Spectral properties of periodic pseudo-differential operators

Irina Pchelintseva

UCL

PhD

Declaration

I, Irina Pchelintseva confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Date: **Signature:**

Abstract

We study elliptic differential and pseudo-differential operators with periodic coefficients. For a wide class of such operators we prove the Bethe-Sommerfeld conjecture, i.e. that the spectrum can have only finitely many gaps. We also study the integrated density of states of periodic Schröedinger operators and prove a lower bound for its variance in the high energy regime. This results in the lower bound for the non-integrated density of states of such operators.

Contents

List of Figures

Chapter 1

Introduction

We will consider periodic elliptic differential and pseudo-differential operators acting in $\mathbb{R}^d$. By periodicity of an operator $H = \mathrm{Op}(h)$ we understand that there exists full ranked lattice Γ such that $h(\mathbf{x}+\boldsymbol{\gamma}, \boldsymbol{\xi}) = h(\mathbf{x}, \boldsymbol{\xi})$ for every $\boldsymbol{\gamma} \in \mathsf{\Gamma}$.

Let $\mathsf{\Gamma} \subset \mathbb{R}^d$ ($d \geq 1$) be a lattice. Denote by $\mathcal{O}$ its fundamental domain: $\mathcal{O} = \mathbb{R}^d/\mathsf{\Gamma}$. For example, for $\mathcal{O}$ one can choose a parallelepiped spanned by a basis of $\mathsf{\Gamma}$.

$\mathsf{\Gamma}^\dagger$ is a lattice dual to $\mathsf{\Gamma}$:

$$\mathsf{\Gamma}^\dagger = \{\boldsymbol{\eta} : \langle \boldsymbol{\eta}, \boldsymbol{\gamma} \rangle \in 2\pi\mathbb{Z}, \forall \boldsymbol{\gamma} \in \mathsf{\Gamma}\}$$

and its fundamental domain is denoted by $\mathcal{O}^\dagger$.

Under very broad conditions, spectra of elliptic differential operators with periodic coefficients in $\mathsf{L}^2(\mathbb{R}^d)$, have a band structure. In order to describe this band structure it is convenient to introduce the Floquet-Bloch decomposition (see, e.g., [28]) and thus express our operator H as a direct integral

$$H = \int_{\mathcal{O}^\dagger} H(\mathbf{k}) d\mathbf{k}.$$

Here $H(\mathbf{k})$ is a pseudo-differential operator, acting on $\mathcal{O}^\dagger$ with a symbol $h(\mathbf{x}, \boldsymbol{\xi}+\mathbf{k})$, where $h(\mathbf{x}, \boldsymbol{\xi})$ is the symbol of the original operator H. The parameter $\mathbf{k}$ is usually called the quasi-momentum.

All the operators $H(\mathbf{k})$ have compact resolvents, therefore their spectra are purely discrete, and by applying a simple perturbation theory argument, one can see that their eigenvalues $\lambda_j(\mathbf{k})$ are continuous functions of the quasi-momentum $\mathbf{k}$. Here by $\lambda_j(\mathbf{k})$ we denote the j-th eigenvalue of the operator $H(\mathbf{k})$ counting the multiplicities and arranged in the increasing order . Now the spectrum of the initial operator H can be represented as the union of the closed intervals $\ell_j = \bigcup_{\mathbf{k}} \lambda_j(\mathbf{k})$ which are called *spectral bands*:

$$\sigma(H) = \bigcup_j \ell_j,$$

possibly separated by spectrum-free intervals (*gaps*) (see [28] and [18]).

The natural first question to ask in this setting is how does the spectrum look like a set.

Since the 30's it has been a general belief among physicists that the number of gaps in the spectrum of the Schrödinger operator $H_V = -\Delta + V$ with a periodic electric potential V in dimension three must be finite. After the classical monograph [2] this belief is known as the Bethe-Sommerfeld conjecture. It is relatively straightforward to see that this conjecture holds for potentials which admit a separation of variables, as shown in [5], p.121. But for general potentials this conjecture turned out to be quite difficult, thus the first rigorous results appeared only in the beginning of the 80's. Though the original conjecture was stated only for three-dimensional space, we will say that 'Bethe-Sommerfeld conjecture holds' for an elliptic pseudo-differential periodic operator in general dimension if the number of gaps in its spectrum is finite.

In the case of the Schrödinger operator H_V it is known that the number of gaps is generically infinite if $d = 1$ (see [28]). For $d \geq 2$ there has been a large number of publications proving the conjecture for H_V under various conditions on the potential and the periodicity lattice.

Define the counting function:

$$N(\lambda, H(\mathbf{k})) = \#\{j : \lambda_j(\mathbf{k}) \leq \lambda\}$$

Due to the mentioned boundedness of the counting function $N(\lambda, H(\mathbf{k}))$, each interval $(-\infty, \lambda]$ has non-empty intersection with finitely many spectral bands. We define two quantitative characteristics of overlapping of the bands:

1. The *multiplicity of overlapping*, which measures the number of bands covering given point λ:

$$\mathfrak{m}(\lambda) = \#\{j : \lambda \in \ell_j\}$$

2. The *overlapping function*, which shows how far the bands penetrate into each other. This function $\zeta(\lambda)$, $\lambda \in \mathbb{R}$ is defined as the maximal number t such that the symmetric interval $[\lambda - t, \lambda + t]$ is entirely contained in one band, i.e.

$$\zeta(\lambda) = \zeta(\lambda; H) = \begin{cases} \max_j \max\{t : [\lambda - t, \lambda + t] \subset \ell_j\}, \ \lambda \in \sigma(H); \\ 0, \ \lambda \notin \sigma(H). \end{cases} \tag{1.0.1}$$

Both these functions were first introduced by M. Skriganov (see [34]). The quantities $\mathbf{m}(\lambda)$ and $\zeta(\lambda)$ can be linked with the counting function $N(\lambda)$ of the operator $H(\mathbf{k})$:

$$\begin{cases} \mathfrak{m}(\lambda) = \max_{\mathbf{k}} N(\lambda, H(\mathbf{k})) - \min_{\mathbf{k}} N(\lambda, H(\mathbf{k})), \\ \zeta(\lambda) = \sup\{t : \min_{\mathbf{k}} N(\lambda + t, H(\mathbf{k})) < \max_{\mathbf{k}} N(\lambda - t, H(\mathbf{k}))\} \end{cases} \tag{1.0.2}$$

Here we assume that λ is not an end point of any ℓ_j.

3. The *integrated density of states (IDS)* for H is defined as

$$\mathbf{N}(\lambda) := \lim_{L\to\infty} L^{-d} N(\lambda; H_D^{(L)}), \quad \lambda \in \mathbb{R}. \tag{1.0.3}$$

Here, $H_D^{(L)}$ is the restriction of H to the cube $[0, L]^d$ with the Dirichlet boundary conditions, and $N(\lambda; \cdot)$ is the counting function of the discrete spectrum below λ. The existence of the limit in (1.0.3) is well known, see e.g. [28], [32]. Moreover there is a nice explicit formula

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

Also we can formulate in terms of the IDS that a point belongs to a gap: λ belongs to a gap if and only if the IDS is a constant in a neighborhood of λ: there exists a number $\varepsilon > 0$ such that for any $\lambda_1 \in (\lambda-\varepsilon, \lambda+\varepsilon)$ it we have $\mathbf{N}(\lambda_1) = \mathbf{N}(\lambda)$. Similarly, to say that λ belong so the band is equivalent to saying that $\mathfrak{m}(\lambda) \geq 1$ or $\zeta(\lambda) > 0$.

Now we briefly discuss a history of the subject until now. The first rigorous results for the Bethe-Sommerfeld conjecture for the Schrödinger operator relied on number-theoretic ideas, and they appeared in [27], [4] ($d = 2$) and [33], [34], [35] ($d \geq 2$). At that time it was found that the complexity of the problem increases with the dimension: the validity of the conjecture for dimensions $d \geq 4$ was established by M. Skriganov only for rational lattices, see [34]. Later, the conjecture for arbitrary lattices was extended to $d = 4$ in the work of B. Helffer and A. Mohamed [8]. The definitive result was obtained in the paper of L. Parnovski [21] where the Bethe-Sommerfeld conjecture was proved for the Schrödinger operator for any periodicity lattice in all dimensions $d \geq 2$, with an arbitrary smooth potential V. We observe that the complexity of the problem increases dramatically when instead of the bounded potential perturbation one introduces in the Schrödinger operator a periodic magnetic potential $\mathbf{a} = \mathbf{a}(x) = (a_1, a_2, \dots, a_d)$, $H = (-i\nabla - \mathbf{a})^2 + V$. Bethe-Sommerfeld conjecture for this operator was proven in the case of $d = 2$, see [19], [14]. The Bethe-Sommerfeld conjecture for the polyharmonic operator in $L^2(\mathbb{R}^d)$

$$H = H_0 + V, \;\; H_0 = (-\Delta)^m, \quad m > 0$$

with a periodic real-valued function V, was also studied by various authors. The first result is due to M. Skriganov (see [34], [35]), who showed that the number of gaps is finite if $2m > d$; $d \geq 3$. Then the polyharmonic operator was studied by Yu. Karpeshina in [10] (see also [12] and references therein) in the framework of the analytic perturbation theory. The high energy asymptotics of the Bloch eigenvalues found in [8] implied the Bethe-Sommerfeld conjecture for $4m > d + 1$; $d \geq 2$.

Later, L. Parnovski and A. Sobolev [22], [23] extended the result for potential perturbations to the case $8m > d+3$, $d \geq 2$. For the Schrödinger case $m = 1$, the latter condition is equivalent to the requirement that $d = 2, 3$ or 4. These are exactly the dimensions for which the conjecture was justified in [8]. In the paper [1] G. Barbatis and L. Parnovski considered the case of a pseudo-differential operator of the form:

$$H = (-\Delta)^m + B, \quad m > 0, \tag{1.0.4}$$

with a pseudo-differential perturbation B of order $a < 2m - 1$ and arbitrary $d \geq 2$ and proved Bethe-Sommerfeld conjecture. Note that this class does not cover the case of magnetic Schrödinger operator. Finally, the paper [24] extended the results of [1] to the case of arbitrary perturbations of order smaller than $2m$. Thus, in particular, they have settled the Bethe-Sommerfeld conjecture for the magnetic Schrödinger operator.

Note that all these operators considered so far are perturbations of the Schrödinger operator $(-\Delta)^m$. One of the aims of this thesis is to extend these results to a slightly wider class of pseudo-differential operators.

Now we want to discuss different methods used by researchers in order to tackle this problem.

Due to its physical relevance, the case of the Schrödinger operator, has been studied better than the general one.

There are three main methods in dimensions $d \geq 2$ of dealing with this type of problems in literature, which lead in one way or another to the justification of the conjecture.

The first one is due to M. Skriganov, and it presents a combination of number-theoretic ideas and analytics tools. The main object of studies in this context is the dependence of the counting function $N(\lambda, H(\mathbf{k}))$ on the quasi-momentum. For the first time this approach was used in [27] by V. Popov and M. Skriganov to give the first rigorous proof of the conjecture for the case $d = 2$. Then M. Skriganov (see [34] and references therein) obtained a proof for all dimensions $d \geq 3$ for rational lattices Γ. For $d = 3$ the result was extended to arbitrary Γ in [15]. A slightly simpler proof in the case $d = 2$ was given by B.E.J. Dahlberg and E. Trubowitz in [4]. The case of rational lattices was revisited in paper [36] by M. Skriganov and A. Sobolev. In [36] the original Skriganov proof [34] was simplified by separating the number-theoretic part from the spectral one.

The other two methods are indirect in the sense that the finiteness of the number of gaps in both of them is inferred from other spectral properties of the operator in question, the study of which presents a separate difficult problem. The first of these indirect approaches was described by Yu. Karpeshina in [12] and based on the high energy asymptotics of the Bloch eigenvalues and corresponding eigenfunctions of the operator $H(\mathbf{k})$. It was first applied by O. Veliev (see also [41]) to prove the validity of the Bethe-Sommerfeld conjecture for $d = 3$. Another proof can be found in the book [12] by Yu. Karpeshina. We point out that in [12] the conjecture was also proved for a wide class of singular potentials, including Coulomb potentials.

The third method was first introduced by L. Parnovski in [21][1]. He has shown that, by obtaining a very precise asymptotics of the eigenvalues and using certain arguments from the geometrical combinatorics it is possible to prove the Bethe-Sommerfeld conjecture without requesting any information on the eigenfunctions. Afterwards in paper [24] this method was improved by combining it with the gauge transformation method.

The fruitful technique of gauge transformation was first introduced by A. Sobolev in [39][2] and [38] and was also used in [24]. This method consists of constructing two pseudo-differential operators, H_1 and H_2. Here, $H_1 = e^{i\Psi} H e^{-i\Psi}$, where Ψ is 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 $H_2(\mathbf{k})$ are "almost diagonal" and, in particular, have a lot of invariant subspaces. These invariant subspaces can be generated by two different types of eigenvalues of $H(\mathbf{k})$. Eigenvalues of the first type are called stable or non-resonant eigenvalues and in order to study them one employes the perturba-

[1]For 2008, L. Parnovski receives the Annales Henri Poincaré Prize for this paper

[2]For 2005, A Sobolev receives the Annales Henri Poincaré Prize for this paper

tion theory of simple eigenvalues. The corresponding invariant subspaces of $H_2(\mathbf{k})$ are one-dimensional. The second type, called unstable or resonant eigenvalues, is much more complicated so that the perturbation theory for multiple eigenvalues is required in order to describe their behaviour, and thus they are not so well controlled. The corresponding invariant subspaces of $H_2(\mathbf{k})$ are generated by clusters of unstable eigenvalues.

The approach suggested in the papers [39] and [38] can be viewed as a variant of the method of "near-similarity" put forward by G. Rosenblum in [29] for the PDO's on the unit circle. Rosenblum's idea was to construct for a given elliptic PDO A a suitable Fourier Integral Operator S such that the operator $S^{-1}AS$ up to some negligible terms coincides with a PDO with constant coefficients.

Then this method was further developed in the the joined work [24] of L. Parnovski and A. Sobolev to prove Bethe-Sommerfeld conjecture for the polyharmonic operator.

Let us give some notational conventions.

We use notation $f \ll g$ or $g \gg f$ for two positive functions f, g, if there is a constant $C > 0$ independent of f, g such that $f \le Cg$. If $f \ll g$ and $g \ll f$, then we write $f \asymp g$.

Now we discuss the results related to the studies of the asymptotic behaviour of functions describing the spectrum.

As we have seen, the condition that λ belongs to the gap can be reformulated in terms of either of 3 functions: N, $\mathfrak{m}$, ζ. Thus, together with proving Bethe-Sommerfeld conjecture, one can try to obtain additional information about the behaviour of this functions. In particular following facts are known: in dimensions $d = 2, 3, 4$ it have been proved in [34], [35], [4], and [23] that for large λ we have

$$\mathfrak{m}(\lambda) \gg \lambda^{\frac{d-1}{4}}$$

$$\zeta(\lambda) \gg \lambda^{\frac{3-d}{4}}$$

In [21] it is shown that in arbitrary dimensions

$$\zeta(\lambda) \gg \lambda^{\frac{1-d}{2}}$$

Notice that when we increase the dimension, estimates for the multiplicity of overlapping improve at least when $d \leq 4$, whereas the estimates for the overlapping function are becoming worse. This explains why it is becoming more difficult to prove the conjecture as the dimension increases.

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

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

where

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

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

$$N(\lambda) \sim \lambda^{d/2}\left(C_d + \sum_{j=1}^{\infty} e_j \lambda^{-j}\right), \tag{1.0.5}$$

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

$$N(\lambda) = \lambda^{d/2}\left(C_d + \sum_{j=1}^{K} e_j \lambda^{-j}\right) + R_K(\lambda) \tag{1.0.6}$$

with $R_K(\lambda) = o(\lambda^{\frac{d}{2}-K})$. In those formulas, e_j are real numbers which depend on the potential b. They can be calculated relatively easily using the heat kernel invariants (computed in [9]); they are equal to certain integrals of the potential b and its derivatives.

Indeed, in the paper [17], all these coefficients were computed; in particular, it turns out that, if d is even, then e_j vanish whenever $j > d/2$. Formula (1.0.5) was proved in the case $d = 1$ in [31]. There were important results due to A. Sobolev, who obtained the complete asymptotic expansion of $N(\lambda)$ in the case of $d = 1$ and pseudo-differential perturbation in [38]. Later, using similar methods, he obtained

3 asymptotical terms of $N(\lambda)$ (i.e. $K = 2$ in (1.0.6)) for the two-dimensional case in [39]. In paper by L. Parnovski and R. Shterenberg [25], the complete asymptotics (1.0.5) was obtained in the case $d = 2$.

In the paper of B. Helffer and A. Mohamed [8] the authors, using microlocal machinery, derived a suitable two-term asymptotic formula for the integrated density of states of the operator H at large energies (i.e. $K = 1$). This was later used to show the validity of the Bethe-Sommerfeld conjecture for $d = 2; 3; 4$.

By Yu. Karpeshina in [13] it was proved that formula (1.0.6) is valid with $K = 1$ and $R(\lambda) = O(\lambda^{-\delta})$ with some small positive δ when $d = 3$ and

$$R(\lambda) = O(\lambda^{-\delta})$$

when $d > 3$.

Finally, the complete asymptotic expansion of the integrated density of states of a Schrödinger operator is proved in the work of L. Parnovski and R. Shterenberg [26].

These were the results in the asymptotical behaviour of the IDS.One can also study the size of the local variation of the IDS. For example, in [21] it was shown that for each $n \in \mathbb{N}$ and $\varepsilon = \lambda^{-n}$ we have

$$\mathbf{N}(\lambda + \varepsilon) - \mathbf{N}(\lambda) \ll \varepsilon \lambda^{(d-2)/2}. \tag{1.0.7}$$

The second objective of this thesis is to obtain a lower bound similar to (1.0.7).

This naturally leads us to the discussion of the results of the thesis.

As we have already noted, all the mentioned results concern the operators with the principal symbol $|\boldsymbol{\xi}|^{2m}$. Naturally, one would think of generalizing to a wider class.

Consider a pseudo-differential operator H with a homogeneous principal symbol $h_0(\boldsymbol{\xi}) : h_0(t\boldsymbol{\xi}) = t^m h_0(\boldsymbol{\xi})$. Consider its level set $M(\boldsymbol{\xi}^*) = \{\boldsymbol{\xi} : h_0(\boldsymbol{\xi}) = h_0(\boldsymbol{\xi}^*)\}$. If this level set is the surface of a cube, then due to result by M. Skriganov in [34] we know, that one can add a perturbation of arbitrarily small order so that the resulting perturbed operator has infinite number of gaps. This phenomenon occurs because for the cube there are points where spectral bands do not intersect but only tough each other, that is to say that there are points where $\zeta = 0$. In view of this, we want to assume the strict convexity for the level set, that is we require all principal curvatures to be positive.

In chapter 2 we extend the results of L. Parnovski and A. Sobolev to this class of operators. When doing this, we apply method similar to those of [24]. Thus, in some steps we will repeat the calculations from the work [24] with minor adjustments, in order to make this thesis self-contained.

On the other hand there are some essential differences that arise if the principal symbol is not $|\boldsymbol{\xi}|^{2m}$. The most tricky part is how to redefine the resonance sets. In particular, due to this difficulty we have to define two different classes of resonance sets ("narrow" and "wide" ones) and play around with their properties. The definitive result on this matter is presented in Chapter 2, see theorems 2.1.1, 2.1.2.

The next question we want to consider is the lower bound for the non-integrated density of states, that is the rate of increase of the IDS. As we have already seen for Schrödinger operators, there is an upper bound, which assumes that the variation is calculated between the points located not too close to each other (i.e. ε can decay like the power of λ but not faster).

It turned out that similar lower bound can be proved for arbitrarily small values of ε. When ε tends to zero we can prove that non-integrated density of states, thought of as a measure, is bounded below.

These results were published in [20], and in chapter 3 we present them similar to the published version. Therefore several definitions of the resonance sets are different in chapter 2 and 3. Also chapter 3 is less self-contained than chapter 2: in chapter 3 we use several results from [21] without proof.

Chapter 2

Bethe-Sommerfeld conjecture for pseudo-differential operators

The exposition of this chapter follows that of [24], in particular sections 2.4 (Properties of periodic PDO's) 2.5 (A "gauge transformation"), 2.8 (Estimates of volumes) follow with minor changes corresponding sections of [24]. The other sections however contain the material, which is completely new.

2.1 Periodic pseudo-differential operators. Main result

2.1.1 Classes of PDO's

Before we define the pseudo-differential operators (PDO's), we introduce the relevant classes of symbols. Let $\Gamma \in \mathbb{R}^d$ be a lattice. Denote by $\mathcal{O}$ its fundamental domain. For example, for $\mathcal{O}$ one can choose a parallelepiped spanned by a basis of Γ. The dual lattice and its fundamental domain are denoted by $\Gamma^\dagger$ and $\mathcal{O}^\dagger$ respectively. Sometimes we reflect the dependence on the lattice and write $\mathcal{O}_\Gamma$ and $\mathcal{O}^\dagger_\Gamma$. In particular, in the case $\Gamma = (2\pi\mathbb{Z})^d$ one has $\Gamma^\dagger = \mathbb{Z}^d$ and it is natural to take $\mathcal{O} = [0, 2\pi)^d$, $\mathcal{O}^\dagger = [0, 1)^d$. For any measurable set $\mathcal{C} \subset \mathbb{R}^d$ we denote by $|\mathcal{C}|$ or

$\mathrm{vol}(\mathcal{C})$ its Lebesgue measure (volume). The volume of the fundamental domain does not depend on its choice, it is called the *determinant of the lattice* Γ and denoted $\mathrm{d}(\Gamma) = |\mathcal{O}|$. By $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d$ we denote the standard orthonormal basis in $\mathbb{R}^d$.

Let $\rho \in \mathbb{R}$ be some value, that will be later chosen large enough.

Now we introduce a notation: let $A_{max} > 1$ be a number such that

$$\frac{1}{A_{max}} < h_0(\boldsymbol{\xi}) < A_{max}, \quad \forall \boldsymbol{\xi} : |\boldsymbol{\xi}| = 1.$$

Clearly, A_{max} exists, due to the compactness of the unit sphere. Denote

$$S(\rho) = \{\boldsymbol{\xi} : \frac{\rho^{2m}}{A_{max}} < h_0(\boldsymbol{\xi}) < A_{max}\rho^{2m}\}. \tag{2.1.1}$$

For $r \in \mathbb{R}$ denote by Θ_r the following set of vectors:

$$\Theta_r = \{\boldsymbol{\theta} \in \Gamma : |\boldsymbol{\theta}| < r\}.$$

Here

$$r < \rho^{\varkappa}, \tag{2.1.2}$$

where $\varkappa < 1$ is some fixed number independent of ρ. The value of $\varkappa$ will be chosen later.

For any $u \in \mathsf{L}^2(\mathcal{O})$ and $f \in \mathsf{L}^2(\mathbb{R}^d)$ define the Fourier coefficients and Fourier transform respectively:

$$\hat{u}(\boldsymbol{\theta}) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \int_{\mathcal{O}} e^{-i\langle \boldsymbol{\theta}, \mathbf{x}\rangle} u(\mathbf{x}) d\mathbf{x},\ \boldsymbol{\theta} \in \Gamma^{\dagger},\ (\mathcal{F}f)(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\langle \boldsymbol{\xi}, \mathbf{x}\rangle} f(\mathbf{x}) d\mathbf{x},\ \boldsymbol{\xi} \in \mathbb{R}^d.$$

Let us now define the periodic symbols and PDO's associated with them. Let $b = b(\mathbf{x}, \boldsymbol{\xi})$, $\mathbf{x}, \boldsymbol{\xi} \in \mathbb{R}^d$, be a Γ-periodic complex-valued function, i.e.

$$b(\mathbf{x} + \boldsymbol{\gamma}, \boldsymbol{\xi}) = b(\mathbf{x}, \boldsymbol{\xi}),\ \forall \boldsymbol{\gamma} \in \Gamma.$$

Let $w : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function such that $w(\boldsymbol{\xi}) \geq 1\ \forall \boldsymbol{\xi} \in \mathbb{R}^d$ and

$$w(\boldsymbol{\xi} + \boldsymbol{\eta}) \leq C w(\boldsymbol{\xi}) \langle \boldsymbol{\eta} \rangle^{\kappa},\ \forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^d, \tag{2.1.3}$$

for some $\kappa \ge 0$. Here we have used the standard notation $\langle \mathbf{t} \rangle = \sqrt{1 + |\mathbf{t}|^2}$, $\forall \mathbf{t} \in \mathbb{R}^d$. We say that the symbol b belongs to the class $\mathbf{S}_\gamma = \mathbf{S}_\gamma(w) = \mathbf{S}_\gamma(w, \Gamma)$, $\gamma \in \mathbb{R}$, if for any $l \ge 0$ and any non-negative $s \in \mathbb{Z}$ the condition

$$\lvert b \rvert^{(\gamma)}_{l,s} := \max_{|\mathbf{s}| \le s} \sup_{\boldsymbol{\xi}, \boldsymbol{\theta}} \langle \boldsymbol{\theta} \rangle^l \; w(\boldsymbol{\xi})^{-\gamma + |\mathbf{s}|} |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})| < \infty, \quad |\mathbf{s}| = s_1 + s_2 + \cdots + s_d, \tag{2.1.4}$$

is fulfilled. Here, of course, $\hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})$ is the Fourier coefficient of the symbol $b(\cdot, \boldsymbol{\xi})$ with respect to the first variable. The quantities (2.1.4) define norms on the class $\mathbf{S}_\gamma$. In the situations when it is not important for us to know the exact values of l, s, we denote the above norm by $\lvert b \rvert^{(\gamma)}$. In this case the inequality $A \le C \lvert b \rvert^{(\gamma)}$ means that there exist values of l and s, and a constant $C > 0$, possibly depending on l, s, such that $A \le C \lvert b \rvert^{(\gamma)}_{l,s}$. Similarly, when we write $\lvert b \rvert^{(\gamma_1)} \le C \lvert g \rvert^{(\gamma_2)}$ for some symbols $b \in \mathbf{S}_{\gamma_1}, g \in \mathbf{S}_{\gamma_2}$, we mean that for any l and s the norm $\lvert b \rvert^{(\gamma_1)}_{l,s}$ is bounded by $\lvert g \rvert^{(\gamma_2)}_{p,n}$ with some p and n depending on l, s, and some constant $C = C_{l,s}$. Here we of course assume that b belongs to some set of symbols ang g is dependent on b; an example of using this convention can be seen in formulas (2.5.20)-(2.5.23). In general, by C, c (with or without indices) we denote various positive constants, whose precise value is unimportant. Throughout this chapter we adopt the following convention. An estimate (or an assertion) is said to be uniform in a symbol $b \in \mathbf{S}_\gamma$ if the constants in the estimate (or assertion) at hand depend only on the constants $C_{l,s}$ in the bounds $\lvert b \rvert^{(\gamma)}_{l,s} \le C_{l,s}$. This is sometimes expressed by saying that an estimate (or assertion) is uniform in the symbol b satisfying $\lvert b \rvert^{(\gamma)} \le C$.

We use the classes $\mathbf{S}_\gamma$ mainly with the weight $w(\boldsymbol{\xi}) = \langle \boldsymbol{\xi} \rangle^\beta$, $\beta \in (0, 1]$, which satisfies (2.1.3) for $\kappa = \beta$. Note that $\mathbf{S}_\gamma$ is an increasing function of γ, i.e. $\mathbf{S}_{\gamma_2} \subset \mathbf{S}_{\gamma_1}$ for $\gamma_2 < \gamma_1$. For later reference we write here the following convenient bounds that follow from definition (2.1.4) and property (2.1.3):

$$|\mathbf{D}^{\mathbf{s}}_{\xi} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})| \le \lvert b \rvert^{(\gamma)}_{l,s} \langle \boldsymbol{\theta} \rangle^{-l} w(\boldsymbol{\xi})^{\gamma - s}, \tag{2.1.5}$$

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\eta}) - \mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})| \le C \lvert b \rvert^{(\gamma)}_{l,s+1} \langle \boldsymbol{\theta} \rangle^{-l} w(\boldsymbol{\xi})^{\gamma - s - 1} \langle \boldsymbol{\eta} \rangle^{\kappa |\gamma - s - 1|} |\boldsymbol{\eta}|, \;\; s = |\mathbf{s}|, \tag{2.1.6}$$

where a constant $C = \max\{1, c^{(\gamma - s - 1)}\}$, where c is the constant from (2.1.2). For a vector $\boldsymbol{\eta} \in \mathbb{R}^d$ introduce the symbol

$$b_{\boldsymbol{\eta}}(\mathbf{x}, \boldsymbol{\xi}) = b(\mathbf{x}, \boldsymbol{\xi} + \boldsymbol{\eta}), \boldsymbol{\eta} \in \mathbb{R}^d, \tag{2.1.7}$$

so that $\hat{b}_{\boldsymbol{\eta}}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\eta})$. The bound (2.1.6) implies that for all $|\boldsymbol{\eta}| \leq C$ we have

$$|b - b_{\boldsymbol{\eta}}|^{(\gamma-1)}_{l,s} \leq C_s |b|^{(\gamma)}_{l,s+1} |\boldsymbol{\eta}|, \tag{2.1.8}$$

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

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

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

the integrals being over $\mathbb{R}^d$. Under the condition $b \in \mathbf{S}_\gamma$ the integral in the r.h.s. 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 $\mathsf{L}^2(\mathbb{R}^d)$, see Proposition 2.4.1. Unless otherwise stated, from now on $\mathsf{S}(\mathbb{R}^d)$ is taken as a natural domain for all PDO's at hand. Observe that the operator $\mathrm{Op}(b)$ is symmetric if its symbol satisfies the condition

$$\hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \overline{\hat{b}(-\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\theta})}. \tag{2.1.9}$$

We shall call such symbols *symmetric*.

Our aim is to study the spectrum of the operator

$$\begin{cases} H = \mathrm{Op}(h), \ h(\mathbf{x}, \boldsymbol{\xi}) = h_0(\boldsymbol{\xi}) + b(\mathbf{x}, \boldsymbol{\xi}), \\ h_0(\boldsymbol{\xi}) = a(\boldsymbol{\xi}') |\boldsymbol{\xi}|^{2m}, \ m > 0, \\ b \in \mathbf{S}_\gamma(\langle \boldsymbol{\xi} \rangle^\beta), \ \gamma\beta < 2m, \end{cases} \tag{2.1.10}$$

with a symmetric symbol b. Here $\boldsymbol{\xi}' = \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}$, a - is a positive smooth (we require the continuity of the first derivative) function, $m \in \mathbb{R}$.

Also we put one more restriction on the function a: the set

$$\{\boldsymbol{\xi} \in \mathbb{R}^d : h_0(\boldsymbol{\xi}) \leq 1\}$$

is strictly convex.

The operator $\mathrm{Op}(b)$ is infinitesimally H_0-bounded, see Lemma 2.4.2, so that H is self-adjoint on the domain $D(H) = D(H_0) = \mathsf{H}^{2m}(\mathbb{R}^d)$. Due to the Γ-periodicity of the symbol b, the operator H commutes with the shifts along the lattice vectors, i.e.

$$H\mathcal{T}_{\boldsymbol{\gamma}} = \mathcal{T}_{\boldsymbol{\gamma}}H,\ \boldsymbol{\gamma} \in \Gamma.$$

with $(\mathcal{T}_{\boldsymbol{\gamma}}u)(\mathbf{x}) = u(\mathbf{x} + \boldsymbol{\gamma})$. This allows us to use the *Floquet decomposition*.

For an arbitrary point $\boldsymbol{\xi}^* \in \mathbb{R}^d \setminus \{0\}$ we consider a level set

$$M(\boldsymbol{\xi}^*) = \{\boldsymbol{\xi} : h_0(\boldsymbol{\xi}) = h_0(\boldsymbol{\xi}^*)\}.$$

Since the function h_0 is smooth and grows in any direction from the origin (i.e. increases when the argument is multiplied by a constant greater than 1) and is zero at the origin, then any level set is a smooth convex surface, homeomorphic to a sphere. Also it is clear that $\boldsymbol{\xi}^* \in M(\boldsymbol{\xi}^*)$.

Now define a function $\psi : \mathbb{R}^d \setminus \{0\} \to S^{d-1}$ at $\boldsymbol{\xi}^*$ as the unit outer normal vector to the set $M(\boldsymbol{\xi}^*)$ at the point $\boldsymbol{\xi}^*$.

2.1.2 Floquet decomposition

We identify the underlying Hilbert space $\mathcal{H} = \mathsf{L}^2(\mathbb{R}^d)$ with the direct integral

$$\mathfrak{G} = \int_{\mathcal{O}^\dagger} \mathfrak{H} d\mathbf{k},\ \mathfrak{H} = \mathsf{L}^2(\mathcal{O}).$$

This identification is implemented by the Gelfand transform

$$(Uu)(\mathbf{x}, \mathbf{k}) = \frac{1}{\sqrt{\mathrm{d}(\Gamma^\dagger)}} e^{-i\langle \mathbf{k}, \mathbf{x}\rangle} \sum_{\boldsymbol{\gamma} \in \Gamma} e^{-i\langle \mathbf{k}, \boldsymbol{\gamma}\rangle} u(\mathbf{x} + \boldsymbol{\gamma}),\ \mathbf{k} \in \mathbb{R}^d, \tag{2.1.11}$$

which is initially defined on $u \in \mathsf{S}(\mathbb{R}^d)$ and extends by continuity to a unitary mapping from $\mathcal{H}$ onto $\mathfrak{G}$. In terms of the Fourier transform the Gelfand transform is defined as follows: $\widehat{(Uu)}(\boldsymbol{\theta}, \mathbf{k}) = (\mathcal{F}u)(\boldsymbol{\theta} + \mathbf{k}),\ \boldsymbol{\theta} \in \Gamma^\dagger$. The unitary operator U reduces $\mathcal{T}_{\boldsymbol{\gamma}}$ to the diagonal form:

$$(U\mathcal{T}_{\boldsymbol{\gamma}}U^{-1}f)(\ \cdot\ , \mathbf{k}) = e^{i\mathbf{k}\cdot\boldsymbol{\gamma}} f(\ \cdot\ , \mathbf{k}),\ \forall \boldsymbol{\gamma} \in \Gamma.$$

Let us consider a self-adjoint operator A in $\mathcal{H}$ which commutes with $\mathcal{T}_{\boldsymbol{\gamma}}$ for all $\boldsymbol{\gamma} \in \Gamma$, i.e. $A\mathcal{T}_{\boldsymbol{\gamma}} = \mathcal{T}_{\boldsymbol{\gamma}}A$. We call such operators (Γ-)periodic. Then A is partially diagonalised by U (see [28]), that is, there exists a measurable family of self-adjoint operators (fibres) $A(\mathbf{k}), \mathbf{k} \in \mathcal{O}^\dagger$ acting in $\mathfrak{H}$, such that

$$UAU^* = \int_{\mathcal{O}^\dagger} A(\mathbf{k})d\mathbf{k}. \qquad (2.1.12)$$

It is easy to show that any periodic operator T, which is A-bounded with relative bound $\epsilon < 1$, can be also decomposed into a measurable set of fibers $T(\mathbf{k})$ in the sense that

$$(UTf)(\,\cdot\,, \mathbf{k}) = T(\mathbf{k})(Uf)(\,\cdot\,, \mathbf{k}), \text{ a.e. } \mathbf{k} \in \mathcal{O}^\dagger,$$

for all $f \in D(A)$. Moreover, the fibers $T(\mathbf{k})$ are $A(\mathbf{k})$-bounded with the bound ϵ, and if T is symmetric, then the operator $A(\mathbf{k}) + T(\mathbf{k})$ is self-adjoint on $D(A(\mathbf{k}))$.

Suppose that the operator A (and hence $A(\mathbf{k})$) is bounded from below and that the spectrum of each $A(\mathbf{k})$ is discrete. Denote by $\lambda_j\big(A(\mathbf{k})\big), j = 1, 2, \dots,$ the eigenvalues of $A(\mathbf{k})$ labeled in the ascending order. Define the counting function in the usual way:

$$N\big(\lambda, A(\mathbf{k})\big) = \#\{j : \lambda_j\big(A(\mathbf{k})\big) \le \lambda\}, \ \lambda \in \mathbb{R}.$$

If $A = \mathrm{Op}(a)$ with a real-valued symbol $a \in \mathsf{L}^\infty_{\mathrm{loc}}(\mathbb{R}^d)$ depending only on $\boldsymbol{\xi}$, then $A(\mathbf{k})$ is a self-adjoint PDO in $\mathfrak{H}$ defined as follows:

$$A(\mathbf{k})u(\mathbf{x}) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \sum_{\mathbf{m}\in\Gamma^\dagger} e^{i\mathbf{m}\cdot\mathbf{x}} a(\mathbf{m}+\mathbf{k})\hat{u}(\mathbf{m}).$$

If $a(\boldsymbol{\xi}) \to \infty$ as $|\boldsymbol{\xi}| \to \infty$, then the spectrum of each $A(\mathbf{k})$ is purely discrete with eigenvalues given by $\lambda^{(\mathbf{m})}(\mathbf{k}) = a(\mathbf{m}+\mathbf{k}), \mathbf{m} \in \Gamma^\dagger$. Consequently, the number of eigenvalues below each $\lambda \in \mathbb{R}$ is bounded from above uniformly in $\mathbf{k} \in \mathcal{O}^\dagger$. If T is a periodic symmetric operator which is A-bounded with a bound $\epsilon < 1$, then the spectrum of $A(\mathbf{k}) + T(\mathbf{k})$ is also purely discrete and the counting function is also bounded uniformly in $\mathbf{k}$. In particular, the above applies to the elliptic operator H defined in (2.1.10). In fact, applying the Gelfand transform (2.1.11) to $\mathrm{Op}(b)$, one finds that, similarly to A considered above, the operator $H(\mathbf{k})$ is a PDO in $\mathfrak{H}$ of the form

$$H(\mathbf{k})u(\mathbf{x}) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \sum_{\mathbf{m}\in\Gamma^\dagger} e^{i\mathbf{m}\cdot\mathbf{x}} h(\mathbf{x}, \mathbf{m}+\mathbf{k})\hat{u}(\mathbf{m}),\ \mathbf{k} \in \mathbb{R}^d. \tag{2.1.13}$$

The values $H(\mathbf{k})$ for $\mathbf{k} \in \mathcal{O}^\dagger$ determine $H(\mathbf{k})$ for all $\mathbf{k} \in \mathbb{R}^d$ due to the following unitary equivalence:

$$H(\mathbf{k}+\mathbf{m}) = e^{-i\mathbf{m}\mathbf{x}} H(\mathbf{k}) e^{i\mathbf{m}\mathbf{x}},\ \mathbf{m} \in \Gamma^\dagger.$$

This implies, in particular, that

$$\lambda_j(H(\mathbf{k}+\mathbf{m})) = \lambda_j(H(\mathbf{k})),\ j = 1, 2, \ldots, \tag{2.1.14}$$

for all $\mathbf{m} \in \Gamma^\dagger$. The images

$$\sigma_j = \bigcup_{\mathbf{k}\in\overline{\mathcal{O}^\dagger}} \lambda_j(H(\mathbf{k})),$$

are called *spectral bands of* H. The spectrum of H is the union

$$\sigma(H) = \bigcup_j \sigma_j.$$

Due to the mentioned boundedness of the counting function $N(\lambda, H(\mathbf{k}))$, each interval $(-\infty, \lambda]$ has non-empty intersection with finitely many spectral bands.

The main result of this section is the following Theorem:

Theorem 2.1.1. *Let* $H = H_0 + \mathrm{Op}(b)$ *where* $H_0 = \mathrm{Op}(a(\boldsymbol{\xi}')|\boldsymbol{\xi}|^{2m})$, *with some* $m > 0$, *and* $b \in \mathbf{S}_\gamma(w)$, $w = \langle\boldsymbol{\xi}\rangle^\beta$, *with some* $\gamma \in \mathbb{R}$ *and* $\beta \in (0, 1)$ *satisfying the condition*

$$2m - 2 > \beta(\gamma - 2). \tag{2.1.15}$$

Here $a(\cdot)$ *- is a positive smooth (the first derivative should be continuous) function and* $\boldsymbol{\xi}' = \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}$. *Then the spectrum of the operator* H *contains a half-line, i.e. there exists a number* $\lambda_0 \in \mathbb{R}$ *such that* $[\lambda_0, \infty) \subset \sigma(H)$. *Moreover, there is a number* $S \in \mathbb{R}$ *and a constant* $c > 0$ *such that for each* $\lambda \geq \lambda_0$ *we have* $\zeta(\lambda; H) \geq c\lambda^S$. *The constant* c *and parameter* λ_0 *are uniform in* b *satisfying* $|b|^{(\gamma)} \leq C$.

If one prefers stating the conditions on b in terms of the "standard" classes $\mathbf{S}_a(\langle\boldsymbol{\xi}\rangle)$, one can re-write Theorem 2.1.1 as follows:

Theorem 2.1.2. *Let $H = H_0 + \mathrm{Op}(b)$ where $H_0 = \mathrm{Op}(a(\boldsymbol{\xi}')|\boldsymbol{\xi}|^{2m})$, with some $m > 0$, and $b \in \mathbf{S}_a(w)$, $w = \langle\boldsymbol{\xi}\rangle$, with some $a < 2m$. Here $a(\cdot)$ - is a positive smooth function and $\boldsymbol{\xi}' = \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}$. Then the spectrum of the operator H contains a half-line, i.e. there exists a number $\lambda_0 \in \mathbb{R}$ such that $[\lambda_0, \infty) \subset \sigma(H)$. Moreover there is a number $S \in \mathbb{R}$ and a constant $c > 0$ such that for each $\lambda \geq \lambda_0$ we have $\zeta(\lambda; H) \geq c\lambda^S$. The constant c and parameter λ_0 are uniform in b satisfying $\boldsymbol{|} b \boldsymbol{|}^{(\gamma)} \leq C$.*

To deduce Theorem 2.1.2 from 2.1.1 it suffices to note that $S_a(\langle\boldsymbol{\xi}\rangle) \subset S_\gamma(\langle\boldsymbol{\xi}\rangle^\beta)$ for any $\beta \in (0, 1)$ and $\gamma = a\beta^{-1}$, and that for this γ the condition (2.1.15) is equivalent to

$$\beta > \frac{a}{2} - m + 1. \tag{2.1.16}$$

2.1.3 Some notational conventions

For any measurable set $\mathcal{C} \subset \mathbb{R}^d$ we denote by $\mathcal{P}(\mathcal{C})$ the operator $\mathrm{Op}(\chi(\,\cdot\,; \mathcal{C}))$, where $\chi(\,\cdot\,; \mathcal{C})$ is the characteristic function of the set $\mathcal{C}$. We denote $\mathcal{H}(\mathcal{C}) = \mathcal{P}(\mathcal{C})\mathcal{H}$, $\mathcal{H} = \mathsf{L}^2(\mathbb{R}^d)$. Accordingly, the fibres $\mathcal{P}(\mathbf{k}, \mathcal{C}), \mathbf{k} \in \mathcal{O}^\dagger$, of $\mathcal{P}(\mathcal{C})$, which act in $\mathfrak{H}$, are PDO's with symbols $\sum_{\mathbf{m}\in\Gamma^\dagger} \chi(\mathbf{m}+\mathbf{k}; \mathcal{C})$. In other words, each $\mathcal{P}(\mathbf{k}; \mathcal{C})$ is a projection in $\mathfrak{H}$ on the linear span of the exponentials

$$E_{\mathbf{m}}(\mathbf{x}) := \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} e^{i\mathbf{m}\cdot\mathbf{x}}, \quad \mathbf{m} \in \Gamma^\dagger : \mathbf{m} + \mathbf{k} \in \mathcal{C}. \tag{2.1.17}$$

The subspace $\mathcal{P}(\mathbf{k}; \mathcal{C})\mathfrak{H}$ of $\mathfrak{H}$ is denoted by $\mathfrak{H}(\mathbf{k}; \mathcal{C})$.

Suppose that $\mathcal{H}(\mathcal{C})$ is an invariant subspace of the operator H defined in (2.1.10), that is $(H - iI)^{-1}\mathcal{H}(\mathcal{C}) \subset \mathcal{H}(\mathcal{C})$. Then the subspace $\mathfrak{H}(\mathbf{k}; \mathcal{C})$, $\mathbf{k} \in \mathcal{O}^\dagger$, is invariant for $H(\mathbf{k})$. We denote by $H(\mathbf{k}; \mathcal{C})$ the part of $H(\mathbf{k})$ in $\mathfrak{H}(\mathbf{k}; \mathcal{C})$, so that

$$H(\mathbf{k}) = H(\mathbf{k}; \mathcal{C}) \oplus H(\mathbf{k}; \mathbb{R}^d \setminus \mathcal{C}),\ \mathbf{k} \in \mathcal{O}^\dagger,$$

where $\oplus$ denotes the orthogonal sum. If $\mathcal{H}(\mathcal{C})$ is invariant for H, then we denote by $N(\lambda, H(\mathbf{k}); \mathcal{C})$ the counting function of $H(\mathbf{k}; \mathcal{C})$ on the subspace $\mathfrak{H}(\mathbf{k}; \mathcal{C})$.

Each $\boldsymbol{\xi} \in \mathbb{R}^d$ can be uniquely represented as the sum $\boldsymbol{\xi} = \mathbf{m} + \mathbf{k}$, where $\mathbf{m} \in \Gamma^\dagger$ and $\mathbf{k} \in \mathcal{O}^\dagger$. We say that $\mathbf{m} =: [\boldsymbol{\xi}]$ is the integer part of $\boldsymbol{\xi}$ and $\mathbf{k} =: \{\boldsymbol{\xi}\}$ is the fractional part of $\boldsymbol{\xi}$.

The notation $B(\mathbf{x}_0, R)$ is used for the open ball in $\mathbb{R}^d$ of radius $R > 0$, centered at $\mathbf{x}_0 \in \mathbb{R}^d$. We also write $B(R)$ for the open ball of radius R centered at 0.

2.2 Elementary estimates for the sets $\Lambda(\boldsymbol{\theta})$

In this section we are going to define the elementary resolvent set $\Lambda(\boldsymbol{\theta})$ and establish its basic properties.

The sets we are interested in will be described in terms of a number ρ. All the considerations will be taken for "sufficiently large" values of ρ. Exact conditions on ρ will be defined on the way.

The set $\Lambda(\boldsymbol{\theta})$ is the set of the points such that when we move from them by the vector $\boldsymbol{\theta}$, the value of the function h_0 changes by *sufficiently small* amount. A precise definition will be given later.

Note the following auxiliary fact.

Lemma 2.2.1. *Let* $b(r) = \min\limits_{\theta_1,..,\theta_k \in \Theta_r} \max\limits_{i=1,..,k} \sin \beta_i$*, where* β_i *is the angle between the vector* $\boldsymbol{\theta}_i$ *and the linear span of the remaining* $k-1$ *vectors and* $\{\theta_1, .., \theta_k\}$ *are linear independent. Then there exists a value* $\varsigma > 0$ *such that* $b(r) \gg r^{\varsigma}$*.*

Proof:

Follows from Lemma 5.1 in [24] □

First we will obtain a necessary condition (in terms of the vector function ψ) for the vector $\boldsymbol{\xi}$ to be in the set $\Lambda(\boldsymbol{\theta}) \cap S(\rho)$ for large enough positive values of $\rho \gg |\boldsymbol{\theta}|^{1/\varkappa}$.

Then some more useful facts will be proven, and the section is finished by an upper bound for the volume of the set $\Lambda(\boldsymbol{\theta}) \cap S(\rho)$.

Consider a positive value α, exact conditions for it will be stated later (in (2.2.13)).

First, let us introduce the definition for $\Lambda(\boldsymbol{\theta})$.

Definition 2.2.2. Let $\boldsymbol{\theta} \in \mathbb{R}^d$ be an arbitrary vector. Then the set $\Lambda(\boldsymbol{\theta}) \subset \mathbb{R}^d$ is defined as follows:

$$\Lambda(\boldsymbol{\theta}) = \{\boldsymbol{\xi} : |\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})| < \rho^{\alpha}|\boldsymbol{\theta}|\}, \tag{2.2.1}$$

where

$$\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) := |h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})|.$$

Lemma 2.2.3. *For any $\boldsymbol{\xi} \in \mathbb{R}^d$ and a positive $t \in \mathbb{R}_+$ we have:*

$$\psi(\boldsymbol{\xi}) = \psi(t\boldsymbol{\xi}).$$

Proof:

Since the function h_0 is homogeneous of order $2m$, it holds that

$$h_0(t\boldsymbol{\xi}) = t^{2m}h_0(\boldsymbol{\xi}),$$

therefore the homothety with the factor t maps the level set $M(\boldsymbol{\xi})$ onto $M(t\boldsymbol{\xi})$. Thus, the directions of the normals at the point $\boldsymbol{\xi}$ (to the level set $M(\boldsymbol{\xi})$) and $t\boldsymbol{\xi}$ (for the level set $M(t\boldsymbol{\xi})$) coincide. □

So the value of $\psi(\boldsymbol{\xi})$ depends only on $\boldsymbol{\xi}' = \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}$, and does not depend on $|\boldsymbol{\xi}|$.

Giver a vector $\gamma \neq 0$, by the derivative $\frac{\partial h_0(\boldsymbol{\xi})}{\partial \gamma}$ we denote the derivative of the function $h_0(\boldsymbol{\xi})$ in the direction of the vector γ at the point $\boldsymbol{\xi}$, that means:

$$\frac{\partial h_0(\boldsymbol{\xi})}{\partial \gamma} = \lim_{t \to 0} \frac{h_0(\boldsymbol{\xi} + t\gamma) - h_0(\boldsymbol{\xi})}{t|\gamma|}.$$

Lemma 2.2.4. *For unit vectors α, β, γ, such that $\gamma = A\alpha + B\beta$, we have that*

$$\frac{\partial h_0(\boldsymbol{\xi})}{\partial \gamma} = A\frac{\partial h_0(\boldsymbol{\xi})}{\partial \alpha} + B\frac{\partial h_0(\boldsymbol{\xi})}{\partial \beta}. \tag{2.2.2}$$

Proof:

$$\frac{\partial h_0(\boldsymbol{\xi})}{\partial \gamma} = \lim_{t \to 0} \frac{h_0(\boldsymbol{\xi} + t\gamma) - h_0(\boldsymbol{\xi})}{t|\gamma|} = \lim_{t \to 0} \frac{h_0(\boldsymbol{\xi} + At\alpha + Bt\beta) - h_0(\boldsymbol{\xi})}{t} =$$

$$= \lim_{t\to 0} \frac{h_0\left(\boldsymbol{\xi} + At\alpha + Bt\beta\right) - h_0\left(\boldsymbol{\xi} + Bt\beta\right) + h_0\left(\boldsymbol{\xi} + Bt\beta\right) - h_0\left(\boldsymbol{\xi}\right)}{t} =$$

$$= \lim_{t\to 0} \frac{h_0\left(\boldsymbol{\xi} + At\alpha + Bt\beta\right) - h_0\left(\boldsymbol{\xi} + Bt\beta\right)}{t} + \lim_{t\to 0} \frac{h_0\left(\boldsymbol{\xi} + Bt\beta\right) - h_0\left(\boldsymbol{\xi}\right)}{t} =$$

(this equality is justified, by decomposing h_0 into Taylor's series)

$$= \lim_{t\to 0} \frac{h_0\left(\boldsymbol{\xi} + At\beta\right) - h_0\left(\boldsymbol{\xi}\right)}{t} + \lim_{t\to 0} \frac{h_0\left(\boldsymbol{\xi} + Bt\beta\right) - h_0\left(\boldsymbol{\xi}\right)}{t} = A\frac{\partial h_0(\boldsymbol{\xi})}{\partial \alpha} + B\frac{\partial h_0(\boldsymbol{\xi})}{\partial \beta}.$$

□

Now we are ready to prove first estimate for the set $\Lambda(\boldsymbol{\theta})$.

Proposition 2.2.5. *There exists a constant C_1, such that for sufficiently large values of ρ and any $\boldsymbol{\theta}$ such that $1 < |\boldsymbol{\theta}| < \rho^{\varkappa}$, the following inclusion holds:*

$$\Lambda(\boldsymbol{\theta}) \cap S(\rho) \subset \Lambda^{(C_1)}(\boldsymbol{\theta}) \cap S(\rho), \tag{2.2.3}$$

where

$$\Lambda^{(C_1)}(\boldsymbol{\theta}) := \left\{ \boldsymbol{\xi} \in S(\rho) : \left| \frac{\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}) \rangle}{|\boldsymbol{\theta}|} \right| < C_1 \rho^{\alpha+1-2m} \right\} \tag{2.2.4}$$

and the set $\Lambda(\boldsymbol{\theta})$ is defined by (2.2.1)

Proof:

Let us transform the difference $h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})$.

Since h_0 is homogeneous of order $2m$,

$$h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) = |\boldsymbol{\xi}|^{2m} h_0\left(\boldsymbol{\xi}' + \frac{\boldsymbol{\theta}}{|\boldsymbol{\xi}|}\right). \tag{2.2.5}$$

Let us decompose the function h_0 in the right hand side of (2.2.5) into Taylor's series (here we use smoothness of h_0, the value $\frac{|\boldsymbol{\theta}|}{|\boldsymbol{\xi}|} < \rho^{\varkappa-1}$ we assume to be sufficiently small)

$$h_0\left(\boldsymbol{\xi}' + \frac{\boldsymbol{\theta}}{|\boldsymbol{\xi}|}\right) = h_0(\boldsymbol{\xi}') + \frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}} |\boldsymbol{\theta}||\boldsymbol{\xi}|^{-1} + O\left(|\boldsymbol{\xi}|^{2\varkappa-2}\right).$$

Therefore, substituting this decomposition into (2.2.5) we obtain

$$h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) = |\boldsymbol{\xi}|^{2m} \left(h_0(\boldsymbol{\xi}') + \frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}} |\boldsymbol{\theta}||\boldsymbol{\xi}|^{-1} + O\left(|\boldsymbol{\xi}|^{2\varkappa-2}\right)\right). \tag{2.2.6}$$

Now denote by $\boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}$ and $\boldsymbol{\theta}^{\perp}_{\psi(\boldsymbol{\xi})}$ vectors of length $|\boldsymbol{\theta}|$, in directions parallel and perpendicular to $\psi(\boldsymbol{\xi})$ correspondingly.

Let us decompose $\boldsymbol{\theta}$ into linear sum of this vectors:

$$\boldsymbol{\theta} = \cos(\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi})))\boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})} + \sin(\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi})))\boldsymbol{\theta}^{\perp}_{\psi(\boldsymbol{\xi})}.$$

Note that since the vector $\boldsymbol{\theta}^{\perp}_{\psi(\boldsymbol{\xi})}$ is tangent to the level set, the derivative $\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{\perp}_{\psi(\boldsymbol{\xi})}} = 0$. Therefore,

$$\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}} = \cos(\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi})))\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}.$$

Applying this expression for the derivative to (2.2.6), we obtain (for the simplicity of the notation, temporally denote $\cos(\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}))) = W$):

$$h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) = |\boldsymbol{\xi}|^{2m}\left(h_0(\boldsymbol{\xi}') + W\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}|\boldsymbol{\theta}||\boldsymbol{\xi}|^{-1} + O\left(|\boldsymbol{\xi}|^{2\varkappa-2}\right)\right). \quad (2.2.7)$$

Now let is transform the formula $\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = |h_0(\boldsymbol{\xi}+\boldsymbol{\theta}) - h_0(\boldsymbol{\xi})|$, using the expression for $h_0(\boldsymbol{\xi} + \boldsymbol{\theta})$ from (2.2.7) and canceling the terms $|\boldsymbol{\xi}|^{2m}h_0(\boldsymbol{\xi}')$:

$$\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = |h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})| = |\boldsymbol{\xi}|^{2m}\left|W\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}|\boldsymbol{\theta}||\boldsymbol{\xi}|^{-1} + O\left(|\boldsymbol{\xi}|^{2\varkappa-2}\right)\right| =$$

$$= \left|W\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}|\boldsymbol{\theta}||\boldsymbol{\xi}|^{2m-1} + O(|\boldsymbol{\xi}|^{2m-2+2\varkappa})\right|. \quad (2.2.8)$$

To obtain a condition that the vector $\boldsymbol{\xi} \in \Lambda(\boldsymbol{\theta})$, use for $\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})$ right hand side of (2.2.8)

$$\Lambda(\boldsymbol{\theta}) \subset \left\{\boldsymbol{\xi} : \left|W\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}|\boldsymbol{\theta}||\boldsymbol{\xi}|^{2m-1} + O(|\boldsymbol{\xi}|^{2m-2+2\varkappa})\right| < \rho^{\alpha}|\boldsymbol{\theta}|\right\}.$$

Dividing by $|\boldsymbol{\xi}|^{2m-1}$ and using that $A_{max}^{-1/m}\rho \le |\boldsymbol{\xi}| \le A_{max}^{1/m}\rho$:

$$\Lambda(\boldsymbol{\theta}) \subset \left\{\boldsymbol{\xi} : \left|W\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}|\boldsymbol{\theta}| + O(\rho^{2\varkappa-1})\right| < A_{max}^2\rho^{\alpha+1-2m}|\boldsymbol{\theta}|\right\}. \quad (2.2.9)$$

Therefore (remembering that $W = \cos(\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}))) = \frac{\langle\boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle}{|\boldsymbol{\theta}|}$), we obtain

$$\Lambda(\boldsymbol{\theta}) \subset \left\{\boldsymbol{\xi} : \left|\frac{\langle\boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle}{|\boldsymbol{\theta}|}\frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}}|\boldsymbol{\theta}| + O(\rho^{2\varkappa-1})\right| < A_{max}^2\rho^{\alpha+1-2m}|\boldsymbol{\theta}|\right\} \Rightarrow$$

$$\Rightarrow \Lambda(\boldsymbol{\theta}) \subset \left\{ \boldsymbol{\xi} : \left| \frac{\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}) \rangle}{|\boldsymbol{\theta}|} \frac{\partial h_0(\boldsymbol{\xi}')}{\partial \boldsymbol{\theta}^{||}_{\psi(\boldsymbol{\xi})}} |\boldsymbol{\theta}| \right| < 2A^2_{max} \rho^{\alpha+1-2m} |\boldsymbol{\theta}| \right\}.$$

The last inequality can be fulfilled by the choice of $\varkappa$. Here $\alpha+1-2m$ is some fixed number in the interval $(-1, 0)$ (we will make in (2.2.13) an assumption about value of α) and the value $2\varkappa - 1$ can be chosen to be as close to -1 as needed.

The statement of this proposition immediately follows if we introduce a notation:

$$\frac{1}{2A^2_{max}} \min_{\boldsymbol{\theta}} \min_{\boldsymbol{\xi}} \left| \frac{1}{|\nabla h_0|} \right| = \frac{1}{C_1}.$$

The internal minimum is nonzero as a minimum of a strictly positive continuous function considered on a compact. The external minimum exists because $|\boldsymbol{\theta}| > 1$. $\square$

Our next aim is to obtain some explicit estimates for the set $\Lambda(\boldsymbol{\theta})$: we will estimate the volume of the intersection $\Lambda(\boldsymbol{\theta})$ with the set $S(\rho)$, defined in (2.1.1).

Proposition 2.2.6. *There exist constants $C_2, C_2' > 0$ independent of ρ such that*

$$C_2|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| \le |\psi(\boldsymbol{\xi}_1) - \psi(\boldsymbol{\xi}_2)| \le C_2'|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|,$$

where $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ are two arbitrary points on the unit sphere S^{d-1}.

Proof:

We will prove the upper bound, the lower one can be proved similarly.

First, assume the proposition is true for the dimension 2. Then we can derive this result for higher dimensions in the following way. For any intersection of a sphere with a plane containing the origin, this constant can be chosen (if the proposition holds for $d = 2$). Therefore, the minimum of these constants taken for different planes will satisfy our statement (since there exists a plain containing our arbitrary points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ and the origin). The minimum exists (and is also non-zero) since the constant continuously depends on the chosen plane, and the set of all such planes is a compact. Thus without loss of generality we can only consider the case $d = 2$.

This means that the points $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ are on a smooth curve on the plane. Also note that the curvature of this curve is positive at every point (since the level sets are strictly convex). Denote by $\kappa^* > 0$ the minimum of the curvature.

Suppose the statement is false. Then (due to the compactness of the set $S^{d-1} \times S^{d-1}$) there exists a convergent sequence of pairs $\{(\boldsymbol{\xi}_1^{(n)}, \boldsymbol{\xi}_2^{(n)})\}_{n\in\mathbb{N}}$, such that the expression $\frac{|\psi(\boldsymbol{\xi}_1^{(n)})-\psi(\boldsymbol{\xi}_2^{(n)})|}{|\boldsymbol{\xi}_1^{(n)}-\boldsymbol{\xi}_2^{(n)}|}$ tends to zero. Denote the limit of the sequences $\{\boldsymbol{\xi}_1^{(n)}\}_{n\in\mathbb{N}}$, $\{\boldsymbol{\xi}_2^{(n)}\}_{n\in\mathbb{N}}$ by $\boldsymbol{\xi}_1^*$ and $\boldsymbol{\xi}_2^*$ correspondingly.

If $\boldsymbol{\xi}_1^* \neq \boldsymbol{\xi}_2^*$, then

$$0 = \lim_{n\to\infty} \frac{|\psi(\boldsymbol{\xi}_1^{(n)}) - \psi(\boldsymbol{\xi}_2^{(n)})|}{|\boldsymbol{\xi}_1^{(n)} - \boldsymbol{\xi}_2^{(n)}|} = \frac{\psi(\boldsymbol{\xi}_1^*) - \psi(\boldsymbol{\xi}_2^*)}{\boldsymbol{\xi}_1^* - \boldsymbol{\xi}_2^*}.$$

But the last expression cannot be zero, since the normal vectors in different points of the strictly convex surface can not coincide. Therefore, $\boldsymbol{\xi}_1^* = \boldsymbol{\xi}_2^*$.

Denote by κ the curvature at the point $\boldsymbol{\xi}_1^*$. Then the value of

$$\lim_{\boldsymbol{\xi}_1\neq\boldsymbol{\xi}_2\to\boldsymbol{\xi}_1^*} \frac{|\psi(\boldsymbol{\xi}_1) - \psi(\boldsymbol{\xi}_2)|}{|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|}$$

for this curve coincides with the same limit for the circle of the radius $\frac{1}{\kappa}$. For the circle this limit can easily be calculated explicitly since for any non-collinear vectors $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ we have $\frac{|\psi(\boldsymbol{\xi}_1)-\psi(\boldsymbol{\xi}_2)|}{|\boldsymbol{\xi}_1-\boldsymbol{\xi}_2|} = \frac{1}{R}$, where R is the radius of the circle. So we can conclude that on the curve that we consider

$$\lim_{\boldsymbol{\xi}_1\neq\boldsymbol{\xi}_2\to\boldsymbol{\xi}_1^*} \frac{|\psi(\boldsymbol{\xi}_1) - \psi(\boldsymbol{\xi}_2)|}{|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|} = \kappa \geq \kappa^* > 0.$$

Thus we arrive at a contradiction. □

This statement has a usefull corollary.

Corollary 2.2.7. *Suppose vectors $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ satisfy the inequality $|\boldsymbol{\xi}| \geq 2|\boldsymbol{\theta}|$. Then we have*

$$|\psi(\boldsymbol{\xi} + \boldsymbol{\theta}) - \psi(\boldsymbol{\xi})| \leq \frac{4C_2'|\boldsymbol{\theta}|}{|\boldsymbol{\xi}|}, \tag{2.2.10}$$

Proof: Applying the Triangle inequality to the left hand side of (2.2.10) we can rewrite it using the homogeneity of the function $\psi(\boldsymbol{\xi})$ as

$$|\psi(\boldsymbol{\xi} + \boldsymbol{\theta}) - \psi(\boldsymbol{\xi})| \leq \left|\psi\left(\frac{(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}|}{|\boldsymbol{\xi}+\boldsymbol{\theta}|}\right) - \psi(\boldsymbol{\xi})\right| + \left|\psi(\boldsymbol{\xi}+\boldsymbol{\theta}) - \psi\left(\frac{(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}|}{|\boldsymbol{\xi}+\boldsymbol{\theta}|}\right)\right| =$$

$$= \left| \psi \left(\frac{(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}|}{|\boldsymbol{\xi}+\boldsymbol{\theta}|} \right) - \psi(\boldsymbol{\xi}) \right|.$$

The second summand equals to zero, since $\psi(\boldsymbol{\xi})$ does not depend on the length of the vector. Due to Proposition 2.2.6 we obtain

$$|\psi(\boldsymbol{\xi}+\boldsymbol{\theta}) - \psi(\boldsymbol{\xi})| \leq \left| \psi \left(\frac{(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}|}{|\boldsymbol{\xi}+\boldsymbol{\theta}|} \right) - \psi(\boldsymbol{\xi}) \right| \leq$$

$$\leq \frac{C_2'}{|\boldsymbol{\xi}|} \left| \frac{(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}|}{|\boldsymbol{\xi}+\boldsymbol{\theta}|} - \boldsymbol{\xi} \right| \leq \frac{C_2'}{|\boldsymbol{\xi}||\boldsymbol{\xi}+\boldsymbol{\theta}|} \left| (\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}| - \boldsymbol{\xi}|\boldsymbol{\xi}+\boldsymbol{\theta}| \right|.$$

Let $|\boldsymbol{\xi}+\boldsymbol{\theta}| = |\boldsymbol{\xi}| + C$, where $|C| \leq |\boldsymbol{\theta}|$. Then

$$|(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}| - \boldsymbol{\xi}|\boldsymbol{\xi}+\boldsymbol{\theta}|| = |(\boldsymbol{\xi}+\boldsymbol{\theta})|\boldsymbol{\xi}| - \boldsymbol{\xi}|\boldsymbol{\xi}| - C\boldsymbol{\xi}| \leq |\boldsymbol{\xi}|\,|C + |\boldsymbol{\theta}|| \leq 2|\boldsymbol{\xi}||\boldsymbol{\theta}|.$$

Thus we conclude that

$$|\psi(\boldsymbol{\xi}+\boldsymbol{\theta}) - \psi(\boldsymbol{\xi})| \leq \frac{2C_2'|\boldsymbol{\theta}|}{|\boldsymbol{\xi}+\boldsymbol{\theta}|} \leq \frac{4C_2'|\boldsymbol{\theta}|}{|\boldsymbol{\xi}|}.$$

□

Let us consider an arbitrary two-dimensional plane $\mathcal{P}$, containing the origin and the vector $\boldsymbol{\theta}$, and consider the unit circle S_1 on this plane. Since the level sets of the function h_0 are convex, there exist exactly two points on this unit circle where the tangent vector to the level curve is collinear to the vector $\boldsymbol{\theta}$. Denote them by $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$.

Now, for an arbitrary point $\boldsymbol{\xi} \neq \boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ on the unit circle, we define the function as follows:

$$F(\boldsymbol{\xi}) := \min\{\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}_1) - \psi(\boldsymbol{\xi})); \angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}_2) - \psi(\boldsymbol{\xi}))\},$$

where $\angle(\delta_1, \delta_2)$ is the acute angle between the lines containing the vectors δ_1 and δ_2.

Lemma 2.2.8. *For the function, constructed as above, the following holds:*

$$\max_{\boldsymbol{\xi}} F(\boldsymbol{\xi}) < \frac{\pi}{2}.$$

Proof:

Denote $\gamma_2 := \max_{\boldsymbol{\xi}} F(\boldsymbol{\xi})$, where the maximum is taken over the unit circle. Obviously, $\gamma_2 \leq \frac{\pi}{2}$.

Let us prove that $\gamma_2 < \frac{\pi}{2}$ by contradiction. Indeed, assume $\gamma_2 = \frac{\pi}{2}$, then there exists a sequence $\{\boldsymbol{\xi}^{(k)}\}_{k=1,2,\ldots}$, such that

$$\lim_{k\to\infty} \min\{\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}_1) - \psi(\boldsymbol{\xi}^{(k)})); \angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}_2) - \psi(\boldsymbol{\xi}^{(k)}))\} = \frac{\pi}{2}.$$

.

We choose a convergent subsequence of this sequence and denote by $\hat{\boldsymbol{\xi}}$ its limit. Then we can state that

$$\angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}_1) - \psi(\hat{\boldsymbol{\xi}})) = \angle(\boldsymbol{\theta}, \psi(\boldsymbol{\xi}_2) - \psi(\hat{\boldsymbol{\xi}})) = \frac{\pi}{2} \Rightarrow$$

$$\Rightarrow \left\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}_1) - \psi(\hat{\boldsymbol{\xi}}) \right\rangle = \left\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}_2) - \psi(\hat{\boldsymbol{\xi}}) \right\rangle = 0.$$

Given that $\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}_2) \rangle = 0$, we can conclude that $\left\langle \boldsymbol{\theta}, \psi(\hat{\boldsymbol{\xi}}) \right\rangle = 0$. This implies that either $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}_1$ or $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}_2$. Neither of this can be true since when we are approaching the point $\boldsymbol{\xi}_1$ (or $\boldsymbol{\xi}_2$) the expression $F(\boldsymbol{\xi})$ tends to zero.

Thus, we conclude that the values of the function $F(\boldsymbol{\xi})$ are separated from $\frac{\pi}{2}$. $\square$

Proposition 2.2.9. *For the set $\Lambda^{(c)}(\boldsymbol{\theta})$ defined in (2.2.4) and sufficiently large value ρ there exists a constant C_4 such that:*

$$\mathrm{vol}(\Lambda^{(c)}(\boldsymbol{\theta}) \cap S(\rho)) \leq C_4 c \rho^{\alpha - 2m + d}.$$

Proof:

Firstly, one should note that

$$\frac{\mathrm{vol}(\Lambda^{(c)}(\boldsymbol{\theta}) \cap S(\rho))}{\mathrm{vol}(S(\rho))} \leq A^2_{max} \frac{\mathrm{vol}_{d-1}(\Lambda^{(c)}(\boldsymbol{\theta}) \cap S^{d-1})}{\mathrm{vol}_{d-1}(S^{d-1})} \tag{2.2.11}$$

(here vol_{d-1} - is a natural measure on S^{d-1}). This follows from the observation that the set $\Lambda^{(c)}(\boldsymbol{\theta})$ is a cone.

Now let us consider a set $D_{\boldsymbol{\theta}} \subset S^{d-1}$ consisting of all the points, where the vector $\boldsymbol{\theta}$ is a tangent vector to the corresponding level set. This set is a $d-2$-dimensional surface embedded into the sphere (the dimension of the surface is one less then the dimension of the sphere).

Let us consider an arbitrary two-dimensional plane $\mathcal{P}$, containing the origin and the vector $\boldsymbol{\theta}$, and consider the intersection of this plane with the unit sphere S^{d-1}(it is a unit circle S_1).

Denote the points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ and the function F according to the notation introduced before Lemma 2.2.8.

By Lemma 2.2.8 we can conclude that the values of the function $F(\boldsymbol{\xi})$ are separated from $\frac{\pi}{2}$.

Using the uniform continuity of $F(\cdot)$ (as a function of $\mathcal{P}$) and the compactness of S^{d-1} we see that there exists a positive $\gamma < \frac{\pi}{2}$, such that for any $\boldsymbol{\xi} \in S^{d-1}$ we have $F(\boldsymbol{\xi}) < \gamma$.

Denote $C_3 := \cos\gamma$. Let us transform the expression $\frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle|}{|\boldsymbol{\theta}|}$ in a following way:

$$\frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle|}{|\boldsymbol{\theta}|} = \frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle - \langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}_i)\rangle|}{|\boldsymbol{\theta}|} = \frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}_i)\rangle|}{|\boldsymbol{\theta}|} \geq$$

$$\geq \frac{|\boldsymbol{\theta}||\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}_i)|\cos\gamma}{|\boldsymbol{\theta}|} = |\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}_i)|\cos\gamma \geq C_2C_3|\boldsymbol{\xi} - \boldsymbol{\xi}_i|.$$

Now applying the formula (2.2.3) we obtain:

$$\boldsymbol{\xi} \in \Lambda^{(c)}(\boldsymbol{\theta}) \cap S(\rho) \Rightarrow \frac{C_2C_3}{c}|\boldsymbol{\xi} - \boldsymbol{\xi}_i| < \rho^{\alpha+1-2m}.$$

so we use similarity transformation with ratio $\frac{1}{\rho}$)

$$\boldsymbol{\xi} \in \Lambda^{(c)}(\boldsymbol{\theta}) \cap S^1 \Rightarrow \frac{C_2C_3}{c}|\boldsymbol{\xi} - \boldsymbol{\xi}_i| < \rho^{\alpha-2m}. \tag{2.2.12}$$

Clearly, there exists a constant C_4 (which does not depend on c), such that

$$\mathrm{vol}_1(\{\boldsymbol{\xi} \in S^1 : \frac{C_2C_3}{c}|\boldsymbol{\xi} - \boldsymbol{\xi}_i| < \rho^{\alpha-2m}\}) \leq \frac{C_4c}{A_{max}^2}\rho^{\alpha-2m} \Rightarrow$$

$$\Rightarrow \frac{\mathrm{vol}_1(\Lambda^{(c)}(\boldsymbol{\theta}) \cap S^1)}{\mathrm{vol}_1(S^1)} \leq \frac{C_4c}{A_{max}^2}\rho^{\alpha-2m}.$$

Integrating in the cylindrical coordinates we arrive at the estimate

$$\frac{\mathrm{vol}_{d-1}(\Lambda^{(c)}(\boldsymbol{\theta}) \cap S^{d-1})}{\mathrm{vol}_{d-1}(S^{d-1})} \leq \frac{C_4c}{A_{max}^2}\rho^{\alpha-2m}.$$

Due to (2.2.11)

$$\frac{\operatorname{vol}(\Lambda^{(c)}(\boldsymbol{\theta}) \cap S(\rho))}{\operatorname{vol}(S(\rho))} \leq C_4 c \rho^{\alpha - 2m}.$$

□

Remark. If $m > \frac{\alpha}{2}$ then the proportion of the volume of $S(\rho)$ occupied by $\Lambda(\boldsymbol{\theta})$, tends to zero when ρ goes to infinity.

We introduce a notation $\alpha_1 = \alpha + 2 - 2m$. The value of the original α we consider to be such that $\alpha_1 \in (0,1)$, so

$$\alpha \in (2m - 2, 2m - 1). \tag{2.2.13}$$

(Obviously, we can increase α without loss of generality)

2.3 Resonance sets

Consider the set Θ_r. For each point $\boldsymbol{\xi}$ we denote by $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$ the set of points that can be reached from $\boldsymbol{\xi}$ in several "steps", with the following rules:

Rules 2.3.1. *1. Each step is a translation by a vector $\pm\boldsymbol{\theta}$ where $\boldsymbol{\theta} \in \Theta_r$.*

2. We are allowed to make a translation step in the direction $\boldsymbol{\theta}$ only if both the starting and the end points of this step belong to the set $\Lambda^{(c)}(\boldsymbol{\theta})$.

This way all the points in $\mathbb{R}^d$ can be divided into equivalence classes $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$, where $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) = \{\boldsymbol{\eta} : \boldsymbol{\eta} \text{ is reachable from } \boldsymbol{\xi}\}$

We say that a subspace $\mathfrak{V} \subset \mathbb{R}^d$ is a *lattice r-subspace* if $\mathfrak{V}$ is spanned by some linearly independent lattice vectors from the set Θ_r. The set of all lattice r-subspaces of dimension n is denoted by $\mathcal{V}(n)$ and $\mathcal{V} = \bigcup_{n=0}^{d} \mathcal{V}(n)$.

The main goal of this section is the construction of the sets $\Xi(\mathfrak{V})$, $\Xi^*(\mathfrak{V})$, $\mathfrak{V} \in \mathcal{V}$ satisfying following properties:

Properties 2.3.2. i. If $\mathfrak{V}_1 \neq \mathfrak{V}_2$, then $\Xi(\mathfrak{V}_1) \cap \Xi(\mathfrak{V}_2) = \emptyset$.

ii. For each $\mathfrak{V}$ $\quad \Xi(\mathfrak{V}) \subset \Xi^*(\mathfrak{V})$.

iii. $\mathbb{R}^d = \bigcup_{\mathfrak{V}} \Xi(\mathfrak{V})$ (the union includes $\mathfrak{V} = \{0\}$).

iv. If $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$, then $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \Xi(\mathfrak{V})$.

v. For any set $\mathfrak{V}$ there exists a direction, such that if $\boldsymbol{\eta}$ is a shift of $\boldsymbol{\xi}$ in this direction and both $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ belong to $\Xi^*(\mathfrak{V})$ then $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) + \boldsymbol{\eta} - \boldsymbol{\xi} \subset \Xi^*(\mathfrak{V})$.

vi. $\mathrm{vol}\left(\bigcup_{\mathfrak{V}:\dim(\mathfrak{V})\geq 1} \Xi^*(\mathfrak{V})\right)$ is not too big.

The construction will be performed in several steps. On each step we will introduce a new class of sets, based on the classes already defined. In the last step we will define the set $\Xi^*(\mathfrak{V})$.

2.3.1 The set $\Xi_1^{(\mathbf{c})}(\mathfrak{V})$

The crucial fact we use here is the statement 2.2.5. It shows us that a point $\boldsymbol{\xi}$ belongs to the set $\Lambda(\boldsymbol{\theta})$ if and only if the scalar product $\langle\psi(\boldsymbol{\xi}), \boldsymbol{\theta}\rangle$ is less then a certain value. It is fairly clear that the set of all $\boldsymbol{\xi}$ defined by bounding $\langle\psi(\boldsymbol{\xi}), \boldsymbol{\theta}\rangle$ from above is constructed in a simple way. Therefore, we would prefer to work with a more convenient set $\Xi_1^{(\mathbf{c})}(\mathfrak{V}^{\boldsymbol{\theta}})$ instead of the set $\Lambda(\boldsymbol{\theta})$, where

$$\mathfrak{V}^{\boldsymbol{\theta}} := span\{\boldsymbol{\theta}\},$$

and the set $\Xi_1^{(\mathbf{c})}(\mathfrak{V}^{\boldsymbol{\theta}})$ is defined as follows (for a lattice subspace $\mathfrak{V}$ let as denote the length of the orthogonal projection of the normal vector $\psi(\boldsymbol{\xi})$ on the space $\mathfrak{V}$ by $(\psi(\boldsymbol{\xi}))_{\mathfrak{V}}$):

Definition 2.3.3. 1. For an arbitrary vector $\boldsymbol{\theta}$ and a number $\mathbf{c} > 0$

$$\Xi_1^{(\mathbf{c})}(\mathfrak{V}^{\boldsymbol{\theta}}) := \{\boldsymbol{\xi} : |\langle\psi(\boldsymbol{\xi}), \boldsymbol{\theta}\rangle| \leq \mathbf{c}|\boldsymbol{\theta}|\}.$$

2. For a lattice subspace $\mathfrak{V} \in \mathcal{V}$

$$\Xi_1^{(\mathbf{c})}(\mathfrak{V}) := \{\boldsymbol{\xi} : |(\psi(\boldsymbol{\xi}))_{\mathfrak{V}}| \leq \mathbf{c}\}.$$

Note that $\mathbf{c}$ here is the parameter describing the width of the zone, which can depend on ρ later on.

For any $\mathbf{c} > 0$ if $\mathfrak{V} = \{0\}$ then $\Xi_1^{(\mathbf{c})}(\mathfrak{V}) = \mathbb{R}^d$.

Let us prove a helpful fact:

Lemma 2.3.4. *Let* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_k$ *be some vectors in a* k*-dimensional space, satisfying the following property: for any* $k-1$*-dimensional plane we can choose* $i \in \{1, .., k\}$ *so that the angle between this plane and the vector* $\boldsymbol{\theta}_i$ *is at least* β^**, where* β^* *is a fixed acute angle.*

Then for any unit vector x *the following inequality holds:*

$$\max_{1\le i\le k} \frac{|\langle x, \boldsymbol{\theta}_i\rangle|}{|\boldsymbol{\theta}_i|} \ge \sin\beta^*.$$

Proof:

Consider an arbitrary unit vector x. Choose a vector $\boldsymbol{\theta}_{i_0}$ from the set $\{\boldsymbol{\theta}_i\}_{i=1..k}$ so that the angle between $\boldsymbol{\theta}_{i_0}$ and the plane orthogonal to x is at least β^*. Then

$$\angle(\boldsymbol{\theta}_{i_0}, x) \le \frac{\pi}{2} - \beta^* \Rightarrow \frac{|\langle x, \boldsymbol{\theta}_i\rangle|}{|\boldsymbol{\theta}_i|} \ge \sin\beta^*. \quad \square$$

Remark. At this point one should note that due to Lemma 2.2.1, for linearly independent vectors from the set Θ_r there exists β such that $\beta^* = 2\beta$ satisfies the conditions of the Lemma 2.3.4, and at the same time following estimate for β holds:

$$\sin\beta \approx r^{-\varsigma}. \tag{2.3.1}$$

Let us keep in mind this condition, though we will still use β, when proving "general" facts. Afterwards, when obtaining final estimates we will replace β, using the expression (2.3.1). Also without loss of generality we can consider the case of $\beta < \frac{\pi}{3}$.

Lemma 2.3.5. *For an arbitrary affine space* $\mathbf{L}$ *(with a condition* $0 \notin \mathbf{L}$*) there exists a unique point* $\boldsymbol{\xi}^*(\mathbf{L}) \in \mathbf{L}$*, such that* $\mathbf{L}$ *is tangent to the level set* $M = \{\boldsymbol{\xi} : h_0(\boldsymbol{\xi}) = h_0(\boldsymbol{\xi}^*)\}$ *at the point* $\boldsymbol{\xi}^*$*.*

Proof:

The fact that the level sets for the function h_0 are strictly convex immediately implies that there can be no more than one point $\boldsymbol{\xi}^*$ satisfying the condition of the lemma. Let us prove that there exists at least one such point.

We will consider two cases.

1. $\dim \mathbf{L} = d-1$. Let $M_A = \{\boldsymbol{\xi} : h_0(\boldsymbol{\xi}) = A\}$ be some level set. It is strictly convex so it has exactly two tangent planes parallel to $\mathbf{L}$. And only one of them can be transformed into $\mathbf{L}$ by a homothetic transformation with a positive scaling factor. Clearly, this homothety transforms the set M_A into some other level set M_B (due to the homogeneity of the function h_0).

Obviously, $\mathbf{L}$ is tangent to M_B which gives us the existence of $\boldsymbol{\xi}^*$.

2. $\dim \mathbf{L} < d-1$. Then we can consider a new linear space $\mathbf{L}_0 = span\{\mathbf{L}, 0\}$. Consider the function h_0 restricted to the space $\mathbf{L}_0$. All its main properties are preserved: homogeneity and strict convexity of the level sets.

Now for this new space $\mathbf{L}_0$ we can apply the result proven in the first case (since the dimension of $\mathbf{L}$ is one less then the dimension of $\mathbf{L}_0$). Consequently there exists a level set $M = \{\boldsymbol{\xi} \in \mathbf{L}_0 : h_0(\boldsymbol{\xi}) = A\}$, such that $\mathbf{L}$ is a tangent plane to M. That means that $\#(\mathbf{L} \cap M) = 1$, and therefore

$$\#(\mathbf{L} \cap \{\boldsymbol{\xi} \in \mathbb{R}^d : h_0(\boldsymbol{\xi}) = A\}) = 1.$$

Also since $\mathbf{L} \subset \mathbf{L}_0$, we have $\mathbf{L} \cap W = \mathbf{L} \cap (W \cap \mathbf{L}_0)$ for any set W.

So we have proven that $\mathbf{L}$ has exactly one common point with M, which means that $\mathbf{L}$ is tangent to it. $\square$

Lemma 2.3.6. *Let $\mathfrak{L}$ be the collection of all planes which do not contain zero. Then the following inequality holds (here $(x)_{\mathbf{L}}$ is the orthogonal projection of the vector x on the plane $\mathbf{L}$):*

$$\tilde{C} = \inf_{\mathbf{L} \in \mathfrak{L}} \inf_{\boldsymbol{\xi} \in \mathbf{L}, \boldsymbol{\xi} \neq \boldsymbol{\xi}^*(\mathbf{L})} \frac{|(\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*(\mathbf{L})))_L|}{|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*(\mathbf{L}))|} > 0,$$

Proof:

For a plane $\mathbf{L}$ which does not contain zero, we define the value $\tilde{C}(\mathbf{L})$ as follows:

$$\tilde{C}(\mathbf{L}) = \inf_{\boldsymbol{\xi} \in \mathbf{L}, \boldsymbol{\xi} \neq \boldsymbol{\xi}^*(\mathbf{L})} \frac{|(\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*(\mathbf{L})))_{\mathbf{L}}|}{|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*(\mathbf{L}))|}.$$

First, let us show that for every $\mathbf{L}$, $\tilde{C}(\mathbf{L}) > 0$.

Let us assume the contrary: $\tilde{C}(\mathbf{L}) = 0$. That would mean that there exists a sequence $\{\boldsymbol{\xi}_k\}_{k\in\mathbb{N}} \subset \mathbf{L}$, such that the sequence of vectors $\psi(\boldsymbol{\xi}_k)$ tends to $\psi(\boldsymbol{\xi}^*(\mathbf{L}))$ or to $-\psi(\boldsymbol{\xi}^*(\mathbf{L}))$. The case of $\psi(\boldsymbol{\xi}^*(\mathbf{L}))$ is trivial and has to be considered separately:

$$\frac{|(\psi(\boldsymbol{\xi}_k) - \psi(\boldsymbol{\xi}^*(\mathbf{L})))_{\mathbf{L}}|}{|\psi(\boldsymbol{\xi}_k) - \psi(\boldsymbol{\xi}^*(\mathbf{L}))|} \geq \frac{\sin(\angle(\psi(\boldsymbol{\xi}_k), \psi(\boldsymbol{\xi}^*(\mathbf{L}))))}{\angle(\psi(\boldsymbol{\xi}_k), \psi(\boldsymbol{\xi}^*(\mathbf{L})))}, \text{ so}$$

$$\psi(\boldsymbol{\xi}_k) \xrightarrow[k\to\infty]{} \psi(\boldsymbol{\xi}^*(\mathbf{L})) \Rightarrow \frac{|(\psi(\boldsymbol{\xi}_k) - \psi(\boldsymbol{\xi}^*(\mathbf{L})))_{\mathbf{L}}|}{|\psi(\boldsymbol{\xi}_k) - \psi(\boldsymbol{\xi}^*(\mathbf{L}))|} \xrightarrow[k\to\infty]{} 1$$

Now consider the case where $\psi(\boldsymbol{\xi}_k)$ tends to $-\psi(\boldsymbol{\xi}^*(\mathbf{L}))$. Since the vector $\psi(\boldsymbol{\xi})$ does not change when we multiply $\boldsymbol{\xi}$ by a positive coefficient, consider a sequence $\{\boldsymbol{\xi}_k^*\}_{k\in\mathbb{N}}$, where $\boldsymbol{\xi}_k^*$ - is the point of the intersection of semi-infinite line starting at the origin and containing the point $\boldsymbol{\xi}_k$ with a unit level set $M(1)$. For this new sequence the property that $\psi(\boldsymbol{\xi}_k^*)$ tends to $-\psi(\boldsymbol{\xi}^*(\mathbf{L}))$ will also hold.

Since the level set is compact, we can choose a convergent subsequence of the sequence $\{\boldsymbol{\xi}_k^*\}_{k\in\mathbb{N}}$. Denote its limit by $\boldsymbol{\xi}^*$. By construction, the half-line starting at 0 and containing $\boldsymbol{\xi}^*$ has either to intersect the plane $\mathbf{L}$ or to be parallel to it. But since $\psi(\boldsymbol{\xi}^*) = -\psi(\boldsymbol{\xi}^*(\mathbf{L}))$, this half-line intersects the plane $-\mathbf{L}$, that means that our assumption was false. Therefore, we have proven that $\tilde{C}(\mathbf{L}) > 0$.

Now we are ready to prove the lemma. Assume that $\tilde{C} = 0$. Then there exists a sequence of planes $\{\mathbf{L}_k\}_{k\in\mathbb{N}}$, such that that corresponding sequence $\tilde{C}(\mathbf{L}_k)$ tends to zero. Therefore we can choose a sequence of planes of same dimension for which this property will also hold (denote this dimension by p). Now, since $\tilde{C}(\mathbf{L}) = \tilde{C}(\upsilon\mathbf{L})$, where υ is a positive number, we can choose a sequence of planes of same dimension tangent to a unit sphere again satisfying the property for $\tilde{C}(\mathbf{L}_k)$.

Now we can choose a convergent subsequence: first we choose subsequence such that the tangency point (on unit sphere) tends to some limit (since the sphere S^d is a compact); and then we choose a convergent subsequence from the of the set planes with the same tangent point on the unit sphere S^{d-1-p} (this can be done since the set of the subspaces tangent to a sphere is compact in the natural metric).

Thus from the condition that $\tilde{C}(\mathbf{L}) > 0$ for each plane $\mathbf{L}$ we see that $\tilde{C}$ can be chosen as a positive constant depending only on the original function h_0.□

Lemma 2.3.7. *Consider a lattice subspace* $\mathfrak{V} \in \mathcal{V}$ *and an arbitrary point* $\boldsymbol{\xi} \in \mathbb{R}^d \backslash \mathfrak{V}$. *We construct a plane* $\mathbf{L} = \boldsymbol{\xi} + \mathfrak{V}$. *Then there exists a point* $\boldsymbol{\xi}^* \in \mathbf{L}$, *such that*

1. *For any* $\boldsymbol{\xi}_0 \in \Xi_1^{(\mathbf{c})}(\mathfrak{V}) \cap \mathbf{L}$ *the following inequality holds:*

$$|\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*)| \leq \frac{\mathbf{c}}{\tilde{C}}.$$

2. *If for some* $\boldsymbol{\xi}_0 \in \mathbf{L}$ *it is true that*

$$|\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*)| \leq \mathbf{c},$$

then $\boldsymbol{\xi}_0 \in \Xi_1^{(\mathbf{c})}(\mathfrak{V}) \cap \mathbf{L}$.

Proof:

Let $\boldsymbol{\xi}^*$ be the point obtained by applying Lemma 2.3.5 to the set $\mathbf{L}$.

Let us prove the first statement of this lemma.

Consider an arbitrary point $\boldsymbol{\xi}_0 \in \Xi_1^{(\mathbf{c})}(\mathfrak{V}) \cap \mathbf{L}$. Denote by ψ_0 the orthogonal projection of the vector $\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*)$ onto L. Note that by the Lemma 2.3.6 $|\psi_0| \geq \tilde{C}|\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*)|$.

But on the other hand

$$|(\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*))_L| = |(\psi(\boldsymbol{\xi}_0))_L| \leq \mathbf{c} \tag{2.3.2}$$

(here we were using the definition of the point $\boldsymbol{\xi}^*$ and the fact that $\boldsymbol{\xi}_0 \in \Xi_1^{(\mathbf{c})}(\mathfrak{V})$).

Consequently applying (2.3.2) we obtain:

$$|\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*)| \leq \frac{|(\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*))_L|}{\tilde{C}} \leq \frac{\mathbf{c}}{\tilde{C}}.$$

This concludes the proof of the first statement of the lemma.

Now let us prove the second part. If $|\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*)| \leq \mathbf{c}$, then for each $\boldsymbol{\theta} \in \mathfrak{V}$ we have

$$|\langle\psi(\boldsymbol{\xi}_0), \boldsymbol{\theta}\rangle| = |\langle\psi(\boldsymbol{\xi}_0) - \psi(\boldsymbol{\xi}^*), \boldsymbol{\theta}\rangle| \leq \mathbf{c}|\boldsymbol{\theta}|,$$

which finishes the proof. □

We denote by $\boldsymbol{\xi}^*(\mathbf{L})$ the point $\boldsymbol{\xi}^*$ constructed for the plane $\mathbf{L}$ using Lemma 2.3.5.

2.3.2 The set $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$

The set $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$ was defined in the beginning of Section 2.3. Here we will obtain several estimates for the vectors belonging to such set.

Consider a k-dimensional lattice subspace $\mathfrak{V}_k \in \mathcal{V}$. Denote the orthogonal projection of the vector $\boldsymbol{\xi}$ on the space $\mathfrak{V}_k$ by $(\boldsymbol{\xi})_{\mathfrak{V}_k}$.

Proposition 2.3.8. *There exists a monotonically increasing sequence of the real numbers $\{c_k\}_{k=1}^d$ independent of ρ, such that if we can reach the point $\boldsymbol{\eta}$ starting from the point $\boldsymbol{\xi}$ following Rules 2.3.1 (1 and 2) by means of translation steps from the set $\mathfrak{V}_k \cap \Theta_r$ (i.e. for each j, the vector $\boldsymbol{\theta}_j$ we use on the j-th step belongs to $\mathfrak{V}_k \cap \Theta_r$) then*

1. $|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\eta})| \le c_k \rho^{\alpha_1 - 1}(\sin\beta)^{(3-2k)_-}$, *where* $A_- = \min\{0, A\}$,
2. $(\psi(\boldsymbol{\xi}))_{\mathfrak{V}_k} \le c_k \rho^{\alpha_1-1}(\sin\beta)^{2-2k}$, $(\psi(\boldsymbol{\eta}))_{\mathfrak{V}_k} \le c_k \rho^{\alpha_1-1}(\sin\beta)^{2-2k}$.

For convenience in writing out some expressions below we will denote $\mathbf{c}_k = c_k \rho^{\alpha_1 - 1}$, where c_k is independent of ρ.

Proof:

The proof goes by induction. Firstly, we consider the base case $k = 1$. In this case the space $\mathfrak{V}_k$ is spanned by a single vector $\boldsymbol{\theta}_1 \in \Theta_r$.

$$(\psi(\boldsymbol{\xi}))_{\mathfrak{V}_1} = \frac{\langle \psi(\boldsymbol{\xi}), \boldsymbol{\theta}_1 \rangle}{|\boldsymbol{\theta}_1|}.$$

The statement that $\boldsymbol{\eta}$ is reachable from the point $\boldsymbol{\xi}$ using only translation steps by $\boldsymbol{\theta}_1$ means that $\boldsymbol{\xi} \in \Lambda(\boldsymbol{\theta}_1)$. Therefore, Proposition 2.2.5 implies that for sufficiently large values of ρ there exists a constant C_0 such that

$$\frac{\langle \psi(\boldsymbol{\xi}), \boldsymbol{\theta}_1 \rangle}{|\boldsymbol{\theta}_1|} \le C_0 \rho^{\alpha_1 - 1},$$

which finishes the proof of the second statement of this lemma in the base case $k = 1$ (one just has to choose $c_1 \ge C_0$).

Now consider the first hypothesis of the lemma in the base case $k = 1$. Consider a point $\boldsymbol{\xi}^* = \boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V}_1)$. By Lemma 2.3.6 applying the fact that $\boldsymbol{\xi}^* \in \Lambda(\boldsymbol{\theta}_1)$

$$|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\eta})| \le |\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*)| + |\psi(\boldsymbol{\xi}^*) - \psi(\boldsymbol{\eta})| \le$$

$$\leq \frac{1}{\tilde{C}}\left(|(\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*))_{\mathfrak{V}_1}| + |(\psi(\boldsymbol{\xi}^*) - \psi(\boldsymbol{\eta}))_{\mathfrak{V}_1}|\right) \leq$$

$$\leq \frac{1}{\tilde{C}}|\psi(\boldsymbol{\xi})_{\mathfrak{V}_1}| + \frac{2}{\tilde{C}}|\psi(\boldsymbol{\xi}^*)_{\mathfrak{V}_1}| + \frac{1}{\tilde{C}}|\psi(\boldsymbol{\eta})_{\mathfrak{V}_1}| \leq \frac{4C_0}{\tilde{C}}\rho^{\alpha_1 - 1}.$$

So choosing the first constant to be $c_1 = 4C_0/\tilde{C}$ we claim both parts of the lemma $k = 1$ to be true in the base case.

Now we move on to the induction. Assume that the statement is true for k and let us prove it for $k+1$ (denote the corresponding lattice subspace by $\mathfrak{V}_{k+1}$). We prove the induction step for the case when $k > 1$. If $k = 1$ all the considerations are similar. Consider the first step when the amount of linear independent vectors $\boldsymbol{\theta}$ used as the translation steps on our way from the point $\boldsymbol{\xi}$ to $\boldsymbol{\eta}$ becomes more then k; that means that there exists a point $\boldsymbol{\eta}_1$ such that we reach $\boldsymbol{\eta}_1$ from $\boldsymbol{\xi}$ only by means of translations by the vectors from the set $\mathfrak{V}_k$, and from the point $\boldsymbol{\eta}_1$ we make a step using for translation the vector $\boldsymbol{\theta}_{k+1} \notin \mathfrak{V}_k$.

The base case of the statement applied to the point $\boldsymbol{\eta}_1$ and the vector $\boldsymbol{\theta}_{k+1}$ gives us

$$(\psi(\boldsymbol{\eta}_1))_{\mathfrak{V}^{\boldsymbol{\theta}_{k+1}}} \leq \mathbf{c}_1. \tag{2.3.3}$$

The inductive hypothesis applied to the point $\boldsymbol{\eta}_1$ and the space $\mathfrak{V}_k$ states that

$$(\psi(\boldsymbol{\eta}_1))_{\mathfrak{V}_k} \leq \mathbf{c}_k(\sin\beta)^{2-2k}. \tag{2.3.4}$$

Due to the monotonicity of the sequence $\{\mathbf{c}_i\}$ we see that the right hand side of (2.3.4) is larger than the right hand side of (2.3.3), thus we can state that

$$\boldsymbol{\eta}_1 \in \Xi_1^{(\mathbf{c})}(\mathfrak{V}_{k+1})$$

with the parameter $\mathbf{c} = \mathbf{c}_k(\sin\beta)^{1-2k}$.

Consider a plane $\mathbf{L} = \boldsymbol{\xi} + \mathfrak{V}_{k+1}$ (note that $\boldsymbol{\eta}_1 \in \mathbf{L}$). Without loss of generality we can assume that $\mathbf{0} \notin \mathbf{L}$ due to the contiuity of the function $\psi(\cdot)$). Applying now the first statement of Lemma 2.3.7 to the plane $\mathbf{L}$, we find a point $\boldsymbol{\xi}^* = \boldsymbol{\xi}^*(\mathbf{L})$ such that

$$|\psi(\boldsymbol{\eta}_1) - \psi(\boldsymbol{\xi}^*)| \leq \frac{\mathbf{c}}{\tilde{C}}.$$

By the induction hypothesis

$$|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\eta}_1)| \leq \mathbf{c}_k (\sin\beta)^{3-2k} \Rightarrow$$

due to the triangle inequality

$$\Rightarrow |\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*)| \leq \mathbf{c}_k (\sin\beta)^{3-2k} + \frac{\mathbf{c}}{\tilde{C}} \Leftrightarrow$$

$$\Leftrightarrow |\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*)| \leq \mathbf{c}_k (\sin\beta)^{3-2k} + \frac{\mathbf{c}_k}{\tilde{C}} (\sin\beta)^{1-2k} \leq$$

(given that $\mathbf{c}_{k+1} \geq 4\mathbf{c}_k(1/\tilde{C} + 1)$)

$$\leq \frac{\mathbf{c}_{k+1}}{2} (\sin\beta)^{1-2k} = \frac{\mathbf{c}_{k+1}}{2} (\sin\beta)^{3-2(k+1)}. \tag{2.3.5}$$

Similarly, it is easy to show that

$$|\psi(\boldsymbol{\eta}) - \psi(\boldsymbol{\xi}^*)| \leq \frac{\mathbf{c}_{k+1}}{2} (\sin\beta)^{3-2(k+1)},$$

so

$$|\psi(\boldsymbol{\eta}) - \psi(\boldsymbol{\xi})| \leq \mathbf{c}_{k+1} (\sin\beta)^{3-2(k+1)}.$$

Therefore, by (2.3.5) and recalling the Definition 2.3.3 we have proven that

$$\boldsymbol{\xi} \in \Xi_1^{(\mathbf{c}')}(\mathfrak{V}_{k+1}),$$

with the parameter $\mathbf{c}' = \mathbf{c}_{k+1}(\sin\beta)^{1-2k}$.

Now for an arbitrary element of $\Xi_1^{(\mathbf{c})}$ and any vector $\boldsymbol{\theta} \in \mathfrak{V}_{k+1} \cap \Theta_r$ we can state that

$$\langle \psi(\boldsymbol{\xi}), \boldsymbol{\theta} \rangle \leq |\boldsymbol{\theta}| \mathbf{c}_{k+1} (\sin\beta)^{1-2k} \Leftrightarrow (\psi(\boldsymbol{\xi}))_{\langle \boldsymbol{\theta} \rangle} \leq \mathbf{c}_{k+1} (\sin\beta)^{1-2k}.$$

We plug the vector $\boldsymbol{\theta} = \boldsymbol{\theta}_{k+1}$ in this inequality and apply it the the second part of the induction hypothesis. Given that we already obtained the upper bounds for the orthogonal projections of the vector $\psi(\boldsymbol{\xi})$ on the linear spaces $\mathfrak{V}_k$ and $\langle \boldsymbol{\theta}_{k+1} \rangle$, we can write down an upper bound for the projection on the space $\mathfrak{V}_{k+1}$ as the maximum of those two projections divided by $\sin\beta$:

$$(\psi(\boldsymbol{\xi}))_{\mathfrak{V}_{k+1}} \leq \mathbf{c}_{k+1} (\sin\beta)^{-2k} = \mathbf{c}_{k+1} (\sin\beta)^{2-2(k+1)}.$$

This way we have proven the induction step for the second part of our statement. The induction step for the first part immediately follows from (2.3.5). □

Corollary 2.3.9. *There exists a constant C_5, such that under the assumptions of the 2.3.8 (and the assumption that $\boldsymbol{\xi}, \boldsymbol{\eta} \in S(\rho)$) we have*

$$|\boldsymbol{\xi} - \boldsymbol{\eta}| \le C_5 \rho^{\alpha_1} (\sin \beta)^{(3-2d)}.$$

Proof:

Follows from Proposition 2.3.8 and Proposition 2.2.6. □

2.3.3 The set $\Xi_2(\mathfrak{V})$

Definition 2.3.10. Let $\mathfrak{V} \in \mathcal{V}$. Then we define $\Xi_2(\mathfrak{V})$ as follows:

$$\Xi_2(\mathfrak{V}) = \bigcup_{\boldsymbol{\xi} \in \Xi_1^{(\hat{\mathbf{c}}_n)}(\mathfrak{V})} \boldsymbol{\Upsilon}(\boldsymbol{\xi}), \text{ for } n = \dim(\mathfrak{V}),$$

with the parameters $\hat{\mathbf{c}}_n = C\rho^{\alpha_n}$, where the sequence $\{\alpha_k\}_{k=1}^d$ is increasing and each $\alpha_n \in (0,1)$. Here C is an arbitrary constant chosen so that $C > c_d$ where c_d is the last term of the sequence in the conditions of the Proposition 2.3.8.

Proposition 2.3.11. *Consider two different lattice subspaces $\mathfrak{V}_1, \mathfrak{V}_2$, $\mathfrak{V}_1 \neq \mathfrak{V}_2$. Then for sufficientlysmall parameter $\varkappa$*

$$\Xi_2(\mathfrak{V}_1) \cap \Xi_2(\mathfrak{V}_2) \subset \Xi_2(\mathfrak{V}_1 + \mathfrak{V}_2).$$

Proof:

If $\mathfrak{V}_1 \subset \mathfrak{V}_2$ the statement turns into the obvious one, so we can assume without loss of generality that neither of these spaces is a subset of the other.

Consider an arbitrary point from the intersection $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V}_1) \cap \Xi_2(\mathfrak{V}_2)$. Denote by k the maximum of the dimensions of $\mathfrak{V}_1$ and $\mathfrak{V}_2$. According to the Proposition 2.3.8:

$$\begin{cases} (\psi(\boldsymbol{\xi}))_{\mathfrak{V}_1} & \le & c_k \rho^{\alpha_1 - 1} (\sin \beta)^{2-2k}, \\ (\psi(\boldsymbol{\xi}))_{\mathfrak{V}_2} & \le & c_k \rho^{\alpha_1 - 1} (\sin \beta)^{2-2k}. \end{cases}$$

Given that we have upper bounds for the projections of the vector $\psi(\boldsymbol{\xi})$ on two linear spaces, we can write down the estimate for the projection on the $\mathfrak{V}_1 + \mathfrak{V}_2$ as the maximum of these two projections divided by $\sin\beta$:

$$(\psi(\boldsymbol{\xi}))_{\mathfrak{V}_1\cup\mathfrak{V}_2} \le c_k\rho^{\alpha_1-1}(\sin\beta)^{1-2k}.$$

Therefore, the point $\boldsymbol{\xi}$ belongs to the set $\Xi_1(\mathfrak{V}_1 + \mathfrak{V}_2)$ (and thus also to the set $\Xi_2(\mathfrak{V}_1 + \mathfrak{V}_2)$) whenever

$$c_k\rho^{\alpha_1-1}(\sin\beta)^{1-2k} < \rho^{\alpha_l-1},$$

where l is the dimension of $\mathfrak{V}_1 \cup \mathfrak{V}_2$. Obviously, $l > k$, and because $C > c_k$

$$c_k\rho^{\alpha_1-1}(\sin\beta)^{1-2k} < C\rho^{\alpha_l-1} \Leftarrow (\sin\beta)^{1-2k} < \rho^{\alpha_l-\alpha_1}.$$

Now one can see that for our goal the sequence $\{\alpha_k\}$ has to be chosen in such a way that

$$(\sin\beta)^{1-2k} < \rho^{\alpha_{k+1}-\alpha_1}. \tag{2.3.6}$$

Apply the formulas (2.1.2) and (2.3.1). Using (2.3.1) we obtain

$$(\sin\beta)^{1-2k} < \rho^{\alpha_{k+1}-\alpha_1} \Leftrightarrow r^{-\varsigma(1-2k)} < C\rho^{\alpha_{k+1}-\alpha_1},$$

and then applying (2.1.2) we conclude that

$$r^{\varsigma(2k-1)} < \rho^{\alpha_{k+1}-\alpha_k} \Leftarrow \rho^{\varkappa\varsigma(2k-1)} < \rho^{\alpha_{k+1}-\alpha_1}.$$

Recall that the value ς is fixed, and the parameter $\varkappa$ can be chosen sufficiently small so that for any k it will be true that

$$\varkappa\varsigma(2k-1) < \alpha_{k+1} - \alpha_1, \tag{2.3.7}$$

since $\{\alpha_k\}$ is strictly increasing. Also as a corollary

$$\varkappa\varsigma(2k-1) < (\alpha_{k+1} - \alpha_1) \Rightarrow \rho^{\varkappa\varsigma(2k-1)} < \rho^{\alpha_{k+1}-\alpha_1}.$$

□

In this proof we obtained all the necessary conditions for the sequence $\{\alpha_k\}$ no other will appear.

Here we have obtained all the conditions on the terms of the sequence $\{\alpha_k\}$ when $k \leq d-1$. Also we will postulate that

$$a_d = a_{d-1} + \varkappa d^2. \tag{2.3.8}$$

Obviously, for sufficiently small value of the parameter $\varkappa$ all the terms of the sequence $\{\alpha_k\}$ will belong to the interval $(0,1)$.

Let us make an easy remark:

Lemma 2.3.12.

$$\boldsymbol{\xi} \in \Xi_2(\mathfrak{V}) \Leftrightarrow \boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \Xi_2(\mathfrak{V}).$$

Proof:

$\Rightarrow$ By definition of $\Xi_2(\mathfrak{V})$, if $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$, then there exist a $\boldsymbol{\eta} \in \Xi_1^{\mathbf{c}_n}(\mathfrak{V})$ such that $\boldsymbol{\xi} \in \boldsymbol{\Upsilon}(\boldsymbol{\eta})$. Therefore $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) = \boldsymbol{\Upsilon}(\boldsymbol{\eta}) \subset \Xi_2(\mathfrak{V})$

$\Leftarrow$ Is also true since $\boldsymbol{\xi} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$.

□

Now we are going to estimate the volume of the union of the sets $\Xi_2(\mathfrak{V})$ (aiming to satisfy point vi of Properties 2.3.2):

Proposition 2.3.13.

$$\mathrm{vol}\left(\bigcup_{\mathfrak{V}:\dim\mathfrak{V}\geq 1} \Xi_2(\mathfrak{V})\right) \ll \rho^{d-3+2\alpha_d+(2d-2)\varkappa\varsigma}.$$

Proof:

Consider the lattice $\mathfrak{V}$ and the set $\Xi_2(\mathfrak{V})$. Let $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$. Then (due to the definition of Ξ_2) there exists a point $\boldsymbol{\eta} \in \Xi_1^{(\mathbf{c}_d)}$, such that $\boldsymbol{\xi} \in \boldsymbol{\Upsilon}(\boldsymbol{\eta})$. (Let $\mathbf{L}$ be a linear space spanned by the lattice $\mathfrak{V}$ and the point $\boldsymbol{\xi}$. Consider a point $\boldsymbol{\xi}^*(\mathbf{L})$, defined in Lemma 2.3.7 (We do not need to consider the case of $\mathbf{0} \in \mathbf{L}$ since we are interested in the volume).

Due to Lemma 2.3.7 following holds:

$$|\psi(\boldsymbol{\eta}) - \psi(\boldsymbol{\xi}^*(\mathbf{L}))| \leq \tilde{C}\mathbf{c}_d. \tag{2.3.9}$$

From Proposition 2.3.8 we have:

$$|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\eta})| \leq \mathbf{c}_d \rho^{(2d-3)\varkappa\varsigma}. \tag{2.3.10}$$

Therefore (right hand side of (2.3.10) is greater the right hand side of (2.3.9))

$$|\psi(\boldsymbol{\xi}) - \psi(\boldsymbol{\xi}^*(\mathbf{L}))| \leq 2\mathbf{c}_d \rho^{(2d-3)\varkappa\varsigma}.$$

Now we can apply again Lemma 2.3.7 and conclude that

$$\boldsymbol{\xi} \in \Xi_1^{(c^*)},$$

where $c^* = 2\mathbf{c}_d \rho^{(2d-3)\varkappa\varsigma}$.

Thus,

$$\bigcup_{\mathfrak{V}:\dim \mathfrak{V} \geq 1} \Xi_2(\mathfrak{V}) \subset \bigcup_{\mathfrak{V}:\dim \mathfrak{V} \geq 1} \Xi_1^{(c^*)}(\mathfrak{V}). \tag{2.3.11}$$

One can easily note that

$$\Xi_1^{(c^*)}(\mathfrak{V}) \subset \Xi_1^{(c^*)}(\mathfrak{V}_1),$$

where $\mathfrak{V}_1$ is a one dimensional lattice subspace of $\mathfrak{V}$.

But for a one dimensional lattice $\mathfrak{V}_1$ due to Proposition 2.2.9 the following inequality holds

$$\mathrm{vol}(\Xi_1^{(c^*)}(\mathfrak{V}_1)) \ll c^* \rho^{d+\alpha_1-2}.$$

The total number of one-dimensional subspaces can be estimated by $\#(\Theta_r) \leq \hat{C}\rho^{\varkappa d}$ for a positive constant $\hat{C}$, therefore, applying (2.3.11), we obtain

$$\mathrm{vol}\left(\bigcup_{\mathfrak{V}:\dim \mathfrak{V} \geq 1} \Xi_2(\mathfrak{V}) \right) \ll \rho^{\varkappa d} \rho^{d+\alpha_1-2} \mathbf{c}_d \rho^{(2d-3)\varkappa\varsigma}.$$

Finally, keeping in mind the fact that $\mathbf{c}_d \asymp \rho^{\alpha_d - 1}$ (see Proposition 2.3.8), we obtain an estimate

$$\mathrm{vol}\left(\bigcup_{\mathfrak{V}:\dim \mathfrak{V} \geq 1} \Xi_2(\mathfrak{V}) \right) \ll \rho^{\varkappa d + d + \alpha_1 - 2 + \alpha_1 - 1 + (2d-3)\varkappa\varsigma} = \rho^{d-3+2\alpha_d+(2d-2)\varkappa\varsigma}.$$

□

2.3.4 The set $\Xi(\mathfrak{V})$

Take a lattice subspace $\mathfrak{V} \in \mathcal{V}$. Then the set $\Xi(\mathfrak{V})$ is defined by the formula

$$\Xi(\mathfrak{V}) := \Xi_2(\mathfrak{V}) \setminus \left(\bigcup_{\mathfrak{V}' \supsetneq \mathfrak{V}} \Xi_2(\mathfrak{V}') \right).$$

We aim to prove that for sufficiently large values of ρ the new defined sets $\Xi(\mathfrak{V})$ do not intersect.

Proposition 2.3.14. *If $\mathfrak{V}_1 \neq \mathfrak{V}_2$, then $\Xi(\mathfrak{V}_1) \cap \Xi(\mathfrak{V}_2) = \emptyset$.*

Proof:

By definition and Proposition 2.3.11:

$$\Xi(\mathfrak{V}_1) \cap \Xi(\mathfrak{V}_2) \subset \Xi_2(\mathfrak{V}_1) \cap \Xi_2(\mathfrak{V}_2) \subset \Xi_2(\mathfrak{V}_1 + \mathfrak{V}_2).$$

But by construction $\Xi(\mathfrak{V}_1) \cap \Xi_2(\mathfrak{V}) = \emptyset$ for any $\mathfrak{V}$, such that $\mathfrak{V}_1 \subset \mathfrak{V}$, thus it also holds for $\mathfrak{V} = \mathfrak{V}_1 + \mathfrak{V}_2$. So

$$\Xi(\mathfrak{V}_1) \cap \Xi(\mathfrak{V}_2) = \Xi(\mathfrak{V}_1) \cap (\Xi(\mathfrak{V}_1) \cap \Xi(\mathfrak{V}_2)) \subset \Xi(\mathfrak{V}_1) \cap \Xi_2(\mathfrak{V}_1 + \mathfrak{V}_2) = \emptyset. \quad \square$$

Proposition 2.3.15. *If $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ then $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \Xi(\mathfrak{V})$.*

Proof:

Suppose the opposite: let $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$ and $\boldsymbol{\eta} \notin \Xi(\mathfrak{V})$. Then there exists some $\mathfrak{V}' \supset \mathfrak{V}$ such that $\boldsymbol{\eta} \in \Xi_2(\mathfrak{V}')$ holds. Now, applying the fact that $\boldsymbol{\xi} \in \boldsymbol{\Upsilon}(\boldsymbol{\eta})$ and Lemma 2.3.12 we can conclude that $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V}')$. Hence $\boldsymbol{\xi} \notin \Xi(\mathfrak{V})$, which contradicts the assumption of our Proposition. $\square$

Lemma 2.3.16. *If $\boldsymbol{\theta}$ and $\mathfrak{V}_0$ are such that $\boldsymbol{\theta} \notin \mathfrak{V}_0$, then*

$$\Lambda(\boldsymbol{\theta}) \cap \Xi(\mathfrak{V}_0) = \emptyset.$$

Proof: Clearly,

$$\Lambda(\boldsymbol{\theta}) \subset \bigcup_{\mathfrak{V}: \boldsymbol{\theta} \in \mathfrak{V}} \Xi(\mathfrak{V}),$$

but for any $\mathfrak{V} : \boldsymbol{\theta} \in \mathfrak{V}$ it holds that $\Xi(\mathfrak{V}) \cap \Xi(\mathfrak{V}_0) = \emptyset$. $\square$

Lemma 2.3.17.

$$\bigcup \Xi(\mathfrak{V}) = \mathbb{R}^d. \tag{2.3.12}$$

Proof: Obvious.

2.3.5 The set $\Xi^*(\mathfrak{V})$

We build a new family of sets in the following way: for each $\mathfrak{V}$ we

1. build the set $\Xi(\mathfrak{V})$,
2. construct the intersection $\Xi(\mathfrak{V}) \cap L_{A_{max}\rho}$, where

$$L_{A_{max}\rho} := \{\boldsymbol{\xi} : |\boldsymbol{\xi}| = A_{max}\rho\}$$

3. "extend" this intersection towards the origin for $\boldsymbol{\xi} \in S(\rho)$.

The set that we obtain as a result of this sequence of steps we denote by $\Xi^*(\mathfrak{V})$.

Let us recall that in the Property 2.3.2 (iii) we red for existence of a direction, such that if $\boldsymbol{\eta}$ is a shift of $\boldsymbol{\xi}$ in this direction and both $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ belong to $\Xi^*(\mathfrak{V})$ then $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) + \boldsymbol{\eta} - \boldsymbol{\xi} \subset \Xi^*(\mathfrak{V})$.

Define the mapping F in the following way:

$$F(\boldsymbol{\xi}) := \boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V}).$$

Taking some point $\boldsymbol{\xi}$ from the set $\Xi(\mathfrak{V})$ we construct the plane $L = \boldsymbol{\xi} + \mathfrak{V}$ and apply Lemma 2.3.7 to obtain the point $\boldsymbol{\xi}^*$.

Now we define $\Xi^*(\mathfrak{V})$. If $\dim(\mathfrak{V}) > 0$, then we have

$$\Xi^*(\mathfrak{V}) := \left\{\boldsymbol{\xi} + \gamma F(\boldsymbol{\xi}) : \ \boldsymbol{\xi} \in \Xi_2(\mathfrak{V}) \cap L_{A_{max}\rho}, \gamma \in \left(-\frac{1}{A_{max}^2}, 0\right)\right\} \cap S(\rho).$$

For $\mathfrak{X} = \{0\}$

$$\Xi^*(\mathfrak{X}) = \mathbb{R}^d \setminus \bigcup_{\dim \mathfrak{V} \geq 1} \Xi^*(\mathfrak{V}).$$

The sets

$$\mathcal{B} := \Xi(\mathfrak{X}) = \mathbb{R}^d \setminus \bigcup_{m \geq 1} \bigcup_{\mathfrak{W} \in \mathcal{V}(m)} \Xi(\mathfrak{W}) \quad \text{and} \quad \mathcal{D} = \mathbb{R}^d \setminus \mathcal{B}. \tag{2.3.13}$$

are called the *non-resonant set* and *resonant set* of $\mathbb{R}^d$ respectively. The sets introduced above obviously depend on the parameter ρ. Whenever necessary, the dependence on ρ is reflected in the notation, e.g. $\Xi(\mathfrak{V};\rho)$, $\mathcal{B}(\rho), \mathcal{D}(\rho)$.

Similarly denote

$$\mathcal{B}_L := \Xi(\mathfrak{X}) = \mathbb{R}^d \setminus \bigcup_{m\geq 1} \bigcup_{\mathfrak{W}\in\mathcal{V}(m)} \Xi^*(\mathfrak{W}) \quad \text{and} \quad \mathcal{D}_L = \mathbb{R}^d \setminus \mathcal{B}_L. \tag{2.3.14}$$

For simplicity we introduce slightly abused notation. Consider a point $\boldsymbol{\xi}$ and a lattice subspace $\mathfrak{V}$. The point $\boldsymbol{\xi}^*(\boldsymbol{\xi}+\mathfrak{V}) \in \boldsymbol{\xi}+\mathfrak{V}$ is the tangency point to the level set containing $\boldsymbol{\xi}$ (as in Lemma 2.3.5).

For an arbitrary vector $\boldsymbol{\mu} \in S(\rho)$ denote by $\boldsymbol{\mu}_{\not\parallel\mathfrak{V}}$ the vector

$$\boldsymbol{\mu}_{\not\parallel\mathfrak{V}} = \boldsymbol{\mu} + A\boldsymbol{\xi}^*(\boldsymbol{\mu}+\mathfrak{V}) \in \mathfrak{V} \tag{2.3.15}$$

(clearly, there exists a unique $A \in \mathbb{R}$ with such a property), and $\boldsymbol{\mu} - \boldsymbol{\mu}_{\not\parallel\mathfrak{V}}$ we denote by $\boldsymbol{\mu}_{\parallel\mathfrak{V}}$.

Note an immediate consequence of the definition of set Ξ_2 and Proposition 2.3.8:

Lemma 2.3.18. *Let $\mathfrak{V} \in \mathcal{V}(n)$, $n = 1, 2, \ldots, d$, and $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$. Then for sufficiently large ρ we have:*

$$|\boldsymbol{\xi}_{\not\parallel\mathfrak{V}}| \leq 2\rho^{\alpha_d - 1}. \tag{2.3.16}$$

Proof: By Definition 2.3.10, $\boldsymbol{\xi} \in \boldsymbol{\Upsilon}(\boldsymbol{\eta})$ for some $\boldsymbol{\eta} \in \Xi_1(\mathfrak{V})$. Thus, by Proposition 2.3.8,

$$|\boldsymbol{\xi}_{\not\parallel\mathfrak{V}}| \leq |\boldsymbol{\eta}_{\not\parallel\mathfrak{V}}| + \max_{\mathbf{m}\in\boldsymbol{\Upsilon}(\boldsymbol{\eta})} |\mathbf{m} - \boldsymbol{\eta}| \ll 2\rho^{\alpha_1 - 1} r^{2d-3} \leq$$
$$\leq 2\rho^{\alpha_1 - 1} r^{d^2} = 2\rho^{\alpha_1 - 1 + \varkappa d^2} \leq 2\rho^{\alpha_{d-1} - 1 + \varkappa d^2} = 2\rho^{\alpha_d - 1}.$$

In view of monotonicity of α_j's and of (2.3.8), this proves (2.3.16).

Lemma 2.3.19. *Let $\mathfrak{V} \in \mathcal{V}(n)$, $n = 1, 2, \ldots, d$, and $\boldsymbol{\xi} \in \Xi^*(\mathfrak{V})$. Then for sufficiently large ρ we have:*

$$|\boldsymbol{\xi}_{\not\parallel\mathfrak{V}}| \leq 2\rho^{\alpha_d - 1}. \tag{2.3.17}$$

The proof can be easily obtained following the proof of the previous Lemma.

This Lemma shows that the resonant sets $\Xi^*(\mathfrak{V}), \mathfrak{V} \neq \mathfrak{X}$, are "small" relative to the non-resonant set $\mathcal{B}^* = \Xi^*(\mathfrak{X})$. More precisely, we show that the resonant set $\mathcal{D}^*$ has a small angular measure. To this end for each $\boldsymbol{\theta} \in \Theta_r$ define

$$\tilde{\Lambda}(\boldsymbol{\theta}) = \{\boldsymbol{\xi} \in \mathbb{R}^d : |\psi(\boldsymbol{\xi}) \cdot \boldsymbol{\theta}| < 2\rho^{\alpha_{d-1}-1}|\boldsymbol{\theta}|\}.$$

By Lemma 2.3.19, for any $\mathfrak{V} \in \mathcal{V}(n), n \leq d-1$ we have

$$\Xi_L(\mathfrak{V}) \subset \bigcup_{\boldsymbol{\theta} \in \mathfrak{V} \cap \Theta_r} \tilde{\Lambda}(\boldsymbol{\theta}),$$

so that

$$\mathcal{D} \setminus B\left(\frac{\rho}{4A_{max}}\right) \subset \bigcup_{\boldsymbol{\theta} \in \Theta_r} \tilde{\Lambda}(\boldsymbol{\theta}) \setminus B\left(\frac{\rho}{4A_{max}}\right).$$

An elementary calculation shows that

$$\tilde{\Lambda}(\boldsymbol{\theta}) \setminus B\left(\frac{\rho}{4A_{max}}\right) \subset S(\boldsymbol{\theta}; \rho) \times \left[\frac{\rho}{4A_{max}}, \infty\right),$$

$$S(\boldsymbol{\theta}; \rho) := \{\boldsymbol{\Omega} \in \mathbb{S}^{d-1} : |\boldsymbol{\psi}(\Omega) \cdot \boldsymbol{\theta}| < 8A_{max}\rho^{\alpha_{d-1}-1}|\boldsymbol{\theta}|\},$$

for all sufficiently large ρ. Let

$$\mathbf{S}(\rho) = \bigcup_{\boldsymbol{\theta} \in \Theta_r} S(\boldsymbol{\theta}; \rho), \quad \mathbf{T}(\rho) = \mathbb{S}^{d-1} \setminus \mathbf{S}(\rho). \tag{2.3.18}$$

Lemma 2.3.20. *Let the sets* $\mathbf{S}(\rho) \subset \mathbb{S}^{d-1}$, $\mathbf{T}(\rho) \subset \mathbb{S}^{d-1}$ *be as defined above. Then*

$$\mathrm{vol}_{\mathbb{S}^{d-1}} \mathbf{S}(\rho) \ll \rho^{\alpha_d - 1}, \; \mathrm{vol}_{\mathbb{S}^{d-1}} \mathbf{T}(\rho) \asymp 1, \tag{2.3.19}$$

for sufficiently large ρ.

Proof: The elementary bound

$$\mathrm{vol}_{\mathbb{S}^{d-1}} S(\boldsymbol{\theta}; \rho) = \int_{\mathbb{S}^{d-2}} \int_{|\cos\omega| \leq 8A_{max}\rho^{\alpha_{d-1}-1}} \sin^{d-2}\omega d\omega d\hat{\boldsymbol{\Omega}} \ll \rho^{\alpha_{d-1}-1},$$

together with the observation that the number of the sets $\tilde{\Lambda}(\boldsymbol{\theta})$ is bounded above by $\mathrm{card}\,\Theta_r \ll r^d$, gives the estimate

$$\mathrm{vol}_{\mathbb{S}^{d-1}} \mathbf{S}(\rho) \ll r^d \rho^{\alpha_{d-1}-1} \ll \rho^{\alpha_d - 1}.$$

The second bound in (2.3.19) immediately follows from the first one by definition (2.3.18).

In what follows, apart from the non-resonant set $\mathcal{B}(\rho)$, the set

$$\tilde{\mathcal{B}}(\rho) = \mathbf{T}(\rho) \times \left[\frac{\rho}{4A_{max}}, \infty\right) \tag{2.3.20}$$

will play an important role. Since $\mathcal{D} \subset \mathbf{S}(\rho) \times \left[\frac{\rho}{4A_{max}}, \infty\right)$, and $\mathcal{B} = \mathbb{R}^d \setminus \mathcal{D}$ (see (2.3.13)), we have

$$\tilde{\mathcal{B}}(\rho) \subset \mathcal{B}(\rho).$$

2.4 Properties of periodic PDO's

In this section we follow the path of the corresponding section in [24]. Main corrections are induced by the change of the considered cut-off functions in the formuli (2.4.8), (2.4.9).

In this section we collect various properties of periodic PDO's to be used in what follows.

2.4.1 Some basic results on the calculus of periodic PDO's

We begin by listing some elementary results for periodic PDO's, some of which can be found in [38].

Recall that $\mathsf{S}(\mathbb{R}^d)$ is taken as a natural domain of $\mathrm{Op}(b)$. Unless otherwise stated, all the symbols are supposed to belong to the class $\mathbf{S}_\gamma = \mathbf{S}_\gamma(w; \Gamma)$, $\gamma \in \mathbb{R}$, with an arbitrary function w satisfying (2.1.3) and a lattice Γ. The functions w and the lattice Γ are usually omitted from the notation.

Proposition 2.4.1. *(See e.g. [38]) Suppose that* $\|b\|^{(0)}_{l,0} < \infty$ *with some* $l > d$. *Then* $B = \mathrm{Op}(b)$ *is bounded in* $\mathcal{H}$ *and* $\|B\| \le C \|b\|^{(0)}_{l,0}$, *with a constant* C *independent of* b.

Since $\mathrm{Op}(b)u \in \mathsf{S}(\mathbb{R}^d)$ for any $b \in \mathbf{S}_\gamma$ and $u \in \mathsf{S}(\mathbb{R}^d)$, the product $\mathrm{Op}(b)\,\mathrm{Op}(g)$, $b \in \mathbf{S}_{\gamma_1}, g \in \mathbf{S}_{\gamma_2}$, is well defined on $\mathsf{S}(\mathbb{R}^d)$. A straightforward calculation gives the following formula for the symbol $b \circ g$ of the product $\operatorname{Op}(b)\operatorname{Op}(g)$:

$$(b \circ g)(\mathbf{x}, \boldsymbol{\xi}) = \frac{1}{\mathrm{d}(\Gamma)} \sum_{\boldsymbol{\theta}, \boldsymbol{\phi}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\phi}) \hat{g}(\boldsymbol{\phi}, \boldsymbol{\xi}) e^{i(\boldsymbol{\theta}+\boldsymbol{\phi})\mathbf{x}},$$

and hence

$$\widehat{(b \circ g)}(\boldsymbol{\chi}, \boldsymbol{\xi}) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \sum_{\boldsymbol{\theta}+\boldsymbol{\phi}=\boldsymbol{\chi}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi} + \boldsymbol{\phi}) \hat{g}(\boldsymbol{\phi}, \boldsymbol{\xi}), \ \boldsymbol{\chi} \in \Gamma^{\dagger}, \ \boldsymbol{\xi} \in \mathbb{R}^d. \tag{2.4.1}$$

Here and below $\boldsymbol{\theta}, \boldsymbol{\phi} \in \Gamma^{\dagger}$. In particular, one sees that $\operatorname{Op}(b)\operatorname{Op}(w^{\delta}) = \operatorname{Op}(bw^{\delta})$ for any $\delta \in \mathbb{R}$. This observation leads to the following Lemma. We remind that the symbol $b_{\boldsymbol{\eta}}$ is defined in (2.1.7).

Lemma 2.4.2. *Let $b \in \mathbf{S}_{\gamma}(w)$ with $w(\boldsymbol{\xi}) = \langle \boldsymbol{\xi} \rangle^{\beta}, \beta \in (0, 1]$. Then for any $u \in \mathsf{S}(\mathbb{R}^d)$ and any $l > d$, we have*

$$\| \operatorname{Op}(b)u\| \le C \,[\![b]\!]^{(\gamma)}_{l,0} \|(H_0 + I)^{\tilde{\gamma}} u\|, \tilde{\gamma} = \frac{\gamma\beta}{2m}, \tag{2.4.2}$$

with a constant C independent of b, u. In particular, if $\gamma\beta < 2m$, then $\operatorname{Op}(b)$ is H_0-bounded with an arbitrarily small relative bound.

Moreover, for any $\boldsymbol{\eta} \in \mathbb{R}^d$ and any $l > d$,

$$\|(\operatorname{Op}(b) - \operatorname{Op}(b_{\boldsymbol{\eta}}))u\| \le C|\boldsymbol{\eta}| \,[\![b]\!]^{(\gamma)}_{l,1} \|(H_0 + I)^{\hat{\gamma}} u\|, \ \hat{\gamma} = \frac{\beta(\alpha - 1)}{2m}, \tag{2.4.3}$$

where the constant C does not depend on b, u, and is uniform in $\boldsymbol{\eta}$: $|\boldsymbol{\eta}| \le \tilde{C}$.

Proof: Define $G = B \operatorname{Op}(w^{-\gamma})$. As we have observed earlier, $G = \operatorname{Op}(g)$ with $g = bw^{-\gamma}$, so that $g \in \mathbf{S}_0(w)$ and $[\![g]\!]^{(0)}_{l,0} = [\![b]\!]^{(\gamma)}_{l,0}$. Hence, by Lemma 2.4.1, $\|G\| \le C [\![b]\!]^{(\gamma)}_{l,0}$ and

$$\| \operatorname{Op}(b)u\| = \|G \operatorname{Op}(w^{\gamma})u\| \le C [\![b]\!]^{(\gamma)}_{l,0} \| \operatorname{Op}(w^{\gamma})u\|. \tag{2.4.4}$$

As $\operatorname{Op}(w^{\gamma}) \le C(H_0 + I)^{\tilde{\gamma}}$, $\tilde{\gamma} = \gamma\beta(2m)^{-1}$, we get (2.4.2).

The bound (2.4.3) follows from (2.4.2) when applied to the symbol $b - b_{\boldsymbol{\eta}}$, and from the estimate (2.1.8).

The bound (2.4.2) allows one to give a proper meaning to the operator (2.1.10), since b is infinitesimally H_0-bounded. The bound (2.4.3) will be useful in the study of the Floquet eigenvalues as functions of the quasi-momentum $\mathbf{k}$.

For general symbols b, g we have the following proposition (see e.g. [38]).

Proposition 2.4.3. *Let $b \in \mathbf{S}_{\gamma_1}$, $g \in \mathbf{S}_{\gamma_2}$. Then $b \circ g \in \mathbf{S}_{\gamma_1+\gamma_2}$ and*

$$\left| b \circ g \right|^{(\gamma_1+\gamma_2)} \le C \left| b \right|^{(\gamma_1)} \left| g \right|^{(\gamma_2)},$$

with a constant C independent of b, g.

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

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

For the sake of convenience we use the notation $\operatorname{ad}(a; \psi_1, \psi_2, \dots, \psi_N)$ and $\operatorname{ad}^N(a, \psi)$ for the symbols of multiple commutators. It follows from (2.4.1) that the Fourier coefficients of the symbol $\operatorname{ad}(b, g)$ are given by

$$\widehat{\operatorname{ad}(b, g)}(\boldsymbol{\chi}, \boldsymbol{\xi}) = \frac{i}{\sqrt{\operatorname{d}(\Gamma)}} \sum_{\boldsymbol{\theta}+\boldsymbol{\phi}=\boldsymbol{\chi}} \big[\hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}+\boldsymbol{\phi})\hat{g}(\boldsymbol{\phi}, \boldsymbol{\xi}) - \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi})\hat{g}(\boldsymbol{\phi}, \boldsymbol{\xi}+\boldsymbol{\theta})\big],$$
$$\boldsymbol{\chi} \in \Gamma^\dagger,\ \boldsymbol{\xi} \in \mathbb{R}^d. \quad (2.4.5)$$

Proposition 2.4.4. *(See e.g. [38]) Let $b \in \mathbf{S}_{\gamma_0}$ and $g_j \in \mathbf{S}_{\gamma_j}$, $j = 1, 2, \dots, N$. Then* $\operatorname{ad}(b; g_1, \dots, g_N) \in \mathbf{S}_\gamma$ *with*

$$\gamma = \sum_{j=0}^{N} (\gamma_j - 1),$$

and

$$\left| \operatorname{ad}(b; g_1, \dots, g_N) \right|^{(\gamma)} \le C \left| b \right|^{(\gamma_0)} \prod_{j=1}^{N} \left| g_j \right|^{(\gamma_j)}, \quad (2.4.6)$$

with a constant C independent of b, g_j.

2.4.2 Partition of the perturbation

From now on the weights in the definition of classes $\mathbf{S}_\gamma = \mathbf{S}_\gamma(w)$ are assumed to be $w(\boldsymbol{\xi}) = \langle \boldsymbol{\xi} \rangle^\beta$ with some $\beta \in (0, 1]$. Here we partition every symbol $b \in \mathbf{S}_\gamma$ into the sum of several symbols, restricted to different parts of the phase space. These symbols depend on the parameter $\rho \geq 1$, but this dependence is usually omitted from the notation.

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

$$0 \leq \iota \leq 1, \quad \iota(z) = \begin{cases} 1, \ z \leq \frac{1}{2A_{max}}; \\ 0, \ z \geq \frac{1}{A_{max}}. \end{cases} \tag{2.4.7}$$

For $L \geq 1$ and $\boldsymbol{\theta} \in \Gamma^\dagger, \boldsymbol{\theta} \neq \mathbf{0}$, define the following C^∞-cut-off functions:

$$\begin{cases} e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = & \iota\left(\left|\frac{|\boldsymbol{\xi}+\boldsymbol{\theta}/2|}{\rho} - 1\right|\right), \\ \ell^{>}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = & 1 - \iota\left(\frac{|\boldsymbol{\xi}+\boldsymbol{\theta}/2|}{\rho} - 1\right), \\ \ell^{<}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = & 1 - \iota\left(1 - \frac{|\boldsymbol{\xi}+\boldsymbol{\theta}/2|}{\rho}\right), \end{cases} \tag{2.4.8}$$

and

$$\begin{cases} \zeta_\theta(\boldsymbol{\xi}; L) = & \iota\left(\frac{|\langle \boldsymbol{\theta}, \hat{\psi}(\boldsymbol{\xi}+\boldsymbol{\theta}/2)\rangle|}{L|\boldsymbol{\theta}|}\right), \\ \varphi_\theta(\boldsymbol{\xi}; L) = & 1 - \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}; L), \end{cases} \tag{2.4.9}$$

where $\hat{\psi}(\boldsymbol{\xi}) = |\boldsymbol{\xi}|\psi(\boldsymbol{\xi})$, and $\psi(\boldsymbol{\xi})$ is a unit normal vector at the point $\boldsymbol{\xi}$ to the surface defined by $\{\varsigma : h_0(\varsigma) = h_0(\boldsymbol{\xi})\}$.

Note that $e_{\boldsymbol{\theta}} + \ell^{>}_{\boldsymbol{\theta}} + \ell^{<}_{\boldsymbol{\theta}} = 1$. The function $\ell^{>}_{\boldsymbol{\theta}}$ is supported on the set $|\boldsymbol{\xi}+\boldsymbol{\theta}/2| > \rho(1+\frac{1}{2A_{max}})$, and $\ell^{<}_{\boldsymbol{\theta}}$ is supported on the set $|\boldsymbol{\xi}+\boldsymbol{\theta}/2| < \rho(1-\frac{1}{2A_{max}})$. The function $e_{\boldsymbol{\theta}}$ is supported in the shell $\rho(1-\frac{1}{A_{max}}) \leq |\boldsymbol{\xi}| \leq \rho(1+\frac{1}{A_{max}})$. Omitting the parameter L and using the notation $\ell_{\boldsymbol{\theta}}$ for any of the functions $\ell^{>}_{\boldsymbol{\theta}}$ or $\ell^{<}_{\boldsymbol{\theta}}$, we point out that

$$\begin{cases} e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = e_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}), \ \ell_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \ell_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}), \\ \varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \varphi_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}), \quad \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \zeta_{-\boldsymbol{\theta}}(\boldsymbol{\xi}+\boldsymbol{\theta}). \end{cases} \tag{2.4.10}$$

Note that there exists constant $C > 0$ (not depending on ρ) such that

$$\left|\frac{\partial \hat{\psi}(\boldsymbol{\xi})}{\partial x_i}\right| \leq C$$

for any vector $\boldsymbol{\xi}$ and component i.

Therefore,

$$|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi};L)|+|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi};L)|\ll L^{-|\mathbf{s}|}C^{|\mathbf{s}|}.$$

Since afterwards we are going to choose $L=\rho^{\beta}$, we can conclude that the above functions satisfy the estimates

$$\begin{cases}|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}e_{\boldsymbol{\theta}}(\boldsymbol{\xi})|+|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\ell_{\boldsymbol{\theta}}(\boldsymbol{\xi})|\ll \rho^{-|\mathbf{s}|},\\ |\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi};L)|+|\mathbf{D}_{\boldsymbol{\xi}}^{\mathbf{s}}\zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi};L)|\ll L^{-|\mathbf{s}|}.\end{cases} \tag{2.4.11}$$

Let

$$\Theta_r=\Theta_r(\Gamma)=\{\boldsymbol{\theta}\in\Gamma^{\dagger}:0<|\boldsymbol{\theta}|\le r\},\ \Theta_r^0=\Theta_r\cup\{\mathbf{0}\}, \tag{2.4.12}$$

with $r=\rho^{\varkappa}$. Using the above cut-off functions, for any symbol $b\in\mathbf{S}_{\gamma}(w)$ we introduce six new symbols $b^{\mathcal{LF}},b^{\mathcal{SE}},b^{o},b^{\mathcal{LE}},b^{\mathcal{NR}},b^{\mathcal{R}}$ in the following way:

$$b^{\mathcal{LF}}(\mathbf{x},\boldsymbol{\xi};\rho)=\frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\sum_{\boldsymbol{\theta}\notin\Theta_r^0}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.4.13}$$

$$b^{\mathcal{LE}}(\mathbf{x},\boldsymbol{\xi};\rho)=\frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\sum_{\boldsymbol{\theta}\in\Theta_r}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\ell_{\boldsymbol{\theta}}^{>}(\boldsymbol{\xi})e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.4.14}$$

$$b^{\mathcal{NR}}(\mathbf{x},\boldsymbol{\xi};\rho)=\frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\sum_{\boldsymbol{\theta}\in\Theta_r}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi};\rho^{\beta})e_{\boldsymbol{\theta}}(\boldsymbol{\xi})e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.4.15}$$

$$b^{\mathcal{R}}(\mathbf{x},\boldsymbol{\xi};\rho)=\frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\sum_{\boldsymbol{\theta}\in\Theta_r}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi};\rho^{\beta})e_{\boldsymbol{\theta}}(\boldsymbol{\xi})e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.4.16}$$

$$b^{\mathcal{SE}}(\mathbf{x},\boldsymbol{\xi};\rho)=\frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\sum_{\boldsymbol{\theta}\in\Theta_r}\hat{b}(\boldsymbol{\theta},\boldsymbol{\xi})\ell_{\boldsymbol{\theta}}^{<}(\boldsymbol{\xi})e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.4.17}$$

$$b^{o}(\mathbf{x},\boldsymbol{\xi};\rho)=b^{o}(\boldsymbol{\xi};\rho)=\frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\hat{b}(0,\boldsymbol{\xi}). \tag{2.4.18}$$

The superscripts here are chosen to mean correspondingly: $\mathcal{LF}$ =‘large Fourier’ (coefficients), $\mathcal{LE}$ = ‘large energy’, $\mathcal{NR}$ = ‘non-resonance’, $\mathcal{R}$ = ‘resonance’, $\mathcal{SE}$ =‘small energy’, o =0-th Fourier coefficient. Sometimes the dependence of the introduced symbols on the parameter ρ is omitted from the notation. The corresponding operators are denoted by

$$B^{\mathcal{LF}}=\mathrm{Op}(b^{\mathcal{LF}}),\ B^{\mathcal{LE}}=\mathrm{Op}(b^{\mathcal{LE}}),\ B^{\mathcal{NR}}=\mathrm{Op}(b^{\mathcal{NR}}),$$

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

By definition (2.4.7),

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

The role of each of these operator is easy to explain. The symbol $b^{\mathcal{LF}}$ contains only Fourier coefficients with $|\boldsymbol{\theta}| > r$, and the remaining symbols contain the Fourier coefficients with $|\boldsymbol{\theta}| \le r$. Note that on the support of the functions $\hat{b}^{\mathcal{NR}}(\boldsymbol{\theta}, \ \cdot\ ; \rho)$ and $\hat{b}^{\mathcal{R}}(\boldsymbol{\theta}, \ \cdot\ ; \rho)$ we have

$$|\boldsymbol{\theta}| \le \rho^{\varkappa}, \ \rho\left(1 - \frac{1}{A_{max}}\right) \le |\boldsymbol{\xi} + \boldsymbol{\theta}/2| \le \rho\left(1 + \frac{1}{A_{max}}\right) \Rightarrow$$

$$\Rightarrow \rho\left(1 - \frac{1}{A_{max}}\right) - \frac{1}{2}\rho^{\varkappa} \le |\boldsymbol{\xi}| \le \rho\left(1 + \frac{1}{A_{max}}\right) + \frac{1}{2}\rho^{\varkappa}. \tag{2.4.19}$$

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

$$\left|\boldsymbol{\xi} + \frac{\boldsymbol{\theta}}{2}\right| \le \rho\left(1 - \frac{1}{2A_{max}}\right), \ |\boldsymbol{\xi}| \le \rho\left(1 - \frac{1}{2A_{max}}\right) + \frac{1}{2}\rho^{\varkappa}. \tag{2.4.20}$$

On the support of $b^{\mathcal{LE}}(\mathbf{x}, \ \cdot\ ; \rho)$ we have

$$\left|\boldsymbol{\xi} + \frac{\boldsymbol{\theta}}{2}\right| \ge \rho\left(1 + \frac{1}{2A_{max}}\right), \ |\boldsymbol{\xi}| \ge \rho\left(1 + \frac{1}{2A_{max}}\right) - \frac{1}{2}\rho^{\varkappa}. \tag{2.4.21}$$

The introduced symbols play a central role in the proof of the Main Theorem 2.1.1. As we show in the course of the proof, due to (2.4.20) and (2.4.21) the symbols $b^{\mathcal{LF}}$, $b^{\mathcal{SE}}$ and $b^{\mathcal{LE}}$ make only a negligible contribution to the spectrum of the operator (2.1.10) near the point $\lambda = \rho^{2m}$. The only significant components of b are the symbols $b^{\mathcal{NR}}, b^{\mathcal{R}}$ and b^o. The symbol $b^{\mathcal{NR}}$ will be transformed in the next Section into another symbol, independent of $\mathbf{x}$.

We will often combine $B^{\mathcal{R}}$, $B^{\mathcal{LE}}$ and $B^{\mathcal{LF}}, B^{\mathcal{SE}}$: for instance $B^{\mathcal{R},\mathcal{LE}} = B^{\mathcal{R}} + B^{\mathcal{LE}}$, $B^{\mathcal{R},\mathcal{LE},\mathcal{LF}} = B^{\mathcal{R},\mathcal{LE}} + B^{\mathcal{LF}}$. A similar convention applies to the symbols. Under the condition $b \in \mathbf{S}_{\gamma}(w)$ the above symbols belong to the same class $\mathbf{S}_{\gamma}(w)$ and the following bounds hold:

$$[\![b^{\mathcal{R}}]\!]^{(\gamma)}_{l,s} + [\![b^{\mathcal{NR}}]\!]^{(\gamma)}_{l,s} + [\![b^{\mathcal{LE}}]\!]^{(\gamma)}_{l,s} + [\![b^{o}]\!]^{(\gamma)}_{l,s} + [\![b^{\mathcal{SE}}]\!]^{(\gamma)}_{l,s} + [\![b^{\mathcal{LF}}]\!]^{(\gamma)}_{l,s} \ll [\![b]\!]^{(\gamma)}_{l,s}. \tag{2.4.22}$$

Indeed, let us check this for the symbol $b^{\mathcal{NR}}$, for instance. According to (2.4.19) and (2.4.11), on the support of the function $\hat{b}^{\mathcal{NR}}(\boldsymbol{\theta}, \ \cdot\ ; \rho)$ we have

$$|\mathbf{D}^{\mathbf{s}}\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}, \rho^{\beta})| \ll \rho^{-\beta|\mathbf{s}|} \ll w^{-|\mathbf{s}|},$$

$$|\mathbf{D}^{\mathbf{s}}\ell^{>}_{\boldsymbol{\theta}}(\boldsymbol{\xi})| + |\mathbf{D}^{\mathbf{s}}\ell^{<}_{\boldsymbol{\theta}}(\boldsymbol{\xi})| + |\mathbf{D}^{\mathbf{s}}e_{\boldsymbol{\theta}}(\boldsymbol{\xi})| \ll \rho^{-|\mathbf{s}|} \ll w^{-|\mathbf{s}|}.$$

This immediately leads to the bound of the form (2.4.22) for the symbol $b^{\mathcal{NR}}$.

The introduced operations also preserve symmetry. Precisely, calculate using (2.4.10):

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

Therefore, by (2.1.9) the operator $B^{\mathcal{R}}$ is symmetric if so is B. 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 Subsection 2.1.3.

Lemma 2.4.5. *Let $b \in \mathbf{S}_{\gamma}(w)$, $w = \langle \boldsymbol{\xi} \rangle^{\beta}$, $\beta \in (0, 1]$ with some $\alpha \in \mathbb{R}$. Then the following hold:*

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

$$\| \operatorname{Op}(b^{\mathcal{SE}}) \| \ll \left[\!\left[b \right]\!\right]^{(\gamma)}_{l,0} \rho^{\beta \max(\gamma, 0)}.$$

Moreover,

$$\left(I - \mathcal{P}\left(B\left(\rho\left(1 - \frac{1}{4A_{max}}\right)\right)\right)\right)\operatorname{Op}(b^{\mathcal{SE}}) =$$

$$= \operatorname{Op}(b^{\mathcal{SE}})\left(I - \mathcal{P}\left(B\left(\rho\left(1 - \frac{1}{4A_{max}}\right)\right)\right)\right) = 0.$$

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

$$\begin{aligned}
&\mathcal{P}\left(B\left(\rho\left(1-\frac{5}{4A_{max}}\right)\right)\right)B^{\mathcal{R}} = B^{\mathcal{R}}\mathcal{P}\left(B\left(\rho\left(1-\frac{5}{4A_{max}}\right)\right)\right) = \\
&= \left(I-\mathcal{P}\left(B\left(\rho\left(1+\frac{5}{4A_{max}}\right)\right)\right)\right)B^{\mathcal{R}} = \\
&= B^{\mathcal{R}}\left(I-\mathcal{P}\left(B\left(\rho\left(1+\frac{5}{4A_{max}}\right)\right)\right)\right) = 0,
\end{aligned} \tag{2.4.23}$$

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

Moreover, for any $\gamma_1 \in \mathbb{R}$ one has $b^{\mathcal{NR}}, b^{\mathcal{R}} \in \mathbf{S}_{\gamma_1}$ and

$$[b^{\mathcal{NR}}]^{(\gamma_1)}_{l,s} + [b^{\mathcal{R}}]^{(\gamma_1)}_{l,s} \ll \rho^{\beta(\gamma-\gamma_1)}[b]^{(\gamma)}_{l,s}, \tag{2.4.24}$$

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

$$\|B^{\mathcal{NR}}\| + \|B^{\mathcal{R}}\| \ll \rho^{\beta\gamma}[b]^{(\gamma)}_{l,0},$$

for any $l > d$.

(iii)

$$\mathcal{P}\left(B\left(\rho\left(1+\frac{1}{4A_{max}}\right)\right)\right)B^{\mathcal{LE}} = B^{\mathcal{LE}}\mathcal{P}\left(B\left(\rho\left(1+\frac{1}{4A_{max}}\right)\right)\right) = 0.$$

(iv) If $R \leq A_{max}\rho$, then

$$\|\mathcal{P}(B(R))B^{\mathcal{LF}}\| + \|B^{\mathcal{LF}}\mathcal{P}(B(R))\| \ll [b]^{(\alpha)}_{l,0} r^{p-l}\rho^{\beta\max(\alpha,0)}, \tag{2.4.25}$$

for any $p > d$ and any $l \geq p$.

In what follows a central role is played by the operator of the form

$$A := H_0 + B^o + B^{\mathcal{R}} \tag{2.4.26}$$

with some symmetric symbol $b \in \mathbf{S}_\gamma$. In the next Theorem we study the continuity of the Floquet eigenvalues $\lambda_j(A(\mathbf{k}))$, $j = 1, 2, \dots$, as functions of the quasi-momentum $\mathbf{k}$. Here, $A(\mathbf{k})$ are the fibers of the operator (2.4.26). For any vector $\boldsymbol{\eta} \in \mathbb{R}^d$ we define the *distance on the torus:*

$$|\boldsymbol{\eta}|_{\mathbb{T}} = \min_{\mathbf{m}\in\Gamma^\dagger} |\boldsymbol{\eta} - \mathbf{m}|. \tag{2.4.27}$$

Theorem 2.4.6. *Suppose that* $\rho \geq 1$ *and*

$$\beta(\gamma - 1) < 2m - 1. \tag{2.4.28}$$

If for some j

$$\lambda_j(A(\mathbf{k})) \asymp \rho^{2m}, \tag{2.4.29}$$

then for any $l > d$ *we have*

$$|\lambda_j(A(\mathbf{k}+\boldsymbol{\eta})) - \lambda_j(A(\mathbf{k}))| \ll \big(1 + [\![b]\!]^{(\gamma)}_{l,1}\big)|\boldsymbol{\eta}|_{\mathbb{T}}\rho^{2m-1}. \tag{2.4.30}$$

The implied constant in (2.4.30) *depends on the constants in* (2.4.29).

Proof: The proof coincides with the proof of Theorem 3.6 in [24].

2.5 A "gauge transformation"

In this and all the subsequent sections we assume that $\mathbf{S}_\gamma = \mathbf{S}_\gamma(w)$ with $w(\boldsymbol{\xi}) = \langle\boldsymbol{\xi}\rangle^\beta$, $\beta \in (0, 1]$. Recall that we study spectral properties of the operator H defined in (2.1.10). Our ultimate goal is to prove that each sufficiently large λ belongs to the spectrum of H. We are going to use the notation from the previous section with the parameter $\rho = \lambda^{\frac{1}{2m}} \geq 1$.

2.5.1 Preparation

Our strategy is to find a unitary operator which reduces $H = H_0 + \mathrm{Op}(b)$ to another PDO, whose symbol, up to some controllable small errors, depends only on $\boldsymbol{\xi}$. The unitary operator sought is constructed in the form $U = e^{iK}$ with a suitable bounded self-adjoint Γ-periodic PDO K. This is why we sometimes call it a "gauge transformation". It is useful to consider e^{iK} as an element of the group

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

We assume that the operator $\mathrm{ad}(H_0, K)$ **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:

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

By induction it is easy to show that

$$A_1 = H + \sum_{j=1}^{M} \frac{1}{j!}\,\mathrm{ad}^j(H;K) + R^{(1)}_{M+1}, \tag{2.5.1}$$

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

The operator K is sought in the form

$$K = \sum_{k=1}^{M} K_k,\ \ K_k = \mathrm{Op}(\kappa_k), \tag{2.5.2}$$

with symbols κ_k from some suitable classes $\mathbf{S}_\sigma, \sigma = \sigma_k$ to be specified later on. Substitute this formula in (2.5.1) and rewrite, regrouping the terms:

$$A_1 = H_0 + B + \sum_{j=1}^{M} \frac{1}{j!} \sum_{l=j}^{M} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(H; K_{k_1}, K_{k_2}, \dots, K_{k_j})$$
$$+R^{(1)}_{M+1} + R^{(2)}_{M+1},$$

$$R^{(2)}_{M+1} := \sum_{j=1}^{M} \frac{1}{j!} \sum_{k_1+k_2+\dots+k_j \ge M+1} \mathrm{ad}(H; K_{k_1}, K_{k_2}, \dots, K_{k_j}). \tag{2.5.3}$$

Changing this expression yet again produces

$$A_1 = H_0 + B + \sum_{l=1}^{M} \mathrm{ad}(H_0; K_l) + \sum_{j=2}^{M} \frac{1}{j!} \sum_{l=j}^{M} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(H_0; K_{k_1}, K_{k_2}, \dots, K_{k_j})$$
$$+ \sum_{j=1}^{M} \frac{1}{j!} \sum_{l=j}^{M} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(B; K_{k_1}, K_{k_2}, \dots, K_{k_j}) + R^{(1)}_{M+1} + R^{(2)}_{M+1}.$$

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

$$A_1 = H_0 + B + \sum_{l=1}^{M} \mathrm{ad}(H_0; K_l) + \sum_{l=2}^{M} \sum_{j=2}^{l} \frac{1}{j!} \sum_{k_1+k_2+\dots+k_j=l} \mathrm{ad}(H_0; K_{k_1}, K_{k_2}, \dots, K_{k_j})$$
$$+ \sum_{l=2}^{M+1} \sum_{j=1}^{l-1} \frac{1}{j!} \sum_{k_1+k_2+\dots+k_j=l-1} \mathrm{ad}(B; K_{k_1}, K_{k_2}, \dots, K_{k_j}) + R^{(1)}_{M+1} + R^{(2)}_{M+1}.$$

Now we introduce the notation

$$B_1 := B,$$

$$B_l := \sum_{j=1}^{l-1} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l-1} \operatorname{ad}(B; K_{k_1}, K_{k_2}, \dots, K_{k_j}),\ l \geq 2, \tag{2.5.4}$$

$$T_l := \sum_{j=2}^{l} \frac{1}{j!} \sum_{k_1+k_2+\cdots+k_j=l} \operatorname{ad}(H_0; K_{k_1}, K_{k_2}, \dots, K_{k_j}),\ l \geq 2. \tag{2.5.5}$$

We emphasize that the operators B_l and T_l depend only on $K_1, K_2, \dots, K_{l-1}$. Let us make one more rearrangement:

$$A_1 = H_0 + B + \sum_{l=1}^{M} \operatorname{ad}(H_0, K_l) + \sum_{l=2}^{M} B_l + \sum_{l=2}^{M} T_l + R_{M+1},$$

$$R_{M+1} = B_{M+1} + R_{M+1}^{(1)} + R_{M+1}^{(2)}. \tag{2.5.6}$$

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

$$\operatorname{ad}(H_0; K_1) + B_1^{\mathcal{NR}} = 0, \tag{2.5.7}$$

$$\operatorname{ad}(H_0; K_l) + B_l^{\mathcal{NR}} + T_l^{\mathcal{NR}} = 0,\ l \geq 2, \tag{2.5.8}$$

and hence

$$\begin{cases} A_1 = A_0 + X_M^{\mathcal{R}} + X_M^{\mathcal{SE},\mathcal{LE},\mathcal{LF}} + R_{M+1}, \\ X_M = \sum_{l=1}^{M} B_l + \sum_{l=2}^{M} T_l, \\ A_0 = H_0 + X_M^{(o)}. \end{cases} \tag{2.5.9}$$

Below we denote by x_M the symbol of the PDO X_M. Recall that by Lemma 2.4.5(ii), the operators $B_l^{\mathcal{NR}}, T_l^{\mathcal{NR}}$ are bounded, and therefore, in view of (2.5.7), (2.5.8), so is the commutator $\operatorname{ad}(H_0; K)$. This justifies the assumption made in the beginning of the formal calculations in this Section.

2.5.2 Commutator equations

Recall that

$$h_0(\boldsymbol{\xi}) = |\boldsymbol{\xi}|^{2m} a\left(\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right)$$

with $m > 0$. Before proceeding to the study of the commutator equations (2.5.7), (2.5.8) note that for $\boldsymbol{\xi}$ in the support of the function $\hat{b}^{\mathcal{NR}}(\boldsymbol{\theta}, \cdot\, ; \rho)$ the symbol

$$\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi}) \tag{2.5.10}$$

can be estimated as follows.

Let us estimate $\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})$ on the support of the $b^{\mathcal{NR}}(x, \boldsymbol{\xi}; \rho)$.

The upper bound follows from the Taylor expansion for the $\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})$ and (2.4.19).

For simplicity we introduce one more notation: $\hat{a}(\boldsymbol{\xi}) = a\left(\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right)$, in terms of this function we can rewrite

$$h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi}) = |\boldsymbol{\xi}|^{2m}\left(\hat{a}(\boldsymbol{\xi} + \boldsymbol{\theta})(|\boldsymbol{\xi} + \boldsymbol{\theta}||\boldsymbol{\xi}|^{-1})^{2m} - \hat{a}(\boldsymbol{\xi})\right) =$$

$$= |\boldsymbol{\xi}|^{2m}\left((\hat{a}(\boldsymbol{\xi} + \boldsymbol{\theta}) - \hat{a}(\boldsymbol{\xi}))(|\boldsymbol{\xi} + \boldsymbol{\theta}||\boldsymbol{\xi}|^{-1})^{2m} + \hat{a}(\boldsymbol{\xi})((|\boldsymbol{\xi} + \boldsymbol{\theta}||\boldsymbol{\xi}|^{-1})^{2m} - 1)\right).$$

Note that $\hat{a}(\boldsymbol{\xi})((|\boldsymbol{\xi} + \boldsymbol{\theta}||\boldsymbol{\xi}|^{-1})^{2m} - 1) \ll 1$. Also, note that there exist a value $B^* > 0$, such that for any $\boldsymbol{\theta}, \boldsymbol{\xi}$ in region defined by (2.4.19)

$$|\hat{a}(\boldsymbol{\xi} + \boldsymbol{\theta}) - \hat{a}(\boldsymbol{\xi})| = \left|\hat{a}\left(\frac{\boldsymbol{\xi} + \boldsymbol{\theta}}{|\boldsymbol{\xi}|}\right) - \hat{a}\left(\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right)\right| \le B^* \frac{|\boldsymbol{\theta}|}{|\boldsymbol{\xi}|}. \tag{2.5.11}$$

Obviously, B^* can be chosen once for all $\rho > \rho_0$ (nothing changes when we simultaneously proportionally increase both vectors $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$, that means that increasing of ρ is equivalent to contracting of the area, that we consider). Then

$$|h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})| \ll |\boldsymbol{\xi}|^{2m}(B^*|\boldsymbol{\theta}||\boldsymbol{\xi}|^{-1} + (|\boldsymbol{\xi} + \boldsymbol{\theta}||\boldsymbol{\xi}|^{-1})^{2m}) \ll |\boldsymbol{\theta}|\rho^{2m-1}.$$

For the lower bound, let us note that on the support of $b^{\mathcal{NR}}(x, \boldsymbol{\xi}; \rho)$,

$$\frac{|\langle \boldsymbol{\theta}, \hat{\psi}(\boldsymbol{\xi} + \boldsymbol{\theta}/2)\rangle|}{\rho^{\beta}|\boldsymbol{\theta}|} > \frac{1}{4} \Rightarrow \frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi} + \boldsymbol{\theta}/2)\rangle|}{|\boldsymbol{\theta}|} \gg \rho^{\beta-1}$$

(because $\hat{\psi}(\boldsymbol{\xi} + \boldsymbol{\theta}/2) \asymp \boldsymbol{\xi}\psi(\boldsymbol{\xi} + \boldsymbol{\theta}/2)$).

By Proposition 2.2.5 if $|\boldsymbol{\theta}| < \rho^{\varkappa}$ then

$$\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) < \rho^{\alpha}|\boldsymbol{\theta}| \implies |\boldsymbol{\theta}\psi(\boldsymbol{\xi})| \ll |\boldsymbol{\theta}|\rho^{\alpha+1-2m}. \tag{2.5.12}$$

Consequently, when considering

$$\alpha + 1 - 2m = \beta - 1 \Leftrightarrow \alpha = 2m - 2 + \beta,$$

we can conclude that

$$\frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle|}{|\boldsymbol{\theta}|} \asymp \frac{|\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi} + \boldsymbol{\theta}/2)\rangle|}{|\boldsymbol{\theta}|} \gg \rho^{\beta-1} \Rightarrow |\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi})\rangle| \gg |\boldsymbol{\theta}|\rho^{\alpha+1-2m},$$

so by (2.5.12)

$$\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi}) \gg \rho^{\alpha}|\boldsymbol{\theta}| = \rho^{2m-2+\beta}|\boldsymbol{\theta}|.$$

Now let us estimate derivatives of $\tau_{\boldsymbol{\theta}}$.

Note that using considerations similar to (2.5.11), for any $\mathbf{s}$ we can obtain

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\hat{a}(\boldsymbol{\xi} + \boldsymbol{\theta}) - \mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\hat{a}(\boldsymbol{\xi})| \leq B_{\mathbf{s}}\frac{|\boldsymbol{\theta}|}{|\boldsymbol{\xi}|}. \tag{2.5.13}$$

Thus, we conclude that

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})| \ll |\boldsymbol{\theta}|\rho^{2m-1-|\mathbf{s}|}.$$

Now for the inverse one can write out:

$$\tau^{-1}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) = \frac{1}{h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})} \Rightarrow$$

for $|\mathbf{s}| = 1$ we obtain

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\tau^{-1}_{\boldsymbol{\theta}}(\boldsymbol{\xi})| = \left|\frac{\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})}{(h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi}))^2}\right| \ll \frac{|\boldsymbol{\theta}|\rho^{2m-2}}{|\boldsymbol{\theta}|^2\rho^{4m-4+2\beta}} = \frac{1}{|\boldsymbol{\theta}|\rho^{2m-2+2\beta}}.$$

Similarly, for any $\mathbf{s}$ we obtain

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\tau^{-1}_{\boldsymbol{\theta}}(\boldsymbol{\xi})| \ll |\boldsymbol{\theta}|^{-1}\rho^{-2m+2-(|\mathbf{s}|+1)\beta} \ll |\boldsymbol{\theta}|^{-1}w^{-(2m-2)\beta^{-1}-1-s}, \tag{2.5.14}$$

for all $\boldsymbol{\xi}$ in the support of the function $\hat{b}^{\mathcal{NR}}(\boldsymbol{\theta}, \ \cdot \ ; \rho)$. This estimate will come in handy in the next lemma.

Lemma 2.5.1. *Let $G = \mathrm{Op}(g)$ be a symmetric PDO with $g \in \mathbf{S}_{\omega}$. Then the PDO K with the Fourier coefficients of the symbol $\kappa(\mathbf{x}, \boldsymbol{\xi}; \rho)$ given by*

$$\begin{cases} \hat{\kappa}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho) = i\dfrac{\hat{g}^{\mathcal{NR}}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho)}{\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})}, \ \boldsymbol{\theta} \neq 0, \\ \hat{\kappa}(\mathbf{0}, \boldsymbol{\xi}; \rho) = 0, \end{cases} \tag{2.5.15}$$

solves the equation

$$\mathrm{ad}(H_0; K) + \mathrm{Op}(g^{\mathcal{NR}}) = 0. \tag{2.5.16}$$

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

$$\|\kappa\|^{(\gamma)}_{l,s} \ll \rho^{\beta(\sigma-\gamma)} \|g\|^{(\omega)}_{l-1,s}, \tag{2.5.17}$$

where

$$\sigma = \omega - (2m-2)\beta^{-1} - 1. \tag{2.5.18}$$

Proof: For brevity we omit ρ from the notation. Let t be the symbol of $\mathrm{ad}(H_0; K)$. The Fourier transform $\hat{t}(\boldsymbol{\theta}, \boldsymbol{\xi})$ is easy to find using (2.4.1):

$$\hat{t}(\boldsymbol{\theta}, \boldsymbol{\xi}) = i\big(h_0(\boldsymbol{\xi} + \boldsymbol{\theta}) - h_0(\boldsymbol{\xi})\big)\hat{\kappa}(\boldsymbol{\theta}, \boldsymbol{\xi}) = i\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})\hat{\kappa}(\boldsymbol{\theta}, \boldsymbol{\xi}).$$

Therefore, by definition (2.4.15), the equation (2.5.16) amounts to

$$i\tau_{\boldsymbol{\theta}}(\boldsymbol{\xi})\hat{\kappa}(\boldsymbol{\theta}, \boldsymbol{\xi}) = -\hat{g}^{\mathcal{NR}}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho) = -\hat{g}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho)\varphi_{\boldsymbol{\theta}}(\boldsymbol{\xi}; \rho^\beta)e_{\boldsymbol{\theta}}(\boldsymbol{\xi}),\ |\boldsymbol{\theta}| \le r.$$

By definition of the functions $\varphi_{\boldsymbol{\theta}}, e_{\boldsymbol{\theta}}$, the function $\hat{\kappa}$ given by (2.5.15) is defined for all $\boldsymbol{\xi}$. Moreover, the symbol $\hat{\kappa}$ satisfies the condition (2.1.9), so that K is a symmetric operator.

In order to prove that $\kappa \in \mathbf{S}_\gamma$ for all $\gamma \in \mathbb{R}$, note that according to (2.4.24) and (2.1.5),

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\hat{g}^{\mathcal{NR}}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho)| \ll \rho^{\beta(\omega-\gamma)} \|g\|^{(\omega)}_{l,s} w^{\gamma-s}|\boldsymbol{\theta}|^{-l}.$$

Together with (2.5.14) this implies that

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}}\hat{\kappa}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho)| \ll \rho^{-\beta\gamma} \|a\|^{(\omega)}_{l,s} w^{\sigma+\gamma-s}|\boldsymbol{\theta}|^{-l-1},$$

so that $\kappa \in \mathbf{S}_\gamma$ and it satisfies (2.5.17).

The estimate (2.5.17) with $\gamma = 0, s = 0$, and Proposition 2.4.1 ensure the boundedness of K.

Let us apply Lemma 2.5.1 to equations (2.5.7) and (2.5.8).

Lemma 2.5.2. *Let $b \in \mathbf{S}_{\gamma^*}$ be a symmetric symbol, $\rho \geq 1$, and let*

$$\begin{cases} \sigma = & \gamma^* - (2m-2)\beta^{-1} - 1, \\ \sigma_j = & j(\sigma - 1) + 1, \\ \epsilon_j = & j(\sigma - 1) + (2m-2)\beta^{-1} + 2, \end{cases} \tag{2.5.19}$$

$j = 1, 2, \ldots$. Then there exists a sequence of self-adjoint bounded PDO's K_j, $j = 1, 2, \ldots$ with the symbols κ_j such that $\kappa_j \in \mathbf{S}_\gamma$ for any $\gamma \in \mathbb{R}$, (2.5.7) *and* (2.5.8) *hold, and*

$$|\kappa_j|^{(\gamma)} \ll \rho^{\beta(\sigma_j - \gamma)} \big(|b|^{(\gamma^*)}\big)^j, \ j \geq 1. \tag{2.5.20}$$

The symbols b_j, t_j of the corresponding operators B_j, T_j belong to $\mathbf{S}_\gamma$ for any $\gamma \in \mathbb{R}$ and

$$|b_j|^{(\gamma)} + |t_j|^{(\gamma)} \ll \rho^{\beta(\epsilon_j - \gamma)} (|b|^{(\gamma^*)})^j, \ j \geq 2. \tag{2.5.21}$$

If $\rho^{\beta(\sigma-1)} |b|^{(\gamma^*)} \ll 1$, then for any M and $\kappa = \sum_{j=1}^M \kappa_j$ the following bounds hold:

$$|\kappa|^{(\gamma)} \ll \rho^{\beta(\sigma-\gamma)} |b|^{(\gamma^*)}, \ \forall \gamma \in \mathbb{R}; \qquad |x_M|^{(\gamma^*)} \ll |b|^{(\gamma^*)}, \tag{2.5.22}$$

$$\|R_{M+1}\| \ll (|b|^{(\gamma^*)})^{M+1} \rho^{\beta \epsilon_{M+1}}; \tag{2.5.23}$$

uniformly in b satisfying $\rho^{\beta(\sigma-1)} |b|^{(\gamma^*)} \ll 1$.

Let us now summarize the results of this section in the following Theorem: the implications of the above Lemma for the operator $H = H_0 + \mathrm{Op}(b)$, defined in (2.1.10).

Theorem 2.5.3. *Let $b \in \mathbf{S}_{\gamma^*}(w)$, $w(\boldsymbol{\xi}) = \langle \boldsymbol{\xi} \rangle^\beta, \beta \in (0, 1]$, $\gamma^* \in \mathbb{R}$ be a symmetric symbol, and let H be the operator defined in* (2.1.10). *Suppose that the condition* (2.1.15) *is satisfied. Then for any positive integer M there exist symmetric symbols $\kappa = \kappa_M$, $x = x_M$, and a self-adjoint bounded operator R_{M+1} satisfying the following properties:*

1. $\kappa \in \mathbf{S}_\gamma$ *for all* $\gamma \in \mathbb{R}$, $x \in \mathbf{S}_{\gamma^*}$, *and*

$$|\kappa|^{(\gamma)} \ll \rho^{\beta(\sigma-\gamma)} |b|^{(\gamma^*)},$$
$$|x|^{(\gamma^*)} \ll |b|^{(\gamma^*)},$$
$$\|R_{M+1}\| \ll (|b|^{(\gamma^*)})^{M+1} \rho^{\beta\epsilon_{M+1}},$$

uniformly in b *satisfying* $|b|^{(\gamma^*)} \ll 1$*;*

2. *The operator* $A_1 = e^{-iK} H e^{iK}, K = \mathrm{Op}(\kappa)$, *has the form*

$$A_1 = A_0 + X^{\mathcal{R}} + X^{\mathcal{SE},\mathcal{LE},\mathcal{LF}} + R_{M+1},\ A_0 = H_0 + X^o. \tag{2.5.24}$$

2.6 Invariant subspaces for the "gauged" operator

In this section we will consider modified sets $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ which are defined as follows:

Definition 2.6.1. Consider a point $\boldsymbol{\xi}$ and the set $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$, defined in the beginning of Section 2.3. Let $\mathfrak{V}$ be the corresponding lattice subspace, then we have following inclusion: $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \boldsymbol{\xi} + \mathfrak{V}$. The point $\boldsymbol{\xi}^*(\boldsymbol{\xi}+\mathfrak{V}) \in \boldsymbol{\xi}+\mathfrak{V}$ is the tangency point to the level set containing $\boldsymbol{\xi}$ (as in Lemma 2.3.5). Denote

$$A_L = \sup\{A : \boldsymbol{\xi} + A\boldsymbol{\xi}^*(\boldsymbol{\xi}+\mathfrak{V}) \in S(\rho)\}$$

(the set S is defined in (2.1.1))and $\boldsymbol{\xi}_L = \boldsymbol{\xi} + A_L\boldsymbol{\xi}^*(\boldsymbol{\xi}+\mathfrak{V})$.

The set $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ is defined as the projection of $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_L)$ on $\boldsymbol{\xi}+\mathfrak{V}$ along the vector $\boldsymbol{\xi}^*(\boldsymbol{\xi}+\mathfrak{V})$.

Let us prove some of the properties of $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$.

Lemma 2.6.2. *1. The set* $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ *is finite.*

2. *If* $|\boldsymbol{\xi}| > \rho/A_{max}$, *then*

$$\max_{\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})} |\boldsymbol{\eta}| \asymp \min_{\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})} |\boldsymbol{\eta}| \asymp |\boldsymbol{\xi}|.$$

3. If $\boldsymbol{\xi}_0 \in \boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$*,* $A \in \mathbb{R}$ *and both points* $\boldsymbol{\xi}_0$ *and* $\boldsymbol{\xi}_0 + A\boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V})$ *belong to the set* $S(\rho)$*, then* $\boldsymbol{\xi}_0 + A\boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V}) \in \boldsymbol{\Upsilon}_L(\boldsymbol{\xi} + A\boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V}))$*.*

4. $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$*.*

Proof:

1. It is sufficient to prove that the set $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ is bounded (since it is a subspace of a lattice). Due to the Proposition 2.3.8 and the Corollary 2.2.10 we can conclude that if $\boldsymbol{\xi}_1 \in \boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$, then

$$|\psi(\boldsymbol{\xi}_1) - \psi(\boldsymbol{\xi})| \leq \mathbf{c}_d(\sin\beta)^{3-2d} \Rightarrow$$
$$\Rightarrow |\boldsymbol{\xi}_1 - \boldsymbol{\xi}| \leq \frac{\mathbf{c}_d}{C_2}(\sin\beta)^{3-2d} = C. \tag{2.6.1}$$

So the set $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ is bounded (all its points are have at most distance C to the point $\boldsymbol{\xi}$), and therefore finite.

2. Immediately follows from (2.6.1).

3. Follows from the fact that $\boldsymbol{\xi}_L = \boldsymbol{\eta}_L$ where $\boldsymbol{\eta} = \boldsymbol{\xi} + A\boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V})$.

4. Follows from the fact that for an arbitrary vector $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$ and a positive number A we have that $\boldsymbol{\xi} + A\boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V}) \in \Xi_2(\mathfrak{V})$ (which is clearly true, since the set $\Xi_2(\mathfrak{V})$ is a convex cone and $\boldsymbol{\xi}^*(\boldsymbol{\xi} + \mathfrak{V}) \in \Xi_2(\mathfrak{V})$). $\square$

The resonant sets $\Xi(\mathfrak{V})$ are designed to describe the invariant subspaces of the periodic PDO's having the form

$$A = H_0 + B^o + B^{\mathcal{R}}, \tag{2.6.2}$$

with the symbols $h_0(\boldsymbol{\xi}) = |\boldsymbol{\xi}|^{2m}$ and $b \in \mathbf{S}_\gamma(w)$, where $\gamma \in \mathbb{R}$, $w(\boldsymbol{\xi}) = \langle\boldsymbol{\xi}\rangle^\beta$, $\beta \in (0,1)$. By (2.1.5) and (2.4.22),

$$|\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} b(\boldsymbol{\theta}, \boldsymbol{\xi})| + |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} \hat{b}^o(\boldsymbol{\theta}, \boldsymbol{\xi})| + |\mathbf{D}^{\mathbf{s}}_{\boldsymbol{\xi}} b^{\mathcal{R}}(\boldsymbol{\theta}, \boldsymbol{\xi})| \ll \lvert b \rvert^{(\gamma)}_{l,s} \langle\boldsymbol{\xi}\rangle^{(\gamma - |\mathbf{s}|)\beta} \langle\boldsymbol{\theta}\rangle^{-l}, \tag{2.6.3}$$

for all $\mathbf{s}$. We always assume that (2.1.15) is satisfied, so that

$$2m > \gamma\beta,\ 2m - 1 > \beta(\gamma - 1), 2m - 2 > \beta(\gamma - 2). \tag{2.6.4}$$

This guarantees that the symbol b and its first two derivatives grow slower than the principal symbol h_0 and its corresponding derivatives respectively.

In order to use the resonant sets $\Xi^*(\mathfrak{V})$ constructed previously, set

$$\alpha_1 = \beta,$$

and assume that the condition (2.3.7) is satisfied. In addition to the symbol (2.4.16), for any lattice subspace $\mathfrak{V} \in \mathcal{V}(n)$, $n = 1, 2, \dots, d$ we define

$$b^{\mathcal{R}}_{\mathfrak{V}}(\mathbf{x}, \boldsymbol{\xi}; \rho) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \sum_{\boldsymbol{\theta} \in \Theta_r \cap \mathfrak{V}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho) \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}; \rho^{\beta}) e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}}. \tag{2.6.5}$$

It is clear that the above symbol retains from $b^{\mathcal{R}}$ only the Fourier coefficients with $\boldsymbol{\theta} \in \mathfrak{V}$. Introduce also the notation for the appropriate reduced version of the model operator (2.6.2):

$$A_{\mathfrak{V}} = H_0 + B^o + B^{\mathcal{R}}_{\mathfrak{V}}, \ B^{\mathcal{R}}_{\mathfrak{V}} = \mathrm{Op}(b^{\mathcal{R}}_{\mathfrak{V}}).$$

Recall (see Subsect. 2.1.3) that for any set $\mathcal{C} \subset \mathbb{R}^d$ we denote by $\mathcal{P}(\mathcal{C})$ the operator $\chi(\mathbf{D}; \mathcal{C})$, where $\chi(\ \cdot \ ; \mathcal{C})$ is the characteristic function of the set $\mathcal{C}$. Accordingly, for the operators in the Floquet decomposition acting on the torus, we define $\mathcal{P}(\mathbf{k}; \mathcal{C})$ to be $\chi(\mathbf{D} + \mathbf{k}; \mathcal{C})$.

In what follows the estimates which we obtain are uniform in the symbol b, satisfying the condition $[\![b]\!]^{(\gamma)} \ll 1$.

Lemma 2.6.3. *Let b be as above. Then for sufficiently large ρ, and any $\mathfrak{V} \in \mathcal{V}(n)$ we have*

$$B^{\mathcal{R}} \mathcal{P}\big(\Xi(\mathfrak{V})\big) = B^{\mathcal{R}}_{\mathfrak{V}} \mathcal{P}\big(\Xi(\mathfrak{V})\big) = \mathcal{P}\big(\Xi(\mathfrak{V})\big) B^{\mathcal{R}}_{\mathfrak{V}} \mathcal{P}\big(\Xi(\mathfrak{V})\big), \tag{2.6.6}$$

and

$$B^{\mathcal{R}}(\mathbf{k}) \mathcal{P}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})) = \mathcal{P}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})) B^{\mathcal{R}}_{\mathfrak{V}}(\mathbf{k}) \mathcal{P}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})), \tag{2.6.7}$$

for any $\mathbf{k} \in \mathcal{O}^{\dagger}$ and any $\boldsymbol{\mu} \in \Xi(\mathfrak{V})$ with $\{\boldsymbol{\mu}\} = \mathbf{k}$.

Proof: Assume without loss of generality that $b^{\mathcal{R}}$ has only one non-zero Fourier coefficient, i.e.

$$b^{\mathcal{R}}(\mathbf{x}, \boldsymbol{\xi}; \rho) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \hat{b}(\boldsymbol{\theta}, \boldsymbol{\xi}; \rho) \zeta_{\boldsymbol{\theta}}(\boldsymbol{\xi}; \rho^{\beta}) e_{\boldsymbol{\theta}}(\boldsymbol{\xi}) e^{i\boldsymbol{\theta}\mathbf{x}}, \tag{2.6.8}$$

so that $b^{\mathcal{R}}_{\mathfrak{V}} = 0$ if $\boldsymbol{\theta} \notin \mathfrak{V}$. According to (2.1.13),

$$(B^{\mathcal{R}}(\mathbf{k})u)(\mathbf{x}) = \frac{1}{\mathrm{d}(\Gamma)} \sum_{\mathbf{m}\in\Gamma^\dagger} \hat{b}^{\mathcal{R}}(\boldsymbol{\theta}, \mathbf{m}+\mathbf{k}) e^{i(\mathbf{m}+\boldsymbol{\theta})\mathbf{x}} \hat{u}(\mathbf{m}),$$

$$\big(B^{\mathcal{R}}(\mathbf{k})\mathcal{P}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu}))u\big)(\mathbf{x}) = \frac{1}{\mathrm{d}(\Gamma)} \sum_{\mathbf{m}:\mathbf{m}+\mathbf{k}\in\boldsymbol{\Upsilon}_L(\boldsymbol{\mu})} \hat{b}^{\mathcal{R}}(\boldsymbol{\theta}, \mathbf{m}+\mathbf{k}) e^{i(\mathbf{m}+\boldsymbol{\theta})\mathbf{x}} \hat{u}(\mathbf{m}), \quad (2.6.9)$$

for any $u \in \mathsf{L}^2(\mathbb{T}^d)$. Observe that by virtue of (2.4.9) for any $\boldsymbol{\xi} := \mathbf{m} + \mathbf{k} \in \operatorname{supp} \zeta_{\boldsymbol{\theta}}(\,\cdot\,; \rho^\beta)$ we have (use Corollary 2.2.10)

$$|\langle \psi(\boldsymbol{\xi}), \boldsymbol{\theta}\rangle| \le |\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}+\boldsymbol{\theta}/2)\rangle| + |\langle \boldsymbol{\theta}, \psi(\boldsymbol{\xi}+\boldsymbol{\theta}/2) - \psi(\boldsymbol{\xi})\rangle| \le$$

$$\le (\rho^{\beta-1}/2 + 2C_2' r/\rho)|\boldsymbol{\theta}| < \rho^{\beta-1}|\boldsymbol{\theta}|,$$

that is $\boldsymbol{\xi} \in \Lambda(\boldsymbol{\theta})$, and a similar calculation shows that $\boldsymbol{\xi} + \boldsymbol{\theta} \in \Lambda(\boldsymbol{\theta})$ as well. By Lemma 2.3.16, $\boldsymbol{\Upsilon}(\boldsymbol{\mu}) \cap \Lambda(\boldsymbol{\theta}) = \varnothing$, if $\boldsymbol{\theta} \notin \mathfrak{V}$, so it follows from (2.6.9) that

$$B^{\mathcal{R}}(\mathbf{k})\mathcal{P}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})) = B^{\mathcal{R}}_{\mathfrak{V}}(\mathbf{k})\mathcal{P}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})) = 0, \quad \text{if } \ \boldsymbol{\theta} \notin \mathfrak{V}.$$

In the case $\boldsymbol{\theta} \in \mathfrak{V}$, by the definition of $\boldsymbol{\Upsilon}$, the points $\boldsymbol{\xi} := \mathbf{m} + \mathbf{k} \in \boldsymbol{\Upsilon}(\boldsymbol{\mu})$ and $\boldsymbol{\xi} + \boldsymbol{\theta}$ belong to $\Lambda(\boldsymbol{\theta})$, so that $\boldsymbol{\xi} + \boldsymbol{\theta} \in \boldsymbol{\Upsilon}(\boldsymbol{\mu})$. This completes the proof of (2.6.7).

Using (2.3.16) we get from (2.6.7):

$$B^{\mathcal{R}}(\mathbf{k})\mathcal{P}(\mathbf{k}; \Xi^*(\mathfrak{V})) = \mathcal{P}(\mathbf{k}; \Xi(\mathfrak{V})) B^{\mathcal{R}}_{\mathfrak{V}}(\mathbf{k})\mathcal{P}(\mathbf{k}; \Xi(\mathfrak{V})).$$

Taking the direct integral in $\mathbf{k}$ yields (2.6.6).

2.6.1 Operator A in the invariant subspaces

Let us denote equivalence relation $\leftrightarrow$ by $\boldsymbol{\xi} \leftrightarrow \boldsymbol{\eta}$ if $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$.

Due to Properties 2.3.2 (i,iii), the formulas (2.6.6) and (2.6.7) imply the following orthogonal decomposition for the Floquet fibres $A(\mathbf{k})$:

$$A(\mathbf{k}) = \bigoplus_{\mathfrak{V}\in\mathcal{V}} A(\mathbf{k}; \Xi(\mathfrak{V})) = \bigoplus_{\mathfrak{V}\in\mathcal{V}} \bigoplus_{\substack{\boldsymbol{\mu}\in\Xi(\mathfrak{V})/\leftrightarrow \\ \{\boldsymbol{\mu}\}=\mathbf{k}}} A_{\mathfrak{V}}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})). \quad (2.6.10)$$

Since card $\boldsymbol{\Upsilon}_L(\boldsymbol{\mu}) < \infty$, see Lemma 2.6.2, for each $\boldsymbol{\mu} \in \Xi^*(\mathfrak{V})$ the operator $A_{\mathfrak{V}}(\mathbf{k}; \boldsymbol{\Upsilon}_L(\boldsymbol{\mu}))$ is finite dimensional (therefore it also holds for the operator $A_{\mathfrak{V}}(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu}))$). Denote $\mathsf{N}(\boldsymbol{\mu}) = \boldsymbol{\Upsilon}(\boldsymbol{\mu}) - \boldsymbol{\mu}$. In the basis

$$E_{[\boldsymbol{\mu}]+\mathbf{m}}(\mathbf{x}),\ \ \mathbf{m} \in \mathsf{N}(\boldsymbol{\mu})$$

(see definition (2.1.17)) of the subspace $\mathfrak{H}\big(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu})\big)$, the operator $A(\mathbf{k}; \boldsymbol{\Upsilon}(\boldsymbol{\mu}))$ reduces to the matrix $\mathcal{A}(\boldsymbol{\mu})$ with the entries

$$\mathcal{A}_{\mathbf{m},\mathbf{n}}(\boldsymbol{\mu}) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\, \hat{a}(\mathbf{m}-\mathbf{n}, \boldsymbol{\mu}+\mathbf{n}; \rho),\ \ \mathbf{m}, \mathbf{n} \in \mathsf{N}(\boldsymbol{\mu}). \tag{2.6.11}$$

(Similarly can be defined a matrix $\mathcal{A}^L$ for the operator $A(\mathbf{k}; \boldsymbol{\Upsilon}_L(\boldsymbol{\mu}))$).

Denote by $\lambda_j(\mathcal{A}^L(\boldsymbol{\mu}))$, $j = 1, 2, \dots, N(\boldsymbol{\mu}) = \operatorname{card} \mathsf{N}(\boldsymbol{\mu})$ the eigenvalues (counting multiplicities) of the matrix $\mathcal{A}^L(\boldsymbol{\mu})$, arranged in descending order. It is easy to check that the matrices $\mathcal{A}^L(\boldsymbol{\mu})$ and $\mathcal{A}^L(\boldsymbol{\mu}')$ with $\boldsymbol{\mu}' \in \boldsymbol{\Upsilon}(\boldsymbol{\mu})$, are unitarily equivalent [1], so that the eigenvalues do not depend on the choice of $\boldsymbol{\mu}$, but only on the set $\boldsymbol{\Upsilon}(\boldsymbol{\mu})$.

Lemma 2.6.4. *Let $\lambda_j(\mathcal{A}^L(\boldsymbol{\mu}))$ be the eigenvalues introduced above. Then for sufficiently large ρ, for all $|\boldsymbol{\mu}| \gg \rho$, and for all $j = 1, 2, \dots, N(\boldsymbol{\mu})$ one has*

$$\lambda_j(\mathcal{A}^L(\boldsymbol{\mu})) \asymp \min_{\boldsymbol{\eta} \in \boldsymbol{\Upsilon}_L(\boldsymbol{\mu})} |\boldsymbol{\eta}|^{2m} \asymp \max_{\boldsymbol{\eta} \in \boldsymbol{\Upsilon}_L(\boldsymbol{\mu})} |\boldsymbol{\eta}|^{2m} \asymp |\boldsymbol{\mu}|^{2m},$$

uniformly in $\boldsymbol{\mu}$.

Proof: The operator A has the form $H_0 + b^{o,\mathcal{R}}$, and since $\alpha\beta < 2m$ (see (2.6.4)), by Lemma 3.2 in [PS], the perturbation $b^{o,\mathcal{R}}$ is infinitesimally H_0-bounded, so that $cH_0 - \tilde{C} \le A \le CH_0 + \tilde{C}$ with some positive constants $C, c, \tilde{C}$. Therefore, the same bounds hold for the fibers $H_0(\mathbf{k})$ and $A(\mathbf{k})$. As a consequence, the restriction of both operators to the subspace $\mathfrak{H}(\mathbf{k}, \boldsymbol{\Upsilon}_L(\boldsymbol{\mu}))$ satisfy the same inequalities:

$$cH_0(\mathbf{k}; \boldsymbol{\Upsilon}_L(\boldsymbol{\mu})) - \tilde{C} \le A(\mathbf{k}; \boldsymbol{\Upsilon}_L(\boldsymbol{\mu})) \le CH_0(\mathbf{k}; \boldsymbol{\Upsilon}_L(\boldsymbol{\mu})) + \tilde{C}.$$

Now the claimed inequalities follow from Lemma 2.6.2.

[1] It is sufficient to note that if $\mathbf{m}' = \mathbf{m} + \boldsymbol{\mu} - \boldsymbol{\mu}'$ and $\mathbf{n}' = \mathbf{n} + \boldsymbol{\mu} - \boldsymbol{\mu}'$, then $\mathbf{m}' - \mathbf{n}' = \mathbf{m} - \mathbf{n}$ and $\boldsymbol{\mu}' + \mathbf{n}' = \boldsymbol{\mu} + \mathbf{n}$.

If $\boldsymbol{\Upsilon}(\boldsymbol{\mu})$ is non-critical[2], the set $\mathsf{N}(\boldsymbol{\mu})$ remains constant in a neighbourhood of $\boldsymbol{\mu}$. Since the entries of the matrix $\mathcal{A}$ depend continuously on $\boldsymbol{\mu}$, we conclude that the eigenvalues $\lambda_j(\mathcal{A}(\ \cdot\ ))$ are continuous in a neighbourhood of such a point $\boldsymbol{\mu}$.

Moreover, by virtue of Lemma 2.6.2, for any $\boldsymbol{\mu} \in \Xi(\mathfrak{V})$ the set $\mathsf{N}(\boldsymbol{\mu})$ remains constant if $\boldsymbol{\mu}_{\not\parallel\mathfrak{V}}$ is kept constant, and hence it makes sense to study the eigenvalues as functions of the component $\boldsymbol{\nu} = \boldsymbol{\mu}_{\parallel\mathfrak{V}}$. Define the matrix

$$\tilde{\mathcal{A}}(t) = \mathcal{A}^L(\boldsymbol{\mu}_{\not\parallel\mathfrak{V}} + t\mathbf{e}(\boldsymbol{\nu})),\ \mathbf{e}(\boldsymbol{\nu}) = \frac{\boldsymbol{\nu}}{|\boldsymbol{\nu}|},$$

with a real-valued parameter $t \geq t_0 := |\boldsymbol{\nu}|$. By (2.6.11) the entries of this matrix are

$$\tilde{\mathcal{A}}_{\mathbf{m},\mathbf{n}}(t) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}}\hat{a}(\mathbf{m}-\mathbf{n}, \boldsymbol{\mu}_{\not\parallel\mathfrak{V}} + \mathbf{n} + t\mathbf{e}(\boldsymbol{\nu}); \rho),\ \ \mathbf{m}, \mathbf{n} \in \mathsf{N}(\boldsymbol{\mu}). \tag{2.6.12}$$

By the definition of $\boldsymbol{\Upsilon}_L$ the matrix $\tilde{\mathcal{A}}$ is well-defined on the interval $[t_0, T_0]$.

Lemma 2.6.5. *Let* (2.6.4) *be satisfied. Suppose that* $\boldsymbol{\mu} \in \Xi^*(\mathfrak{V})$ *and* $|\boldsymbol{\mu}| \asymp \rho$. *Then*

$$\lambda_j(\tilde{\mathcal{A}}(t_2)) - \lambda_j(\tilde{\mathcal{A}}(t_1)) \asymp \rho^{2m-1}(t_2 - t_1), \tag{2.6.13}$$

for any $t_1, t_2 \asymp t_0,\ t_0 \leq t_1 < t_2$, *uniformly in* $j = 1, 2, \dots, \mathsf{N}(\boldsymbol{\mu})$, $\boldsymbol{\mu}$ *and* $\mathfrak{V}$.

Proof: Clearly, $\boldsymbol{\nu} \in S(\rho)$, so that $t_0 = |\boldsymbol{\nu}| \asymp |\boldsymbol{\mu}| \asymp \rho$. By elementary perturbation theory, it would suffice to establish for the matrix

$$\tilde{\mathcal{A}}(t_1, t_2) = \tilde{\mathcal{A}}(t_2) - \tilde{\mathcal{A}}(t_1)$$

the relation

$$(\tilde{\mathcal{A}}(t_1, t_2)u, u) \asymp \rho^{2m-1}(t_2 - t_1)\|u\|^2, \quad t_1, t_2 \asymp \rho, t_2 > t_1 \geq t_0,$$

for all $u \in \mathfrak{H}$. The entries of this matrix are

$$\int_{t_1}^{t_2} \mathcal{Y}_{\mathbf{m},\mathbf{n}}(t)dt, \quad \mathcal{Y}_{\mathbf{m},\mathbf{n}}(t) = \frac{d}{dt}\tilde{\mathcal{A}}_{\mathbf{m},\mathbf{n}}(t).$$

[2] That means that for a small perturbation of the vector $\boldsymbol{\mu}$ by a vector $\boldsymbol{\Delta}$ the set $\boldsymbol{\Upsilon}(\boldsymbol{\mu})$ is moved by the same vector. Clearly, the set of all critical points is a zero measure set.

We show that the matrix $\mathcal{Y}(t)$ satisfies

$$(\mathcal{Y}(t)u, u) \asymp \rho^{2m-1}\|u\|^2, \ t \asymp \rho, \tag{2.6.14}$$

for all $u \in \mathfrak{H}$, uniformly in $\boldsymbol{\mu}$, $\mathfrak{V}$, and the symbol b. Denote

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{\not\parallel\mathfrak{V}} + t\mathbf{e}(\boldsymbol{\nu}), \boldsymbol{\nu} = \boldsymbol{\mu}_{\parallel\mathfrak{V}}.$$

Lemma 2.3.16 implies that $\mathsf{N}(\boldsymbol{\xi}) \in \mathfrak{V}$. Therefore, $\mathcal{Y}(t)$ is the sum of the matrix with diagonal entries

$$\frac{d}{dt} h_0(\boldsymbol{\mu}_t + \mathbf{m}) \asymp |\boldsymbol{\mu}_t + \mathbf{m}|^{2m-2} t, \ \mathbf{m} \in \mathsf{N}(\boldsymbol{\mu}),$$

and the matrix $\mathcal{Z}(t)$ with the entries

$$\mathcal{Z}_{\mathbf{m},\mathbf{n}}(t) = \frac{1}{\sqrt{\mathrm{d}(\Gamma)}} \nabla_{\boldsymbol{\xi}} \hat{b}^{o,\mathcal{R}}(\mathbf{m} - \mathbf{n}, \boldsymbol{\xi}; \rho) \cdot \mathbf{e}(\boldsymbol{\nu})\Big|_{\boldsymbol{\xi}=\boldsymbol{\mu}_t+\mathbf{n}}.$$

By Proposition 2.3.8, $|\boldsymbol{\mu}_t + \mathbf{m}| \ll \rho$, and hence

$$\rho \ll t \le |\boldsymbol{\mu}_t + \mathbf{m}| \ll \rho, \ \mathbf{m} \in \mathsf{N}(\boldsymbol{\mu}). \tag{2.6.15}$$

Thus

$$\frac{d}{dt} h_0(\boldsymbol{\mu}_t + \mathbf{m}) \asymp \rho^{2m-1}. \tag{2.6.16}$$

Also, by (2.6.3),

$$|\mathcal{Z}_{\mathbf{m},\mathbf{n}}(t)| \ll \langle \mathbf{m} - \mathbf{n} \rangle^{-l} |\boldsymbol{\mu}_t + \mathbf{n}|^{(\alpha-1)\beta} \ll \langle \mathbf{m} - \mathbf{n} \rangle^{-l} \rho^{(\alpha-1)\beta},$$

for any $l > 0$. Assuming that $l > d$, from here we get:

$$\begin{aligned} \|\mathcal{Z}(t)\| \le & \max_{\mathbf{n}} \sum_{\mathbf{m}} |\mathcal{Z}_{\mathbf{m},\mathbf{n}}(t)| \\ & \le C_l \rho^{(\alpha-1)\beta} \sup_{\mathbf{n}} \max_{\mathbf{m}} \langle \mathbf{m} - \mathbf{n} \rangle^{-l} \ll C \rho^{(\alpha-1)\beta}. \end{aligned}$$

This, together with (2.6.4) and (2.6.16), leads to (2.6.14), which implies (2.6.13), as required.

2.7 Global description of the eigenvalues of the operator $A(\mathbf{k})$

In this section we continue the study of the discrete spectrum of the fibres $A(\mathbf{k})$. Our aim is to construct a function $g : \mathbb{R}^d \to \mathbb{R}$, which establishes a one-to-one correspondence between the points of $\mathbb{R}^d$ and the eigenvalues of $A(\mathbf{k})$. More precisely, we seek a function g such that

1. for every $\boldsymbol{\xi} \in \mathbb{R}^d$ such that $|\boldsymbol{\xi}| \asymp \rho$ the value $g(\boldsymbol{\xi})$ is an eigenvalue of the operator $A(\mathbf{k})$, $\mathbf{k} = \{\boldsymbol{\xi}\}$, and

2. for every $j \in \mathbb{N}$ such that $\lambda_j(A(\mathbf{k})) \asymp \rho$ there exists a uniquely defined point $\boldsymbol{\xi}$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$ such that $g(\boldsymbol{\xi}) = \lambda_j(A(\mathbf{k}))$.

In other words, we intend to label the eigenvalues of $A(\mathbf{k})$ by the points of the lattice $\Gamma^\dagger$, shifted by $\mathbf{k}$. The construction of the convenient function g is conducted using the decomposition (2.6.10), individually in the invariant subspaces generated by the sets $\Xi^*(\mathfrak{V})$.

Since the sets Ξ^* may be intersected, we introduce an auxiliary function $\hat{g}(\cdot)$ on each of the sets $\Xi^*(\mathfrak{V})$ (we omit the dependance on Ξ^* in notation), and then will consider the possible situation of it being multivalued.

2.7.1 Construction of the function $\hat{g}(\cdot, \cdot)$

We begin with the non-resonant set $\mathcal{B} = \Xi(\mathfrak{X}), \mathfrak{X} = \{0\}$. On the subspace $\mathcal{H}(\Xi(\mathfrak{X}))$ the symbol of the operator A is $\mathbf{x}$-independent, and it is $a^o(\boldsymbol{\xi}) = h_0(\boldsymbol{\xi}) + b^o(\boldsymbol{\xi})$. Therefore the eigenvalues of the operator $A(\mathbf{k})$ are given by $a^o(\boldsymbol{\mu}+\mathbf{k}), \boldsymbol{\mu} \in \Gamma^\dagger, \boldsymbol{\mu}+\mathbf{k} \in \Xi(\mathfrak{X})$. Clearly, it is natural to label the eigenvalues by lattice points. Let us define

$$\hat{g}(\boldsymbol{\xi}, \mathfrak{X}) = a^o(\boldsymbol{\xi}),\ \boldsymbol{\xi} \in \Xi(\mathfrak{X}) = \mathcal{B}.$$

According to (2.6.3),

$$\hat{g}(\boldsymbol{\xi}, \mathfrak{X}) = h_0(\boldsymbol{\xi}) + b^o(\boldsymbol{\xi}), \quad \left|\frac{\partial}{\partial|\boldsymbol{\xi}|} b^o(\boldsymbol{\xi})\right| \le C\langle\boldsymbol{\xi}\rangle^{(\alpha-1)\beta}, \quad \left|\frac{\partial^2}{\partial|\boldsymbol{\xi}|^2} b^o(\boldsymbol{\xi})\right| \le C\langle\boldsymbol{\xi}\rangle^{(\alpha-2)\beta} \tag{2.7.1}$$

for all $\boldsymbol{\xi} \in \mathcal{B}$.

Suppose now that $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ with some non-trivial lattice subspace $\mathfrak{V}$. Let us label all points $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$, $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ in the increasing order of the difference $|\boldsymbol{\eta} - \boldsymbol{\xi}^*(\boldsymbol{\eta} + \mathfrak{V})|$ by natural numbers from the set $\{1, 2, \ldots, N(\boldsymbol{\xi})\}$; if there are two different vectors $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}} \in \boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ with $|\boldsymbol{\eta} - \boldsymbol{\xi}^*(\boldsymbol{\eta} + \mathfrak{V})| = |\tilde{\boldsymbol{\eta}} - \boldsymbol{\xi}^*(\tilde{\boldsymbol{\eta}} + \mathfrak{V})|$, we label them in the lexicographic order of coordinates of point $\boldsymbol{\eta}_{\not\mathfrak{V}}$, $\tilde{\boldsymbol{\eta}}_{\not\mathfrak{V}}$, i.e. we put $\boldsymbol{\eta}_{\not\mathfrak{V}}$ (with $\boldsymbol{\eta}_{\not\mathfrak{V}} - \boldsymbol{\xi}^*(\boldsymbol{\eta} + \mathfrak{V}) = (\eta_1, \eta_2, \ldots, \eta_d)$) before $\tilde{\boldsymbol{\eta}}_{\not\mathfrak{V}}$ (with $\tilde{\boldsymbol{\eta}}_{\not\mathfrak{V}} - \boldsymbol{\xi}^*(\tilde{\boldsymbol{\eta}} + \mathfrak{V}) = (\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_d)$) if either $\eta_1 < \tilde{\eta}_1$, or $\eta_1 = \tilde{\eta}_1$ and $\eta_2 < \tilde{\eta}_2$, etc. Such a labeling associates in a natural way with each point $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ a positive integer $\hat{\ell} = \hat{\ell}(\boldsymbol{\eta}) \leq N(\boldsymbol{\xi})$. Clearly, this number does not depend on the choice of the point $\boldsymbol{\xi}$ as long as $\boldsymbol{\xi}$ remains within the same $\boldsymbol{\Upsilon}_L$ set. In particular,

$$|\boldsymbol{\eta}|^{2m} = \lambda_{\hat{\ell}(\boldsymbol{\eta})}(H_0(\mathbf{k}; \boldsymbol{\Upsilon}_L(\boldsymbol{\eta}))). \tag{2.7.2}$$

Now for every $\boldsymbol{\eta} \in \mathbb{R}^d$ we define

$$\hat{g}(\boldsymbol{\eta}, \mathfrak{V}) := \lambda_{\hat{\ell}(\boldsymbol{\eta})}(\mathcal{A}^L(\boldsymbol{\eta})).$$

Note that in view of Lemma 2.6.4

$$\hat{g}(\boldsymbol{\eta}, \mathfrak{V}) \asymp |\boldsymbol{\eta}|^{2m}, \ |\boldsymbol{\eta}| \gg \rho, \tag{2.7.3}$$

for sufficiently large ρ.

2.7.2 Definition of function $g(\cdot)$

Our goal is to construct a function uniquely matching an eigenvalue to a point. The function $\hat{g}(\cdot, \mathfrak{V})$ depends on the set $\mathfrak{V}$ and can assign two or more eigenvalues to one point, when we change $\mathfrak{V}$. So we introduce one more function $g(\cdot)$.

This function will not possess same nice properties as $\hat{g}(\cdot, \mathfrak{V})$, since $\boldsymbol{\Upsilon}$ does not obey some of properties of $\boldsymbol{\Upsilon}_L$ (i.e. changes other way under small shifts in the direction $\xi^*(\mathfrak{V})$, see Lemma 2.6.2). But we will derive required properties of $g(\cdot)$ from those of $\hat{g}(\cdot, \mathfrak{V})$.

Let us define the function $g(\cdot)$ similarly to the function $\hat{g}(\cdot,\cdot)$, replacing $\boldsymbol{\Upsilon}_L(\boldsymbol{\xi})$ by $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$:

Following previous consideration we label the points $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$, $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$, associating a number $\ell(\boldsymbol{\eta})$ to each point $\boldsymbol{\eta}$ and denote

$$g(\boldsymbol{\eta}) := \lambda_{\ell(\boldsymbol{\eta})}(\mathcal{A}(\boldsymbol{\eta})).$$

Similarly,

$$g(\boldsymbol{\eta}) \asymp |\boldsymbol{\eta}|^{2m},\ |\boldsymbol{\eta}| \gg \rho, \tag{2.7.4}$$

for sufficiently large ρ.

2.7.3 Properties of the function $\hat{g}(\cdot,\cdot)$

In order to analyze the continuity of $\hat{g}(\cdot,\cdot)$, we assume that $\boldsymbol{\eta}$ is a non-critical point, i.e. the set $\mathsf{N}(\ \cdot\ )$ remains constant in a neighbourhood of $\boldsymbol{\eta}$. Furthermore, in the non-critical set, $\ell(\boldsymbol{\eta})$ remains constant, if the point $\boldsymbol{\xi}^*(\mathfrak{V})$ stays away from the Voronoi hyper-planes associated with pairs of points from the set $\boldsymbol{\Upsilon}_L(\boldsymbol{\eta})$. Recall that the Voronoi hyper-plane for a pair $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in \mathbb{R}^d$ is the set of all points $\mathbf{z} \in \mathbb{R}^d$ such that $|\boldsymbol{\eta}_1 - \mathbf{z}| = |\boldsymbol{\eta}_2 - \mathbf{z}|$. Thus, the function $\hat{g}(\cdot,\cdot)$ is continuous on an open set of full measure in $\mathbb{R}^d$.

In each set $\Xi^*(\mathfrak{V})$ the labeling function ℓ possesses the following important property.

Lemma 2.7.1. *Let $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}} \in \Xi^*(\mathfrak{V})$ satisfy $\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta} := \boldsymbol{\nu}$ is proportional to $\boldsymbol{\xi}^*(\boldsymbol{\eta} + \mathfrak{V})$. Then $\ell(\tilde{\boldsymbol{\eta}}) = \ell(\boldsymbol{\eta})$.*

Proof: Recall that by Lemma 2.6.2 $\boldsymbol{\Upsilon}_L(\boldsymbol{\eta}) + \boldsymbol{\nu} = \boldsymbol{\Upsilon}_L(\tilde{\boldsymbol{\eta}})$.

Let us note that, if we move the point $\boldsymbol{\eta}$ by a vector $\boldsymbol{\nu} \parallel \boldsymbol{\xi}^*(\boldsymbol{\eta}+\mathfrak{V})$, the set $\boldsymbol{\Upsilon}_L(\boldsymbol{\eta})$ and the point $\boldsymbol{\xi}^*(\boldsymbol{\eta} + \mathfrak{V})$ are also moved by the vector $\boldsymbol{\nu}$. Therefore, the distance we are interested in does not change:

$$|\tilde{\boldsymbol{\eta}} - \boldsymbol{\xi}^*(\tilde{\boldsymbol{\eta}} + \mathfrak{V})| = |\boldsymbol{\eta} - \boldsymbol{\xi}^*(\boldsymbol{\eta} + \mathfrak{V})|.$$

Similar equality holds for any point in $\boldsymbol{\Upsilon}_L(\boldsymbol{\eta})$. Thus, when arranging points of the set $\boldsymbol{\Upsilon}_L(\tilde{\boldsymbol{\eta}})$ by the distance from $\boldsymbol{\xi}^*(\tilde{\boldsymbol{\eta}}+\mathfrak{V})$, the point $\tilde{\boldsymbol{\eta}}$ will obtain the same number as the point $\boldsymbol{\eta}$ when arranging the points of the set $\boldsymbol{\Upsilon}_L(\boldsymbol{\eta})$ according to their distance to the point $\boldsymbol{\xi}^*(\boldsymbol{\eta}+\mathfrak{V})$.

Also note that, similarly, the order would not change when we have several points from $\boldsymbol{\Upsilon}_L(\boldsymbol{\eta})$ equidistant from the point $\boldsymbol{\xi}^*(\boldsymbol{\eta}+\mathfrak{V})$, since for any point $\boldsymbol{\mu} \in \boldsymbol{\Upsilon}_L(\boldsymbol{\eta})$ we have $\boldsymbol{\mu}_{\not\parallel\mathfrak{V}} = \tilde{\boldsymbol{\mu}}_{\not\parallel\mathfrak{V}}$, with $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu}+\boldsymbol{\nu}$. Recall that we have defined $\boldsymbol{\mu}_{\not\parallel\mathfrak{V}}$ in (2.3.15).

Therefore, $\ell(\boldsymbol{\eta}) = \ell(\boldsymbol{\eta}+\boldsymbol{\nu})$.

The next lemma allows us to establish smoothness of the function $\hat{g}$ with respect to the variable $\boldsymbol{\eta}_{\parallel\mathfrak{V}}$.

Lemma 2.7.2. *Let $\mathfrak{V} \in \mathcal{V}(n), 1 \le n \le d-1$ and let (2.6.4) be satisfied. Suppose that $\boldsymbol{\eta} \in \Xi^*(\mathfrak{V})$ and $|\boldsymbol{\eta}| \asymp \rho$, and let $\boldsymbol{\nu} = \boldsymbol{\eta}_{\parallel\mathfrak{V}}$. Then for sufficiently large ρ, on the interval $[t_0, T_0)$, $t_0 := |\boldsymbol{\nu}|$, $T_0 : \frac{T_0}{t_0}\boldsymbol{\eta} \in S(\rho)$, the function $\tilde{g}(t;\boldsymbol{\eta}) := \hat{g}(\boldsymbol{\eta}_1(\mathfrak{V}) + t\mathbf{e}(\boldsymbol{\nu}), \mathfrak{V})$ satisfies*

$$\tilde{g}(t_2,\boldsymbol{\eta}) - \tilde{g}(t_1,\boldsymbol{\eta}) \asymp \rho^{2m-1}(t_2-t_1), \tag{2.7.5}$$

for any $t_1, t_2 \in [t_0, T_0]$ such that $t_1 < t_2$ and $t_1, t_2 \asymp t_0$, uniformly in $\boldsymbol{\eta} \in \Xi^*(\mathfrak{V})$ and $\mathfrak{V}$.

Proof: Let us remind that by construction of the set $\Xi^*(\mathfrak{V})$ we have $\boldsymbol{\eta}_1(\mathfrak{V}) + t\mathbf{e}(\boldsymbol{\nu}) \in \Xi(\mathfrak{V})$ for all $t \in [t_0, T_0]$. Thus, by Lemma 2.7.1, $\ell(\boldsymbol{\eta}) = \ell(\boldsymbol{\eta}_{\not\parallel\mathfrak{V}} + t\mathbf{e}(\boldsymbol{\nu})) =: \ell$ for $t \in [t_0, T_0)$. Therefore,

$$\tilde{g}(t;\boldsymbol{\eta}) = \lambda_\ell(\mathcal{A}^L(\boldsymbol{\eta}_{\not\parallel\mathfrak{V}} + t\mathbf{e}(\boldsymbol{\nu}))).$$

It remains to apply Lemma 2.6.5.

For the following Lemma recall that the distance on the torus $|\cdot|_{\mathbb{T}}$ is defined in (2.4.27).

Lemma 2.7.3. *Let $\mathfrak{V} \in \mathcal{V}(n), 1 \le n \le d-1$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ be such that $|\mathbf{a}| \asymp \rho$. Then there exists a vector $\mathbf{n} \in \Gamma^\dagger$ such that*

$$|\hat{g}(\mathbf{b}+\mathbf{n},\mathfrak{V}) - \hat{g}(\mathbf{a},\mathfrak{V})| \ll \rho^{2m-1}|\mathbf{b}-\mathbf{a}|_{\mathbb{T}}, \tag{2.7.6}$$

for sufficiently large ρ.

Suppose in addition, that $\mathbf{m} \in \Gamma^\dagger, \mathbf{m} \neq \mathbf{0}$, is a vector such that $|\mathbf{a}+\mathbf{m}| \asymp \rho$. Then there exists a $\tilde{\mathbf{n}} \in \Gamma^\dagger$, such that $\mathbf{n} \neq \tilde{\mathbf{n}}$ and

$$|\hat{g}(\mathbf{b}+\tilde{\mathbf{n}}) - \hat{g}(\mathbf{a}+\mathbf{m})| \ll \rho^{2m-1}|\mathbf{b}-\mathbf{a}|_{\mathbb{T}}, \tag{2.7.7}$$

for sufficiently large ρ.

Proof: As $\lambda_{J(\mathbf{a})}(A(\mathbf{k})) = \hat{g}(\mathbf{a}, \mathfrak{V})$, by (2.7.3), we have

$$|\lambda_{J(\mathbf{a})}(A(\mathbf{k}))| \asymp \rho^{2m}.$$

Denote $\mathbf{k} = \{\mathbf{a}\}$, $\mathbf{k}_1 = \{\mathbf{b}\}$. Recall that the condition (2.4.28) is satisfied due to (2.6.4), so by Theorem 2.4.6

$$|\lambda_{J(\mathbf{a})}(A(\mathbf{k})) - \lambda_{J(\mathbf{a})}(A(\mathbf{k}_1))| \ll \rho^{2m-1}|\mathbf{k}-\mathbf{k}_1|_{\mathbb{T}} = \rho^{2m-1}|\mathbf{b}-\mathbf{a}|_{\mathbb{T}}. \tag{2.7.8}$$

Let $\mathbf{p} \in \mathbb{R}^d$ be a vector such that $\{\mathbf{p}\} = \mathbf{k}_1$ and $\hat{g}(\mathbf{p}, \mathfrak{V}) = \lambda_{J(\mathbf{a})}(A(\mathbf{k}_1))$. Now (2.7.8) implies (2.7.6) with $\mathbf{n} = \mathbf{p} - \mathbf{b}$.

In order to prove (2.7.7), we use (2.7.8) with $\mathbf{a}+\mathbf{m}$ instead of $\mathbf{a}$. Then, as above, one can find a vector $\tilde{\mathbf{p}}$ such that $\{\tilde{\mathbf{p}}\} = \mathbf{k}_1$ and $\hat{g}(\tilde{\mathbf{p}}, \mathfrak{V}) = \lambda_{J(\mathbf{a}+\mathbf{m})}(A(\mathbf{k}_1))$. Since J is one-to-one, we have $J(\mathbf{a}+\mathbf{m}) \neq J(\mathbf{a})$, and hence $\mathbf{p} \neq \tilde{\mathbf{p}}$. As a consequence, $\tilde{\mathbf{n}} = \tilde{\mathbf{p}} - \mathbf{b} \neq \mathbf{n}$, as required, and (2.7.8) again leads to (2.7.7).

2.8 Estimates of volumes

In this section we continue the investigation of the operator of the form (2.6.2), with a symbol $b \in \mathbf{S}_\gamma(w)$, $w(\boldsymbol{\xi}) = \langle\boldsymbol{\xi}\rangle^\beta$, with parameters γ, β, satisfying the conditions (2.6.4). Let $\hat{g} : \mathbb{R}^d \to \mathbb{R}$ be the function defined in the previous section, and let $\mathcal{B}(\rho), \mathcal{D}(\rho)$ and $\tilde{\mathcal{B}}(\rho)$ be the sets introduced in Section 2.3 respectively.

Let $\delta \in (0, \lambda/4]$, $\lambda = \rho^{2m}$, and let

$$\begin{cases} \mathcal{A}_L(\rho,\delta) = \mathcal{A}_L(\hat{g};\rho,\delta) := \bigcup_{\mathfrak{V}} \hat{g}(\cdot,\mathfrak{V})^{-1}([\lambda-\delta,\lambda+\delta]), \\ \mathcal{B}_L(\rho,\delta) = \mathcal{B}_L(\hat{g};\rho,\delta) := \mathcal{A}_L(\rho,\delta) \cap \mathcal{B}_L(\rho), \\ \mathcal{D}_L(\rho,\delta) = \mathcal{D}_L(\hat{g};\rho,\delta) := \mathcal{A}_L(\rho,\delta) \cap \mathcal{D}_L(\rho), \\ \tilde{\mathcal{B}}_L(\rho,\delta) = \tilde{\mathcal{B}}_L(\hat{g};\rho,\delta) := \mathcal{A}_L(\rho,\delta) \cap \tilde{\mathcal{B}}_L(\rho). \end{cases} \tag{2.8.1}$$

Also

$$\begin{cases} \mathcal{A}(\rho,\delta) = \mathcal{A}(g;\rho,\delta) := g^{-1}([\lambda-\delta,\lambda+\delta]), \\ \mathcal{B}(\rho,\delta) = \mathcal{B}(g;\rho,\delta) := \mathcal{A}(\rho,\delta) \cap \mathcal{B}(\rho), \\ \mathcal{D}(\rho,\delta) = \mathcal{D}(g;\rho,\delta) := \mathcal{A}(\rho,\delta) \cap \mathcal{D}(\rho), \\ \tilde{\mathcal{B}}(\rho,\delta) = \tilde{\mathcal{B}}(g;\rho,\delta) := \mathcal{A}(\rho,\delta) \cap \tilde{\mathcal{B}}(\rho). \end{cases} \tag{2.8.2}$$

The estimates for the volumes of the above sets are very important for our argument.

First, let us prove an auxiliary statement.

Lemma 2.8.1. *For an arbitrary* $\mathfrak{V} : \dim \mathfrak{V} \geq 1$ *it holds that*

$$\mathrm{vol}(g^{-1}([\lambda-\delta,\lambda+\delta]) \cap \mathfrak{V}) \leq \mathrm{vol}(\hat{g}(\cdot,\mathfrak{V})^{-1}([\lambda-\delta,\lambda+\delta])).$$

Proof:

Denote $I := [\lambda-\delta,\lambda+\delta]$. We are going to show that for any $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ there exists a unique vector $\boldsymbol{\eta} \in \Xi^*(\mathfrak{V})$ such that

$$\begin{cases} g(\boldsymbol{\xi}) = \hat{g}(\boldsymbol{\eta},\mathfrak{V}) \\ \boldsymbol{\xi} - \boldsymbol{\eta} \in \mathbb{Z}^d \end{cases} \tag{2.8.3}$$

Indeed, there exists i, such that $g(\boldsymbol{\xi}) = \lambda_i(\mathcal{A}(\boldsymbol{\xi}))$. Note that $\mathcal{A}(\boldsymbol{\xi})$ is unitarily equivalent to PHP where P is the projection onto an invariant subspace of H. Therefore $\mathcal{A}(\boldsymbol{\xi})$ is unitarily equivalent to $P\mathcal{A}^L(\boldsymbol{\xi})P$. Thus, there exists some number l and $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$, such that $\lambda_i(\mathcal{A}(\boldsymbol{\xi})) = \lambda_l(\mathcal{A}^L(\boldsymbol{\eta}))$ and (2.8.3) holds.

Denote for each $\boldsymbol{\xi}$: $\gamma(\boldsymbol{\xi}) = \boldsymbol{\xi} - \boldsymbol{\eta}(\boldsymbol{\xi})$, where $\boldsymbol{\eta}(\boldsymbol{\xi})$ - is the vector from the relation (2.8.3).

Therefore,

$$\mathrm{vol}(g^{-1}(I)\cap\mathfrak{V}) = \sum_{\boldsymbol{\mu}\in\mathbb{Z}^d} \mathrm{vol}\{\boldsymbol{\xi}\in g^{-1}(I) : \gamma(\boldsymbol{\xi}) = \boldsymbol{\mu}\} =$$

$$= \sum_{\boldsymbol{\mu}\in\mathbb{Z}^d} \mathrm{vol}\{\boldsymbol{\eta}(\boldsymbol{\xi})\in \hat{g}(\cdot,\mathfrak{V})^{-1}(I) : \gamma(\boldsymbol{\xi}) = \boldsymbol{\mu}\} \le \mathrm{vol}(g(\cdot,\mathfrak{V})^{-1}(I))$$

Corollary 2.8.2. *It holds that*

$$\mathrm{vol}(\mathcal{D}(\rho,\delta)) \le \mathrm{vol}(\mathcal{D}_L(\rho,\delta)).$$

Lemma 2.8.3. *Let A be the operator* (2.6.2)*, and let α,β satisfy the conditions* (2.6.4)*. Then for all sufficiently large ρ and any $\delta\in(0,\rho^{2m}/4]$ the following estimates hold*

$$\mathrm{vol}\,\tilde{\mathcal{B}}_L(\rho,\delta) \asymp \delta\rho^{d-2m}, \tag{2.8.4}$$

and

$$\mathrm{vol}(\mathcal{D}_L(\rho,\delta)) \ll \delta\rho^{d-1-2m+\alpha_d}. \tag{2.8.5}$$

Here $\alpha_d\in(0,1)$ is defined in (2.3.8) above.

Before proving this lemma we find a convenient representation of the set $\tilde{\mathcal{B}}_L(\rho,\delta)$. Since $\tilde{\mathcal{B}}_L(\rho,\delta)\subset\mathcal{B}_L(\rho)$, for all $\boldsymbol{\xi}\in\tilde{\mathcal{B}}_L(\rho,\delta)$ the function $\hat{g}$ is defined by the formula (2.7.1), and in particular, it is continuous. For all $\boldsymbol{\Omega}\in T(\rho)$ (see (2.3.18) for definition), we introduce the subsets of the real line defined as follows:

$$I(\boldsymbol{\Omega};\rho,\delta) = \{t>0 : \rho^{2m}-\delta \le \hat{g}(t\boldsymbol{\Omega},\mathfrak{X}) \le \rho^{2m}+\delta\}. \tag{2.8.6}$$

By (2.7.1), for $t\in I(\boldsymbol{\Omega};\rho,\delta)$ we have $\rho/A_{max} < t < A_{max}\rho$, and hence $t\boldsymbol{\Omega}\in\tilde{\mathcal{B}}_L(\rho,\delta)$. If ρ is sufficiently large, by virtue of (2.7.1), for these values of t the function $\hat{g}(t\boldsymbol{\Omega},\mathfrak{X})$ is strictly increasing, and hence $I(\boldsymbol{\Omega};\rho,\delta)$ is a closed interval. Moreover, (2.7.1) implies the relation

$$|I(\boldsymbol{\Omega};\rho,\delta)| \asymp \delta\rho^{1-2m} \tag{2.8.7}$$

for its length, uniformly in $\boldsymbol{\Omega}$. By construction,

$$\tilde{\mathcal{B}}_L(\rho,\delta) = \bigcup_{\boldsymbol{\Omega}\in T(\rho)} I(\boldsymbol{\Omega};\rho,\delta)\boldsymbol{\Omega}. \tag{2.8.8}$$

Proof:[Proof of Lemma 2.8.3] In view of Lemma 2.3.20 and of the bound (2.8.7), we obtain from (2.8.8):

$$\operatorname{vol}\tilde{\mathcal{B}}_L(\rho,\delta) = \int_{T(\rho)} \int_{I(\boldsymbol{\Omega};\rho,\delta)} t^{d-1} dt d\boldsymbol{\Omega} \asymp \delta\rho^{d-2m}.$$

This proves (2.8.4).

Proof of (2.8.5). By definition (2.3.13) and relation (2.3.12),

$$\mathcal{D}(\rho) = \bigcup_{\mathfrak{V}\subset\mathcal{V}(n),1\le n\le d} \Xi(\mathfrak{V};\rho) \subset \bigcup_{\mathfrak{V}\subset\mathcal{V}(n),1\le n\le d} \Xi^*(\mathfrak{V};\rho).$$

Let us estimate the volume of each intersection $\Xi^*(\mathfrak{V};\rho)\cap\mathcal{A}_L(\rho,\delta)$. Clearly, $\Xi^*(\mathbb{R}^d)\cap\mathcal{A}_L(\rho,\delta)=\varnothing$, thus we assume that $n\le d-1$.

Consider a vector $\boldsymbol{\chi}\in\mathfrak{V}$ and a vector $\boldsymbol{\Omega}$, $|\boldsymbol{\Omega}|=1$, collinear with $\boldsymbol{\xi}^*(\mathfrak{V})$. Denote

$$S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho) = \{t\ge 0 : \boldsymbol{\chi}+t\boldsymbol{\Omega}\in\Xi^*(\mathfrak{V};\rho)\}.$$

Due to Corollary 2.3.9 it holds that if $S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho)\ne\varnothing$, then $|\boldsymbol{\chi}|<2\rho^{\alpha_{d-1}}$. Consider the subset

$$S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho,\delta) = \{t\in S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho) : \rho^{2m}-\delta \le \hat{g}(\boldsymbol{\chi}+t\boldsymbol{\Omega},\mathfrak{V}) \le \rho^{2m}+\delta\}. \qquad (2.8.9)$$

In view of (2.7.3), $t\asymp\rho$. By (2.7.5), the function $\tilde{g}(t)=\hat{g}(\boldsymbol{\chi}+t\boldsymbol{\Omega},\mathfrak{V})$ is strictly increasing and continuous, and hence, $S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho,\delta)$ is an interval. The bound (2.7.5) also guarantees the upper bound

$$|S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho,\delta)| \ll \delta\rho^{1-2m},$$

for the length of this interval, uniformly in $\boldsymbol{\chi}$ and $\boldsymbol{\Omega}$. Now we can estimate the volume of the intersection:

$$\begin{aligned}\operatorname{vol}(\Xi(\mathfrak{V};\rho)\cap\mathcal{A}_L(\rho,\delta)) &= \int\limits_{|\boldsymbol{\chi}|<2\rho^{\alpha_{d-1}}} \int\limits_{\mathbb{S}^{d-n-1}} \int\limits_{S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho,\delta)} t^{d-n-1} dt d\boldsymbol{\Omega}\, d\boldsymbol{\chi} \\ &\ll \rho^{d-n-1} \int\limits_{|\boldsymbol{\chi}|<2\rho^{\alpha_{d-1}}} \int\limits_{\mathbb{S}^{d-n-1}} |S(\boldsymbol{\chi},\boldsymbol{\Omega};\rho,\delta)| d\boldsymbol{\Omega}\, d\boldsymbol{\chi} \\ &\ll \delta\rho^{1-2m}\rho^{d-n-1}(\rho^{\alpha_{d-1}})^n \ll \delta\rho^{d-1-2m+\alpha_{d-1}}.\end{aligned}$$

To prove the last step it is enough to show that

$$-n + \alpha_{d-1} n < -1 + \alpha_{d-1} \Leftarrow$$

$$\Leftrightarrow (-1 + \alpha_{d-1}) n < -1 + \alpha_{d-1}.$$

Recall that the number of distinct subspaces $\mathfrak{V} \subset \mathcal{V}$ does not exceed Cr^{d^2} with some universal constant C, so that

$$\operatorname{vol} \mathcal{D}_L(\rho, \delta) \ll \delta \rho^{d-1-2m+\alpha_{d-1}} r^{d^2} \ll \delta \rho^{d-1-2m+\alpha_d},$$

where we have used the conditions (2.3.8). Now (2.8.11) is proved.

Applying the fact that on non-resonance sets the values of functions g and $\hat{g}$ coincide and Corollary 2.8.2, we obtain

Lemma 2.8.4. *Let A be the operator* (2.6.2)*, and let α, β satisfy the conditions* (2.6.4)*. Then for any $\delta \in (0, \rho^{2m}/4]$ and for all sufficiently large ρ, the following estimates hold*

$$\operatorname{vol} \tilde{\mathcal{B}}(\rho, \delta) \asymp \delta \rho^{d-2m}, \tag{2.8.10}$$

and

$$\operatorname{vol}(\mathcal{D}(\rho, \delta)) \ll \delta \rho^{d-1-2m+\alpha_d}. \tag{2.8.11}$$

Here, $\alpha_d \in (0,1)$ is the number defined together with $\alpha_1, \alpha_2, \ldots, \alpha_{d-1}$ above.

The next section is devoted to the most important statements about the volumes of our sets.

2.9 More subtle volume estimates

We aim to construct an upper bound for the volume of an intersection set:

$$\mathcal{B}(\rho, \delta) \cap (\mathcal{B}(\rho, \delta) + \mathbf{b}).$$

First we need to introduce extra notation.

Denote

$$b(\boldsymbol{\xi}') = (a(\boldsymbol{\xi}'))^{\frac{1}{m}} \tag{2.9.1}$$

and

$$h_1(\boldsymbol{\xi}) = b(\boldsymbol{\xi}')|\boldsymbol{\xi}|^2 = (h_0(\boldsymbol{\xi}))^{\frac{1}{m}}. \tag{2.9.2}$$

We consider the function b defined on the unit sphere only, and the function h_1 - in the whole of $\mathbb{R}^d$.

Let $\mathbf{b}$ be a vector such that

$$1 \ll |\mathbf{b}| \ll \rho. \tag{2.9.3}$$

Due to the definition of the function g (see Subsection 2.7.2), in the non-resonant set $\mathcal{B}(\rho)$ it can be expressed in a following form, with some nice properties on G:

$$g(\boldsymbol{\xi}) = \hat{g}(\boldsymbol{\xi}, \mathfrak{X}) = h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi}).$$

Since g satisfies properties (2.7.1) and (2.6.4) (to be more precise, the property (2.7.1) is true for $\hat{g}$, but on the non-resonant set the functions g and $\hat{g}$ coincide), we can conclude that the following properties hold for G:

$$|G(\boldsymbol{\xi})| \ll \rho^{\kappa}, \text{ where } \kappa < 2m, \tag{2.9.4}$$

$$|\nabla G(\boldsymbol{\xi})| \ll \rho^{\omega}, \text{ where } \omega < 2m-1, \tag{2.9.5}$$

$$|H(G(\boldsymbol{\xi}))| \ll \rho^{\gamma}, \text{ where } \gamma < 2m-2 \text{ and } H \text{ is the Hessian}, \tag{2.9.6}$$

(i.e, all second derivatives of G are $O(\rho^{\gamma})$).

Remark. Note that (2.9.5) implies the formula

$$|G(\boldsymbol{\xi} + \mathbf{b}) - G(\boldsymbol{\xi})| \ll |\mathbf{b}|\rho^{\sigma}, \text{ where } \sigma < 2m-1. \tag{2.9.7}$$

Also, we will be assuming that the values of ρ and δ are constrained by the following relation:

$$\delta < \rho^{\varkappa},\ \varkappa < 2m - 1. \tag{2.9.8}$$

Consider a set $\mathcal{A}_*(\rho, \delta) = \{\boldsymbol{\xi} \in \Xi(\mathfrak{X}) : |h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi}) - \rho^{2m}| < \delta\}$ and a set $\mathcal{A}_{\mathbf{b}}(\rho, \delta)$, denoted as follows:

$$\mathcal{A}_{\mathbf{b}}(\rho, \delta) = \{\boldsymbol{\xi} \in \mathcal{A}_*(\rho, \delta) : \boldsymbol{\xi} + \mathbf{b} \in \mathcal{A}_*(\rho, \delta)\}.$$

Obviously,

$$\mathcal{B}(\rho, \delta) \subset \mathcal{A}_*(\rho, \delta), \tag{2.9.9}$$

$$\mathcal{B}(\rho, \delta) \cap (\mathcal{B}(\rho, \delta) + \mathbf{b}) \subset \mathcal{A}_{\mathbf{b}}(\rho, \delta). \tag{2.9.10}$$

Our aim is to estimate the volume of the set $\mathcal{A}_{\mathbf{b}}(\rho, \delta)$ for "sufficiently large ρ" and arbitrary δ, satisfying the condition (2.9.8).

First we will consider the case $d = 2$, and then derive the general result from that. Consider a plane with orthonormal basis x_1, x_2. Without loss of generality we can assume that x_1 is collinear to $\mathbf{b}$. We emphasize that we will be estimating only the part of the set $\mathcal{A}_{\mathbf{b}}(\rho, \delta)$ which belongs to the upper halfplane. A similar estimate can be done for the lower "half" of the set.

First let us prove several auxiliary facts.

Lemma 2.9.1. If $|H(G(\boldsymbol{\xi}))| \ll \rho^{\gamma}$ for some $\gamma < 2m - 2$, then the function $h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi})$ is convex (all its level sets are convex) for sufficiently large values of ρ.

Proof:

The function $h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi})$ is convex if and only if the Hessian

$$H(h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi})) = H(h_0(\boldsymbol{\xi})) + H(G(\boldsymbol{\xi}))$$

is a positive matrix.

Due to Euler's homogeneous function theorem, for a homogeneous function f of degree m the first-order partial derivatives are homogeneous of degree $m - 1$.

Therefore, its second-order partial derivatives are homogeneous of degree $m-2$. Thus, the following equality holds:

$$H(h_0(\boldsymbol{\xi})) = |\boldsymbol{\xi}|^{m-2} H(h_0(\boldsymbol{\xi}')).$$

Now we use the strict convexity of the $h_0(\boldsymbol{\xi}')$. This property implies that for any non-zero vector $\mathbf{x} \in \mathbb{R}^d$, $|\mathbf{x}| = 1$,

$$\mathbf{x}^*(H(h_0(\boldsymbol{\xi}')))\mathbf{x} > 0.$$

Due to the compactness of the unit sphere, there exists a positive constant $\nu > 0$ such that for any vector $\mathbf{x}$, $|\mathbf{x}| = 1$, it holds

$$\mathbf{x}^*(H(h_0(\boldsymbol{\xi}')))\mathbf{x} > \nu.$$

Therefore,

$$\mathbf{x}^* H(h_0(\boldsymbol{\xi})))\mathbf{x} \gg \rho^{2m-2}.$$

If $|H(G(\boldsymbol{\xi}))| \ll \rho^{\gamma}$, where $\gamma < 2m-2$, then for sufficiently large ρ

$$\mathbf{x}^*(H(h_0(\boldsymbol{\xi})) + H(G(\boldsymbol{\xi})))\mathbf{x} \gg \rho^{2m-2},$$

thus the matrix $H(h_0(\boldsymbol{\xi})) + H(G(\boldsymbol{\xi}))$ is positive. □

Lemma 2.9.2. There exists a positive constant C_{11} such that for an arbitrary $\boldsymbol{\theta}$ the intersection of the line passing through the origin and $\boldsymbol{\theta}$ and the set $\mathcal{A}_*(\rho, \delta)$ is a segment of the length not greater then $C_{11}\rho^{1-2m}\delta$ (where C_{11} does not depend on $\boldsymbol{\theta}$).

Proof:

Consider the function $h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi})$ on the line containing the origin as a function of one variable and differentiate it in the direction $\boldsymbol{\theta}$:

$$\frac{\partial(h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi}))}{\partial\boldsymbol{\theta}} = \frac{\partial h_0(\boldsymbol{\xi})}{\partial\boldsymbol{\theta}} + \langle \nabla G(\boldsymbol{\xi}), \boldsymbol{\theta}\rangle \geq C_{12}|\boldsymbol{\xi}|^{2m-1} + |\nabla G(\boldsymbol{\xi})| > C_{13}\rho^{2m-1}$$

(here we used (2.9.5), and denoted by C_{12} maximum of the function a).

Therefore, this line makes a cut in the set $\mathcal{A}_*(\rho,\delta)$ of length not greater than $C_{11}\rho^{1-2m}\delta$. □

Now we are going to prove one more auxiliary statement. Recall that we consider only the upper halfplane.

Lemma 2.9.3. For any $\varepsilon > 0$ there exists a value ρ_ε such that for any $\rho > \rho_\varepsilon$, $\mathbf{b} : |\mathbf{b}| \geq 1$ and δ satisfying (2.9.8), we have that the set $\mathcal{A}_{\mathbf{b}}(\rho,\delta)$ can be covered by a ball of the radius $\rho\varepsilon$.

Proof:

Obviously, it is sufficient to prove the statement for the maximal possible value of δ, that is $\delta = \rho^{\varkappa}$. Let us take a look at the whole picture scaled down by a factor ρ. After this rescaling we will have 1 instead of linear size ρ, $\delta_{new} = \rho^{\varkappa-2m}$ instead of $\delta_{old} = \rho^{\varkappa}$, and we will be interested in a ball of radius ε instead of $\rho\varepsilon$. This way we can reformulate our statement: we have to prove that there exists a large enough value of ρ such that the set $\mathcal{A}_{\mathbf{b}}(\rho^{\varkappa-2m},1)$ can be covered by a ball of radius ε for any vector $\mathbf{b}$ of length greater than or equal to $\frac{1}{\rho}$.

For convenience let us consider a horizontal vector $\mathbf{b}$. Denote by $\boldsymbol{\xi}_0$ the highest point of the curve U defined by $U = \{\boldsymbol{\xi} : h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi}) = 1\}$. Let us consider a ball of radius $\frac{\varepsilon}{2}$ with center at the point $\boldsymbol{\xi}_0$.

In Figure 2.1 we show the layer $\mathcal{A}_*(\rho^{\varkappa-2m},1)$, the unit level set U in the middle of the layer, the point $\boldsymbol{\xi}_0$ and the ball around it. Denote by T the distance between the points of intersection of the ball and U. Let y_0 be the least of the ordinates of these two points.

Let us prove the following fact.

Fact: There exists a value ρ_0 such that for any $\rho > \rho_0$ and any vector $\mathbf{b}$ satisfying $\frac{1}{\rho} \leq |\mathbf{b}| \leq \frac{T}{2}$, the set $\mathcal{A}_{\mathbf{b}}(\rho^{e-2m},1)$ lies inside the ball $B_\varepsilon(\boldsymbol{\xi}_0)$ of the radius ε and the center $\boldsymbol{\xi}_0$.

Define the function $v(\cdot)$ as follows: the value $v(y)$ equals the sum of lengths of the sections cut by a horizontal line of height y on our layer $\mathcal{A}_*(\rho^{\varkappa-2m},1)$.

Denote by Y the right one of the points of the intersection of $B_\varepsilon(\boldsymbol{\xi}_0)$ with U.

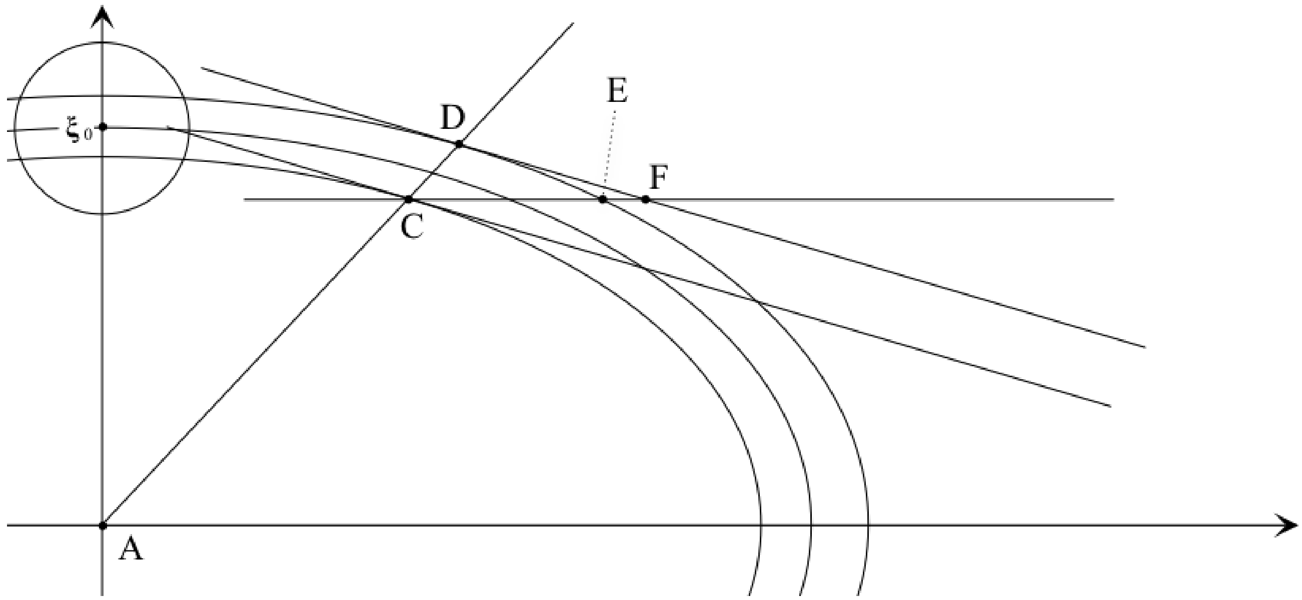

Figure 2.1: Layer

Denote by γ the acute angle between the horizontal line and the tangent to the curve at the point Y. It is clear that for all the other points on U in the top right quarter of the plane outside the ball, this angle (between the horizontal line and the tangent one) will be greater than or equal to γ.

Let us show that for any $y < y_0$ the function $v(y)$ has an upper bound: $v(y) \leq \frac{C\rho^{\varkappa-2m}}{\sin\gamma}$. This can be seen on the picture. The distance we are interested in for the drawn horizontal line is the length of the interval CE. It can be estimated from above by $|CF|$. For the triangle CDF we can apply the sine theorem and recall that the value of the angle CFD is at least γ.

Thus we obtain that $|CF| \leq \frac{|CD|\sin\angle CDF}{\sin\gamma}$. Obviously $\sin\angle CDF \leq 1$, and we know that $|CD| \asymp \rho^{\varkappa-2m}$.

Now let us finally prove our **Fact**. Choose a value ρ_0, so that for any $\rho > \rho_0$ we will have

$$\frac{C\rho^{\varkappa-2m}}{\sin\gamma} < \frac{1}{\rho}.$$

This value exists due to $\varkappa - 2m < -1$ (according to 2.9.8). This concludes the proof of the Fact.

Now let us consider a horizontal vector $\mathbf{b}$, such that $\frac{1}{\rho} \leq |\mathbf{b}| \leq \frac{T}{2}$ and shift our layer by $\mathbf{b}$. Below the height y_0 the distance between the left and the right parts of the layer is greater then $\frac{T}{2}$ (if not, just increase ρ, this condition can be satisfied independently of the previous one). Also, below the height y_0 the horizontal line cuts in the layer sections of total length less than $\frac{1}{\rho}$. This means that all the set $\mathcal{A}_{\mathbf{b}}(\rho^{\varkappa-2m}, 1)$ has to be above the height y_0.

Now let us note that we can choose a value ρ_1, so that for any $\rho > \rho_1$ all the points of the layer lying above y_0, will be inside the ball with the center $\boldsymbol{\xi}_0$ and the radius ε (recall that up to this moment we were using the ball of the radius $\frac{\varepsilon}{2}$).

Now we can apply the uniform continuity argument. Assume that the statement of the lemma is false. Then there exist sequences $\rho_k \to \infty$ and $\mathbf{b}_k$, such that for any k the set $\mathcal{A}_{\mathbf{b}_k}(\rho_k^{\varkappa-2m}, 1)$ can not be covered by a ball of the radius ε. Vectors $\mathbf{b}_k$ belong to a compact set, therefore we can choose a converging subsequence of this sequence. Thus without loss of generality in order to keep notation simple, we will assume that the sequence $\{\mathbf{b}_k\}$ converges. Take its limit. Consider two cases:

1. If this limit is a zero vector, there exists a number k^*, such that for any $k > k^*$ we have $\frac{1}{\rho_k} \leq |\mathbf{b}_k| \leq \frac{T}{2}$. Then for sufficiently large values of ρ_k we arrive at a contradiction to the **Fact**.

2. If this limit is a nonzero vector, it means that the set $\mathcal{A}_{\mathbf{b}_k}(\rho_k^{\varkappa-2m}, 1)$ tends to a point, and again we arrive at a contradiction.

□

Fix the values of ρ, δ and $\mathbf{b}$. Consider the angle between $\boldsymbol{\xi}$ and $\boldsymbol{\xi} + \mathbf{b}$ for all $\boldsymbol{\xi} \in \mathcal{A}_{\mathbf{b}}(\rho, \delta)$ and its sine. Denote by $\beta^*(\mathbf{b})(= \beta^*(\mathbf{b}, \rho, \delta))$ the angle for which this sine is minimal. Obviously, this angle depends also on the values of ρ and δ, but for convenience of the notation we will omit this dependence.

Proposition 2.9.4. *If $\sin \beta(\boldsymbol{b}) > 0$, then there exists a constant $C > 0$ independent of $\boldsymbol{b}$ such that for any sufficiently large ρ and arbitrary value of δ satisfying the condition (2.9.8) we have:*

$$\mathrm{vol}(\mathcal{A}_{\boldsymbol{b}}(\rho, \delta)) \leq \frac{C\delta^2 \rho^{2-4m}}{\sin \beta(\boldsymbol{b})}.$$

Proof:

We construct the proof of this statement according to the following plan:

1. For each point $\boldsymbol{\xi} : h_0(\boldsymbol{\xi})+G(\boldsymbol{\xi}) = \rho^{2m}$ and a number ε we construct a rectangle $\mathbf{Rec}(\boldsymbol{\xi}, \varepsilon)$, whose side lengths depend on ρ.

2. Choose ε^* so that for sufficiently large values of ρ the set $\mathcal{A}_{\mathbf{b}}(\rho, \delta)$ is covered by the rectangle $\mathbf{Rec}(\boldsymbol{\xi}, \varepsilon^*)$.

3. Then we will show that the set $\mathcal{A}_{\mathbf{b}}(\rho, \delta)$ is covered by the rectangle $\mathbf{Rec}(\boldsymbol{\xi} + \mathbf{b}, \varepsilon^*) - \mathbf{b}$.

4. We will estimate the area of the intersection for the rectangles from the 2nd and 3rd steps by the area of the intersection of two infinite strips of width $2C_{11}\delta\rho^{1-2m}$ (the constant C_{11} has been defined in Lemma 2.9.2) and with the angle between them bounded from below by the value $\beta(\mathbf{b})$.

So let us proceed according to plan.

1. Fix $\delta > 0$. Take an arbitrary value $\varepsilon < 1$. For each point $\boldsymbol{\xi} : h_0(\boldsymbol{\xi}) + G(\boldsymbol{\xi}) = \rho^{2m}$ denote by $\mathbf{Rec}(\boldsymbol{\xi}, \varepsilon)$ a rectangle with the sides $\varepsilon\rho$ and $2C_{11}\delta\rho^{1-2m}$, such that the side of length $2C_{11}\delta\rho^{1-2m}$ is parallel to the vector $\boldsymbol{\xi}$ and the point $\boldsymbol{\xi}$ is the center of the rectangle.

2. Denote by $\varepsilon(\boldsymbol{\xi})$ the maximal width of the rectangle $\mathbf{Rec}(\boldsymbol{\xi}, \varepsilon)$, such that its sides of length $\varepsilon\rho$ do not intersect with the set $\mathcal{A}_*(\rho, \delta)$. The function $\varepsilon(\boldsymbol{\xi})$ is continuous and thus attains its minimum on the compact. Denote:

$$\varepsilon^* = \min_{\boldsymbol{\xi}} \varepsilon(\boldsymbol{\xi})$$

(so far we were doing everything for a fixed value of ρ).

Clearly, for any $\boldsymbol{\xi}$ we have $\varepsilon(\boldsymbol{\xi}) > 0$. Therefore also $\varepsilon^* > 0$ (as a minimum of a uniformly continuous strictly positive function defined on a compact set).

For this ε^* the following holds: for any rectangle $\mathbf{Rec}(\boldsymbol{\xi}, \varepsilon^*)$ its sides of length $\varepsilon^*\rho$ do not intersect with the set $\mathcal{A}_*(\rho, \delta)$. Moreover, clearly, this remains true when we increase the value of ρ.

Consider a positive $\epsilon^* < \frac{\varepsilon^*}{4}$ and a value ρ_{ϵ^*}, defined above.

For any value $\rho > \rho_{\epsilon^*}$ and any point $\boldsymbol{\xi} \in \mathcal{A}_{\mathbf{b}}(\rho, \delta)$ we have:

$$\mathcal{A}_{\mathbf{b}}(\rho, \delta) \subset \mathbf{Rec}(\boldsymbol{\xi}, \varepsilon^*).$$

3. One can easily see that

$$\mathcal{A}_{\mathbf{b}}(\rho, \delta) + \mathbf{b} \subset \mathbf{Rec}(\boldsymbol{\xi} + \mathbf{b}, \varepsilon^*) \Rightarrow \mathcal{A}_{\mathbf{b}}(\rho, \delta) \subset \mathbf{Rec}(\boldsymbol{\xi} + \mathbf{b}, \varepsilon^*) - \mathbf{b}.$$

4. Obviously, if $\boldsymbol{\xi} \in \mathcal{A}_{\mathbf{b}}(\rho, \delta)$ and $\boldsymbol{\eta} = \boldsymbol{\xi} + \mathbf{b}$, we have:

$$\mathcal{A}_{\mathbf{b}}(\rho, \delta) \subset \mathbf{Rec}(\boldsymbol{\xi}, \varepsilon^*) \cap (\mathbf{Rec}(\boldsymbol{\eta}, \varepsilon^*) - \mathbf{b}) \Rightarrow$$

$$\Rightarrow \mathrm{vol}(\mathcal{A}_{\mathbf{b}}(\rho, \delta)) \leq \mathrm{vol}(\mathbf{Rec}(\boldsymbol{\xi}, \varepsilon^*) \cap (\mathbf{Rec}(\boldsymbol{\eta}, \varepsilon^*) - \mathbf{b}))$$
$$\leq \frac{(2C_{11}\delta\rho^{1-2m})^2}{\sin\angle(\boldsymbol{\xi}, \boldsymbol{\eta})} \leq \frac{C\delta^2\rho^{2-4m}}{\sin\beta(\mathbf{b})},$$

which finishes the proof.

□

Consider the function $b(\cdot)$ defined in (2.9.1) on the unit sphere. It is continuous, smooth, and defined on a compact. Consequently, there exists a positive constant s such that

$$|b(\boldsymbol{\xi}') - b(\boldsymbol{\eta}')| \leq s|\angle(\boldsymbol{\xi}, \boldsymbol{\eta})| \quad \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^d \setminus \{0\}. \tag{2.9.11}$$

Let us divide the set $\mathcal{A}_{\mathbf{b}}(\rho, \delta)$ into two parts: $\mathcal{A}_{\mathbf{b}}(\rho, \delta) = \mathcal{A}_1 \sqcup \mathcal{A}_2$, where

$$\mathcal{A}_1 = \{\boldsymbol{\xi} : \boldsymbol{\xi} \in \mathcal{A}_{\mathbf{b}}(\rho, \delta), |\boldsymbol{\xi}_2| < \rho^{-u}\}, \quad \mathcal{A}_2 = \mathcal{A}_{\mathbf{b}}(\rho, \delta) \setminus \mathcal{A}_1$$

(here $(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ are the coordinates of the vector $\boldsymbol{\xi}$ along the axes x_1, x_2 and $u > 0$ is a constant to be determined later).

Lemma 2.9.5. $\mathrm{vol}(\mathcal{A}_1) \ll \delta\rho^{1-2m-u}$.

Proof:

Let us estimate the volume of the set $\mathcal{A}_1$. Suppose, $\boldsymbol{\xi} \in \mathcal{A}_1$.

Applying (2.9.4), we conclude that $h_0(\boldsymbol{\xi}) \gg \rho^{2m}$.

Consider the function $b(\cdot)$ and denote by b and B its lower and upper bounds correspondingly: $0 < b < b(\cdot) < B$.

$$h_1(\boldsymbol{\xi}) \gg \rho^2 \Rightarrow B(\boldsymbol{\xi}_1^2 + \boldsymbol{\xi}_2^2) \gg \rho^2,$$

and since $\boldsymbol{\xi}_2 \ll \rho^{-u}$, we have

$$\boldsymbol{\xi}_1^2 \gg \rho^2. \tag{2.9.12}$$

Consider a function

$$h_2(\boldsymbol{\xi}) = \frac{h_1(\boldsymbol{\xi})}{\boldsymbol{\xi}^2}.$$

Let us note that $h_2(\boldsymbol{\xi}) = h_2(\boldsymbol{\xi}') = b(\boldsymbol{\xi}')$.

Then for the derivative we write put the following expression:

$$\frac{\partial h_0}{\partial x}(\boldsymbol{\xi}) = m(b(\boldsymbol{\xi}'))^{m-1}|\boldsymbol{\xi}|^{2(m-1)}\left(\frac{\partial h_2(\boldsymbol{\xi}')}{\partial x}|\boldsymbol{\xi}^2| + 2h_2(\boldsymbol{\xi}')\boldsymbol{\xi}_1\right). \tag{2.9.13}$$

Consider the value $\frac{\partial h_2(\boldsymbol{\xi}')}{\partial x}$. Let $\mathbf{x}$ be a unit vector in the direction of the axis x, and Δ be a real number. Take the triangle ABC with the vertices $A = \boldsymbol{\xi}'$, $B = \boldsymbol{\xi}' + \Delta\mathbf{x}$ and $C = \mathbf{0}$. This derivative can be written as the following limit:

$$\frac{\partial h_2}{\partial x}(\boldsymbol{\xi}') = \lim_{\Delta\to 0}\frac{h_2(\boldsymbol{\xi}' + \Delta\mathbf{x}) - h_2(\boldsymbol{\xi}')}{\Delta} = \lim_{\Delta\to 0}\frac{h_2((\boldsymbol{\xi}' + \Delta\mathbf{x})') - h_2(\boldsymbol{\xi}')}{\Delta}.$$

Recalling (2.9.11), we obtain an inequality:

$$h_2((\boldsymbol{\xi}' + \Delta\mathbf{x})') - h_2(\boldsymbol{\xi}') \le s\angle(\boldsymbol{\xi}' + \Delta\mathbf{x}, \boldsymbol{\xi}') \Rightarrow \frac{\partial b}{\partial x}(\boldsymbol{\xi}') \le \lim_{\Delta\to 0}\frac{s\angle(\boldsymbol{\xi}' + \Delta\mathbf{x}, \boldsymbol{\xi}')}{\Delta}.$$

Consider a fraction $\frac{s\angle(\boldsymbol{\xi}'+\Delta\mathbf{x},\boldsymbol{\xi}')}{\Delta}$. Applying the law of sines to our triangle and keeping in mind that $\angle(\boldsymbol{\xi}' + \Delta\mathbf{x}, \boldsymbol{\xi}') \approx \sin(\boldsymbol{\xi}' + \Delta\mathbf{x}, \boldsymbol{\xi}')$:

$$\frac{\sin(\boldsymbol{\xi}' + \Delta\mathbf{x}, \boldsymbol{\xi}')}{\Delta} = \frac{\sin\angle C}{|AB|} = \frac{\sin\angle B}{|AC|}.$$

By the definition of the set $\mathcal{A}_1$ and the property (2.9.12), the value $\sin\angle B$ does not exceed ρ^{-1-u}, and the side $|AC|$ has length of order ρ, therefore

$$\frac{\sin\angle B}{|AC|} \le \rho^{-u-2} \Rightarrow$$

$$\Rightarrow \frac{\partial h_2(\boldsymbol{\xi}')}{\partial x} \ll \rho^{-2-u} \Rightarrow \frac{\partial h_2(\boldsymbol{\xi}')}{\partial x}|\boldsymbol{\xi}^2| + 2h_2(\boldsymbol{\xi}')\boldsymbol{\xi}_1 \gg \rho \Rightarrow \frac{\partial h_0}{\partial x}(\boldsymbol{\xi}) \gg \rho^{2m-1}.$$

Using (2.9.5), we conclude that $\frac{\partial(h_0+G)}{\partial x}(\boldsymbol{\xi}) \gg \rho^{2m-1}$.

Consequently, the volume of the set $\mathcal{A}_1$ does not exceed $\delta\rho^{1-2m-u}$ (we estimate the area as the product of the width ($\delta\rho^{1-2m}$) and the height (ρ^{-u})), which finishes the proof of the Lemma. □

Now let us estimate the volume of the set $\mathcal{A}_2$, using the Proposition 2.9.4. Note that for any vector $\boldsymbol{\xi} \in \mathcal{A}_2$ it holds that $\sin\angle(\boldsymbol{\xi}, \boldsymbol{\xi} + \mathbf{b}) \gg \rho^{-1-u}$, and thus

$$\mathrm{vol}(\mathcal{A}_2) \ll \delta^2\rho^{3-4m+u} \Rightarrow \mathrm{vol}(\mathcal{A}_{\mathbf{b}}(\rho,\delta)) \ll \delta^2\rho^{3-4m+u} + \delta\rho^{1-2m-u}.$$

Now we will derive a result for higher dimensions from the proven facts for $d = 2$ (Proposition 2.9.4 and Lemma 2.9.5):

Theorem 2.9.6. *If $d \geq 2$ then for any $u > 0$*

$$\mathrm{vol}(\mathcal{A}_{\boldsymbol{b}}(\rho,\delta)) \ll \delta^2\rho^{1+d-4m+u} + \delta\rho^{1-2m-u(d-1)}.$$

Proof:

Let us estimate the volumes of the sets $\mathcal{A}_1$ and $\mathcal{A}_2$ separately.

Consider the set $\mathcal{A}_1$. It is contained in a cylinder of radius ρ^{-u}, so the two-dimensional result we have to multiply by $\rho^{-u(d-2)}$:

$$\mathrm{vol}(\mathcal{A}_1) \ll \delta\rho^{1-2m-u}\rho^{-u(d-2)} = \delta\rho^{1-2m-u(d-1)}.$$

Now consider the set $\mathcal{A}_2$. Similarly it is contained in a cylinder of radius ρ and thus we multiply the two-dimensional result by ρ^{d-2}:

$$\mathrm{vol}(\mathcal{A}_2) \ll \delta^2\rho^{3-4m+u}\rho^{(d-2)} = \delta^2\rho^{1+d-4m+u}.$$

Therefore,

$$\mathrm{vol}(\mathcal{A}_{\mathbf{b}}(\rho,\delta)) \ll \delta^2\rho^{1+d-4m+u} + \delta\rho^{1-2m-u(d-1)}\ .\square$$

Corollary 2.9.7. *As soon as conditions (2.9.3) - (2.9.8) are satisfied, for any $u > 0$ we have*

$$\mathrm{vol}(\mathcal{B}(\rho,\delta) \cap (\mathcal{B}(\rho,\delta) + \boldsymbol{b})) \ll \delta^2\rho^{1+d-4m+u} + \delta\rho^{1-2m-u(d-1)}. \tag{2.9.14}$$

Moreover,

$$\operatorname{vol} \bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{0\}} \Big(\mathcal{B}(\rho,\delta)\cap\big(\mathcal{B}(\rho,\delta)+\mathbf{n}\big)\Big) \ll \delta^2\rho^{1+2d-4m+u}+\delta\rho^{1-2m+d-u(d-1)}. \quad (2.9.15)$$

Another important ingredient is the following estimate on the volumes:

Lemma 2.9.8. *Let* $\mathcal{B}(\rho,\delta)$, $\mathcal{D}(\rho,\delta)$, $\delta\in(0,\rho^{2m}/4]$, *be as defined in* (2.8.2). *Let* $u>0$ *be some number, then*

$$\operatorname{vol} \bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{0\}} \Big(\mathcal{B}(\rho,\delta)\cap\big(\mathcal{D}(\rho,\delta)+\mathbf{n}\big)\Big)$$

$$\ll \delta^2\rho^{1+2d-4m+u}+\delta\rho^{1-2m+d-u(d-1)}+\delta\rho^{d-1-2m+\alpha_d}. \quad (2.9.16)$$

Proof: Let us split $\mathcal{D}(\rho,\delta)$ in three disjoint sets:

$$\mathcal{D}_0(\rho,\delta)=\{\boldsymbol{\xi}\in\mathcal{D}(\rho,\delta):\boldsymbol{\xi}+\mathbf{n}\notin\mathcal{B}(\rho,\delta),\ \text{for all}\ \ \mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}\},$$

$$\mathcal{D}_1(\rho,\delta)=\{\boldsymbol{\xi}\in\mathcal{D}(\rho,\delta):$$
$$\text{there exists a } \underline{\text{unique}}\ \ \mathbf{n}=\mathbf{n}(\boldsymbol{\xi})\in\Gamma^\dagger\setminus\{\mathbf{0}\} \text{ such that } \boldsymbol{\xi}+\mathbf{n}\in\mathcal{B}(\rho,\delta)\},$$

$$\mathcal{D}_2(\rho,\delta)=\mathcal{D}(\rho,\delta)\setminus\Big(\mathcal{D}_0(\rho,\delta)\bigcup\mathcal{D}_1(\rho,\delta)\Big).$$

The definition of $\mathcal{D}_0(\rho,\delta)$ immediately implies that

$$\mathcal{B}(\rho,\delta)\bigcap\Big(\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\big(\mathcal{D}_0(\rho,\delta)+\mathbf{n}\big)\Big)=\varnothing. \quad (2.9.17)$$

For the set $\mathcal{D}_2(\rho,\delta)$ we have the inclusion

$$\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\big(\mathcal{D}_2(\rho,\delta)+\mathbf{n}\big)\subset\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\big(\mathcal{B}(\rho,\delta)+\mathbf{n}\big). \quad (2.9.18)$$

Indeed,for each $\boldsymbol{\xi}\in\mathcal{D}_2(\rho,\delta)$ there are at least two distinct lattice vectors $\mathbf{n}_1,\mathbf{n}_2\neq 0$ such that $\boldsymbol{\xi}+\mathbf{n}_1\in\mathcal{B}(\rho,\delta)$ and $\boldsymbol{\xi}+\mathbf{n}_2\in\mathcal{B}(\rho,\delta)$, so that any lattice vector $\mathbf{m}\neq 0$ is distinct either from $\mathbf{n}_1$ or from $\mathbf{n}_2$. Thus, assuming for definiteness that $\mathbf{m}\neq\mathbf{n}_1$, we get

$$\boldsymbol{\xi}+\mathbf{m}=\boldsymbol{\xi}+\mathbf{n}_1+(\mathbf{m}-\mathbf{n}_1)\in\big(\mathcal{B}(\rho,\delta)+\mathbf{m}-\mathbf{n}_1\big)\subset\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\big(\mathcal{B}(\rho,\delta)+\mathbf{n}\big).$$

This proves (2.9.18).

Now observe that by definition of $\mathcal{D}_1(\rho,\delta)$ the sets $\mathcal{D}_1(\rho,\delta)\cap\big(\mathcal{B}(\rho,\delta)+\mathbf{n}\big)$ are disjoint for different $\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}$. Therefore

$$\operatorname{vol}\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\Big(\big(\mathcal{D}_1(\rho,\delta)+\mathbf{n}\big)\cap\mathcal{B}(\rho,\delta)\Big)=\sum_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\operatorname{vol}\Big(\mathcal{D}_1(\rho,\delta)\cap\big(\mathcal{B}(\rho,\delta)+\mathbf{n}\big)\Big)$$

$$\le\operatorname{vol}\mathcal{D}_1(\rho,\delta)\le\operatorname{vol}\mathcal{D}(\rho,\delta).$$

Together with (2.9.17) and (2.9.18) this produces the bound

$$\operatorname{vol}\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\Big(\mathcal{B}(\rho,\delta)\cap\big(\mathcal{D}(\rho,\delta)+\mathbf{n}\big)\Big)\le\operatorname{vol}\mathcal{D}(\rho,\delta)+\operatorname{vol}\bigcup_{\mathbf{n}\in\Gamma^\dagger\setminus\{\mathbf{0}\}}\Big(\mathcal{B}(\rho,\delta)\cap\big(\mathcal{B}(\rho,\delta)+\mathbf{n}\big)\Big).$$

The estimate (2.9.16) follows from (2.9.15) and (2.8.11).

2.10 Proof of the Bethe-Sommerfeld Conjecture

In this section, we prove Theorem 2.1.1. We do it in several steps. First we prove it for the model operator A defined by (2.6.2) with conditions (2.6.4) satisfied. After that we invoke Theorem 2.5.3, which states that the original operator H can be reduced to the model operator up to controllable error terms. At the second step we show that these errors do not destroy the spectral band overlap, obtained for the model operator.

2.10.1 Theorem 2.1.1 for the model operator (2.6.2)

Our proof of the spectral band overlap for the operator A relies on the following elementary Intermediate Value Theorem type result for the function $g(\boldsymbol{\xi})$ defined in Section 2.7. As before we assume that $\lambda=\rho^{2m}$.

Lemma 2.10.1. *Let $\boldsymbol{\xi}=\boldsymbol{\xi}(t)\subset\mathcal{B}, t\in[t_1,t_2], t_1<t_2$, be a continuous path. Suppose that $g(\boldsymbol{\xi}(t_1))\le\lambda-\delta$, $g(\boldsymbol{\xi}(t_2))\ge\lambda+\delta$ with some $\delta\in(0,\lambda/4)$, and for each $t\in[t_1,t_2]$ the number $g(\boldsymbol{\xi}(t))$ is a simple eigenvalue of $A(\mathbf{k}),\mathbf{k}=\{\boldsymbol{\xi}(t)\}$. Then there exists a $t_0\in(t_1,t_2)$ such that $\lambda=g(\boldsymbol{\xi}(t_0))$, so that $\lambda\in\sigma(A)$. Moreover, $\zeta(\lambda;A)\ge\delta$.*

Proof: Since $g(\boldsymbol{\xi}(t))$ is a simple eigenvalue of $A(\mathbf{k})$ and $\{\boldsymbol{\xi}(t)\} = \mathbf{k}$ for each $t \in [t_1, t_2]$, we have $g(\boldsymbol{\xi}(t)) = \lambda_j(A(\mathbf{k}))$ with j independent of the choice of t. Since g is continuous on $\mathcal{B}$, the function $g(\boldsymbol{\xi}(t))$ is a continuous function of $t \in [t_1, t_2]$, and hence the intermediate value theorem implies that there is a $t_0 \in (t_1, t_2)$ such that $\lambda_j(A(\{\boldsymbol{\xi}(t_0)\})) = \lambda$. The bound $\zeta(\lambda; A) \geq \delta$ follows from the definition (1.0.1) of $\zeta(\lambda; A)$.

Our next step is to prove that there is a path with the properties required in Lemma 2.10.1. In fact we shall prove that the required properties will hold for an interval $I(\boldsymbol{\Omega}; \rho, \delta) \subset (0, \infty)$ (see (2.8.6)) with some $\boldsymbol{\Omega} \in T(\rho)$.

Lemma 2.10.2. *There exists a constant $Z \geq 1$ with the following property. Suppose that for some $\boldsymbol{\Omega} \in T(\rho)$ and some $t \in I(\boldsymbol{\Omega}; \rho, \delta), \delta \in (0, \rho^{2m}/4]$, the number $g(\boldsymbol{\eta})$, $\boldsymbol{\eta} = t\boldsymbol{\Omega}$ is a multiple eigenvalue of $A(\mathbf{k}), \mathbf{k} = \{\boldsymbol{\eta}\}$. Then for any $\tau \in I(\boldsymbol{\Omega}; \rho, \delta)$ there exists a vector $\mathbf{n} \in \Gamma^\dagger \setminus \{\mathbf{0}\}$ such that $\tau\boldsymbol{\Omega} + \mathbf{n} \in \mathcal{A}(\rho, Z\delta)$.*

Proof: Since the number $g(\boldsymbol{\eta}), \boldsymbol{\eta} = t\boldsymbol{\Omega}$, is a multiple eigenvalue, by definition of the function $g(\ \cdot\ )$, there is a vector $\mathbf{p} \in \Gamma^\dagger \setminus \{\mathbf{0}\}$ such that $g(\boldsymbol{\eta}) = g(\boldsymbol{\eta} + \mathbf{p})$. In view of (2.7.3), $|\boldsymbol{\eta} + \mathbf{p}| \asymp \rho$. Since on the non-resonant sets the functions g and $\hat{g}$ coincide, by Lemma 2.7.3, for any $\tau \in I(\boldsymbol{\Omega}; \rho, \delta)$ there exist two vectors $\mathbf{m}_1, \mathbf{m}_2 \in \Gamma^\dagger$, $\mathbf{m}_1 \neq \mathbf{m}_2$ such that, with $\boldsymbol{\xi} = \tau\boldsymbol{\Omega}$,

$$\begin{cases} |g(\boldsymbol{\eta}) - g(\boldsymbol{\xi} + \mathbf{m}_1)| \ll \rho^{2m-1}|\boldsymbol{\eta} - \boldsymbol{\xi}| \ll \delta, \\ |g(\boldsymbol{\eta} + \mathbf{p}) - g(\boldsymbol{\xi} + \mathbf{m}_2)| \ll \rho^{2m-1}|\boldsymbol{\eta} - \boldsymbol{\xi}| \ll \delta. \end{cases} \tag{2.10.1}$$

Here we have used the bound $|t - \tau| \ll \delta\rho^{1-2m}$, which follows from (2.8.7). As $\mathbf{m}_1 \neq \mathbf{m}_2$, one of these vectors is not zero. Denote this vector by $\mathbf{n}$. Since $g(\boldsymbol{\eta}) = g(\boldsymbol{\eta} + \mathbf{p})$, it follows from (2.10.1) that

$$|g(\boldsymbol{\xi} + \mathbf{n}) - g(\boldsymbol{\eta})| \ll \delta,$$

so that $\boldsymbol{\xi} + \mathbf{n} \in \mathcal{A}(\rho, Z\delta)$ with some constant Z independent of $\boldsymbol{\xi}$ and ρ, as required.

The next Lemma is the key point of our argument: it shows that at least for one $\boldsymbol{\Omega} \in T(\rho)$ the interval $I(\boldsymbol{\Omega}; \rho, \delta)$ consists entirely of the points t such that $g(t\boldsymbol{\Omega})$ is a simple eigenvalue.

Lemma 2.10.3. *There exists a vector* $\boldsymbol{\Omega} \in T(\rho)$ *and a number* $c_3 > 0$ *such that for* $\delta = c_3 \rho^{2m-1-d-2(d-1)^{-1}}$ *and each* $t \in I(\boldsymbol{\Omega}; \rho, \delta)$ *the number* $g(\boldsymbol{\xi})$, $\boldsymbol{\xi} = t\boldsymbol{\Omega}$ *is a simple eigenvalue of* $A(\{\boldsymbol{\xi}\})$. *Moreover,* $\zeta(\rho^{2m}; A) \geq \delta$.

Proof: Suppose the contrary, i.e. if ρ is sufficiently large, then for any $\boldsymbol{\Omega} \in T(\rho)$ there is a $t \in I(\boldsymbol{\Omega}; \rho, \delta)$ such that $g(t\boldsymbol{\Omega})$ is a multiple eigenvalue of $A(t\boldsymbol{\Omega})$. Then due to formula (2.8.8), Lemma 2.10.2 implies that

$$\tilde{\mathcal{B}}(\rho, \delta) \subset \bigcup_{\mathbf{n} \in \Gamma^\dagger \setminus \{0\}} (\mathcal{A}(\rho, \delta_1) + \mathbf{n}) \tag{2.10.2}$$

with $\delta_1 := Z\delta$. Since $\tilde{\mathcal{B}}(\rho, \delta) \subset \mathcal{B}(\rho, \delta_1)$, we can re-write (2.10.2) as

$$\begin{aligned} \tilde{\mathcal{B}}(\rho, \delta) &\subset \bigcup_{\mathbf{n} \in \Gamma^\dagger \setminus \{\mathbf{0}\}} \Big(\big(\mathcal{A}(\rho, \delta_1) + \mathbf{n}\big) \bigcap \mathcal{B}(\rho, \delta_1) \Big) \\ &= \bigcup_{\mathbf{n} \in \Gamma^\dagger \setminus \{\mathbf{0}\}} \Big(\big(\mathcal{B}(\rho, \delta_1) + \mathbf{n}\big) \bigcap \mathcal{B}(\rho, \delta_1) \Big) \bigcup \bigcup_{\mathbf{n} \in \Gamma^\dagger \setminus \{\mathbf{0}\}} \Big(\big(\mathcal{D}(\rho, \delta_1) + \mathbf{n}\big) \bigcap \mathcal{B}(\rho, \delta_1) \Big). \end{aligned} \tag{2.10.3}$$

Let us estimate the volumes of sets on both sides of this inclusion. For a fixed $u > 0$, whose value is chosen a few lines down, we can use (2.9.14) and (2.9.16) for the volume of the right hand side. For the left hand side we use (2.8.10), so that (2.10.3) results in the estimate

$$\delta \rho^{d-2m} \ll \delta^2 \rho^{1-4m+2d+u} + \delta \rho^{1-2m+d-u(d-1)} + \delta \rho^{d-1-2m+\alpha_d},$$

which simplifies to

$$1 \ll \delta \rho^{1-2m+d+u} + \rho^{1-u(d-1)} + \rho^{-1+\alpha_d}.$$

Choose $u = 2(d-1)^{-1}$ and $\delta = c_3 \rho^{2m-1-d-u}$ with a suitably small c_3. Then for large ρ the right hand side is less than the left hand side, which produces a contradiction, thus proving the Lemma.

2.10.2 Proof of the Main Theorem

We assume that the conditions of Theorem 2.1.1 are satisfied. The proof uses the reduction of the operator H to A_1, established in Theorem 2.5.3. The first step is to show that the spectrum of A_1 is well approximated by that of the model operator (2.6.2) with B replaced with X, i.e.

$$A = H_0 + X^o + X^{\mathcal{R}}.$$

Let numbers $\alpha_j < 1, j = 1, 2, \dots, d$ be as defined in Subsection 2.3.3.

Lemma 2.10.4. *Suppose that the conditions of Theorem 2.1.1 are satisfied. Let A_1 be the operator* (2.5.24)*, and let $r = \rho^{\varkappa}$ with a number $\varkappa > 0$, satisfying* (2.1.2) *and the inequality*

$$d^2\varkappa < (2m - \gamma\beta)\alpha_d. \tag{2.10.4}$$

Then for any $L > 0$ there exists an M (i.e. the number of steps in Theorem 2.5.3) such that

$$N(\mu - \rho^{-L}, A(\mathbf{k})) \le N(\mu, A_1(\mathbf{k})) \le N(\mu + \rho^{-L}, A(\mathbf{k})) \tag{2.10.5}$$

for all $\mu \in \big((1-c_4)^{2m}\rho^{2m}, (1+c_4)^{2m}\rho^{2m}\big)$ with any $c_4 < \frac{1}{16A_{max}}$.

Proof: By Theorem 2.5.3, $\|R_{M+1}\| \ll \rho^{\beta\epsilon_{M+1}}$, uniformly in $b : |b|^{(\gamma)} \ll 1$ (see (2.5.19) for definition of ϵ_{M+1}). The condition (2.1.15) is equivalent to $\sigma < 1$, so that $\epsilon_j \to -\infty$ as $j \to \infty$. Thus for sufficiently large $M = M(L)$ we have $\|R_{M+1}\| \ll \rho^{-L}/2$. As a consequence,

$$N\big(\mu - \rho^{-L}/2, \tilde{A}_1(\mathbf{k})\big) \le N(\mu, A_1(\mathbf{k})) \le N\big(\mu + \rho^{-L}/2, \tilde{A}_1(\mathbf{k})\big),$$
$$\tilde{A}_1 = A + X^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}, \tag{2.10.6}$$

for all $\mu \in \mathbb{R}$. Due to (2.6.10), the operator $\tilde{A}_1$ can be represented in the block-matrix form:

$$\tilde{A}_1 = \bigoplus_{\mathfrak{V}\in\mathcal{V}} \mathcal{P}(\Xi(\mathfrak{V}))A_{\mathfrak{V}}\mathcal{P}(\Xi(\mathfrak{V})) + \bigoplus_{\mathfrak{V},\mathfrak{W}\in\mathcal{V},} \mathcal{P}(\Xi(\mathfrak{V}))X^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}\mathcal{P}(\Xi(\mathfrak{W})).$$

Since the number of distinct subspaces $\mathfrak{W} \in \mathcal{V}$ is bounded above by Cr^{d^2} with some universal constant $C > 0$, the second term satisfies the two-sided estimate

$$-Cr^{d^2}\bigoplus_{\mathfrak{V}\in\mathcal{V}}\mathcal{P}(\Xi(\mathfrak{V}))|X|^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}\mathcal{P}(\Xi(\mathfrak{V})) \le \bigoplus_{\mathfrak{V},\mathfrak{W}\in\mathcal{V},} \mathcal{P}(\Xi(\mathfrak{V}))X^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}\mathcal{P}(\Xi(\mathfrak{W}))$$
$$\le Cr^{d^2}\bigoplus_{\mathfrak{V}\in\mathcal{V}}\mathcal{P}(\Xi(\mathfrak{V}))|X|^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}\mathcal{P}(\Xi(\mathfrak{V})).$$

Here we have denoted $|X|^{\mathcal{SE},\mathcal{LE},\mathcal{LF}} = |X^{\mathcal{SE}}| + |X^{\mathcal{LE}}| + |X^{\mathcal{LF}}|$. Consequently,

$$\tilde{A}_- \leq \tilde{A}_1 \leq \tilde{A}_+$$

with

$$\tilde{A}_\pm = \bigoplus_{\mathfrak{V}\in\mathcal{V}} \mathcal{P}(\Xi(\mathfrak{V}))\big(A_{\mathfrak{V}} \pm Cr^{d^2}|X|^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}\big)\mathcal{P}(\Xi(\mathfrak{V})).$$

Since $\tilde{A}_\pm$ are orthogonal sums, the problem is reduced to estimating the counting functions of $\tilde{A}_\pm(\mathbf{k})$ on each invariant subspace $\mathfrak{H}\big(\mathbf{k};\Xi(\mathfrak{V})\big)$. From now on we assume that $\mathfrak{V}$ is fixed and omit it from the notation.

If $\mathfrak{V} \in \mathcal{V}(r, d)$, i.e. $\mathfrak{V} = \mathbb{R}^d$, then $\Xi = \Xi(\mathfrak{V}) \subset B(0, 2\rho^{\alpha_d})$, see Lemma 2.3.18. Clearly, $\|H_0\mathcal{P}(\Xi)\| \leq \rho^{2m\alpha_d}$. Also, by (2.4.22),

$$|\!|x^o|\!|^{(\gamma)} + |\!|x^{\mathcal{SE}}|\!|^{(\gamma)} + |\!|x^{\mathcal{LE})}|\!|^{(\gamma)} + |\!|x^{\mathcal{LF}}|\!|^{(\gamma)} \ll |\!|b|\!|^{(\gamma)},$$

and hence, by Lemma 2.4.2,

$$\|X^o\mathcal{P}(\Xi)\| + r^{d^2}\|\mathcal{P}(\Xi)|X|^{\mathcal{SE},\mathcal{LE},\mathcal{LF}}\mathcal{P}(\Xi)\| \ll r^{d^2}\rho^{\gamma\beta\alpha_d}.$$

In view of (2.10.4), the right hand side of the last inequality does not exceed $\rho^{2m\alpha_d}$. Consequently, $\|\tilde{A}_\pm\mathcal{P}(\Xi)\| \ll \rho^{2m\alpha_d}$, which implies that $N(\mu, \tilde{A}_\pm(\mathbf{k}); \Xi) = 0$ for all $\mu \geq (\rho/2)^{2m}$.

Now, let us fix $\mathfrak{V} \in \mathcal{V}(r, n), n \leq d - 1$, and prove the bounds

$$N(\mu - \rho^{-L}/2, A_{\mathfrak{V}}(\mathbf{k}); \Xi) \leq N(\mu, \tilde{A}_\pm(\mathbf{k}); \Xi) \leq N(\mu + \rho^{-L}/2, A_{\mathfrak{V}}(\mathbf{k}); \Xi), \quad (2.10.7)$$

for sufficiently large ρ. Split Ξ into three disjoint sets:

$$\Xi = \mathcal{C}_< \cup \mathcal{C}_0 \cup \mathcal{C}_>,$$

$$\mathcal{C}_0 = \left\{\boldsymbol{\xi} \in \Xi : \rho\left(1 - \frac{1}{4A_{max}}\right) \leq |\boldsymbol{\xi}_{/\!/\mathfrak{V}}| \leq \rho\left(1 + \frac{1}{8A_{max}}\right)\right\},$$

$$\mathcal{C}_< = \left\{\boldsymbol{\xi} \in \Xi : |\boldsymbol{\xi}_{/\!/\mathfrak{V}}| < \rho\left(1 - \frac{1}{4A_{max}}\right)\right\}, \quad \mathcal{C}_> = \left\{\boldsymbol{\xi} \in \Xi : \rho\left(1 + \frac{1}{8A_{max}}\right) < |\boldsymbol{\xi}_{/\!/\mathfrak{V}}|\right\}.$$

Note that by definition of the operator $A_{\mathfrak{V}}$ (see (2.6.5)) all three subspaces $\mathcal{H}(\mathcal{C}_0)$, $\mathcal{H}(\mathcal{C}_<)$, $\mathcal{H}(\mathcal{C}_>)$ (see Subsection 2.1.3) are invariant for $A_{\mathfrak{V}}$. Since $|\boldsymbol{\xi}_{\mathfrak{V}}| < 2\rho^{\alpha_{d-1}}$ (see

Lemma 2.3.18), we have

$$\Xi \cap B\left(0, \rho\left(1-\frac{1}{4A_{max}}\right)\right) \subset \mathcal{C}_< \subset B\left(0, \rho\left(1-\frac{3}{16A_{max}}\right)\right),$$
$$\Xi \cap B\left(\rho\left(1+\frac{1}{8A_{max}}\right)\right) \subset (\mathcal{C}_< \cup \mathcal{C}_0) \subset B\left(0, \rho\left(1+\frac{1}{4A_{max}}\right)\right).$$

Therefore, by Lemma 2.4.5,

$$\mathcal{P}(\Xi)|X^{\mathcal{SE}}|\mathcal{P}(\Xi) = \mathcal{P}(\mathcal{C}_<)|X^{\mathcal{SE}}|\mathcal{P}(\mathcal{C}_<),\ \mathcal{P}(\Xi)|X^{\mathcal{LE}}|\mathcal{P}(\Xi) = \mathcal{P}(\mathcal{C}_>)|X^{\mathcal{LE}}|\mathcal{P}(\mathcal{C}_>).$$

Thus, $\tilde{A}_\pm \mathcal{P}(\Xi)$ can be rewritten as

$$\tilde{A}_\pm \mathcal{P}(\Xi) = F_\pm \pm C r^{d^2}\big(\mathcal{P}(\Xi)|X|^{\mathcal{LF}}\mathcal{P}(\Xi) - \mathcal{P}(\mathcal{C}_>)|X|^{\mathcal{LF}}\mathcal{P}(\mathcal{C}_>)\big),$$

with

$$F_\pm = \mathcal{P}(\mathcal{C}_<)\big(A_{\mathfrak{V}} \pm C r^{d^2}|X^{\mathcal{SE}}|\big)\mathcal{P}(\mathcal{C}_<) \oplus \mathcal{P}(\mathcal{C}_0)A_{\mathfrak{V}}\mathcal{P}(\mathcal{C}_0)$$
$$\oplus\, \mathcal{P}(\mathcal{C}_>)\big(A_{\mathfrak{V}} \pm C r^{d^2}|X|^{\mathcal{LF},\mathcal{LE}}\big)\mathcal{P}(\mathcal{C}_>).$$

By (2.4.25),

$$r^{d^2}\|\mathcal{P}(\mathcal{C}_< \cup \mathcal{C}_0)|X^{\mathcal{LF}}|\| + r^{d^2}\||X^{\mathcal{LF}}|\mathcal{P}(\mathcal{C}_< \cup \mathcal{C}_0)\| \ll r^{d^2+p-l}\rho^{\beta\max(\gamma,0)},$$

for any $p > d$ and $l \geq p$ uniformly in b satisfying $[\![b]\!]^{(\gamma)} \ll 1$. As $r = \rho^{\varkappa}$, $\varkappa > 0$, by choosing a sufficiently large l, we can guarantee that the right hand side is bounded by $\rho^{-L}/2$. This leads to the bounds

$$N(\mu - \rho^{-L}/2, F_\pm(\mathbf{k}); \Xi) \leq N(\mu, \tilde{A}_\pm(\mathbf{k}); \Xi) \leq N(\mu + \rho^{-L}/2, F_\pm(\mathbf{k}); \Xi), \tag{2.10.8}$$

for all $\mu \in \mathbb{R}$. Consequently, (2.10.7) will be proved if we show that

$$N(\mu, F_\pm; \Xi) = N(\mu, A_{\mathfrak{V}}; \Xi),\ \left(\rho\left(1-\frac{1}{8A_{max}}\right)\right)^{2m} \leq \mu \leq \left(\rho\left(1+\frac{1}{16A_{max}}\right)\right)^{2m}. \tag{2.10.9}$$

To this end note first that the definition of $\mathcal{C}_<$ and $\mathcal{C}_>$ implies

$$H_0\mathcal{P}(\mathcal{C}_<) \leq \left(\rho\left(1-\frac{3}{16A_{max}}\right)\right)^{2m}\mathcal{P}(\mathcal{C}_<),\ H_0\mathcal{P}(\mathcal{C}_>) \geq \left(\rho\left(1+\frac{1}{8A_{max}}\right)\right)^{2m}\mathcal{P}(\mathcal{C}_>). \tag{2.10.10}$$

Also, by Lemma 2.4.5,

$$\|X^o\| + \|X^{\mathcal{R}}_{\mathfrak{V}}\| + Cr^{d^2}\|X^{\mathcal{SE}}\| \ll r^{d^2}\rho^{\beta\max(\gamma,0)}.$$

Under the condition (2.10.4) the right hand side of this estimate is bounded by $o(\rho^{2m})$, $\rho \to \infty$ uniformly in b. Together with (2.10.10), this implies

$$N(\mu, A_{\mathfrak{V}} \pm Cr^{d^2}\mathcal{P}(\mathcal{C}_<)|X^{\mathcal{SE}}|\mathcal{P}(\mathcal{C}_<); \mathcal{C}_<) = N(\mu, A_{\mathfrak{V}}; \mathcal{C}_<), \quad \mu \geq \left(\rho\left(1 - \frac{1}{8A_{max}}\right)\right)^{2m}. \tag{2.10.11}$$

Furthermore, in view of (2.4.22) and (2.4.2),

$$\mathcal{P}(\mathcal{C}_>)(|X^o| + |X^{\mathcal{R}}_{\mathfrak{V}}| + Cr^{d^2}|X|^{\mathcal{LF},\mathcal{LE}})\mathcal{P}(\mathcal{C}_>) \ll r^{d^2}(H_0 + I)^{\gamma^*}\mathcal{P}(\mathcal{C}_>), \ \gamma^* = \frac{\gamma\beta}{2m}.$$

Using again (2.10.4) and remembering (2.10.10), we conclude that the right hand side is bounded above by $o(1)H_0\mathcal{P}(\mathcal{C}_>)$, $\rho \to \infty$, uniformly in b. Together with (2.10.10), this implies that

$$N(\mu, A_{\mathfrak{V}} \pm Cr^{d^2}\mathcal{P}(\mathcal{C}_>)|X|^{\mathcal{LF},\mathcal{LE}}\mathcal{P}(\mathcal{C}_>); \mathcal{C}_>) = 0, \quad \mu \leq \left(\rho\left(1 + \frac{1}{16A_{max}}\right)\right)^{2m}. \tag{2.10.12}$$

Putting together (2.10.11) and (2.10.12), we arrive at (2.10.9). In combination with (2.10.8) this leads to (2.10.7). Together with (2.10.6) they yield (2.10.5).

Proof: [Proof of the Main Theorem] By Theorem 2.5.3, it suffices to prove that $\zeta(\rho^{2m}, A_1) > c\rho^S$ with some S for sufficiently large ρ. It follows from Lemma 2.10.3 that $\zeta(\rho^{2m}; A) \geq c\rho^S$ with $S = 2m - 4 - d - 12(d-1)^{-1}$. Using the bounds (2.10.5) with $L > -S$, we get the required estimate $\zeta(\rho^{2m}, A_1) \gg \rho^S$ from the definition (1.0.2). This completes the proof of Theorem 2.1.1.

Chapter 3

Lower Bound on the Density of States for Periodic Schrödinger Operators

3.1 Introduction

Since now we will consider a different problem we would like to refresh various definitions and the framework of the subject. Let $H = -\Delta + V$ be a Schrödinger operator in $L_2(\mathbb{R}^d)$ with a smooth periodic potential V. We will assume throughout that $d \geqslant 2$. The *integrated density of states (IDS)* was defined in (1.0.3). The existence of the limit in (1.0.3) is well known, see e.g. [28, 32]. For $H_0 := -\Delta$ the IDS can be easily computed explicitly (e.g. using the representation (3.2.6) below):

$$\mathbf{N}_0(\lambda) = \begin{cases} (2\pi)^{-d} d^{-1} \omega_d \lambda^{d/2}, & \lambda > 0; \\ 0, & \lambda \leqslant 0. \end{cases} \tag{3.1.1}$$

Here $\omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of the unit sphere S^{d-1} in $\mathbb{R}^d$.

This chapter concerns the high–energy behaviour of the Radon–Nikodym derivative of the IDS

$$g := \mathrm{d}\mathbf{N}/\,\mathrm{d}\lambda,$$

which is called the *density of states (DOS)* (see [28]). Our main result is that for large values of λ

$$g(\lambda) \geqslant g_0(\lambda)\big(1 - o(1)\big), \tag{3.1.2}$$

where

$$g_0(\lambda) = \mathrm{d}\mathbf{N}_0(\lambda)/\,\mathrm{d}\lambda = (2\pi)^{-d}\omega_d\lambda^{(d-2)/2}/2.$$

We remark that (3.1.2) should be understood in the sense of measures; in particular, we do not claim that $g(\lambda)$ is everywhere differentiable.

It has been proved in [21] that the spectrum of H contains a semi-axis $[\lambda_0, +\infty)$. This result has an obvious reformulation in terms of the IDS: each point $\lambda \geqslant \lambda_0$ is a point of growth of N. It was also proved in [21] that for each $n \in \mathbb{N}$ and $\varepsilon = \lambda^{-n}$ we have

$$\mathbf{N}(\lambda + \varepsilon) - \mathbf{N}(\lambda) \ll \varepsilon\lambda^{(d-2)/2}. \tag{3.1.3}$$

Later, Yu. Karpeshina suggested that using the technique from that [21], one should be able to prove the opposite bound

$$\mathbf{N}(\lambda + \varepsilon) - \mathbf{N}(\lambda) \gg \varepsilon\lambda^{(d-2)/2} \tag{3.1.4}$$

when λ is sufficiently large, not just with $\varepsilon = \lambda^{-n}$ (when the proof is relatively straightforward given [21]), but also uniformly over all $\varepsilon \in (0, 1]$. In this Chapter we prove that for large λ

$$\mathbf{N}(\lambda + \varepsilon) - \mathbf{N}(\lambda) \geqslant \frac{\omega_d}{2(2\pi)^d}\varepsilon\lambda^{(d-2)/2}\big(1 - o(1)\big). \tag{3.1.5}$$

Note that (3.1.5) implies the claimed bound (3.1.2).

The main result of this Chapter is

Theorem 3.1.1. *For sufficiently large λ and any $\varepsilon > 0$ the integrated density of states of H satisfies* (3.1.5)*.*

The proof of Theorem 3.1.1 is heavily based on the technique of [21] and uses various statements proved therein. In order to minimize the size of the Chapter, we will try to quote as many results as we can from [21], possibly with some minor modifications where necessary.

The chapter is organized as follows. In Section 3.2 we introduce the necessary notation and quote the results of [21] which we need for the proof of Theorem 3.1.1. Sections 3.3 and 3.4 contain some auxiliary results, and the proof is finished in Section 3.5.

3.2 Preliminaries

We study the Schrödinger operator

$$H = -\Delta + V(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d \tag{3.2.1}$$

with the potential V being infinitely smooth and periodic with the lattice of periods Λ. We denote the lattice dual to Λ by $\Lambda^\dagger$; the fundamental cells of these lattices are denoted by Ω and $\Omega^\dagger$, respectively. We introduce

$$Q := \sup\left\{|\boldsymbol{\xi}|\big|\boldsymbol{\xi} \in \Omega^\dagger\right\}. \tag{3.2.2}$$

Let

$$\mathbf{D} := -i\nabla, \quad \mathbf{D}(\mathbf{k}) := \mathbf{D} + \mathbf{k}. \tag{3.2.3}$$

The Floquet-Bloch decomposition allows to represent our operator (3.2.1) as a direct integral:

$$H = \int_{\Omega^\dagger} \oplus H(\mathbf{k})\, \mathrm{d}\mathbf{k}, \tag{3.2.4}$$

where

$$H(\mathbf{k}) = \mathbf{D}(\mathbf{k})^2 + V(\mathbf{x}) \tag{3.2.5}$$

is the family of 'fibre' operators acting in $L_2(\Omega)$. The domain $\mathfrak{D}$ of each $H(\mathbf{k})$ is the set of periodic functions from $H^2(\Omega)$. The spectrum of H is the union over $\mathbf{k} \in \Omega^\dagger$ of the spectra of the operators (3.2.5). Let $\{\lambda_j(\mathbf{k})\}$, $j \in \mathbb{N}$ be the set of eigenvalues of $H(\mathbf{k})$ (counting with multiplicities). Then the integrated density of states (1.0.3) admits the following representation:

$$\mathbf{N}(\lambda) := (2\pi)^{-d} \int_{\Omega^\dagger} \#\big\{j : \lambda_j(\mathbf{k}) < \lambda\big\}\, \mathrm{d}\mathbf{k}, \tag{3.2.6}$$

see e.g. [28].

We denote by $|\cdot|_\circ$ the surface area Lebesgue measure on the unit sphere S^{d-1} in $\mathbb{R}^d$ and put $\omega_d := |S^{d-1}|_\circ = 2\pi^{d/2}/\Gamma(d/2)$. By $\mathrm{vol}(\cdot)$ we denote the Lebesgue measure in $\mathbb{R}^d$. We write $B(R)$ for the ball of radius R centered at the origin. The identity matrix is denoted by $\mathbf{I}$. By $\lambda = \rho^2$ we denote a point on the spectral axis. We also denote by v the L_∞–norm of the potential V, and put $J := [\lambda - 20v, \lambda + 20v]$.

Any vector $\boldsymbol{\xi} \in \mathbb{R}^d$ can be uniquely decomposed as $\boldsymbol{\xi} = \mathbf{n} + \mathbf{k}$ with $\mathbf{n} \in \Lambda^\dagger$ and $\mathbf{k} \in \Omega^\dagger$. We call $\mathbf{n} =: [\boldsymbol{\xi}]$ the 'integer part' of $\boldsymbol{\xi}$ and $\mathbf{k} =: \{\boldsymbol{\xi}\}$ the 'fractional part' of $\boldsymbol{\xi}$. For $\boldsymbol{\xi} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ we define $r = r(\boldsymbol{\xi}) := |\boldsymbol{\xi}|$ and $\boldsymbol{\xi}' := \boldsymbol{\xi}/|\boldsymbol{\xi}|$. For any $h \in L_2(\Omega)$ we introduce its Fourier coefficients

$$h_{\mathbf{n}} := (\mathrm{vol}\,\Omega)^{-1/2} \int_\Omega h(\mathbf{x}) \exp\big(-i\langle \mathbf{n}, \mathbf{x}\rangle\big)\, \mathrm{d}\mathbf{x}, \quad \mathbf{n} \in \Lambda^\dagger. \tag{3.2.7}$$

Given two positive functions f and g, we say that $f \gg g$, or $g \ll f$, or $g = O(f)$ if the ratio g/f is bounded. We say $f \asymp g$ if $f \gg g$ and $f \ll g$. Whenever we use O, o, $\gg$, $\ll$, or $\asymp$ notation, the constants involved can depend on d and norms of the potential in various Sobolev spaces H^s; the same is also the case when we use the expression 'sufficiently large'.

Let

$$\mathcal{A} := \Big\{\boldsymbol{\xi} \in \mathbb{R}^d,\ \big||\boldsymbol{\xi}|^2 - \lambda\big| \leqslant 40v\Big\}. \tag{3.2.8}$$

Notice that the definition of $\mathcal{A}$ obviously implies that if $\boldsymbol{\xi} \in \mathcal{A}$, then $\big||\boldsymbol{\xi}| - \rho\big| \ll \rho^{-1}$. We put

$$R = R(\rho) := \rho^{1/(36d^2(d+2))} \tag{3.2.9}$$

(so that the condition stated after equation (5.15) in [21] is satisfied). For $j \in \mathbb{N}$ let

$$\Theta'_j := \Lambda^\dagger \cap B(jR) \setminus \{\mathbf{0}\}.$$

Let $M := 5d^2 + 7d$. We introduce the set

$$\mathcal{B} := \Big\{\boldsymbol{\xi} \in \mathcal{A} \Big| \big|\langle \boldsymbol{\xi}, \boldsymbol{\eta}'\rangle\big| > \rho^{1/2},\ \text{for all } \boldsymbol{\eta} \in \Theta'_{6M}\Big\}. \tag{3.2.10}$$

In other words, $\mathcal{B}$ consists of all points $\boldsymbol{\xi} \in \mathcal{A}$ the projections of which to the directions of all vectors $\boldsymbol{\eta} \in \Theta'_{6M}$ have lengths larger than $\rho^{1/2}$. We also denote $\mathcal{D} := \mathcal{A} \setminus \mathcal{B}$.

The main result we will need follows from Corollary 7.15 of [21]:

Proposition 3.2.1. *There exist mappings $f, g : \mathcal{A} \to \mathbb{R}$ which satisfy the following properties:*

(i) $f(\boldsymbol{\xi})$ is an eigenvalue of $H(\mathbf{k})$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$; $\big|f(\boldsymbol{\xi}) - |\boldsymbol{\xi}|^2\big| \leqslant 2v$. f is an injection (if we count all eigenvalues with multiplicities) and all eigenvalues of $H(\mathbf{k})$ inside J are in the image of f.

(ii) If $\boldsymbol{\xi} \in \mathcal{A}$, then $\big|f(\boldsymbol{\xi}) - g(\boldsymbol{\xi})\big| \leqslant \rho^{-d-3}$.

(iii) For any $\boldsymbol{\xi} \in \mathcal{B}$

$$g(\boldsymbol{\xi}) = |\boldsymbol{\xi}|^2 + \sum_{j=1}^{2M} \sum_{\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_j \in \Theta'_M} \sum_{2 \leqslant n_1 + \cdots + n_j \leqslant 2M} C_{n_1, \ldots, n_j} \langle \boldsymbol{\xi}, \boldsymbol{\eta}_1 \rangle^{-n_1} \ldots \langle \boldsymbol{\xi}, \boldsymbol{\eta}_j \rangle^{-n_j}. \tag{3.2.11}$$

Remark 3.2.2. Formula (3.2.11) implies that

$$\partial g / \partial r(\boldsymbol{\xi}) \asymp \rho, \quad \text{for any} \quad \boldsymbol{\xi} \in \mathcal{B}. \tag{3.2.12}$$

For each positive $\delta \leqslant v$ we denote by $\mathcal{A}(\delta)$, $\mathcal{B}(\delta)$, and $\mathcal{D}(\delta)$ the intersections of $g^{-1}\big([\rho^2 - \delta, \rho^2 + \delta]\big)$ with $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{D}$, respectively.

It is proved in Lemma 8.1 of [21] that

$$\operatorname{vol}\big(\mathcal{D}(\delta)\big) \ll \rho^{d-7/3}\delta. \tag{3.2.13}$$

The following statement (Corollary 8.5 of [21]) gives a sufficient condition for the continuity of f:

Lemma 3.2.3. *There exists a constant C_1 with the following properties. Let*

$$I := \big\{\boldsymbol{\xi}(t) : t \in [t_{min}, t_{max}]\big\} \subset \mathcal{B}(v).$$

be a straight interval of length $L < \rho^{-1}\delta$. *Suppose that there is a point* $t_0 \in [t_{min}, t_{max}]$ *with the property that for each non-zero* $\mathbf{n} \in \Lambda^\dagger$ $g\big(\boldsymbol{\xi}(t_0) + \mathbf{n}\big)$ *is either outside the interval*

$$\Big[g\big(\boldsymbol{\xi}(t_0)\big) - C_1\rho^{-d-3} - C_1\rho L,\ g\big(\boldsymbol{\xi}(t_0)\big) + C_1\rho^{-d-3} + C_1\rho L\Big]$$

or not defined. Then $f\big(\boldsymbol{\xi}(t)\big)$ *is a continuous function of* t.

By inspection of the proof of Lemma 8.3 of [21] we obtain

Lemma 3.2.4. *For large enough* ρ *and* $\delta < \rho^{-1}$ *the following estimates hold uniformly over* $\mathbf{a} \in \Lambda^\dagger \setminus \{\mathbf{0}\}$*: if* $d \geqslant 3$,

$$\operatorname{vol}\Big(\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta) + \mathbf{a}\big)\Big) \ll (\delta^2\rho^{d-3} + \delta\rho^{-d}); \tag{3.2.14}$$

if $d = 2$,

$$\operatorname{vol}\Big(\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta) + \mathbf{a}\big)\Big) \begin{cases} \ll \delta^{3/2}, & |\mathbf{a}| \leqslant 2\rho - 1, \\ \ll \delta^{3/2} + \delta\rho^{-2}, & \big||\mathbf{a}| - 2\rho\big| < 1, \\ = 0, & |\mathbf{a}| \geqslant 2\rho + 1. \end{cases} \tag{3.2.15}$$

3.3 Prevalence of regular directions

In this section we prove that for most directions $\boldsymbol{\xi}'$ the image of the function f of Proposition 3.2.1 is an isolated eigenvalue of $H\big(\{\boldsymbol{\xi}\}\big)$ continuously depending on $|\xi|$ if it belongs to a neighborhood of ρ^2.

Lemma 3.3.1. *For* ρ *big enough and*

$$0 < \delta \leqslant \rho^{-d-3}$$

there exists a set $\mathcal{F} = \mathcal{F}(\rho)$ *on the unit sphere* S^{d-1} *in* $\mathbb{R}^d$ *with*

$$|\mathcal{F}|_\circ \geqslant \omega_d\big(1 - o(1)\big) \tag{3.3.1}$$

such that $f(\boldsymbol{\xi})$ *is a simple eigenvalue of* $H\big(\{\boldsymbol{\xi}\}\big)$ *continuously depending on* $r := |\boldsymbol{\xi}|$ *for every* $\boldsymbol{\xi} = (r, \boldsymbol{\xi}') \in f^{-1}\big([\rho^2 - \delta, \rho^2 + \delta)\big)$ *with* $\boldsymbol{\xi}' := \boldsymbol{\xi}/|\boldsymbol{\xi}| \in \mathcal{F}$.

Proof: It is enough to consider $\delta := \rho^{-d-3}$. For each $\boldsymbol{\xi}' \in S^{d-1}$ let

$$I_{\boldsymbol{\xi}'}(\delta) := \{r\boldsymbol{\xi}', r > 0\} \cap \mathcal{B}(\delta). \tag{3.3.2}$$

Let $\mathcal{F}_1 := \big\{\boldsymbol{\xi}' \in S^{d-1} \big| I_{\boldsymbol{\xi}'}(\delta) \neq \emptyset,\ \overline{I_{\boldsymbol{\xi}'}(\delta)} \cap \mathcal{D}(\delta) = \emptyset\big\}$.

For any $\boldsymbol{\eta} \in \Theta'_{6M}$ the area of the set of points $\boldsymbol{\xi}' \in S^{d-1}$ satisfying

$$\big|\langle r\boldsymbol{\xi}', \boldsymbol{\eta}'\rangle\big| \leqslant \rho^{1/2}$$

is evidently $O(\rho^{-1/2})$ if $r \geqslant \rho/2$ (the latter is true for all $r\boldsymbol{\xi}' \in \mathcal{A}$). Since the number of elements in Θ'_{6M} is $O(R^d)$, by (3.2.9) and (3.2.10) we have

$$|S^{d-1} \setminus \mathcal{F}_1|_\circ = o(1). \tag{3.3.3}$$

By definition $\mathcal{B}(\delta) = \mathcal{B} \cap g^{-1}\big([\rho^2 - \delta, \rho^2 + \delta]\big)$, hence (3.2.12) implies that for big ρ the length $l_{\boldsymbol{\xi}'}(\delta)$ of $I_{\boldsymbol{\xi}'}(\delta)$ satisfies

$$l_{\boldsymbol{\xi}'}(\delta) \asymp \delta\rho^{-1}, \quad \boldsymbol{\xi}' \in \mathcal{F}_1. \tag{3.3.4}$$

Let

$$\mathcal{F} := \big\{\boldsymbol{\xi}' \in \mathcal{F}_1 \,\big|\, f \text{ is continuous on } I_{\boldsymbol{\xi}'}(\delta)\big\},$$

and

$$\mathcal{E}(\delta) := \big\{\boldsymbol{\xi} \in \mathcal{B}(\delta) \big| \boldsymbol{\xi}' \in \mathcal{F}_1 \setminus \mathcal{F}\big\}.$$

Lemma 3.2.3 tells us that for each point $\boldsymbol{\xi} \in \mathcal{E}(\delta)$ there is a non-zero vector $\mathbf{n} \in \Lambda^\dagger$ such that

$$\big|g(\boldsymbol{\xi} + \mathbf{n}) - g(\boldsymbol{\xi})\big| \leqslant C_1\big(\rho^{-d-3} + \rho l_{\boldsymbol{\xi}'}(\delta)\big) \ll (\rho^{-d-3} + \delta). \tag{3.3.5}$$

Since $\big|g(\boldsymbol{\xi}) - \rho^2\big| \leqslant \delta$, this implies

$$\big|g(\boldsymbol{\xi} + \mathbf{n}) - \rho^2\big| \leqslant C_2(\rho^{-d-3} + \delta) =: \delta_1 \ll \rho^{-d-3} = \delta,$$

and thus $\boldsymbol{\xi} + \mathbf{n} \in \mathcal{A}(\delta_1)$; notice that $C_2 > 1$ and so $\delta_1 > \delta$. Therefore, each point $\boldsymbol{\xi} \in \mathcal{E}(\delta)$ also belongs to the set $\big(\mathcal{A}(\delta_1) - \mathbf{n}\big)$ for a non-zero $\mathbf{n} \in \Lambda^\dagger$; obviously, $|\mathbf{n}| \ll \rho$. In other words,

$$\mathcal{E}(\delta) \subset \bigcup_{\mathbf{n} \in \Lambda^\dagger \cap B(C\rho),\, \mathbf{n} \neq 0} \big(\mathcal{A}(\delta_1) - \mathbf{n}\big) = \bigcup_{\mathbf{n} \neq 0} \big(\mathcal{B}(\delta_1) - \mathbf{n}\big) \cup \bigcup_{\mathbf{n} \neq 0} \big(\mathcal{D}(\delta_1) - \mathbf{n}\big). \tag{3.3.6}$$

To proceed further, we need more notation. Denote $\mathcal{D}_0(\delta_1)$ to be the set of all points $\boldsymbol{\nu}$ from $\mathcal{D}(\delta_1)$ for which there is no non-zero $\mathbf{n} \in \Lambda^\dagger$ satisfying $\boldsymbol{\nu} - \mathbf{n} \in \mathcal{B}(\delta)$; $\mathcal{D}_1(\delta_1)$ to be the set of all points $\boldsymbol{\nu}$ from $\mathcal{D}(\delta_1)$ for which there is a unique non-zero $\mathbf{n} \in \Lambda^\dagger$ satisfying $\boldsymbol{\nu} - \mathbf{n} \in \mathcal{B}(\delta)$; and $\mathcal{D}_2(\delta_1)$ to be the rest of the points from $\mathcal{D}(\delta_1)$ (i.e. $\mathcal{D}_2(\delta_1)$ consists of all points $\boldsymbol{\nu}$ from $\mathcal{D}(\delta_1)$ for which there exist at least two different non-zero vectors $\mathbf{n}_1, \mathbf{n}_2 \in \Lambda^\dagger$ satisfying $\boldsymbol{\nu} - \mathbf{n}_j \in \mathcal{B}(\delta)$). Then Lemma 8.7 of [21] implies that we can rewrite (3.3.6) as

$$\mathcal{E}(\delta) \subset \bigcup_{\mathbf{n}\neq\mathbf{0}} \big(\mathcal{B}(\delta_1) - \mathbf{n}\big) \cup \bigcup_{\mathbf{n}\neq\mathbf{0}} \big(\mathcal{D}_1(\delta_1) - \mathbf{n}\big). \tag{3.3.7}$$

From this we conclude that

$$\mathcal{E}(\delta) \subset \bigcup_{\mathbf{n}\neq\mathbf{0}} \Big(\big(\mathcal{B}(\delta_1) - \mathbf{n}\big) \cap \mathcal{B}(\delta)\Big) \cup \bigcup_{\mathbf{n}\neq\mathbf{0}} \Big(\big(\mathcal{D}_1(\delta_1) - \mathbf{n}\big) \cap \mathcal{B}(\delta)\Big), \tag{3.3.8}$$

since $\mathcal{E}(\delta) \subset \mathcal{B}(\delta)$.

The definition of the set $\mathcal{D}_1(\delta_1)$ and (3.2.13) imply that

$$\begin{aligned} &\operatorname{vol}\Bigg(\bigcup_{\mathbf{n}\neq\mathbf{0}} \Big(\big(\mathcal{D}_1(\delta_1) - \mathbf{n}\big) \cap \mathcal{B}(\delta)\Big)\Bigg) \leqslant \operatorname{vol}\big(\mathcal{D}_1(\delta_1)\big) \\ &\leqslant \operatorname{vol}\big(\mathcal{D}(\delta_1)\big) \ll \delta_1 \rho^{d-7/3} \ll \delta \rho^{d-7/3}. \end{aligned} \tag{3.3.9}$$

For $d \geqslant 3$ Lemma 3.2.4, inequality $\delta < \delta_1$, and the fact that the union in (3.3.8) consists of no more than $C\rho^d$ terms imply

$$\operatorname{vol}\Bigg(\bigcup_{\mathbf{n}\neq\mathbf{0}} \Big(\big(\mathcal{B}(\delta_1) - \mathbf{n}\big) \cap \mathcal{B}(\delta)\Big)\Bigg) \ll \rho^d(\delta_1^2 \rho^{d-3} + \delta_1 \rho^{-d}) \ll \delta(\rho^{d-6} + 1). \tag{3.3.10}$$

For $d = 2$ we obtain by Lemma 3.2.4

$$\begin{aligned} &\operatorname{vol}\Bigg(\bigcup_{\mathbf{n}\in\Lambda^\dagger\setminus\{\mathbf{0}\}} \Big(\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta_1) + \mathbf{n}\big)\Big)\Bigg) \\ &\leqslant \sum_{\substack{\mathbf{n}\in\Lambda^\dagger\setminus\{\mathbf{0}\} \\ |\mathbf{n}|\leqslant 2\rho - 1}} \operatorname{vol}\Big(\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta_1) + \mathbf{n}\big)\Big) \\ &+ \sum_{\substack{\mathbf{n}\in\Lambda^\dagger\setminus\{\mathbf{0}\} \\ ||\mathbf{n}| - 2\rho| < 1}} \operatorname{vol}\Big(\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta_1) + \mathbf{n}\big)\Big) \\ &\ll \delta_1^{3/2} \rho^2 + \rho(\delta_1^{3/2} + \delta_1 \rho^{-2}) \ll \delta \rho^{-1/2}, \end{aligned} \tag{3.3.11}$$

where we have used that

$$\#\Big\{\mathbf{n}\in\Lambda^\dagger\Big|\big||\mathbf{n}|-2\rho\big|<1\Big\}\ll\rho.$$

Applying (3.3.9), (3.3.10), and (3.3.11) to (3.3.8) we obtain for all $d\geqslant 2$

$$\operatorname{vol}\mathcal{E}(\delta)\ll\delta\rho^{d-7/3}. \tag{3.3.12}$$

By definition,

$$\mathcal{E}(\delta)=\bigcup_{\boldsymbol{\xi}'\in\mathcal{F}_1\setminus\mathcal{F}} I_{\boldsymbol{\xi}'}(\delta).$$

Hence by (3.3.4)

$$|\mathcal{F}_1\setminus\mathcal{F}|_\circ\ll\delta^{-1}\rho^{2-d}\operatorname{vol}\mathcal{E}(\delta). \tag{3.3.13}$$

Combining (3.3.12) and (3.3.13) we conclude that for big ρ

$$|\mathcal{F}_1\setminus\mathcal{F}|_\circ=o(1). \tag{3.3.14}$$

We have

$$|S^{d-1}\setminus\mathcal{F}|_\circ=|S^{d-1}\setminus\mathcal{F}_1|_\circ+|\mathcal{F}_1\setminus\mathcal{F}|_\circ. \tag{3.3.15}$$

Substituting (3.3.3) and (3.3.14) into (3.3.15) we obtain (3.3.1).

Now we notice that for every $\boldsymbol{\xi}'\in\mathcal{F}$ the interval $I_{\boldsymbol{\xi}'}(\delta)$ has the following property: for each point $\boldsymbol{\xi}\in I_{\boldsymbol{\xi}'}(\delta)$ and each non-zero vector $\mathbf{n}\in\Lambda^\dagger$ such that $\boldsymbol{\xi}+\mathbf{n}\in\mathcal{A}$ we have $\big|g(\boldsymbol{\xi}+\mathbf{n})-g(\boldsymbol{\xi})\big|>2\rho^{-d-3}$. This implies $f(\boldsymbol{\xi}+\mathbf{n})-f(\boldsymbol{\xi})\neq 0$. Therefore, $f(\boldsymbol{\xi})$ is a simple eigenvalue of $H\big(\{\boldsymbol{\xi}\}\big)$ for each $\boldsymbol{\xi}\in I_{\boldsymbol{\xi}'}(\delta)$. The lemma is proved.

3.4 Some properties of operators on the fibers

In this section we discuss some properties of operators $H(\mathbf{k})$, $\mathbf{k}\in\Omega^\dagger$. In Lemma 3.4.1 we prove that the Fourier coefficients of the eigenfunctions of these operators satisfy certain decay estimates if the corresponding eigenvalues are big enough. Using this, we obtain an estimate on the rate of change of such eigenvalues with $\mathbf{k}$ in Lemma 3.4.2.

For $m \in \mathbb{R}$ let

$$V^{(m)} := \Big(\sum_{\mathbf{n}\in\Lambda^\dagger} |\mathbf{n}|^{2m} |V_\mathbf{n}|^2 \Big)^{1/2}.$$

Since V is smooth, $V^{(m)}$ is finite for any $m \geqslant 0$. Recall that Q is defined by (3.2.2).

Lemma 3.4.1. *Fix $m \in \mathbb{N}$ and $\varkappa \in (0,1)$. For $\mathbf{k} \in \Omega^\dagger$ let ψ be a normalized eigenfunction of $H(\mathbf{k})$:*

$$H(\mathbf{k})\psi = \zeta\psi \tag{3.4.1}$$

with the eigenvalue

$$\zeta \geqslant \max\big\{36Q^2\varkappa^{-2}, (1+m\varkappa)^{2/(d-1)}\varkappa^{-2d/(d-1)}\big\}. \tag{3.4.2}$$

Then there exists $M_m = M_m(d,\Lambda,V) \in \mathbb{R}_+$ such that for all $\mathbf{n} \in \Lambda^\dagger$ with

$$|\mathbf{n}| \geqslant (1+m\varkappa)\sqrt{\zeta} \tag{3.4.3}$$

the Fourier coefficients of ψ satisfy

$$|\psi_\mathbf{n}| < M_m \varkappa^{-m} |\mathbf{n}|^{-(3m+1)/2}. \tag{3.4.4}$$

Proof: We proceed by induction. Suppose that either $m = 1$, or $m > 1$ and the statement is proved for $m-1$. Substituting the Fourier series

$$\psi(\mathbf{x}) = (\operatorname{vol}\Omega)^{-1/2} \sum_{\mathbf{n}\in\Lambda^\dagger} \psi_\mathbf{n} \exp\big(i\langle \mathbf{n},\mathbf{x}\rangle\big), \quad \mathbf{x} \in \Omega$$

into (3.4.1) and equating the coefficients at $\exp\big(i\langle\mathbf{n},\mathbf{x}\rangle\big)$ on both sides, we obtain by (3.2.5):

$$|\mathbf{n}+\mathbf{k}|^2\psi_\mathbf{n} + \sum_{\mathbf{l}\in\Lambda^\dagger} V_{\mathbf{n}-\mathbf{l}}\psi_\mathbf{l} = \zeta\psi_\mathbf{n}. \tag{3.4.5}$$

Since $|\mathbf{k}| \leqslant Q$, by (3.4.2) and (3.4.3) we have

$$2|\mathbf{n}||\mathbf{k}| \leqslant \varkappa|\mathbf{n}|^2/6 + 6\varkappa^{-1}Q^2 \leqslant \varkappa|\mathbf{n}|^2/3. \tag{3.4.6}$$

For $\varkappa \in (0,1)$, it follows from (3.4.3) that

$$|\mathbf{n}|^2 - \zeta \geqslant \big(1-(1+\varkappa)^{-2}\big)|\mathbf{n}|^2 = \varkappa(2+\varkappa)(1+\varkappa)^{-2}|\mathbf{n}|^2 \geqslant \varkappa|\mathbf{n}|^2/2. \tag{3.4.7}$$

Combining (3.4.6) and (3.4.7) we obtain

$$|\mathbf{n}+\mathbf{k}|^2-\zeta \geqslant |\mathbf{n}|^2-2|\mathbf{n}||\mathbf{k}|-\zeta \geqslant \varkappa|\mathbf{n}|^2/6,$$

and thus by (3.4.5)

$$|\psi_{\mathbf{n}}| < 6\varkappa^{-1}|\mathbf{n}|^{-2}\sum_{\mathbf{l}\in\Lambda^\dagger}|V_{\mathbf{n}-\mathbf{l}}\psi_{\mathbf{l}}|. \tag{3.4.8}$$

If $m=1$ we estimate the sum on the r. h. s. by $V^{(0)}$ using Cauchy–Schwarz inequality (since ψ is normalized) and obtain (3.4.4) with $M_1 := 6V^{(0)}$.

If $m>1$, we estimate

$$\sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}-\mathbf{n}|\leqslant|\mathbf{n}|^{1/d}}|V_{\mathbf{n}-\mathbf{l}}\psi_{\mathbf{l}}| \leqslant \sup_{\mathbf{m}:\,|\mathbf{m}|\geqslant|\mathbf{n}|-|\mathbf{n}|^{1/d}}|\psi_{\mathbf{m}}|\sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}|\leqslant|\mathbf{n}|^{1/d}}|V_{\mathbf{l}}|. \tag{3.4.9}$$

By (3.4.3), (3.4.2), and monotonicity of the function $q(t)=t-t^{1/d}$ for $t>1$ we have

$$|\mathbf{n}|-|\mathbf{n}|^{1/d} \geqslant (1+m\varkappa)\sqrt{\zeta}-\big((1+m\varkappa)\sqrt{\zeta}\big)^{1/d} \geqslant \big(1+(m-1)\varkappa\big)\sqrt{\zeta}.$$

According to the induction hypothesis

$$\sup_{\mathbf{m}:\,|\mathbf{m}|\geqslant|\mathbf{n}|-|\mathbf{n}|^{1/d}}|\psi_{\mathbf{m}}| \leqslant \varkappa^{1-m}M_{m-1}\big(1-|\mathbf{n}|^{(1-d)/d}\big)^{1-3m/2}|\mathbf{n}|^{1-3m/2}. \tag{3.4.10}$$

Since $\varkappa\in(0,1)$, from (3.4.3) and (3.4.2) we conclude

$$|\mathbf{n}| \geqslant (1+m\varkappa)\sqrt{\zeta} \geqslant (\varkappa^{-1}+m)^{d/(d-1)} > 2^{d/(d-1)},$$

hence

$$\big(1-|\mathbf{n}|^{(1-d)/d}\big)^{1-3m/2} < 2^{3m/2-1}. \tag{3.4.11}$$

Let

$$W := \sup_{r>1} r^{-d}\#\big\{\mathbf{l}\in\Lambda^\dagger\big||\mathbf{l}|\leqslant r\big\}.$$

Clearly, $W<\infty$. By Cauchy–Schwarz inequality

$$\sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}|\leqslant|\mathbf{n}|^{1/d}}|V_{\mathbf{l}}| \leqslant W^{1/2}V^{(0)}|\mathbf{n}|^{1/2}. \tag{3.4.12}$$

Substituting (3.4.10), (3.4.11), and (3.4.12) into (3.4.9) we get

$$\sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}-\mathbf{n}|\leqslant|\mathbf{n}|^{1/d}}|V_{\mathbf{n}-\mathbf{l}}\psi_{\mathbf{l}}| < 2^{3m/2-1}\varkappa^{1-m}W^{1/2}V^{(0)}M_{m-1}|\mathbf{n}|^{3(1-m)/2}. \tag{3.4.13}$$

On the other hand, since $\|\psi\| = 1$, applying Cauchy–Schwarz inequality we obtain

$$\begin{aligned}
\sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}-\mathbf{n}|>|\mathbf{n}|^{1/d}} |V_{\mathbf{n}-\mathbf{l}}\psi_{\mathbf{l}}| &< |\mathbf{n}|^{3(1-m)/2} \sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}|>|\mathbf{n}|^{1/d}} |\mathbf{l}|^{3(m-1)d/2}|V_{\mathbf{l}}||\psi_{\mathbf{n}-\mathbf{l}}| \\
&\leqslant |\mathbf{n}|^{3(1-m)/2}\Big(\sum_{\mathbf{l}\in\Lambda^\dagger:\,|\mathbf{l}|>|\mathbf{n}|^{1/d}} |\mathbf{l}|^{3(m-1)d}|V_{\mathbf{l}}|^2\Big)^{1/2} \leqslant V^{(3(m-1)d/2)}|\mathbf{n}|^{3(1-m)/2}.
\end{aligned} \tag{3.4.14}$$

Inserting (3.4.13) and (3.4.14) into (3.4.8) we arrive at (3.4.4) with

$$M_m := 6(2^{3m/2-1}W^{1/2}V^{(0)}M_{m-1} + V^{(3(m-1)d/2)}).$$

Lemma 3.4.2. *For any $\eta \in (0,1)$ there exists $\zeta_0 > 0$ such that if $\zeta(\mathbf{k}) \geqslant \zeta_0$ is a simple eigenvalue of $H(\mathbf{k})$ for some $\mathbf{k} \in \Omega^\dagger$ then*

$$|\nabla_{\mathbf{k}}\zeta| \leqslant 2(1+\eta)\sqrt{\zeta}. \tag{3.4.15}$$

Proof: Let $\psi(\mathbf{k})$ be the eigenfunction corresponding to $\zeta(\mathbf{k})$ with

$$\big\|\psi(\mathbf{k})\big\| = 1. \tag{3.4.16}$$

Then

$$\nabla_{\mathbf{k}}\zeta(\mathbf{k}) = \nabla_{\mathbf{k}}\big(\psi(\mathbf{k}), H(\mathbf{k})\psi(\mathbf{k})\big) = \Big(\psi(\mathbf{k}), \big(\nabla_{\mathbf{k}}H(\mathbf{k})\big)\psi(\mathbf{k})\Big). \tag{3.4.17}$$

By (3.2.5) and (3.2.3),

$$\nabla_{\mathbf{k}}H(\mathbf{k}) = 2\mathbf{D}(\mathbf{k}).$$

Substituting this into (3.4.17) we obtain:

$$\big|\nabla_{\mathbf{k}}\zeta(\mathbf{k})\big| \leqslant 2\big\|\mathbf{D}(\mathbf{k})\psi(\mathbf{k})\big\| = 2\Big(\sum_{\mathbf{n}\in\Lambda^\dagger} |\mathbf{n}+\mathbf{k}|^2\big|\psi_{\mathbf{n}}(\mathbf{k})\big|^2\Big)^{1/2}. \tag{3.4.18}$$

Let

$$m := \big[(d+1)/3\big] + 1 \tag{3.4.19}$$

and

$$\varkappa := \eta/(2m+1). \tag{3.4.20}$$

We assume that

$$\zeta := \zeta(\mathbf{k}) \geqslant \max\big\{36Q^2\varkappa^{-2}, (1+m\varkappa)^{2/(d-1)}\varkappa^{-2d/(d-1)}\big\}. \tag{3.4.21}$$

Since by (3.2.2) $|\mathbf{k}| \leqslant Q$, by (3.4.16), (3.4.21), and (3.4.20) we have

$$\sum_{|\mathbf{n}|<(1+m\varkappa)\sqrt{\zeta}} |\mathbf{n}+\mathbf{k}|^2 \big|\psi_{\mathbf{n}}(\mathbf{k})\big|^2 < \big(1+(m+1/6)\varkappa\big)^2 \zeta < (1+\eta/2)^2 \zeta. \qquad (3.4.22)$$

For $|\mathbf{n}| \geqslant (1+m\varkappa)\sqrt{\zeta}$ we apply Lemma 3.4.1 obtaining

$$\sum_{|\mathbf{n}|\geqslant(1+m\varkappa)\sqrt{\zeta}} |\mathbf{n}+\mathbf{k}|^2 \big|\psi_{\mathbf{n}}(\mathbf{k})\big|^2 \leqslant M_m^2 \varkappa^{-2m} \sum_{|\mathbf{n}|\geqslant(1+m\varkappa)\sqrt{\zeta}} |\mathbf{n}+\mathbf{k}|^2 \big|\mathbf{n}\big|^{-3m-1}. \qquad (3.4.23)$$

By (3.4.19) the r. h. s. of (3.4.23) is finite and is $O(\zeta^{-1/2})$. Thus, choosing ζ_0 big enough, by (3.4.18), (3.4.22), and (3.4.23) we obtain (3.4.15).

3.5 Proof of Theorem 3.1.1

We are now ready to finish the proof of the main result. It is enough to prove

Theorem 3.5.1. *For any $\alpha \in (0,1)$ there exists $\rho_0 > 0$ big enough such that for all $\rho \geqslant \rho_0$*

$$\mathbf{N}(\rho^2+\delta) - \mathbf{N}(\rho^2-\delta) \geqslant (1-\alpha)(2\pi)^{-d}\omega_d \delta \rho^{d-2} \qquad (3.5.1)$$

for any

$$0 < \delta \leqslant \rho^{-d-3}. \qquad (3.5.2)$$

Indeed, the original statement of Theorem 3.1.1 can be obtained by partitioning of the interval $[\lambda, \lambda+\varepsilon]$ into subintervals with lengths not exceeding $2\lambda^{(-d-3)/2}$ and adding up estimates (3.5.1) on these subintervals (with ρ^2 being respective middle points).

Proof: We first express the growth of IDS in terms of the function f of Proposition 3.2.1(i) using (3.2.6):

$$\mathbf{N}(\rho^2+\delta) - \mathbf{N}(\rho^2-\delta) = (2\pi)^{-d} \operatorname{vol}\big(f^{-1}[\rho^2-\delta, \rho^2+\delta)\big). \qquad (3.5.3)$$

We can write

$$\operatorname{vol}\big(f^{-1}[\rho^2-\delta, \rho^2+\delta)\big) = \int_{S^{d-1}} \int_0^\infty \chi(r, \boldsymbol{\xi}') r^{d-1} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{\xi}', \qquad (3.5.4)$$

where χ is the indicator function of $f^{-1}\big([\rho^2-\delta,\rho^2+\delta)\big)$. To obtain a lower bound we can restrict the integration in (3.5.4) to $\boldsymbol{\xi}'\in\mathcal{F}$ defined in Lemma 3.3.1. Then for any $\eta\in(0,1)$ there exists $\rho_0>0$ such that for any $\rho\geqslant\rho_0$ we have

$$|\mathcal{F}|_\circ\geqslant(1-\eta)\omega_d, \tag{3.5.5}$$

and for any $\boldsymbol{\xi}'\in\mathcal{F}$ the support of $\chi(\cdot,\boldsymbol{\xi}')$ contains an interval $[r_1,r_2]$ with

$$(1-\eta)\rho\leqslant r_1<r_1+(1-\eta)\rho^{-1}\delta\leqslant r_2. \tag{3.5.6}$$

Indeed, the first inequality in (3.5.6) follows from Proposition 3.2.1(ii),(iii). The last inequality in (3.5.6) follows from Lemmata 3.3.1 and 3.4.2.

Thus for all $\rho\geqslant\rho_0$ by (3.5.5) and (3.5.6) we obtain

$$\begin{aligned}\int_{S^{d-1}}\int_0^\infty\chi(r,\boldsymbol{\xi}')r^{d-1}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}'&\geqslant\int_{\mathcal{F}}(1-\eta)^d\rho^{d-2}\delta\,\mathrm{d}\boldsymbol{\xi}'\\&\geqslant(1-\eta)^{d+1}\omega_d\rho^{d-2}\delta,\end{aligned} \tag{3.5.7}$$

Combining (3.5.3), (3.5.4), and (3.5.7), and choosing η small enough we arrive at (3.5.1). The theorem is proved.

Bibliography

[1] G. Barbatis and L. Parnovski, *Bethe-Sommerfeld conjecture for pseudo-differential perturbation*, Comm.P.D.E. **34**, Issue 4(2009), 383 – 418.

[2] A. Sommerfeld and H. Bethe, *Elektronentheorie der Metalle*, in H. Geiger and K. Scheel, eds., Handbuch der Physik, **24**, Part 2, 333-622 (1933). Later edition: *Elektronentheorie der Metalle*, Springer, 1967.

[3] J.W.S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1959.

[4] B.E.J. Dahlberg, E. Trubowitz, *A remark on two dimensional periodic potentials*, Comment. Math. Helvetici **57** (1982), 130–134.

[5] M.S.P. Eastham, *The spectral theory of periodic differential equations*, Scottish Academice Press, 1973.

[6] J. Feldman, H. Knörrer, E. Trubowitz, *The perturbatively stable spectrum of a periodic Schrödinger operator*, Invent. Math. **100** (1990), 259–300.

[7] J. Feldman, H. Knörrer, E. Trubowitz, *Perturbatively unstable eigenvalues of a periodic Schrödinger operator*, Comment. Math. Helvetici **66** (1991), 557-579.

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

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

[10] Yu. Karpeshina, *Analytic perturbation theory for a periodic potential*, Izv. Akad. Nauk SSSR Ser. Mat., 53 (1989), No 1, 45-65; English transl.: Math. USSR Izv., **34** (1990), No 1, 43 – 63.

[11] Yu. Karpeshina, *Perturbation series for the Schrödinger operator with a periodic potential near planes of diffraction*, Comm. Anal. Geom. **4** (1996), 339–413.

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

[13] Yu. Karpeshina, *On the density of states for the periodic Schrödinger operator*, Ark. Mat. **38** (2000), no. 1, 111–137.

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

[15] Yu. Karpeshina, Y.-R.Lee *Spectral Properties of Polyharmonic Operators with Limit-Periodic Potential in Dimension Two*, Journal d'Analyse Mathematique, **102** (2007), 225 – 310.

[16] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin, 1980.

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

[18] P. Kuchment, *Floquet theory for partial differential equations*, Birkhäuser, Basel, 1993.

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

[20] S. Morozov, L. Parnovski, and I. Pchelintseva, *Lower bound on the density of states for periodic Schrödinger operators*, Operator theory and its applications: in memory of V. B. Lidskii (1924-2008), **231**, (2010), 161-172

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

[22] L. Parnovski and A. Sobolev, *Bethe-Sommerfeld conjecture for polyharmonic operators*, Duke Math. J. **107** no.2 (2001), 209–238.

[23] L. Parnovski and A. Sobolev, *Perturbation theory and the Bethe-Sommerfeld conjecture*, Annales H. Poincaré **2**(2001), 573–581.

[24] L. Parnovski, A. Sobolev, *Bethe-Sommerfeld conjecture for periodic operators with strong perturbations*, Inventiones Mathematicae, **181** (2010), 467–540.

[25] L. Parnovski and R. Shterenberg, *Asymptotic expansion of the integrated density of states of a two-dimensional periodic Schrödinger operator*, Invent. Math. **176** (2009), no. 2, 275–323.

[26] L. Parnovski and R. Shterenberg, *Complete asymptotic expansion of the integrated density of states of multidimensional almost-periodic Schrödinger operators*, to appear in Ann. of Math.

[27] V.N. Popov and M. Skriganov, *A remark on the spectral structure of the two dimensional Schrödinger operator with a periodic potential*, Zap. Nauchn. Sem. LOMI AN SSSR **109** (1981), 131–133 (Russian).

[28] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, IV*, Academic Press, New York, 1975.

[29] G. Rozenbljum, *Near-similarity of operators and the spectral asymptotic behavior of pseudodifferential operators on the circle*, Trudy Moskov. Mat. Obshch. **36** (1978), 59–84 (Russian).

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

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

[32] M. Shubin, *Spectral theory and the index of elliptic operators with almost-periodic coefficients*, Uspekhi Mat. Nauk **34** (1979), no. 2(206), 95–135.

[33] M. Skriganov, *Proof of the Bethe-Sommerfeld conjecture in dimension two*, Soviet Math. Dokl. **20** no. 1 (1979), 89–90.

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

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

[36] M. Skriganov and A. Sobolev, *Asymptotic estimates for spectral bands of periodic Schrödinger operators*, St Petersburg Math. J. **17** no. 1 (2006), 207–216.

[37] M. Skriganov and A. Sobolev, *Variation of the number of lattice points in large balls*, Acta Arith. **120**(3)(2005), 245–267.

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

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

[40] A. Sobolev, *Recent results on the Bethe-Sommerfeld conjecture*, Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon's 60th birthday, ix–xii, Proc. Sympos. Pure Math., **76**, Part 1, Amer. Math. Soc., Providence, RI, 2007.

[41] O. Veliev, *Asymptotic formulas for the eigenvalues of the periodic Schrödinger operator and the Bethe-Sommerfeld conjecture*, Functional Anal. Appl. **21** no. 2 (1987), 87–100.

[42] O. Veliev, *The spectrum of multidimensional periodic operators*, Teor. Funktsi **49** (1988), 17–34.

[43] O. Veliev, *Perturbation theory for the periodic multidimensional Schrödinger operator and the Bethe-Sommerfeld Conjecture*, Int. J. Contemp. Math. Sci., **2** no. 2 (2007), 19–87.

[44] N. Yakovlev, *Asymptotic estimates of the densities of lattices k-packing and k-covering, and the structure of the spectrum of the Schrödinger operator with a periodic potential*, Dokl. Akad. Nauk SSSR, **276** (1984); English transl. in Sovit Math. Dokl. **29** (1984), 457–460

[45] N. Yakovlev, *On spectra of multi-dimensional pseudo-differential periodic operators*, Vestn. Mosk. Uni. Ser. 1 Mat, Mekh, **3**, 80-81 (1985) (Russian)