BETHE-SOMMERFELD CONJECTURE

LEONID PARNOVSKI

ABSTRACT. We consider Schrödinger operator $-\Delta + V$ in $\mathbb{R}^d$ ($d \geq 2$) with smooth periodic potential V and prove that there are only finitely many gaps in its spectrum.

Dedicated to the memory of B.M.Levitan

1. INTRODUCTION

This paper is devoted to proving the Bethe-Sommerfeld conjecture which states that number of gaps in the spectrum of a Schrödinger operator

$$-\Delta + V(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^d \tag{1.1}$$

with a periodic potential V is finite whenever $d \geq 2$. We prove the conjecture for smooth potentials in all dimensions greater than one and for arbitrary lattices of periods. The conjecture so far was proved by V.Popov and M.Skriganov [9] (see also [11]) in dimension 2, by M.Skriganov [12], [13] in dimension 3, and by B.Helffer and A.Mohamed [3] in dimension 4; M.Skriganov [12] has also shown the conjecture to hold in arbitrary dimension under the assumption that the lattice of periods is rational. In the case $d = 3$ the conjecture was proved in [5] for non-smooth or even singular potentials (admitting Coulomb and even stronger singularities). An interesting approach to proving the conjecture was presented by O.A.Veliev in [15].

There is a number of problems closely related to the Bethe-Sommerfeld conjecture on which extensive work has been done; the relevant publications include, but are by no means restricted to, [2], [5] (and references therein), [7], [8]. Methods used to tackle these problems range from number theory ([12], [13], [7], [8]) to microlocal analysis in [3] and perturbation theory in [5], [14] and [15]. The approach used in the present paper consists, mostly, of perturbation theoretical arguments with a bit of geometry and geometrical combinatorics thrown in at the end.

There are certain parallels between the approach of our paper and the approach used in [15]. In particular, there are several important intermediate results in our paper and in [15] which look rather similar to each other. Examples of such similarities are: precise asymptotic formulae for eigenvalues in the non-resonance regions and some, although not very precise, formulae in the resonance regions; proving that the eigenvalue is simple when we move the dual parameter $\boldsymbol{\xi}$ along a certain short interval, and, finally, the use of geometrical combinatorics. However, here the similarities end, because the detailed methods used on each step are completely different. For example, paper [15] makes a heavy use of the asymptotic formulae for the eigenfunctions, whereas in our paper they are not needed. On the other hand, we prove that each eigenvalue close to λ is described by exactly one asymptotic formula (i.e. the mapping f constructed in our paper is a bijection in a certain sense), and this plays an essential role in our proof, but in [15] this property is not required at all. In [15] a very important role is played by the isoenergetic surface, whereas we don't need it. This list can be continued, but it is probably better to stop here and state once again: the methods of [15] and our paper are different, despite the similarity of some intermediate results. It is also worthwhile to mention that asymptotic expressions for eigenfunctions as well as asymptotic formulae for isoenergetic surfaces were obtained by Yu.Karpeshina (see for example [5]).

In many of the papers mentioned above, proving the conjecture in special cases comes together with obtaining lower bounds for either of the functions describing the band structure of the spectrum: the multiplicity of overlapping $m(\lambda)$ and the overlapping function $\zeta(\lambda)$ (we will give a definition of these functions in the next section). For example, in dimensions $d = 2, 3, 4$ it has been proved in [12], [13], [2], and [8] that for large λ we have

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

and

$$\zeta(\lambda) \gg \lambda^{\frac{3-d}{4}}; \qquad (1.2)$$

however, these estimates do not seem likely to hold in high dimensions. The estimates of the present paper are rather weaker, but they hold in all dimensions. Unfortunately, our approach does not allow to say anything stronger than $m(\lambda) \geq 1$ for large λ (this inequality is equivalent to the finiteness of the number of spectral gaps). However, it is possible to give a nontrivial lower bound for the overlapping function:

we will show that in all dimensions for sufficiently large λ

$$\zeta(\lambda) \gg \lambda^{\frac{1-d}{2}}. \tag{1.3}$$

The rest of the introduction is devoted to the informal discussion of the proof. Since the proof of the main theorem 2.1 is rather complicated and technically involved, the major ideas are outlined here.

After an affine change of coordinates, we can re-write our operator (1.1) as

$$H = H_0 + V(\mathbf{x}), \qquad H_0 = \mathbf{DGD}, \tag{1.4}$$

with the potential V being smooth and periodic with the lattice of periods of V equal $(2\pi\mathbb{Z})^d$ ($\mathbf{D} = i\nabla$ and $\mathbf{G} = \mathbf{F}^2$ is $d \times d$ positive matrix, where the matrix $\mathbf{F}$ is also assumed to be positive). Without loss of generality, we assume that the average of the potential V over the cell $[0, 2\pi]^d$ is zero (otherwise we simply subtract this average from the potential, which will just shift the spectrum of the problem). Let us fix a sufficiently large value of spectral parameter $\lambda = \rho^2$; we will prove that λ is inside the spectrum of H.

The first step of the proof, as usual, consists in performing the Floquet-Bloch decomposition to our operator (1.4):

$$H = \int_\oplus H(\mathbf{k})d\mathbf{k}, \tag{1.5}$$

where $H(\mathbf{k}) = H_0 + V(\mathbf{x})$ is the family of 'twisted' operators with the same symbol as H acting in $L^2(T^d)$ where $T^d := \{\mathbf{x} \in \mathbb{R}^d, |x_j| \leq \pi, 1 \leq j \leq d\}$. The domain $\mathfrak{D}(\mathbf{k})$ of $H(\mathbf{k})$ consists of functions $f \in H^2(T^d)$ satisfying the boundary conditions $f\big|_{x_j=\pi} = e^{i2\pi k_j} f\big|_{x_j=-\pi}$, $\frac{\partial f}{\partial x_j}\big|_{x_j=\pi} = e^{i2\pi k_j} \frac{\partial f}{\partial x_j}\big|_{x_j=-\pi}$. These auxiliary operators are labelled by the quasi-momentum $\mathbf{k} \in \mathbb{R}^d/\mathbb{Z}^d$; see [10] for more details about this decomposition. The next step is to assume that the potential V is a finite trigonometric polynomial whose Fourier coefficients $\hat{V}(\mathbf{m})$ vanish when $|\mathbf{m}| > R$. The justification of the fact that it is enough to prove the conjecture in this case is not too difficult once we keep careful control of the dependence of all the estimates on R. The main part of the argument consists of finding an asymptotic formula for all sufficiently large eigenvalues of all operators $H(\mathbf{k})$, with an arbitrarily small power of the energy in the remainder estimate. In order to be able to write such a formula, however, we have to abandon the traditional way of labelling eigenvalues of each $H(\mathbf{k})$ in the non-decreasing order. Instead, we will label eigenvalues by means of the integer vectors $\mathbf{n} \in \mathbb{Z}^d$. Consider, for example, the unperturbed operator $H_0(\mathbf{k})$. Its eigenfunctions

and eigenvalues are

$$\{e^{i(\mathbf{n}+\mathbf{k})\mathbf{x}}\}_{\mathbf{n}\in\mathbb{Z}^d}$$

and

$$\{|\mathbf{F}(\mathbf{n}+\mathbf{k})|^2\}_{\mathbf{n}\in\mathbb{Z}^d} \tag{1.6}$$

correspondingly. However, despite our precise knowledge of eigenvalues, it is extremely difficult to write them in increasing order or, indeed, even to derive the one-term asymptotic formula for the j-th eigenvalue with the precise remainder estimate. It is rather convenient to introduce one parameter which takes care of both the quasi-momentum $\mathbf{k}$ and the integer vector $\mathbf{n}$ which labels eigenvalues in (1.6). We denote $\boldsymbol{\xi} := \mathbf{n}+\mathbf{k}$ (notation indicates that $\boldsymbol{\xi}$ can be thought of as being a dual variable) so that $\mathbf{n} = [\boldsymbol{\xi}]$ and $\mathbf{k} = \{\boldsymbol{\xi}\}$ (integer and fractional parts, respectively). Then we can reformulate formula (1.6) for the unperturbed eigenvalues as follows: there is a mapping $f : \mathbb{R}^d \to \mathbb{R}$, given by the formula $f(\boldsymbol{\xi}) = |\mathbf{F}(\mathbf{n}+\mathbf{k})|^2$ such that for each $\mathbf{k}$ the restriction of f to $\{\boldsymbol{\xi} \in \mathbb{R}^d : \{\boldsymbol{\xi}\} = \mathbf{k}\}$ is a bijection onto the set of all eigenvalues of $H(\mathbf{k})$ (counting multiplicities). We want to give an analogue of this (trivial) statement in the general case. Let us define the spherical layer

$$\mathcal{A} := \{\boldsymbol{\xi} \in \mathbb{R}^d, \, \big| \, |\mathbf{F}\boldsymbol{\xi}|^2 - \lambda \, \big| \leq 40v\}$$

(v is the L_∞-norm of V). Let $N \in \mathbb{N}$ be a fixed number. We will construct two mappings $f, g : \mathcal{A} \to \mathbb{R}$ which satisfy the following properties:

(I) for each $\mathbf{k}$ the first mapping f is an injection from the set of all $\boldsymbol{\xi}$ with fractional part equal to $\mathbf{k}$ into the spectrum of $H(\mathbf{k})$ (counting multiplicities). Each eigenvalue of $H(\mathbf{k})$ inside $J := [\lambda - 20v, \lambda + 20v]$ has a pre-image $\boldsymbol{\xi} \in \mathcal{A}$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$. The perturbation inequality $|f(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \leq 2v$ holds for all $\boldsymbol{\xi} \in \mathcal{A}$.

(II) for $\boldsymbol{\xi} \in \mathcal{A}$ satisfying $|\mathbf{F}\boldsymbol{\xi}|^2 \in J$ we have: $|f(\boldsymbol{\xi}) - g(\boldsymbol{\xi})| < \rho^{-N}$;

(III) one can divide the domain of g in two parts: $\mathcal{A} = \mathcal{B} \cup \mathcal{D}$ (both $\mathcal{B}$ and $\mathcal{D}$ are intersections of $\mathcal{A}$ with some cones centered at the origin – at least modulo very small sets) such that $g(\boldsymbol{\xi})$ is given by an explicit formula when $\boldsymbol{\xi} \in \mathcal{B}$, we have some control over $g(\boldsymbol{\xi})$ when $\boldsymbol{\xi} \in \mathcal{D}$, and the ratio of volumes of $\mathcal{B}$ and $\mathcal{D}$ goes to infinity when $\rho \to \infty$.

The set $\mathcal{B}$, called the non-resonance set, contains, among others, all points $\boldsymbol{\xi} \in \mathcal{A}$ which satisfy the inequality

$$|\langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\theta}\rangle| \geq \rho^{1/3}|\mathbf{F}\boldsymbol{\theta}| \tag{1.7}$$

for all non-zero integer vectors $\boldsymbol{\theta}$ with $|\boldsymbol{\theta}| \ll RN$. The precise formula for g will imply, in particular, that when $\boldsymbol{\xi} \in \mathcal{B}$ we have $g(\boldsymbol{\xi}) = |\mathbf{F}\boldsymbol{\xi}|^2 + G(\boldsymbol{\xi})$ with all partial derivatives of G being $O(\rho^{-\epsilon})$ for some $\epsilon > 0$. When $\boldsymbol{\xi}$ belongs to the resonance set $\mathcal{D}$, we can give good estimates

only of the partial derivative of g along one direction; this direction has a small angle with the direction of $\boldsymbol{\xi}$. The behaviour of g along all other directions is much worse. Indeed, by considering potentials V which allow to perform the separation of variables, one can see that the function g can not, in general, be made even continuous in the resonance set. However, we still have some (although rather weak) control over the behaviour of g along all directions inside the resonance set; see lemma 7.11 for the precise formulation of these properties.

One should mention that asymptotic formulae of non-resonance eigenvalues (i.e. the function $g(\boldsymbol{\xi})$ for $\boldsymbol{\xi} \in \mathcal{B}$ in our notation) and some resonance eigenvalues were obtained before in certain cases, using completely different methods, by O. A. Veliev, [14] and [15] and Y. E. Karpeshina (see [4], [5] and references therein). However, as has been already mentioned, there are certain distinctions between the settings of [15] and [5] and the settings of our paper. Because of this, and in order to make our paper self-contained, it seems sensible to include an independent proof of the asymptotic formula for eigenvalues.

Before describing how to construct these mappings, we explain first how to prove the Bethe-Sommerfeld conjecture using them. Put $\delta = \rho^{-N}$. For each $\boldsymbol{\eta} \in \mathbb{R}^d$ of unit length we denote $I_{\boldsymbol{\eta}}$ the interval consisting of points $\boldsymbol{\xi} = t\boldsymbol{\eta}$, $t > 0$ satisfying $g(\boldsymbol{\xi}) \in [\lambda - \delta, \lambda + \delta]$; we will consider only vectors $\boldsymbol{\eta}$ for which $I_{\boldsymbol{\eta}} \in \mathcal{B}$. Suppose we have found an interval $I_{\boldsymbol{\eta}}$ on which the mapping f is continuous. Then property (II) above together with the intermediate value theorem would imply that there is a point $\boldsymbol{\xi}(\lambda) \in I_{\boldsymbol{\eta}}$ satisfying $f(\boldsymbol{\xi}(\lambda)) = \lambda$, which would mean that λ is in the spectrum of H. Thus, if we assume that λ belongs to the spectral gap, this would imply that the mapping f is not continuous on each of the intervals $I_{\boldsymbol{\eta}}$. A simple argument shows that in this case for each point $\boldsymbol{\xi} \in \mathcal{B}$ with $|g(\boldsymbol{\xi}) - \lambda| \le \delta$ there exists another point $\boldsymbol{\xi}_1 \ne \boldsymbol{\xi}$ with $\boldsymbol{\xi} - \boldsymbol{\xi}_1 \in \mathbb{Z}^d$ and $|g(\boldsymbol{\xi}_1) - \lambda| \ll \delta$. The existence of such a point $\boldsymbol{\xi}_1$ (which we call a conjugate point to $\boldsymbol{\xi}$) is a crucial part of the proof; it seems that similar arguments based on the existence of conjugate points could be helpful in analogous problems. Afterwards, a geometrical combinatorics argument shows that for sufficiently small δ ($\delta \asymp \rho^{1-d}$ would do) some (moreover, most) of the points $\boldsymbol{\xi} \in \mathcal{B} \cap g^{-1}([\lambda - \delta, \lambda + \delta])$ have no conjugate points; the important part in the proof is played by the fact that the surface $g^{-1}(\lambda) \cap \mathcal{B}$ has positive curvature in each direction.

Now let us discuss how to construct mappings f and g with properties described above. This is done in several steps. First, we prove lemma 3.2 which states that under certain conditions it is possible instead of

studying eigenvalues of the operator $H = H_0 + V$, to study eigenvalues of the operator

$$\sum_j P^j H P^j, \tag{1.8}$$

where P^j are spectral projections of H_0; the error of this approximation is small. This result can be applied to the operators $H(\mathbf{k})$ from the direct integral (1.5). We want therefore to study the spectrum of the (direct) sum (1.8) where P^j are projections 'localized' in some domains of the $\boldsymbol{\xi}$-space. The geometrical structure of these projections will depend on whether the localization happens inside or outside the resonance regions. The case of a projection P^j 'localized' around a point $\boldsymbol{\xi} \in \mathcal{B}$ is relatively simple: the rank of such projections does not depend on ρ or the 'localization point' $\boldsymbol{\xi}$. Thus, in this case we will need to compute the eigenvalue of the finite matrix $P^j H(\mathbf{k}) P^j$. This can be done by computing the characteristic polynomial of this matrix and then using the iterative process based on the Banach fixed point theorem to find the root of this characteristic polynomial. It is much more difficult to construct projections P^j corresponding to the points $\boldsymbol{\xi}$ located inside the resonance set $\mathcal{D}$. The form of projections will depend on, loosely speaking, how many linearly independent integer vectors $\boldsymbol{\theta}$ for which (1.7) is not satisfied are there. The construction of such projections is the most technically difficult part of the paper. Once these projections are constructed, it turns out that the eigenvalues of $P^j H P^j$ with large ρ can be easily expressed in terms of the eigenvalues of the operator pencil $rA + B$ where A and B are fixed and $r \sim \rho$ is a large parameter. The rest is a relatively simple perturbation theory.

The approach used in this paper can be applied to various related problems. For example, it seems possible to obtain several new terms of the asymptotics of the integrated density of states using these methods. It might even be possible to obtain the complete asymptotic formula; however, this would require much more careful analysis of the mapping g in the resonance set. As an immediate 'free' corollary of our results we obtain the theorem 7.17 which seems too be new. Loosely speaking, it states that there are no 'simultaneous clusters' of eigenvalues of all operators $H(\mathbf{k})$.

The approach of this paper works almost without changes for the polyharmonic operators $(-\Delta)^l + V$ with a smooth periodic potential V. Another possible field of applications of the results of this paper is studying the structure of the (complex) Bloch and Fermi varieties.

The rest of the paper is constructed as follows: section 2 has all necessary preliminaries; also, in this section for the convenience of the

reader we, taking into account the size of the paper, give references to the definitions of all major objects in the paper. Section 3 proves the abstract result allowing to reduce computation of the spectrum of $H = H_0 + V$ to the computation of the spectrum of $\sum_j P^j H P^j$, P^j being the spectral projections of H_0. Section 4 proves various estimates of angles between lattice points which are needed to keep track on the dependence of all results on R – the size of the support of the potential. In section 5 we apply the abstract lemma from section 3 to our case and perform the reduction of H to the sum of simpler operators. In section 6 we compute the eigenvalues of these simpler operators corresponding to the non-resonance set; we also give the formula for $g(\boldsymbol{\xi})$ when $\boldsymbol{\xi} \in \mathcal{B}$. Section 7 is devoted to the study of the properties of these simpler operators and the mapping g restricted to the resonance set $\mathcal{D}$. Finally, in section 8 we prove the Bethe-Sommerfeld conjecture.

When this manuscript was ready, I have learned that another article of Veliev [16] was published recently.

2. Preliminaries

We study the Schrödinger operator

$$H = H_0 + V(\mathbf{x}), \qquad H_0 = \mathbf{DGD}, \tag{2.1}$$

with the potential V being infinitely smooth and periodic with the lattice of periods equal $(2\pi\mathbb{Z})^d$. Here, $\mathbf{D} = i\nabla$, and $\mathbf{G} = \mathbf{F}^2$ is $d \times d$ positive matrix; $\mathbf{F}$ is also taken to be positive.

Throughout the paper we use the following notation. If A is a bounded below self-adjoint operator with compact resolvent, then we denote by $\{\mu_j(A)\}$ $(j = 1, 2, \dots)$ the set of eigenvalues of A written in non-decreasing order, counting multiplicities.

As we have already mentioned, the spectrum of H is the union over $\mathbf{k} \in \mathbb{R}^d/\mathbb{Z}^d$ of the spectra of the operators $H(\mathbf{k})$, the domain of each

$H(\mathbf{k})$ is $\mathfrak{D}(\mathbf{k})$ and $H(\mathbf{k}) := \mathbf{DGD} + V(\mathbf{x})$. By $\mathfrak{H} := L^2(T^d)$ we denote the Hilbert space in which all the operators $H(\mathbf{k})$ act. We also denote by $H_0(\mathbf{k})$ the operator $\mathbf{DGD}$ with the domain $\mathfrak{D}(\mathbf{k})$. Let $\lambda_j(\mathbf{k}) = \mu_j(H(\mathbf{k}))$ be the jth eigenvalue of $H(\mathbf{k})$. Then it is well-known (see, for example, [10]) that each function $\lambda_j(\cdot)$ is continuous and piecewise smooth. Denote by ℓ_j the image of $\lambda_j(\cdot)$. Then ℓ_j is called the jth *spectral band*. We also define, for each $\lambda \in \mathbb{R}$, the following functions:

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

is the *multiplicity of overlapping* ($\#$ denotes the number of elements in a set). The *overlapping function* $\zeta(\lambda)$ is defined as the maximal number t such that the symmetric interval $[\lambda - t, \lambda + t]$ is entirely contained in one of the bands ℓ_j:

$$\zeta(\lambda) = \max_j \max\{t \in \mathbb{R} : [\lambda - t, \lambda + t] \subset \ell_j\}.$$

Finally,

$$N(\lambda) = \int_{[0,1]^d} \#\{j : \lambda_j(\mathbf{k}) < \lambda\} d\mathbf{k} \qquad (2.2)$$

is the *integrated density of states* of the operator (1.4). For technical reasons sometimes it will be convenient to assume that the dimension d is at least 3 (in the 2-dimensional case the conjecture has been proved already, so this assumption does not restrict generality). The main result of the paper is the following:

Theorem 2.1. *Let $d \geq 3$. Then all sufficiently large points $\lambda = \rho^2$ are inside the spectrum of H. Moreover, there exists a positive constant Z such that for large enough ρ the whole interval $[\rho^2 - Z\rho^{1-d}, \rho^2 + Z\rho^{1-d}]$ lies inside some spectral band.*

Without loss of generality we always assume that $\int_{[0,2\pi]^d} V(\mathbf{x}) d\mathbf{x} = 0$. Abusing the notation slightly, we will denote by V both the potential itself and the operator of multiplication by V.

By $B(R)$ we denote a ball of radius R centered at the origin. By C or c we denote positive constants, depending only on d, $\mathbf{G}$, and norms of the potential in various Sobolev spaces H^s. In section 5 we will introduce parameters p, q_j and M; constants are allowed to depend on the values of these parameters as well. The exact value of constants can be different each time they occur in the text, possibly even each time they occur in the same formula. On the other hand, the constants which are labelled (like C_1, c_3, etc) have their values being fixed throughout the text. Whenever we use O, o, $\gg$, $\ll$, or $\asymp$ notation, the constants involved will also depend on d, $\mathbf{G}$, M, and norms of the potential; the same is also the case when we use the expression 'sufficiently large'.

Given two positive functions f and g, we say that $f \gg g$, or $g \ll f$, or $g = O(f)$ if the ratio $\frac{g}{f}$ is bounded. We say $f \asymp g$ if $f \gg g$ and $f \ll g$. By $\lambda = \rho^2$ we denote a point on the spectral axis. We will always assume that λ is sufficiently large. We also denote by v the L_∞-norm of the potential V, and $J := [\lambda - 20v, \lambda + 20v]$. Finally,

$$\mathcal{A} := \{\boldsymbol{\xi} \in \mathbb{R}^d, \ \big| \ |\mathbf{F}\boldsymbol{\xi}|^2 - \lambda \ \big| \le 40v\} \tag{2.3}$$

and

$$\mathcal{A}_1 := \{\boldsymbol{\xi} \in \mathbb{R}^d, \ \big| \ |\mathbf{F}\boldsymbol{\xi}|^2 - \lambda \ \big| \le 20v\}. \tag{2.4}$$

Notice that the definition of $\mathcal{A}$ obviously implies that if $\boldsymbol{\xi} \in \mathcal{A}$, then $\big| \ |\mathbf{F}\boldsymbol{\xi}| - \rho \ \big| \ll \rho^{-1}$.

Given several vectors $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n \in \mathbb{R}^d$, we denote by $R(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ the linear subspace spanned by these vectors, and by $Z(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ the lattice generated by them (i.e. the set of all linear combinations of $\boldsymbol{\eta}_j$ with integer coefficients; we will use this notation only when these vectors are linearly independent). We denote by $\mathbf{M}(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ the $d \times n$ matrix whose jth column equals $\boldsymbol{\eta}_j$. Given any lattice Γ, we denote by $|\Gamma|$ the volume of the cell of Γ, so that if $\Gamma = Z(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_d)$, then $|\Gamma|$ is the absolute value of the determinant of $\mathbf{M}(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_d)$. We also denote, for any linear space $\mathfrak{V} \subset \mathbb{R}^d$, $B(\mathfrak{V}; R) := \mathfrak{V} \cap B(R)$. For any non-zero vector $\boldsymbol{\xi} \in \mathbb{R}^d$ we denote $n(\boldsymbol{\xi}) := \frac{\boldsymbol{\xi}}{|\mathbf{F}\boldsymbol{\xi}|}$. Any vector $\boldsymbol{\xi} \in \mathbb{R}^d$ can be uniquely decomposed as $\boldsymbol{\xi} = \mathbf{n} + \mathbf{k}$ with $\mathbf{n} \in \mathbb{Z}^d$ and $\mathbf{k} \in [0,1)^d$. We call $\mathbf{n} = [\boldsymbol{\xi}]$ the integer part of $\boldsymbol{\xi}$ and $\mathbf{k} = \{\boldsymbol{\xi}\}$ the fractional part of $\boldsymbol{\xi}$.

Whenever P is a projection and A is an arbitrary operator acting in a Hilbert space $\mathcal{H}$, the expression PAP means, slightly abusing the notation, the operator $PAP : P\mathcal{H} \to P\mathcal{H}$.

Throughout the paper we use the following convention: vectors are denoted by bold lowercase letters; matrices by bold uppercase letters; sets (subsets of $\mathbb{R}^d$) by calligraphic uppercase letters; linear subspaces by gothic uppercase letters. By $\text{vol}(\mathcal{C})$ we denote the Lebesgue measure of the set $\mathcal{C}$. If $\mathcal{C}_j \subset \mathbb{R}^d$, $j = 1, 2$ are two subsets of $\mathbb{R}^d$, their sum is defined in the usual way:

$$\mathcal{C}_1 + \mathcal{C}_2 = \{\boldsymbol{\xi} \in \mathbb{R}^d : \boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2, \ \boldsymbol{\xi}_j \in \mathcal{C}_j\}.$$

Finally, for the benefit of the reader we will list here either the definitions of the major objects introduced later in the paper or references to the formulas in which they are defined.

$f, g : \mathcal{A} \to \mathbb{R}$ are mappings satisfying properties listed in theorem 7.13 (if the Fourier transform of V has compact support) and in corollary 7.15 for general potentials. The sets Θ_j and Θ'_j are defined in (5.5).

The projections $\mathcal{P}^{(\mathbf{k})}(\mathcal{C})$ are defined immediately before Lemma 5.14. $\mathcal{V}(n)$, $\boldsymbol{\xi}_{\mathfrak{V}}$, $\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}$, and $\Theta(\mathfrak{V})$ are defined at the beginning of subsection 7.1. The sets $\Xi(\mathfrak{V})$ and $\Xi_j(\mathfrak{V})$ ($j = 0, \dots, 3$) are defined in formulas (5.7)-(5.11); the sets $\boldsymbol{\Upsilon}_j(\boldsymbol{\xi})$, $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$, $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$, and $\boldsymbol{\Upsilon}(\boldsymbol{\xi}; U)$ are defined by formulas (7.1), (7.2), (7.11), and (7.15) correspondingly. The numbers p and q_n are defined in (5.15), $K = \rho^p$ and $L_n = \rho^{q_n}$. The projection $P(\boldsymbol{\xi})$ and the operator $H'(\boldsymbol{\xi})$ are defined in (7.3) and (7.4) correspondingly. The sets $\mathcal{B}$ and $\mathcal{D}$ are defined in (5.13) and (5.12). $r(\boldsymbol{\xi})$ and $\boldsymbol{\xi}_{\mathfrak{V}}'$ are defined by formula (7.8). Operators A and B are defined by (7.30) and (7.31). Finally, the sets $\mathcal{A}(\delta)$, $\mathcal{B}(\delta)$, and $\mathcal{D}(\delta)$ are defined before Lemma 8.1.

3. Reduction to invariant subspaces: general result

The key tool in finding a good approximation of the eigenvalues of $H(\mathbf{k})$ will be the following two lemmas.

Lemma 3.1. *Let H_0, V and A be self-adjoint operators such that H_0 is bounded below and has compact resolvent, and V and A are bounded. Put $H = H_0 + V$ and $\hat{H} = H_0 + V + A$ and denote by $\mu_l = \mu_l(H)$ and $\hat{\mu}_l = \mu_l(\hat{H})$ the sets of eigenvalues of these operators. Let $\{P_j\}$ ($j = 0, \dots, n$) be a collection of orthogonal projections commuting with H_0 such that $\sum P_j = I$, $P_j V P_k = 0$ for $|j - k| > 1$, and $A = P_n A$. Let l be a fixed number. Denote by a_j the distance from μ_l to the spectrum of $P_j H_0 P_j$ Assume that for $j \geq 1$ we have $a_j > 4a$, where $a := ||V|| + ||A||$. Then $|\hat{\mu}_l - \mu_l| \leq 2^{2n} a^{2n+1} \prod_{j=1}^n (a_j - 2a)^{-2}$.*

Proof. Let $H_t = H + tA$, $0 \leq t \leq 1$ and let $\mu(t) = \mu_l(H(t))$ be the corresponding family of eigenvalues. We also choose the family of corresponding normalized eigenfunctions $\phi(t) = \phi_l(t)$. We will skip writing the index l in the rest of the proof. Elementary perturbation theory (see, e.g., [6]) implies that $\mu(t)$ is piecewise differentiable and

$$\frac{d\mu(t)}{dt} = (A\phi(t), \phi(t)). \tag{3.1}$$

Let $\Phi_j = \Phi_j(t) := P_j \phi(t)$, and let $V_{kj} := P_k V P_j$ (so that $V_{jk} = 0$ if $|j - k| > 1$). Then the eigenvalue equation for $\phi(t)$ can be written in the following way:

$$H_0 \Phi_0 + V_{00} \Phi_0 + V_{01} \Phi_1 = \mu(t) \Phi_0$$
$$H_0 \Phi_j + V_{j\,j-1} \Phi_{j-1} + V_{j\,j} \Phi_j + V_{j\,j+1} \Phi_{j+1} = \mu(t) \Phi_j, \ 1 \leq j < n \quad (3.2)$$
$$H_0 \Phi_n + V_{n\,n-1} \Phi_{n-1} + V_{n\,n} \Phi_n + tA \Phi_n = \mu(t) \Phi_n.$$

Indeed, let us apply P_k to both sides of equation $(H_0 + V + tA)\phi = \mu\phi$. We will obtain

$$H_0\Phi_k + P_kV\phi + tP_kA\phi = \mu\Phi_k. \tag{3.3}$$

Now we use the following identities:

$$P_kV\phi = P_kV\sum_j P_j\phi = \sum_{j,|j-k|\leq 1} P_kVP_j\phi = \sum_{j,|j-k|\leq 1} V_{kj}\Phi_j \tag{3.4}$$

and $A = P_nA = (P_nA)^* = AP_n = P_nAP_n$, so

$$P_kA\phi = \delta_{k,n}A\Phi_n. \tag{3.5}$$

Identities (3.3)–(3.5) imply (3.2).

Let us now prove, using the backwards induction, that for all k, $1 \leq k \leq n$ and all $t \in [0,1]$ we have

$$||\Phi_k(t)|| \leq \frac{2a}{a_k - 2a}||\Phi_{k-1}(t)||. \tag{3.6}$$

Indeed, from the last equation in (3.2) we see that

$$\Phi_n(t) = -[P_n(H_0 + V_{nn} + tA - \mu(t))P_n]^{-1}V_{nn-1}\Phi_{n-1}(t). \tag{3.7}$$

Since $|\mu - \mu(t)| \leq a$, the distance from $\mu(t)$ to the spectrum of $P_nH_0P_n$ is at least $a_n - a$. Since $||V_{nn} + tA|| \leq a$, this implies

$$||[P_n(H_0 + V_{nn} + tA - \mu(t))P_n]^{-1}|| \leq \frac{1}{a_n - 2a}.$$

Thus, (3.7) implies $||\Phi_n(t)|| \leq \frac{a}{a_n - 2a}||\Phi_{n-1}(t)||$, and (3.6) holds for $k = n$. Assume now that we have proved (3.6) for all k between $j+1$ and n, $1 \leq j < n$. Then, analogously to (3.7), we have:

$$\Phi_j(t) = -[P_j(H_0 + V_{jj} - \mu(t))P_j]^{-1}[V_{jj-1}\Phi_{j-1}(t) + V_{jj+1}\Phi_{j+1}(t)], \tag{3.8}$$

so

$$\begin{aligned}||\Phi_j(t)|| &\leq \frac{a}{a_j - 2a}(||\Phi_{j-1}(t)|| + ||\Phi_{j+1}(t)||) \\ &\leq \frac{a}{a_j - 2a}||\Phi_{j-1}(t)|| + \frac{2a^2}{(a_j - 2a)(a_{j+1} - 2a)}||\Phi_j(t)||,\end{aligned} \tag{3.9}$$

where we have used the validity of (3.6) for $k = j+1$. This shows that (3.6) holds for $k = j$, since

$$\frac{2a^2}{(a_j - 2a)(a_{j+1} - 2a)} < \frac{1}{2}.$$

Using (3.6) and the fact that $||\Phi_1|| \leq 1$, we see that

$$||\Phi_n|| \leq \frac{2^n a^n}{\prod_{j=1}^n(a_j - 2a)}.$$

Since the RHS of (3.1) equals

$$(P_n A P_n \phi(t), \phi(t)) = (A\Phi_n(t), \Phi_n(t)),$$

this finishes the proof $\qquad\square$

Now we formulate the immediate corollary of lemma 3.1 which we will be using throughout.

Lemma 3.2. *Let H_0 and V be self-adjoint operators such that H_0 is bounded below and has compact resolvent and V is bounded. Let $\{P^m\}$ ($m = 0,\ldots,T$) be a collection of orthogonal projections commuting with H_0 such that if $m \neq n$ then $P^m P^n = P^m V P^n = 0$. Denote $Q := I - \sum P^m$. Suppose that each P^m is a further sum of orthogonal projections commuting with H_0: $P^m = \sum_{j=0}^{j_m} P_j^m$ such that $P_j^m V P_l^m = 0$ for $|j - l| > 1$ and $P_j^m V Q = 0$ if $j < j_m$. Let $v := ||V||$ and let us fix an interval $J = [\lambda_1, \lambda_2]$ on the spectral axis which satisfies the following properties: spectra of the operators $Q H_0 Q$ and $P_j^k H_0 P_j^k$, $j \geq 1$ lie outside J; moreover, the distance from the spectrum of $Q H_0 Q$ to J is greater than $6v$ and the distance from the spectrum of $P_j^k H_0 P_j^k$ ($j \geq 1$) to J, which we denote by a_j^k, is greater than $16v$. Denote by $\mu_p \leq \cdots \leq \mu_q$ all eigenvalues of $H = H_0 + V$ which are inside J. Then the corresponding eigenvalues $\tilde{\mu}_p, \ldots, \tilde{\mu}_q$ of the operator*

$$\tilde{H} := \sum_m P^m H P^m + Q H_0 Q$$

are eigenvalues of $\sum_m P^m H P^m$, and they satisfy

$$|\tilde{\mu}_r - \mu_r| \leq \max_m \left[(6v)^{2j_m+1} \prod_{j=1}^{j_m} (a_j^m - 6v)^{-2} \right];$$

all other eigenvalues of $\tilde{H}$ are outside the interval $[\lambda_1 + 2v, \lambda_2 - 2v]$.

Proof. Assumptions of the lemma imply that

$$H = \tilde{H} + (Q + \sum_m P_{j_m}^m) V (Q + \sum_m P_{j_m}^m) - (\sum_m P_{j_m}^m) V (\sum_m P_{j_m}^m).$$

Therefore, $\tilde{H} - 2v(Q + \sum_m P_{j_m}^m) \leq H \leq \tilde{H} + 2v(Q + \sum_m P_{j_m}^m)$, and the elementary perturbation theory implies that for all l

$$\mu_l(\tilde{H} - 2v(Q + \sum_m P_{j_m}^m)) \leq \mu_l(H) \leq \mu_l(\tilde{H} + 2v(Q + \sum_m P_{j_m}^m)). \quad (3.10)$$

The operators $\tilde{H} \pm 2v(Q + \sum_m P_{j_m}^m)$ split into the sum of invariant operators $Q H_0 Q \pm 2vQ$ and $P^m H P^m \pm 2v P_{j_m}^m$ ($m = 0, \ldots, n$). The spectrum of operators $Q H_0 Q \pm 2vQ$ is outside $[\lambda_1 - 4v, \lambda_2 + 4v]$ due to the

assumptions of the lemma. Therefore, since the shift of an eigenvalue is at most the norm of the perturbation, for $p \le l \le q$, $\mu_l(\tilde{H} \pm 2v(Q + \sum_m P_{j_m}^m))$ is an eigenvalue of one of the operators $P^m H P^m \pm 2v P_{j_m}^m$. If we now apply lemma 3.1 to each of the operators $P^m H P^m \pm 2v P_{j_m}^m$ with $A := \pm 2v P_{j_m}^m$ and $a = 3v$, we will obtain

$$|\mu_k(P^m H P^m \pm 2v P_{j_m}^m) - \mu_k(P^m H P^m)|$$

$$\le 6^{2j_m+1} v^{2j_m+1} \max_m \prod_{j=1}^{j_m} (a_j^m - 6v)^{-2}$$

$$\le \max_m \left[(6v)^{2j_m+1} \prod_{j=1}^{j_m} (a_j^m - 6v)^{-2} \right] =: \tau,$$

$$(3.11)$$

provided $\mu_k(P^m H P^m) \in [\lambda_1 - 4v, \lambda_2 + 4v]$. Let us now define the bijection F mapping the set of all eigenvalues of $\tilde{H}$ to the set of all eigenvalues of $\{\mu_l(\tilde{H} + 2v(Q + \sum_m P_{j_m}^m))\}$ (counting multiplicities) in the following way. Suppose, μ is an eigenvalue of $\tilde{H}$. Then either $\mu = \mu_k(Q H_0 Q)$, or $\mu = \mu_k(P^m H P^m)$ for some k, m. We define $F(\mu) := \mu_k(Q H_0 Q + 2vQ)$ in the former case, and $F(\mu) := \mu_k(P^m H P^m + 2v P_{j_m}^m)$ in the latter case. Then the mapping F satisfies the following properties:

$$|F(\mu) - \mu| \le 2v; \qquad (3.12)$$

moreover, if $\mu \in [\lambda_1 - 4v, \lambda_2 + 4v]$, then

$$|F(\mu) - \mu| \le \tau \qquad (3.13)$$

(this follows from (3.11)). A little thought shows that this implies

$$|\mu_l(\tilde{H} + 2v(Q + \sum_m P_{j_m}^m)) - \mu_l(\tilde{H})| \le \tau \qquad (3.14)$$

for $p \le l \le q$. Indeed, suppose that (3.14) is not satisfied for some l, say

$$\mu_l(\tilde{H} + 2v(Q + \sum_m P_{j_m}^m)) - \mu_l(\tilde{H}) > \tau; \qquad (3.15)$$

in particular, this implies $\mu_l(\tilde{H} + 2v(Q + \sum_m P_{j_m}^m)) > \lambda_1 - 2v$. Then the pigeonhole principle shows that F maps at least one of the eigenvalues $\mu_k(\tilde{H})$, $k \le l$ to $\mu_t(\tilde{H} + 2v(Q + \sum_m P_{j_m}^m))$ with $t \ge l$. If $\mu_k(\tilde{H}) < \lambda_1 - 4v$, this contradicts (3.12), and if $\mu_k(\tilde{H}) \ge \lambda_1 - 4v$, this contradicts (3.13). These contradictions prove (3.14). Similarly, we prove that

$$|\mu_l(\tilde{H} - 2v(Q + \sum_m P_{j_m}^m)) - \mu_l(\tilde{H})| \le \tau. \qquad (3.16)$$

Estimates (3.14) and (3.16) together with (3.10) prove the lemma. □

Corollary 3.3. *If all conditions of lemma 3.2 are satisfied, there exists an injection G defined on a set of eigenvalues of the operator $\sum_m P^m H P^m$ (all eigenvalues are counted according to their multiplicities) and mapping them to a subset of the set of eigenvalues of H (again considered counting multiplicities) such that:*
 (i) all eigenvalues of H inside J have a pre-image,
 (ii) If $\mu_j \in [\lambda_1 + 2v, \lambda_2 - 2v]$ is an eigenvalue of $\sum_m P^m H P^m$, then

$$|G(\mu_j) - \mu_j| \leq \max_m \left[(6v)^{2j_m+1} \prod_{j=1}^{j_m} (a_j^m - 6v)^{-2} \right],$$

 and
 (iii) $G(\mu_j(\sum_m P^m H P^m)) = \mu_{j+l}(H)$, where l is the number of eigenvalues of $Q H_0 Q$ which are smaller than λ_1.

Proof. Statements (i) and (ii) follow immediately from lemma 3.2, and to prove (iii) we just notice that if $\mu_j(\sum_m P^m H P^m) \in J$, then

$$\mu_j(\sum_m P^m H P^m) = \mu_{j+l}(Q H_0 Q + \sum_m P^m H P^m).$$

□

4. Lattice points

In this section, we prove various auxiliary estimates of angles between integer vectors.

Lemma 4.1. *Let $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n \in \mathbb{Z}^d$ be linearly independent. Let $\Gamma = Z(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ and suppose that $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_{n-1} \in \Gamma \cap B(R)$. Then there exists a vector $\boldsymbol{\theta} \in \Gamma$, $\boldsymbol{\theta} \neq 0$ orthogonal to all $\boldsymbol{\nu}_j$'s, such that*

$$|\boldsymbol{\theta}| \leq 2^n |\Gamma| \prod_{j=1}^{n-1} |\boldsymbol{\nu}_j| \tag{4.1}$$

and, therefore, $|\boldsymbol{\theta}| \leq 2^n |\Gamma| R^{n-1}$.

Proof. For $r > 1$ let $\mathcal{A}_r \subset R(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ be the set

$\mathcal{A}_r = \{\boldsymbol{\xi} \in R(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n) : |\langle \boldsymbol{\xi}, \boldsymbol{\nu}_j \rangle| < 1, j = 1, 2, \ldots, n-1, \ \& \ |\boldsymbol{\xi}| < r\}.$

This set is obviously convex and symmetric about the origin. Moreover,

$$\mathrm{vol}(\mathcal{A}_r) > r \prod_{j=1}^{n-1} |\boldsymbol{\nu}_j|^{-1}$$

By Minkowski's convex body theorem (see, e.g., [1], §III.2.2 Theorem II), under the condition $\mathrm{vol}(\mathcal{A}_r) > |\Gamma|2^n$ the set $\mathcal{A}_r$ contains at least two non-zero points $\pm\boldsymbol{\theta} \in \Gamma$. The above condition is satisfied if $r \prod_{j=1}^{n-1} |\boldsymbol{\nu}_j|^{-1} \geq 2^d|\Gamma|$, that is if $r \geq 2^d|\Gamma| \prod_{j=1}^{n-1} |\boldsymbol{\nu}_j|$. Since $\boldsymbol{\nu}_j$'s and $\boldsymbol{\theta}$ are integer vectors, the condition $|\langle\boldsymbol{\theta},\boldsymbol{\nu}_j\rangle| < 1$ is equivalent to $\langle\boldsymbol{\theta},\boldsymbol{\nu}_j\rangle = 0$. This implies the required result. $\qquad\square$

Lemma 4.2. *Let $\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n,\boldsymbol{\mu} \in \mathbb{Z}^d \cap B(R)$ be linearly independent. Then the angle between $\boldsymbol{\mu}$ and $R(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)$ is $\gg R^{-n-1}$.*

Proof. Suppose this angle is smaller than R^{-n-1}. Then the lattice $\Gamma = Z(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n,\boldsymbol{\mu})$ has $|\Gamma| \leq 1$. Lemma 4.1 then implies that there exists a vector $\boldsymbol{\theta} \in \Gamma$, $\boldsymbol{\theta} \perp \boldsymbol{\theta}_j$, $|\boldsymbol{\theta}| \ll R^n$. Then, since $\boldsymbol{\theta}$ and $\boldsymbol{\mu}$ are non-orthogonal integer vectors, we have: $|\langle\boldsymbol{\mu},\boldsymbol{\theta}\rangle| \geq 1$, and sin of the angle between $\boldsymbol{\mu}$ and $R(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)$, which equals cos of the angle between $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$, is bounded below by $|\boldsymbol{\theta}|^{-1}|\boldsymbol{\mu}|^{-1} \gg R^{-n-1}$. $\qquad\square$

Corollary 4.3. *Let $\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n,\boldsymbol{\mu} \in \mathbb{Z}^d \cap B(R)$ be linearly independent. Then the angle between $\mathbf{F}\boldsymbol{\mu}$ and $R(\mathbf{F}\boldsymbol{\theta}_1,\ldots,\mathbf{F}\boldsymbol{\theta}_n)$ is $\gg R^{-n-1}$.*

Proof. This is equivalent to saying that for each $\boldsymbol{\xi} \in R(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)$ the distance between $\mathbf{F}(n(\boldsymbol{\mu}))$ and $\mathbf{F}\boldsymbol{\xi}$ is larger than cR^{-n-1}. But the distance between $\mathbf{F}(n(\boldsymbol{\mu}))$ and $\mathbf{F}\boldsymbol{\xi}$ is not greater than the largest eigenvalue of $\mathbf{F}$ times the distance between $n(\boldsymbol{\mu})$ and $\boldsymbol{\xi}$. Now the statement follows from lemma 4.2. $\qquad\square$

It is possible to generalize lemma 4.2 a bit: if we talk about distance from a vector to a linear sub-space instead of the angle between a vector and a subspace, we can drop the assumption that $|\mu| \leq R$:

Lemma 4.4. *Let $\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n \in \mathbb{Z}^d \cap B(R)$ and $\boldsymbol{\mu} \in \mathbb{Z}^d$ be linearly independent. Then the distance between $\boldsymbol{\mu}$ and $R(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)$ is $\gg R^{-n}$.*

Proof. The distance between $\boldsymbol{\mu}$ and $R(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)$ equals $\frac{|Z(\mu,\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)|}{|Z(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n)|}$. The square of the denominator of this fraction is the determinant of the $n \times n$ matrix A with $A_{jk} := \langle\boldsymbol{\theta}_j,\boldsymbol{\theta}_k\rangle = O(R^2)$, so the denominator is $O(R^n)$. Similarly, the square of the numerator is the determinant of $(n+1) \times (n+1)$ non-singular matrix with integer entries. Therefore, the absolute value of the numerator is at least 1. This proves our statement. $\qquad\square$

The following result is a generalization of lemma 4.1 and the proof is similar:

Lemma 4.5. *Let Γ be as above and let $\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_m \in \Gamma \cap B(R)$ ($m < n$) Then there exist linearly independent vectors $\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_{n-m} \in \Gamma$ such*

that each $\boldsymbol{\theta}_l$ is orthogonal to each $\boldsymbol{\nu}_j$ and

$$\prod_{l=1}^{n-m} |\boldsymbol{\theta}_l| \ll |\Gamma| \prod_{j=1}^{m} |\boldsymbol{\nu}_j| \leq |\Gamma| R^m \qquad (4.2)$$

Proof. Applying lemma 4.1 $n-m$ times, we see that the set of vectors from Γ which are orthogonal to $\boldsymbol{\nu}_j$ form a lattice Γ_{n-m} of dimension $n-m$. Let $\boldsymbol{\theta}_j$ $(j=1,\ldots,n-m)$ be successive minimal vectors of Γ_{n-m}. That means that $\boldsymbol{\theta}_1$ is the smallest nonzero vector in Γ_{n-m}; $\boldsymbol{\theta}_2 \in \Gamma_{n-m}$ is the smallest vector linearly independent of $\boldsymbol{\theta}_1$; $\boldsymbol{\theta}_3 \in \Gamma_{n-m}$ is the smallest vector linearly independent of $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$, etc.

For $r > |\boldsymbol{\theta}_1|$ let $\mathcal{A}_r \subset R(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ be the set

$$\mathcal{A}_r = \{\boldsymbol{\xi} \in R(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n) : |\langle \boldsymbol{\xi}, \boldsymbol{\nu}_j \rangle| < 1, j = 1, 2, \ldots, m, \ \& \ |\boldsymbol{\xi}| < r\}.$$

This set is obviously convex and symmetric about the origin. Moreover,

$$\mathrm{vol}(\mathcal{A}_r) \gg r^{n-m} \prod_{j=1}^{m} |\boldsymbol{\nu}_j|^{-1}. \qquad (4.3)$$

Applying again Minkowski's convex body theorem, we find that the set $\mathcal{A}_r$ contains at least

$$N = [2^{-n}|\Gamma|^{-1} \mathrm{vol}(\mathcal{A}_r)] \gg |\Gamma|^{-1} r^{n-m} \prod_{j=1}^{m} |\boldsymbol{\nu}_j|^{-1} \qquad (4.4)$$

pairs of points $\pm\boldsymbol{\mu}_k \in \Gamma$, $k = 1, \ldots, N$. Obviously, each $\boldsymbol{\mu}_k$ is orthogonal to each $\boldsymbol{\nu}_j$. Suppose, $r < |\boldsymbol{\theta}_{n-m}|$. Then, obviously, $|\boldsymbol{\theta}_p| \leq r < |\boldsymbol{\theta}_{p+1}|$ for some $p \leq n - m - 1$. The dimension of $R(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N)$ is then $\leq p$, and each $\boldsymbol{\mu}_k$ is a linear combination of $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p$ with integer coefficients. Denote $\Gamma_p := Z(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p)$. Minkowski's second theorem (see, e.g., [1], §VIII.2, Theorem I) shows that

$$\prod_{l=1}^{p} |\boldsymbol{\theta}_l| \ll |\Gamma_p|. \qquad (4.5)$$

A simple packing argument shows that $N|\Gamma_p|$ is smaller than the volume of the ball of radius $(p+1)r$ in $R(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p)$, i.e. $N|\Gamma_p| \ll r^p$. Estimate (4.5) implies

$$N \prod_{l=1}^{p} |\boldsymbol{\theta}_l| \ll r^p.$$

Therefore, if the condition

$$N \prod_{l=1}^{p} |\boldsymbol{\theta}_l| > C r^p \qquad (4.6)$$

is satisfied, where C is sufficiently large, this implies that $r \geq |\boldsymbol{\theta}_{p+1}|$.

Estimate (4.4) shows that if

$$r > C|\Gamma| \prod_{j=1}^{m} |\boldsymbol{\nu}_j| \prod_{l=1}^{n-m-1} |\boldsymbol{\theta}_l|^{-1}, \tag{4.7}$$

then condition (4.6) with $p = n - m - 1$ will be satisfied and this would guarantee that $r > |\boldsymbol{\theta}_{n-m}|$. In other words, if r is greater than the RHS of (4.7), then $r > |\boldsymbol{\theta}_{n-m}|$. This implies

$$|\boldsymbol{\theta}_{n-m}| \ll |\Gamma| \prod_{j=1}^{m} |\boldsymbol{\nu}_j| \prod_{l=1}^{n-m-1} |\boldsymbol{\theta}_l|^{-1},$$

which finishes the proof. $\square$

Let $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n \in \mathbb{R}^d$ $(n \leq d)$. We denote by $\mathbf{M} = \mathbf{M}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n)$ the $d \times n$ matrix whose jth column equals $\boldsymbol{\nu}_j$. We also denote

$$\|\boldsymbol{\nu}_1 \wedge \cdots \wedge \boldsymbol{\nu}_n\|_2 := \sqrt{\det(\mathbf{M}^*\mathbf{M})} \tag{4.8}$$

($\mathbf{M}^*\mathbf{M}$ is obviously non-negative, and so is the determinant). The reason for the notation is that we can think of $\|\boldsymbol{\nu}_1 \wedge \cdots \wedge \boldsymbol{\nu}_n\|_2$ as being the Hilbert-Schmidt norm of the tensor $\boldsymbol{\nu}_1 \wedge \cdots \wedge \boldsymbol{\nu}_n$.

Lemma 4.6. *Let $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m \in \mathbb{R}^d$. Let $V_1 = R(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n)$ and $V_2 = R(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m)$. Let α be the angle between V_1 and V_2. Then the following inequality holds:*

$$\sin \alpha \geq \frac{\|\boldsymbol{\nu}_1 \wedge \cdots \wedge \boldsymbol{\nu}_n \wedge \boldsymbol{\mu}_1 \wedge \cdots \wedge \boldsymbol{\mu}_m\|_2}{\|\boldsymbol{\nu}_1 \wedge \cdots \wedge \boldsymbol{\nu}_n\|_2 \|\boldsymbol{\mu}_1 \wedge \cdots \wedge \boldsymbol{\mu}_m\|_2} \tag{4.9}$$

Proof. If we multiply matrix $\mathbf{M}$ from the right by a non-singular $n \times n$ matrix $\mathbf{B}$, the expression (4.8) is multiplied by $\det \mathbf{B}$. This observation shows that elementary transformations of the set of vectors $\boldsymbol{\nu}$ (i.e. multiplying $\boldsymbol{\nu}_j$ by a non-zero scalar, adding $\boldsymbol{\nu}_j$ to $\boldsymbol{\nu}_k$, etc) do not change both sides of (4.9); the same is the case for elementary transformations of the vectors $\boldsymbol{\mu}$. Thus, we may assume that vectors $\boldsymbol{\nu}$ form an orthonormal basis of V_1, vectors $\boldsymbol{\mu}$ form an orthonormal basis of V_2, and the angle between $\boldsymbol{\nu}_1$ and $\boldsymbol{\mu}_1$ equals α. Notice that now the denominator of the RHS of (4.9) equals 1. Next, we notice that an orthogonal change of coordinates results in multiplying $\mathbf{M}$ from the left by a $d \times d$ orthogonal matrix and thus doesn't change (4.8) and the RHS of (4.9); the LHS of (4.9) is obviously invariant under an orthogonal change of coordinates as well. Assume, without loss of generality, that $n \geq m$. Then, applying an orthogonal change of coordinates, we can make our vectors to have the following form: $\boldsymbol{\nu}_j = \mathbf{e}_j$ $(j = 1, \ldots, n$, where $\mathbf{e}_j$ are

standard basis vectors), $\boldsymbol{\mu}_j = p_j \mathbf{e}_j + q_j \mathbf{e}_{n+j}$, $(p_j, q_j \geq 0, p_j^2 + q_j^2 = 1)$. Elementary geometry implies $\cos \alpha = p_1$, and so $\sin \alpha = q_1$. Computing the determinant, we obtain:

$$\|\boldsymbol{\nu}_1 \wedge \cdots \wedge \boldsymbol{\nu}_n \wedge \boldsymbol{\mu}_1 \wedge \cdots \wedge \boldsymbol{\mu}_m\|_2 = \prod_{j=1}^{n} q_j \leq q_1.$$

The lemma is proved. $\square$

Lemma 4.7. *Let* $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m \in \mathbb{Z}^d \cap B(R)$ *be linearly independent. Let* $V_1 = R(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n)$ *and* $V_2 = R(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m)$*. Then the angle between* V_1 *and* V_2 *is* $\gg \prod_{j=1}^{n} |\boldsymbol{\nu}_j|^{-1} \prod_{l=1}^{m} |\boldsymbol{\mu}_l|^{-1} \geq R^{-n-m}$*.*

Remark 4.8. It is not difficult to see that the power $-n - m$ in lemma 4.7 is optimal.

Proof. We use the inequality (4.9) and notice that the numerator of the RHS is a square root of an integer number (since all vectors involved are integer) and is non-zero (since the vectors are linearly independent). Therefore, the numerator is at least 1. The denominator is, obviously, $\ll \prod_{j=1}^{n} |\boldsymbol{\nu}_j| \prod_{l=1}^{m} |\boldsymbol{\mu}_l|$. This finishes the proof. $\square$

Using the same argument we have used while proving Corollary 4.3, we can prove the following

Corollary 4.9. *Let* $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m \in \mathbb{Z}^d \cap B(R)$ *be linearly independent. Let* $V_1 = R(\mathbf{F}\boldsymbol{\nu}_1, \ldots, \mathbf{F}\boldsymbol{\nu}_n)$ *and* $V_2 = R(\mathbf{F}\boldsymbol{\mu}_1, \ldots, \mathbf{F}\boldsymbol{\mu}_m)$*. Then the angle between* V_1 *and* V_2 *is* $\gg R^{-n-m}$

Lemma 4.10. *Let* $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n \in \mathbb{Z}^d \cap B(R)$ *and* $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m \in \mathbb{Z}^d \cap B(R)$ *be two sets. We assume that each set consists of linearly independent vectors (but the union of two sets is not necessary linearly independent). Let* $V_1 = R(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n)$ *and* $V_2 = R(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m)$*. Suppose,* $\dim(V_1 \cap V_2) = l$*. Then there are* l *integer linearly independent vectors* $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l \in (\mathbb{Z}^d \cap V_1 \cap V_2)$ *such that* $|\boldsymbol{\theta}_j| \ll R^{m+n-l+1}$*. Moreover, the angle between orthogonal complements to* $(V_1 \cap V_2)$ *in* V_1 *and* V_2 *is bounded below by* $CR^{-\alpha}$*,* $\alpha = \alpha(n, m, l) = n + m + 2l(m + n - l + 1)$*.*

Proof. Denote $\mathbf{M} = \mathbf{M}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n, -\boldsymbol{\mu}_1 \ldots, -\boldsymbol{\mu}_m)$. The rank of $\mathbf{M}$ equals $k := m + n - l$. Without loss of generality we can assume that the top left $k \times k$ minor of this matrix is non-zero (otherwise we just change the order of the vectors $-\boldsymbol{\mu}_j$ or the order of the coordinates x_j). In order to find the basis of the intersection $V_1 \cap V_2$ we have to solve the system of equations

$$\mathbf{M}\mathbf{t} = 0. \tag{4.10}$$

Indeed, if $\mathbf{t} = (t_1, \ldots, t_{n+m})^T$ is a solution of (4.10), then $\sum_{p=1}^n t_p \boldsymbol{\nu}_p = \sum_{q=1}^m t_{n+q} \boldsymbol{\mu}_q \in V_1 \cap V_2$. Now the simple linear algebra tells us that the basis of solutions of (4.10) is formed by the vectors of the form $(s_1, \ldots, s_k, 1, 0, \ldots, 0)$, $(t_1, \ldots, t_k, 0, 1, \ldots, 0)$,..., $(\tau_1, \ldots, \tau_k, 0, \ldots, 0, 1)$. Using Cramer's rule, we find that each of the numbers s_j, t_j, τ_j, etc is a ratio of two determinants, each of them an integer number $\ll R^k$; moreover, the denominator is the same for all of the numbers s_j, t_j, etc. After multiplication by the denominator, we obtain an integer basis of solutions of (4.10) with entries $\ll R^k$. For any such solution $\mathbf{t}$ the following estimate holds: $|\sum_{p=1}^n t_p \boldsymbol{\nu}_p| \ll R^{k+1}$. This proves the first statement of lemma. To prove the second statement, we first use Lemma 4.5 to construct integer bases $\{\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_{n-l}\}$ and $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{m-l}\}$ of the orthogonal complements to $V_1 \cap V_2$ in V_1 and V_2 correspondingly with properties

$$\prod_{j=1}^{n-l} |\boldsymbol{\eta}_j| \ll R^{n+l(m+n-l+1)}$$

and

$$\prod_{j=1}^{m-l} |\boldsymbol{\xi}_j| \ll R^{m+l(m+n-l+1)}.$$

Now Lemma 4.7 produces the required estimate. This finishes the proof. $\qquad\square$

Using the same argument we have used while proving Corollary 4.3, we can prove the following

Corollary 4.11. *Let* $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_n, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m \in \mathbb{Z}^d \cap B(R)$ *be two linearly independent families of vectors. Let* $V_1 = R(\mathbf{F}\boldsymbol{\nu}_1, \ldots, \mathbf{F}\boldsymbol{\nu}_n)$ *and* $V_2 = R(\mathbf{F}\boldsymbol{\mu}_1, \ldots, \mathbf{F}\boldsymbol{\mu}_m)$. *Then the angle between orthogonal complements to* $V_1 \cap V_2$ *in* V_1 *and* V_2 *is* $\gg R^{-\alpha(n,m,l)}$.

5. REDUCTION TO INVARIANT SUBSPACES

Let $\lambda = \rho^2$ be a large real number. In this section, we use lemma 3.2 to construct the family of operators $\tilde{H}(\mathbf{k})$ the spectrum of which (or at least the part of the spectrum near λ) is close to the spectrum of $H(\mathbf{k})$. Consider the truncated potential

$$V'(\mathbf{x}) = \sum_{\mathbf{m} \in B(R) \cap \mathbb{Z}^d} \hat{V}(\mathbf{m}) e_{\mathbf{m}}(\mathbf{x}), \tag{5.1}$$

where

$$e_{\mathbf{m}}(\mathbf{x}) := \frac{1}{(2\pi)^{d/2}} e^{i\langle \mathbf{m}, \mathbf{x} \rangle}, \quad \mathbf{m} \in \mathbb{Z}^d$$

and

$$\hat{V}(\mathbf{m}) = \int_{[0,2\pi]^d} V(\mathbf{x})e_{-\mathbf{m}}(\mathbf{x}) \tag{5.2}$$

are the Fourier coefficients of V. R is a large parameter the precise value of which will be chosen later; at the moment we just state that $R \sim \rho^\gamma$ with $\gamma > 0$ being small. Throughout the text, we will prove various statements which will hold under conditions of the type $R < \rho^{\gamma_j}$. After each statement of this type, we will always assume, without possibly specifically mentioning, that these conditions are always satisfied in what follows; at the end, we will choose $\gamma = \min \gamma_j$.

Since V is smooth, for each n we have

$$\sup_{\mathbf{x}\in\mathbb{R}^d} |V(\mathbf{x}) - V'(\mathbf{x})| < C_n R^{-n}. \tag{5.3}$$

This implies that if we denote $H'(\mathbf{k}) := H_0(\mathbf{k}) + V'$ with the domain $\mathcal{D}(\mathbf{k})$, the following estimate holds for all n:

$$|\mu_j(H(\mathbf{k})) - \mu_j(H'(\mathbf{k}))| < C_n R^{-n}. \tag{5.4}$$

Throughout this and the next two sections, we will work with the truncated operators $H'(\mathbf{k})$. These sections will be devoted to the construction of mappings f, g with properties specified in the introduction. Let $M \in \mathbb{N}$ be a fixed number. For each natural j we denote

$$\Theta_j := \mathbb{Z}^d \cap B(jR), \ \Theta_0 := \{0\}, \ \Theta'_j := \Theta_j \setminus \{0\}; \tag{5.5}$$

Let $\mathfrak{V} \subset \mathbb{R}^d$ be a linear subspace of dimension n and $r > 0$. We say that $\mathfrak{V}$ is an integer r-subspace if $\mathfrak{V} = R(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$ and each $\boldsymbol{\theta}_j$ is an integer vector with length smaller than r. The set of all integer r-subspaces of dimension n will be denoted by $\mathcal{V}(r, n)$. We mostly will be dealing with $\mathcal{V}(6MR, n)$; for brevity we will denote $\mathcal{V}(n) := \mathcal{V}(6MR, n)$. If $\boldsymbol{\xi} \in \mathbb{R}^d$ and $\mathfrak{V} \in \mathcal{V}(n)$, we denote $\boldsymbol{\xi}_{\mathfrak{V}}$ and $\boldsymbol{\xi}_{\mathfrak{V}}^\perp$ vectors such that

$$\boldsymbol{\xi} = \boldsymbol{\xi}_{\mathfrak{V}} + \boldsymbol{\xi}_{\mathfrak{V}}^\perp, \ \boldsymbol{\xi}_{\mathfrak{V}} \in \mathfrak{V}, \ \mathbf{G}\boldsymbol{\xi}_{\mathfrak{V}}^\perp \perp \mathfrak{V}. \tag{5.6}$$

If $\mathfrak{V} \in \mathcal{V}(n)$, we put $\Theta(\mathfrak{V}) := \Theta_{6M} \cap \mathfrak{V}$, $\Theta'(\mathfrak{V}) := \Theta(\mathfrak{V}) \setminus \{0\}$. By p, q_n ($n = 1, \ldots, d$) we denote positive constants smaller than $1/3$; the precise value of these constants will be specified later; we also denote $K = \rho^p$ and $L_n = \rho^{q_n}$.

Let $\mathfrak{V} \in \mathcal{V}(n)$. We denote

$$\Xi_0(\mathfrak{V}) := \{\boldsymbol{\xi} \in \mathcal{A}, |\boldsymbol{\xi}_{\mathfrak{V}}| < L_n\}, \tag{5.7}$$

$$\Xi_1(\mathfrak{V}) := \big(\Xi_0(\mathfrak{V}) + \mathfrak{V}\big) \cap \mathcal{A}, \tag{5.8}$$

$$\Xi_2(\mathfrak{V}) := \Xi_1(\mathfrak{V}) \setminus \big(\cup_{m=n+1}^d \cup_{\mathfrak{W}\in\mathcal{V}(m):\,\mathfrak{V}\subset\mathfrak{W}} \Xi_1(\mathfrak{W})\big), \tag{5.9}$$

$$\Xi_3(\mathfrak{V}) := \Xi_2(\mathfrak{V}) + B(\mathfrak{V}, K), \tag{5.10}$$

and finally,

$$\Xi(\mathfrak{V}) := \Xi_3(\mathfrak{V}) + \Theta_M. \tag{5.11}$$

These objects (especially $\Xi_3(\mathfrak{V})$ and $\Xi(\mathfrak{V})$) play a crucial role in what follows; the pictures of them are shown in Figures 1-4 in the case $d = 2$ (here, the integer subspaces $\mathfrak{V}$ are 1-dimensional, so $\mathfrak{V} = R(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \Theta'$; we have called $\Xi_j(\boldsymbol{\theta}) := \Xi_j(R(\boldsymbol{\theta}))$).
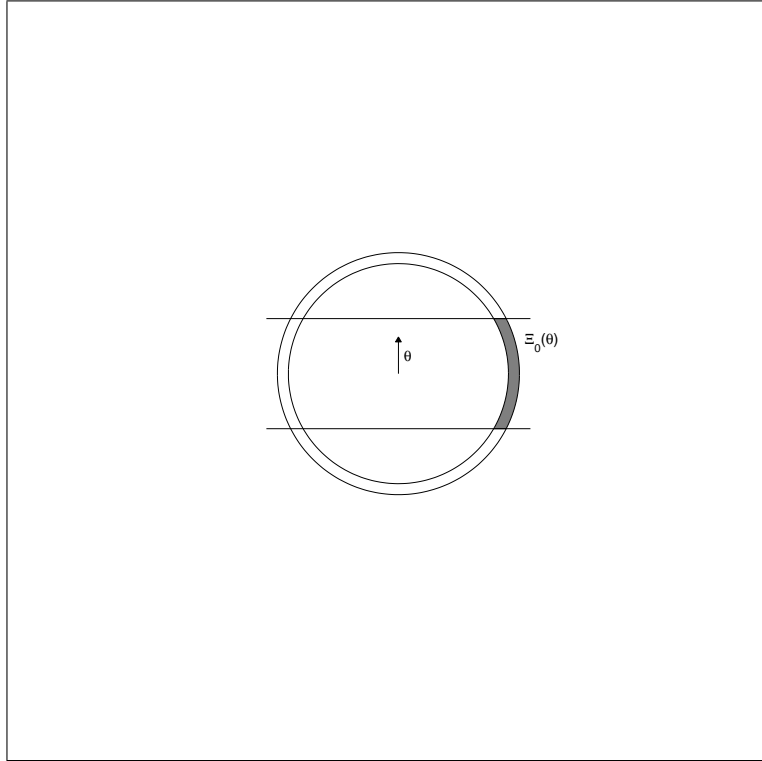
FIGURE 1. The set $\Xi_0(\boldsymbol{\theta})$ in the 2-dimensional case

It may seem that the definition of these objects is overcomplicated; for example, one may be tempted to define $\Xi_3(\mathfrak{V})$ by Figure 5. This definition is indeed simpler and it would work in the 2-dimensional case; however, if we try to extend this definition to higher dimensions, we would find out that lemma 5.12 no longer holds. One more remark concerning the definitions of the sets Ξ is that it is very difficult to make a mental picture of them in high dimensions (even when $d = 3$). A good approach to working with these sets is to do it on a purely formal level, without trying to imagine how they look like.

We also put

$$\mathcal{D} := \cup_{m=1}^{d} \cup_{\mathfrak{W} \in \mathcal{V}(m)} \Xi_1(\mathfrak{W}) \tag{5.12}$$
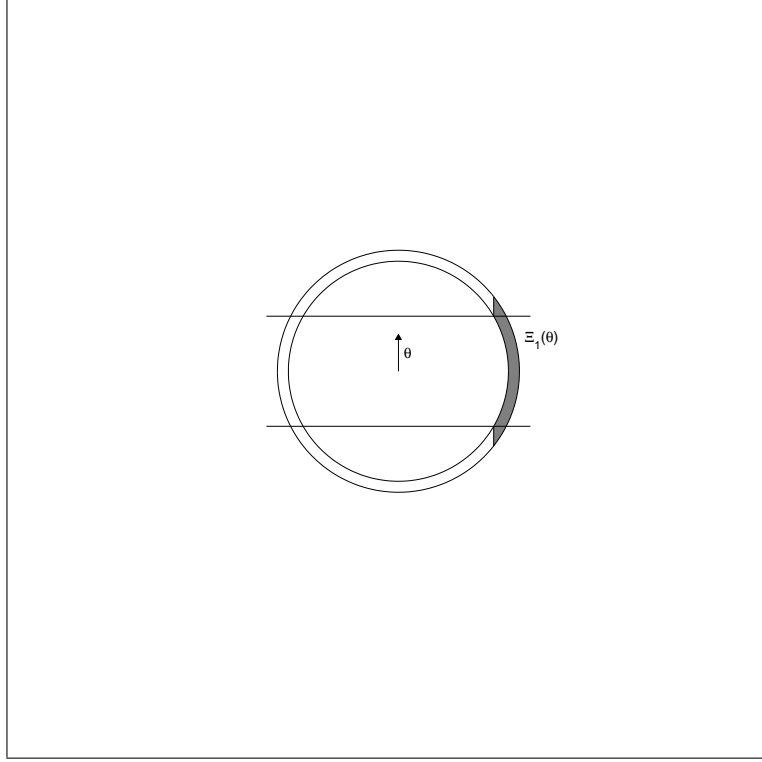
FIGURE 2. The set $\Xi_1(\boldsymbol{\theta}) = \Xi_2(\boldsymbol{\theta})$ in the 2-dimensional case

and

$$\mathcal{B} := \mathcal{A} \setminus \mathcal{D}. \tag{5.13}$$

We will often call the set $\mathcal{D}$ *the resonance region* and the set $\mathcal{B}$ *the non-resonance region*.

Note that the definitions (5.7)–(5.11) make sense for the subspace $\mathfrak{U}_0 := \{0\} \in \mathcal{V}(0)$. In particular, we have $\Xi_0(\mathfrak{U}_0) = \Xi_1(\mathfrak{U}_0) = \mathcal{A}$, $\Xi_2(\mathfrak{U}_0) = \Xi_3(\mathfrak{U}_0) = \mathcal{B}$, and

$$\Xi(\mathfrak{U}_0) = \mathcal{B} + \Theta_M. \tag{5.14}$$

Let us now formulate several properties of the sets Ξ_j. In what follows, we always assume that ρ and R are sufficiently large. We also assume that $L_n = \rho^{q_n}$ with $q_{n+1} \geq q_n + 3p$ for all n, $q_d \leq 1/3$, and $K = \rho^p$ with $q_1 \geq 3p > 0$. We also put $q_0 = 0$ so that $L_0 = 1$. From now on, we fix the values p and q_n satisfying these conditions; say, we put

$$q_n = 3np, \qquad\qquad p = (9d)^{-1}. \tag{5.15}$$
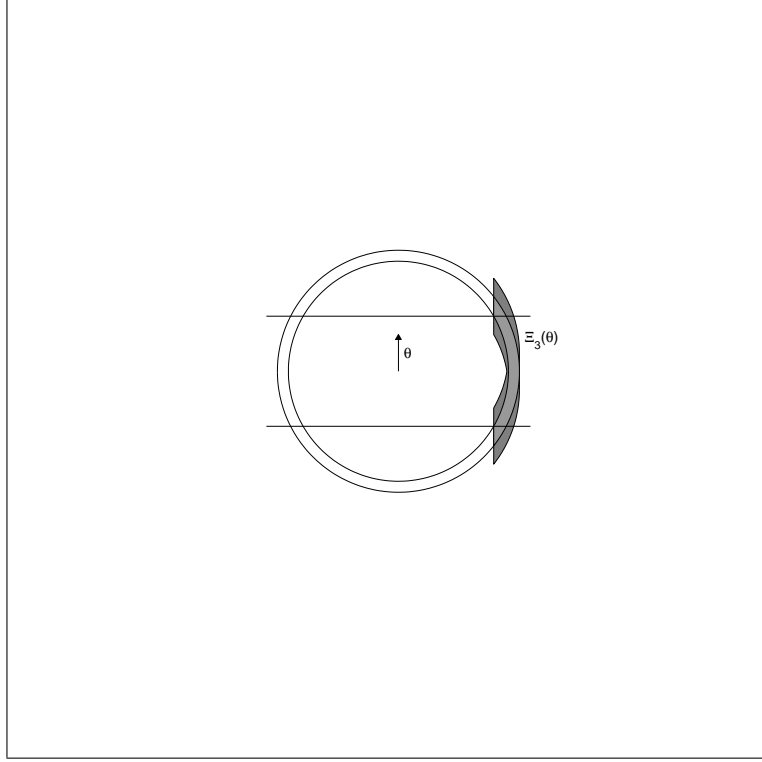
FIGURE 3. The set $\Xi_3(\boldsymbol{\theta})$ in the 2-dimensional case

Finally, we assume that $M > 2$ and that $\rho^p > R^{2\beta}$, where β is the maximal possible value the exponent $\alpha(n, m, l)$ from lemma 4.10 can attain.

Lemma 5.1. $\Xi_0(\mathbb{R}^d) = \emptyset$.

Proof. This statement is obvious since if $\mathfrak{V} = \mathbb{R}^d$, then for each $\boldsymbol{\xi}$ we have $\boldsymbol{\xi} = \boldsymbol{\xi}_{\mathfrak{V}}$; therefore one cannot have a point $\boldsymbol{\xi} \in \mathcal{A}$ with $|\boldsymbol{\xi}_{\mathfrak{V}}| < L_d \leq \rho^{1/3}$. $\qquad\square$

Lemma 5.2. *Let* $\mathfrak{V} \in \mathcal{V}(n)$, $0 \leq n < d$, *and* $\boldsymbol{\xi} \in \Xi_1(\mathfrak{V})$. *Then* $|\boldsymbol{\xi}_{\mathfrak{V}}| < 2L_n$.

Proof. The condition $\boldsymbol{\xi} \in \Xi_1(\mathfrak{V})$ means that $\boldsymbol{\xi} \in \mathcal{A}$ and there exists $\boldsymbol{\xi}' \in \mathcal{A}$, $|\boldsymbol{\xi}'_{\mathfrak{V}}| < L_n$ such that $\boldsymbol{\xi} - \boldsymbol{\xi}' \in \mathfrak{V}$. These conditions imply

$$||\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}|^2 - |\mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}}|^2| = ||\mathbf{F}\boldsymbol{\xi}|^2 - |\mathbf{F}\boldsymbol{\xi}'|^2| \ll 1.$$

Now the statement is obvious. $\qquad\square$

Corollary 5.3. *If* $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$, *then* $|\boldsymbol{\xi}_{\mathfrak{V}}| \ll L_n$.
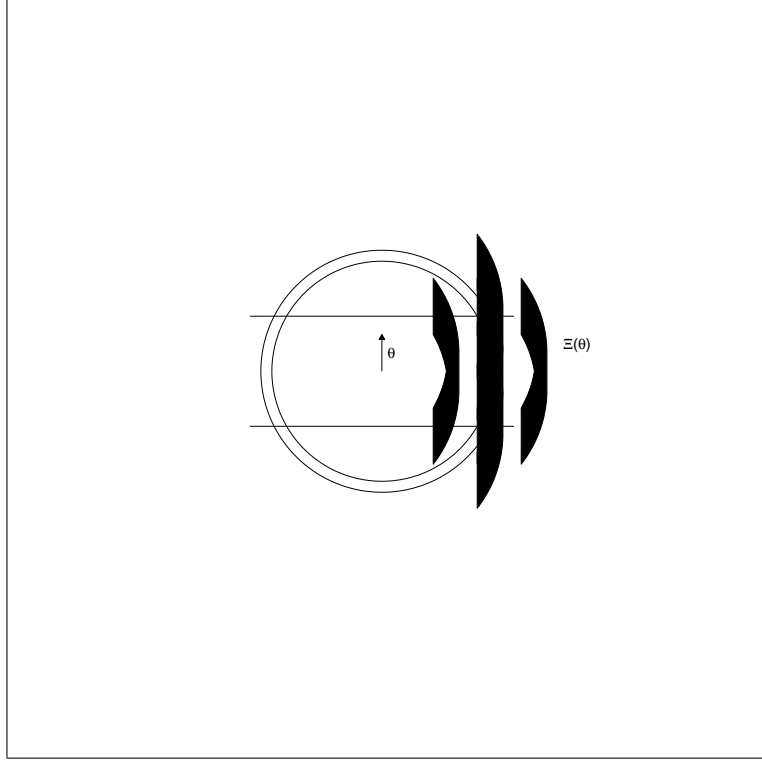
FIGURE 4. The set $\Xi(\boldsymbol{\theta})$ in the 2-dimensional case; here, $\Theta = \{(0,0), (\pm 1, 0), (0, \pm 1)\}$ consists of five elements.

Lemma 5.4. *Suppose, $\mathfrak{V}_1 \in \mathcal{V}(n_1)$ and $\mathfrak{V}_2 \in \mathcal{V}(n_2)$ are two subspaces such that neither of them is contained in the other one. Let $\boldsymbol{\xi}_j \in \Xi_2(\mathfrak{V}_j)$. Then $|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| > L_1$.*

Proof. The conditions of lemma imply $|(\boldsymbol{\xi}_j)_{\mathfrak{V}_j}| \ll L_{n_j}$, $j = 1, 2$. Let $\mathfrak{W} = \mathfrak{V}_1 + \mathfrak{V}_2$, $\mathfrak{U} = \mathfrak{V}_1 \cap \mathfrak{V}_2$, $\dim \mathfrak{U} = l$. Then $\mathfrak{W}$ is an integer $6MR$-subspace, say $\mathfrak{W} \in \mathcal{V}(m)$. Also, conditions of lemma imply that $\mathfrak{W} \neq \mathfrak{V}_j$, so $m > n_j$. Suppose, the statement of lemma does not hold, i.e. $|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| \le L_1$. Then $|(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)_{\mathfrak{V}_2}| \ll L_1$ and thus $|(\boldsymbol{\xi}_1)_{\mathfrak{V}_2}| \ll L_{n_2}$. By corollary 4.11, the angle between $\mathbf{F}\mathfrak{V}_1$ and $\mathbf{F}\mathfrak{V}_2$ is at least $CR^{-\alpha(n_1, n_2, l)}$. Since the projections of $\boldsymbol{\xi}_1$ onto $\mathfrak{V}_1$ and $\mathfrak{V}_2$ are smaller than L_{n_1} and $2L_{n_2}$ respectively, it is a simple geometry to deduce that $|(\boldsymbol{\xi}_1)_{\mathfrak{W}}| \ll (L_{n_1} + L_{n_2})R^{\alpha(n_1, n_2, l)}$. Due to the conditions stated before lemma 5.1, this implies $|(\boldsymbol{\xi}_1)_{\mathfrak{W}}| < L_m$. Therefore, $\boldsymbol{\xi}_1 \in \Xi_1(\mathfrak{W})$. Now definition (5.9) implies that $\boldsymbol{\xi}_1 \notin \Xi_2(\mathfrak{V}_1)$, which contradicts our assumptions. Thus, $|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| > L_1$. $\square$
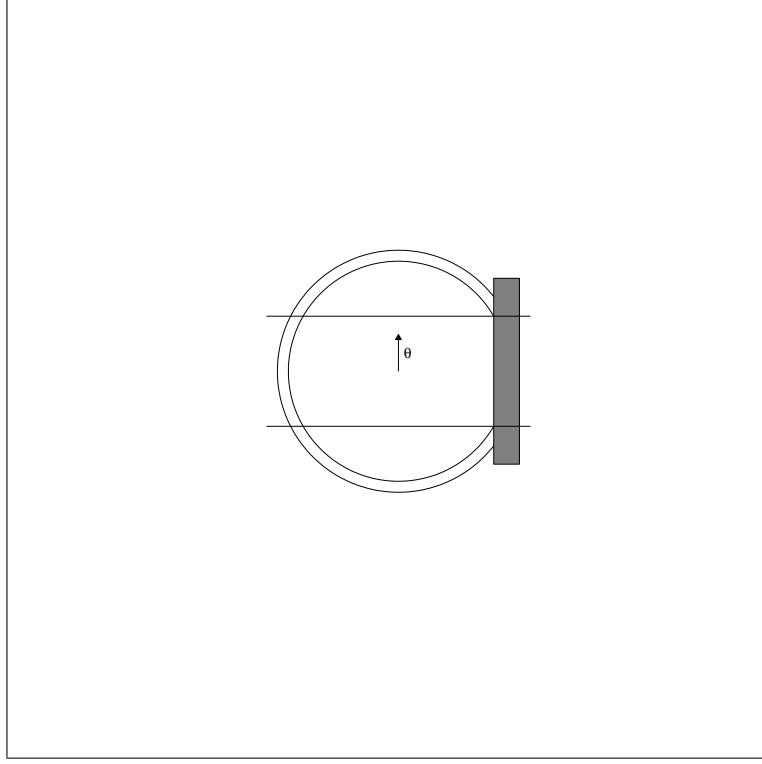
FIGURE 5. Bad definition of the set $\Xi_3(\boldsymbol{\theta})$ in the 2-dimensional case

Corollary 5.5. *Suppose,* $\mathfrak{V}_1 \in \mathcal{V}(n_1)$ *and* $\mathfrak{V}_2 \in \mathcal{V}(n_2)$ *are two subspaces such that neither of them is contained in the other one. Let* $\boldsymbol{\xi}_j \in \Xi(\mathfrak{V}_j)$. *Then* $|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| \gg L_1$.

Lemma 5.6. *Let* $\mathfrak{V} \in \mathcal{V}(n)$ *and* $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$. *Then* $||\mathbf{F}\boldsymbol{\xi}|^2 - \rho^2| \ll KL_n$ *and* $||\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|^2 - \rho^2| \ll L_n^2$.

Proof. The assumption of lemma imply that there exists $\boldsymbol{\eta} \in \Xi_2(\mathfrak{V})$ such that $\boldsymbol{\xi} - \boldsymbol{\eta} \in \mathfrak{V}$ and $|\boldsymbol{\xi} - \boldsymbol{\eta}| < K$. Lemma 5.2 implies $|\boldsymbol{\xi}_{\mathfrak{V}}| \ll L_n$, and thus $||\mathbf{F}\boldsymbol{\xi}|^2 - |\mathbf{F}\boldsymbol{\eta}|^2| = ||\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}|^2 - |\mathbf{F}\boldsymbol{\eta}_{\mathfrak{V}}|^2| \ll KL_n$. The first statement now follows from the fact that $\boldsymbol{\eta} \in \mathcal{A}$. Now we compute:

$$\rho^2 - |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|^2 = |\mathbf{F}\boldsymbol{\xi}|^2 + O(KL_n) - |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|^2 = |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}|^2 + O(KL_n) = O(L_n^2)$$

by corollary 5.3. $\square$

Factorizing the LHS's of the estimates from this lemma, we immediately obtain the following

Corollary 5.7. *Let* $\mathfrak{V} \in \mathcal{V}(n)$ *and* $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$. *Then* $||\mathbf{F}\boldsymbol{\xi}| - \rho| \ll \rho^{p+q_n-1}$ *and* $||\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}| - \rho| \ll \rho^{2q_n-1}$.

Lemma 5.8. *Let $\mathfrak{V} \in \mathcal{V}(n)$ and $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$. Suppose, for some $\boldsymbol{\eta} \in \mathcal{A}$ we have $\boldsymbol{\xi} - \boldsymbol{\eta} \in \mathfrak{V}$. Then $\boldsymbol{\eta} \in \Xi_2(\mathfrak{V})$.*

Proof. Definition (5.8) implies that $\boldsymbol{\eta} \in \Xi_1(\mathfrak{V})$. Therefore, in order to prove our lemma, we need to show that for any $\mathfrak{W} \in \mathcal{V}(m)$ $(m > n)$, $\mathfrak{V} \subset \mathfrak{W}$, we have $\boldsymbol{\eta} \notin \Xi_1(\mathfrak{W})$. Suppose, this is not the case and $\boldsymbol{\eta} \in \Xi_1(\mathfrak{W})$. Then the fact that $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$ means that there exists a vector $\tilde{\boldsymbol{\xi}} \in \Xi_2(\mathfrak{V})$ with $\boldsymbol{\xi} - \tilde{\boldsymbol{\xi}} \in \mathfrak{V}$. But then $\boldsymbol{\eta} - \tilde{\boldsymbol{\xi}} \in \mathfrak{V} \subset \mathfrak{W}$. Therefore, $\tilde{\boldsymbol{\xi}} \in \Xi_1(\mathfrak{W})$. This contradicts the assumption $\tilde{\boldsymbol{\xi}} \in \Xi_2(\mathfrak{V})$. The lemma is proved. $\qquad\qquad\square$

Lemma 5.9. *Let $\mathfrak{V} \in \mathcal{V}(n)$ and $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$. Suppose, for some $\boldsymbol{\eta} \in \mathfrak{V}$ we have $\boldsymbol{\alpha} := \boldsymbol{\xi} + \boldsymbol{\eta} \notin \Xi_3(\mathfrak{V})$. Then $||\mathbf{F}\boldsymbol{\alpha}|^2 - \rho^2| \gg K^2$.*

Proof. Let $\tilde{\boldsymbol{\alpha}}$ be the point which satisfies the following conditions: $\tilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha} \in \mathfrak{V}$, $\tilde{\boldsymbol{\alpha}} \in \mathcal{A}$, and the vector $\boldsymbol{\alpha}_{\mathfrak{V}}$ is a non-negative multiple of $\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}}$ (a simple geometrical argument shows that such a point always exists). Then lemma 5.8 implies that $\tilde{\boldsymbol{\alpha}} \in \Xi_2(\mathfrak{V})$. Therefore, since $\boldsymbol{\alpha} \notin \Xi_3(\mathfrak{V})$, we have $||\mathbf{F}\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}}| - |\mathbf{F}\boldsymbol{\alpha}_{\mathfrak{V}}|| = |\mathbf{F}\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}} - \mathbf{F}\boldsymbol{\alpha}_{\mathfrak{V}}| \gg K$. Moreover,

$$||\mathbf{F}\boldsymbol{\alpha}|^2 - |\mathbf{F}\tilde{\boldsymbol{\alpha}}|^2| = ||\mathbf{F}\boldsymbol{\alpha}_{\mathfrak{V}}|^2 - |\mathbf{F}\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}}|^2|$$

$$= ||\mathbf{F}\boldsymbol{\alpha}_{\mathfrak{V}}| - |\mathbf{F}\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}}|| \, ||\mathbf{F}\boldsymbol{\alpha}_{\mathfrak{V}}| + |\mathbf{F}\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}}|| \geq \left(|\mathbf{F}\boldsymbol{\alpha}_{\mathfrak{V}}| - |\mathbf{F}\tilde{\boldsymbol{\alpha}}_{\mathfrak{V}}|\right)^2 \gg K^2.$$

This finishes the proof, since $||\mathbf{F}\tilde{\boldsymbol{\alpha}}|^2 - \rho^2| \ll 1$. $\qquad\qquad\square$

Lemma 5.10. *Let $\mathfrak{V} \in \mathcal{V}(n)$ and $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$. Suppose, $\boldsymbol{\theta} \in \Theta'_{6M}$, $\boldsymbol{\theta} \notin \mathfrak{V}$. Denote $\boldsymbol{\eta} := \boldsymbol{\xi} + \boldsymbol{\theta}$. Then $||\mathbf{F}\boldsymbol{\eta}|^2 - \rho^2| \gg K^2 L_n$.*

Proof. Let $\mathfrak{W}$ be the linear span of $\mathfrak{V}$ and $\boldsymbol{\theta}$, and let $\mathfrak{U} := R(\boldsymbol{\theta})$ be the one-dimensional subspace.

Assume first that $|\boldsymbol{\xi}_{\mathfrak{U}}| \leq K^2 L_n$. Then, since $|\boldsymbol{\xi}_{\mathfrak{V}}| \leq L_n$, the geometrical argument similar to the one used in the proof of lemma 5.4 implies that $|\boldsymbol{\xi}_{\mathfrak{W}}| < L_{n+1}/2$ (recall that the assumption we have made on the exponents p and q_n imply that $L_{n+1} \geq K^3 L_n$). Since $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$, there exists a vector $\tilde{\boldsymbol{\xi}} \in \Xi_2(\mathfrak{V})$, $|\boldsymbol{\xi} - \tilde{\boldsymbol{\xi}}| < K$. Therefore, $|\tilde{\boldsymbol{\xi}}_{\mathfrak{W}}| \leq |\boldsymbol{\xi}_{\mathfrak{W}}| + |(\tilde{\boldsymbol{\xi}} - \boldsymbol{\xi})_{\mathfrak{W}}| < L_{n+1}$, which implies $\tilde{\boldsymbol{\xi}} \in \Xi_1(\mathfrak{W})$. This contradicts the condition $\tilde{\boldsymbol{\xi}} \in \Xi_2(\mathfrak{V})$.

Therefore, we must have $|\boldsymbol{\xi}_{\mathfrak{U}}| > K^2 L_n$. This implies

$$||\mathbf{F}\boldsymbol{\eta}|^2 - |\mathbf{F}\boldsymbol{\xi}|^2| = ||\mathbf{F}(\boldsymbol{\xi}_{\mathfrak{U}} + \boldsymbol{\theta})|^2 - |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{U}}|^2| \gg K^2 L_n.$$

Now it remains to notice that lemma 5.6 implies that $||\mathbf{F}\boldsymbol{\xi}|^2 - \rho^2| \ll KL_n$. This finishes the proof. $\qquad\qquad\square$

Corollary 5.11. *Let $\mathfrak{V} \in \mathcal{V}(n)$ and $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$. Suppose, $\boldsymbol{\theta} \in \Theta'_{6M}$ and $\boldsymbol{\eta} = \boldsymbol{\xi} + \boldsymbol{\theta} \notin \Xi_3(\mathfrak{V})$. Then $||\mathbf{F}\boldsymbol{\eta}|^2 - \rho^2| \gg K^2$.*

Proof. If $\boldsymbol{\theta} \in \mathfrak{V}$, then the statement follows from lemma 5.9, and if $\boldsymbol{\theta} \notin \mathfrak{V}$, the statement follows from lemma 5.10. $\square$

Lemma 5.12. *For each two different integer subspaces $\mathfrak{V}_j \in \mathcal{V}(n_j)$, $j = 1, 2$, $0 \leq n_j < d$ we have $(\Xi(\mathfrak{V}_1) + \Theta_1) \cap (\Xi(\mathfrak{V}_2) + \Theta_1) = \emptyset$.*

Proof. Suppose, $\boldsymbol{\xi} \in (\Xi(\mathfrak{V}_1) + \Theta_1) \cap (\Xi(\mathfrak{V}_2) + \Theta_1)$. Then corollary 5.5 implies that one of the subspaces $\mathfrak{V}_j$ is inside the other, say $\mathfrak{V}_1 \subset \mathfrak{V}_2$. Moreover, there exist two points, $\boldsymbol{\xi}_1 \in \Xi_3(\mathfrak{V}_1)$ and $\boldsymbol{\xi}_2 \in \Xi_3(\mathfrak{V}_2)$ such that $\boldsymbol{\theta}_j := \boldsymbol{\xi}_j - \boldsymbol{\xi} \in \Theta_{M+1}$. Then $\boldsymbol{\theta} := \boldsymbol{\xi}_1 - \boldsymbol{\xi}_2 = \boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 \in \Theta_{3M}$.

There are two possibilities: either $\boldsymbol{\theta} \in \mathfrak{V}_2$, or $\boldsymbol{\theta} \notin \mathfrak{V}_2$.

Assume first that $\boldsymbol{\theta} \in \mathfrak{V}_2$. Since $\boldsymbol{\xi}_j \in \Xi_3(\mathfrak{V}_j)$, there exist points $\tilde{\boldsymbol{\xi}}_j \in \Xi_2(\mathfrak{V}_j)$ such that $\tilde{\boldsymbol{\xi}}_j - \boldsymbol{\xi}_j \in \mathfrak{V}_j$, $|\tilde{\boldsymbol{\xi}}_j - \boldsymbol{\xi}_j| < K$. But then $\tilde{\boldsymbol{\xi}}_1 - \tilde{\boldsymbol{\xi}}_2 \in \mathfrak{V}_2$. Since $\tilde{\boldsymbol{\xi}}_2 \in \Xi_2(\mathfrak{V}_2) \subset \Xi_1(\mathfrak{V}_2)$, according to definition (5.8) this means that $\tilde{\boldsymbol{\xi}}_1 \in \Xi_1(\mathfrak{V}_2)$. Now definition (5.9) implies $\tilde{\boldsymbol{\xi}}_1 \notin \Xi_2(\mathfrak{V}_1)$ which contradicts our assumption.

Assume now $\boldsymbol{\theta} \notin \mathfrak{V}_2$. Then lemma 5.10 implies

$$||\mathbf{F}\boldsymbol{\xi}_1|^2 - \rho^2| = ||\mathbf{F}(\boldsymbol{\xi}_2 + \boldsymbol{\theta})|^2 - \rho^2| \gg K^2 L_{n_2}.$$

However, this contradicts the inequality $||\mathbf{F}\boldsymbol{\xi}_1|^2 - \rho^2| \ll K L_{n_1}$ which was established in lemma 5.6. $\square$

Corollary 5.13. *Each point $\boldsymbol{\xi} \in \mathcal{A}$ belongs to precisely one of the sets $\Xi(\mathfrak{V})$.*

Proof. Indeed, definitions (5.7)–(5.14) imply that each point $\boldsymbol{\xi} \in \mathcal{A}$ belongs to at least one of the sets $\Xi(\mathfrak{V})$. The rest follows from lemma 5.12. $\square$

Let us introduce more notation. Let $\mathcal{C} \subset \mathbb{R}^d$ be a measurable set. We denote by $\mathcal{P}^{(\mathbf{k})}(\mathcal{C})$ the orthogonal projection in $\mathfrak{H} = L^2([0, 2\pi]^d)$ onto the subspace spanned by the exponentials $e_{\boldsymbol{\xi}}(\mathbf{x})$, $\boldsymbol{\xi} \in \mathcal{C}$, $\{\boldsymbol{\xi}\} = \mathbf{k}$.

Lemma 5.14. *For arbitrary set $\mathcal{C} \subset \mathbb{R}^d$ and arbitrary $\mathbf{k}$ we have:*

$$V'\mathcal{P}^{(\mathbf{k})}(\mathcal{C}) = \mathcal{P}^{(\mathbf{k})}(\mathcal{C} + \Theta_1)V'\mathcal{P}^{(\mathbf{k})}(\mathcal{C}) \tag{5.16}$$

Proof. This follows from the obvious observation that if $\boldsymbol{\xi} = \mathbf{m} + \mathbf{k} \in \mathcal{C}$ and $|\mathbf{n}| \leq R$, then $\boldsymbol{\xi} + \mathbf{n} \in (\mathcal{C} + \Theta_1)$. $\square$

We are going to apply lemma 3.2 and now we will specify what are the projections P_j^l. The construction will be the same for all values of quasi-momenta, so often we will skip $\mathbf{k}$ from the superscripts. For each $\mathfrak{V} \in \mathcal{V}(n)$, $n = 0, 1, \ldots, d-1$ we put $P(\mathfrak{V}) := \mathcal{P}^{(\mathbf{k})}(\Xi(\mathfrak{V}))$. We also define $P_j(\mathfrak{V}) := \mathcal{P}^{(\mathbf{k})}((\Xi_3(\mathfrak{V}) + \Theta_j) \setminus (\Xi_3(\mathfrak{V}) + \Theta_{j-1}))$, $j = 1, \ldots, M$, $P_0(\mathfrak{V}) = \mathcal{P}^{(\mathbf{k})}(\Xi_3(\mathfrak{V}))$. We also denote $Q := I - (\sum_{\mathfrak{V}} P(\mathfrak{V}))$ (the sum

is over all integer $6MR$-subspaces of dimension $n = 0, 1, \ldots, d-1$).
Now we apply lemma 3.2 with the set of projections being $\{P(\mathfrak{V})\}$,
$J := [\lambda - 20v, \lambda + 20v]$, and $H_0 = H_0(\mathbf{k})$. Let us check that all the
conditions of lemma 3.2 are satisfied assuming, as before, that all the
conditions before lemma 5.1 are fulfilled. Indeed, lemmas 5.12 and
5.14 imply that $P(\mathfrak{V}_1)P(\mathfrak{V}_2) = 0$ and $P(\mathfrak{V}_1)VP(\mathfrak{V}_2) = 0$ for differ-
ent subsets $\mathfrak{V}_1$ and $\mathfrak{V}_2$ (in particular, Q is also a projection). Prop-
erties $P(\mathfrak{V}) = \sum_{j=0}^{M} P_j(\mathfrak{V})$, $P_j(\mathfrak{V})VP_l(\mathfrak{V}) = 0$ for $|j - l| > 1$ and
$P_j(\mathfrak{V})VQ = 0$ for $j < M$ follow from the construction of the projections
$P_j(\mathfrak{V})$ and lemma 5.14. Since $\mathcal{A} \subset \cup_{\mathfrak{V}} \Xi_3(\mathfrak{V})$, the distance between the
spectrum of QH_0Q and J is greater than $6v$. Corollary 5.11 implies
that the distances between the spectra of $P_j(\mathfrak{V})H_0P_j(\mathfrak{V})$, $j = 1, \ldots, M$
and J are $\gg K^2$. All these remarks imply that we can apply lemma
3.2 (or rather corollary 3.3) and, instead of studying eigenvalues in-
side J of $H'(\mathbf{k})$, study eigenvalues of $\tilde{H}(\mathbf{k}) := \sum_{\mathfrak{V}} P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$;
the distance between any eigenvalue of $H'(\mathbf{k})$ lying inside J and the
corresponding eigenvalue of $\tilde{H}(\mathbf{k})$ is $\ll \rho^{-4Mp}$.

To be more precise, we do the following. Assume $\boldsymbol{\xi} = \mathbf{n} + \mathbf{k} \in \mathcal{A}$.
Then $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ for some uniquely defined $\mathfrak{V} \in \mathcal{V}(n)$. In the following
sections, we will define a mapping $\tilde{g} : \boldsymbol{\xi} \mapsto \mu_{\tau(\boldsymbol{\xi})}(P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V}))$,
where $\tau = \tau(\boldsymbol{\xi})$ is a function with values in $\mathbb{N}$. The mapping $\tilde{g}$
will be an injection and any eigenvalue of $P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$ inside J
will have a pre-image under $\tilde{g}$. Then, $\tilde{g}(\boldsymbol{\xi})$ is also an eigenvalue of
$\sum_{\mathfrak{V}} P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V}) + QH'(\mathbf{k})Q$, say

$$\tilde{g}(\boldsymbol{\xi}) = \mu_{\tau_1(\boldsymbol{\xi})}\Big(\sum_{\mathfrak{V}} P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V}) + QH'(\mathbf{k})Q\Big).$$

Then lemma 3.2 implies that $|\tilde{g}(\boldsymbol{\xi}) - \mu_{\tau_1(\boldsymbol{\xi})}(H'(\mathbf{k}))| \ll \rho^{-4Mp}$. We
then define $f(\boldsymbol{\xi}) := \mu_{\tau_1(\boldsymbol{\xi})}(H(\mathbf{k}))$ so that $|\tilde{g}(\boldsymbol{\xi}) - f(\boldsymbol{\xi})| \ll \rho^{-4Mp}$. In
order to construct the mapping g, we compute $\tilde{g}$ (or at least obtain
an asymptotic formula for it) and then, roughly speaking, throw away
terms which are sufficiently small for our purposes.

In the next two sections, we discuss how to obtain an asymptotic
formula for $\tilde{g}$ when ρ is large. We will consider separately the case
$\boldsymbol{\xi} \in \Xi_2(\mathfrak{U}_0) = \mathcal{B}$ (recall that $\mathfrak{U}_0 = \{0\} \in \mathcal{V}(0)$ and we have called $\mathcal{B}$
the non-resonance region) and the case of $\boldsymbol{\xi}$ lying inside the resonance
region $\mathcal{D}$. We start by looking at the case $\boldsymbol{\xi} \in \mathcal{B}$.

6. Computation of the eigenvalues outside resonance layers

First of all, we notice that lemma 5.10 implies that the operator $P(\mathfrak{U}_0)H'(\mathbf{k})P(\mathfrak{U}_0)$ splits into the direct sum of operators. Namely,

$$P(\mathfrak{U}_0)H'(\mathbf{k})P(\mathfrak{U}_0) = \bigoplus \mathcal{P}^{(\mathbf{k})}(\boldsymbol{\xi} + \Theta_M)H'(\mathbf{k})\mathcal{P}^{(\mathbf{k})}(\boldsymbol{\xi} + \Theta_M), \qquad (6.1)$$

the sum being over all $\boldsymbol{\xi} \in \mathcal{B}$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$. We denote by $\tilde{g}(\boldsymbol{\xi})$ the eigenvalue of $\mathcal{P}^{(\mathbf{k})}(\boldsymbol{\xi} + \Theta_M)H'(\mathbf{k})\mathcal{P}^{(\mathbf{k})}(\boldsymbol{\xi} + \Theta_M)$ which lies within the distance v from $|\mathbf{F}\boldsymbol{\xi}|^2$ (Lemma 5.10 implies that this eigenvalue is unique). Our next task is to compute $\tilde{g}(\boldsymbol{\xi})$. In this section we will prove the following lemma:

Lemma 6.1. *Let $R < \rho^{pd^{-1}/2}$. Then the following asymptotic formula holds:*

$$\tilde{g}(\boldsymbol{\xi}) \sim |\mathbf{F}\boldsymbol{\xi}|^2$$

$$+ \sum_{r=1}^{\infty} \sum_{\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_r \in \Theta'_M} \sum_{n_1+\cdots+n_r \geq 2} A_{n_1,\ldots,n_r} \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_1 \rangle^{-n_1} \ldots \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_r \rangle^{-n_r} \qquad (6.2)$$

in the sense that for each $m \in \mathbb{N}$ we have

$$\tilde{g}(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2$$

$$- \sum_{r=1}^{m} \sum_{\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_r \in \Theta'_M} \sum_{2 \leq n_1+\cdots+n_r \leq m} A_{n_1,\ldots,n_r} \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_1 \rangle^{-n_1} \ldots \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_r \rangle^{-n_r}$$

$$= O(\rho^{-(m+1)p}).$$

$$(6.3)$$

uniformly over $R < \rho^{pd^{-1}/2}$. Here, $A_{n_1,\ldots,n_p}$ is a polynomial of the Fourier coefficients $\hat{V}(\boldsymbol{\eta}_j)$ and $\hat{V}(\boldsymbol{\eta}_j - \boldsymbol{\eta}_l)$ of the potential.

Proof. Let us denote

$$a(\boldsymbol{\eta}) = |\mathbf{F}(\boldsymbol{\xi} + \boldsymbol{\eta})|^2. \qquad (6.4)$$

The matrix of $P(\boldsymbol{\xi} + \Theta_M)H'(\mathbf{k})P(\boldsymbol{\xi} + \Theta_M)$ has the following form:

$$\begin{pmatrix} a(0) & \hat{V}(\boldsymbol{\eta}_1) & \hat{V}(\boldsymbol{\eta}_2) & \ldots & \hat{V}(\boldsymbol{\eta}_n) & \ldots \\ \hat{V}(\boldsymbol{\eta}_1) & a(\boldsymbol{\eta}_1) & \hat{V}(\boldsymbol{\eta}_2 - \boldsymbol{\eta}_1) & \ldots & \hat{V}(\boldsymbol{\eta}_n - \boldsymbol{\eta}_1) & \ldots \\ \hat{V}(\boldsymbol{\eta}_2) & \hat{V}(\boldsymbol{\eta}_2 - \boldsymbol{\eta}_1) & a(\boldsymbol{\eta}_2) & \ldots & \hat{V}(\boldsymbol{\eta}_n - \boldsymbol{\eta}_2) & \ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ldots \\ \hat{V}(\boldsymbol{\eta}_n) & \hat{V}(\boldsymbol{\eta}_n - \boldsymbol{\eta}_1) & \hat{V}(\boldsymbol{\eta}_n - \boldsymbol{\eta}_2) & \ldots & a(\boldsymbol{\eta}_n) & \ldots \\ \vdots & \vdots & \vdots & \ldots & \vdots & \ddots \end{pmatrix} \qquad (6.5)$$

The diagonal elements of this matrix equal $|\mathbf{F}(\boldsymbol{\xi} + \boldsymbol{\eta})|^2$ (with $\boldsymbol{\eta}$ running over Θ_M) and off-diagonal elements are Fourier coefficients of the potential (and are thus bounded). Let L be the number of columns of this matrix; obviously, $L \asymp R^d$.

Let us compute the characteristic polynomial $p(\mu)$ of (6.5). The definition of the determinant implies

$$p(\mu) = \Big(\prod_{\boldsymbol{\eta} \in \Theta_M} (a(\boldsymbol{\eta}) - \mu) \Big) + \sum_{m=2}^{L} J_m(\mu), \qquad (6.6)$$

where J_m consists of products of exactly $(L - m)$ diagonal terms of (6.5) and m off-diagonal terms. Put $J_m = J'_m + J''_m$, where J'_m (resp. J''_m) consists of all terms, not containing (resp. containing) $(a(0) - \mu)$. Then we can re-write (6.6) as

$$p(\mu) = \Big(\prod_{\boldsymbol{\eta} \in \Theta'_M} (a(\boldsymbol{\eta}) - \mu) \Big) \Big((a(0) - \mu) + I(\mu) \Big), \qquad (6.7)$$

where $I(\mu) := \sum_{m=1}^{L-1} I_m(\mu) + \sum_{m=2}^{L} \tilde{I}_m(\mu)$ with

$$I_m := \frac{J'_{m+1}}{\prod_{\boldsymbol{\eta} \in \Theta'_M} (a(\boldsymbol{\eta}) - \mu)}$$

and

$$\tilde{I}_m := \frac{J''_m}{\prod_{\boldsymbol{\eta} \in \Theta'_M} (a(\boldsymbol{\eta}) - \mu)}.$$

We can easily compute the first several terms:

$$I_1(\mu) := - \sum_{\boldsymbol{\eta} \in \Theta'_M} \frac{|\hat{V}(\boldsymbol{\eta})|^2}{a(\boldsymbol{\eta}) - \mu}, \qquad (6.8)$$

$$I_2(\mu) := - \sum_{\boldsymbol{\eta}, \boldsymbol{\eta}' \in \Theta'_M, \boldsymbol{\eta} \neq \boldsymbol{\eta}'} \frac{2\Re(\hat{V}(\boldsymbol{\eta}) \hat{V}(\boldsymbol{\eta} - \boldsymbol{\eta}') \overline{\hat{V}(\boldsymbol{\eta}')})}{(a(\boldsymbol{\eta}) - \mu)(a(\boldsymbol{\eta}') - \mu)},$$

$$\tilde{I}_2(\mu) := -(a(0) - \mu) \sum_{\boldsymbol{\eta}, \boldsymbol{\eta}' \in \Theta'_M, \boldsymbol{\eta} \neq \boldsymbol{\eta}'} \frac{|\hat{V}(\boldsymbol{\eta} - \boldsymbol{\eta}')|^2}{(a(\boldsymbol{\eta}) - \mu)(a(\boldsymbol{\eta}') - \mu)}. \qquad (6.9)$$

Overall, I_m is the sum of $O(R^{dn})$ terms of the form

$$\frac{W_m(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)}{(a(\boldsymbol{\eta}_1) - \mu) \ldots (a(\boldsymbol{\eta}_n) - \mu)}, \qquad (6.10)$$

and $\tilde{I}_m$ is the sum of $O(R^{dn})$ terms of the form

$$(a(0) - \mu) \frac{\tilde{W}_m(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)}{(a(\boldsymbol{\eta}_1) - \mu) \ldots (a(\boldsymbol{\eta}_n) - \mu)}. \qquad (6.11)$$

Here, $W_m(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ and $\tilde{W}_m(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n)$ are some polynomials of $\hat{V}(\boldsymbol{\eta}_j)$ and $\hat{V}(\boldsymbol{\eta}_j - \boldsymbol{\eta}_l)$.

On the interval $[a(0) - v, a(0) + v]$ the equation $p(\mu) = 0$ has a unique solution, which we have called $\tilde{g}(\boldsymbol{\xi})$; this is the solution of the equation $a(0) - \mu + I(\mu) = 0$. After denoting $F(\mu) := a(0) + I(\mu)$, this equation becomes equivalent to $\mu = F(\mu)$. Throughout the rest of the section we will assume that $\mu \in [a(0) - v, a(0) + v]$. Then, since $\boldsymbol{\xi} \in \mathcal{B}$, lemma 5.10 guarantees that $|a(\boldsymbol{\eta}) - a(0)| \gg \rho^{2p}$ for $\boldsymbol{\eta} \in \Theta'_M$. This implies $I_n(\mu) = O(R^{dn}\rho^{-2np}) = O(\rho^{-np})$; similarly, $\tilde{I}_n(\mu) = O(\rho^{-np})$. Computing the derivatives, we see that $\frac{d}{d\mu}I_n(\mu)$ and $\frac{d}{d\mu}\tilde{I}_n(\mu)$ are $O(\rho^{-np})$ as well. Slightly more careful analysis shows that in fact $I_1 = O(R^d\rho^{-4p}) = O(\rho^{-2p})$ and $\frac{d}{d\mu}I_1 = O(R^d\rho^{-4p}) = O(\rho^{-2p})$. Indeed, we have:

$$
\begin{aligned}
I_1(\mu) &= -\sum_{\boldsymbol{\eta}\in\Theta'_M} \frac{|\hat{V}(\boldsymbol{\eta})|^2}{a(\boldsymbol{\eta}) - \mu} \\
&= -\frac{1}{2}\sum_{\boldsymbol{\eta}\in\Theta'_M} |\hat{V}(\boldsymbol{\eta})|^2 \left(\frac{1}{a(\boldsymbol{\eta}) - \mu} + \frac{1}{a(-\boldsymbol{\eta}) - \mu}\right) \\
&= -\frac{1}{2}\sum_{\boldsymbol{\eta}\in\Theta'_M} |\hat{V}(\boldsymbol{\eta})|^2 \left(\frac{a(\boldsymbol{\eta}) + a(-\boldsymbol{\eta}) - 2\mu}{(a(\boldsymbol{\eta}) - \mu)(a(-\boldsymbol{\eta}) - \mu)}\right) \\
&= -\sum_{\boldsymbol{\eta}\in\Theta'_M} |\hat{V}(\boldsymbol{\eta})|^2 \left(\frac{a(0) + |\mathbf{F}\boldsymbol{\eta}|^2 - \mu}{(a(\boldsymbol{\eta}) - \mu)(a(-\boldsymbol{\eta}) - \mu)}\right),
\end{aligned}
\tag{6.12}
$$

and it remains to notice that $\sum_{\boldsymbol{\eta}\in\Theta'_M} |\hat{V}(\boldsymbol{\eta})|^2$ is bounded by the square of the L_2-norm of V. These estimates show that when $R < \rho^{pd-1/2}$, we have $I(\mu) = O(\rho^{-2p})$ and $\frac{d}{d\mu}F(\mu) = \frac{d}{d\mu}I(\mu) = O(\rho^{-2p})$. We will find $\tilde{\lambda}(\boldsymbol{\xi})$ using a sequence of approximations. We define a sequence μ_k in the following way: $\mu_0 = a(0)$, $\mu_{k+1} = F(\mu_k) = a(0) + I(\mu_k)$. Since $|\mu_{k+1} - \tilde{g}(\boldsymbol{\xi})| = |F(\mu_k) - F(\tilde{g}(\boldsymbol{\xi}))| = |\mu_k - \tilde{g}(\boldsymbol{\xi})|O(\rho^{-2p})$ and $|\mu_0 - \tilde{g}(\boldsymbol{\xi})| = O(1)$, we have:

$$
|\mu_k - \tilde{g}(\boldsymbol{\xi})| = O(\rho^{-2kp}). \tag{6.13}
$$

Therefore, we will prove the lemma if we show that for all $k \geq 1$ the approximation μ_k enjoys the same asymptotic behaviour (6.2), at least up to an error $O(\rho^{-kp})$. This computation is straightforward. For example, we have

$$
\mu_1 = |\mathbf{F}\boldsymbol{\xi}|^2 + I_1(|\mathbf{F}\boldsymbol{\xi}|^2) + I_2(|\mathbf{F}\boldsymbol{\xi}|^2) + \tilde{I}_2(|\mathbf{F}\boldsymbol{\xi}|^2) + O(\rho^{-3p}),
$$

and, using (6.12), we obtain:

$$
\begin{aligned}
I_1(a(0)) &= -\sum_{\boldsymbol{\eta}\in\Theta_M'}|V(\boldsymbol{\eta})|^2\Big(\frac{|\mathbf{F}\boldsymbol{\eta}|^2}{(a(\boldsymbol{\eta})-a(0))(a(-\boldsymbol{\eta})-a(0))}\Big) \\
&= \sum_{\boldsymbol{\eta}\in\Theta_M'}|V(\boldsymbol{\eta})|^2\Big(\frac{|\mathbf{F}\boldsymbol{\eta}|^2}{(2\langle\mathbf{F}\boldsymbol{\xi},\mathbf{F}\boldsymbol{\eta}\rangle+|\mathbf{F}\boldsymbol{\eta}|^2)(2\langle\mathbf{F}\boldsymbol{\xi},\mathbf{F}\boldsymbol{\eta}\rangle-|\mathbf{F}\boldsymbol{\eta}|^2)}\Big) \\
&= \sum_{\boldsymbol{\eta}\in\Theta_M'}|V(\boldsymbol{\eta})|^2\Big(\frac{|\mathbf{F}\boldsymbol{\eta}|^2}{4\langle\mathbf{F}\boldsymbol{\xi},\mathbf{F}\boldsymbol{\eta}\rangle^2-|\mathbf{F}\boldsymbol{\eta}|^4}\Big) \\
&= \sum_{\boldsymbol{\eta}\in\Theta_M'}|V(\boldsymbol{\eta})|^2\Big(\frac{4^{-1}|\mathbf{F}\boldsymbol{\eta}|^2\langle\mathbf{F}\boldsymbol{\xi},\mathbf{F}\boldsymbol{\eta}\rangle^{-2}}{1-4^{-1}|\mathbf{F}\boldsymbol{\eta}|^4\langle\mathbf{F}\boldsymbol{\xi},\mathbf{F}\boldsymbol{\eta}\rangle^{-2}}\Big) \\
&= \sum_{\boldsymbol{\eta}\in\Theta_M'}|V(\boldsymbol{\eta})|^2\sum_{n=1}^{\infty}4^{-n}|\mathbf{F}\boldsymbol{\eta}|^{4n-2}\langle\mathbf{F}\boldsymbol{\xi},\mathbf{F}\boldsymbol{\eta}\rangle^{-2n} \\
&= \sum_{\boldsymbol{\eta}\in\Theta_M'}|V(\boldsymbol{\eta})|^2\sum_{n=1}^{\infty}4^{-n}|\mathbf{F}\boldsymbol{\eta}|^{4n-2}\langle\boldsymbol{\xi},\mathbf{G}\boldsymbol{\eta}\rangle^{-2n}.
\end{aligned}
\tag{6.14}
$$

Computations of $I_2(a(0))$ are similar (and, obviously, $\tilde{I}_2(a(0)) = 0$), only now the result will have terms which involve inner products of ξ with two different $\boldsymbol{\eta}$'s. Thus,

$$
\mu_1 = |\mathbf{F}\boldsymbol{\xi}|^2 + \sum_{\boldsymbol{\eta}_1,\boldsymbol{\eta}_2\in\Theta_M'}\sum_{n_1,n_2}A_{n_1,n_2}\langle\boldsymbol{\xi},\mathbf{G}\boldsymbol{\eta}_1\rangle^{-n_1}\langle\boldsymbol{\xi},\mathbf{G}\boldsymbol{\eta}_2\rangle^{-n_2}+O(\rho^{-3p}),
\tag{6.15}
$$

the sum being over all n_1, n_2 with $n_1 + n_2 \geq 2$ (in fact, we can take the sum over $n_1 + n_2 = 2$, since other terms will be $O(\rho^{-3p})$). Using induction, it is easy to prove now that

$$
\begin{aligned}
\mu_k = {}& |\mathbf{F}\boldsymbol{\xi}|^2 \\
& + \sum_{r=1}^{k+1}\sum_{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_r\in\Theta_M'}\sum_{n_1,\dots,n_r}A_{n_1,\dots,n_r}\langle\boldsymbol{\xi},\mathbf{G}\boldsymbol{\eta}_1\rangle^{-n_1}\dots\langle\boldsymbol{\xi},\mathbf{G}\boldsymbol{\eta}_r\rangle^{-n_r} \\
& + O(\rho^{-(k+2)p}),
\end{aligned}
\tag{6.16}
$$

the sum being over $2 \leq \sum_{j=1}^{r} n_j \leq k + 1$; $A_{n_1,\dots,n_p}$ is a polynomial of $\{\hat{V}(\boldsymbol{\eta}_j)\}$ and $\{\hat{V}(\boldsymbol{\eta}_j - \boldsymbol{\eta}_l)\}$. Indeed, if μ_k satisfies (6.16), then a calculation similar to (6.14) shows that for each $\boldsymbol{\eta} \in \Theta_M'$ the fraction $\frac{1}{a(\boldsymbol{\eta})-\mu_k}$ can be decomposed as a sum of products of negative powers

of $\langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_j \rangle$. Therefore, all functions $I_n(\mu_k)$ (and, thus, $I(\mu_k)$) admit similar decomposition. This implies that the next approximation $\mu_{k+1} = |\mathbf{F}\boldsymbol{\xi}|^2 + I(\mu_k)$ also satisfies (6.16).

Estimate (6.13) now shows that the asymptotic formula (6.3) holds.
□

We now define $g(\boldsymbol{\xi})$ as the finite part of the RHS of the expansion (6.2), namely

$$g(\boldsymbol{\xi}) = |\mathbf{F}\boldsymbol{\xi}|^2$$

$$+ \sum_{r=1}^{4M} \sum_{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_r \in \Theta'_M} \sum_{n_1+\cdots+n_r \geq 2} A_{n_1,\dots,n_r} \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_1 \rangle^{-n_1} \dots \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_r \rangle^{-n_r}.$$
(6.17)

Lemma 6.2. *We have:*

$$|g(\boldsymbol{\xi}) - \tilde{g}(\boldsymbol{\xi})| \ll \rho^{-4Mp}. \tag{6.18}$$

Proof. This follows from lemma 6.1.   □

7. COMPUTATION OF THE EIGENVALUES INSIDE RESONANCE LAYERS

Now let us fix $\mathfrak{V} \in \mathcal{V}(n)$, $1 \leq n \leq d-1$, and try to study the eigenvalues of $P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$. Let $\boldsymbol{\xi} = \mathbf{n} + \mathbf{k} \in \Xi_3(\mathfrak{V})$. We denote

$$\boldsymbol{\Upsilon}_j = \boldsymbol{\Upsilon}_j(\boldsymbol{\xi}) := \big((\boldsymbol{\xi} + (\mathfrak{V} \cap \mathbb{Z}^d)) \cap \Xi_j(\mathfrak{V}) \big), \qquad j = 0, 1, 2, 3, \quad (7.1)$$

$$\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}(\boldsymbol{\xi}) := \boldsymbol{\Upsilon}_3(\boldsymbol{\xi}) + \Theta_M, \tag{7.2}$$

$$P(\boldsymbol{\xi}) := \mathcal{P}^{(\mathbf{k})}\big(\boldsymbol{\Upsilon}(\boldsymbol{\xi})\big), \tag{7.3}$$

$$H'(\boldsymbol{\xi}) := P(\boldsymbol{\xi})H'(\mathbf{k})P(\boldsymbol{\xi}), \tag{7.4}$$

$$H_0(\boldsymbol{\xi}) := P(\boldsymbol{\xi})H_0(\mathbf{k})P(\boldsymbol{\xi}), \tag{7.5}$$

and

$$V'_{\boldsymbol{\xi}} := P(\boldsymbol{\xi})V'P(\boldsymbol{\xi}). \tag{7.6}$$

Out of all sets denoted by the letter $\boldsymbol{\Upsilon}$, we will mostly use $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi})$ and $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$; see Figures 6-7 for an illustration of these sets when $d = 2$. Let us establish some simple properties of these sets.

Lemma 7.1. *Suppose, $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}) \setminus \boldsymbol{\Upsilon}_3(\boldsymbol{\xi})$. Then $||\mathbf{F}\boldsymbol{\eta}|^2 - \lambda| \gg K^2$ (in particular, $\boldsymbol{\eta} \notin \mathcal{A}$).*

Proof. The assumptions of the lemma imply that $\boldsymbol{\eta} = \tilde{\boldsymbol{\xi}} + \boldsymbol{\theta}$ with $\tilde{\boldsymbol{\xi}} \in \boldsymbol{\Upsilon}_3(\boldsymbol{\xi})$ and $\boldsymbol{\theta} \in \Theta_M$. If $\boldsymbol{\theta} \notin \mathfrak{V}$, the statement follows from lemma 5.10. Assume $\boldsymbol{\theta} \in \mathfrak{V}$. Then $\boldsymbol{\eta} \notin \Xi_3(\mathfrak{V})$ (otherwise we had $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}_3(\boldsymbol{\xi})$). Now the statement follows from lemma 5.9.   □
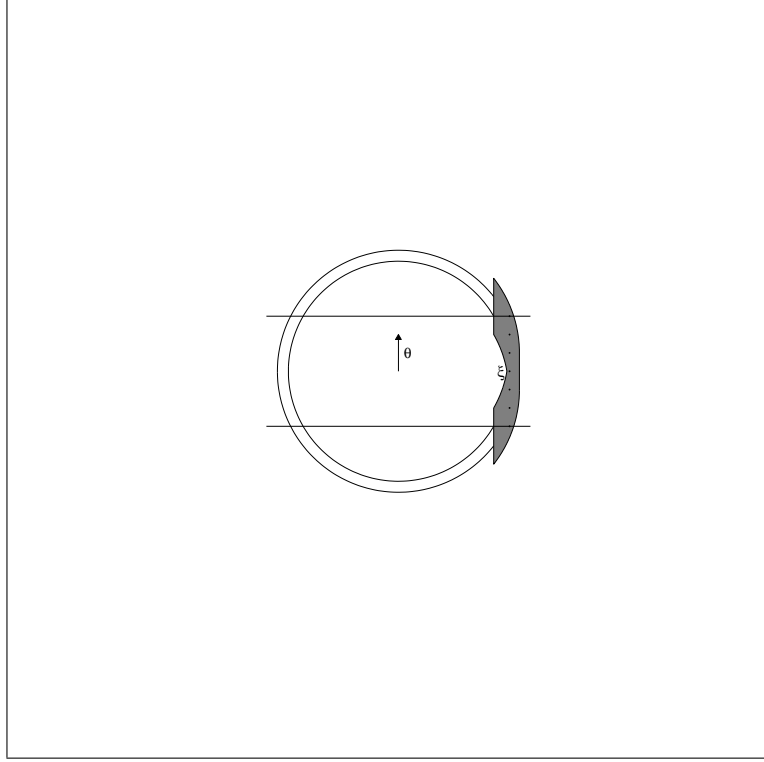
FIGURE 6. The sets $\Xi_3(\boldsymbol{\theta})$ and $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi})$

Lemma 7.2. *We have $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}) \subset \Xi_3(\mathfrak{V})$ and $\boldsymbol{\Upsilon}(\boldsymbol{\xi}) \subset \Xi(\mathfrak{V})$. If $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$, then $\boldsymbol{\eta} - \boldsymbol{\xi} \in \mathbb{Z}^d$. If for some $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi_3(\mathfrak{V})$ we have $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) \cap \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2) \neq \emptyset$, then $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2)$.*

Proof. The first three statements follow immediately from the definitions. Assume $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) \cap \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2) \neq \emptyset$, say $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) \cap \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2)$. Then $\boldsymbol{\eta} = \tilde{\boldsymbol{\xi}}_j + \boldsymbol{\theta}_j$, $j = 1, 2$, with $\tilde{\boldsymbol{\xi}}_j \in \Xi_3(\mathfrak{V})$ and $\boldsymbol{\theta}_j \in \Theta_M$. Then $\tilde{\boldsymbol{\xi}}_1 = \tilde{\boldsymbol{\xi}}_2 + (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)$. Since $\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1 \in \Theta_{2M}$, lemmas 5.10 and 5.6 imply that $\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1 \in \mathfrak{V}$. Therefore, $\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1 \in (\mathfrak{V} \cap \mathbb{Z}^d)$, so $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2)$. $\square$

Lemma 7.2 implies that the operator $P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$ splits into the direct sum:

$$P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V}) = \bigoplus H'(\boldsymbol{\xi}), \qquad (7.7)$$

the sum being over all classes of equivalence of $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$. Two vectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are equivalent if $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2)$.

Remark 7.3. The programme formulated at the end of section 5 requires to put into correspondence to each point $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$ a number $\tilde{g}(\boldsymbol{\xi})$ which is an eigenvalue of $P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$. It is natural to choose
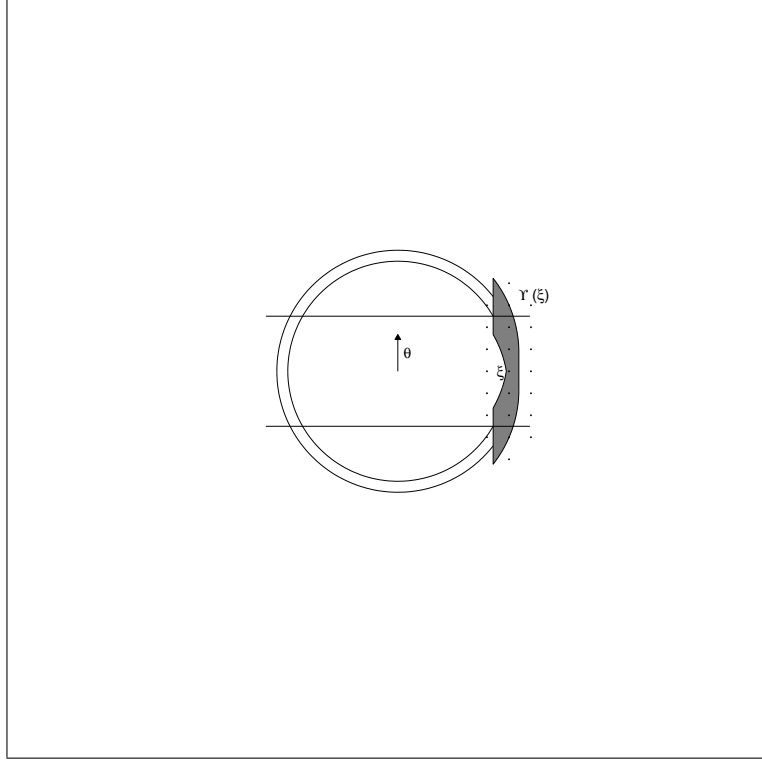
FIGURE 7. The sets $\Xi_3(\boldsymbol{\theta})$ and $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$; here, $\Theta = \{(0,0), (\pm 1, 0), (0, \pm 1)\}$ consists of five elements.

$\tilde{g}(\boldsymbol{\xi})$ to be an eigenvalue of $H'(\boldsymbol{\xi})$, say $\tilde{g}(\boldsymbol{\xi}) = \mu_j(H'(\boldsymbol{\xi}))$, where $j = j(\boldsymbol{\xi})$ is some natural number, and the mapping $j : \boldsymbol{\Upsilon}(\boldsymbol{\xi}) \to \mathbb{N}$ is (at least) an injection. There are certain technical problems with defining the function j. The first problem is that the sets $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1)$ and $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_2)$ can have different number of elements for different $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi_3(\mathfrak{V})$ (as Figure 8 illustrates), and the mapping j obviously has to take care of this fact. The second problem is that the mapping j cannot possibly be continuous (otherwise, since it takes only natural values, it would be constant and therefore not an injection), so $\tilde{g}$ as well cannot be continuous. Finally, we want $\tilde{g}(\boldsymbol{\xi})$ not to change too much when we change $\boldsymbol{\xi}$ a little. We cannot exactly achieve this (since, as we mentioned above, $\tilde{g}$ must be discontinuous), but we can achieve some weaker version of this (see lemma 7.11 for the precise statement).

Hence, we will study operators $H'(\mathbf{k})$ for each $\boldsymbol{\xi} = \mathbf{m} + \mathbf{k} \in \Xi(\mathfrak{V})$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$. Recall that we have denoted by $\boldsymbol{\xi}_{\mathfrak{V}}$ and $\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}$ vectors such
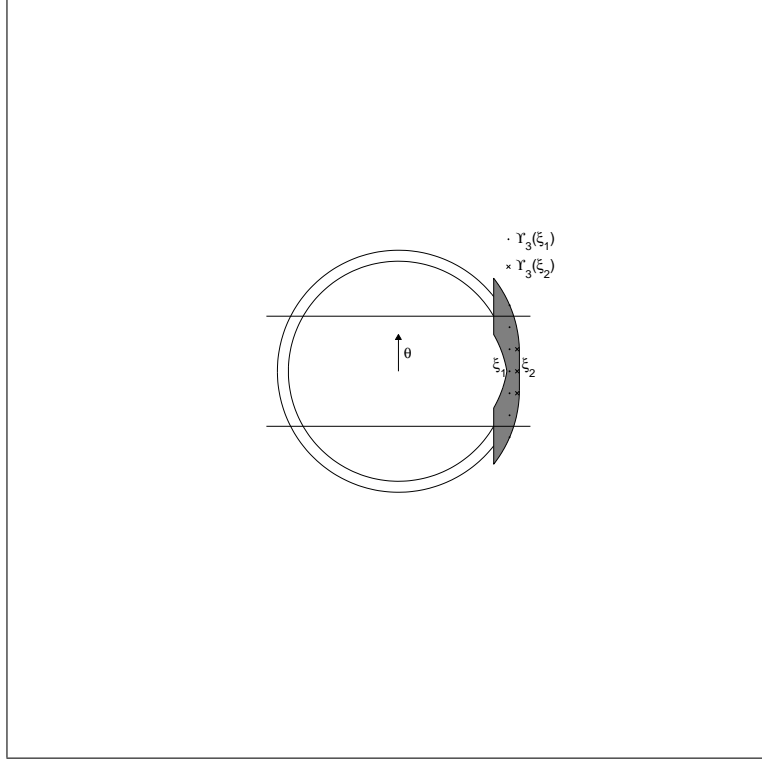
FIGURE 8. The sets $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1)$ (dots) and $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_2)$ (crosses) have different number of elements

that $\boldsymbol{\xi} = \boldsymbol{\xi}_{\mathfrak{V}} + \boldsymbol{\xi}_{\mathfrak{V}}^{\perp}$, $\boldsymbol{\xi}_{\mathfrak{V}} \in \mathfrak{V}$, $\mathbf{G}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp} \perp \mathfrak{V}$. Let us also define

$$r = r(\boldsymbol{\xi}) := |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|, \ \boldsymbol{\xi}_{\mathfrak{V}}' := \boldsymbol{\xi}_{\mathfrak{V}}^{\perp}/r, \qquad (7.8)$$

so that $|\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}'| = 1$. We can think of the triple $(r, \boldsymbol{\xi}_{\mathfrak{V}}', \boldsymbol{\xi}_{\mathfrak{V}})$ as the cylindrical coordinates on $\Xi(\mathfrak{V})$. Corollary 5.3 implies that $|\boldsymbol{\xi}_{\mathfrak{V}}| \ll \rho^{q_n}$; corollary 5.7 implies

$$|r - \rho| = O(\rho^{2q_n - 1}) = O(\rho^{-1/3}), \qquad (7.9)$$

since $q_n \leq 1/3$; in particular, we have $r > 0$. The current objective is to express the asymptotic behaviour of eigenvalues of $H'(\boldsymbol{\xi})$ inside J in terms of r. In order to do this, we want to compare the eigenvalues of $H'(\boldsymbol{\xi}_1)$ and $H'(\boldsymbol{\xi}_2)$ when $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi(\mathfrak{V})$ are two points which are close to each other. Since the operators $H'(\boldsymbol{\xi}_1)$ and $H'(\boldsymbol{\xi}_2)$ act in different Hilbert spaces $P(\boldsymbol{\xi}_j)\mathfrak{H}$, we first need to map these Hilbert spaces onto each other. A natural idea is to employ the mapping $F_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2} : P(\boldsymbol{\xi}_1)\mathfrak{H} \to P(\boldsymbol{\xi}_2)\mathfrak{H}$ defined in the following way:

$$F_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2}(e_{\boldsymbol{\eta}}) = e_{\boldsymbol{\eta} + \boldsymbol{\xi}_2 - \boldsymbol{\xi}_1}. \qquad (7.10)$$

This mapping is 'almost' an isometry, except for the fact that it is not well-defined, i.e. it could happen for example that $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1)$, but $(\boldsymbol{\eta} + \boldsymbol{\xi}_2 - \boldsymbol{\xi}_1) \notin \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2)$ (Figure 8 illustrates how this can happen). In order to avoid this, we will extend the sets $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$. We do this in the following way. First, for $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi_2(\mathfrak{V})$ we define

$$\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) := \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) \cup \big(\boldsymbol{\Upsilon}(\boldsymbol{\xi}_2) - \boldsymbol{\xi}_2 + \boldsymbol{\xi}_1\big) \qquad (7.11)$$

and, similarly,

$$\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) := \boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1) \cup \big(\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_2) - \boldsymbol{\xi}_2 + \boldsymbol{\xi}_1\big)$$

(the set $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_2; \boldsymbol{\xi}_1)$ is shown on Figure 9). We also define

$$P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) := \mathcal{P}^{(\mathbf{k}_1)}\big(\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)\big), \qquad (7.12)$$

where $\mathbf{k}_1 := \{\boldsymbol{\xi}_1\}$;

$$H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) := P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) H'(\boldsymbol{\xi}_1) P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2), \qquad (7.13)$$

and

$$H_0(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) := P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) H_0(\boldsymbol{\xi}_1) P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2). \qquad (7.14)$$

Suppose also that $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$ and let $U \subset \Xi_2(\mathfrak{V})$ be a set containing $\boldsymbol{\xi}$ of diameter $\ll \rho^{-1}$. Denote

$$\boldsymbol{\Upsilon}(\boldsymbol{\xi}; U) := \cup_{\boldsymbol{\eta} \in U} \boldsymbol{\Upsilon}(\boldsymbol{\xi}; \boldsymbol{\eta}), \qquad (7.15)$$

$$\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}; U) := \cup_{\boldsymbol{\eta} \in U} \boldsymbol{\Upsilon}_3(\boldsymbol{\xi}; \boldsymbol{\eta}),$$

$$P(\boldsymbol{\xi}; U) = \mathcal{P}^{(\mathbf{k})}\big(\boldsymbol{\Upsilon}(\boldsymbol{\xi}; U)\big),$$

$$H'(\boldsymbol{\xi}; U) := P(\boldsymbol{\xi}; U) H'(\boldsymbol{\xi}) P(\boldsymbol{\xi}; U), \qquad (7.16)$$

and

$$H_0(\boldsymbol{\xi}; U) := P(\boldsymbol{\xi}; U) H_0(\boldsymbol{\xi}) P(\boldsymbol{\xi}; U). \qquad (7.17)$$

Notice that $\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2\})$.

Now if we define the mapping $F_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2} : P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)\mathfrak{H} \to P(\boldsymbol{\xi}_2; \boldsymbol{\xi}_1)\mathfrak{H}$ by formula (7.10), this mapping will be a bijection and an isometry, since obviously

$$\boldsymbol{\Upsilon}(\boldsymbol{\xi}_2; \boldsymbol{\xi}_1) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) + \boldsymbol{\xi}_2 - \boldsymbol{\xi}_1.$$

Similarly, if U is any set containing $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, then the mapping $F_{\boldsymbol{\xi}_1; U} : P(\boldsymbol{\xi}_1; U)\mathfrak{H} \to P(\boldsymbol{\xi}_2; U)\mathfrak{H}$ defined by (7.10) will be a bijection and an isometry, since

$$\boldsymbol{\Upsilon}(\boldsymbol{\xi}_2; U) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; U) + \boldsymbol{\xi}_2 - \boldsymbol{\xi}_1.$$

Note also that if $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; U)$, then $\boldsymbol{\eta} - \boldsymbol{\xi} \in \mathbb{Z}^d$.

The problem, of course, is that in general the spectra of $H'(\boldsymbol{\xi}_1)$ and $H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$ (or $H'(\boldsymbol{\xi}_1; U)$) can be quite far from each other. However, we can give sufficient conditions which guarantee that the spectra of $H'(\boldsymbol{\xi}_1)$ and $H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$ (or rather the parts of the spectra lying inside J)
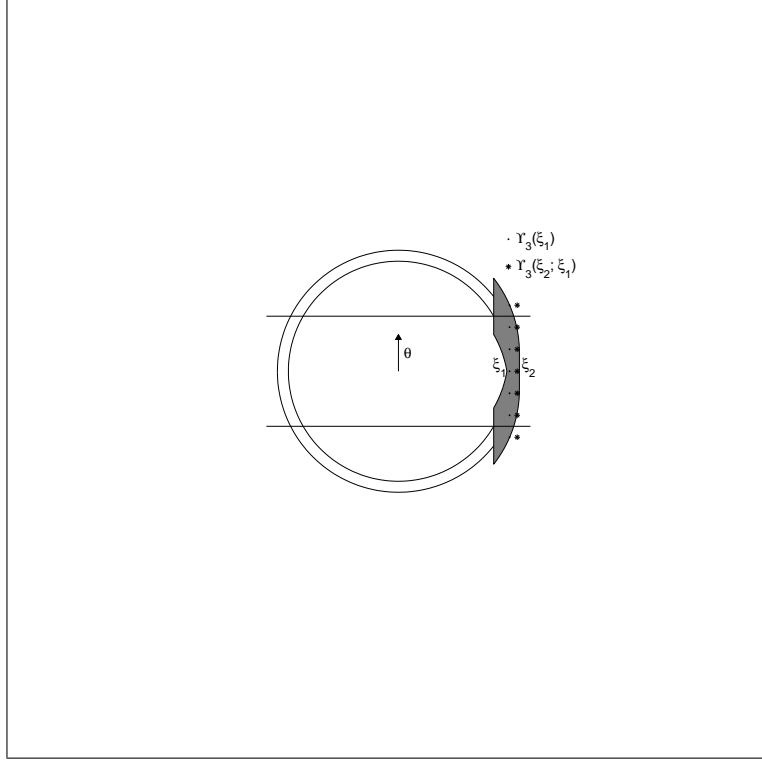
FIGURE 9. Now the sets $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) = \boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1)$ (dots) and $\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_2; \boldsymbol{\xi}_1)$ (stars) have the same number of elements

are within a small distance (of order $O(\rho^{-4Mp})$) from each other. The following statement is a straightforward corollary of lemma 3.2.

Lemma 7.4. *a) Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi_2(\mathfrak{V}) \subset \mathcal{A}$ satisfy $|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| < \rho^{-1}$. Then there exists a bijection $G = G_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2}$ defined on a subset of the set of all eigenvalues of $H'(\boldsymbol{\xi}_1)$ and mapping them to a subset of the set of all eigenvalues of $H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$ (eigenvalues in both sets are counted including multiplicities) satisfying the following properties:*

(i) all eigenvalues of $H'(\boldsymbol{\xi}_1)$ (resp. $H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$) inside J are in the domain (resp. range) of $G_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2}$;

(ii) for any eigenvalue $\mu_j(H'(\boldsymbol{\xi}_1)) \in J$ (and thus in the domain of $G_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2}$) we have:

$$|\mu_j(H'(\boldsymbol{\xi}_1)) - G(\mu_j(H'(\boldsymbol{\xi}_1)))| \ll \rho^{-4Mp}. \qquad (7.18)$$

b) Suppose $\boldsymbol{\xi} \in U \subset \Xi_2(\mathfrak{V})$ and the diameter of U is $\ll \rho^{-1}$. Then there exists a bijection $G = G_{\boldsymbol{\xi},U}$ defined on a subset of the set of all eigenvalues of $H'(\boldsymbol{\xi})$ and mapping them to a subset of the set of all

eigenvalues of $H'(\boldsymbol{\xi}; U)$ (eigenvalues in both sets are counted including multiplicities) satisfying the following properties:

(i) all eigenvalues of $H'(\boldsymbol{\xi})$ (resp. $H'(\boldsymbol{\xi}; U)$) inside J are in the domain (resp. range) of $G_{\boldsymbol{\xi}, U}$;

(ii) for any eigenvalue $\mu_j(H'(\boldsymbol{\xi})) \in J$ (and thus in the domain of $G_{\boldsymbol{\xi}, U}$) we have:

$$|\mu_j(H'(\boldsymbol{\xi})) - G(\mu_j(H'(\boldsymbol{\xi})))| \ll \rho^{-4Mp}. \tag{7.19}$$

Proof. Let us prove part a) of this lemma; part b) is proved analogously. Suppose, $\boldsymbol{\xi} \in (\boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1))$. Let us prove that then

$$|\,|\mathbf{F}\boldsymbol{\xi}|^2 - \rho^2| \gg K^2. \tag{7.20}$$

Indeed, we obviously have $\tilde{\boldsymbol{\xi}} := \boldsymbol{\xi} + \boldsymbol{\xi}_2 - \boldsymbol{\xi}_1 \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2) \subset \Xi(\mathfrak{V})$. Then definitions (5.10), (5.11), (7.1) and (7.2) imply that $\tilde{\boldsymbol{\xi}} = \tilde{\boldsymbol{\eta}} + \mathbf{a} + \boldsymbol{\theta}$, $\boldsymbol{\eta} \in (\Xi_2(\mathfrak{V}) \cap (\boldsymbol{\xi}_2 + \mathfrak{V}))$, $\mathbf{a} \in B(\mathfrak{V}, K)$, $\boldsymbol{\theta} \in \Theta_M$. If $\boldsymbol{\theta} \notin \mathfrak{V}$, (7.20) follows from lemma 5.10 and the inequality

$$|\,|\mathbf{F}\boldsymbol{\xi}|^2 - |\mathbf{F}\tilde{\boldsymbol{\xi}}|^2| \ll 1,$$

which in turn follows from the conditions of lemma. Suppose $\boldsymbol{\theta} \in \mathfrak{V}$. Then $\boldsymbol{\xi} - \boldsymbol{\xi}_1 = \mathbf{a} + \boldsymbol{\theta} + (\boldsymbol{\eta} - \boldsymbol{\xi}_2) \in \mathfrak{V}$ and, since $\boldsymbol{\xi} \notin \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1)$, we have $\boldsymbol{\xi} \notin \boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1)$, which in turn implies $\boldsymbol{\xi} \notin \Xi_3(\mathfrak{V})$. Now (7.20) follows from lemma 5.9.

Inequality (7.20) shows that as before we can apply lemma 3.2, or rather its corollary 3.3. This time, we apply this lemma with $H = H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$, $H_0 = H_0(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$, $n = 0$, $P^0 = P(\boldsymbol{\xi}_1)$, $Q = P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) - P(\boldsymbol{\xi}_1)$, $P_0^0 = \mathcal{P}^{(\mathbf{k}_1)}(\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1))$, $P_j^0(\mathfrak{V}) := P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)\mathcal{P}^{(\mathbf{k}_1)}\big((\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1) + \Theta_j) \setminus (\boldsymbol{\Upsilon}_3(\boldsymbol{\xi}_1) + \Theta_{j-1})\big)P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$, $j = 1, \ldots, M$. The fulfillment of all conditions of lemma 3.2 follows from (7.20) and lemma 5.14. Now the statement of lemma immediately follows from corollary 3.3. $\qquad\square$

Remark 7.5. Part (iii) of corollary 3.3 shows that the bijection $G_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2}$ is given by the following formula. Let $l = l(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ be the number of points

$$\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1) \tag{7.21}$$

with $|\mathbf{F}\boldsymbol{\eta}|^2 < \lambda$ (notice that if $\boldsymbol{\eta}$ satisfies (7.21), then $|\mathbf{F}\boldsymbol{\eta}| \notin J$). Then $G_{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2}(\mu_j(H'(\boldsymbol{\xi}_1))) = \mu_{j+l}(H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)))$. Similarly, if $l = l(\boldsymbol{\xi}, U)$ is the number of points

$$\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; U) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi}) \tag{7.22}$$

with $|\mathbf{F}\boldsymbol{\eta}|^2 < \lambda$, then $G_{\boldsymbol{\xi}, U}(\mu_j(H'(\boldsymbol{\xi}))) = \mu_{j+l}(H'(\boldsymbol{\xi}; U)))$.

The next lemma shows that the eigenvalues of $H'(\boldsymbol{\xi}; U)$ do not change much if we increase U; this lemma is an immediate corollary of lemma 7.4.

Lemma 7.6. *Let $\boldsymbol{\xi} \in U_1 \subset U_2 \subset \Xi_2(\mathfrak{V})$ and let the diameter of U_2 be $\ll \rho^{-1}$. Denote by $l = l(\boldsymbol{\xi}; U_1, U_2)$ the number of points*

$$\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; U_2) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi}; U_1) \tag{7.23}$$

with $|\mathbf{F}\boldsymbol{\eta}|^2 < \lambda$. Then:

a) for any eigenvalue $\mu_j(H'(\boldsymbol{\xi}; U_1)) \in J$ we have:

$$|\mu_j(H'(\boldsymbol{\xi}; U_1)) - \mu_{j+l}(H'(\boldsymbol{\xi}; U_2))| \ll \rho^{-4Mp}; \tag{7.24}$$

b) the number $l(\boldsymbol{\xi}; U_1, U_2)$ does not depend on $\boldsymbol{\xi}$, i.e. if $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in U_1$, then $l(\boldsymbol{\xi}_1; U_1, U_2) = l(\boldsymbol{\xi}_2; U_1, U_2)$.

Proof. Part a) of lemma follows from lemma 7.4 and remark 7.5, since $l(\boldsymbol{\xi}; U_1, U_2) = l_2 - l_1$, $l_j := l(\boldsymbol{\xi}; U_j)$, and we have

$$|\mu_j(H'(\boldsymbol{\xi})) - \mu_{j+l_j}(H'(\boldsymbol{\xi}; U_j))| \ll \rho^{-4Mp}, \qquad j = 1, 2. \tag{7.25}$$

Let us prove part b). Suppose, $\boldsymbol{\eta}_1 \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; U_2) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; U_1)$. Then, in the same way as we have proved (7.20), we can show that $||\mathbf{F}\boldsymbol{\eta}_1|^2 - \lambda| \gg K^2$. Denote $\boldsymbol{\eta}_2 := \boldsymbol{\eta}_1 + (\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1)$. The definitions of the sets $\boldsymbol{\Upsilon}$ imply that $\boldsymbol{\eta}_2 \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2; U_2) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi}_2; U_1)$. Since $|\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1| \ll \rho^{-1}$, we have $||\mathbf{F}\boldsymbol{\eta}_2|^2 - |\mathbf{F}\boldsymbol{\eta}_1|^2| \ll 1$. Therefore, the inequality $|\mathbf{F}\boldsymbol{\eta}_1|^2 < \lambda$ is satisfied if and only if the inequality $|\mathbf{F}\boldsymbol{\eta}_2|^2 < \lambda$ is satisfied. This proves that $l(\boldsymbol{\xi}_1; U_1, U_2) = l(\boldsymbol{\xi}_2; U_1, U_2)$. $\qquad\square$

As we have already mentioned, if $\boldsymbol{\xi}, \boldsymbol{\eta} \in U$, we have $\boldsymbol{\Upsilon}(\boldsymbol{\eta}; U) = \boldsymbol{\Upsilon}(\boldsymbol{\xi}; U) + (\boldsymbol{\eta} - \boldsymbol{\xi})$, which implies that the mapping $F_{\boldsymbol{\xi}, \boldsymbol{\eta}} : P(\boldsymbol{\xi}; U)\mathfrak{H} \to P(\boldsymbol{\eta}; U)\mathfrak{H}$ defined by (7.10) is an isometry. Thus, by considering the sets $\boldsymbol{\Upsilon}(\boldsymbol{\xi}; U)$ instead of $\boldsymbol{\Upsilon}(\boldsymbol{\xi})$ we have overcome the first difficulty mentioned in remark 7.3. Now we will try to face the other problems mentioned there.

Let $\boldsymbol{\eta}_0 = 0, \boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_p$ be the complete system of representatives of Θ_M modulo $\mathfrak{V}$ (we assume of course that $\boldsymbol{\eta}_j \in \Theta_M$). That means that each vector $\boldsymbol{\theta} \in \Theta_M$ has a unique representation $\boldsymbol{\theta} = \boldsymbol{\eta}_j + \mathbf{a}$, $\mathbf{a} \in \mathfrak{V}$. Denote $\boldsymbol{\Psi}_j = \boldsymbol{\Psi}_j(\boldsymbol{\xi}) := \big(\boldsymbol{\xi} + \boldsymbol{\eta}_j + (\mathfrak{V} \cap \mathbb{Z}^d)\big) \cap \boldsymbol{\Upsilon}(\boldsymbol{\xi})$. Then

$$\boldsymbol{\Upsilon}(\boldsymbol{\xi}) = \bigcup_j \boldsymbol{\Psi}_j, \tag{7.26}$$

and this is a disjoint union (on Figure 7, the set $\boldsymbol{\Psi}_0$ is the middle column of dots, and $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ are the left and right columns).

Let us compute diagonal elements of $H'(\boldsymbol{\xi})$. Let $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi})$. Then $\boldsymbol{\eta}$ can be uniquely decomposed as

$$\boldsymbol{\eta} = \boldsymbol{\xi} + \boldsymbol{\mu} + \boldsymbol{\eta}_j \tag{7.27}$$

with $\boldsymbol{\mu} \in \mathfrak{V} \cap \mathbb{Z}^d$. Recall that $H'(\boldsymbol{\xi}) = H_0(\boldsymbol{\xi}) + V'_{\boldsymbol{\xi}}$ and $H_0(\boldsymbol{\xi})e_{\boldsymbol{\eta}} = |\mathbf{F}\boldsymbol{\eta}|^2 e_{\boldsymbol{\eta}}$ whenever $\boldsymbol{\eta} \in \Upsilon(\boldsymbol{\xi})$. Since $\mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}} \perp \mathbf{F}\boldsymbol{\mu}$ and $\mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}} \perp \mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}$, we have:

$$
\begin{aligned}
|\mathbf{F}\boldsymbol{\eta}|^2 &= |\mathbf{F}(\boldsymbol{\xi} + \boldsymbol{\eta}_j + \boldsymbol{\mu})|^2 = |\mathbf{F}(r\boldsymbol{\xi}'_{\mathfrak{V}} + \boldsymbol{\xi}_{\mathfrak{V}} + \boldsymbol{\eta}_j + \boldsymbol{\mu})|^2 \\
&= r^2 + 2\langle \mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}}, \mathbf{F}\boldsymbol{\eta}_j \rangle r + |\mathbf{F}((\boldsymbol{\xi} + \boldsymbol{\eta}_j)_{\mathfrak{V}} + \boldsymbol{\mu})|^2.
\end{aligned}
\tag{7.28}
$$

This simple computation implies that

$$
H'(\boldsymbol{\xi}) = r^2 I + rA + B.
\tag{7.29}
$$

Here, $A = A(\boldsymbol{\xi}) = A(\boldsymbol{\xi}_{\mathfrak{V}}, \boldsymbol{\xi}'_{\mathfrak{V}}, r)$ and $B = B(\boldsymbol{\xi}) = B(\boldsymbol{\xi}_{\mathfrak{V}}, \boldsymbol{\xi}'_{\mathfrak{V}}, r)$ are self-adjoint operators acting in $P(\boldsymbol{\xi})\mathfrak{H}$ in the following way:

$$
A = 2\sum_{j=0}^{p} \langle \mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}}, \mathbf{F}\boldsymbol{\eta}_j \rangle \mathcal{P}^{(\mathbf{k})}(\boldsymbol{\Psi}_j);
$$

in other words,

$$
Ae_{\boldsymbol{\eta}} = 2\langle \mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}}, \mathbf{F}\boldsymbol{\eta}_j \rangle e_{\boldsymbol{\eta}} = 2\langle \mathbf{F}\boldsymbol{\xi}'_{\mathfrak{V}}, \mathbf{F}(\boldsymbol{\eta} - \boldsymbol{\xi}) \rangle e_{\boldsymbol{\eta}},
\tag{7.30}
$$

and

$$
Be_{\boldsymbol{\eta}} = |\mathbf{F}((\boldsymbol{\xi} + \boldsymbol{\eta}_j)_{\mathfrak{V}} + \boldsymbol{\mu})|^2 e_{\boldsymbol{\eta}} + V'_{\boldsymbol{\xi}} e_{\boldsymbol{\eta}} = (|\mathbf{F}\boldsymbol{\eta}_{\mathfrak{V}}|^2 + V'_{\boldsymbol{\xi}}) e_{\boldsymbol{\eta}}
\tag{7.31}
$$

for all $\boldsymbol{\eta} \in \boldsymbol{\Psi}_j(\boldsymbol{\xi})$ with $\boldsymbol{\eta}_j$ and $\boldsymbol{\mu}$ being defined by (7.27). These definitions imply that $\ker A = \mathcal{P}^{(\mathbf{k})}(\boldsymbol{\Psi}_0)\mathfrak{H}$. Notice that

$$
\|A(\boldsymbol{\xi})\| \ll R < \rho^{1/3}
\tag{7.32}
$$

and

$$
\|B(\boldsymbol{\xi})\| \ll L_n^2 < \rho^{2/3}
\tag{7.33}
$$

due to our assumptions made before lemma 5.1; see also corollary 5.3.

The dependence of the operator pencil $H' = r^2 I + rA + B$ on r is two-fold: together with the obvious quadratic dependence, the coefficients A and B depend on r as well. However, as we will show in lemma 7.7, the second type of dependence is rather weak. Put

$$
D(\boldsymbol{\xi}) := r(\boldsymbol{\xi})A(\boldsymbol{\xi}) + B(\boldsymbol{\xi}).
$$

By $\{\nu_j(\boldsymbol{\xi})\}$ we denote the eigenvalues of $D(\boldsymbol{\xi})$. Then according to (7.29) the eigenvalues of $H'(\boldsymbol{\xi})$ are equal to

$$
\lambda_j(\boldsymbol{\xi}) = r^2(\boldsymbol{\xi}) + \nu_j(\boldsymbol{\xi}).
\tag{7.34}
$$

If $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi_3(\mathfrak{V})$, then we can define the operator $A(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$ as the operator defined by (7.30) with the domain $P(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)\mathfrak{H}$. Similarly, if U is a set containing $\boldsymbol{\xi}$ of diameter $\ll \rho^{-1}$, then we define the operator $A(\boldsymbol{\xi}; U)$ as the operator defined by (7.30) with the domain $P(\boldsymbol{\xi}; U)\mathfrak{H}$. In the same way, we can define $B(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$, $B(\boldsymbol{\xi}; U)$ (they are defined by

means of (7.31)), $D(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) = r(\boldsymbol{\xi}_1)A(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) + B(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$, and $D(\boldsymbol{\xi}; U)$. We also denote by $\nu_j(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$ the eigenvalues of $D(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$ and by

$$\lambda_j(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2) = r^2(\boldsymbol{\xi}_1) + \nu_j(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$$

the eigenvalues of $H'(\boldsymbol{\xi}_1; \boldsymbol{\xi}_2)$; $\nu_j(\boldsymbol{\xi}; U)$ and $\lambda_j(\boldsymbol{\xi}; U)$ are defined analogously.

Let us now study how the eigenvalues change under the change of r.

Lemma 7.7. *Let $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$, $r = r(\boldsymbol{\xi})$. Let U be a set of diameter $\ll \rho^{-1}$ containing $\boldsymbol{\xi}$. Let t be a real number with $|t - r| \ll \rho^{-1}$ and $\mathbf{a} = \mathbf{a}(t) \in \Xi_2(\mathfrak{V})$ be a unique point satisfying $(\mathbf{a})_{\mathfrak{V}} = (\boldsymbol{\xi})_{\mathfrak{V}}$, $(\mathbf{a})'_{\mathfrak{V}} = (\boldsymbol{\xi})'_{\mathfrak{V}}$, and $r(\mathbf{a}) = t$ (thus, when we vary t, the path $\mathbf{a}(t)$ is a straight interval which goes along the $\mathbf{F}$-perpendicular dropped from the point $\boldsymbol{\xi}$ onto $\mathfrak{V}$). Suppose, $\mathbf{a} \in U$. Let $\nu_j(t)$ (resp. $\lambda_j(t)$) denote the eigenvalues of $D(\mathbf{a}(t); U)$ (resp. $H'(\mathbf{a}(t); U)$). Then*

$$\frac{d\nu_j(t)}{dt} = O(\rho^{1/3}) \tag{7.35}$$

and

$$\frac{d\lambda_j(t)}{dt} = 2t + O(\rho^{1/3}) \tag{7.36}$$

Proof. Let t_1, t_2 be real numbers satisfying $|t_j - r| \ll \rho^{-1}$ and $\mathbf{a}_1 = \mathbf{a}(t_1)$, $\mathbf{a}_2 = \mathbf{a}(t_2)$ be the corresponding points inside $\Xi_3(\mathfrak{V}) \cap U$. First of all, we notice that the mapping $F_{\mathbf{a}_1, \mathbf{a}_2}$ defined by (7.10) is an isometry which maps $P(\mathbf{a}_1; U)\mathfrak{H}$ onto $P(\mathbf{a}_2; U)\mathfrak{H}$. Moreover, the definitions of the operators A and B imply that

$$A(\mathbf{a}_1; U) = F_{\mathbf{a}_2, \mathbf{a}_1} A(\mathbf{a}_2; U) F_{\mathbf{a}_1, \mathbf{a}_2};$$

similarly,

$$B(\mathbf{a}_1; U) = F_{\mathbf{a}_2, \mathbf{a}_1} B(\mathbf{a}_2; U) F_{\mathbf{a}_1, \mathbf{a}_2}.$$

These unitary equivalencies show that the eigenvalues $\nu_j(t)$ are in fact the eigenvalues of the linear operator pencil $tA + B$, with A and B being any of the operators $A(\mathbf{a}; U)$ and $B(\mathbf{a}; U)$ with $\mathbf{a}$ satisfying $(\mathbf{a})_{\mathfrak{V}} = (\boldsymbol{\xi})_{\mathfrak{V}}$ and $(\mathbf{a})'_{\mathfrak{V}} = (\boldsymbol{\xi})'_{\mathfrak{V}}$; it does not matter which particular point $\mathbf{a}$ we have chosen, since all corresponding operators are unitarily equivalent. For example, we can choose $A = A(\boldsymbol{\xi}; U)$ and $B = B(\boldsymbol{\xi}; U)$. Now an elementary perturbation theory shows that

$$\frac{d\nu_j}{dt} = \langle Au_j, u_j \rangle, \tag{7.37}$$

where u_j is the eigenvector of D corresponding to the eigenvalue ν_j. The estimate (7.32) shows that $d\nu_j/dt = O(\rho^{1/3})$. This proves (7.35). The estimate (7.36) follows from this and the identity $\lambda_j(t) = t^2 + \nu_j(t)$. $\quad\square$

Using similar perturbative argument, we can study how the eigenvalues change when we change the other variables, namely, $\boldsymbol{\xi}_{\mathfrak{V}}$ and $\boldsymbol{\xi}'_{\mathfrak{V}}$.

Lemma 7.8. *Let $\boldsymbol{\xi} \in \Xi_3(\mathfrak{V})$, and let $\mathbf{a} \in \Xi_3(\mathfrak{V})$ be the point satisfying $r(\mathbf{a}) = r(\boldsymbol{\xi})$, $|\mathbf{a} - \boldsymbol{\xi}| \ll \rho^{-1}$. Suppose, $\boldsymbol{\xi}, \mathbf{a} \in U$. Then*

$$|\lambda_j(\mathbf{a}; U) - \lambda_j(\boldsymbol{\xi}; U)| = |\nu_j(\mathbf{a}; U) - \nu_j(\boldsymbol{\xi}; U)| \ll |\mathbf{a} - \boldsymbol{\xi}| \rho^{1/3}. \quad (7.38)$$

Proof. Formula (7.34) and the condition $r(\mathbf{a}) = r(\boldsymbol{\xi})$ imply that $\lambda_j(\mathbf{a}; U) - \lambda_j(\boldsymbol{\xi}; U) = \nu_j(\mathbf{a}; U) - \nu_j(\boldsymbol{\xi}; U)$. Moreover, definitions (7.30) and (7.31) imply that

$$||A(\boldsymbol{\xi}; U) - F_{\mathbf{a}, \boldsymbol{\xi}} A(\mathbf{a}; U) F_{\boldsymbol{\xi}, \mathbf{a}}|| \ll |\boldsymbol{\xi}'_{\mathfrak{V}} - \mathbf{a}'_{\mathfrak{V}}| R \ll |\mathbf{a} - \boldsymbol{\xi}| \rho^{-2/3} \quad (7.39)$$

and

$$||B(\boldsymbol{\xi}; U) - F_{\boldsymbol{\xi}, \mathbf{a}} B(\mathbf{a}; U) F_{\mathbf{a}, \boldsymbol{\xi}}|| \ll |\boldsymbol{\xi}_{\mathfrak{V}} - \mathbf{a}_{\mathfrak{V}}|(|\boldsymbol{\xi}_{\mathfrak{V}}| + L_n) \ll |\mathbf{a} - \boldsymbol{\xi}| \rho^{1/3}. \quad (7.40)$$

Indeed, let us check for example (7.39). Suppose, $\boldsymbol{\eta} \in \Upsilon$ (say, $\boldsymbol{\eta} \in \Psi_j$). Then we have:

$$A(\boldsymbol{\xi}; U) \epsilon_{\boldsymbol{\eta}} = 2 \langle \mathbf{F} \boldsymbol{\xi}'_{\mathfrak{V}}, \mathbf{F} \boldsymbol{\eta}_j \rangle e_{\boldsymbol{\eta}}$$

and

$$F_{\mathbf{a}, \boldsymbol{\xi}} A F_{\boldsymbol{\xi}, \mathbf{a}} \epsilon_{\boldsymbol{\eta}} = 2 \langle \mathbf{F} \mathbf{a}'_{\mathfrak{V}}, \mathbf{F} \boldsymbol{\eta}_j \rangle e_{\boldsymbol{\eta}}.$$

Since

$$|\mathbf{a}'_{\mathfrak{V}} - \boldsymbol{\xi}'_{\mathfrak{V}}| \ll |\mathbf{a} - \boldsymbol{\xi}| \rho^{-1}$$

and

$$|\boldsymbol{\eta}_j| \ll R,$$

we have (7.39). The estimate (7.40) can be proved analogously.

Therefore, since $r(\boldsymbol{\xi}) \sim \rho$, we have

$$||D(\boldsymbol{\xi}; U) - F_{\boldsymbol{\xi}, \mathbf{a}} D(\mathbf{a}; U) F_{\mathbf{a}, \boldsymbol{\xi}}|| \ll |\mathbf{a} - \boldsymbol{\xi}| \rho^{1/3}.$$

Since the spectra of $D(\mathbf{a}; U)$ and $F_{\boldsymbol{\xi}, \mathbf{a}} D(\mathbf{a}; U) F_{\mathbf{a}, \boldsymbol{\xi}}$ coinside, this implies

$$|\nu_j(\boldsymbol{\xi}; U) - \nu_j(\mathbf{a}; U)| \ll |\mathbf{a} - \boldsymbol{\xi}| \rho^{1/3},$$

which finishes the proof. $\square$

Let us summarize the information about the spectra of $H'(\boldsymbol{\xi})$ we have obtained so far. Recall that $\mathcal{A}_1$ is a slightly 'slimmed down' version of $\mathcal{A}$; it consists of all points $\boldsymbol{\xi}$ with $|\mathbf{F} \boldsymbol{\xi}|^2 \in J$.

Lemma 7.9. *Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in U \subset \Xi_2(\mathfrak{V}) \cap \mathcal{A}_1$ with the diameter of U being ρ^{-1}. Assume that $\mu_j(H'(\boldsymbol{\xi}_1; U)) \in J$. Then*

$$|\mu_j(H'(\boldsymbol{\xi}_1; U)) - \mu_j(H'(\boldsymbol{\xi}_2; U))| \ll \rho |\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| + \rho^{-4Mp}.$$

If we assume, moreover, that $(\boldsymbol{\xi}_1)_{\mathfrak{V}} = (\boldsymbol{\xi}_2)_{\mathfrak{V}}$ and $(\boldsymbol{\xi}_1)'_{\mathfrak{V}} = (\boldsymbol{\xi}_2)'_{\mathfrak{V}}$, then

$$\mu_j(H'(\boldsymbol{\xi}_1; U)) - \mu_j(H'(\boldsymbol{\xi}_2; U)) = (2\rho + O(\rho^{1/3}))(r(\boldsymbol{\xi}_1) - r(\boldsymbol{\xi}_2)) + O(\rho^{-4Mp}).$$

Finally, if $(\boldsymbol{\xi}_1)_{\mathfrak{V}} = (\boldsymbol{\xi}_2)_{\mathfrak{V}}$, $(\boldsymbol{\xi}_1)'_{\mathfrak{V}} = (\boldsymbol{\xi}_2)'_{\mathfrak{V}}$, and U contains the interval I joining $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, then

$$\mu_j(H'(\boldsymbol{\xi}_1; U)) - \mu_j(H'(\boldsymbol{\xi}_2; U)) = (2\rho + O(\rho^{1/3}))(r(\boldsymbol{\xi}_1) - r(\boldsymbol{\xi}_2)).$$

Proof. The last statement follows directly from lemma 7.7. Assume now that $(\boldsymbol{\xi}_1)_{\mathfrak{V}} = (\boldsymbol{\xi}_2)_{\mathfrak{V}}$ and $(\boldsymbol{\xi}_1)'_{\mathfrak{V}} = (\boldsymbol{\xi}_2)'_{\mathfrak{V}}$. Denote $U_1 := U$, $U_2 := U \cup I$, where I is the interval joining $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, and $l = l(\boldsymbol{\xi}_1; U_1, U_2) = l(\boldsymbol{\xi}_2; U_1, U_2)$ (the last equality follows from lemma 7.6). Then lemma 7.6 implies that

$$|\mu_j(H'(\boldsymbol{\xi}_m; U_1)) - \mu_{j+l}(H'(\boldsymbol{\xi}_m; U_2))| \ll \rho^{-4Mp}, \qquad m = 1, 2.$$

Now the statement follows from lemma 7.7. If $r(\boldsymbol{\xi}_1) = r(\boldsymbol{\xi}_2)$, the statement follows in a similar way from lemmas 7.6 and 7.8. In the general case, we join $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ by a path consisting of intervals falling into either of the two cases above. $\qquad\square$

Now we will 'globalize' the local mappings constructed so far, in other words, we will define the function $j : \Upsilon(\boldsymbol{\xi}) \to \mathbb{N}$ mentioned in the remark 7.3. Let $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ and $\{\boldsymbol{\xi}\} = \mathbf{k}$. Then the set of eigenvalues $\{\mu_j(H'_0(\boldsymbol{\xi}))\}$ of the unperturbed operator $H'_0(\boldsymbol{\xi})$ coincides with the set $\{|\mathbf{F}\boldsymbol{\eta}|^2, \boldsymbol{\eta} \in \Upsilon(\boldsymbol{\xi})\}$. Let us label all numbers $\{|\mathbf{F}\boldsymbol{\eta}|^2, \boldsymbol{\eta} \in \Upsilon(\boldsymbol{\xi})\}$ in the increasing order; if there are two different vectors $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}} \in \Upsilon(\boldsymbol{\xi})$ with $|\mathbf{F}\boldsymbol{\eta}|^2 = |\mathbf{F}\tilde{\boldsymbol{\eta}}|^2$, we label them in the lexicographic order of their coordinates (i.e. we put $\boldsymbol{\eta}$ before $\tilde{\boldsymbol{\eta}}$ if either $\eta_1 < \tilde{\eta}_1$, or $\eta_1 = \tilde{\eta}_1$ and $\eta_2 < \tilde{\eta}_2$, etc.) Then to each point $\boldsymbol{\eta} \in \Upsilon(\boldsymbol{\xi})$ we have put into correspondence a natural number $j = j(\boldsymbol{\eta})$ such that

$$|\mathbf{F}\boldsymbol{\eta}|^2 = \mu_j(H'_0(\boldsymbol{\xi})). \tag{7.41}$$

Next we define

$$\tilde{g}(\boldsymbol{\eta}) := \mu_{j(\boldsymbol{\eta})}(H'(\boldsymbol{\xi})).$$

This mapping is well-defined and satisfies the following obvious property: $|\tilde{g}(\boldsymbol{\eta}) - |\mathbf{F}\boldsymbol{\eta}|^2| \le v$ (recall that $v = ||V||_\infty$).

The problem with the mapping $\tilde{g}$ defined in this way is that we cannot apply lemma 7.7 to it, since lemma 7.7 treats not the eigenvalues of $H'(\boldsymbol{\xi})$, but the eigenvalues of $H'(\boldsymbol{\xi}; U)$ with the set U containing certain intervals perpendicular to $\mathfrak{V}$. Thus, we need to introduce a different definition which takes care of lemma 7.7 and at the same time is reasonably canonical.

Let $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$. Denote

$$X = X(\boldsymbol{\xi}) := \{\boldsymbol{\eta} \in \Xi_2(\mathfrak{V}) : \boldsymbol{\eta}_{\mathfrak{V}} = \boldsymbol{\xi}_{\mathfrak{V}}, \, \boldsymbol{\eta}'_{\mathfrak{V}} = \boldsymbol{\xi}'_{\mathfrak{V}}\}.$$

Simple geometry implies that $X(\boldsymbol{\xi})$ is an interval of length $\ll \rho^{-1}$. Similarly to our actions when we were defining $\tilde{g}$, we notice that the set of eigenvalues $\{\mu_j(H_0'(\boldsymbol{\xi}; X(\boldsymbol{\xi})))\}$ coincides with the set $\{|\mathbf{F}\boldsymbol{\eta}|^2, \boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; X)\}$. Let us label all numbers $\{|\mathbf{F}\boldsymbol{\eta}|^2, \boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; X)\}$ in the increasing order; if there are two different vectors $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; X)$ with $|\mathbf{F}\boldsymbol{\eta}_1|^2 = |\mathbf{F}\boldsymbol{\eta}_2|^2$, we label them in the lexicographic order of their coordinates. Then to the point $\boldsymbol{\xi}$ we have put into correspondence a natural number $i = i(\boldsymbol{\xi})$ such that

$$|\mathbf{F}\boldsymbol{\xi}|^2 = \mu_i(H_0'(\boldsymbol{\xi}; X)). \tag{7.42}$$

Next we define $g(\boldsymbol{\xi}) := \mu_{i(\boldsymbol{\xi})}(H'(\boldsymbol{\xi}; X))$. This mapping is well-defined and satisfies the property $|g(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \le v$.

Lemma 7.10. *Let $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V}) \cap \mathcal{A}_1$. Then the following properties are satisfied:*

(i) $|g(\boldsymbol{\xi}) - \tilde{g}(\boldsymbol{\xi})| \ll \rho^{-4Mp}$;

(ii) $g(\boldsymbol{\xi}) = r^2 + s$, where $r = r(\boldsymbol{\xi})$ and $s = s(\boldsymbol{\xi}) = s(\boldsymbol{\xi}_{\mathfrak{V}}, \boldsymbol{\xi}'_{\mathfrak{V}}, r)$ is a function which smoothly depends on r with $\frac{\partial s}{\partial r} = O(\rho^{1/3})$.

Proof. Let us prove the first statement. First, we notice that the difference $i(\boldsymbol{\xi}) - j(\boldsymbol{\xi})$ is equal to the number of points $\boldsymbol{\eta} \in (\boldsymbol{\Upsilon}(\boldsymbol{\xi}; X(\boldsymbol{\xi})) \setminus \boldsymbol{\Upsilon}(\boldsymbol{\xi})$ satisfying $|\mathbf{F}\boldsymbol{\eta}|^2 < \lambda$. Now the statement follows from lemma 7.4 and remark 7.5.

Let us now prove the second statement. Suppose, $\boldsymbol{\xi}_1 \in X(\boldsymbol{\xi})$ and $\boldsymbol{\eta} \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}; X)$. Then

$$\boldsymbol{\eta}_1 := \boldsymbol{\eta} + (\boldsymbol{\xi}_1 - \boldsymbol{\xi}) \in \boldsymbol{\Upsilon}(\boldsymbol{\xi}_1; X).$$

Note that $(\boldsymbol{\xi})_{\mathfrak{V}} = (\boldsymbol{\xi}_1)_{\mathfrak{V}}$ and therefore $(\boldsymbol{\eta})_{\mathfrak{V}} = (\boldsymbol{\eta}_1)_{\mathfrak{V}}$. Let us assume that

$$|\mathbf{F}\boldsymbol{\xi}| \ge |\mathbf{F}\boldsymbol{\eta}| \tag{7.43}$$

and prove that this implies $|\mathbf{F}\boldsymbol{\xi}_1| \ge |\mathbf{F}\boldsymbol{\eta}_1|$. Indeed, there are two possible cases:

(i) $\boldsymbol{\xi} - \boldsymbol{\eta} \in \mathfrak{V}$. Then, since $\boldsymbol{\xi}_1 - \boldsymbol{\eta}_1 = \boldsymbol{\xi} - \boldsymbol{\eta} \in \mathfrak{V}$, we have:

$$|\mathbf{F}\boldsymbol{\xi}_1|^2 - |\mathbf{F}\boldsymbol{\eta}_1|^2 = |\mathbf{F}(\boldsymbol{\xi}_1)_{\mathfrak{V}}|^2 - |\mathbf{F}(\boldsymbol{\eta}_1)_{\mathfrak{V}}|^2$$
$$= |\mathbf{F}(\boldsymbol{\xi})_{\mathfrak{V}}|^2 - |\mathbf{F}(\boldsymbol{\eta})_{\mathfrak{V}}|^2 = |\mathbf{F}\boldsymbol{\xi}|^2 - |\mathbf{F}\boldsymbol{\eta}|^2 \ge 0$$

(ii) $\boldsymbol{\xi} - \boldsymbol{\eta} \notin \mathfrak{V}$. Then, in the same way we have proved estimate (7.20), using lemma 5.10 we can show that $||\mathbf{F}\boldsymbol{\eta}|^2 - \lambda| \gg K^2$. But since $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V}) \subset \mathcal{A}$, we have $||\mathbf{F}\boldsymbol{\xi}|^2 - \lambda| \ll 1$. Thus, (7.43) implies $\lambda - |\mathbf{F}\boldsymbol{\eta}|^2 \gg K^2$. Since $|\boldsymbol{\eta}_1 - \boldsymbol{\eta}| = |\boldsymbol{\xi}_1 - \boldsymbol{\xi}| \ll \rho^{-1}$, we have $\lambda - |\mathbf{F}\boldsymbol{\eta}_1|^2 \gg K^2$. Since $\boldsymbol{\xi}_1 \in \mathcal{A}$, this implies $|\mathbf{F}\boldsymbol{\xi}_1| \ge |\mathbf{F}\boldsymbol{\eta}_1|$.

Thus, we have proved that the inequality $|\mathbf{F}\boldsymbol{\xi}_1| \ge |\mathbf{F}\boldsymbol{\eta}_1|$ is equivalent to $|\mathbf{F}\boldsymbol{\xi}| \ge |\mathbf{F}\boldsymbol{\eta}|$. This implies that $i(\boldsymbol{\xi}_1) = i(\boldsymbol{\xi})$, where i is the function

defined by (7.42). Now the second statement of lemma follows from lemma 7.7. □

This lemma shows that the mapping g behaves in a nice way as a function of r. Unfortunately, the dependence on other variables is not quite so nice. In fact, this mapping is not continuous, even modulo $O(\rho^{-4Mp})$, because the functions $i(\boldsymbol{\xi})$ are not continuous; moreover, a little thought shows that we cannot, in general, define the mapping g to have all properties formulated in the introduction and be continuous at the same time. Indeed, if the function $i = i(\boldsymbol{\xi})$ were continuous, it would necessary have been a constant. Thus, the function i has discontinuities, and the function g may have discontinuities at the same points as i. However, lemmas 7.4, 7.7, and 7.8 show that for each small neighbourhood U in the space of quasi-momenta we can find a family of representatives of the functions g which is 'almost' smooth. Namely, the following statement holds:

Lemma 7.11. *Let $I = [\mathbf{a}, \mathbf{b}] \subset \Xi_2(\mathfrak{V}) \cap \mathcal{A}_1$ be a straight interval of length $L := |\mathbf{b} - \mathbf{a}| \ll \rho^{-1}$. Then there exists an integer vector $\mathbf{n}$ such that $|g(\mathbf{b} + \mathbf{n}) - g(\mathbf{a})| \ll L\rho + \rho^{-4Mp}$. Moreover, suppose in addition that there exists an integer vector $\mathbf{m} \neq 0$ such that the interval $I + \mathbf{m}$ is entirely inside $\Xi_2(\mathfrak{V}) \cap \mathcal{A}_1$. Then there exist two different integer vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ such that $|g(\mathbf{b} + \mathbf{n}_1) - g(\mathbf{a})| \ll L\rho + \rho^{-4Mp}$ and $|g(\mathbf{b} + \mathbf{n}_2) - g(\mathbf{a} + \mathbf{m})| \ll L\rho + \rho^{-4Mp}$.*

Proof. Lemmas 7.4 and 7.10 show that $g(\mathbf{a}) = \lambda_l(\mathbf{a}; I) + O(\rho^{-4Mp})$ for some integer l. Lemma 7.9 now implies that

$$|\lambda_l(\mathbf{a}; I) - \lambda_l(\mathbf{b}; I)| \ll L\rho + \rho^{-4Mp}. \tag{7.44}$$

Once again using lemma 7.4, we deduce that $\lambda_l(\mathbf{b}; I) = g(\boldsymbol{\eta}) + O(\rho^{-4Mp})$ for some $\boldsymbol{\eta} \in \Upsilon(\mathbf{b}; I)$; in particular, we have $\boldsymbol{\eta} = \mathbf{b} + \mathbf{n}$ for some integer vector $\mathbf{n}$. This proves the first statement.

Let us prove the second statement. Conditions of lemma imply that $g(\mathbf{a}) = \lambda_j(\mathbf{a}; I) + O(\rho^{-4Mp})$ and $g(\mathbf{a} + \mathbf{m}) = \lambda_l(\mathbf{a} + \mathbf{m}; I + \mathbf{m}) + O(\rho^{-4Mp})$ for some integers j, l. Moreover, if $\mathbf{a} + \mathbf{m} \in \Upsilon(\mathbf{a}; I)$ (so that $\Upsilon(\mathbf{a}; I) = \Upsilon(\mathbf{a} + \mathbf{m}; I + \mathbf{m})$), then, since $\mathbf{m} \neq 0$ we have

$$j \neq l. \tag{7.45}$$

Lemma 7.9 now implies that together with (7.44) we have

$$|\lambda_l(\mathbf{a} + \mathbf{m}; I + \mathbf{m}) - \lambda_l(\mathbf{b} + \mathbf{m}; I + \mathbf{m})| \ll L\rho + \rho^{-4Mp}. \tag{7.46}$$

Once again using lemma 7.4, we deduce that $\lambda_j(\mathbf{b}; I) = g(\boldsymbol{\eta}_1) + O(\rho^{-4Mp})$ and $\lambda_l(\mathbf{b} + \mathbf{m}; I + \mathbf{m}) = g(\boldsymbol{\eta}_2) + O(\rho^{-4Mp})$ for different points $\boldsymbol{\eta}_1 \in \Upsilon(\mathbf{b}; I)$ and $\boldsymbol{\eta}_2 \in \Upsilon(\mathbf{b} + \mathbf{m}; I + \mathbf{m})$ (the points $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are different

because of (7.45)). In particular, these inclusions imply $\boldsymbol{\eta}_1 - \mathbf{b} \in \mathbb{Z}^d$ and $\boldsymbol{\eta}_2 - \mathbf{b} \in \mathbf{m} + \mathbb{Z}^d = \mathbb{Z}^d$. This proves the second statement. $\square$

Thus, we have proved the following lemma, which is the main result of this section:

Lemma 7.12. *Let $\mathfrak{V} \in \mathcal{V}(n)$. Then there are two mappings $\tilde{g}, g :$ $\Xi_2(\mathfrak{V}) \to \mathbb{R}$ which satisfy the following properties:*
(i) $\tilde{g}(\boldsymbol{\xi})$ is an eigenvalue of $P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$ with $\{\boldsymbol{\xi}\} = \mathbf{k}$. All eigenvalues of $P(\mathfrak{V})H'(\mathbf{k})P(\mathfrak{V})$ inside J are in the image of $\tilde{g}$.
(ii) If $\boldsymbol{\xi} \in \mathcal{A}_1$, then $|\tilde{g}(\boldsymbol{\xi}) - g(\boldsymbol{\xi})| \le C\rho^{-4Mp}$ and $|g(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \le 2v$.
(iii) $g(\boldsymbol{\xi}) = r^2 + s(\boldsymbol{\xi})$ with $r := |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|$ and $\frac{\partial s}{\partial r} = O(\rho^{1/3})$.

Proof. The only statement which has not been checked so far is that $|g(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \le 2v$. This follows immediately from the second statement of this lemma together with the inequality $|\tilde{g}(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \le v$ and $|\tilde{g}(\boldsymbol{\xi}) - g(\boldsymbol{\xi})| \le C\rho^{-4Mp}$. $\square$

Now, we can put together the results of the previous sections

Theorem 7.13. *Suppose, R is sufficiently large, all conditions before lemma 5.1 are satisfied, and $R < \rho^{pd^{-1}/2}$. Then there are two 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 $\{\boldsymbol{\xi}\} = \mathbf{k}$; $|f(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \le 2v$. 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 $\boldsymbol{\xi} \in \mathcal{A}_1$, then $|f(\boldsymbol{\xi}) - g(\boldsymbol{\xi})| \le C\rho^{-4Mp}$.
(iii) We can decompose the domain of g into the disjoint union: $\mathcal{A} = \mathcal{B} \cup \bigcup_{n=1}^{d-1} \bigcup_{\mathfrak{V} \in \mathcal{V}(6MR,n)} \Xi_2(\mathfrak{V})$. For any $\boldsymbol{\xi} \in \mathcal{B}$

$$g(\boldsymbol{\xi}) = |\mathbf{F}\boldsymbol{\xi}|^2$$
$$+ \sum_{j=1}^{2M} \sum_{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_j \in \Theta'_M} \sum_{2 \le n_1 + \dots + n_j \le 2M} C_{n_1,\dots,n_j} \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_1 \rangle^{-n_1} \dots \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_j \rangle^{-n_j}.$$
$$(7.47)$$

For any $\boldsymbol{\xi} \in \Xi_2(\mathfrak{V})$
$$g(\boldsymbol{\xi}) = r^2(\boldsymbol{\xi}) + s(\boldsymbol{\xi}), \qquad (7.48)$$
with $r(\boldsymbol{\xi}) = |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|$, $s(\boldsymbol{\xi}) = s(r, \boldsymbol{\xi}_{\mathfrak{V}}, \boldsymbol{\xi}'_{\mathfrak{V}})$, $\frac{\partial s}{\partial r} = O(\rho^{1/3})$.

Proof. We have described the construction of the mapping f at the end of section 5. Mapping g is constructed in sections 6 and 7. $\square$

Let us formulate an important property of the mapping g, which is a global version of lemma 7.11.

Lemma 7.14. *Let $I = [\mathbf{a}, \mathbf{b}] \subset \mathcal{A}_1$ be a straight interval of length $L := |\mathbf{b} - \mathbf{a}| \ll \rho^{-1}$. Then there exists an integer vector $\mathbf{n}$ such that $|g(\mathbf{b} + \mathbf{n}) - g(\mathbf{a})| \ll L\rho + \rho^{-4Mp+d}$. Moreover, suppose in addition that there exists an integer vector $\mathbf{m} \neq 0$ such that the interval $I + \mathbf{m}$ is entirely inside $\mathcal{A}_1$. Then there exist two different integer vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ such that $|g(\mathbf{b} + \mathbf{n}_1) - g(\mathbf{a})| \ll L\rho + \rho^{-4Mp+d}$ and $|g(\mathbf{b} + \mathbf{n}_2) - g(\mathbf{a} + \mathbf{m})| \ll L\rho + \rho^{-4Mp+d}$.*

Proof. Let us parametrise the interval I so that

$$I = \{\boldsymbol{\xi}(t), t \in [t_{min}, t_{max}]\},$$

$\mathbf{a} = \boldsymbol{\xi}(t_{min})$, $\mathbf{b} = \boldsymbol{\xi}(t_{max})$. Let us prove the first statement. If the interval $\boldsymbol{\xi}(t)$ lies entirely inside $\mathcal{B}$, then the statement is obvious since the length of the gradient of g inside $\mathcal{B}$ is $\ll \rho$, so we can take $\mathbf{n} = 0$. If the interval $\boldsymbol{\xi}(t)$ lies entirely inside $\Xi_2(\mathfrak{V})$ for $\mathfrak{V} \in \mathcal{V}(n)$, the statement has been proved in lemma 7.11. Consider the general case. Denote by $y_j(\mathbf{k}) := \mu_j(H'(\mathbf{k}))$ the j-th eigenvalue of $H'(\mathbf{k})$. Then the definition of the mapping f implies that if $y_j(\mathbf{k}) \in J$, then

$$y_j(\mathbf{k}) = f(\mathbf{n} + \mathbf{k})$$

for some integer vector $\mathbf{n}$; the opposite is also true, namely if $f(\mathbf{n}+\mathbf{k}) \in J$, then $f(\mathbf{n} + \mathbf{k}) = y_j(\mathbf{k})$ for some j. Notice also that for each j the function y_j is continuous.

Now let us return to the study of the behaviour of the function $g(\boldsymbol{\xi}(t))$. Suppose for definiteness that $\boldsymbol{\xi}(t_{min}) \in \mathcal{B}$. Then, as we mentioned in the beginning of proof, since the gradient of g has length $\ll \rho$, we have $|g(\boldsymbol{\xi}(t)) - g(\boldsymbol{\xi}(t_{min}))| \ll |\boldsymbol{\xi}(t) - \boldsymbol{\xi}(t_{min})|\rho$ as soon as $\boldsymbol{\xi}(t)$ stays inside $\mathcal{B}$. Suppose that t_1 is the point at which $\boldsymbol{\xi}(t)$ crosses the boundary of $\mathcal{B}$. Then

$$|g(\boldsymbol{\xi}(t_1 - 0)) - g(\boldsymbol{\xi}(t_{min}))| \ll |\boldsymbol{\xi}(t) - \boldsymbol{\xi}(t_{min})|\rho \qquad (7.49)$$

According to the relationship between the mapping f and functions y_j stated above, there exists an index j such that $f(\boldsymbol{\xi}(t_1-0)) = y_j(\{\boldsymbol{\xi}(t_1 - 0)\})$ (recall that if $\boldsymbol{\xi} = \mathbf{n} + \mathbf{k}$, then we call $\mathbf{k} = \{\boldsymbol{\xi}\}$ the fractional part of $\boldsymbol{\xi}$). Since y_j is continuous function, $y_j(\{\boldsymbol{\xi}(t_1 - 0)\}) = y_j(\{\boldsymbol{\xi}(t_1 + 0)\})$. Using the relationship between the mapping f and functions y_j again, we deduce that there exists an integer vector $\mathbf{n}_1$ such that $y_j(\{\boldsymbol{\xi}(t_1 + 0))\}) = f(\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1)$. Property (ii) of theorem 7.13 implies that $f(\boldsymbol{\xi}(t_1 - 0)) = g(\boldsymbol{\xi}(t_1 - 0)) + O(\rho^{-4Mp})$ and, similarly, $f(\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1) = g(\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1) + O(\rho^{-4Mp})$. All these estimates imply

$$g(\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1) = g(\boldsymbol{\xi}(t_1 - 0)) + O(\rho^{-4Mp}). \qquad (7.50)$$

Since $\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1 \in \mathcal{A}_1$, we have either $\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1 \in \Xi_2(\mathfrak{V})$ or $\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1 \in \mathcal{B}$. Assume the former. Let $t_2 > t_1$ be the smallest value of t at which $\boldsymbol{\xi}(t) + \mathbf{n}_1$ crosses the boundary of $\Xi_2(\mathfrak{V})$. Then lemma 7.11 implies that thee exists an integer vector $\mathbf{n}_2$ such that

$$|g(\boldsymbol{\xi}(t_2 - 0) + \mathbf{n}_2) - g(\boldsymbol{\xi}(t_1 + 0) + \mathbf{n}_1)| \ll |\boldsymbol{\xi}(t_2) - \boldsymbol{\xi}(t_1)|\rho + O(\rho^{-4Mp}).$$
(7.51)

Now repeating the argument we have already used at the moment t_1, we deduce that there exists an integer $\mathbf{n}_3$ such that

$$g(\boldsymbol{\xi}(t_2 + 0) + \mathbf{n}_3) = g(\boldsymbol{\xi}(t_2 - 0) + \mathbf{n}_2) + O(\rho^{-4Mp}).$$
(7.52)

Now we repeat the process and increase t beginning from t_2 until we hit another piece of boundary of some $\Xi_2(\mathfrak{W})$ at $t = t_3$, etc. The shift of the function g at each of the points t_j of hitting the boundary is $O(\rho^{-4Mp})$. The number of such points is $\ll \rho^d$, since for each fixed integer vector $\mathbf{m}$ the number of intersections of the interval $(\boldsymbol{\xi}(t) + \mathbf{m})$ $(t \in [t_{min}, t_{max}])$ with the boundaries of all sets $\Xi_2(\mathfrak{V})$ is finite, and the number of possible integer vectors $\mathbf{m}$ allowed here is $\ll \rho^d$ (obviously, the length of each of these integer vectors is $\ll \rho$). Now formulas (7.49)–(7.52) lead to the desired result.

The proof of the second statement is similar and can be derived from the proof of the first statement in the same way as the proof of the second part of lemma 7.11 follows from the proof of the first part of that lemma. $\square$

Now it remains to extend the above results to the 'full' operator $H(\mathbf{k})$.

Corollary 7.15. *For each natural N 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 $\{\boldsymbol{\xi}\} = \mathbf{k}$; $|f(\boldsymbol{\xi}) - |\mathbf{F}\boldsymbol{\xi}|^2| \le 2v$. 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 $\boldsymbol{\xi} \in \mathcal{A}_1$, then $|f(\boldsymbol{\xi}) - g(\boldsymbol{\xi})| \le \rho^{-N}$.

(iii) We can decompose the domain of g into the disjoint union: $\mathcal{A} = \mathcal{B} \cup \bigcup_{n=1}^{d-1} \bigcup_{\mathfrak{V} \in \mathcal{V}(n)} \Xi_2(\mathfrak{V})$. For any $\boldsymbol{\xi} \in \mathcal{B}_\rho$

$$g(\boldsymbol{\xi}) = |\mathbf{F}\boldsymbol{\xi}|^2$$

$$+ \sum_{j=1}^{2M} \sum_{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_j \in \Theta'_M} \sum_{2 \le n_1 + \dots + n_j \le 2M} C_{n_1, \dots, n_j} \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_1 \rangle^{-n_1} \dots \langle \boldsymbol{\xi}, \mathbf{G}\boldsymbol{\eta}_j \rangle^{-n_j}$$
(7.53)

with $M = [(N + d)(4p)^{-1}] + 1$. *For any* $\boldsymbol{\xi} \in (\Xi_2(\mathfrak{V}) \cap \mathcal{A}_1)$

$$g(\boldsymbol{\xi}) = r^2 + s(\boldsymbol{\xi}), \qquad (7.54)$$

with $r := |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|$, $s(\boldsymbol{\xi}) = s(r, \boldsymbol{\xi}_{\mathfrak{V}}, \boldsymbol{\xi}_{\mathfrak{V}}')$ *and* $\frac{\partial s}{\partial r} = O(\rho^{1/3})$.

(iv) Let $I = [\mathbf{a}, \mathbf{b}] \subset \mathcal{A}_1$ *be a straight interval of length* $L := |\mathbf{b} - \mathbf{a}| \ll \rho^{-1}$. *Then there exists an integer vector* $\mathbf{n}$ *such that* $|g(\mathbf{b} + \mathbf{n}) - g(\mathbf{a})| \ll L\rho + \rho^{-N}$. *Moreover, suppose* $\mathbf{m} \neq 0$ *is a given integer vector such that the interval* $I + \mathbf{m}$ *is entirely inside* $\mathcal{A}_1$. *Then there exist two different integer vectors* $\mathbf{n}_1$ *and* $\mathbf{n}_2$ *such that* $|g(\mathbf{b} + \mathbf{n}_1) - g(\mathbf{a})| \ll L\rho + \rho^{-N}$ *and* $|g(\mathbf{b} + \mathbf{n}_2) - g(\mathbf{a} + \mathbf{m})| \ll L\rho + \rho^{-N}$.

Proof. We use theorem 7.13 for the operator $H'(\mathbf{k})$ with $M = [(N + d)(4p)^{-1}] + 1$. Estimate (5.4) implies that $|\mu_j(H(\mathbf{k})) - \mu_j(H'(\mathbf{k}))| < \rho^{-N-1}$, so that all the required properties are fulfilled. $\qquad\square$

Remark 7.16. The function f is not necessarily continuous.

Before we continue with the proof of the Bethe-Sommerfeld conjecture, let us formulate a theorem which immediately follows from our results, just to illustrate their usefulness. Recall that by $N(\lambda)$ we have denoted the integrated density of states of the operator (1.4) defined in (2.2).

Theorem 7.17. *For each natural n we have the following estimate:* $N(\lambda + \lambda^{-n}) - N(\lambda - \lambda^{-n}) = O(\lambda^{d/2-n-1})$.

Proof. We use corollary 7.15 with $N = 2n + 1$. Then

$$\begin{aligned} N(\lambda + \lambda^{-n}) - N(\lambda - \lambda^{-n}) &= \mathrm{vol}(f^{-1}([\lambda - \lambda^{-n}, \lambda + \lambda^{-n}])) \\ &\leq \mathrm{vol}(g^{-1}([\lambda - 2\lambda^{-n}, \lambda + 2\lambda^{-n}])) = O(\lambda^{d/2-n-1}), \end{aligned} \qquad (7.55)$$

the last equality being an easy geometric exercise (which will anyway be established in the next section). $\qquad\square$

Remark 7.18. As it was pointed out to the author by Yu.Karpeshina, it seems possible that using the results of this paper (including the results from the next section) one can prove the following lower bound:

$$N(\lambda + \varepsilon) - N(\lambda) \gg \varepsilon\lambda^{(d-2)/2},$$

uniformly over $\varepsilon < 1$ as $\lambda \to \infty$ (in particular, ε does not have to be a negative power of λ). We will not prove this estimate in our paper though.

8. Proof of the Bethe-Sommerfeld conjecture

Throughout this section we keep the notation from the previous section. Without specific mentioning, we always assume that ρ is sufficiently large; the precise value of the power N will be chosen later. In what follows, it will be convenient to consider a slightly slimmed down resonance set. Namely, we introduce the set

$$\tilde{\mathcal{B}} := \{\boldsymbol{\xi} \in \mathcal{A}_1 : |\boldsymbol{\xi}_{\mathfrak{U}}| > \rho^{1/2}, \forall \mathfrak{U} \in \mathcal{V}(1)\}.$$

In other words, $\tilde{\mathcal{B}}$ consists of all points $\boldsymbol{\xi} \in \mathcal{A}_1$ the $\mathbf{F}$-projections of which to all vectors $\boldsymbol{\eta} \in \Theta'_{6M}$ has $\mathbf{F}$-length larger than $\rho^{1/2}$. Obviously, $\tilde{\mathcal{B}} \subset \mathcal{B}$. We also denote $\tilde{\mathcal{D}} := \mathcal{A}_1 \setminus \tilde{\mathcal{B}}$.

Now we will study various properties of mappings f and g. We begin with the function g.

For each positive $\delta \leq v$ denote $\mathcal{A}(\delta)$, $\mathcal{B}(\delta)$, and $\mathcal{D}(\delta)$ to be intersections of $g^{-1}([\rho^2 - \delta, \rho^2 + \delta])$ with $\mathcal{A}_1$, $\tilde{\mathcal{B}}$, and $\tilde{\mathcal{D}}$ correspondingly. The following is a simple geometry:

Lemma 8.1. *The following estimates hold:*

$$\mathrm{vol}(\mathcal{A}(\delta)) \asymp \rho^{d-2}\delta, \tag{8.1}$$

$$\mathrm{vol}(\mathcal{B}(\delta)) \asymp \rho^{d-2}\delta, \tag{8.2}$$

and

$$\mathrm{vol}(\mathcal{D}(\delta)) \ll \rho^{(3d-7)/3}\delta. \tag{8.3}$$

Proof. Let $\boldsymbol{\xi} = r\boldsymbol{\xi}' \in \mathcal{B}$, $|\mathbf{F}\boldsymbol{\xi}'| = 1$ Then the definition of g implies that

$$\frac{\partial g}{\partial r} \asymp \rho \tag{8.4}$$

uniformly over $\boldsymbol{\xi}'$. Therefore, for each fixed $\boldsymbol{\xi}'$ the intersection of $g^{-1}([\rho^2 - \delta, \rho^2 + \delta])$ with the set $\{r\boldsymbol{\xi}', r > 0\}$ is an interval of length $\asymp \delta\rho^{-1}$. Integrating over $\boldsymbol{\xi}'$, we obtain (8.2). Estimate (8.3) is obtained in a similar way, only for $\boldsymbol{\xi} \in \Xi(\mathfrak{V})$ we put $r := |\mathbf{F}\boldsymbol{\xi}_{\mathfrak{V}}^{\perp}|$. Then the estimate (8.4) is still valid. Let $\boldsymbol{\eta} \in \Theta'_M$. Then (8.4) implies that the set of all points $\boldsymbol{\xi} \in \mathcal{A}(\delta)$ such that the $\mathbf{F}$-projection of $\boldsymbol{\xi}$ onto $\boldsymbol{\eta}$ has $\mathbf{F}$-length smaller than $\rho^{1/2}$ has volume $O(\rho^{(2d-5)/2}\delta)$. Since the number of elements in Θ'_M is $O(R^d) = O(\rho^{p/2})$, we have

$$\mathrm{vol}(\mathcal{D}(\delta)) \ll \rho^{(2d-5+p)/2}\delta \ll \rho^{(3d-7)/3}\delta,$$

since $p < 1/3$. Finally, (8.1) is the sum of (8.2) and (8.3). □

Remark 8.2. Putting $\delta = 2\lambda^{-n}$ in (8.1), we establish the last equality in (7.55).

The next estimate is more subtle.

Lemma 8.3. *Let $d \geq 3$. Then for large enough ρ and $\delta < \rho^{-1}$ the following estimate holds uniformly over $\mathbf{a} \in \mathbb{R}^d$ with $|\mathbf{a}| > 1$:*

$$\operatorname{vol}\big(\mathcal{B}(\delta) \cap (\mathcal{B}(\delta) + \mathbf{a})\big) \ll (\delta^2 \rho^{d-3} + \delta \rho^{-d}). \qquad (8.5)$$

If $d = 2$, similar estimate holds with $\delta^{3/2} + \delta \rho^{-2}$ in the RHS.

Proof. After making the substitution $\boldsymbol{\nu} = \mathbf{F}\boldsymbol{\xi}$, the function g in new coordinates will have the form $h(\boldsymbol{\nu}) = |\boldsymbol{\nu}|^2 + G(\boldsymbol{\nu})$, with

$$G(\boldsymbol{\nu}) = O(|\boldsymbol{\nu}|^{-1/2}) \qquad (8.6)$$

and

$$\frac{\partial G}{\partial \nu_j} \leq C_2 |\boldsymbol{\nu}|^{-1} \qquad (8.7)$$

for all $j = 1, \ldots, d$, provided $\boldsymbol{\nu} \in \mathbf{F}\tilde{\mathcal{B}}$; these estimates follow from lemma 6.1. We need to estimate the volume of the set

$$\mathcal{X} = \{\boldsymbol{\nu} \in \big(\mathbf{F}(\mathcal{B}) \cap (\mathbf{F}(\mathcal{B}) + \mathbf{Fa})\big), h(\boldsymbol{\nu}) \in [\rho^2 - \delta, \rho^2 + \delta],$$
$$h(\boldsymbol{\nu} - \mathbf{Fa}) \in [\rho^2 - \delta, \rho^2 + \delta]\}. \qquad (8.8)$$

Indeed, we have $\mathcal{X} = \mathbf{F}\big(\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta) + \mathbf{a}\big)\big)$, so the volume of $\mathcal{X}$ equals $\det \mathbf{F}$ times the volume of the set $\mathcal{B}(\delta) \cap \big(\mathcal{B}(\delta) + \mathbf{a}\big)$. Denote $\mathbf{b} := \mathbf{Fa}$. First, we will estimate the 2-dimensional area of the intersection of $\mathcal{X}$ with arbitrary 2-dimensional plane containing the origin and vector $\mathbf{b}$; the volume of $\mathcal{X}$ then can be obtained using the integration in cylindrical coordinates. So, let $\mathfrak{V}$ be any 2-dimensional plane containing the origin and $\mathbf{b}$, and let us estimate the area of $\mathcal{X}_{\mathfrak{V}} := \mathfrak{V} \cap \mathcal{X}$. Let us introduce cartesian coordinates in $\mathfrak{V}$ so that $\boldsymbol{\nu} \in \mathfrak{V}$ has coordinates (ν_1, ν_2) with ν_1 going along $\mathbf{b}$, and ν_2 being orthogonal to $\mathbf{b}$. For any $\boldsymbol{\nu} \in \mathcal{X}_{\mathfrak{V}}$ estimate (8.6) implies

$$h(\boldsymbol{\nu}) = \nu_1^2 + \nu_2^2 + O(\rho^{-1/2}),$$

and so

$$2\delta \geq |h(\boldsymbol{\nu}) - h(\boldsymbol{\nu} - \mathbf{b})| = |\nu_1^2 - (|\mathbf{b}| - \nu_1)^2| + O(\rho^{-1/2}).$$

This implies that

$$\frac{|\mathbf{b}|}{3} < \nu_1 < \frac{2|\mathbf{b}|}{3} \qquad (8.9)$$

when ρ is sufficiently large, and therefore

$$\frac{\partial h(\boldsymbol{\nu})}{\partial \nu_1} \gg |\mathbf{b}| \qquad (8.10)$$

whenever $\boldsymbol{\nu} \in \mathcal{X}_{\mathfrak{V}}$. Thus, for any fixed $t \in \mathbb{R}$, the intersection of the line $\nu_2 = t$ with $\mathcal{X}_{\mathfrak{V}}$ is an interval of length $\ll |\mathbf{b}|^{-1}\delta$.

Let us cut $\mathcal{X}_{\mathfrak{V}}$ into two parts: $\mathcal{X}_{\mathfrak{V}} = \mathcal{X}_{\mathfrak{V}}^1 \cup \mathcal{X}_{\mathfrak{V}}^2$ with $\mathcal{X}_{\mathfrak{V}}^1 := \{\boldsymbol{\nu} \in \mathcal{X}_{\mathfrak{V}}, |\nu_2| \le 2C_2\rho^{-1}\}$, $\mathcal{X}_{\mathfrak{V}}^2 = \mathcal{X}_{\mathfrak{V}} \setminus \mathcal{X}_{\mathfrak{V}}^1$, and estimate the volumes of these sets (C_2 is the constant from (8.7)). A simple geometrical argument shows that if $\mathcal{X}_{\mathfrak{V}}^1$ is nonempty, then $|\mathbf{b}| \gg \rho$. This, together with the remark after (8.10), implies that the area of $\mathcal{X}_{\mathfrak{V}}^1$ is $\ll \rho^{-2}\delta$. Now we define the 'rotated' set $\mathcal{X}^1$ which consists of the points from $\mathcal{X}$ which belong to $\mathcal{X}_{\mathfrak{V}}^1$ for some $\mathfrak{V}$. Computing the volume of this set using integration in the cylindrical coordinates, we obtain

$$\mathrm{vol}(\mathcal{X}^1) \ll \rho^{-d}\delta. \qquad (8.11)$$

Now consider $\mathcal{X}_{\mathfrak{V}}^2$. Let us decompose $\mathcal{X}_{\mathfrak{V}}^2 = \overline{\mathcal{X}_{\mathfrak{V}}^2} \cup \underline{\mathcal{X}_{\mathfrak{V}}^2}$, where

$$\overline{\mathcal{X}_{\mathfrak{V}}^2} = \{\boldsymbol{\nu} \in \mathcal{X}_{\mathfrak{V}}^2 : \nu_2 > 0\}$$

and

$$\underline{\mathcal{X}_{\mathfrak{V}}^2} = \{\boldsymbol{\nu} \in \mathcal{X}_{\mathfrak{V}}^2 : \nu_2 < 0\}.$$

Notice that for any $\boldsymbol{\nu} \in \overline{\mathcal{X}_{\mathfrak{V}}^2}$, formula (8.7) implies

$$\frac{\partial h(\boldsymbol{\nu})}{\partial \nu_2} \gg \nu_2. \qquad (8.12)$$

Let $\boldsymbol{\nu}^l = (\nu_1^l, \nu_2^l)$ be the point in the closure of $\overline{\mathcal{X}_{\mathfrak{V}}^2}$ with the smallest value of the first coordinate: $\nu_1^l \le \nu_1$ for any $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \overline{\mathcal{X}_{\mathfrak{V}}^2}$. Analogously, we define $\boldsymbol{\nu}^r$ to be the point in the closure of $\overline{\mathcal{X}_{\mathfrak{V}}^2}$ with the largest first coordinate, $\boldsymbol{\nu}^t$ the point with the largest second coordinate, and $\boldsymbol{\nu}^b$ the point with the smallest second coordinate (see Figure 10 for an illustration). Note that $\nu^t \ll \rho$.

Let us prove that

$$\nu_1^r - \nu_1^l \ll \delta. \qquad (8.13)$$

Indeed, suppose first that $\nu_2^r \ge \nu_2^l$. Let $\boldsymbol{\nu}^{rl} := (\nu_1^r, \nu_2^l)$. Then, since h is an increasing function of ν_2 when $\nu_2 > 2C_2\rho^{-1}$, we have $h(\boldsymbol{\nu}^{rl}) \le h(\boldsymbol{\nu}^r) \le \rho^2 + \delta$. Therefore, $h(\boldsymbol{\nu}^{rl}) - h(\boldsymbol{\nu}^l) \le 2\delta$. Estimate (8.10) then implies (8.13).

Suppose now that $\nu_2^r \le \nu_2^l$. Let $\boldsymbol{\nu}^{lr} := (\nu_1^l, \nu_2^r)$. Then $h(\boldsymbol{\nu}^{lr} - \mathbf{b}) \le h(\boldsymbol{\nu}^l - \mathbf{b}) \le \rho^2 + \delta$. Therefore, $h(\boldsymbol{\nu}^{lr} - \mathbf{b}) - h(\boldsymbol{\nu}^r - \mathbf{b}) \le 2\delta$. Now, (8.9) and (8.10) imply (8.13).

Thus, we have estimated the width of $\mathcal{X}_{\mathfrak{V}}^2$. Let us estimate its hight (i.e. $\nu_2^t - \nu_2^b$). Let us assume that $\nu_1^t \ge \nu_1^b$; otherwise, we use the same trick as in the previous paragraph and consider $h(\cdot - \mathbf{b})$ instead of h. Let $\boldsymbol{\nu}^{bt} := (\nu_1^b, \nu_2^t)$. Then $h(\boldsymbol{\nu}^{bt}) \le h(\boldsymbol{\nu}^t) \le \rho^2 + \delta$. Therefore,
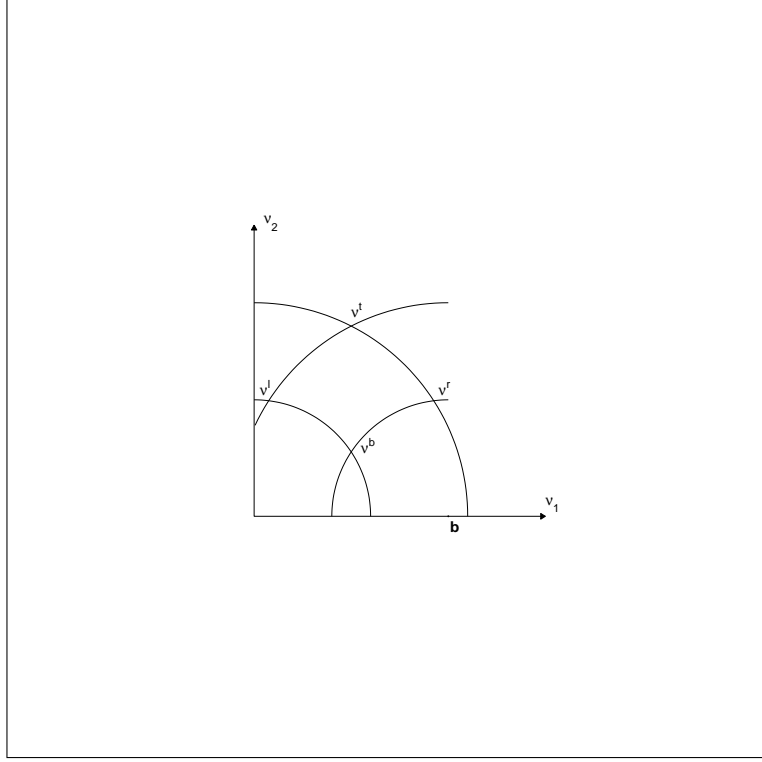
FIGURE 10. The set $\overline{\mathcal{X}^2_{\mathfrak{Y}}}$ (the area bounded by four arcs)

$h(\boldsymbol{\nu}^{bt}) - h(\boldsymbol{\nu}^b) \leq 2\delta$. Now, (8.12) implies

$$(\nu_2^t)^2 - (\nu_2^b)^2 = 2\int_{\nu_2^b}^{\nu_2^t} \nu_2 d\nu_2 \ll \int_{\nu_2^b}^{\nu_2^t} \frac{\partial h}{\partial \nu_2}(\nu_1^b, \nu_2)d\nu_2 \leq 2\delta. \qquad (8.14)$$

Therefore, we have the following estimate for the hight of $\mathcal{X}^2_{\mathfrak{Y}}$:

$$\nu_2^t - \nu_2^b \ll \frac{\delta}{\nu_2^t + \nu_2^b}. \qquad (8.15)$$

Now, we can estimate the volume of $\mathcal{X}^2 := \mathcal{X} \setminus \mathcal{X}^1$ using estimates (8.13) and (8.15). The cylindrical integration produces the following:

$$\mathrm{vol}(\mathcal{X}^2) \ll \frac{\delta^2}{\nu_2^t + \nu_2^b}(\nu_2^t)^{d-2} \leq \delta^2(\nu_2^t)^{d-3} \leq \delta^2\rho^{d-3}. \qquad (8.16)$$

Equations (8.11) and (8.16) imply (8.5). If $d = 2$, we have to notice that (8.14) implies $\nu_2^t - \nu_2^b \ll \delta^{1/2}$ and then use (8.11) and (8.13). $\qquad \square$

As was mentioned already, the function f is not necessarily continuous. We now give a sufficient condition for its continuity. Recall that v is the L_∞-norm of the potential V.

Lemma 8.4. *Let $\boldsymbol{\xi} \in \mathcal{B}(v)$ be a point of discontinuity of f. Then there is a non-zero vector $\mathbf{n} \in \mathbb{Z}^d$ such that*

$$|g(\boldsymbol{\xi} + \mathbf{n}) - g(\boldsymbol{\xi})| \leq 2\rho^{-N}. \tag{8.17}$$

Proof. If $\boldsymbol{\xi} = \mathbf{m} + \mathbf{k} \in \mathcal{B}(v)$ is a point of discontinuity of a bounded function f, there exist two sequences $\{\boldsymbol{\xi}_j\}$ and $\{\tilde{\boldsymbol{\xi}}_j\}$ which both converge to $\boldsymbol{\xi}$, such that the limits $\lambda(\boldsymbol{\xi}) := \lim f(\boldsymbol{\xi}_j)$ and $\tilde{\lambda}(\boldsymbol{\xi}) := \lim f(\tilde{\boldsymbol{\xi}}_j)$ exist and are different. Since the points $f(\boldsymbol{\xi}_j)$ are eigenvalues of $H(\{\boldsymbol{\xi}_j\})$, the limit λ is an eigenvalue of $H(\mathbf{k})$ (it is well-known that the spectrum of $H(\mathbf{k})$ is continuously dependent on $\mathbf{k}$). The same argument implies that $\tilde{\lambda}$ is also an eigenvalue of $H(\mathbf{k})$. Since $\lambda \neq \tilde{\lambda}$, at most one of these points can be equal to $f(\boldsymbol{\xi})$. Say, $\tilde{\lambda} \neq f(\boldsymbol{\xi})$. But since $\tilde{\lambda}$ is inside J, it must belong to the image of f, say $\tilde{\lambda} = f(\tilde{\boldsymbol{\xi}})$, $\{\tilde{\boldsymbol{\xi}}\} = \{\boldsymbol{\xi}\}$. Thus, $\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi} + \mathbf{n}$ with $\mathbf{n} \in \mathbb{Z}^d$. Since the function g is continuous in $\mathcal{B}$, $\lim g(\tilde{\boldsymbol{\xi}}_j) = g(\boldsymbol{\xi})$, and so

$$|g(\boldsymbol{\xi}) - \tilde{\lambda}| = \lim |g(\tilde{\boldsymbol{\xi}}_j) - f(\tilde{\boldsymbol{\xi}}_j)| \leq \rho^{-N}.$$

But we also have $|g(\tilde{\boldsymbol{\xi}}) - \tilde{\lambda}| = |g(\tilde{\boldsymbol{\xi}}) - f(\tilde{\boldsymbol{\xi}})| \leq \rho^{-N}$. The last two inequalities imply (8.17). $\qquad\square$

Corollary 8.5. *There is a constant C_3 with the following properties. Let*

$$I := \{\boldsymbol{\xi}(t) : t \in [t_{min}, t_{max}]\} \subset \mathcal{B}(v).$$

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 $\mathbf{n} \in \mathbb{Z}^d$ $g(\boldsymbol{\xi}(t_0) + \mathbf{n})$ is either outside the interval

$$[g(\boldsymbol{\xi}(t_0)) - C_3\rho^{-N} - C_3\rho L, g(\boldsymbol{\xi}(t_0)) + C_3\rho^{-N} + C_3\rho L]$$

or not defined. Then $f(\boldsymbol{\xi}(t))$ is a continuous function of t.

Proof. Suppose not. Then previous lemma implies that there is a point $t_1 \in [t_{min}, t_{max}]$ and a non-zero vector $\mathbf{n} \in \mathbb{Z}^d$ such that $|g(\boldsymbol{\xi}(t_1) + \mathbf{n}) - g(\boldsymbol{\xi}(t_1))| \leq 2\rho^{-N}$. Since $|\boldsymbol{\xi}(t_1) - \boldsymbol{\xi}(t_0)| \leq |\boldsymbol{\xi}(t_{max}) - \boldsymbol{\xi}(t_{min})| \leq L$, it follows that $(I + \mathbf{n}) \subset \mathcal{A}_1$, and now lemma 7.14 implies that for two different integer vectors $\mathbf{m}_1$ and $\mathbf{m}_2$ we have $|g(\boldsymbol{\xi}(t_0) + \mathbf{m}_1) - g(\boldsymbol{\xi}(t_1) + \mathbf{n})| \ll \rho L + \rho^{-N}$ and $|g(\boldsymbol{\xi}(t_0) + \mathbf{m}_2) - g(\boldsymbol{\xi}(t_1))| \ll \rho L + \rho^{-N}$. Since $\boldsymbol{\xi}(t) \in \mathcal{B}$ for all t and the length of the gradient of g is $\ll \rho$ in $\mathcal{B}$, we also have $|g(\boldsymbol{\xi}(t_1)) - g(\boldsymbol{\xi}(t_0))| \ll \rho L$. Thus, we have $|g(\boldsymbol{\xi}(t_0) + \mathbf{m}_j) -$

$g(\boldsymbol{\xi}(t_0))| \leq C\rho^{-N} + C\rho L$ $(j = 1, 2)$. Since at least one of vectors $\mathbf{m}_j$ is non-zero, this contradicts the assumption of the corollary. $\qquad\square$

Now we are ready to prove the Bethe-Sommerfeld conjecture. Since in the two-dimensional case it has been proved, we will assume that $d \geq 3$.

Theorem 8.6. *Let $d \geq 3$. Then all sufficiently large points $\lambda = \rho^2$ are inside the spectrum of H. Moreover, there exists a positive constant c_4 such that for large enough ρ the whole interval $[\rho^2 - c_4\rho^{1-d}, \rho^2 + c_4\rho^{1-d}]$ lies inside some spectral band.*

Proof. Put $N = d$ in the corollary 7.15. Also put $\delta = c_4\rho^{1-d}$(the precise value of c_4 will be chosen later). For each unit vector $\boldsymbol{\eta} \in \mathbb{R}^d$ we denote $I_{\boldsymbol{\eta}}$ to be the intersection of $\{r\boldsymbol{\eta}, \ r > 0\}$ with $\mathcal{A}(\delta)$. We will consider only vectors $\boldsymbol{\eta}$ for which $I_{\boldsymbol{\eta}} \subset \tilde{\mathcal{B}}$. As was mentioned in the proof of lemma 8.1, the length L of any interval $I_{\boldsymbol{\eta}}$ satisfies $L \asymp \delta\rho^{-1}$. Let us prove that f is continuous on at least one of the intervals $I_{\boldsymbol{\eta}} \subset \tilde{\mathcal{B}}$. Suppose this is not the case. Then corollary 8.5 tells us that for each point $\boldsymbol{\xi} \in \mathcal{B}(\delta)$ there is a non-zero integer vector $\mathbf{n}$ such that

$$|g(\boldsymbol{\xi} + \mathbf{n}) - g(\boldsymbol{\xi})| \leq C_3(\rho^{-d} + \rho L) \ll (\rho^{-d} + \delta). \qquad (8.18)$$

Since $|g(\boldsymbol{\xi}) - \rho^2| \leq \delta$, this implies $|g(\boldsymbol{\xi} + \mathbf{n}) - \rho^2| \leq C_5(\rho^{-d} + \delta) =: \delta_1$, and thus $\boldsymbol{\xi} + \mathbf{n} \in \mathcal{A}(\delta_1)$; notice that $C_5 > 1$ and so $\delta_1 > \delta$. Therefore, each point $\boldsymbol{\xi} \in \mathcal{B}(\delta)$ also belongs to the set $\big(\mathcal{A}(\delta_1) - \mathbf{n}\big)$ for a non-zero integer $\mathbf{n}$; obviously, $|\mathbf{n}| \ll \rho$. In other words,

$$\begin{aligned}
\mathcal{B}(\delta) \subset &\bigcup_{\mathbf{n}\in\mathbb{Z}^d\cap B(C\rho),\mathbf{n}\neq 0} \big(\mathcal{A}(\delta_1) - \mathbf{n}\big) \\
&= \bigcup_{\mathbf{n}\neq 0}\big(\mathcal{B}(\delta_1) - \mathbf{n}\big) \cup \bigcup_{\mathbf{n}\neq 0}\big(\mathcal{D}(\delta_1) - \mathbf{n}\big)
\end{aligned} \qquad (8.19)$$

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 $\mathbf{n} \in \mathbb{Z}^d$ 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 $\mathbf{n} \in \mathbb{Z}^d$ 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 $\mathbf{n}_1, \mathbf{n}_2 \in \mathbb{Z}^d$ satisfying $\boldsymbol{\nu} - \mathbf{n}_j \in \mathcal{B}(\delta)$). Then a little thought shows that we can replace $\mathcal{D}(\delta_1)$ by $\mathcal{D}_1(\delta_1)$ in the RHS of (8.19). Indeed, this is shown in the following lemma.

Lemma 8.7. *The following formulae hold:*

$$\mathcal{B}(\delta) \bigcap \left(\bigcup_{\mathbf{n} \neq 0} \left(\mathcal{D}_0(\delta_1) - \mathbf{n} \right) \right) = \emptyset, \tag{8.20}$$

and

$$\left(\bigcup_{\mathbf{n} \neq 0} \left(\mathcal{D}_2(\delta_1) - \mathbf{n} \right) \right) \subset \bigcup_{\mathbf{n} \neq 0} \left(\mathcal{B}(\delta_1) - \mathbf{n} \right) \tag{8.21}$$

Proof. The first formula is an immediate corollary of the definition of $\mathcal{D}_0(\delta_1)$. Let us prove the second formula. Suppose, $\boldsymbol{\nu} \in \mathcal{D}_2(\delta_1)$. Then there exist two integer vectors, $\mathbf{n}_1$ and $\mathbf{n}_2$ such that $\boldsymbol{\nu} - \mathbf{n}_j \in \mathcal{B}(\delta)$. Let $\mathbf{m}$ be an integer vector. Then $\mathbf{m}$ is different from either $\mathbf{n}_1$ or $\mathbf{n}_2$, say $\mathbf{m} \neq \mathbf{n}_1$. Since $\delta_1 \geq \delta$, this implies:

$$\boldsymbol{\nu} - \mathbf{m} = \boldsymbol{\nu} - \mathbf{n}_1 - (\mathbf{m} - \mathbf{n}_1) \in \left(\mathcal{B}(\delta) - (\mathbf{m} - \mathbf{n}_1) \right) \subset \bigcup_{\mathbf{n} \neq 0} \left(\mathcal{B}(\delta_1) - \mathbf{n} \right).$$

This finishes the proof of the lemma. $\square$

This lemma shows that we can re-write (8.19) as

$$\mathcal{B}(\delta) \subset \bigcup_{\mathbf{n} \neq 0} \left(\mathcal{B}(\delta_1) - \mathbf{n} \right) \bigcup \bigcup_{\mathbf{n} \neq 0} \left(\mathcal{D}_1(\delta_1) - \mathbf{n} \right). \tag{8.22}$$

This, obviously, implies

$$\mathcal{B}(\delta) = \bigcup_{\mathbf{n} \neq 0} \left(\left(\mathcal{B}(\delta_1) - \mathbf{n} \right) \cap \mathcal{B}(\delta) \right)$$
$$\bigcup \bigcup_{\mathbf{n} \neq 0} \left(\left(\mathcal{D}_1(\delta_1) - \mathbf{n} \right) \cap \mathcal{B}(\delta) \right). \tag{8.23}$$

Now let us compare volumes of the sets in both sides of (8.23). The volume of the LHS we already know from (8.2): it is $\gg \rho^{d-2}\delta$. The definition of the set $\mathcal{D}_1$ implies that

$$\mathrm{vol}\left(\bigcup_{\mathbf{n}} \left(\left(\mathcal{D}_1(\delta_1) - \mathbf{n} \right) \cap \mathcal{B}(\delta) \right) \right)$$
$$\ll \mathrm{vol}(\mathcal{D}_1(\delta_1)) \ll \rho^{(3d-7)/3}\delta_1 \ll \rho^{(3d-7)/3}(\rho^{-d} + \delta). \tag{8.24}$$

Finally, lemma 8.3, inequality $\delta < \delta_1$ and the fact that the union in (8.23) consists of no more than $C\rho^d$ terms imply

$$\mathrm{vol}\left(\bigcup_{\mathbf{n}} \left(\left(\mathcal{B}(\delta_1) - \mathbf{n} \right) \cap \mathcal{B}(\delta) \right) \right) \ll \rho^d(\delta_1^2 \rho^{d-3} + \delta_1 \rho^{-d})$$
$$\ll \rho^d \left((\rho^{-d} + \delta)^2 \rho^{d-3} + (\rho^{-d} + \delta)\rho^{-d} \right) \ll \delta^2 \rho^{2d-3} + \delta \rho^{d-3} + \rho^{-3}. \tag{8.25}$$

Putting all these inequalities together, we get

$$\rho^{d-2}\delta < C_6(\delta^2\rho^{2d-3} + \delta\rho^{(3d-7)/3} + \rho^{-7/3}). \qquad (8.26)$$

It is time to recall that $\delta = c_4\rho^{1-d}$. Plugging this into (8.26), we obtain

$$c_4\rho^{-1} < C_6(c_4^2\rho^{-1} + c_4\rho^{-4/3} + \rho^{-7/3}). \qquad (8.27)$$

Now, if we choose c_4 to be small enough (i.e. $c_4 < C_6^{-1}$), the inequality (8.27) will not be satisfied for sufficiently large ρ. Thus, our assumption that function f is discontinuous on every interval $I_{\boldsymbol{\eta}} \subset \mathcal{B}(\delta)$ leads to a contradiction (provided we have chosen small enough c_4). Therefore, there is an interval $I_{\boldsymbol{\eta}} \subset \mathcal{B}(\delta)$ on which f is continuous. Since the value of f on one end of this interval is $\leq \rho^2 - c_4\rho^{1-d}$, and the value on the other end is $\geq \rho^2 + c_4\rho^{1-d}$, the point ρ^2 must be in the range of f. The first part of the theorem is proved. In order to prove the second part of the theorem, we notice that the interval $I_{\boldsymbol{\eta}}$ which we found satisfies the following condition: for each point $\boldsymbol{\xi} \in I_{\boldsymbol{\eta}}$ and each non-zero integer vector $\mathbf{n}$ such that $\boldsymbol{\xi} + \mathbf{n} \in \mathcal{A}_1$ we have $|g(\boldsymbol{\xi} + \mathbf{n}) - g(\boldsymbol{\xi})| > 2\rho^{-N}$. This implies $f(\boldsymbol{\xi} + \mathbf{n}) - f(\boldsymbol{\xi}) \neq 0$. Therefore, $f(\boldsymbol{\xi})$ is a simple eigenvalue of $H(\{\boldsymbol{\xi}\})$ for each $\boldsymbol{\xi} \in I_{\boldsymbol{\eta}}$. This implies that the interval $[\rho^2 - c_4\rho^{1-d}, \rho^2 + c_4\rho^{1-d}]$ is inside the spectral band. The theorem is proved. $\qquad\square$

References

1. J.W.S.Cassels *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1959.
2. B.E.J. Dahlberg, E. Trubowitz, *A remark on two dimensional periodic potentials*, Comment. Math. Helvetici **57** (1982), 130–134.
3. B. Helffer, A. Mohamed, *Asymptotics of the density of states for the Schrödinger operator with periodic electric potential*, Duke Math. J. **92** (1998), 1–60.
4. Y. E. Karpeshina, *Perturbation series for the Schrdinger operator with a periodic potential near planes of diffraction*, Comm. Anal. Geom. **4** (1996), no. 3, 339–413.
5. Y. E. Karpeshina, *Perturbation theory for the Schrödinger operator with a periodic potential*, Lecture Notes in Math. vol 1663, Springer Berlin 1997.
6. T. Kato *Perturbation theory for linear operators*, Springer-Verlag, Berlin, 1980.
7. L. Parnovski, A. V. Sobolev, *Bethe-Sommerfeld conjecture for polyharmonic operators*, Duke Math. J., 2001.
8. L. Parnovski, A. V. Sobolev, *Perturbation theory and the Bethe-Sommerfeld conjecture*, Annals H. Poincare, 2001.
9. V.N. Popov, 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).
10. M. Reed, B. Simon, *Methods of modern mathematical physics, IV*, Academic Press, New York, 1975.

11. M. Skriganov, *Proof of the Bethe-Sommerfeld conjecture in dimension two*, Soviet Math. Dokl. **20** (1979), 1, 89–90.
12. M. Skriganov, *Geometrical and arithmetical methods in the spectral theory of the multi-dimensional periodic operators*, Proc. Steklov Math. Inst. Vol. 171, 1984.
13. M.Skriganov, *The spectrum band structure of the three-dimensional Schrödinger operator with periodic potential*, Inv. Math. **80** (1985), 107–121.
14. O.A.Veliev, *Asymptotic formulas for the eigenvalues of the periodic Schrdinger operator and the Bethe-Sommerfeld conjecture*, Functional Anal. Appl. **21** (1987), no. 2, 87–100.
15. O.A.Veliev, *On the spectrum of multidimensional periodic operators*, Theory of Functions, functional analysis and their applications, Kharkov University, **49**, (1988), 17–34 (in Russian).
16. O.A.Veliev, *Perturbation theory for the periodic multidimensional Schrödinger operator and the Bethe-Sommerfeld Conjecture*, Int. J. Contemp. Math. Sci., **2** (2007), no.2, 19–87.

Department of Mathematics, University College London, Gower Street, London, WC1E 6BT UK

E-mail address: Leonid@math.ucl.ac.uk