, \dots, J_p\}.$$

In this section, we give only a brief description of how to prove the main statements, since the complete proof would be too long and tedious. The complexity of the proof is mostly caused by the fact that we need to describe a procedure working in all dimensions. If one is interested only in the cases $d = 2$ or $d = 3$, the proofs become substantially simpler. (Indeed, the case $d = 2$ has already been proved, since then we have $K = 1$ or $K = 0$, and any polyhedron in dimensions 0 or 1 is obviously a simplex.)

Consider first the case when $\boldsymbol{\Xi}(\mathfrak{V})_p$ is a 'cone,' i.e., there is a point $\mathbf{a} = \mathbf{a}(p) \in \mathfrak{V}^\perp$ such that $\langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_j(p)\rangle = L_{m+1}$, $j = 1, \dots, J_p$. (Such point, if it exists, is always unique, and the existence of it is automatic in the simplex case, i.e., when $J_p = K + 1$.) Then we can introduce the coordinates $(r, \tilde{\Phi})$ by formulas (7.6) and (7.7) with $\tilde{\Phi} \in M_p$, where M_p is still given by (7.8).

Lemma 11.1. *Suppose $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}}$. Then we can write $\boldsymbol{\theta}_{\mathfrak{V}^\perp} = \sum_q \tilde{b}_q \tilde{\boldsymbol{\mu}}_q(p)$, where either all $\tilde{b}_q$ are nonpositive, or all of them are nonnegative.* (*But such a decomposition is not necessarily unique.*)

Proof. Our assumptions imply that for each $\boldsymbol{\eta} \in \tilde{\boldsymbol{\Xi}}(\mathfrak{V})_p$ (where $\tilde{\boldsymbol{\Xi}}(\mathfrak{V})_p$ is defined by (7.4)), we have $\langle \boldsymbol{\eta}, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle \neq 0$. Assume for definiteness that $\langle \boldsymbol{\eta}, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle > 0$. This property can be reformulated like this: whenever $\mathbf{z} \in \mathfrak{V}^\perp$ is a vector with $\langle \mathbf{z}, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle < 0$, there is at least one vector $\tilde{\boldsymbol{\mu}}_j(p)$ such that $\langle \mathbf{z}, \tilde{\boldsymbol{\mu}}_j(p)\rangle < 0$. This, in turn, is equivalent to saying that whenever $\mathbf{z} \in \mathfrak{V}^\perp$ is a vector with $\langle \mathbf{z}, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle > 0$, there is at least one vector $\tilde{\boldsymbol{\mu}}_j(p)$ such that $\langle \mathbf{z}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > 0$. Consider the set $S := \{\mathbf{z} = \sum_q \tilde{b}_q \tilde{\boldsymbol{\mu}}_q(p), \tilde{b}_q \ge 0\}$. We need to prove that $\boldsymbol{\theta}_{\mathfrak{V}^\perp} \in S$. Suppose not. Let $\mathbf{z}_0$ be a nearest to $\boldsymbol{\theta}_{\mathfrak{V}^\perp}$ point from S. Then $\langle \boldsymbol{\theta}_{\mathfrak{V}^\perp} - \mathbf{z}_0, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle > 0$, because otherwise the point $|\boldsymbol{\theta}_{\mathfrak{V}^\perp}|^2 \langle \mathbf{z}_0, \boldsymbol{\theta}_{\mathfrak{V}^\perp}\rangle^{-1} \mathbf{z}_0 \in S$ is closer to $\boldsymbol{\theta}_{\mathfrak{V}^\perp}$ than $\mathbf{z}_0$. Thus, there is a value of j, say $j = 1$, so that $\langle \boldsymbol{\theta}_{\mathfrak{V}^\perp} - \mathbf{z}_0, \tilde{\boldsymbol{\mu}}_1(p)\rangle > 0$. But then for sufficiently small $\epsilon > 0$, the vector $\mathbf{z}_0 + \epsilon \tilde{\boldsymbol{\mu}}_1(p) \in S$ is closer to $\boldsymbol{\theta}_{\mathfrak{V}^\perp}$ than $\mathbf{z}_0$. This contradiction proves our lemma. □

This result shows that there is a big similarity between cases of the cone and the simplex. The only difference from the simplex case is that the number of sides of M_p is now greater than $K+1$. This, however, means that if we introduce the angular coordinates Φ in the same way as in Section 7, we will have difficulties trying to get rid of several of them. Instead, we will follow a different strategy. We will cut the spherical polyhedron M_p into several simplexes,

$$M_p = \sqcup_q {}^q M_p, \tag{11.2}$$

and then perform integration over each simplex ${}^q M_p$ in the same way as we did in Sections 7–10. The only thing we need to make sure of is that the lengths of all sides (edges) of ${}^q M_p$, as well as all nonzero angles between two sides of any dimensions of ${}^q M_p$, are $\gg \rho_n^{0-}$. (Let us call this *angles and sides property.*) However, we know that the angles and sides property holds for the original polyhedron M_p because of Lemma 7.1. (Strictly speaking, Lemma 7.1 was proved for simplexes, rather than for cones, but the proof is the same.) Thus, the only problem we face is how to cut a polyhedron M_p into simplexes ${}^q M_p$ without drastically decreasing sides or angles. We do it by induction in K. For $K=1$, the statement is obvious. (Each 1-dimensional connected polyhedron is a simplex, i.e., an interval.) Assume that K is arbitrary. Also assume for simplicity that M_p is not a spherical, but a Euclidean polyhedron. (We can achieve this by projecting M_p onto any hyperplane tangent to it; obviously, this projection keeps the angles and sides property invariant.)

Step I. We find a simplex $\hat{M}_p \subset M_p$ satisfying the angles and sides property. To do this, we consider a ball centered at any vertex $\mathbf{v}$ of M_p of radius $\gg \rho_n^{0-}$, but sufficiently small so that the intersection of this ball with M_p is a cone. The intersection of the boundary of this ball with M_p is a polyhedron N_p of dimension $K-1$. Running the induction argument, we can find a $(K-1)$-dimensional simplex $\hat{N}_p \subset N_p$ satisfying the angles and sides property. Now we define $\hat{M}_p$ as a convex hull of $\mathbf{v}$ and vertexes of $\hat{N}_p$. A straightforward geometrical argument implies that $\hat{M}_p$ satisfies the angles and sides property. In particular, the volume of $\hat{M}_p$ is $\gg \rho_n^{0-}$.

Step II. We find a point $\boldsymbol{\eta}^* \in M_p$ such that the distance from $\boldsymbol{\eta}^*$ to each of the $(K-1)$-dimensional sides of M_p is $\gg \rho_n^{0-}$. (The distance to the side is the length of the perpendicular dropped to the hyperplane containing this side.) This point can be chosen to be the center of gravity of $\hat{M}_p$. Indeed, the distance from $\boldsymbol{\eta}^*$ to a $(K-1)$-dimensional side of M_p is the average of the distances from the vertexes of $\hat{M}_p$ to this side. Thus, we need to show that the distance from at least one of the vertexes of $\hat{M}_p$ to this side is $\gg \rho_n^{0-}$. But if this were not the case, then the breadth of $\hat{M}_p$ in the direction orthogonal to that side would be $\ll \rho_n^{0-}$, and so the volume of $\hat{M}_p$ would be $\ll \rho_n^{0-}$, which would contradict estimates from Step I.

Step III. Now we use the inductive assumption and cut each $(K-1)$-dimensional side of M_p into simplexes. Taking convex hulls of $\boldsymbol{\eta}^*$ with these simplexes, we obtain the required decomposition of M_p into simplexes ${}^qM_p$. It is a geometric exercise to check that simplexes ${}^qM_p$ constructed in this way satisfy the angles and sides property.

Let us denote by ${}^q\boldsymbol{\Xi}(\mathfrak{V})_p$ the infinite cone with the vertex $\mathbf{a}$ and a cross-section ${}^qM_p$, i.e.,

$$(11.3) \qquad {}^q\boldsymbol{\Xi}(\mathfrak{V})_p := \mathbf{a} + \{\boldsymbol{\xi} \in \mathbb{R}^d, \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V}) \text{ and } \mathbf{n}(\boldsymbol{\xi}_{\mathfrak{V}^\perp}) \in {}^qM_p\}.$$

Then we obviously have

$$(11.4) \qquad \boldsymbol{\Xi}(\mathfrak{V})_p = \sqcup_q {}^q\boldsymbol{\Xi}(\mathfrak{V})_p.$$

Now let us discuss how to perform the integration over ${}^q\boldsymbol{\Xi}(\mathfrak{V})_p$. Let us fix q and p and denote by $\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_{K+1}$ the interior unit normal vectors to the faces of ${}^q\boldsymbol{\Xi}(\mathfrak{V})_p$. We denote, as before, $\Phi_q := \frac{\pi}{2} - \phi(\boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}, \boldsymbol{\nu}_q(p))$, $q = 1, \dots, K+1$. In order to perform the integration, we need to check that Lemma 7.5 is still valid in the cone case. So, let $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\tilde{k}}$. Applying Lemma 11.1, we deduce that

$$(11.5) \qquad \boldsymbol{\theta}_{\mathfrak{V}^\perp} = \sum_q \tilde{b}_q \tilde{\boldsymbol{\mu}}_q(p),$$

where either all $\tilde{b}_q$ are nonpositive, or all of them are nonnegative; assume for definiteness that all of them are nonnegative. We also have (applying, for example, the same lemma) that each vector $\tilde{\boldsymbol{\mu}}_q(p)$ admits a decomposition $\tilde{\boldsymbol{\mu}}_q(p) = \sum_{l=1}^{K+1} \hat{b}_{ql} \boldsymbol{\nu}_l$ with all coefficients $\hat{b}_{ql}$ being nonnegative. Now (denoting, as usual, $\boldsymbol{\eta} := \boldsymbol{\xi}_{\mathfrak{V}^\perp} - \mathbf{a}$ and putting $b_l := \sum_{q=1}^{J_p} \tilde{b}_q \hat{b}_{ql} \geq 0$, $l = 1, \dots, K+1$), we have

$$(11.6) \qquad \begin{aligned} \langle \boldsymbol{\xi}, \boldsymbol{\theta} \rangle &= \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \langle \mathbf{a}, \boldsymbol{\theta}_{\mathfrak{V}^\perp} \rangle + \langle \boldsymbol{\eta}, \boldsymbol{\theta}_{\mathfrak{V}^\perp} \rangle \\ &= \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \sum_{q=1}^{J_p} \tilde{b}_q \langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_q \rangle + \sum_{l=1}^{K+1} b_l \langle \boldsymbol{\eta}, \boldsymbol{\nu}_l \rangle \\ &= \langle X, \boldsymbol{\theta}_{\mathfrak{V}} \rangle + \sum_q \tilde{b}_q L_{m+1} + r \sum_l b_l \sin \Phi_l. \end{aligned}$$

The estimates $\rho_n^{0-} \leq |b_q| \leq \rho_n^{0+}$ can be proved in the same way as Lemma 2.8. Finally, (11.5) implies that $\sum_q \tilde{b}_q \gg \rho_n^{0-}$ for any $\boldsymbol{\theta} \notin \mathfrak{V}$. Multiplying (11.5) by $\mathbf{a}$, we deduce that $\sum_q \tilde{b}_q \ll \rho_n^{0+}$.

This finishes the proof for the cone case. Now let us discuss the general case.

Let $\mathbf{a}$ be any point inside $\Xi(\mathfrak{V})_p$ such that for all j, we have $L_{m+1} \le \langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_j(p)\rangle \ll L_{m+1}\rho_n^{0+}$. For each $l = 0, \dots, J_p$, we define

$$\begin{aligned}\Xi(\mathfrak{V})_p^l := \{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V}) \text{ and } \\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > \langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_j(p)\rangle,\ j = 1, \dots, l, \quad \text{and} \\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\mu}}_j(p)\rangle > L_{m+1},\ j = l+1, \dots, J_p\}\end{aligned} \tag{11.7}$$

and

$$\check{\Xi}(\mathfrak{V})_p^l := \Xi(\mathfrak{V})_p^l \setminus \Xi(\mathfrak{V})_p^{l+1}. \tag{11.8}$$

Then we obviously have

$$\Xi(\mathfrak{V})_p = \Xi(\mathfrak{V})_p^{J_p} \sqcup (\sqcup_{l=0}^{J_p-1} \check{\Xi}(\mathfrak{V})_p^l) \tag{11.9}$$

(as usual, modulo boundary points). The domain $\Xi(\mathfrak{V})_p^{J_p}$ is a cone, so we already know how to deal with it. Now let us consider $\check{\Xi}(\mathfrak{V})_p^l$ for some l. Let us introduce a coordinate $t = t(\boldsymbol{\xi}) := \langle \boldsymbol{\xi}, \tilde{\boldsymbol{\mu}}_{l+1}(p)\rangle$. Then for $\boldsymbol{\xi} \in \check{\Xi}(\mathfrak{V})_p^l$, we have $t \in [L_{m+1}, \langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_{l+1}(p)\rangle]$. For each $t \in [L_{m+1}, \langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_{l+1}(p)\rangle]$, we denote $O^l(t) := \{\boldsymbol{\xi} \in \check{\Xi}(\mathfrak{V})_p^l,\ \langle \boldsymbol{\xi}, \tilde{\boldsymbol{\mu}}_{l+1}(p)\rangle = t\}$. It is easy to see that the domain $O^l(t)$ is of the same type as the domain $\Xi(\mathfrak{V})_p$; i.e., it can be written as

$$\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \boldsymbol{\xi}_{\mathfrak{V}} \in \Omega(\mathfrak{V}), \langle \boldsymbol{\xi}, \tilde{\boldsymbol{\mu}}_{l+1}(p)\rangle = t \ \text{ and } \ \forall j \ne l+1,\ \langle \boldsymbol{\xi}_{\mathfrak{V}^\perp}, \tilde{\boldsymbol{\nu}}_j\rangle > s_j\}, \tag{11.10}$$

where $\tilde{\boldsymbol{\nu}}_j$ are, of course, normalized projections of some $\tilde{\boldsymbol{\mu}}_j(p)$ onto the plane orthogonal to $\tilde{\boldsymbol{\mu}}_{l+1}(p)$. Our aim is to compute

$$\operatorname{vol} \hat{A}^+ \cap \check{\Xi}(\mathfrak{V})^l - \operatorname{vol} \hat{A}^- \cap \check{\Xi}(\mathfrak{V})^l \tag{11.11}$$

(or at least to prove that this expression admits an asymptotic expansion in powers of ρ and $\ln \rho$). But (11.11) is obviously equal to

$$\int_{L_{m+1}}^{\langle \mathbf{a}, \tilde{\boldsymbol{\mu}}_{l+1}(p)\rangle} \big(\operatorname{vol}(\hat{A}^+ \cap O^l(t)) - \operatorname{vol}(\hat{A}^- \cap O^l(t))\big) dt. \tag{11.12}$$

If $\dim O^l(t) = 1$, or, more generally, if $O^l(t)$ is a simplex, then we can perform integration over $O^l(t)$ as described above, since the formula (11.6) would still be valid, in the sense that

$$\langle \boldsymbol{\xi}, \boldsymbol{\theta}\rangle = C(X, t) + r \sum_l b_l \sin \Phi_l, \tag{11.13}$$

where $L_{m+1}\rho_n^{0-} \ll C(X, t) \ll L_{m+1}\rho_n^{0+}$ and the coordinates $(r, \{\Phi_l\})$ are the shifted polar coordinates in $O^l(t)$. Results of Section 10 show that for each fixed t, the expression $\big(\operatorname{vol}(\hat{A}^+ \cap O^l(t)) - \operatorname{vol}(\hat{A}^- \cap O^l(t))\big)$ admits the asymptotic expansion in powers of ρ and $\ln \rho$, with coefficients being uniformly bounded in t (and X). Now it remains to integrate this expansion against dt (and dX). If $O^l(t)$ is a cone, we cut it onto simplexes as described earlier in this section, and then integrate over each simplex separately. Finally, if $O^l(t)$ is not a cone, we continue the process of reducing dimension until the dimension of $O^l(t)$ becomes equal one.

References

[1] G. Barbatis and L. Parnovski, Bethe-Sommerfeld conjecture for pseudodifferential perturbation, *Comm. Partial Differential Equations* **34** (2009), 383–418. MR 2530702. Zbl 1185.35349. http://dx.doi.org/10.1080/03605300902769006.

[2] B. Helffer and A. Mohamed, Asymptotic of the density of states for the Schrödinger operator with periodic electric potential, *Duke Math. J.* **92** (1998), 1–60. MR 1609321. Zbl 0951.35104. http://dx.doi.org/10.1215/S0012-7094-98-09201-8.

[3] M. Hitrik and I. Polterovich, Regularized traces and Taylor expansions for the heat semigroup, *J. London Math. Soc.* **68** (2003), 402–418. MR 1994690. Zbl 1167.35335. http://dx.doi.org/10.1112/S0024610703004538.

[4] ———, Resolvent expansions and trace regularizations for Schrödinger operators, in *Advances in Differential Equations and Mathematical Physics* (Birmingham, AL, 2002), *Contemp. Math.* **327**, Amer. Math. Soc., Providence, RI, 2003, pp. 161–173. MR 1991539. Zbl 1109.35030.

[5] Y. E. Karpeshina, *Perturbation theory for the Schrödinger operator with a periodic potential*, *Lecture Notes in Math.* **1663**, Springer-Verlag, New York, 1997. MR 1472485. Zbl 0883.35002.

[6] ———, On the density of states for the periodic Schrödinger operator, *Ark. Mat.* **38** (2000), 111–137. MR 1749362. Zbl 1021.35027. http://dx.doi.org/10.1007/BF02384494.

[7] T. Kato, *Perturbation Theory for Linear Operators*, second ed., *Grundlehren Math. Wiss.* **132**, Springer-Verlag, New York, 1976. MR 0407617. Zbl 0342.47009.

[8] E. Korotyaev and A. Pushnitski, On the high-energy asymptotics of the integrated density of states, *Bull. London Math. Soc.* **35** (2003), 770–776. MR 2000023. Zbl 1075.35542. http://dx.doi.org/10.1112/S0024609303002467.

[9] M. A. Naĭmark, *Normed Rings*, P. Noordhoff N. V., Groningen, 1959. MR 0110956. Zbl 0089.10102.

[10] L. Parnovski, Bethe-Sommerfeld conjecture, *Ann. Henri Poincaré* **9** (2008), 457–508. MR 2419769. Zbl 1201.81054. http://dx.doi.org/10.1007/s00023-008-0364-x.

[11] L. Parnovski and R. Shterenberg, Asymptotic expansion of the integrated density of states of a two-dimensional periodic Schrödinger operator, *Invent. Math.* **176** (2009), 275–323. MR 2495765. Zbl 1171.35092. http://dx.doi.org/10.1007/s00222-008-0164-4.

[12] L. Parnovski and A. V. Sobolev, Bethe-Sommerfeld conjecture for periodic operators with strong perturbations, *Invent. Math.* **181** (2010), 467–540. MR 2660451. Zbl 1200.47067. http://dx.doi.org/10.1007/s00222-010-0251-1.

[13] M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV. Analysis of Operators*, Academic Press, New York, 1978. MR 0493422.

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

[15] D. Shenk and M. Shubin, Asymptotic expansion of the state density and the spectral function of a Hill operator, *Math. USSR Sbornik* **56** (1987), 473–490. Zbl 0624.34018. http://dx.doi.org/10.1070/SM1987v056n02ABEH003047.

[16] M. A. Shubin, Almost periodic functions and partial differential operators, *Uspehi Mat. Nauk* **33** (1978), 3–47, translated in *Russian Math. Surveys* **33** (1978), 1–52. MR 0636410. Zbl 0408.47039. http://dx.doi.org/10.1070/RM1978v033n02ABEH002303.

[17] ———, Spectral theory and the index of elliptic operators with almost-periodic coefficients, *Uspekhi Mat. Nauk* **34** (1979), 95–135, translated in *Russian Math. Surveys* **34** (1979), 109–157. MR 0535710. Zbl 0448.47032. http://dx.doi.org/10.1070/RM1979v034n02ABEH002908.

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

[19] A. V. Sobolev, Integrated density of states for the periodic Schrödinger operator in dimension two, *Ann. Henri Poincaré* **6** (2005), 31–84. MR 2119355. Zbl 1065.81051. http://dx.doi.org/10.1007/s00023-005-0198-8.

[20] ———, Asymptotics of the integrated density of states for periodic elliptic pseudo-differential operators in dimension one, *Rev. Mat. Iberoam.* **22** (2006), 55–92. MR 2267313. Zbl 1121.35149. http://dx.doi.org/10.4171/RMI/449.

[21] O. A. Veliev, The spectrum of multidimensional periodic operators, *Teor. Funkts.*, *Funkts. Anal. i Prilozhen.* **49** (1988), 17–34, translated in *J. Math. Sci.* **4** (1990), 1045–1058. MR 0963619. Zbl 0664.47005. http://dx.doi.org/10.1007/BF02216095.

(Received: May 3, 2010)

University College London, London, UK
E-mail: Leonid@math.ucl.ac.uk

University of Alabama at Birmingham, Birmingham AL
E-mail: shterenb@math.uab.edu