Complete Spectral Asymptotics for Periodic and Almost Periodic Perturbations of Constant Coefficients Operators and Bethe-Sommerfeld Conjecture in Semiclassical Settings

Analysis Seminar, Einstein Institute of Mathematics, Hebrew University of Jerusalem, Israel

Victor Ivrii

Department of Mathematics, University of Toronto

June 3, 2020
# Table of Contents

1. Introduction
2. Main results
   - Complete Spectral Asymptotics
   - Conditions (A)-(D)
   - Bethe-Sommerfeld Conjecture
3. Reduction of operator
   - Resonant and non-resonant points
   - Classification of resonant points
   - Finally, reduction
   - Gauge transformation and proof of Theorem 5
4. Complete spectral asymptotics
   (sketch of proof)
   - Microhyperbolicity, dynamics and remainder estimate
   - Super-long-term propagation
5. Bethe-Sommerfeld conjecture
   (sketch of proof)
   - Reduction
   - Structure of operator $\mathcal{A}$
   - Non-resonant points
   - Almost antipodal points
   - Resonant points
6. Discussion
7. References
This talk represents two works. The first of them is inspired by several remarkable papers of L. Parnovski and R. Shterenberg [PSh1, PSh2, PS3], S. Morozov, L. Parnovski and R. Shterenberg [MPSh] and earlier papers by A. Sobolev [So1, So2]. I wanted to understand the approach of the authors and, combining their ideas with my own approach, generalize their results.
This talk represents two works. The first of them is inspired by several remarkable papers of L. Parnovski and R. Shterenberg [PSh1, PSh2, PS3], S. Morozov, L. Parnovski and R. Shterenberg [MPSh] and earlier papers by A. Sobolev [So1, So2]. I wanted to understand the approach of the authors and, combining their ideas with my own approach, generalize their results.

In these papers the complete asymptotic expansion of the integrated density of states $N(\lambda)$ for operators $\Delta + V$ was derived as $\lambda \to +\infty$; here $\Delta$ is a positive Laplacian and $V$ is a periodic or almost periodic potential (satisfying certain conditions). In [MPSh] more general operators were considered.
Introduction

This talk represents two works. The first of them is inspired by several remarkable papers of L. Parnovski and R. Shterenberg [PSh1, PSh2, PS3], S. Morozov, L. Parnovski and R. Shterenberg [MPSh] and earlier papers by A. Sobolev [So1, So2]. I wanted to understand the approach of the authors and, combining their ideas with my own approach, generalize their results.

In these papers the complete asymptotic expansion of the integrated density of states $N(\lambda)$ for operators $\Delta + V$ was derived as $\lambda \to +\infty$; here $\Delta$ is a positive Laplacian and $V$ is a periodic or almost periodic potential (satisfying certain conditions). In [MPSh] more general operators were considered.

Further, in [PS3] the complete asymptotic expansion of $e(x, x, \lambda)$ was derived, where $e(x, y, \lambda)$ is the Schwartz kernel of the spectral projector.
This talk represents two works. The first of them is inspired by several remarkable papers of L. Parnovski and R. Shterenberg [PSh1, PSh2, PS3], S. Morozov, L. Parnovski and R. Shterenberg [MPSh] and earlier papers by A. Sobolev [So1, So2]. I wanted to understand the approach of the authors and, combining their ideas with my own approach, generalize their results.

In these papers the complete asymptotic expansion of the integrated density of states $N(\lambda)$ for operators $\Delta + V$ was derived as $\lambda \to +\infty$; here $\Delta$ is a positive Laplacian and $V$ is a periodic or almost periodic potential (satisfying certain conditions). In [MPSh] more general operators were considered.

Further, in [PS3] the complete asymptotic expansion of $e(x, x, \lambda)$ was derived, where $e(x, y, \lambda)$ is the Schwartz kernel of the spectral projector.

The second work is inspired by a paper [PSo] by L. Parnovski and A. Sobolev, in which a classical Bethe-Sommerfeld conjecture was proven for operators $\Delta^m + B(x, D)$ with operator $B$ of order $\leq 2m$. 
In the first work I borrowed from these papers Conditions (A)–(D) and the special gauge transformation and added the non-stationary semiclassical Schrödinger operator method—[Ivr1] and extremely long propagation of singularities. I believe that this is a simpler and more powerful approach. Also, in contrast to those papers I consider more general semiclassical asymptotics.
In the first work I borrowed from these papers Conditions (A)–(D) and the special gauge transformation and added the non-stationary semiclassical Schrödinger operator method—[Ivr1] and extremely long propagation of singularities. I believe that this is a simpler and more powerful approach. Also, in contrast to those papers I consider more general semiclassical asymptotics.

Similar ideas were used in the second work with the classification of resonant points borrowed from the same papers.
Consider a scalar self-adjoint $h$-pseudo-differential operator

$$A_h := A(x, hD) = A^0(hD) + \varepsilon B(x, hD), \quad (1)$$

in $\mathbb{R}^d$ where $A^0(\xi)$ is smooth and elliptic

$$|D_\xi^\beta A(\xi)| \leq c_\beta (|\xi| + 1)^m \quad \forall \xi \in \mathbb{R}^d \ \forall \beta, \quad (2)$$

and

$$A(\xi) \geq c_0 |\xi|^m - C_0 \quad \forall \xi \in \mathbb{R}^d \quad (3)$$
Consider a scalar self-adjoint $h$-pseudo-differential operator

$$A_h := A(x, hD) = A^0(hD) + \varepsilon B(x, hD),$$

in $\mathbb{R}^d$ where $A^0(\xi)$ is smooth and elliptic

$$|D^\beta_\xi A(\xi)| \leq c_\beta (|\xi| + 1)^m \quad \forall \xi \in \mathbb{R}^d \quad \forall \beta,$$

and

$$A(\xi) \geq c_0 |\xi|^m - C_0 \quad \forall \xi \in \mathbb{R}^d$$

and on energy level $\tau$ satisfies microhyperbolicity and strong convexity conditions:

$$|A^0(\xi) - \tau| + |\nabla_\xi A^0(\xi)| \geq \varepsilon_0$$

and

$$\pm \sum_{j,k} A^0_{\xi_j \xi_k}(\xi) \eta_j \eta_k \geq \varepsilon_0 |\eta|^2 \quad \forall \xi : A^0(\xi) = \tau \quad \forall \eta : \sum_j A^0_{\xi_j}(\xi) \eta_j = 0.$$
Meanwhile $B(x, \xi)$ is smooth

$$|D_x^\alpha D_\xi^\beta B(x, \xi)| \leq c_{\alpha \beta} (|\xi| + 1)^m \quad \forall \alpha, \beta$$

(6)

and almost periodic

$$B(x, \xi) = \sum_{\theta \in \Theta} b_\theta(\xi) e^{i\langle \theta, x \rangle}$$

(7)

with discrete (i.e. without any accumulation points) frequency set $\Theta$, $\varepsilon > 0$ is a small parameter: $h \leq \varepsilon \leq h^\gamma$ with $\gamma > 0$. 
Meanwhile $B(x, \xi)$ is smooth

$$|D_x^\alpha D_\xi^\beta B(x, \xi)| \leq c_{\alpha\beta}(|\xi| + 1)^m \quad \forall \alpha, \beta$$

and almost periodic

$$B(x, \xi) = \sum_{\theta \in \Theta} b_\theta(\xi) e^{i\langle\theta, x\rangle}$$

with discrete (i.e. without any accumulation points) frequency set $\Theta$, $\varepsilon > 0$ is a small parameter: $h \leq \varepsilon \leq h^{\varkappa}$ with $\varkappa > 0$.

Then it is semibounded from below. Let $e_{h, \varepsilon}(x, y, \lambda)$ be the Schwartz kernel of its spectral projector $E(\lambda) = \theta(\lambda - A)$. 
First main theorem

**Theorem 1.**

Let $A$ be a self-adjoint operator (1), where $A^0$ satisfies (2)–(5) and $B$ satisfies (6) and (7).
**Theorem 1.**

Let $A$ be a self-adjoint operator (1), where $A^0$ satisfies (2)–(5) and $B$ satisfies (6) and (7).

Let Conditions (A)–(D) below be fulfilled (they are fulfilled automatically if $B$ is periodic, i.e. $\Theta$ is a non-degenerate lattice). Then for $|\tau - \lambda| < \epsilon$, $\epsilon \leq h^\kappa$, $\kappa > 0$

$$e_{h,\epsilon}(x, x, \tau) \sim \sum_{n \geq 0} \kappa_n(x, \tau; \epsilon) h^{-d+n} \quad \text{as } h \to +0.$$  (8)
Corollary 2.

In the framework of Theorem 1

\[ N_{h,\varepsilon}(\tau) \sim \sum_{n \geq 0} \overline{K}_n(\tau; \varepsilon) h^{-d+n} \quad \text{as } h \to +0. \] (9)

Here

\[ N_h(\lambda) = M[e(x, x, \lambda)] := \lim_{\ell \to \infty} (\text{mes}(\ell X))^{-1} \int_{\ell X} e(x, x, \lambda) \, dx, \] (10)

where \( 0 \in X \) is an open domain in \( \mathbb{R}^d \). The latter expression in the cases we are interested in does not depend on \( X \) and is called Integrated Density of States.
Conditions (A)-(D)

Without any loss of generality we assume that $\Theta$ spans $\mathbb{R}^d$, contains 0 and is symmetric about 0.
Without any loss of generality we assume that $\Theta$ spans $\mathbb{R}^d$, contains 0 and is symmetric about 0.

**Condition (A).**

For each $\theta_1, \ldots, \theta_d \in \Theta$ either $\theta_1, \ldots, \theta_d$ are linearly independent over $\mathbb{R}$ or they linearly dependent over $\mathbb{Z}$. 
Assume also that

**Condition (B).**

For any arbitrarily large $L$ and for any sufficiently large real number $r$ there are a finite symmetric about 0 set $\Theta' := \Theta'(L,r) \subset (\Theta \cap B(0, r))$ (with $B(\xi, r)$ the ball of the radius $r$ and center $\xi$) and a “cut-off” coefficients $b'_{\theta} := b'_{\theta,(L,r)}$, such that

$$B' := B'_{(L,r)}(x, \xi) := \sum_{\theta \in \Theta'} b'_{\theta}(\xi) e^{i \langle \theta, x \rangle}$$

satisfies

$$\| D_{x}^{\alpha} D_{\xi}^{\beta} (B - B') \|_{L^{\infty}} \leq r^{-L} (|\xi| + 1)^{m} \quad \forall \alpha, \beta: |\alpha| \leq L, |\beta| \leq L.$$
Remark 1.

Then

$$|D^\beta_\xi b_\theta| = O(|\theta|^{-\infty}(|\xi| + 1)^m) \quad \text{as} \quad |\theta| \to \infty$$  \hspace{1cm} (13)

and

$$|D^\beta_\xi (b_\theta - b'_\theta)| = O(r^{-\infty}(|\xi| + 1)^m).$$  \hspace{1cm} (14)

Indeed, one suffices to observe that $b_\theta(\xi) = M(B(\mathbf{x}, \xi)e^{-i\langle \theta, \mathbf{x} \rangle})$ etc.
Remark 1.

Then

\[ |D_{\xi}^\beta b_\theta| = O(|\theta|^{-\infty}(|\xi| + 1)^m) \quad \text{as} \quad |\theta| \to \infty \]  \hspace{1cm} (13)

and

\[ |D_{\xi}^\beta (b_\theta - b'_\theta)| = O(r^{-\infty}(|\xi| + 1)^m). \]  \hspace{1cm} (14)

Indeed, one suffices to observe that \( b_\theta(\xi) = M(B(x, \xi)e^{-i\langle \theta, x \rangle}) \) etc.

On the other hand, under additional assumption

\[ \#\{\theta \in \Theta, |\theta| \leq r\} = O(r^p) \quad \text{as} \quad r \to \infty \]  \hspace{1cm} (15)

for some \( p \), (13) implies Condition (B) with \( \Theta'_{(L,r)} := \Theta \cap B(0, r) \). However we will need \( \Theta'_{(L,r)} \) in the next condition.
Remark 1 (Continued).

We need only to estimate the operator norm of the difference between $B(x, hD)$ and $B'(x, hD)$ (from $\mathcal{H}^m$ to $L^2$); therefore for differential operators we can weaken (12).
Remark 1 (Continued).

3. We need only to estimate the operator norm of the difference between \( B(x, hD) \) and \( B'(x, hD) \) (from \( \mathcal{H}^m \) to \( L^2 \)); therefore for differential operators we can weaken (12).

4. While Condition (B) is Condition B of [PS3], adopted to our case, Condition (A) and Conditions (C), (D) below are borrowed without any modifications (except changing notations).
We need to impose a the Diophantine condition on the frequencies of $B$. We need some definitions.
We need to impose a Diophantine condition on the frequencies of $B$. We need some definitions. We fix a natural number $K$ (the choice of $K$ will be determined later by how many terms in the asymptotic decomposition of $e(x, x, \lambda)$ we want to obtain) and consider $\Theta'_K$, which here and below denotes the algebraic sum of $K$ copies of $\Theta'$:

$$\Theta'_K := \sum_{1 \leq i \leq K} \Theta.$$  \hspace{1cm} (16)
We need to impose a Diophantine condition on the frequencies of $B$. We need some definitions. We fix a natural number $K$ (the choice of $K$ will be determined later by how many terms in the asymptotic decomposition of $e(x, x, \lambda)$ we want to obtain) and consider $\Theta'_K$, which here and below denotes the algebraic sum of $K$ copies of $\Theta'$:

$$\Theta'_K := \sum_{1 \leq i \leq K} \Theta.$$  

(16)

We say that $\mathcal{V}$ is a quasi-lattice subspace of dimension $q$, if $\mathcal{V}$ is a linear span of $q$ linear independent vectors $\theta_1, \ldots, \theta_q \in \Theta'_K \setminus 0$. Obviously, the zero space is a quasi-lattice subspace of dimension 0 and $\mathbb{R}^d$ is a quasi-lattice subspace of dimension $d$.

We denote by $\mathcal{V}_q$ the collection of all quasi-lattice subspaces of dimension $q$ and also $\mathcal{V} := \bigcup_{q \geq 0} \mathcal{V}_q$. 

Victor Ivrii (Math., Toronto)  Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture  June 3, 2020  13 / 63
Consider $\mathcal{V}, \mathcal{U} \in \mathcal{V}$. We say that these subspaces are strongly distinct, if neither of them is a subspace of the other one. Next, let $(\mathcal{V}, \mathcal{U}) \in [0, \pi/2]$ be the angle between them, i.e. the angle between $\mathcal{V} \ominus \mathcal{W}$ and $\mathcal{U} \ominus \mathcal{W}$, $\mathcal{W} = \mathcal{U} \cap \mathcal{V}$. This angle is positive iff $\mathcal{V}$ and $\mathcal{U}$ are strongly distinct.
Consider $\mathcal{V}, \mathcal{U} \in \mathcal{V}$. We say that these subspaces are **strongly distinct**, if neither of them is a subspace of the other one. Next, let $(\mathcal{V}, \mathcal{U}) \in [0, \pi/2]$ be the angle between them, i.e. the angle between $\mathcal{V} \ominus \mathcal{W}$ and $\mathcal{U} \ominus \mathcal{W}$, $\mathcal{W} = \mathcal{U} \cap \mathcal{V}$. This angle is positive iff $\mathcal{V}$ and $\mathcal{U}$ are strongly distinct.

**Condition (C).**

For each fixed $L$ and $K$ the sets $\Theta'_{(L,r)}$ satisfying (11) and (12) can be chosen in such a way that for sufficiently large $r$ we have

$$S(r) = S(\Theta'_K) := \inf_{\mathcal{V}, \mathcal{U} \in \mathcal{V}} \sin((\mathcal{V}, \mathcal{U})) \geq r^{-1}$$

(17)

and

$$R(r) := \inf_{\theta \in \Theta'_K \setminus 0} |\theta| \geq r^{-1},$$

(18)

where the implied constant (how large should $r$ be) depends on $L$ and $K$. 

Victor Ivrii (Math., Toronto)  Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture  June 3, 2020  14 / 63
Let $\mathcal{V}$ be the span of $\theta_1, \ldots, \theta_q \in \Theta'_K \setminus 0$. Then due to Condition (A) each element of the set $\Theta'_K \cap \mathcal{V}$ is a linear combination of $\theta_1, \ldots, \theta_q$ with rational coefficients. Since the set $\Theta'_K \cap \mathcal{V}$ is finite, this implies that the set $\Theta'_\infty \cap \mathcal{V}$ is discrete and is, therefore, a lattice in $\mathcal{V}$. We denote this lattice by $\Gamma(r; \mathcal{V})$. 
Let $V$ be the span of $\theta_1, \ldots, \theta_q \in \Theta'_K \setminus 0$. Then due to Condition (A) each element of the set $\Theta'_K \cap V$ is a linear combination of $\theta_1, \ldots, \theta_q$ with rational coefficients. Since the set $\Theta'_K \cap V$ is finite, this implies that the set $\Theta'_\infty \cap V$ is discrete and is, therefore, a lattice in $V$. We denote this lattice by $\Gamma(r; V)$.

Our final condition states that this lattice cannot be too dense.

**Condition (D).**

We can choose $\Theta'_{(L;r)}$, satisfying Conditions (B) and (C) in such a way that for sufficiently large $r$ and for each $V \in \mathcal{V}$, $V \neq \mathbb{R}^d$, we have

$$\text{vol}(V/\Gamma(r; V)) \geq r^{-1}. \quad (19)$$
**Remark 2.**

See Section 2 of [PS3] for discussion of these conditions.
Remark 2.

See Section 2 of [PS3] for discussion of these conditions.

1. In particular, if \( \Theta \) is a non-degenerate lattice, then Conditions (A)–(D) are fulfilled.
Remark 2.

See Section 2 of [PS3] for discussion of these conditions.

1. In particular, if \( \Theta \) is a non-degenerate lattice, then Conditions (A)–(D) are fulfilled.

2. Further, if \( \Theta \) is a finite set and Condition (A) is fulfilled, then \( \Theta_\infty := \bigcup_{K \geq 1} \Theta_K \) is a lattice and Conditions (B)–(D) are fulfilled.
Remark 2.

See Section 2 of [PS3] for discussion of these conditions.

1. In particular, if $\Theta$ is a non-degenerate lattice, then Conditions (A)–(D) are fulfilled.

2. Further, if $\Theta$ is a finite set and Condition (A) is fulfilled, then $\Theta_\infty := \bigcup_{K \geq 1} \Theta_K$ is a lattice and Conditions (B)–(D) are fulfilled.

3. Furthermore, the same is true, if $\Theta$ is an arithmetic sum of a finite set and a lattice.
Assume now that $B$ is periodic with respect to non-degenerate lattice $\Gamma$:

$$A(x + y, \xi) = A(x, \xi), \quad \forall x \in \mathbb{R}^n \quad \forall y \in \Gamma.$$  \hspace{1cm} (20)
Assume now that $B$ is periodic with respect to non-degenerate lattice $\Gamma$:

$$A(x + y, \xi) = A(x, \xi) \quad \forall x \in \mathbb{R}^n \quad \forall y \in \Gamma. \quad (20)$$

Let us denote by $\Gamma^*$ the dual lattice:

$$\gamma \in \Gamma^* \iff \langle \gamma, y \rangle \in 2\pi \mathbb{Z} \quad \forall y \in \Gamma; \quad (21)$$

since we use $\Gamma^*$ and its elements in the paper much more often, than $\Gamma$ and its elements, it is more convenient for us to reserve notation $\gamma$ for elements of $\Gamma^*$. 

Assume now that $B$ is periodic with respect to non-degenerate lattice $\Gamma$:

$$A(x + y, \xi) = A(x, \xi) \quad \forall x \in \mathbb{R}^n \quad \forall y \in \Gamma.$$  \hfill (20)

Let us denote by $\Gamma^*$ the dual lattice:

$$\gamma \in \Gamma^* \iff \langle \gamma, y \rangle \in 2\pi \mathbb{Z} \quad \forall y \in \Gamma;$$  \hfill (21)

since we use $\Gamma^*$ and it’s elements in the paper much more often, than $\Gamma$ and it’s elements, it is more convenient for us to reserve notation $\gamma$ for elements of $\Gamma^*$.

Also let $\mathcal{O} = \mathbb{R}^d / \Gamma$ and $\mathcal{O}^* = \mathbb{R}^d / \Gamma^*$ be fundamental domains; we identify them with domains in $\mathbb{R}^d$. 
It is well-known that $\text{Spec}(A_h)$ has a \textbf{band-structure}. Namely, consider in $L^2(\mathcal{O})$ operator $A_h(\xi) = A(x, hD)$ with the \textit{quasi-periodic boundary condition}:

$$u(x + y) = e^{i\langle y, \xi \rangle} u(x) \quad \forall x \in \mathcal{O} \quad \forall y \in \Gamma$$

(22)

with $\xi \in \mathcal{O}^*$; it is called a \textit{quasimomentum}. 
It is well-known that $\text{Spec}(A_h)$ has a **band-structure**. Namely, consider in $L^2(\mathcal{O})$ operator $A_h(\xi) = A(x, hD)$ with the **quasi-periodic boundary condition**:

$$u(x + y) = e^{i\langle y, \xi \rangle} u(x) \quad \forall x \in \mathcal{O} \quad \forall y \in \Gamma$$

(22)

with $\xi \in \mathcal{O}^*$; it is called a **quasimomentum**. Then $\text{Spec}(A_h(\xi))$ is discrete

$$\text{Spec}(A_h(\xi)) = \bigcup_n \lambda_{n,h}(\xi)$$

(23)

and depends on $\xi$, continuously.

---

Victor Ivrii (Math., Toronto)  Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture  June 3, 2020  18 / 63
It is well-known that $\text{Spec}(A_h)$ has a band-structure. Namely, consider in $L^2(\mathcal{O})$ operator $A_h(\xi) = A(x, hD)$ with the quasi-periodic boundary condition:

$$u(x + y) = e^{i\langle y, \xi \rangle} u(x) \quad \forall x \in \mathcal{O} \quad \forall y \in \Gamma$$

(22)

with $\xi \in \mathcal{O}^*$; it is called a quasimomentum. Then $\text{Spec}(A_h(\xi))$ is discrete

$$\text{Spec}(A_h(\xi)) = \bigcup_{\xi} \lambda_{n,h}(\xi)$$

(23)

and depends on $\xi$ continuously. Further,

$$\text{Spec}(A_h) = \bigcup_{\xi \in \mathcal{O}^*} \text{Spec}(A_h(\xi)) =: \bigcup_n \Lambda_{n,h},$$

(24)

with the spectral bands $\Lambda_{n,h} := \bigcup_{\xi \in \mathcal{O}^*} \{\lambda_{n,h}(\xi)\}$. 
One can prove that the width of the spectral band near energy level $\tau$ is $O(h)$. Spectral bands could overlap but they also could leave uncovered intervals, called spectral gaps.
One can prove that the width of the spectral band near energy level $\tau$ is $O(h)$. Spectral bands could overlap but they also could leave uncovered intervals, called \textit{spectral gaps}.

It follows from Theorem 1 that in our assumptions the width of the spectral gaps near energy level $\tau$ is $O(h^\infty)$. Bethe-Sommerfeld conjecture in the semiclassical settings claims that there are no spectral gaps near energy level $\tau$ (in the corresponding assumptions, which include $d \geq 2$).
Second main theorem

Theorem 3.

Let $d \geq 2$ and let operator $A_h$ be given by (1) with $\varepsilon = O(h^\kappa)$ with arbitrary $\kappa > 0$ and with $A_h^0 = A^0(hD)$ satisfying (2)–(5) and $B(x, \xi)$ satisfying (7) and (8) with $\Theta = \Gamma^*$ where $\Gamma$ is a non-degenerate lattice of periods.
Second main theorem

**Theorem 3.**

Let $d \geq 2$ and let operator $A_h$ be given by (1) with $\varepsilon = O(h^\kappa)$ with arbitrary $\kappa > 0$ and with $A^0_h = A^0(hD)$ satisfying (2)–(5) and $B(\chi, \xi)$ satisfying (7) and (8) with $\Theta = \Gamma^*$ where $\Gamma$ is a non-degenerate lattice of periods. Furthermore, assume that there exists $\xi \in \Sigma_\tau$ such that

**Condition (E).**

For every $\eta \in \Sigma_\tau$, $\eta \neq \xi$, such that $\nabla_\eta A^0(\eta)$ is parallel to $\nabla_\xi A^0(\xi)$ (we call $\eta$ antipodal point) $\Sigma_\tau$, intersected with some vicinity of $\eta$ and shifted by $(\xi - \eta)$, coincides in the vicinity of $\xi$ with $\{\zeta: \zeta_k = g(\zeta_\hat{k})\}$ and $\Sigma_\tau$ coincides in the vicinity of $\xi$ with $\{\zeta: \zeta_k = f(\zeta_\hat{k})\}$ and $\nabla^\alpha(f - g)(0) \neq 0$ for some $\alpha: |\alpha| = 2$. 
Second main theorem

**Theorem 3.**

Let $d \geq 2$ and let operator $A_h$ be given by (1) with $\varepsilon = O(h^\kappa)$ with arbitrary $\kappa > 0$ and with $A_h^0 = A^0(hD)$ satisfying (2)–(5) and $B(\chi, \xi)$ satisfying (7) and (8) with $\Theta = \Gamma^*$ where $\Gamma$ is a non-degenerate lattice of periods. Furthermore, assume that there exists $\xi \in \Sigma_\tau$ such that

**Condition (E).**

For every $\eta \in \Sigma_\tau$, $\eta \neq \xi$, such that $\nabla_\eta A^0(\eta)$ is parallel to $\nabla_\xi A^0(\xi)$ (we call $\eta$ antipodal point) $\Sigma_\tau$, intersected with some vicinity of $\eta$ and shifted by $(\xi - \eta)$, coincides in the vicinity of $\xi$ with $\{\zeta: \zeta_k = g(\hat{\zeta}_k)\}$ and $\Sigma_\tau$ coincides in the vicinity of $\xi$ with $\{\zeta: \zeta_k = f(\hat{\zeta}_k)\}$ and $\nabla^\alpha(f - g)(0) \neq 0$ for some $\alpha: |\alpha| = 2$.

Then $\text{Spec}(A_h) \supset [\tau - \epsilon, \tau + \epsilon]$ for sufficiently small $\epsilon > 0$. 
Remark 3.

1 If $\Sigma_\tau$ is strongly convex and connected then for every $\xi \in \Sigma_\tau$ there exists exactly one antipodal point $\eta \in \Sigma_\tau$ and $\nabla_\eta A^0(\eta) \parallel \nabla_\xi A^0(\xi)$ and Condition (E) is fulfilled.
Remark 3.

1. If $\Sigma_\tau$ is strongly convex and connected then for every $\xi \in \Sigma_\tau$ there exists exactly one antipodal point $\eta \in \Sigma_\tau$ and $\nabla_\eta A^0(\eta) \parallel \nabla_\xi A^0(\xi)$ and Condition (E) is fulfilled.

2. If $\Sigma_\tau$ is strongly convex and consists of $p$ connected components, then the set $Z(\xi) = \{\eta \in \Sigma_\tau, \eta \neq \xi : \nabla_\eta A^0(\eta) \parallel \nabla_\xi A^0(\xi)\}$ contains exactly $2p - 1$ elements, and for $p$ antipodal points $\nabla_\eta A^0(\eta) \parallel \nabla_\xi A^0(\xi)$ and Condition (E) is fulfilled for sure, while for $(p - 1)$ of them $\nabla_\eta A^0(\eta) \parallel \nabla_\xi A^0(\xi)$ and Condition (E) needs to be satisfied.
One needs to understand, how gaps could appear, why they appear if $d = 1$ and why it is not the case if $d \geq 2$. Note that for $A^0_h$ one can use instead of $\lambda_{n,h}(\xi)$ functions $\lambda^0_{\gamma,h}(\xi) = A^0((h(\gamma + \xi))$. 

Observe that $\lambda_{n,h}(\xi)$ can be identified with some $\lambda^0_{\gamma,h}(\xi)$ only locally, if $\lambda^0_{\gamma,h}(\xi)$ is sufficiently different from $\lambda^0_{\gamma,h}(\xi)$ for any $\gamma \neq \gamma'$. Indeed, in the basis of eigenfunctions of $A^0_\xi(x)$ (consisting of $\exp(i\langle x, \gamma + \xi \rangle)$ perturbation $\epsilon_B(x,hD)$ can contain out-of-diagonal elements $\epsilon_b_{\gamma - \gamma'}(\xi)$ and such identification is possible only if $|\lambda^0_{\gamma,h}(\xi) - \lambda^0_{\gamma,h}(\xi)|$ is larger than the size of such element. If $d = 1$, $A^0_\xi(\xi) = \xi^2$ and $\epsilon \ll h$ and $\tau \approx 1$, it can happen only if $\gamma'$ coincides with $-\gamma$ or with one of two adjacent points in $\Gamma^*$ and $|\xi - 1/2(\gamma + \gamma')| = O(\epsilon h)$.

This exclude from possible values of either $\lambda^0_{\gamma,h}(\xi)$ or $\lambda^0_{\gamma',h}(\xi)$ the interval of the width $O(h)$ and on such interval can happen (and really happens for a generic perturbation) the realignment:
One needs to understand, how gaps could appear, why they appear if $d = 1$ and why it is not the case if $d \geq 2$. Note that for $A^0_h$ one can use instead of $\lambda_{n,h}(\xi)$ functions $\lambda^0_{\gamma,h}(\xi) = A^0((h(\gamma + \xi))$.

Observe that $\lambda_n(\xi)$ can be identified with some $\lambda^0_{\gamma}(\xi)$ only locally, if $\lambda^0_{\gamma}(\xi)$ is sufficiently different from $\lambda^0_{\gamma'}(\xi)$ for any $\gamma' \neq \gamma$. 
One needs to understand, how gaps could appear, why they appear if \( d = 1 \) and why it is not the case if \( d \geq 2 \). Note that for \( A^0_\hbar \) one can use instead of \( \lambda_{n, \hbar}(\xi) \) functions \( \lambda^0_{\gamma, \hbar}(\xi) = A^0((h(\gamma + \xi)) \).

Observe that \( \lambda_n(\xi) \) can be identified with some \( \lambda^0_{\gamma}(\xi) \) only locally, if \( \lambda^0_{\gamma}(\xi) \) is sufficiently different from \( \lambda^0_{\gamma'}(\xi) \) for any \( \gamma' \neq \gamma \).

Indeed, in the basis of eigenfunctions of \( A^0_\xi(hD) \) (consisting of \( \exp(i\langle x, \gamma + \xi \rangle) \)) perturbation \( \varepsilon B(x, hD) \) can contain out-of-diagonal elements \( \varepsilon b_{\gamma - \gamma'}(\xi) \) and such identification is possible only if \( |\lambda^0_{\gamma}(\xi) - \lambda^0_{\gamma'}(\xi)| \) is larger than the size of such element.
One needs to understand, how gaps could appear, why they appear if $d = 1$ and why it is not the case if $d \geq 2$. Note that for $A^0_h$ one can use instead of $\lambda_{n,h}(\xi)$ functions $\lambda^0_{\gamma,h}(\xi) = A^0((h(\gamma + \xi))$.

Observe that $\lambda_n(\xi)$ can be identified with some $\lambda^0_\gamma(\xi)$ only locally, if $\lambda^0_\gamma(\xi)$ is sufficiently different from $\lambda^0_{\gamma'}(\xi)$ for any $\gamma' \neq \gamma$.

Indeed, in the basis of eigenfunctions of $A^0_\xi(hD)$ (consisting of $\exp(i\langle x, \gamma + \xi \rangle)$) perturbation $\varepsilon B(x, hD)$ can contain out-of-diagonal elements $\varepsilon b_{\gamma - \gamma'}(\xi)$ and such identification is possible only if $|\lambda^0_\gamma(\xi) - \lambda^0_{\gamma'}(\xi)|$ is larger than the size of such element.

If $d = 1$, $A^0(\xi) = \xi^2$ and $\varepsilon \ll h$ and $\tau \asymp 1$, it can happen only if $\gamma'$ coincides with $-\gamma$ or with one of two adjacent points in $\Gamma^*$ and $|\xi - \frac{1}{2}(\gamma + \gamma')| = O(\varepsilon h^\infty)$. 
One needs to understand, how gaps could appear, why they appear if $d = 1$ and why it is not the case if $d \geq 2$. Note that for $A^0_h$ one can use instead of $\lambda_{n,h}(\xi)$ functions $\lambda^0_{\gamma,h}(\xi) = A^0((h(\gamma + \xi))$.

Observe that $\lambda_n(\xi)$ can be identified with some $\lambda^0_\gamma(\xi)$ only locally, if $\lambda^0_\gamma(\xi)$ is sufficiently different from $\lambda^0_{\gamma'}(\xi)$ for any $\gamma' \neq \gamma$.

Indeed, in the basis of eigenfunctions of $A^0_\xi(hD)$ (consisting of $\exp(i\langle x, \gamma + \xi \rangle)$) perturbation $\varepsilon B(x, hD)$ can contain out-of-diagonal elements $\varepsilon b_{\gamma - \gamma'}(\xi)$ and such identification is possible only if $|\lambda^0_\gamma(\xi) - \lambda^0_{\gamma'}(\xi)|$ is larger than the size of such element.

If $d = 1$, $A^0(\xi) = \xi^2$ and $\varepsilon \ll h$ and $\tau \asymp 1$, it can happen only if $\gamma'$ coincides with $-\gamma$ or with one of two adjacent points in $\Gamma^*$ and $|\xi - \frac{1}{2}(\gamma + \gamma')| = O(\varepsilon h^\infty)$. This exclude from possible values of either $\lambda^0_\gamma(\xi)$ or $\lambda^0_{\gamma'}(\xi)$ the interval of the width $O(h^\infty)$ and on such interval can happen (and really happens for a generic perturbation) the realignment:
Main results

Bethe-Sommerfeld Conjecture

If $d \geq 2$ the picture becomes more complicated: there are much more opportunities for $\lambda_0^\gamma(\xi)$ and $\lambda_0^{\gamma'}(\xi)$ to become close, even if $\gamma$ and $\gamma'$ are not that far away; on the other hand, there is much more opportunities for us to select $\zeta = h(\gamma + \xi) \in \Sigma$ and then to adjust $\xi$ so that $\zeta = h(\gamma + \xi)$ remains on $\Sigma$ but $\eta = h(\gamma' + \xi)$ moves away from $\Sigma$ sufficiently far away and then tune-up $\xi$ once again so that $\tau \in \text{Spec}(A_h(\xi))$. 

Victor Ivrii (Math., Toronto)
If $d \geq 2$ the picture becomes more complicated: there are much more opportunities for $\lambda^0_\gamma(\xi)$ and $\lambda^0_{\gamma'}(\xi)$ to become close, even if $\gamma$ and $\gamma'$ are not that far away; on the other hand, there is a much more opportunities for us to select $\xi = h(\gamma + \xi) \in \Sigma_\tau$ and then to adjust $\xi$, so that $\xi = h(\gamma + \xi)$ remains on $\Sigma_\tau$ but $\eta = h(\gamma' + \xi)$ moves away from $\Sigma_\tau$ sufficiently far away and then tune-up $\xi$, once again so that $\tau \in \text{Spec}(A_h(\xi))$. 
Theorem 3 follows from

**Theorem 4.**

*In the framework of Theorem 3 there exist \( n \) and \( \xi^* \) such that \( \lambda_n(\xi^*) = \tau \) and \( \lambda_n(\xi) \) covers interval \([\tau - \nu h, \tau + \nu h]\) when \( \xi \) runs ball \( B(\xi^*, \nu) \) while \( |\lambda_m(\xi) - \tau| \geq \epsilon \nu h \) for all \( m \neq n \) and \( \xi \in B(\xi^*, \nu) \).*
Theorem 3 follows from

**Theorem 4.**

*In the framework of Theorem 3 there exist $n$ and $\xi^*$ such that $\lambda_n(\xi^*) = \tau$ and $\lambda_n(\xi)$ covers interval $[\tau - \nu h, \tau + \nu h]$ when $\xi$ runs ball $B(\xi^*, \nu)$ while $|\lambda_m(\xi) - \tau| \geq \epsilon \nu h$ for all $m \neq n$ and $\xi \in B(\xi^*, \nu)$. Here*

$$
u = \epsilon \begin{cases} h^{(d-1)^2} \min(1, \epsilon^{-3(d-1)/2} h^{(d-1)+\sigma}) & d \geq 3, \\ h \min(|\log h|^{-1}, \epsilon^{-3/2} h^{\sigma}) & d = 2 \end{cases} \quad (25)$$

*with arbitrarily small exponent $\sigma > 0$.***
Despite Theorems 1 and 4 are of very different nature, their proofs have a common element: reduction of operator to a canonical form in the vicinity of $\Sigma_\tau$ ($0 < \nu$ is very small):

$$\Omega_\tau := \{ \xi : |A^0(\xi) - \tau| \leq C \varepsilon h^{-\nu} \}$$

(26)
Recall that $\Theta' \subset (\Theta \cap B(0, r))$ and we select $r = h^{-\nu}$. 
Recall that $\Theta' \subset (\Theta \cap B(0, r))$ and we select $r = h^{-\nu}$. We call point $\xi$ non-resonant if

$$|\langle \nabla_\xi A^0(\xi), \theta \rangle| \geq \rho \quad \forall \theta \in \Theta'_K \setminus 0$$

(27)

with $\rho \in [\varepsilon^{1/2} h^{-\delta}, h^{\delta}]$ with arbitrarily small $\delta > 0$. 
Recall that $\Theta' \subset (\Theta \cap B(0, r))$ and we select $r = h^{-\nu}$. We call point $\xi$ non-resonant if

$$|\langle \nabla_\xi A^0(\xi), \theta \rangle| \geq \rho \quad \forall \theta \in \Theta'_K \setminus 0$$

(27)

with $\rho \in [\varepsilon^{1/2} h^{-\delta}, h^\delta]$ with arbitrarily small $\delta > 0$. Otherwise we call it resonant. More precisely

$$\Lambda := \bigcup_{\theta \in \Theta'_K \setminus 0} \Lambda(\theta),$$

(28)

where $\Lambda(\theta)$ is the set of $\xi$, violating (27) for given $\theta \in \Theta'_K \setminus 0$. 

Victor Ivrii (Math., Toronto) Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture June 3, 2020
Recall that $\Theta' \subset (\Theta \cap B(0, r))$ and we select $r = h^{-\nu}$. We call point $\xi$ non-resonant if

$$\left| \langle \nabla_\xi A^0(\xi), \theta \rangle \right| \geq \rho \quad \forall \theta \in \Theta'_K \setminus 0$$

(27)

with $\rho \in [\varepsilon^{1/2} h^{-\delta}, h^{\delta}]$ with arbitrarily small $\delta > 0$. Otherwise we call it resonant. More precisely

$$\Lambda := \bigcup_{\theta \in \Theta'_K \setminus 0} \Lambda(\theta),$$

(28)

where $\Lambda(\theta)$ is the set of $\xi$, violating (27) for given $\theta \in \Theta'_K \setminus 0$.

It follows from the microhyperbolicity and strong convexity assumptions (4) and (5) that

**Proposition 4.**

$\mu_\tau$-measure ($\mu_\tau = d\xi : dA^0(\xi)$ is a natural measure on $\Sigma_\tau$) of $\Lambda \cap \Sigma_\tau$, does not exceed $C_0 r^{d-1} \rho$ and Euclidean measure of $\Lambda \cap \{\xi : |A^0(\xi) - \tau| \leq \varsigma\}$ does not exceed $C_0 r^{d-1} \rho \varsigma$.
Classification of resonant points

We start from the case $d = 2$. Then we have only one kind of resonant points $\Xi_1 = \Lambda$. 
We start from the case $d = 2$. Then we have only one kind of resonant points $\Xi_1 = \Lambda$.

If $d \geq 3$ then there are $(d - 1)$ kinds of resonant points.
Classification of resonant points

We start from the case $d = 2$. Then we have only one kind of resonant points $\Xi_1 = \Lambda$.

If $d \geq 3$ then there are $(d - 1)$ kinds of resonant points.

First, following [PSo] consider lattice spaces $\mathcal{V}$ spanned by $n$ linearly independent elements $\theta_1, \ldots, \theta_n \in \Gamma^* \cap B(0, r)$. Recall that $\mathcal{V}_n$ is the set of all such spaces.
Classification of resonant points

We start from the case $d = 2$. Then we have only one kind of resonant points $\Xi_1 = \Lambda$.

If $d \geq 3$ then there are $(d - 1)$ kinds of resonant points.

First, following [PSo] consider lattice spaces $\mathcal{V}$ spanned by $n$ linearly independent elements $\theta_1, \ldots, \theta_n \in \Gamma^* \cap B(0, r)$. Recall that $\mathcal{V}_n$ is the set of all such spaces.

Fix $0 < \delta_1 < \ldots < \delta_n$ arbitrarily small and for $\mathcal{V} \in \mathcal{V}_n$ let us introduce

$$\Lambda(\mathcal{V}, \rho_n) := \{\xi \in \Omega_\tau : |\langle \nabla_\xi A^0(\xi), \theta \rangle| \leq \rho_n|\theta| \quad \forall \theta \in \mathcal{V}\}$$

(29)

with $\rho_n = \varepsilon^\frac{1}{2} h^{-\delta_n}$. 
We define $\Xi_n$ by induction. First, $\Xi_d = \emptyset$. Assume that we defined $\Xi_d, \ldots, \Xi_{n+1}$. Then we define

$$
\Xi_n := \bigcup_{\mathcal{V} \in \mathcal{V}_n, \xi \in \Lambda(\mathcal{V}) \cap \Omega_\tau} (\xi + \mathcal{V}) \cap \Omega_\tau.
$$

(30)
We define $\Xi_n$ by induction. First, $\Xi_d = \emptyset$. Assume that we defined $\Xi_d, \ldots, \Xi_{n+1}$. Then we define

$$\Xi_n := \bigcup_{\mathfrak{V} \in \mathfrak{V}_n, \xi \in \Lambda(\mathfrak{V}) \cap \Omega_\tau} (\xi + \mathfrak{V}) \cap \Omega_\tau.$$  

(30)

**Proposition 5.**

*Let $\nu > 0$ in the definitions of $\Theta'$ and $\Omega_\tau$ be sufficiently small. Then for sufficiently small $h$*
We define $\Xi_n$ by induction. First, $\Xi_d = \emptyset$. Assume that we defined $\Xi_d, \ldots, \Xi_{n+1}$. Then we define

$$\Xi_n := \bigcup_{\mathcal{V} \in \mathcal{V}_n, \xi \in \Lambda(\mathcal{V}) \cap \Omega_\tau} (\xi + \mathcal{V}) \cap \Omega_\tau.$$  

(30)

**Proposition 5.**

Let $\nu > 0$ in the definitions of $\Theta'$ and $\Omega_\tau$ be sufficiently small. Then for sufficiently small $h$

1. $\Xi_n \subset \bigcup_{\mathcal{V} \in \mathcal{V}_n} \Lambda(\mathcal{V}, 2\rho_n)$. 

Victor Ivrii (Math., Toronto) Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture June 3, 2020 28 / 63
We define $\Xi_n$ by induction. First, $\Xi_d = \emptyset$. Assume that we defined $\Xi_d, \ldots, \Xi_{n+1}$. Then we define

$$\Xi_n := \bigcup_{\mathcal{V} \in \mathcal{V}_n, \xi \in \Lambda(\mathcal{V}) \cap \Omega_\tau} (\xi + \mathcal{V}) \cap \Omega_\tau. \quad (30)$$

**Proposition 5.**

Let $\nu > 0$ in the definitions of $\Theta'$ and $\Omega_\tau$ be sufficiently small. Then for sufficiently small $h$

1. $\Xi_n \subset \bigcup_{\mathcal{V} \in \mathcal{V}_n} \Lambda(\mathcal{V}, 2\rho_n)$.

2. If $\xi \notin \Xi_{n+1}$ and $\xi \in \xi' + \mathcal{V}$, $\xi \in \xi'' + \mathcal{W}$ for $\xi' \in \Lambda(\mathcal{V})$, $\xi'' \in \Lambda(\mathcal{W})$ with $\mathcal{V}, \mathcal{W} \in \mathcal{V}_n$, then $\mathcal{V} = \mathcal{W}$. 
We define $\Xi_n$ by induction. First, $\Xi_d = \emptyset$. Assume that we defined $\Xi_d, \ldots, \Xi_{n+1}$. Then we define

$$\Xi_n := \bigcup_{\mathcal{V} \in \mathcal{V}_n, \xi \in \Lambda(\mathcal{V}) \cap \Omega_\tau} (\xi + \mathcal{V}) \cap \Omega_\tau.$$ (30)

**Proposition 5.**

Let $\nu > 0$ in the definitions of $\Theta'$ and $\Omega_\tau$ be sufficiently small. Then for sufficiently small $h$

1. $\Xi_n \subset \bigcup_{\mathcal{V} \in \mathcal{V}} \Lambda(\mathcal{V}, 2 \rho_n)$.

2. If $\xi \notin \Xi_{n+1}$ and $\xi \in \xi' + \mathcal{V}$, $\xi \in \xi'' + \mathcal{W}$ for $\xi' \in \Lambda(\mathcal{V})$, $\xi'' \in \Lambda(\mathcal{W})$ with $\mathcal{V}, \mathcal{W} \in \mathcal{V}_n$, then $\mathcal{V} = \mathcal{W}$.

3. For each $\xi \in \Xi_n \setminus \Xi_{n+1}$ is defined just one $\mathcal{V} = \mathcal{V}(\xi)$ such that $\xi \in \xi' + \mathcal{V}$ for some $\xi' \in \Lambda(\mathcal{V})$. 
We slightly change definition of $\Xi_n$: $\xi = h(\gamma + \xi) \in \Xi_{n,\text{new}}$ iff $h\gamma \in \Xi_n$. From now on $\Xi_n := \Xi_{n,\text{new}}$. 
We slightly change definition of $\Xi_n$: $\xi = h(\gamma + \xi) \in \Xi_{n,\text{new}}$ iff $h\gamma \in \Xi_n$. From now on $\Xi_n := \Xi_{n,\text{new}}$.

Consider $\xi', \xi'' \in \Xi_n \setminus \Xi_{n+1}$. We say that $\xi' \equiv \xi''$ if there exists $\xi \in \mathcal{V}$, $\mathcal{V} \in \mathcal{V}$ such that $\xi', \xi'' \in \xi + \mathcal{V}$ and if $\xi' - \xi'' \in \Gamma$. 
We slightly change definition of $\Xi_n$: $\xi = h(\gamma + \xi) \in \Xi_{n,\text{new}}$ iff $h\gamma \in \Xi_n$. From now on $\Xi_n := \Xi_{n,\text{new}}$.

Consider $\xi', \xi'' \in \Xi_n \setminus \Xi_{n+1}$. We say that $\xi' \equiv \xi''$ if there exists $\xi \in \mathcal{V}$, $\mathcal{V} \in \mathcal{V}$ such that $\xi', \xi'' \in \xi + \mathcal{V}$ and if $\xi' - \xi'' \in \Gamma$.

This relation is reflexive, symmetric and transitive.
We slightly change definition of $\Xi_n$: $\xi = h(\gamma + \xi) \in \Xi_{n,\text{new}}$ iff $h\gamma \in \Xi_n$.

From now on $\Xi_n := \Xi_{n,\text{new}}$.

Consider $\xi', \xi'' \in \Xi_n \setminus \Xi_{n+1}$. We say that $\xi' \cong \xi''$ if there exists $\xi \in \mathcal{V}$, $\mathcal{V} \in \mathcal{V}$ such that $\xi', \xi'' \in \xi + \mathcal{V}$ and if $\xi' - \xi'' \in \Gamma$.

This relation is reflexive, symmetric and transitive.

For $\xi \in \Xi_n$ we define

$$\mathcal{X}(\xi) = \{\xi': \xi' \cong \xi\}. \quad (31)$$

Then

$$\text{diam}(\mathcal{X}(\xi)) \leq C\rho_{d-1}. \quad (32)$$
Theorem 5.

Let assumptions (4) and (5) be fulfilled.
Theorem 5.

Let assumptions (4) and (5) be fulfilled.

Then there exists a pseudodifferential operator \( P = P(x, hD) \) such that

\[
(e^{-i\varepsilon h^{-1}P} A e^{i\varepsilon h^{-1}P} - \mathcal{A}) Q \equiv 0
\]

with

\[
\mathcal{A} = A^0(hD) + \varepsilon B'(hD) + \varepsilon B''(x, hD)
\]

modulo operator from \( \mathcal{H}^m \) to \( \mathcal{L}^2 \) with the operator norm \( O(h^M) \) with \( M \) arbitrarily large and \( K = K(M, d, \delta) \) in the definition of non-resonant point provided \( Q = Q(hD) \) has a symbol, supported in \( \{\xi : |A^0(\xi) - \tau| \leq 2C\varepsilon h^{-\nu}\} \).
Finally, reduction

Theorem 5 (continuation).

Here $P(x, hD)$, $B'(hD)$ and $B''(x, hD)$ are operator with Weyl symbols of the same form (7) albeit such that

$$|D_\xi^\alpha D_x^\beta P| \leq c_{\alpha\beta} \rho^{-1-|\alpha|} \quad \forall \alpha, \beta,$$

$$|D_\xi^\alpha D_x^\beta B''| \leq c'_{\alpha\beta} \rho^{-|\alpha|} \quad \forall \alpha, \beta,$$

and symbol of $B'$ also satisfies (36).
Theorem 5 (continuation).

Here $P(x, hD)$, $B'(hD)$ and $B''(x, hD)$ are operator with Weyl symbols of the same form (7) albeit such that

\[ |D_\xi^\alpha D_x^\beta P| \leq c_{\alpha\beta} \rho^{-1-|\alpha|} \quad \forall \alpha, \beta, \tag{35} \]

\[ |D_\xi^\alpha D_x^\beta B''| \leq c'_{\alpha\beta} \rho^{-|\alpha|} \quad \forall \alpha, \beta, \tag{36} \]

and symbol of $B'$ also satisfies (36).

Further,

\[ \xi \not\in \Lambda(\theta) \implies b''_{\theta}(\xi) = 0. \tag{37} \]

and $B'(\xi)$ coincides with $b_0(\xi)$ modulo $O(\varepsilon \rho^{-2})$.  

Reduction of operator

Finally, reduction

**Theorem 5 (continuation).**

Here \( P(x, hD) \), \( B'(hD) \) and \( B''(x, hD) \) are operator with Weyl symbols of the same form (7) albeit such that

\[
|D_\xi^\alpha D_x^\beta P| \leq c_{\alpha\beta}\rho^{1-|\alpha|} \quad \forall \alpha, \beta, \tag{35}
\]

\[
|D_\xi^\alpha D_x^\beta B''| \leq c'_{\alpha\beta}\rho^{-|\alpha|} \quad \forall \alpha, \beta, \tag{36}
\]

and symbol of \( B' \) also satisfies (36).

2. **Further,**

\[\xi \notin \Lambda(\theta) \implies b''_\theta(\xi) = 0.\] \( \tag{37} \)

and \( B'(\xi) \) coincides with \( b_0(\xi) \) modulo \( O(\varepsilon\rho^{-2}) \).

3. **Finally, if \( B \) is periodic, so are \( P, B', B'' \).**
**Theorem 5 (continuation).**

Here $P(x, hD)$, $B'(hD)$ and $B''(x, hD)$ are operator with Weyl symbols of the same form (7) albeit such that

\[
|D_\xi^\alpha D_x^\beta P| \leq c_{\alpha \beta} \rho^{-1-|\alpha|} \quad \forall \alpha, \beta,
\]

\[
|D_\xi^\alpha D_x^\beta B''| \leq c'_{\alpha \beta} \rho^{-|\alpha|} \quad \forall \alpha, \beta,
\]

and symbol of $B'$ also satisfies (36).

1. **Further,**

   \[
   \xi \notin \Lambda(\theta) \implies b''_\theta(\xi) = 0.
   \]

   and $B'(\xi)$ coincides with $b_0(\xi)$ modulo $O(\varepsilon \rho^{-2})$.

2. **Finally,** if $B$ is periodic, so are $P$, $B'$, $B''$.

In what follows

\[
\mathcal{A}^0(hD) := A^0(hD) + \varepsilon B'(hD) \quad \text{and} \quad \mathcal{B} := B''(x, hD).
\]
Gauge transformation and proof of Theorem 5

First of all, replace $B$ by $B' := B'_{(L,r)}$ from Condition (B) with $r = h^{-\nu}$, arbitrarily small $\nu > 0$ and $L = 3M/\nu$. 

\[ A \rightarrow e^{-i \frac{1}{\nu} P A e^{i \frac{1}{\nu} P}} \]

Consider now the "gauge" transformation

\[ e^{-i \frac{1}{\nu} P A e^{i \frac{1}{\nu} P}} \]

Observe that

\[ e^{-i \frac{1}{\nu} P A e^{i \frac{1}{\nu} P}} = A - i \frac{1}{\nu} [P, A] + \sum_{2 \leq n \leq \infty} \frac{1}{n!} - i \frac{1}{\nu} P A^n + \int_0^1 K^{-1} \left( -i \frac{1}{\nu} P A^s \right) ds \]

where $A_d P (A) = A$ and $A_{d+1} P (A) = [P, A_d P (A)]$ for $d = 0, 1, \ldots$. 

Victor Ivrii (Math., Toronto)
First of all, replace $B$ by $B' := B'_{(L,r)}$ from Condition (B) with $r = h^{-\nu}$, arbitrarily small $\nu > 0$ and $L = 3M/\nu$.

Consider now the “gauge” transformation $A \mapsto e^{-i\varepsilon h^{-1}P}Ae^{i\varepsilon h^{-1}P}$ with $h$-pseudodifferential operator $P$. 
Gauge transformation and proof of Theorem 5

First of all, replace $B$ by $B' := B'_{(L,r)}$ from Condition (B) with $r = h^{-\nu}$, arbitrarily small $\nu > 0$ and $L = 3M/\nu$.

Consider now the “gauge” transformation $A \mapsto e^{-i\varepsilon h^{-1}P}A e^{i\varepsilon h^{-1}P}$ with $h$-pseudodifferential operator $P$. Observe that

$$e^{-i\varepsilon h^{-1}P}A e^{i\varepsilon h^{-1}P} = A - i\varepsilon h^{-1}[P, A] + \sum_{2 \leq n \leq K-1} \frac{1}{n!}(-i\varepsilon h^{-1})^n \text{Ad}^n_P(A)$$

$$+ \int_0^1 \frac{1}{(K-1)!} (1 - s)^{K-1}(-i\varepsilon h^{-1})^K e^{-i\varepsilon h^{-1}sP} \text{Ad}^K_P(A) e^{i\varepsilon h^{-1}sP} ds, \quad (39)$$

where $\text{Ad}^0_P(A) = A$ and $\text{Ad}_P^{n+1}(A) = [P, \text{Ad}_P^n(A)]$ for $n = 0, 1, \ldots$. 
Then formally we can compensate $\varepsilon B$, taking

$$P = \sum_{\theta} i\hbar (A^0(\xi + \theta \hbar/2) - A^0(\xi - \theta \hbar/2))^{-1} b_\theta(\xi) e^{i(\theta, x)}, \quad (40)$$

so that

$$i\hbar^{-1} [P, A^0] = B \implies i\hbar^{-1} [P, A] = B + i\varepsilon h^{-1} [P, B]. \quad (41)$$
Then formally we can compensate $\varepsilon B$, taking

$$
P = \sum_{\theta} i\hbar (A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2))^{-1} b_\theta(\xi) e^{i\langle\theta, x\rangle}, \quad (40)$$

so that

$$
i\hbar^{-1} [P, A^0] = B \implies i\hbar^{-1} [P, A] = B + i\varepsilon h^{-1} [P, B]. \quad (41)$$

Then perturbation $\varepsilon B$ is replaced by $\varepsilon^2 B'$, which is the right hand expression in (39) minus $A^0$, i.e.

$$
B' = -i\hbar^{-1} [P, B] + \sum_{2 \leq n \leq K-1} \frac{1}{n!} \varepsilon^{n-2} (-i\hbar^{-1})^n \text{Ad}_P^n(A), \quad (42)
$$

where we ignored the remainder.
Then formally we can compensate $\varepsilon B$, taking

$$P = \sum_{\theta} i\hbar (A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2))^{-1} b_\theta(\xi) e^{i\langle \theta, x \rangle}, \quad (40)$$

so that

$$i\hbar^{-1} [P, A^0] = B \implies i\hbar^{-1} [P, A] = B + i\varepsilon h^{-1} [P, B]. \quad (41)$$

Then perturbation $\varepsilon B$ is replaced by $\varepsilon^2 B'$, which is the right hand expression in (39) minus $A^0$, i.e.

$$B' = -i\hbar^{-1} [P, B] + \sum_{2 \leq n \leq K-1} \frac{1}{n!} \varepsilon^{n-2} (-i\hbar^{-1})^n \text{Ad}_P^n(A), \quad (42)$$

where we ignored the remainder.

New perturbation, again formally, has a magnitude of $\varepsilon^2$. Repeating this process we will make a perturbation negligible.
Remark 6.

However, we need to address the following issues:

1. **Denominator**

\[
h^{-1}(A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2)) = \langle \nabla_\xi A^0, \theta \rangle + O(h^{1-\sigma})\]

could be small.
Remark 6.

However, we need to address the following issues:

1. **Denominator**
   \[
   h^{-1}(A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2)) = \langle \nabla_\xi A^0, \theta \rangle + O(h^{1-\sigma}) \]
   could be small.

2. **In set \(B'\) increases**:
   \[
   \epsilon^2 B' = \epsilon^2 B'_2 + \epsilon^3 B'_3 + \ldots + \epsilon^M B'_M, \]
   where for \(B'_j\) the frequency set is \(\Theta'_j\) (the arithmetic sum of \(j\) copies of \(\Theta'\)).
Remark 6.

However, we need to address the following issues:

1. Denominator
\[ h^{-1}(A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2)) = \langle \nabla_\xi A^0, \theta \rangle + O(h^{1-\sigma}) \] could be small.

2. In $B'$ set $\Theta'$ increases: $\epsilon^2 B' = \epsilon^2 B'_2 + \epsilon^3 B'_3 + \ldots + \epsilon^M B'_M$, where for $B'_j$ the frequency set is $\Theta'_j$ (the arithmetic sum of $j$ copies of $\Theta'$).

3. We need to prove that the remainder is negligible.
**Remark 6.**

However, we need to address the following issues:

1. **Denominator**
   
   $$ h^{-1}(A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2)) = \langle \nabla_\xi A^0, \theta \rangle + O(h^{1-\sigma}) $$

   This denominator could be small.

2. **In $B'$ set $\Theta'$ increases**:
   
   $$ \varepsilon^2 B' = \varepsilon^2 B'_2 + \varepsilon^3 B'_3 + \ldots + \varepsilon^M B'_M, $$

   where for $B'_j$ the frequency set is $\Theta'_j$ (the arithmetic sum of $j$ copies of $\Theta'$).

3. **We need to prove that the remainder is negligible.**

4. **This transformation was used in Section 9 of [PS3] (etc); in contrast to these papers we use Weyl quantization instead of $pq$-quantization, and have therefore**
   
   $$ (A^0(\xi + \theta h/2) - A^0(\xi - \theta h/2)) $$

   instead of
   
   $$ (A^0(\xi + \theta h) - A^0(\xi)). $$
One can see easily that $\xi$ is non-resonant (inequality (27) holds for all $\theta \in \Theta'_K \setminus 0$), then the terms could be estimated by $h^{\delta n}$ and our construction works with $K = 3M/\delta$. 
One can see easily that $\xi$ is non-resonant (inequality (27) holds for all $\theta \in \Theta'_K \setminus 0$), then the terms could be estimated by $h^{\delta n}$ and our construction works with $K = 3M/\delta$.

Indeed, if $P = P(x, hD)$ has the symbol, satisfying

$$|D_\xi^\alpha D_x^\beta P| \leq C_{\alpha \beta} \rho^{-1-|\alpha|} \quad \forall \alpha, \beta,$$

then $B' = \varepsilon h^{-1}[P, B]$ has a symbol, satisfying

$$|D_\xi^\alpha D_x^\beta B'| \leq c'_{\alpha \beta} \varepsilon \rho^{-2-|\alpha|} \quad \forall \alpha, \beta,$$

so indeed $\varepsilon' = \varepsilon h^{-1} \rho^{-2}$. 
One can see easily that $\xi$ is non-resonant (inequality (27) holds for all $\theta \in \Theta'_{K} \setminus 0$), then the terms could be estimated by $h^{\delta n}$ and our construction works with $K = 3M/\delta$.

Indeed, if $P = P(x, hD)$ has the symbol, satisfying

$$|D_{\xi}^{\alpha} D_{x}^{\beta} P| \leq C_{\alpha \beta} \rho^{-1-|\alpha|} \quad \forall \alpha, \beta,$$  \hspace{1cm} (43)

then $B' = \varepsilon h^{-1}[P, B]$ has a symbol, satisfying

$$|D_{\xi}^{\alpha} D_{x}^{\beta} B'| \leq c'_{\alpha \beta} \varepsilon \rho^{-2-|\alpha|} \quad \forall \alpha, \beta,$$  \hspace{1cm} (44)

so indeed $\varepsilon' = \varepsilon h^{-1} \rho^{-2}$.

Then we can eliminate a perturbation completely, save terms with the frequency 0, both old and new.
One can see easily that $\xi$ is non-resonant (inequality (27) holds for all $\theta \in \Theta'_K \setminus 0$), then the terms could be estimated by $h^{\delta n}$ and our construction works with $K = 3M/\delta$.

Indeed, if $P = P(x, hD)$ has the symbol, satisfying

$$|D_\xi^\alpha D_x^\beta P| \leq C_{\alpha\beta} \rho^{1-|\alpha|} \forall \alpha, \beta, \quad (43)$$

then $B' = \varepsilon h^{-1}[P, B]$ has a symbol, satisfying

$$|D_\xi^\alpha D_x^\beta B'| \leq c'_{\alpha\beta} \varepsilon \rho^{-2-|\alpha|} \forall \alpha, \beta, \quad (44)$$

so indeed $\varepsilon' = \varepsilon h^{-1} \rho^{-2}$.

Then we can eliminate a perturbation completely, save terms with the frequency 0, both old and new.

On the other hand, if $\xi$ is resonant we can eliminate only frequencies $\theta$ which satisfy (27). This concludes the proof of Theorem 5. $\square$
Complete spectral asymptotics

The main idea behind the proof of Theorem 1 is a very long range (up to time $T \asymp h^{-M}$) propagation of singularities.

It is known (see Chapter 4 of [Ivr1]) that under microhyperbolicity condition (4) for $|\tau - \lambda| < \epsilon$ the following complete asymptotics holds:

$$F_t \to h^{-1} \tau \left[\bar{\chi}_T(t) (Q_2(x,u)_h(x,y,t) Q_1(y)) |y = x\right] \sim \sum_{n \geq 0} \nu'_{n,Q_1,Q_2}(x, \tau) h^{1-d+n},$$

where $u_h(x,y,t)$ is the Schwartz kernel of the propagator $e^{ih^{-1}tA}$, $\bar{\chi} \in C_0^\infty([-1,1])$, $\bar{\chi}(t) = 1$ at $[-1/2, 1/2]$, $\bar{\chi}_T(t) = \bar{\chi}(t/T)$, $T \in [h^{1-\delta}, T^*]$, $T^*$ is a small constant here and $Q_j = Q_j(x, hD)$ are $h$-pseudo-differential operators; we write operators, acting with respect to $y$ on Schwartz kernels, to the right of it.
Complete spectral asymptotics

The main idea behind the proof of Theorem 1 is a very long range (up to time $T \asymp h^{-M}$) propagation of singularities.

It is known (see Chapter 4 of [Ivr1]) that under microhyperbolicity condition (4) for $|\tau - \lambda| < \epsilon$ the following complete asymptotics holds:

$$F_{t \rightarrow h^{-1}T} \left[ \bar{\chi}_T(t) \left( Q_2 x u_h(x, y, t)^t Q_1 y \right) |_{y=x} \right] \sim \sum_{n \geq 0} \kappa'_n, Q_1, Q_2(x, \tau) h^{1-d+n}, \ (45)$$

where $u_h(x, y, t)$ is the Schwartz kernel of of the propagator $e^{ih^{-1}tA}$, $ar{\chi} \in C_0^\infty([-1, 1])$, $ar{\chi}(t) = 1$ at $[-\frac{1}{2}, \frac{1}{2}]$, $ar{\chi}_T(t) = \bar{\chi}(t/T)$, $T \in [h^{1-\delta}, T^*]$, $T^*$ is a small constant here and $Q_j = Q_j(x, hD)$ are $h$-pseudo-differential operators; we write operators, acting with respect to $y$ on Schwartz kernels to the right of it.
This equality (45) plus Hörmander’s Tauberian theorem imply the remainder estimates $O(h^{1-d} T^*)$ for $(Q_{2x} e_h(x, y, \tau)^t Q_{1y})|_{x=y}$. 
This equality (45) plus Hörmander’s Tauberian theorem imply the remainder estimates $O(h^{1-d} T^{*-1})$ for $(Q_2 e_h(x, y, \tau)^t Q_1 y)|_{x=y}$.

On the other hand, if we can improve (45) by increasing $T^*$, we can improve the remainder estimate to $O(T^{*-1} h^{1-d})$. 
This equality (45) plus Hörmander’s Tauberian theorem imply the remainder estimates $O(h^{1-d} T^*)$ for $(Q_2 x e_h(x, y, \tau)^t Q_1 y)|_{x=y}$.

On the other hand, if we can improve (45) by increasing $T^*$, we can improve the remainder estimate to $O(T^* h^{1-d})$.

It is a very general result, and under certain assumptions to the dynamics it could be improved.
This equality (45) plus Hörmander’s Tauberian theorem imply the remainder estimates $O(h^{1-d} T^{*-1})$ for $(Q_{2x} e^{h(x, y, \tau) t} Q_{1y})|_{x=y}$.

On the other hand, if we can improve (45) by increasing $T^*$, we can improve the remainder estimate to $O(T^{*-1} h^{1-d})$.

It is a very general result, and under certain assumptions to the dynamics it could be improved.

Furthermore, for operator (1) with $h \varepsilon \leq \varepsilon_0$ the equality (45) holds with $T^* = \varepsilon_1 \varepsilon^{-1}$ where $\varepsilon_j$ are small constants and we assume that $\varepsilon \geq h^M$ for some $M$. 

Victor Ivrii (Math., Toronto) Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture June 3, 2020 37 / 63
This equality (45) plus Hörmander’s Tauberian theorem imply the remainder estimates $O(h^{1-d} T^{*1})$ for $(Q_2 e_h(x, y, \tau) t Q_1)\big|_{x=y}$.

On the other hand, if we can improve (45) by increasing $T^*$, we can improve the remainder estimate to $O(T^{*-1} h^{1-d})$.

It is a very general result, and under certain assumptions to the dynamics it could be improved.

Furthermore, for operator (1) with $h \varepsilon \leq \varepsilon_0$ the equality (45) holds with $T^* = \varepsilon_1 \varepsilon^{-1}$ where $\varepsilon_j$ are small constants and we assume that $\varepsilon \geq h^M$ for some $M$.

Indeed, one can prove, that for such time $\xi(t)$ is confined to $C \varepsilon_1$-vicinity of $\xi(0)$. One needs to understand this claim in the sense of quantum mechanics (or microlocal analysis):

$$F_{t \rightarrow h^{-1} t} \left[ \overline{\chi} T(t) Q_1(hD) e^{ih^{-1} t A} Q_2(hD) \right] = O(h^M) \quad (46)$$

provided $\text{dist}(\text{supp}(Q_1), \text{supp}(Q_2)) \geq C \varepsilon T$. 
Then $x$ propagates for this time approximately in the direction of $\nabla_\xi A^0(\xi)|_{\xi=\xi(0)}$, again in the sense of quantum mechanics:

$$F_{t \to h^{-1}\tau} [\chi_T(\pm t)\psi_1(x)e^{ih^\tau_1t}A\psi_2(x)Q(hD)] \equiv 0 \mod O(h^M),$$ (47)

provided $Q$ is supported in $\epsilon$-vicinity of $\xi^0$, and

$$\text{dist}(\text{supp}(\psi_1), \text{supp}(\psi_2) \pm T\nabla_\xi A^0(\xi^0)) \geq C_0\epsilon T,$$

where $\chi \in \mathcal{C}_0^\infty([1 - \epsilon, 1 + \epsilon]).$
Then $x$ propagates for this time approximately in the direction of $\nabla_\xi A^0(\xi)|_{\xi=\xi(0)}$, again in the sense of quantum mechanics:

$$F_{t \to h^{-1}\tau}[\chi_T(\pm t)\psi_1(x)e^{ih^{-1}tA}\psi_2(x)Q(hD)] \equiv 0 \mod O(h^M), \quad (47)$$

provided $Q$ is supported in $\epsilon$-vicinity of $\xi^0$, and

$$\text{dist}(\text{supp}(\psi_1), \text{supp}(\psi_2) \pm T\nabla_\xi A^0(\xi^0)) \geq C_0\epsilon T,$$

where $\chi \in \mathcal{C}_0^\infty([1-\epsilon, 1+\epsilon])$.

Therefore considering its kernel and making some partitions with respect to $x$, $y$, $hD_x$ and $hD_y$ we arrive to

$$F_{t \to h^{-1}\tau}[\chi_T(t)(Q_{2x}u_h(x, y, t)^t Q_{1y})|_{y=x}] = O(h^M) \quad (48)$$
Then $x$ propagates for this time approximately in the direction of $\nabla_{\xi} A^0(\xi)|_{\xi = \xi(0)}$, again in the sense of quantum mechanics:

$$
F_{t \to h^{-1} \tau} \left[ \chi_{T}(\pm t) \psi_1(x) e^{i h^{-1} t A} \psi_2(x) Q(hD) \right] \equiv 0 \mod O(h^M),
$$

(47)

provided $Q$ is supported in $\epsilon$-vicinity of $\xi^0$, and

$$
\text{dist}(\text{supp}(\psi_1), \text{supp}(\psi_2) \pm T \nabla_{\xi} A^0(\xi^0)) \geq C_0 \epsilon T,
$$

where $\chi \in \mathcal{C}_0^\infty([1 - \epsilon, 1 + \epsilon])$.

Therefore considering its kernel and making some partitions with respect to $x$, $y$, $hD_x$ and $hD_y$ we arrive to

$$
F_{t \to h^{-1} \tau} \left[ \chi_{T}(t) (Q_{2x} u_h(x, y, t)^t Q_{1y}) |_{y = x} \right] = O(h^M)
$$

(48)

and using partition with respect to $t$ we conclude that (45) holds with

$$
T^* = C_0^{-1} \epsilon^{-1}.
$$
Then $x$ propagates for this time approximately in the direction of
$\nabla_\xi A^0(\xi)|_{\xi=\xi(0)}$, again in the sense of quantum mechanics:

$$
F_{t\to h^{-1}\tau} \left[ \chi_T(\pm t) \psi_1(x) e^{ih^{-1}tA} \psi_2(x) Q(hD) \right] \equiv 0 \mod O(h^M), \quad (47)
$$

provided $Q$ is supported in $\varepsilon$-vicinity of $\xi^0$, and

$$
\text{dist} (\text{supp}(\psi_1), \text{supp}(\psi_2) \pm T \nabla_\xi A^0(\xi^0)) \geq C_0 \varepsilon T,
$$

where $\chi \in \mathcal{C}_0^\infty([1-\varepsilon, 1+\varepsilon])$.

Therefore considering its kernel and making some partitions with respect to $x, y, hD_x$ and $hD_y$ we arrive to

$$
F_{t\to h^{-1}\tau} \left[ \chi_T(t) (Q_{2x} u_h(x, y, t)^t Q_{1y})|_{y=x} \right] = O(h^M) \quad (48)
$$

and using partition with respect to $t$ we conclude that (45) holds with

$$
T^* = C_0^{-1} \varepsilon^{-1}.
$$

Then the remainder estimate is $O(h^{1-d} T^*^{-1}) = O(\varepsilon h^{1-d})$. 
Long-term propagation

Proposition 7.

In the framework of Theorem 1 equalities (46)--(48) hold with \( T^* = h^{-M} \).
Long-term propagation

Proposition 7.

In the framework of Theorem 1 equalities (46)–(48) hold with $T^* = h^{-M}$.

Then (45) also holds with $T^* = h^{-M}$ and the remainder estimate is $O(h^{1-d+M})$ which implies Theorem 1.
Complete spectral asymptotics (sketch of proof)

Long-term propagation

**Proposition 7.**

*In the framework of Theorem 1 equalities (46)–(48) hold with $T^* = h^{-M}$. Then (45) also holds with $T^* = h^{-M}$ and the remainder estimate is $O(h^{1-d+M})$ which implies Theorem 1.

The crucial element is the proof of (46), the rest is easy.*
Long-term propagation

**Proposition 7.**

*In the framework of Theorem 1 equalities (46)–(48) hold with \( T^* = h^{-M} \).*

Then (45) also holds with \( T^* = h^{-M} \) and the remainder estimate is \( O(h^{1-d+M}) \) which implies Theorem 1.

The crucial element is the proof of (46), the rest is easy. And the crucial element in the proof of (46) plays a canonical form, established in Theorem 5.
Long-term propagation

Proposition 7.

*In the framework of Theorem 1 equalities* (46)–(48) *hold with* $T^* = h^{-M}$.

Then (45) also holds with $T^* = h^{-M}$ and the remainder estimate is $O(h^{1-d+M})$ which implies Theorem 1.

The crucial element is the proof of (46), the rest is easy. And the crucial element in the proof of (46) plays a canonical form, established in Theorem 5.

Indeed, for operator $\mathcal{A}$ singularity at non-resonant point does not propagate at all,
Proposition 7.

In the framework of Theorem 1 equalities (46)–(48) hold with $T^* = h^{-M}$.

Then (45) also holds with $T^* = h^{-M}$ and the remainder estimate is $O(h^{1-d+M})$ which implies Theorem 1.

The crucial element is the proof of (46), the rest is easy. And the crucial element in the proof of (46) plays a canonical form, established in Theorem 5.

Indeed, for operator $\mathcal{A}$ singularity at non-resonant point does not propagate at all, while singularity at resonant point $\xi$ can propagate only along $\mathcal{X}(\xi)$. 
Proposition 7.

In the framework of Theorem 1 equalities (46)–(48) hold with $T^* = h^{-M}$.

Then (45) also holds with $T^* = h^{-M}$ and the remainder estimate is $O(h^{1-d+M})$ which implies Theorem 1.

The crucial element is the proof of (46), the rest is easy. And the crucial element in the proof of (46) plays a canonical form, established in Theorem 5.

Indeed, for operator $A$ singularity at non-resonant point does not propagate at all, while singularity at resonant point $\xi$ can propagate only along $X(\xi)$ and therefore does not go far away due to (32) $\text{diam}(X(\xi)) \leq C\rho_{d-1}$ which is due to strong convexity:
First of all, let us replace operator $\mathcal{A}$ defined by (34) by operator

$$\mathcal{A}' = \mathcal{A}^0(hD) + \varepsilon \mathcal{B}'(x, hD), \quad \mathcal{B}'(x, hD) = S(hD)\mathcal{B}S(hD)$$

with $S(hD)$ operator with symbol $T(\xi)$ which is a characteristic function of $\Omega_\tau$ defined by (26) with $C = 6$. Then (33) holds.
First of all, let us replace operator $\mathcal{A}$ defined by (34) by operator

$$\mathcal{A}' = \mathcal{A}^0(hD) + \varepsilon \mathcal{B}'(x, hD), \quad \mathcal{B}'(x, hD) = S(hD)\mathcal{B}S(hD) \quad (49)$$

with $S(hD)$ operator with symbol $T(\xi)$ which is a characteristic function of $\Omega_\tau$ defined by (26) with $C = 6$. Then (33) holds. From now on $\mathcal{A} := \mathcal{A}'$ and $\mathcal{B} := \mathcal{B}'$. 
It would be sufficient to prove Theorem 4 for operator $\mathcal{A}$:
It would be sufficient to prove Theorem 4 for operator $\mathcal{A}$:

**Proposition 8.**

1. For each point $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$
   \[ \text{dist}(\lambda, \text{Spec}(\mathcal{A}(\xi))) \leq Ch^M. \]
It would be sufficient to prove Theorem 4 for operator $\mathcal{A}$:

**Proposition 8.**

1. For each point $\lambda \in \text{Spec}(A(\xi)) \cap \{ |\lambda - \tau| \leq \varepsilon h^{-\delta} \}$
   \[ \text{dist}(\lambda, \text{Spec}(A(\xi))) \leq Ch^M. \]

2. Conversely, for each point $\lambda \in \text{Spec}(A(\xi)) \cap \{ |\lambda - \tau| \leq \varepsilon h^{-\delta} \}$
   \[ \text{dist}(\lambda, \text{Spec}(A(\xi))) \leq Ch^M. \]
It would be sufficient to prove Theorem 4 for operator $A$:

**Proposition 8.**

1. For each point $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$
   \[
   \text{dist}(\lambda, \text{Spec}(A(\xi))) \leq C h^M.
   \]

2. Conversely, for each point $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$
   \[
   \text{dist}(\lambda, \text{Spec}(A(\xi))) \leq C h^M.
   \]

3. Furthermore, if $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$ is a simple eigenvalue separated from the rest of $\text{Spec}(A(\xi))$ by a distance at least $2h^{M-1}$, then there exists $\lambda' \in \text{Spec}(A(\xi)) \cap \{||\lambda' - \lambda| \leq C h^M\}$ separated from the rest of $\text{Spec}(A(\xi))$ by a distance at least $h^{M-1}$. 
It would be sufficient to prove Theorem 4 for operator $\mathcal{A}$:

**Proposition 8.**

1. For each point $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$
   \[\text{dist}(\lambda, \text{Spec}(\mathcal{A}(\xi))) \leq Ch^M.\]

2. Conversely, for each point $\lambda \in \text{Spec}(\mathcal{A}(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$
   \[\text{dist}(\lambda, \text{Spec}(A(\xi))) \leq Ch^M.\]

3. Furthermore, if $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$ is a simple eigenvalue separated from the rest of $\text{Spec}(A(\xi))$ by a distance at least $2h^{M-1}$, then there exists $\lambda' \in \text{Spec}(\mathcal{A}(\xi)) \cap \{|\lambda' - \lambda| \leq Ch^M\}$ separated from the rest of $\text{Spec}(A(\xi))$ by a distance at least $h^{M-1}$.

4. Conversely, if $\lambda' \in \text{Spec}(\mathcal{A}(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\}$ is a simple eigenvalue separated from the rest of $\text{Spec}(A(\xi))$ by a distance at least $2h^{M-1}$, then there exists $\lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda' - \lambda| \leq Ch^M\}$ separated from the rest of $\text{Spec}(A(\xi))$ by a distance at least $h^{M-1}$. 
Structure of operator $\mathcal{A}$

For $\xi \in \Xi_n \setminus \Xi_{n+1}$ denote by $\mathcal{H}(\xi)$ the subspace $L^2(O)$ consisting of functions of the form

$$\sum_{\xi' \in \mathcal{X}(\xi)} c_{\xi'} e^{i\langle x, \xi' \rangle}.$$  \hfill (50)
Structure of operator $\mathcal{A}$

For $\xi \in \Xi_n \setminus \Xi_{n+1}$ denote by $\mathcal{H}(\xi)$ the subspace $L^2(O)$ consisting of functions of the form

$$\sum_{\xi' \in \mathcal{X}(\xi)} c_{\xi'} e^{i\langle x, \xi' \rangle}.$$  \hfill (50)

In virtue of the properties of $\mathcal{A}$ and $\mathcal{B}$ and of resonant sets we arrive to

**Proposition 9.**

Let $\nu > 0$ in the definition of $\Theta'$ and $\delta > 0$ in the definition of $\Omega_\tau$ be sufficiently small. Let $h$ be sufficiently small.

Then for $\xi \in \Xi_n \setminus \Xi_{n+1}$ operators $\mathcal{B}$ and $\mathcal{A}$ transform $\mathcal{H}(\xi)$ into $\mathcal{H}(\xi)$. 
Structure of operator $\mathcal{A}$

For $\xi \in \Xi_n \setminus \Xi_{n+1}$ denote by $\mathcal{H}(\xi)$ the subspace $\mathcal{L}^2(\mathcal{O})$ consisting of functions of the form

$$\sum_{\xi' \in \mathcal{X}(\xi)} c_{\xi'} e^{i\langle x, \xi' \rangle}.$$  \hfill (50)

In virtue of the properties of $\mathcal{A}$ and $\mathcal{B}$ and of resonant sets we arrive to

**Proposition 9.**

*Let $\nu > 0$ in the definition of $\Theta'$ and $\delta > 0$ in the definition of $\Omega_\tau$ be sufficiently small. Let $h$ be sufficiently small. Then for $\xi \in \Xi_n \setminus \Xi_{n+1}$ operators $\mathcal{B}$ and $\mathcal{A}$ transform $\mathcal{H}(\xi)$ into $\mathcal{H}(\xi).$

Let us denote by $\mathcal{A}_\gamma(\xi)$ and $\mathcal{B}_\gamma(\xi)$ restrictions of $\mathcal{A}$ and $\mathcal{B}$ to $\mathcal{H}(h(\gamma + \xi)).$ Here for $n = 0$ we consider $\Xi_0$ to be the set of all non-resonant points and $\mathcal{X}(\xi) = \{\xi\}$ for $\xi \in \Xi_0.$
Then due to Propositions 8 and 9 we arrive to

**Proposition 10.**

1. For each point \( \lambda \in \text{Spec}(A(\xi)) \cap \{ |\lambda - \tau| \leq \varepsilon h^{-\delta} \} \) exists \( \gamma \in \Gamma^* \) such that \( \xi = h(\gamma + \xi) \in \Omega_\tau \) and \( \text{dist}(\lambda, \text{Spec}(A_\gamma(\xi))) \leq Ch^M \).
Then due to Propositions 8 and 9 we arrive to

**Proposition 10.**

1. For each point \( \lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\} \) exists \( \gamma \in \Gamma^* \) such that \( \xi = h(\gamma + \xi) \in \Omega_\tau \) and \( \text{dist}(\lambda, \text{Spec}(A_{\gamma}(\xi))) \leq Ch^M. \)

2. Conversely, for each point \( \lambda \in \text{Spec}(A_{\gamma}(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\} \) and \( \xi = h(\xi + \gamma) \), \( \text{dist}(\lambda, \text{Spec}(A(\xi))) \leq Ch^M. \)
Then due to Propositions 8 and 9 we arrive to

**Proposition 10.**

1. For each point \( \lambda \in \text{Spec}(A(\xi)) \cap \{ |\lambda - \tau| \leq \varepsilon h^{-\delta} \} \) exists \( \gamma \in \Gamma^* \) such that \( \xi = h(\gamma + \xi) \in \Omega_\tau \) and \( \text{dist}(\lambda, \text{Spec}(A_\gamma(\xi))) \leq Ch^M \).

2. Conversely, for each point \( \lambda \in \text{Spec}(A_\gamma(\xi)) \cap \{ |\lambda - \tau| \leq \varepsilon h^{-\delta} \} \) and \( \xi = h(\xi + \gamma) \), \( \text{dist}(\lambda, \text{Spec}(A(\xi))) \leq Ch^M \).

3. Further, if \( \lambda \in \text{Spec}(A_\gamma(\xi)) \cap \{ |\lambda - \tau| \leq \varepsilon h^{-\delta} \} \) is a simple eigenvalue separated from the rest of \( \text{Spec}(A(\xi)) \) by a distance at least \( 2h^{M-1} \), then there exist \( \gamma \) and \( \lambda' \), such that for \( \xi = h(\gamma + \xi) \), \( \lambda' \in \text{Spec}(A(\xi)) \cap \{ |\lambda' - \lambda| \leq Ch^M \} \), separated from the rest of \( \text{Spec}(A_\gamma(\xi)) \) by a distance at least \( h^{M-1} \) and from \( \bigcup_{\gamma' \in \Gamma^*, \gamma' \neq \gamma} \text{Spec}(A_{\gamma'}(\xi)) \) by a distance at least \( h^{M-1} \) as well.
Proposition 10 (continued).

Conversely, if \( \lambda' \in \text{Spec}(\mathcal{A}_\gamma(\xi)) \cap \{|\lambda - \tau| \leq \varepsilon h^{-\delta}\} \) is a simple eigenvalue separated from the rest of \( \text{Spec}(\mathcal{A}_\gamma(\xi)) \) by a distance at least \( 2h^{M-1} \), and also separated from \( \bigcup_{\gamma' \in \Gamma^*, \gamma' \neq \gamma} \text{Spec}(\mathcal{A}_{\gamma'}(\xi)) \) by a distance at least \( 2h^{M-1} \), then there exists \( \lambda \in \text{Spec}(A(\xi)) \cap \{|\lambda' - \lambda| \leq Ch^M\} \) separated from the rest of \( \text{Spec}(A(\xi)) \) by a distance at least \( h^{M-1} \).
Preparatory

We are looking for an appropriate point $\xi^* = h(\gamma^* + \xi^*)$. 

Preparatory

We are looking for an appropriate point $\xi^* = h(\gamma^* + \xi^*)$. The first approximation is $\xi^* \in \Sigma_\tau$ satisfying Condition (E) and such that $\text{dist}(\xi^*, \partial\mathcal{O}) \geq \epsilon_0$. Any $\xi \in \Sigma_\tau$ in $\epsilon'$-vicinity of $\xi^*$ also fits provided $\epsilon' > 0$ is sufficiently small.
Preparatory

We are looking for an appropriate point $\xi^* = h(\gamma^* + \xi^*)$. The first approximation is $\xi^* \in \Sigma_\tau$ satisfying Condition (E) and such that $\text{dist}(\xi^*, \partial \mathcal{O}) \geq \epsilon_0$. Any $\xi \in \Sigma_\tau$ in $\epsilon'$-vicinity of $\xi^*$ also fits provided $\epsilon' > 0$ is sufficiently small.

In virtue of Proposition 4 we can select $\xi_{\text{new}}^* \in \Sigma_\tau$ such that $|\xi_{\text{new}}^* - \xi^*| \leq h^\delta$ and $\xi_{\text{new}}^*$ satisfies (27) with $\rho = \gamma := h^\delta$. Here $\delta > 0$ is arbitrarily small and $\nu = \nu(\delta)$. 
Preparatory

We are looking for an appropriate point \( \xi^* = h(\gamma^* + \xi^*) \). The first approximation is \( \xi^* \in \Sigma_\tau \) satisfying Condition (E) and such that \( \text{dist}(\xi^*, \partial \Omega) \geq \epsilon_0 \). Any \( \xi \in \Sigma_\tau \) in \( \epsilon' \)-vicinity of \( \xi^* \) also fits provided \( \epsilon' > 0 \) is sufficiently small.

In virtue of Proposition 4 we can select \( \xi_{\text{new}}^* \in \Sigma_\tau \) such that \( |\xi_{\text{new}}^* - \xi^*| \leq h^\delta \) and \( \xi_{\text{new}}^* \) satisfies (27) with \( \rho = \gamma := h^\delta \). Here \( \delta > 0 \) is arbitrarily small and \( \nu = \nu(\delta) \). From now on \( \xi^* := \xi_{\text{new}}^* \).
Preparatory

We are looking for an appropriate point \( \xi^* = h(\gamma^* + \xi^*) \). The first approximation is \( \xi^* \in \Sigma_\tau \) satisfying Condition (E) and such that \( \text{dist}(\xi^*, \partial \mathcal{O}) \geq \epsilon_0 \). Any \( \xi \in \Sigma_\tau \) in \( \epsilon' \)-vicinity of \( \xi^* \) also fits provided \( \epsilon' > 0 \) is sufficiently small.

In virtue of Proposition 4 we can select \( \xi^*_{\text{new}} \in \Sigma_\tau \) such that \( |\xi^*_{\text{new}} - \xi^*| \leq h^\delta \) and \( \xi^*_{\text{new}} \) satisfies (27) with \( \rho = \gamma := h^\delta \). Here \( \delta > 0 \) is arbitrarily small and \( \nu = \nu(\delta) \). From now on \( \xi^* := \xi^*_{\text{new}} \).

Then, according to Theorem 5 we can diagonalize operator in \( \gamma \)-vicinity of \( \xi^* \) and there \( \rho = \gamma \). Then there

\[
|\nabla^\alpha (A^0 - A^0)| \leq C_\alpha (\epsilon + \epsilon^2 \rho^{-2-|\alpha|}) \tag{51}
\]

and in particular

\[
|\nabla^\alpha (A^0 - A^0)| \leq Ch^\delta \quad \text{for} \quad |\alpha| \leq 2. \tag{52}
\]

Let

\[
\Sigma'_\tau = \{ \xi : A^0(\xi) = \tau \}. \tag{53}
\]
Observe that in the non-resonant points we are interested in functions \( \lambda_\gamma(\xi) = A^0(h(\gamma + \xi)) \) rather than in \( \lambda_\gamma(\xi) = A^0(h(\gamma + \xi)) \).
Observe that in the non-resonant points we are interested in functions \( \lambda_\gamma(\xi) = A^0(h(\gamma + \xi)) \) rather than in \( \lambda_\gamma^0(\xi) = A^0(h(\gamma + \xi)) \).

Let \( \xi^* =: h(\gamma^* + \xi^*) \) be a point we selected. Then values in the nearby points are sufficiently separated:

\[
|\lambda_\gamma(\xi) - \lambda_\gamma^*(\xi)| \geq \epsilon h^{1+\delta} \quad \forall \gamma: |\gamma - \gamma^*| \leq Kh^{-\nu} \quad \forall \xi \in \mathcal{O}^*. \tag{54}
\]
Observe that in the non-resonant points we are interested in functions 
\( \lambda_\gamma(\xi) = A^0(h(\gamma + \xi)) \) rather than in 
\( \lambda_\gamma^0(\xi) = A^0(h(\gamma + \xi)) \).

Let \( \xi^* = h(\gamma^* + \xi^*) \) be a point we selected. Then values in the nearby points are sufficiently separated:

\[
|\lambda_\gamma(\xi) - \lambda_{\gamma^*}(\xi)| \geq \epsilon h^{1+\delta} \quad \forall \gamma: |\gamma - \gamma^*| \leq Kh^{-\nu} \quad \forall \xi \in \mathcal{O}^*. \tag{54}
\]

Indeed, \( |\gamma - \gamma^*| \leq Kh^{-\nu} \) implies that \( (\gamma - \gamma^*) \in \Theta'_K \) and then

\[
|\langle \nabla A^0(\xi^*), \gamma - \gamma^* \rangle| \geq \gamma
\]

while

\[
|\lambda_\gamma(\xi) - \lambda_{\gamma^*}(\xi) - h\langle \nabla A^0(\xi^*), \gamma - \gamma^* \rangle| \leq Ch^{3-3\nu}.
\]
Consider other non-resonant points (with $\rho = \varepsilon^{1/2} h^{-\delta}$). Let us determine how $\lambda_\gamma(\xi)$ changes when we change $\xi$. Due to (52)

$$\delta \lambda_\gamma := \lambda_\gamma(\xi + \delta \xi) - \lambda_\gamma(\xi) = h\langle \nabla A^0(\xi), \delta \xi \rangle + O(h^2|\delta \xi|^2). \quad (55)$$
Consider other non-resonant points (with $\rho = \varepsilon^{1/2} h^{-\delta}$). Let us determine how $\lambda_\gamma(\xi)$ changes when we change $\xi$. Due to (52)

$$
\delta \lambda_\gamma := \lambda_\gamma(\xi + \delta \xi) - \lambda_\gamma(\xi) = h \langle \nabla A^0(\xi), \delta \xi \rangle + O(h^2 |\delta \xi|^2).
$$

To preserve $\lambda_\gamma^*(\xi) = \tau$ in the linearized settings we need to shift $\xi$ by $\delta \xi$, which is orthogonal to $\nabla_\xi A(\xi^*)$. 
Consider other non-resonant points (with $\rho = \varepsilon^{1/2} h^{-\delta}$). Let us determine how $\lambda_\gamma(\xi)$ changes when we change $\xi$. Due to (52)

$$
\delta \lambda_\gamma := \lambda_\gamma(\xi + \delta \xi) - \lambda_\gamma(\xi) = h \langle \nabla A_0(\xi), \delta \xi \rangle + O(h^2 |\delta \xi|^2).
$$

(55)

To preserve $\lambda_{\gamma^*}(\xi) = \tau$ in the linearized settings we need to shift $\xi$ by $\delta \xi$, which is orthogonal to $\nabla_\xi A(\xi^*)$.

Let us take $\delta \xi = t \eta$

$$
\ell: |\eta| = 1, \quad \langle \nabla A_0(\xi^*), \eta \rangle = 0.
$$

(56)

Then in all non-resonant points $\xi$ the shift will be $\langle \nabla_\xi A(\xi), \delta \xi \rangle$ with an absolute value $|\langle \nabla_\xi A(\xi), \eta \rangle| \cdot |t|$. 

Case $d = 2$.

Let us start from the easiest case $d = 2$. Without any loss of the generality we assume that $\xi^*$ is strictly inside $\mathcal{O}^*$ (at the distance at least $C\epsilon^*$ from the border). Then there is just one tangent direction $\eta$ and

$$
|\langle \nabla_\xi A^0(\xi)|_{\xi=h\gamma}, \eta \rangle| \asymp |\sin \varphi(\gamma^*, \gamma)| \asymp h \min_{1 \le k \le 2p} |\gamma - \gamma^*_k| \quad (57)
$$

where $\varphi(\gamma^*, \gamma)$ is an angle between $\nabla_\xi A^0(\xi)|_{\xi=h\gamma^*}$ and $\nabla_\xi A^0(\xi)|_{\xi=h\gamma}$, and $\xi^*_1, \ldots, \xi^*_{2p-1}$ are antipodal points, and $\xi^*_{2p} = \xi^*$. 
Case $d = 2$.

Let us start from the easiest case $d = 2$. Without any loss of the generality we assume that $\xi^*$ is strictly inside $O^*$ (at the distance at least $C\epsilon^*$ from the border). Then there is just one tangent direction $\eta$ and

$$|\langle \nabla_\xi A^0(\xi)|_{\xi=h\gamma}, \eta \rangle| \simeq |\sin \varphi(\gamma^*, \gamma)| \simeq h \min_{1 \leq k \leq 2p} |\gamma - \gamma_k^*|$$

(57)

where $\varphi(\gamma^*, \gamma)$ is an angle between $\nabla_\xi A^0(\xi)|_{\xi=h\gamma^*}$ and $\nabla_\xi A^0(\xi)|_{\xi=h\gamma}$, and $\xi_1^*, \ldots, \xi_{2p-1}^*$ are antipodal points, and $\xi_{2p}^* = \xi^*$.

As long as $\min_{1 \leq k \leq 2p} |\gamma - \gamma_k^*| \gtrsim h^{1-\nu}$ we may replace here $\xi = h(\gamma + \xi_*)$ by $\xi = h\gamma$ and $A^0$ by $A$. 
Case $d = 2$.

Let us start from the easiest case $d = 2$. Without any loss of the generality we assume that $\xi^*$ is strictly inside $O^*$ (at the distance at least $C\epsilon^*$ from the border). Then there is just one tangent direction $\eta$ and

$$\left| \langle \nabla_\xi \mathcal{A}^0(\xi)|_{\xi = h\gamma}, \eta \rangle \right| \asymp \left| \sin \varphi(\gamma^*, \gamma) \right| \asymp h \min_{1 \leq k \leq 2p} |\gamma - \gamma_k^*|$$

where $\varphi(\gamma^*, \gamma)$ is an angle between $\nabla_\xi \mathcal{A}^0(\xi)|_{\xi = h\gamma^*}$ and $\nabla_\xi \mathcal{A}^0(\xi)|_{\xi = h\gamma}$, and $\xi_1^*, \ldots, \xi_{2p-1}^*$ are antipodal points, and $\xi_{2p}^* = \xi^*$.

As long as $\min_{1 \leq k \leq 2p} |\gamma - \gamma_k^*| \gtrsim h^{1-\nu}$ we may replace here $\xi = h(\gamma + \xi)$ by $\xi = h\gamma$ and $\mathcal{A}^0$ by $\mathcal{A}$.

In the nonlinear settings to ensure that

$$\lambda_{\gamma^*}(\xi^* + \delta\xi(t)) = \tau$$

we need to include in $\delta\xi(t)$ a correction: $\delta\xi(t) = t\eta + O(t^2)$ but still

$$\frac{d}{dt} \lambda_{\gamma}(\xi^* + \delta\xi(t)) \asymp h \langle \nabla_\xi \mathcal{A}(\xi)|_{\xi = h\gamma}, \eta \rangle^{-1}.$$

(59)
Then the set $\mathcal{T}(\xi) := \{ t : |t| \leq \epsilon_0, |A^0(\xi(t)) - \tau| \leq \nu h \}$ is an interval of the length $\asymp \nu$ and then the union of such sets over $\xi = h\gamma + \xi$, with indicated $\gamma$ does not exceed $R\nu$ with

$$R := \sum_{\gamma} |\langle \nabla_{\xi} A(\xi) |_{\xi=h\gamma}, \eta \rangle|^{-1}, \quad (60)$$

where we sum over set

$$\{ \gamma : |\gamma - \gamma^*| \gtrsim h^{-\nu} \ & |\lambda_{\gamma}(h\gamma) - \tau| \leq 2Ch \}.$$
Then the set $\mathcal{T}(\xi) := \{t: |t| \leq \epsilon_0, |\mathcal{A}^0(\xi(t)) - \tau| \leq \nu h\}$ is an interval of the length $\asymp \nu$ and then the union of such sets over $\xi = h\gamma + \xi$, with indicated $\gamma$ does not exceed $R\nu$ with

$$R := \sum_{\gamma} |\langle \nabla_\xi \mathcal{A}(\xi)|_{\xi=h\gamma, \eta}\rangle|^{-1},$$

(60)

where we sum over set

$$\{\gamma: |\gamma - \gamma^*| \gtrsim h^{-\nu} \& |\lambda_\gamma(h\gamma) - \tau| \leq 2Ch\}.$$ 

The last restriction is due to the fact that $\mathcal{T}(\xi) \neq \emptyset$ only for points with $|\lambda_\gamma(h\gamma) - \tau| \leq 2Ch$. 


Then the set $\mathcal{T}(\xi) := \{ t : |t| \leq \varepsilon_0, |A^0(\xi(t)) - \tau| \leq v h \}$ is an interval of the length $\asymp v$ and then the union of such sets over $\xi = h \gamma + \xi$, with indicated $\gamma$ does not exceed $R v$ with

$$R := \sum_{\gamma} |\langle \nabla_{\xi} A(\xi)|_{\xi = h \gamma}, \eta \rangle|^{-1},$$

(60)

where we sum over set

$$\{ \gamma : |\gamma - \gamma^*| \gtrsim h^{-v} \& |\lambda_\gamma(h \gamma) - \tau| \leq 2 Ch \}.$$

The last restriction is due to the fact that $\mathcal{T}(\xi) \neq \emptyset$ only for points with $|\lambda_\gamma(h \gamma) - \tau| \leq 2 Ch$.

One can see easily that $R \asymp h^{-1} \log h$. Then, as $R v \leq \varepsilon'$ the set $[-\varepsilon_0, \varepsilon_0] \setminus \bigcup_{\gamma} \mathcal{T}(h(\gamma + \xi))$ contains an interval of the length $\ell = v$ and for all $t$, belonging to this interval,

$$|\lambda_\gamma(h(\gamma + \xi + \delta \xi(t))) - \tau| \geq \varepsilon v h.$$  

(61)
Then the set $\mathcal{T}(\xi) := \{ t : |t| \leq \epsilon_0, |A^0(\xi(t)) - \tau| \leq v h \}$ is an interval of the length $\asymp v$ and then the union of such sets over $\xi = h\gamma + \xi$, with indicated $\gamma$ does not exceed $R v$ with

$$R := \sum_{\gamma} |\langle \nabla_\xi A(\xi) \rangle_{\xi=h\gamma, \eta}|^{-1}, \quad (60)$$

where we sum over set

$$\{ \gamma : |\gamma - \gamma^*| \gtrsim h^{-\nu} & |\lambda_\gamma(h\gamma) - \tau| \leq 2Ch \}.\]$$

The last restriction is due to the fact that $\mathcal{T}(\xi) \neq \emptyset$ only for points with $|\lambda_\gamma(h\gamma) - \tau| \leq 2Ch$.

One can see easily that $R \asymp h^{-1}|\log h|$. Then, as $R v \leq \epsilon'$ the set $[-\epsilon_0, \epsilon_0] \setminus \bigcup_\gamma \mathcal{T}(h(\gamma + \xi))$ contains an interval of the length $\ell = v$ and for all $t$, belonging to this interval,

$$|\lambda_\gamma(h(\gamma + \xi + \delta\xi(t))) - \tau| \geq \epsilon vh. \quad (61)$$

Then we need to take $\nu = \epsilon R^{-1} = \epsilon h|\log h|^{-1}$ and for $d = 2$ as far as non-resonant are concerned, Theorem 4 is almost proven.
Case $d \geq 3$

In this case we need to be more subtle and to make $(d - 1)$ steps. We start from the point $\xi^* = h(\gamma^* + \xi^*)$; again without any loss of the generality we assume that $\xi^*$ is strictly inside $\mathcal{O}^*$ (at the distance at least $C\varepsilon^*$ from the border). Then after each step below it still will be the case (with decreasing constant).
Case $d \geq 3$

In this case we need to be more subtle and to make $(d - 1)$ steps. We start from the point $\xi^* = h(\gamma^* + \xi^*)$; again without any loss of the generality we assume that $\xi^*$ is strictly inside $\mathcal{O}^*$ (at the distance at least $C\varepsilon^*$ from the border). Then after each step below it still will be the case (with decreasing constant).

1. **On Step 1** we select $\eta = \eta_1$ and consider only $\gamma$ such that (57) holds; more precisely, the left-hand expression needs to be greater than the right-hand expression, multiplied by $\varepsilon$ (one can see easily, that the opposite inequality holds). Then $R \asymp h^{1-d}$ and therefore exists $\xi^*$ such that $\lambda_{\gamma^*}(\xi^*) = \tau$ and $|\lambda_{\gamma}(\xi^*) - \tau| \geq \varepsilon v_1 h$ with $v_1 = \varepsilon h^{d-1}$ for all $\gamma$ indicated above.
Case $d \geq 3$

In this case we need to be more subtle and to make $(d - 1)$ steps. We start from the point $\xi^* = h(\gamma^* + \xi^*)$; again without any loss of the generality we assume that $\xi^*$ is strictly inside $\mathcal{O}^*$ (at the distance at least $C\epsilon^*$ from the border). Then after each step below it still will be the case (with decreasing constant).

1. **On Step 1** we select $\eta = \eta_1$ and consider only $\gamma$ such that (57) holds; more precisely, the left-hand expression needs to be greater than the right-hand expression, multiplied by $\epsilon$ (one can see easily, that the opposite inequality holds). Then $R \asymp h^{1-d}$ and therefore exists $\xi^*$ such that $\lambda_{\gamma^*}(\xi^*) = \tau$ and $|\lambda_{\gamma}(\xi^*) - \tau| \geq \epsilon v_1 h$ with $v_1 = \epsilon h^{d-1}$ for all $\gamma$ indicated above.

2. **On Step 2** we select $\eta = \eta_2$ perpendicular to $\eta_1$. To preserve inequality (61) (with smaller constant $\epsilon$) for $\gamma$, already covered by Step 1, we need to take $|\delta \xi| \leq \epsilon' v_1$ and consider $\delta \xi = t\eta_2 + O(t^2)$. 
Then the same arguments as before results in inequality (61) with
\[ \nu := \nu_2 = \epsilon R^{-1} \nu_1 \]
for a new bunch of points. Then for \( d = 3 \) as far as non-resonant are concerned, Theorem 4 is almost proven.
Then the same arguments as before results in inequality (61) with
\( v := v_2 = \epsilon R^{-1} v_1 \) for a new bunch of points. Then for \( d = 3 \) as far as
non-resonant are concerned, Theorem 4 is almost proven.

Next steps. Continuing this process, on \( k \)-th step we select \( \eta_k \)
orthogonal to \( \eta_1, \ldots, \eta_{k-1} \). Then we get \( v_k = \epsilon R^{-1} v_{k-1} \) and on the
last \( (d - 1) \)-th step we achieve a separation at least \( v_{d-1} = \epsilon R^{1-d} \).
Almost antipodal points

We need to cover points with $|\xi - \xi^*_k| \leq h^{1-\kappa}$ for $k = 1, \ldots, 2p - 1$ and as we already know for each $k$ (and fixed $\xi$) there exists no more than one such point $\xi = h(\gamma + \xi)$ with $|\lambda_\gamma(\xi) - \tau| \lesssim h^{1+\delta}$. 
Almost antipodal points

We need to cover points with $|\xi - \xi_k^*| \leq h^{1-\kappa}$ for $k = 1, \ldots, 2p - 1$ and as we already know for each $k$ (and fixed $\xi$) there exists no more than one such point $\xi = h(\gamma + \xi)$ with $|\lambda_\gamma(\xi) - \tau| \lesssim h^{1+\delta}$.

We take care of such points during Step 1. Observe that during this step we automatically take care of any point with

$$|\nabla_\xi A^0(\xi), \eta_1| \geq \epsilon h,$$  \hspace{1cm} (62)

assuming that $|t| \leq \epsilon_0$ with sufficiently small $\epsilon_0 = \epsilon_0(\epsilon)$. 
Almost antipodal points

We need to cover points with $|\xi - \xi_k| \leq h^{1-\kappa}$ for $k = 1, \ldots, 2p - 1$ and as we already know for each $k$ (and fixed $\xi$) there exists no more than one such point $\xi = h(\gamma + \xi)$ with $|\lambda_\gamma(\xi) - \tau| \lesssim h^{1+\delta}$.

We take care of such points during Step 1. Observe that during this step we automatically take care of any point with

$$|\nabla_\xi A^0(\xi), \eta_1| \geq \epsilon h,$$  \hspace{1cm} (62)

assuming that $|t| \leq \epsilon_0$ with sufficiently small $\epsilon_0 = \epsilon_0(\epsilon)$.

Let us select $\eta_1$ so that on $\eta_1$ quadratic forms at points $\xi_1^*, \ldots, \xi_{2p-1}^*$ in Condition (E) are different from one at point $\xi^*$ by at least $\epsilon_0$. Then for each $j = 1, \ldots, 2p - 1$ the measure of the set

$$\{t: |t| \leq \epsilon_0, |\lambda_{\gamma_j}(\xi + \delta\xi(t))| \leq \nu h\}$$

does not exceed $Ch^{-1}(\nu h)^{1/2}$, and then the measure of the union of such sets (by $j$) also does not exceed it.
Therefore for $\nu_1 = \epsilon_1 h^{d-1}$ (for $d \geq 3$) and $\nu_1 = \epsilon_1 h|\log h|^{-1}$ (for $d = 2$) with sufficiently small $\epsilon_1$ we can find $t : |t| \leq \epsilon_0$ so that condition (57) is fulfilled for all non-resonant points.
Resonant points

Next on this step we need to separate $\lambda_{\gamma^*}(\xi)$ from all $\lambda_n(\xi)$ (save one, coinciding with it) by the distance at least $\nu h$ by choosing $\xi$. We can during the same steps as described in the previous section: let $\lambda_{\gamma,j}(\xi)$ denote eigenvalues of $A_\gamma(\xi)$ with $j = \#\mathcal{X}(\gamma h)$. 
Next on this step we need to separate $\lambda_\gamma^*(\xi)$ from all $\lambda_n(\xi)$ (save one, coinciding with it) by the distance at least $\nu h$ by choosing $\xi$. We can during the same steps as described in the previous section: let $\lambda_{\gamma,j}(\xi)$ denote eigenvalues of $\mathcal{A}_\gamma(\xi)$ with $j = \#\mathcal{X}(\gamma h)$.

Observe that both $\mathcal{A}_\gamma(\xi)$ and $\#\mathcal{X}(\gamma h)$ depend on the equivalency class $[\gamma]$ of $\gamma$ rather than on $\gamma$ itself and that

$$\sum_{[\gamma]} \#\mathcal{X}(\gamma h) = \sum_{1 \leq n \leq d-1} \#(\Xi_n) = O(h^{1-d+\sigma'} + \varepsilon^{3/2} h^{-d-\sigma}), \quad (63)$$

where on the left $[\gamma]$ runs over all equivalency classes with $\gamma \in \bigcup_{1 \leq n \leq d-1} \Xi_n$. 

Victor Ivrii (Math., Toronto) Complete Spectral Asymptotics and Bethe-Sommerfeld Conjecture June 3, 2020 54 / 63
We also observe that for resonant points

$$|\sin \varphi(\xi, \xi^*)| \geq \epsilon h^\delta$$  \hspace{1cm} (64)

and therefore for $\lambda'_\gamma$, which are eigenvalues of $A^0(h(\gamma + \xi))$ (59) holds and signs are the same for $\gamma$ in the same block.
We also observe that for resonant points

$$| \sin \varphi(\xi, \xi^*) | \geq \epsilon h^\delta$$  \hspace{1cm} (64)

and therefore for $\lambda'_\gamma$, which are eigenvalues of $\mathcal{A}^0(h(\gamma + \xi))$ \hspace{1cm} (59) holds

and signs are the same for $\gamma$ in the same block.

On the other hand,

$$| \frac{d}{dt} B(\gamma + \xi^* + \delta \xi(t)) | \leq C \epsilon h \ll h^{1+\delta'}$$  \hspace{1cm} (65)

and therefore for $\lambda_{\gamma,j}(t)$ which are eigenvalues of $\mathcal{A}_{[\gamma]}(\xi)$ \hspace{1cm} (59) sill holds.
We also observe that for resonant points

$$|\sin \varphi(\xi, \xi^*)| \geq \epsilon h^\delta$$

and therefore for $\lambda'_\gamma$, which are eigenvalues of $A^0(h(\gamma + \xi))$ (59) holds and signs are the same for $\gamma$ in the same block.

On the other hand,

$$\left| \frac{d}{dt} B(h(\gamma + \xi^* + \delta \xi(t))) \right| \leq C \epsilon h \ll h^{1+\delta'}$$

and therefore for $\lambda_{\gamma,j}(t)$ which are eigenvalues of $A_{[\gamma]}(\xi)$ (59) still holds.

Therefore the arguments of each Steps 1, 2 etc extends to resonant points as well. However the number of new points to be taken into account on each step is given by the right-hand expression of (63) and therefore $R$ needs to be redefined

$$R := h^{1-d} + \epsilon^{3/2} h^{-d-\sigma}.$$
We also observe that for resonant points

\[ |\sin \varphi(\xi, \xi^*)| \geq \epsilon h^\delta \]  

(64)

and therefore for \( \lambda'_\gamma \), which are eigenvalues of \( \mathcal{A}^0(h(\gamma + \xi)) \) (59) holds and signs are the same for \( \gamma \) in the same block.

On the other hand,

\[ |\frac{d}{dt} \mathcal{B}(h(\gamma + \xi^* + \delta \xi(t)))| \leq C\epsilon h \ll h^{1+\delta'} \]  

(65)

and therefore for \( \lambda_{\gamma,j}(t) \) which are eigenvalues of \( \mathcal{A}_{[\gamma]}(\xi) \) (59) sill holds.

Therefore the arguments of each Steps 1, 2 etc extends to resonant points as well. However the number of new points to be taken into account on each step is given by the right-hand expression of (63) and therefore \( R \) needs to be redefined

\[ R := h^{1-d} + \epsilon^{3/2} h^{-d-\sigma}. \]  

(66)

This leads to the final expression (25) for \( \nu \). Theorem 4 is proven.
Remark 11.

It also follows from Corollary 2 that

$$\frac{1}{\varsigma} \left[ N_{h,\varepsilon}(\tau + \varsigma) - N_{h,\varepsilon}(\tau) \right] = \frac{1}{\varsigma} \left[ \mathcal{N}_{h,\varepsilon}(\tau + \varsigma) - \mathcal{N}_{h,\varepsilon}(\tau) \right] + O(h^\infty) \quad (67)$$

provided $\varsigma \geq h^M$, where $\mathcal{N}_{h,\varepsilon}(\tau)$ is the right-hand expression of (9).
Remark 11.

1 It also follows from Corollary 2 that

\[
\frac{1}{\varsigma} \left[ N_{h,\varepsilon}(\tau + \varsigma) - N_{h,\varepsilon}(\tau) \right] = \frac{1}{\varsigma} \left[ \mathcal{N}_{h,\varepsilon}(\tau + \varsigma) - \mathcal{N}_{h,\varepsilon}(\tau) \right] + O(h^{\infty}) \quad (67)
\]

provided \( \varsigma \geq h^M \), where \( \mathcal{N}_{h,\varepsilon}(\tau) \) is the right-hand expression of (9).

2 The question remains, if (67) holds for smaller \( \varsigma \), in particular, if it holds in \( \varsigma \to 0 \) limit? If the latter holds, then

\[
\frac{\partial}{\partial \tau} N_{h,\varepsilon}(\tau) = \frac{\partial}{\partial \tau} \mathcal{N}_{h,\varepsilon}(\tau) + O(h^{\infty}) \quad (68)
\]

and we call the left-hand expression the density of states.
Remark 11 (Continued).

It definitely is not necessarily true, at least in dimension 1. From now on we consider only asymptotics with respect to $\tau \to +\infty$. Let $A = \Delta + V(x)$ with periodic $V$. It is well-known that for $d = 1$ and generic periodic $V$ all spectral gaps are open which contradicts to

$$\frac{\partial}{\partial \tau} N(\tau) = \frac{\partial}{\partial \tau} N'(\tau) + O(\tau^{-\infty}).$$

(69)
Remark 11 (Continued).

3. It definitely is not necessarily true, at least in dimension 1. From now on we consider only asymptotics with respect to $\tau \to +\infty$. Let $A = \Delta + V(x)$ with periodic $V$. It is well-known that for $d = 1$ and generic periodic $V$ all spectral gaps are open which contradicts to

$$\frac{\partial}{\partial \tau} N(\tau) = \frac{\partial}{\partial \tau} \mathcal{N}(\tau) + O(\tau^{-\infty}).$$

4. On the other hand, this objection does not work in case $d \geq 2$ since only several the lowest spectral gaps are open (Bethe-Sommerfeld conjecture, proven in [PSo]).
Remark 11 (End).

Further, one can differentiate $e(x, x, \tau^2)$ if $d \geq 2$ and $V$ is compactly supported. Some generalizations are considered in [Ivr4].
Remark 11 (End).

5 Further, one can differentiate $e(x, x, \tau^2)$ if $d \geq 2$ and $V$ is compactly supported. Some generalizations are considered in [Ivr4].

6 Moreover, we can differentiate complete asymptotics of the Birman-Krein spectral shift function

$$
\zeta(\tau) := \int (e(x, x, \tau^2) - e^0(x, x, \tau^2)) \, dx \sim \sum_{n \geq 0} \bar{\kappa}_n \tau^{d-n}, \quad (70)
$$

with

$$
\bar{\kappa}_n := \int (\kappa_n(x) - \kappa_n^0) \, dx, \quad (71)
$$

where $e^0(x, y, \tau)$ and $\kappa_n^0$ correspond to $A^0 = \Delta$. 
Remark 11 (End).

Further, one can differentiate $e(x, x, \tau^2)$ if $d \geq 2$ and $V$ is compactly supported. Some generalizations are considered in [Ivr4].

Moreover, we can differentiate complete asymptotics of the Birman-Krein spectral shift function

$$\zeta(\tau) := \int (e(x, x, \tau^2) - e^0(x, x, \tau^2)) \, dx \sim \sum_{n \geq 0} \bar{\kappa}_n \tau^{d-n}, \quad (70)$$

with

$$\bar{\kappa}_n := \int (\kappa_n(x) - \kappa^0_n) \, dx, \quad (71)$$

where $e^0(x, y, \tau)$ and $\kappa^0_n$ correspond to $A^0 = \Delta$. In the case of $A = \Delta$ in the exterior of smooth, compact and non-trapping obstacle and $A^0 = \Delta$ in $\mathbb{R}^d$ such asymptotics was derived in [PP].
Discussion: Bethe-Sommerfeld conjecture for almost periodic perturbations

Remark 12.

While both the proof of Bethe-Sommerfeld conjecture and the statement of Theorem 4 rely upon periodicity, the conjecture itself (as stated in Theorem 3) does not.
Discussion: Bethe-Sommerfeld conjecture for almost periodic perturbations

**Remark 12.**

While both the proof of Bethe-Sommerfeld conjecture and the statement of Theorem 4 rely upon periodicity, the conjecture itself (as stated in Theorem 3) does not.

It is only natural to try to prove it for almost periodic perturbations.


