iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Optimally localized Wannier functions for quasi one-dimensional nonperiodic insulators

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 J. Phys. A: Math. Theor. 41 125202

(http://iopscience.iop.org/1751-8121/41/12/125202)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.147
The article was downloaded on 03/06/2010 at 06:38

Please note that terms and conditions apply.

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 125202 (15pp) doi:10.1088/1751-8113/41/12/125202

Optimally localized Wannier functions for quasi one-dimensional nonperiodic insulators

H D Cornean[1], **A Nenciu**[1] **and G Nenciu**[3,4]

[1] Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg, Denmark
[2] Faculty of Applied Sciences University 'Politehnica' of Bucharest, Splaiul Independentei 313, RO-060042 Bucharest, Romania
[3] Faculty of Physics, University of Bucharest, PO Box MG 11, RO-077125 Bucharest, Romania
[4] Institute of Mathematics of the Romanian Academy, PO Box 1-764, RO-014700 Bucharest, Romania

Received 21 September 2007, in final form 13 February 2008
Published 10 March 2008
Online at stacks.iop.org/JPhysA/41/125202

Abstract
It is proved that for general, not necessarily periodic, quasi one-dimensional systems the band position operator corresponding to an isolated part of the energy spectrum has discrete spectrum and its eigenfunctions have the same spatial localization as the corresponding spectral projection. As a consequence, an eigenbasis of the band position operator provides a basis of optimally localized (generalized) Wannier functions for quasi one-dimensional systems, and this proves the strong Marzari–Vanderbilt conjecture. If the system has some translation symmetries (e.g. usual translations, screw transformations), they are 'inherited' by the Wannier basis.

PACS numbers: 73.21.Cd, 71.15.Ap, 71.20.−b, 71.23.An, 02.30.5a

1. Introduction

Wannier functions (WF) were introduced by Wannier in 1937 [1] as bases in subspaces of states corresponding to energy bands in solids, bases consisting of exponentially localized functions (localized orbitals). For periodic crystals they are defined as Fourier transform of Bloch functions of the corresponding bands. Since then WF proved to be a key tool in quantum theory of solids as they provide a tight-binding description of the electronic band structure of solids. At the conceptual level they lay at the foundation of all effective mass-type theories, e.g., the famous Peierls–Onsager substitution describing the dynamics of Bloch electrons in the presence of an external magnetic field (see, e.g., [2]and references therein). At the quantitative level, especially after the seminal paper by Marzari and Vanderbilt [3], WF become an effective tool in *ab initio* computational studies of electronic properties of materials. Moreover during the last decades WF proved to be an essential ingredient in the study of low dimensional nanostructures such as linear chains of atoms, nanowires, nanotubes, etc (see, e.g.,

 © 2008 IOP Publishing Ltd Printed in the UK

[4, 5]). In particular, WF are essential for most formulations of transport phenomena using real space Green's function method based on Landauer–Büttiker formalism both at rigorous [6] and computational levels [7, 4].

A few remarks are in order here. The first one is that realistic low dimensional systems are not strictly one- (two)-dimensional but rather quasi one- (two)-dimensional and one has to take into account the (restricted) motion along perpendicular directions. This adds specific features such as for example the screw symmetry in nanotubes and nanowires is absent in strictly one-dimensional systems. The second one is that realistic systems, due to the presence of defects, boundaries, randomness, etc, do not have usually full translation symmetry and this asks for a theory of WF not based on Bloch formalism. Finally, let us recall that contrary to a widespread opinion (see, e.g., the discussion in [2]) that WF always exist for isolated band in solids, this is not true. More precisely, in more than one dimension there are subtle topological obstructions and these are related to the QHE [8–10]: a band for which WF are known to exist gives no contribution to the quantum Hall current. It is then crucial to have rigorous proofs of the existence of exponentially localized WF.

For one-dimensional periodic systems the existence of exponentially localized WF has been proved by Kohn in his classic paper [11] about analytic structure of Bloch functions. An extension of Kohn analysis to quasi one-dimensional systems has been done recently by Prodan [12]. As for higher dimensions it was known since the work by des Cloizeaux [13, 14] that there are obstructions to the existence of exponentially localized WF and that these obstructions are of topological origin (more precisely as explicitly stated in [15] these obstructions are connected to the topology of a vector bundle of orthogonal projections). The fact that for simple bands of time-reversal invariant systems the obstructions are absent was proved by des Cloizeaux [13, 14] under the additional condition of the existence of center of inversion and by Nenciu [15] in the general case. While the proofs in [13–15] did not use the vector bundle theory it was suggested in [2, 16] that the characteristic classes theory in combination with some deep results in the theory of analytic functions of several complex variables (Oka principle) can be used to give alternative proof of the above results and to extend them to composite bands of time-reversal symmetric systems. This has been substantiated recently in [17, 10] where the existence of exponentially localized Wannier functions has been proved for composite bands of time-reversal symmetric systems in two and three dimensions settling in the affirmative a long standing conjecture. In conclusion, the situation is satisfactory as far as periodic time-reversal symmetric Hamiltonians are considered (as already mentioned for Hamiltonians which are not time-reversal symmetric, exponentially localized Wannier functions might not exist).

As already said above both the theory and applications of Wannier functions boosted since Marzari and Vanderbilt [3] introduced studied and proposed methods to compute the so-called maximally localized Wannier functions (MLWFs) defined by the fact that they minimize the position mean-square deviation. It was conjectured in [3] that they can be chosen to be real functions and that they have 'optimal' exponential localization in the sense that they have the same exponential localization as the integral kernel of the projection operator of the corresponding band. MLWFs proved to be an invaluable tool in the theory of electronic properties of periodic media especially in the modern theory of electronic polarizability (see, e.g., [18] and references therein).

In the one-dimensional case the theory of MLWFs is much more developed. It is known [3] that MLWFs are identical to the eigenfunctions of the 'band position' operator and so they are unique (up to uninteresting phases) and can be chosen to be real functions. Moreover, the phases of the corresponding Bloch functions are related to the parallel transport procedure [3, 19]. Recently a detailed study of Wannier functions, including their exponential decay, emphasizing the difference between the cases with and without inversion symmetry appeared in [20]. In the same paper the situations in which the Wannier functions could decay slower than the kernel of the projector are pointed out, which shows that choosing the optimal phase is not a trivial task. Our results show that by choosing the right phase one must always obtain an optimal decay.

Motivated by the great interest in nonperiodic structures much effort has been devoted to extend the results about the existence of exponentially localized bases for isolated bands in nonperiodic systems. The basic difficulty stems from the fact that for nonperiodic systems one cannot define Wannier functions as Fourier transforms of the Bloch functions. One way out of the difficulty is to start from the periodic case or tight-binding limit where the Wannier functions are known to exist and use perturbation or 'continuity' arguments. The basic idea is that since the obstructions are of topological origin the existence of exponentially localized WF is stable against perturbations. Indeed, along these lines it has been possible to prove the existence of (generalized) WF for a variety of nonperiodic systems [2, 16, 21–23]. Since in the periodic case the obstructions to the existence of exponentially localized WF are absent [13–15] in one dimension it was natural to conjecture [16, 24] that in one dimension WF exist for all isolated bands irrespective of periodicity properties.

The first problem to be solved was to find an alternative definition of WF. The basic idea goes back to Kivelson [25], who proposed to *define* the generalized WF as the eigenfunctions of the 'band position' operator. To substantiate the idea one has to prove that the band position operator is self-adjoint, has discrete spectrum and its eigenfunctions are exponentially localized. For the particular case of a periodic one-dimensional crystal with one defect Kivelson proved that the eigenfunctions of the band position operator are indeed exponentially localized and asked for a general proof. In the general case, by a bootstrap argument, Niu [24] argued that the eigenfunctions of the band position operator (if they exist) are at least polynomially localized. In full generality the fact that for all isolated parts of the spectrum the band position operator is self-adjoint, has discrete spectrum and its eigenfunctions are exponentially localized has been proved in [26].

In this paper, we extend the results in [26] to quasi one-dimensional systems, i.e., three-dimensional systems for which the motion extends to infinity only in one direction. In addition, we add the result (which is new even in the strictly one-dimensional case) that (see theorem 2 for details) the 'density' of WF is uniformly bounded. While the main ideas of the proof are the same as in [26] there are major differences both at the technical and physical levels. In particular for quasi one-dimensional systems with screw symmetry the constructed WF inherits this symmetry, a property which is very useful in computational applications. Finally let us point out that as in the periodic case, generalized WF defined as the eigenfunctions of the band position operator have very nice properties, e.g., they are (up to uninteresting phases) uniquely defined and for real (i.e. time-reversal invariant) Hamiltonians they can be chosen to be real functions and this solves for the general quasi one-dimensional case the 'strong conjecture' in section V of [3]. As for their exponential localization we have the following 'optimality' result (see proposition 3 for a precise statement) which seems to be new even in the one-dimensional periodic case: the eigenfunctions of the band position operator have the same exponential localization as the integral kernel of the projection operator of the corresponding band.

2. The results

Consider in $L^2(\mathbb{R}^3)$ the following Hamiltonian describing a particle subjected to a scalar potential V:

$$H = \mathbf{P}^2 + V, \qquad \mathbf{P} = -\mathrm{i}\nabla, \qquad \sup_{\mathbf{x}\in\mathbb{R}^3} \int_{|\mathbf{x}-\mathbf{y}|\leqslant 1} |V(\mathbf{y})|^2 \, \mathrm{d}\mathbf{y} < \infty \tag{2.1}$$

which, as is well known (see [27]), is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3)$. We have already said in the introduction that we are interested in potentials V which tend to zero as the distance from the Ox_1-axis tends to infinity. Let us now be more precise. The notation $\mathbf{x} = (x_1, \mathbf{x}_\perp)$ will be used throughout the paper. For any $R > 0$, define

$$I_V(R) := \sup_{x_1\in\mathbb{R}, |\mathbf{x}_\perp|\geqslant R} \int_{|\mathbf{x}-\mathbf{y}|\leqslant 1} |V(\mathbf{y})|^2 \, \mathrm{d}\mathbf{y}. \tag{2.2}$$

The decay assumption for V will be

$$\lim_{R\to\infty} I_V(R) = 0. \tag{2.3}$$

It is easy to see that $[0, \infty) \subset \sigma(H)$ (using a Weyl sequence argument), thus the only region where H might have an isolated spectral island is below zero. Now suppose that σ_0 is such an isolated part of the spectrum and define

$$-E_+ := \sup\{E : E \in \sigma_0\} < 0. \tag{2.4}$$

If Γ is a positively oriented contour of finite length enclosing σ_0, then the spectral subspace corresponding to σ_0 is

$$\mathcal{K} := \mathrm{Ran}(P_0), \qquad P_0 = \frac{\mathrm{i}}{2\pi} \int_\Gamma (H - z)^{-1} \, \mathrm{d}z. \tag{2.5}$$

At a heuristic level, due to the fact that the wave packets from $\mathcal{K}$ cannot propagate in the classically forbidden region (see (2.4) and (2.3), at negative energies the motion is confined near the Ox_1-axis, i.e. the system has a quasi one-dimensional behavior.

2.1. *The technical results*

The following proposition states the 'localization' properties of P_0. On the one hand, this gives a precise meaning to the previously discussed quasi one-dimensional character, and on the other hand it provides some key ingredients to the proof of exponential localization of the eigenfunctions of the band position operator.

Let $a \in \mathbb{R}$, and let $\langle X_{\parallel,a}\rangle$ be the multiplication operator corresponding to

$$g_a(\mathbf{x}) := \sqrt{(x_1 - a)^2 + 1}, \tag{2.6}$$

and $\langle X_\perp\rangle$ be the multiplication operator given by

$$g_\perp(\mathbf{x}) := \sqrt{|\mathbf{x}_\perp|^2 + 1}. \tag{2.7}$$

Proposition 1. *There exist* $\alpha_\parallel > 0, \alpha_\perp > 0, M < \infty$ *such that*

$$\sup_{a\in\mathbb{R}} \| \mathrm{e}^{\alpha_\parallel\langle X_{\parallel,a}\rangle} P_0 \, \mathrm{e}^{-\alpha_\parallel\langle X_{\parallel,a}\rangle} \| \leqslant M \tag{2.8}$$

and

$$\| \mathrm{e}^{\alpha_\perp\langle X_\perp\rangle} P_0 \, \mathrm{e}^{\alpha_\perp\langle X_\perp\rangle} \| \leqslant M. \tag{2.9}$$

The proof of proposition 1 will also give values for $\alpha_\parallel$ and $\alpha_\perp$. In particular, $\alpha_\perp$ can be any number strictly smaller than $\sqrt{E_+}$.

We now can formulate the main technical result of this paper. To emphasize its generality we stress that its proof only uses the decay condition (2.3) and the existence of an isolated part of the spectrum satisfying (2.4).

Theorem 2. *Let $X_\parallel$ be the operator of multiplication with x_1 in $L^2(\mathbb{R}^3)$ and consider in $\mathcal{K}$ the operator*

$$\hat{X}_\parallel := P_0 X_\parallel P_0 \tag{2.10}$$

defined on $\mathcal{D}(\hat{X}_\parallel) = \mathcal{D}(X_\parallel) \cap \mathcal{K}$. Then

(i) $\hat{X}_\parallel$ is self-adjoint on $\mathcal{D}(\hat{X})$;

(ii) $\hat{X}_\parallel$ has purely discrete spectrum;

(iii) Let $g \in G := \sigma(\hat{X}_\parallel)$ be an eigenvalue, m_g its multiplicity and $\{W_{g,j}\}_{1\leqslant j\leqslant m_g}$ an orthonormal basis in the eigenspace of $\hat{X}$ corresponding to g. Then for all $\beta \in [0, 1]$, there exists $M_1 < \infty$ independent of g, j and β such that

$$\int_{\mathbb{R}^3} \mathrm{e}^{2(1-\beta)\alpha_\parallel |x_1-g|}\, \mathrm{e}^{2\beta\alpha_\perp |\mathbf{x}_\perp|} |W_{g,j}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} \leqslant M_1, \tag{2.11}$$

where $\alpha_\parallel$ and $\alpha_\perp$ are the same exponents as those provided by the proof of proposition 1;

(iv) Let $a \in \mathbb{R}$ and $L \geqslant 1$. Denote by $N(a, L)$ the total multiplicity of the spectrum of $\hat{X}_\parallel$ contained in $[a - L, a + L]$. Then there exists $M_2 < \infty$ such that

$$N(a, L) \leqslant M_2 \cdot L. \tag{2.12}$$

Finally, we turn to the question of optimal localization properties of our Wannier functions. Theorem 2 provides an optimal exponential decay in the transverse direction, but in the parallel direction it only implies a decay which is bound by the maximal decay of the resolvent in the gap. The conjecture on optimal exponential decay, as stated in section V of [3], is whether $W_{g,j}$'s have the same exponential decay as the integral kernel $\mathcal{P}_0(\mathbf{x}, \mathbf{y})$ of P_0 (which can be larger than the maximal decay of the resolvent in the gap; we are indebted to one of the referees for pointing this to us). Concerning this issue, we have the following result showing the optimality of the 'parallel' decay of $W_{g,j}$ at the exponential level.

Proposition 3. *Assume that for all $\alpha < \alpha_0$ we are given an a priori bound*

$$\sup_{a\in\mathbb{R}} \| \mathrm{e}^{\alpha\langle X_{\parallel,a}\rangle} P_0\, \mathrm{e}^{-\alpha\langle X_{\parallel,a}\rangle} \| < \infty. \tag{2.13}$$

Then for all $\alpha < \alpha_0$ there exists $M_1(\alpha)$, independent of g and j, such that

$$\int_{\mathbb{R}^3} \mathrm{e}^{2\alpha|x_1-g|} |W_{g,j}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} \leqslant M_1(\alpha). \tag{2.14}$$

Remark. Here α_0 is the 'exact' exponential decay of $\mathcal{P}_0(\mathbf{x}, \mathbf{y})$. In certain particular periodic cases one might obtain a power-like asymptotic behavior of $\mathrm{e}^{\alpha_0|x_1-y_1|}\mathcal{P}_0(\mathbf{x}, \mathbf{y})$ with the variables x_1, y_1. We cannot say anything about an eventual asymptotic behavior of $\mathrm{e}^{\alpha_0|x_1-g|} W_{g,j}(\mathbf{x})$. But due to the generality of the setting, we consider our result to be optimal.

2.2. Further properties of the Wannier basis

We come now to the case when V (hence H) has additional symmetries. The point here is that although the Wannier functions are not eigenfunctions of H, one would like them to inherit in some sense the symmetries of H. The reason is that usually the Wannier basis is used in order to write an effective Hamiltonian in $\mathcal{K}$, and one would like this effective Hamiltonian to inherit as much as possible the symmetries of H.

First, we comment on time-reversal invariance. Since $V(\mathbf{x})$ is real, H commutes with the anti-unitary operator induced by complex conjugation. It follows (see (2.5)) that P_0 and

$\hat{X}_\parallel$ are also real, thus the eigenfunctions of $\hat{X}_\parallel$ can be chosen to be real. Hence, theorem 2 provides us with a Wannier basis which is time-reversal invariant.

Second, we consider the so-called screw symmetry along the Ox_1-axis, of much interest in the physics of carbon nanotubes. Namely, writing

$$\mathbf{x}_\perp = (r, \theta), \qquad r \geqslant 0, \qquad \theta \in [0, 2\pi), \tag{2.15}$$

one assumes that for some $\theta_0 \in [0, 2\pi)$ we have

$$V(x_1, r, \theta) = V(x_1 + 1, r, \theta + \theta_0). \tag{2.16}$$

Here $\theta + \theta_0$ has to be understood as modulo 2π. Defining the screw-symmetry operators $T_n^{\theta_0}$ by

$$\left(T_n^{\theta_0} f\right)(x_1, r, \theta) := f(x_1 - n, r, \theta - n\theta_0), \tag{2.17}$$

one has a (unitary!) representation of $\mathbb{Z}$ in $L^2(\mathbb{R}^3)$. Taking into account (2.16) and the fact that $\left[-\Delta, T_n^{\theta_0}\right] = 0$ (use cylindrical coordinates to prove this), one obtains

$$\left[H, T_n^{\theta_0}\right] = 0, \tag{2.18}$$

and then from functional calculus and (2.5)

$$\left[P_0, T_n^{\theta_0}\right] = 0. \tag{2.19}$$

In particular, this implies that the family $\left\{T_n^{\theta_0}\right\}_{n\in\mathbb{Z}}$ induces a unitary representation of $\mathbb{Z}$ in $\mathcal{K}$. Moreover, from (2.10) and (2.19) one obtains

$$\left[T_n^{\theta_0}, \hat{X}_\parallel\right] = nT_n^{\theta_0}. \tag{2.20}$$

Let $p < \infty$ be the number of eigenvalues of $\hat{X}_\parallel$ in the interval $[0, 1)$, and let $\{g_j\}_{j=1}^p$ be the distinct eigenvalues (each with multiplicity $m_j < \infty$). We have

$$\hat{X}_\parallel W_{g_j,\alpha_j} = g_j W_{g_j,\alpha_j}, \qquad \alpha_j = 1, 2, \ldots, m_{g_j}. \tag{2.21}$$

From (2.20)) and (2.21) one obtains that for all $g_j, \alpha_j, n \in \mathbb{Z}$:

$$\hat{X}_\parallel T_n^{\theta_0} W_{g_j,\alpha_j} = (g_j + n) T_n^{\theta_0} W_{g_j,\alpha_j}. \tag{2.22}$$

Conversely, for every other $g \in \sigma(\hat{X}_\parallel)$, choose an eigenvector W_g. We can find $n \in \mathbb{Z}$ such that $g + n \in [0, 1)$. Since $\hat{X}_\parallel T_n^{\theta_0} W_g = (g + n) T_n^{\theta_0} W_g$, it means that $g + n$ must be one of the g_j's considered above. Therefore we proved the following corollary:

Corollary 4. *The spectrum of $\hat{X}_\parallel$ consists of a union of p ladders:*

$$G = \cup_{j=1}^p G_j, \qquad G_j = \{g : g = g_j + n, n \in \mathbb{Z}\}, \qquad j \in \{1, 2, \ldots, p\}, \tag{2.23}$$

and an orthonormal basis in $\mathcal{K}$ can be chosen as

$$\begin{aligned} &W_{n,g_j,\alpha_j} := W_{g_j+n,\alpha_j} := T_n^{\theta_0} W_{g_j,\alpha_j}, \\ &n \in \mathbb{Z}, \ j \in \{1, 2, \ldots, p\}, \alpha_j \in \{1, 2, \ldots, m_{g_j}\}. \end{aligned} \tag{2.24}$$

It is interesting to express the effective Hamiltonian $P_0 H P_0$ as an infinite matrix with the help of the Wannier basis. For notational simplicity, we relabel the pair (g_j, α_j) as $l \in \left\{1, 2, \ldots, N_c = \sum_{j=1}^p m_{g_j}\right\}$ and write the Wannier basis as $\{W_{n,l}\}_{n\in\mathbb{Z}, l\in\{1,2,\ldots,N_c\}}$. Note that N_c is nothing but the number of Wannier functions per unit cell $[0, 1)$. Let

$$h_{l,k}^{\theta_0}(m, n) := \langle W_{m,l}, H W_{n,k} \rangle. \tag{2.25}$$

The important fact is that in spite of a rotation with an angle θ_0 for which it might happen that $\frac{\theta_0}{2\pi}$ to be irrational, from (2.18) and (2.24) one obtains (with the usual abuse of notation)

$$h_{l,k}^{\theta_0}(m, n) = h_{l,k}^{\theta_0}(m - n). \tag{2.26}$$

Then a standard computation gives the effective Hamiltonian as an operator in $(l^2)^{N_c}$ which is of *standard translation invariant* tight-binding type:

$$\left(h_{\mathrm{eff}}^{\theta_0}\phi\right)_l(m) := \sum_{k,n} h_{l,k}^{\theta_0}(m-n)\phi_k(n). \tag{2.27}$$

This is another consequence of the quasi one-dimensional character of the motion for negative energies. More precisely, it reflects the fact that for arbitrary values of θ_0, since $T_n^{\theta_0}$ is a unitary representation of $\mathbb{Z}$, one can still develop a Bloch-type analysis but with a more complicated form of 'Bloch' functions:

$$\Psi_k(\mathbf{x}) = \mathrm{e}^{\mathrm{i}kx_1}u_k(\mathbf{x}), \qquad u_k(\mathbf{x}) = T_n^{\theta_0}u_k(\mathbf{x}). \tag{2.28}$$

However, due to the complicated symmetry of the resulting Bloch functions (which does not allow us to represent the fiber Hamiltonian as a differential operator on the unit cell with 'simple' boundary conditions), the analysis gets much harder. The Bloch analysis reduces to the standard one (with a larger unit cell) for rational values of $\frac{\theta_0}{2\pi}$.

3. Proofs

This section is devoted to the proof of proposition 1, theorem 2 and proposition 3. A certain number of unimportant finite positive constants appearing during the proof will be denoted by M.

One of the key ingredients in the proofs is the exponential decay of the integral kernel of the resolvent of Schrödinger operators. This is an elementary result in the Combes–Thomas–Agmon theory of weighted estimates. We summarize the needed result in

Lemma 5. *Let W be a potential such that $\sup_{\mathbf{x}\in\mathbf{R}^3}\int_{|\mathbf{x}-\mathbf{y}|\leqslant 1}|W(\mathbf{y})|^2\,\mathrm{d}\mathbf{y} < \infty$. Define $K := \mathbf{P}^2 + W(\mathbf{x})$ as an operator sum, and let h be a real function satisfying*

$$h \in C^\infty(\mathbb{R}^3), \qquad \sup_{\mathbf{x}\in\mathbb{R}^3}\{|\nabla h(\mathbf{x})| + |\Delta h(\mathbf{x})|\} = m < \infty. \tag{3.1}$$

Fix $z \in \rho(H)$. Then there exists $\alpha_z > 0$ such that

$$\|\,\mathrm{e}^{\alpha_z h}(K-z)^{-1}\,\mathrm{e}^{-\alpha_z h}\| \leqslant M, \tag{3.2}$$

$$\|\,\mathrm{e}^{\alpha_z h}P_j(K-z)^{-1}\,\mathrm{e}^{-\alpha_z h}\| \leqslant M, \tag{3.3}$$

where $P_j = -\mathrm{i}\frac{\partial}{\partial x_j}$, $j \in \{1,2,3\}$.

Without giving the details of the proof of lemma 5, for later use we write a key identity in (3.5): under the condition

$$1 + \alpha_z(\pm\mathrm{i}\mathbf{P}\cdot\nabla h \pm \mathrm{i}\nabla h\cdot\mathbf{P} - \alpha_z|\nabla h|^2)(K-z)^{-1} \qquad \text{invertible} \tag{3.4}$$

one has

$$\mathrm{e}^{\pm\alpha_z h}(K-z)^{-1}\,\mathrm{e}^{\mp\alpha_z h} = (K-z)^{-1}[1 + \alpha_z(\pm\mathrm{i}\mathbf{P}\cdot\nabla h \pm \mathrm{i}\nabla h\cdot\mathbf{P} - \alpha_z|\nabla h|^2)(K-z)^{-1}]^{-1}. \tag{3.5}$$

Then (3.4) holds true if for example $\alpha_z > 0$ is small enough.

3.1. Proof of proposition 1

Consider Γ in (2.5) as a contour of finite length enclosing σ_0 and satisfying

$$\mathrm{dist}(\Gamma, \sigma(H)) = \tfrac{1}{2}\mathrm{dist}(\sigma_0, \sigma(H)\backslash\sigma_0). \tag{3.6}$$

Then since $|\nabla g_a| \leqslant 1$, $|\Delta g_a|^2 \leqslant 2$, the estimate (2.8) follows directly from lemma 5 by taking $\alpha_\parallel$ sufficiently small such that for all $z \in \Gamma$

$$\|\alpha_\parallel(\mathrm{i}\mathbf{P}\cdot\nabla g_a + \mathrm{i}\nabla g_a\cdot\mathbf{P} - \alpha_\parallel|\nabla g_a|^2)(K-z)^{-1}\| \leqslant b < 1.$$

We now prove (2.9). If $R > 0$, define

$$H_R = -\Delta + (1-\chi_R)V, \tag{3.7}$$

where

$$\chi_R(\mathbf{x}) = \begin{cases} 1 & \text{for} \quad |\mathbf{x}_\perp| \leqslant R \\ 0 & \text{for} \quad |\mathbf{x}_\perp| > R \end{cases}. \tag{3.8}$$

From (2.3) it follows that

$$\lim_{R\to\infty} \inf \sigma(H_R) = 0.$$

In particular, for sufficiently large R, $(H_R - z)^{-1}$ is analytic inside Γ. Since $H - H_R = \chi_R V$, then using resolvent identities we obtain

$$\begin{aligned}(H-z)^{-1} = {} & (H_R - z)^{-1} - (H_R-z)^{-1}\chi_R V (H_R - z)^{-1} \\ & + (H_R-z)^{-1}\chi_R V(H-z)^{-1}\chi_R V(H_R-z)^{-1}.\end{aligned} \tag{3.9}$$

From (2.5), (3.9) and the fact that $(H_R - z)^{-1}$ is analytic inside Γ one has

$$P_0 = \frac{\mathrm{i}}{2\pi}\int_\Gamma (H_R-z)^{-1}\chi_R V(H-z)^{-1}\chi_R V(H_R-z)^{-1}. \tag{3.10}$$

Note that for all $\alpha > 0$

$$\sup_{\mathbf{x}\in\mathbf{R}^3}\int_{|\mathbf{x}-\mathbf{y}|\leqslant 1} |(\mathrm{e}^{\alpha g_\perp}\chi_R V)(\mathbf{y})|^2\,\mathrm{d}\mathbf{y} < \infty. \tag{3.11}$$

Take now $\alpha_\perp > 0$ such that (3.4) holds true for all $z \in \Gamma$, $K = H_R$, $h = g_\perp$ and $\alpha_z = \alpha_\perp$. That is let us suppose that

$$1 + \alpha_\perp(\pm\mathrm{i}\mathbf{P}\cdot\nabla g_\perp \pm \mathrm{i}\nabla g_\perp\cdot\mathbf{P} - \alpha_\perp|\nabla g_\perp|^2)(H_R - z)^{-1} \qquad \text{is invertible} \tag{3.12}$$

uniformly on Γ. Then we can rewrite P_0 as

$$\begin{aligned} P_0 = \mathrm{e}^{-\alpha_\perp\langle X_\perp\rangle}\Bigg\{ & \frac{\mathrm{i}}{2\pi}\int_\Gamma [\mathrm{e}^{\alpha_\perp\langle X_\perp\rangle}(H_R - z)^{-1}\,\mathrm{e}^{-\alpha_\perp\langle X_\perp\rangle}] \\ & \times\ [\mathrm{e}^{\alpha_\perp g_\perp}\chi_R V(H-z)^{-1}][\mathrm{e}^{\alpha_\perp g_\perp}\chi_R V(H_R-z)^{-1}] \\ & \times\ [1 + \alpha_\perp(-\mathrm{i}\mathbf{P}\cdot\nabla g_\perp - \mathrm{i}\nabla g_\perp\cdot\mathbf{P} - \alpha_\perp|\nabla g_\perp|^2)(H_R-z)^{-1}]^{-1}\,\mathrm{d}z\Bigg\}\,\mathrm{e}^{-\alpha_\perp\langle X_\perp\rangle}. \end{aligned} \tag{3.13}$$

Due to (3.11) the operator under the integral sign is uniformly bounded in z and the proof of proposition 1 is completed provided we can show why we can choose $\alpha_\perp$ as close to $\sqrt{E_+}$ as we want. The argument is as follows. Choose $0 \leqslant \alpha_\perp < \sqrt{E_+}$. Choose a contour Γ which is very close to σ_0, at a distance $\delta > 0$, infinitesimally small. Using the spectral theorem (or in this case the Plancherel theorem), there exists δ small enough such that the following estimates hold true:

$$\sup_{z\in\Gamma}\|(\mathbf{P}^2 - z)^{-1}\| \leqslant \text{const}, \qquad \sup_{z\in\Gamma}\max_{j\in\{1,2,3\}}\|P_j(\mathbf{P}^2-z)^{-1}\| \leqslant \text{const}. \tag{3.14}$$

Hence, we can find δ small enough and R large enough such that the operator in (3.12) is invertible if

$$1 + \alpha_\perp(\pm \mathrm{i}\mathbf{P}\cdot\nabla g_\perp \pm \mathrm{i}\nabla g_\perp\cdot\mathbf{P} - \alpha_\perp|\nabla g_\perp|^2)(\mathbf{P}^2 - \mathrm{Re}(z))^{-1} \quad \text{is invertible} \tag{3.15}$$

uniformly on Γ. Now the operator in (3.15) is invertible if

$$\begin{aligned} 1 \pm \mathrm{i}\alpha_\perp(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}(\mathbf{P}\cdot\nabla h + \nabla h\cdot\mathbf{P})(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}} \\ - \alpha_\perp^2(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}|\nabla h|^2(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}} \end{aligned} \tag{3.16}$$

is invertible (by a resummation of the Neumann series and analytic continuation). Now assume that uniformly on Γ we have

$$0 < \alpha_\perp^2(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}|\nabla h|^2(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}} \leqslant \frac{\alpha_\perp^2}{-\mathrm{Re}(z)} < 1,$$

which can be achieved if $\alpha_\perp^2 < E_+$ and δ is chosen to be small enough. Define

$$S := \left(1 - \alpha_\perp^2(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}|\nabla h|^2(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}\right)^{-\frac{1}{2}},$$

and

$$T = T^* := S(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}(\mathbf{P}\cdot\nabla h + \nabla h\cdot\mathbf{P})(\mathbf{P}^2 - \mathrm{Re}(z))^{-\frac{1}{2}}S.$$

Then the operator in (3.16) is invertible if $1 \pm \mathrm{i}\alpha_\perp T$ is invertible, which is always the case:

$$(1 \pm \mathrm{i}\alpha_\perp T)^{-1} = (1 \mp \mathrm{i}\alpha_\perp T)(1 + \alpha_\perp^2 T^2)^{-1}.$$

Therefore proposition 1 is proved.

3.2. Proof of theorem 2

Proof of (i). First we recall an older result (see, e.g., [2, 28, 29]), according to which the commutator $[X_\parallel, P_0]$ defined on $\mathcal{D}(X_\parallel)$ has a bounded closure on $L^2(\mathbb{R}^3)$. We seek an approximate resolvent of $\hat{X}_\parallel$ by defining for $\mu > 0$ the operator

$$\hat{R}_{\pm\mu} = P_0(X_\parallel \pm \mathrm{i}\mu)^{-1}P_0. \tag{3.17}$$

Since one can rewrite $\hat{R}_{\pm\mu}$ as

$$\hat{R}_{\pm\mu} = (X_\parallel \pm \mathrm{i}\mu)^{-1}P_0 + (X_\parallel \pm \mathrm{i}\mu)^{-1}[X_\parallel, P_0](X_\parallel \pm \mathrm{i}\mu)^{-1}P_0$$

it follows that $\hat{R}_{\pm\mu}\mathcal{K} \subset D(\hat{X}_\parallel)$ and by a straightforward computation (as operators in $\mathcal{K}$)

$$(\hat{X}_\parallel \pm \mathrm{i}\mu)\hat{R}_{\pm\mu} = P_0(X_\parallel \pm \mathrm{i}\mu)P_0(X_\parallel \pm \mathrm{i}\mu)^{-1}P_0 = 1_{\mathcal{K}} + \hat{A}_{\pm\mu} \tag{3.18}$$

with

$$\hat{A}_{\pm\mu} = P_0[X_\parallel, P_0](X_\parallel \pm \mathrm{i}\mu)^{-1}P_0. \tag{3.19}$$

Since $[X_\parallel, P_0]$ is bounded and $\|(X_\parallel \pm \mathrm{i}\mu)^{-1}\| \leqslant \frac{1}{\mu}$, it follows that for sufficiently large μ:

$$\|\hat{A}_{\pm\mu}\| \leqslant \tfrac{1}{2}. \tag{3.20}$$

Then again as operators in $\mathcal{K}$:

$$(\hat{X} \pm \mathrm{i}\mu)\hat{R}_{\pm\mu}(1_{\mathcal{K}} + \hat{A}_{\pm\mu})^{-1} = 1_{\mathcal{K}}. \tag{3.21}$$

This implies that $\hat{X} \pm \mathrm{i}\mu$ is surjective on $\hat{R}_{\pm\mu}(1_{\mathcal{K}} + \hat{A}_{\pm\mu})^{-1}\mathcal{K} \subset D(\hat{X})$. By the fundamental criterion of self-adjointness [27] $\hat{X}$ is self-adjoint in $\mathcal{K}$ on $\mathcal{D}(\hat{X})$. In addition, from (3.21) one obtains the following formula for the resolvent of $\hat{X}_\parallel$:

$$(\hat{X}_\parallel \pm \mathrm{i}\mu)^{-1} = \hat{R}_{\pm\mu}(1_{\mathcal{K}} + \hat{A}_{\pm\mu})^{-1}. \tag{3.22}$$

Proof of (ii). We will show that $\hat{R}_{\pm\mu}$ is compact in $\mathcal{K}$ which implies (see (3.22) that $\hat{X}_{\|}$ has compact resolvent, thus purely discrete spectrum. Consider a cut-off function ϕ_N which equals 1 if $|\mathbf{x}| \leqslant N$ and is zero if $|\mathbf{x}| \geqslant 2N$. For $N \geqslant 1$ we can decompose

$$\hat{R}_{\pm\mu} = P_0(X_{\|} \pm \mathrm{i}\mu)^{-1}\phi_N P_0 + P_0(X_{\|} \pm \mathrm{i}\mu)^{-1}(1-\phi_N)P_0. \tag{3.23}$$

Writing

$$\phi_N P_0 = \{\phi_N(\mathbf{P}^2+1)^{-1}\}\{(\mathbf{P}^2+1)P_0\}$$

we see that $\phi_N P_0$ is compact (even Hilbert–Schmidt) in $L^2(\mathbb{R}^3)$ (the first factor is Hilbert–Schmidt while the second one is bounded). Now if $0 < \alpha$ is small enough, we know that $\mathrm{e}^{\alpha g_\perp} P_0$ is bounded (see (2.9)). Since

$$\lim_{N\to\infty} \|(X_{\|} \pm \mathrm{i}\mu)^{-1}(1-\phi_N)\,\mathrm{e}^{-\alpha g_\perp}\| = 0,$$

we have shown

$$\lim_{N\to\infty} \|\hat{R}_{\pm\mu} - P_0(X_{\|} \pm \mathrm{i}\mu)^{-1}\phi_N P_0\| = 0,$$

thus $\hat{R}_{\pm\mu}$ equals the norm limit of a sequence of compact operators, therefore it is compact. Accordingly, since the self-adjoint operator $\hat{X}_{\|}$ has compact resolvent it has purely discrete spectrum [27]:

$$\sigma(\hat{X}_{\|}) = \sigma_{\mathrm{disc}}(\hat{X}_{\|}) =: G, \tag{3.24}$$

and the proof of the second part of theorem 2 is completed.

Proof of (iii). Now we will consider the exponential localization of the eigenfunctions of $\hat{X}_{\|}$. Let $g \in G$ be an eigenvalue, m_g its multiplicity and $W_{g,j}$, $1 \leqslant j \leqslant m_g$, be an orthonormal basis in the eigenspace of $\hat{X}_{\|}$ corresponding to g. We shall prove that uniformly in g and j

$$\|\,\mathrm{e}^{\alpha_{\|}\langle X_{\|},g\rangle} W_{g,j}\| \leqslant M \tag{3.25}$$

and

$$\|\,\mathrm{e}^{\alpha_\perp\langle X_\perp\rangle} W_{g,j}\| \leqslant M. \tag{3.26}$$

Taking (3.25) and (3.26) as given, one can easily obtain (2.11) by a simple convexity argument: the function $f(x) = a^{1-x}b^x$, $a, b > 0$, is convex on $\mathbb{R}$, and for $0 \leqslant \beta \leqslant 1$ one has

$$\beta\,\mathrm{e}^{2\alpha_{\|}g_a(\mathbf{x})} + (1-\beta)\,\mathrm{e}^{2\alpha_\perp g_\perp} \geqslant \mathrm{e}^{2(1-\beta)\alpha_{\|}g_a(\mathbf{x})}\,\mathrm{e}^{2\beta\alpha_\perp g_\perp}, \tag{3.27}$$

which together with (3.25) and (3.26) proves (2.11) with $M_1 = M^2$. Since (3.26) follows directly from (2.9) and $W_{g,j} = P_0 W_{g,j}$ we are left with the proof of (3.25).

Although the proof of (3.25) mimics closely the proof in the one-dimensional case [26], we give it here for completeness. In order to emphasize the main idea of the proof let us recall one of the simplest proofs of the exponential decay of the eigenfunctions of Schrödinger operators corresponding to discrete eigenvalues (assuming that the potential V is bounded and has compact support). Namely assume that for some $E > 0$ we have $(-\Delta + V + E)\Psi = 0$, which can be rewritten as

$$\Psi = -(-\Delta + E)^{-1}V\Psi. \tag{3.28}$$

Since for $|\alpha| < \sqrt{E}$, $\mathrm{e}^{\alpha|\cdot|}(-\Delta+E)^{-1}\,\mathrm{e}^{-\alpha|\cdot|}$ and $\mathrm{e}^{\alpha|\cdot|}V$ are bounded:

$$\Psi = -\mathrm{e}^{-\alpha|\cdot|}\{\mathrm{e}^{\alpha|\cdot|}(-\Delta+E)^{-1}\,\mathrm{e}^{-\alpha|\cdot|}\}\,(\mathrm{e}^{\alpha|\cdot|}V)\Psi$$

which proves the exponential localization of Ψ. The main idea in proving (3.25) is to rewrite the eigenvalue equation for $\hat{X}_{\|}$ in a form similar to (3.28) and then to use (2.8).

Let us start with some notation. If $b > 0$ (sufficiently large) and $a \in \mathbb{R}$, define

$$f_{a,b}(\mathbf{x}) := bf\left(\frac{x_1 - a}{b}\right) \tag{3.29}$$

where f is a real $C_0^\infty(\mathbb{R})$ cut-off function satisfying $0 \leqslant f(y) \leqslant 1$ and

$$f(y) = \begin{cases} 1 & \text{for} \quad |y| \leqslant \frac{1}{2} \\ 0 & \text{for} \quad |y| \geqslant 1 \end{cases}.$$

Define the function $h_{a,b}$ by

$$h_{a,b}(\mathbf{x}) := x_1 - a + \mathrm{i} f_{a,b}(\mathbf{x}). \tag{3.30}$$

Note that by construction, $h_{a,b}$ depends only on x_1 and obeys

$$|h_{a,b}(\mathbf{x})| \geqslant \frac{b}{2}. \tag{3.31}$$

Moreover, its first two derivatives are uniformly bounded:

$$\sup_{\mathbf{x}\in\mathbb{R}^3} \sup_{a\in\mathbb{R}} \sup_{b\geqslant 1} \{|\nabla h_{a,b}(\mathbf{x})| + |\Delta h_{a,b}(\mathbf{x})|\} = K < \infty. \tag{3.32}$$

The eigenvalue equation for $W_{g,j}$ reads $P_0(\hat{X}_\parallel - g)P_0 W_{g,j} = 0$. Using (3.30) it can be rewritten as

$$P_0 h_{g,b} P_0 W_{g,j} = \mathrm{i} P_0 f_{g,b} P_0 W_{g,j}. \tag{3.33}$$

We now prove that $P_0 h_{g,b} P_0$ is invertible. Like in the proof of self-adjointness of $\hat{X}_\parallel$ we compute

$$P_0 h_{g,b}^{-1} P_0 P_0 h_{g,b} P_0 = 1_{\mathcal{K}} + P_0 h_{g,b}^{-1} [P_0, h_{g,b}] P_0. \tag{3.34}$$

The key remark is that $\left[P_0, h_{g,b}\right]$ is bounded. Indeed, we have the identity

$$\begin{aligned} [P_0, h_{g,b}] &= -\frac{1}{2\pi} \int_\Gamma (H - z)^{-1} \{\mathbf{P} \cdot \nabla h_{g,b} + \nabla h_{g,b} \cdot \mathbf{P}\} (H - z)^{-1} \, \mathrm{d}z \\ &= -\frac{1}{2\pi} \int_\Gamma (H - z)^{-1} \{-\mathrm{i}\Delta h_{g,b} + 2\nabla h_{g,b} \cdot \mathbf{P}\} (H - z)^{-1} \, \mathrm{d}z. \end{aligned} \tag{3.35}$$

It follows that $[P_0, h_{g,b}]$ is uniformly bounded in $g \in \mathbb{R}$ and $b \geqslant 1$ (see (3.32)). Taking into account (3.31) one obtains that the operator

$$\hat{B}_{g,b} = P_0 h_{g,b}^{-1} [P_0, h_{g,b}] P_0 : \mathcal{K} \to \mathcal{K} \tag{3.36}$$

satisfies

$$\|\hat{B}_{g,b}\| \leqslant \tfrac{1}{2} \tag{3.37}$$

if $b \geqslant b_0$ for some large enough $b_0 < \infty$. It follows that $1 + \hat{B}_{g,b}$ is invertible and then the eigenvalue equation (see (3.33), (3.34) and (3.36)) takes the form

$$W_{g,j} = \mathrm{i}(1 + \hat{B}_{g,b})^{-1} P_0 h_{g,b}^{-1} P_0 f_{g,b} P_0 W_{g,j} \tag{3.38}$$

which is the analog of (3.28). By construction (see the definition of $f_{g,b}$ in (3.29))

$$\| \, \mathrm{e}^{\alpha_\parallel \langle X_{\parallel,g} \rangle} f_{g,b} \| \leqslant b \, \mathrm{e}^{\alpha_\parallel (b+1)}.$$

Moreover,

$$\mathrm{e}^{\alpha_\parallel \langle X_{\parallel,g} \rangle} P_0 h_{g,b}^{-1} P_0 \, \mathrm{e}^{-\alpha_\parallel \langle X_{\parallel,g} \rangle} = \{\mathrm{e}^{\alpha_\parallel \langle X_{\parallel,g} \rangle} P_0 \, \mathrm{e}^{-\alpha_\parallel \langle X_{\parallel,g} \rangle}\} h_{g,b}^{-1} \{\mathrm{e}^{\alpha_\parallel \langle X_{\parallel,g} \rangle} P_0 \, \mathrm{e}^{-\alpha_\parallel \langle X_{\parallel,g} \rangle}\}$$

is bounded due to (2.8). Thus, the only thing that remains to be proved is the existence of a b large enough such that the following bound holds:

$$\sup_{g\in\mathbb{R}} \| \mathrm{e}^{\alpha_\| \langle X_{\|,g}\rangle} (1 + \hat{B}_{g,b})^{-1} \mathrm{e}^{-\alpha_\| \langle X_{\|,g}\rangle} \| < \infty. \tag{3.39}$$

Using the Neumann series for $(1 + \hat{B}_{g,b})^{-1}$, it follows that it suffices to prove that

$$\lim_{b\to\infty} \sup_{g\in\mathbb{R}} \| \mathrm{e}^{\alpha_\| \langle X_{\|,g}\rangle} \hat{B}_{g,b} \, \mathrm{e}^{-\alpha_\| \langle X_{\|,g}\rangle} \| = 0. \tag{3.40}$$

Since (see (3.31) $\lim_{b\to\infty} \left\| h_{g,b}^{-1} \right\| = 0$ (uniformly in $g \in \mathbb{R}$), for (3.40) to hold true it is sufficient to show

$$\sup_{g\in\mathbb{R}} \| \mathrm{e}^{\alpha_\| \langle X_{\|,g}\rangle} [P_0, h_{g,b}] \, \mathrm{e}^{-\alpha_\| \langle X_{\|,g}\rangle} \| \leqslant \text{const}. \tag{3.41}$$

But this easily follows from (3.35), (3.32), (3.2) and (3.3) where we take $K = H, \alpha_z = \alpha_\|$ and $h = g_g$. The proof of (iii) is concluded.

Proof of (iv). We start with a technical result:

Lemma 6. *Fix* $0 \leqslant \alpha_\perp < \sqrt{E_+}$. *Then there exists a bounded operator* D *such that*

$$P_0 = \mathrm{e}^{-\alpha_\perp \langle X_\perp \rangle} (\mathbf{P}^2 + 1)^{-1} D. \tag{3.42}$$

Proof. We use the notation and ideas of proposition 1, and we rewrite P_0 in a convenient form. First, for $R > 0$ we have

$$(H - z)^{-1} = (H_R - z)^{-1} - (H_R - z)^{-1} \chi_R V (H - z)^{-1}.$$

Second, choose Γ close enough to σ_0 and R large enough such that $(H_R - z)^{-1}$ becomes analytic inside Γ and (3.12) holds true for all $z \in \Gamma$. Then we can write

$$P_0 = -\mathrm{e}^{-\alpha_\perp \langle X_\perp \rangle} \frac{\mathrm{i}}{2\pi} \int_\Gamma (H_R - z)^{-1} [1 + \alpha_\perp (\mathrm{i}\mathbf{P} \cdot \nabla g_\perp + \mathrm{i} \nabla g_\perp \cdot \mathbf{P} - \alpha_\perp |\nabla g_\perp|^2)(H_R - z)^{-1}]^{-1}$$
$$\times \, \mathrm{e}^{\alpha_\perp g_\perp} \chi_R V (H - z)^{-1} \, \mathrm{d}z.$$

Now by the closed graph theorem we have that $(\mathbf{P}^2 + 1)(H_R + 1)^{-1}$ is bounded (here R is large enough such that $(-\infty, -1/2) \subset \rho(H_R)$), and together with the spectral theorem

$$\sup_{z\in\Gamma} \| (\mathbf{P}^2 + 1)(H_R - z)^{-1} \| < \infty.$$

Use this in (3.43) and we are done.

We now have all the necessary ingredients for proving the last statement of our theorem. For every $L > 0$ and $a \in \mathbb{R}$, denote by $\chi_{L,a}$ the characteristic function of the slab $\{\mathbf{x} : |x_1 - a| \leqslant L\}$. Then define the operator $B := \chi_{L,a} P_0$. Using (3.42) let us show that B is Hilbert–Schmidt and, moreover, uniformly in $a \in \mathbb{R}$ we have

$$\|B\|_2^2 \leqslant M \cdot L, \tag{3.44}$$

for some $M < \infty$. Indeed, since $B = \chi_{L,a} \, \mathrm{e}^{-\alpha_\perp \langle X_\perp \rangle} (-\Delta + 1)^{-1} D$, a direct computation using the explicit formula for the integral kernel of the free Laplacian gives

$$\| \chi_{L,a} \, \mathrm{e}^{-\alpha_\perp \langle X_\perp \rangle} (\mathbf{P}^2 + 1)^{-1} \|_2^2 \leqslant \text{const} \cdot L.$$

It follows that the operator $\chi_{L,a} P_0 \chi_{L,a} = B B^*$ is trace class and

$$|\mathrm{Tr}(\chi_{L,a} P_0 \chi_{L,a})| \leqslant \|B\|_2^2 \leqslant M \cdot L \tag{3.45}$$

for some $M < \infty$ independent of L and a.

Now let $P_0^{L,a}$ be the orthogonal projection onto the subspace spanned by those $W_{g,j}$ for which $g \in [a-L, a+L]$:

$$P_0^{L,a} := \sum_{|g-a|\leqslant L} \sum_{j=1}^{m_g} \langle \cdot, W_{g,j} \rangle W_{g,j}. \tag{3.46}$$

We can choose A sufficiently large such that (3.25) implies

$$\int_{|x_1-a|\geqslant A} |W_{g,j}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} \leqslant \frac{1}{2} \tag{3.47}$$

uniformly in a and $g \in [a-L, a+L]$. Since $P_0 \geqslant P_0^{L,a}$, from (3.45) one obtains

$$\begin{aligned} M \cdot (L+A) &\geqslant \mathrm{Tr}(\chi_{L+A,a} P_0 \chi_{L+A,a}) \geqslant \mathrm{Tr}\big(\chi_{L+A,a} P_0^{L,a} \chi_{L+A,a}\big) \\ &= \sum_{|g-a|\leqslant L} \sum_{j=1}^{m_g} \int_{\mathbb{R}^3} \chi_{L+A,a}(\mathbf{x}) |W_{g,j}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} \\ &\geqslant \sum_{|g-a|\leqslant L} \sum_{j=1}^{m_g} \frac{1}{2} = \frac{1}{2} N(a, L), \end{aligned} \tag{3.48}$$

where in the last inequality we used (3.47). In particular, if $L \geqslant 1$, then uniformly in $a \in \mathbb{R}$ we have

$$N(a, L) \leqslant 2M \cdot (1+A)L$$

and the proof is completed.

3.3. *Proof of proposition 3*

The only thing we have to prove is that (3.41) holds true for $\alpha_{\|}$ replaced by any $\alpha < \alpha_0$, where α_0 is the *a priori* given, 'exact' exponential localization.

We introduce the multiplication operator given by $\{\mathrm{e}^{\alpha|\cdot - t|} f\}(\mathbf{x}) := \mathrm{e}^{\alpha|x_1 - t|} f(\mathbf{x})$. We start by noticing that due to the bound $\mathrm{e}^{\pm\alpha(\sqrt{s^2+1}-|s|)} \leqslant \mathrm{e}^{\alpha}$ we can replace (2.13) with

$$\sup_{t\in\mathbb{R}} \| \mathrm{e}^{\alpha|\cdot-t|} P_0 \,\mathrm{e}^{-\alpha|\cdot-t|} \| < \infty. \tag{3.49}$$

The same replacement can be done in (3.41). Now the integral kernel $\mathcal{A}(\mathbf{x}, \mathbf{y})$ of the operator $A := \mathrm{e}^{\alpha|\cdot - g|}[P_0, h_{g,b}]\mathrm{e}^{-\alpha|\cdot - g|}$ equals

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = \mathcal{P}_0(\mathbf{x}, \mathbf{y}) \,\mathrm{e}^{\alpha(|x_1-g|-|y_1-g|)} (h_{g,b}(\mathbf{y}) - h_{g,b}(\mathbf{x})). \tag{3.50}$$

We consider A as an operator on $L^2(\mathbb{R}^3) = \bigoplus_{p\in\mathbb{Z}} L^2([p, p+1] \times \mathbb{R}^2)$. Let χ_p be the characteristic function of the slab $[p, p+1] \times \mathbb{R}^2$. We have that $A_{pp'} := \chi_p A \chi_{p'}$ is a bounded operator between $L^2([p', p'+1] \times \mathbb{R}^2)$ and $L^2([p, p+1] \times \mathbb{R}^2)$, and we can write $A = \{A_{pp'}\}_{p,p'\in\mathbb{Z}}$. We will bound the norm of A with a Schur–Holmgren-type estimate (see lemma 7):

$$\|A\| \leqslant \left(\sup_{p'\in\mathbb{Z}} \sum_{p\in\mathbb{Z}} \|A_{pp'}\| \right)^{\frac{1}{2}} \left(\sup_{p\in\mathbb{Z}} \sum_{p'\in\mathbb{Z}} \|A_{pp'}\| \right)^{\frac{1}{2}}. \tag{3.51}$$

For $0 \leqslant x_1, y_1 \leqslant 1$, the kernel of $A_{pp'}$ can be written as

$$\begin{aligned} \mathcal{A}_{pp'}(\mathbf{x}, \mathbf{y}) &= \mathcal{P}_0(x_1 + p, \mathbf{x}_\perp; y_1 + p', \mathbf{y}_\perp) \,\mathrm{e}^{\alpha(|x_1+p-g|-|y_1+p'-g|)} (h_{g,b}(y_1 + p') - h_{g,b}(x_1 + p)) \\ &= \mathcal{P}_0(x_1 + p, \mathbf{x}_\perp; y_1 + p', \mathbf{y}_\perp) \,\mathrm{e}^{\alpha(|x_1+p-g|-|y_1+p'-g|)} (h_{g,b}(p') - h_{g,b}(p)) \end{aligned}$$

$$+\,\mathcal{P}_0(x_1+p,\mathbf{x}_\perp;y_1+p',\mathbf{y}_\perp)\,\mathrm{e}^{\alpha(|x_1+p-g|-|y_1+p'-g|)}(h_{g,b}(y_1+p')-h_{g,b}(p'))$$
$$+\,\mathcal{P}_0(x_1+p,\mathbf{x}_\perp;y_1+p',\mathbf{y}_\perp)\,\mathrm{e}^{\alpha(|x_1+p-g|-|y_1+p'-g|)}(-h_{g,b}(x_1+p)+h_{g,b}(p))$$
$$=:\mathcal{A}^{(1)}_{pp'}(\mathbf{x},\mathbf{y})+\mathcal{A}^{(2)}_{pp'}(\mathbf{x},\mathbf{y})+\mathcal{A}^{(3)}_{pp'}(\mathbf{x},\mathbf{y}). \tag{3.52}$$

The last two kernels can be analyzed with the same methods as the first one, thus we only estimate the norm of $A^{(1)}_{pp'}$. The crucial observation is that we can write this operator as a product of three operators having the corresponding kernels:

$$\mathcal{A}^{(1)}_{pp'}(\mathbf{x},\mathbf{y})=\mathrm{e}^{\alpha(|x_1+p-g|-|p-g|)}$$
$$\cdot\,\mathrm{e}^{\alpha(|p-g|-|p'-g|)}\mathcal{P}_0(x_1+p,\mathbf{x}_\perp;y_1+p',\mathbf{y}_\perp)(h_{g,b}(p')-h_{g,b}(p))$$
$$\cdot\,\mathrm{e}^{-\alpha(|y_1+p'-g|-|p'-g|)}. \tag{3.53}$$

The kernel in the middle corresponds to the operator $\chi_p P_0\chi_p'$ times some coefficients depending on p, p'.

Using the triangle inequality to bound the exponentials, and (3.32) in order to write $|h_{g,b}(\mathbf{y})-h_{g,b}(\mathbf{x})|\leqslant K|x_1-y_1|$, we have

$$\left\|A^{(1)}_{pp'}\right\|\leqslant K\,\mathrm{e}^{2\alpha}\,\mathrm{e}^{\alpha|p-p'|}|p-p'|\cdot\|\chi_p P_0\chi_p'\|.$$

Using $t=p'$ and $(\alpha+\alpha_0)/2$ in (3.49) we obtain

$$\|\chi_p P_0\chi_p'\|\leqslant C\,\mathrm{e}^{-(\alpha+\alpha_0)|p-p'|/2},$$

thus

$$\left\|A^{(1)}_{pp'}\right\|\leqslant C'|p-p'|\mathrm{e}^{-(\alpha_0-\alpha)|p-p'|/2}$$

which is summable in the sense of (3.51). The same strategy can be applied in the case of $A^{(2)}_{pp'}$ and $A^{(3)}_{pp'}$. The last thing to be done is to prove the Schur–Holmgren estimate.

Lemma 7. *The estimate (3.51) holds true.*

Proof. Let $\psi\in L^2(\mathbb{R}^3)$ with compact support and $\|\psi\|=1$. We write

$$\|A\psi\|^2=\sum_{p\in\mathbb{Z}}\|\chi_p A\psi\|^2. \tag{3.54}$$

But

$$\|\chi_p A\psi\|\leqslant\sum_{p'\in\mathbb{Z}}\sqrt{\|A_{pp'}\|}\sqrt{\|A_{pp'}\|}\|\chi_{p'}\psi\|\leqslant\left\{\sum_{p'\in\mathbb{Z}}\|A_{pp'}\|\right\}^{\frac12}\left\{\sum_{p'\in\mathbb{Z}}\|A_{pp'}\|\|\chi_{p'}\psi\|^2\right\}^{\frac12}$$
$$\leqslant\left\{\sup_{s\in\mathbb{Z}}\sum_{t\in\mathbb{Z}}\|A_{st}\|\right\}^{\frac12}\left\{\sum_{p'\in\mathbb{Z}}\|A_{pp'}\|\|\chi_{p'}\psi\|^2\right\}^{\frac12} \tag{3.55}$$

where in the second inequality we used Cauchy–Schwarz with respect to p'. Introduce this in (3.54) and the bound follows after the use of $\sum_{p'\in\mathbb{Z}}\|\chi_{p'}\psi\|^2=1$. □

Acknowledgments

Part of this work was carried out during a visit of G Nenciu at the Department of Mathematical Sciences, Aalborg University; both hospitality and financial support are gratefully acknowledged. H Cornean acknowledges support from Danish F.N.U. grant *Mathematical Physics and Partial Differential Equations*. A Nenciu and G Nenciu were partially supported by CEEX Grant 05-D11-45/2005. We also thank the first referee for his/her most valuable comments regarding proposition 3.

References

[1] Wannier G H 1937 *Phys. Rev.* **52** 191–7
[2] Nenciu G 1991 *Rev. Mod. Phys.* **63** 91–127
[3] Marzari N and Vanderbilt D 1997 *Phys. Rev.* B **56** 12847–65
[4] Calzolari A, Marzari N, Souza I and Nardelli M B 2004 *Phys. Rev.* B **69** 03518
[5] Chang E, Bussi G, Ruini A and Molinari E 2004 *Phys. Rev. Lett.* **92** 196401
[6] Cornean H, Jensen A and Moldoveanu V 2005 *J. Math. Phys.* **46** 042106
[7] Nardelli M B 1999 *Phys. Rev.* B **60** 7828–33
[8] Thouless D J 1984 *J. Phys. C: Solid State Phys.* **17** L325–7
[9] Cornean H, Nenciu G and Pedersen T G 2006 *J. Math. Phys.* **47** 013511
[10] Brouder C, Panati G, Calandra M, Mourougane Ch and Marzari N 2007 *Phys. Rev. Lett.* **98** 046402
[11] Kohn W 1959 *Phys. Rev.* **115** 809–21
[12] Prodan E 2006 *Phys. Rev.* B **73** 035128
[13] des Cloizeaux J 1964 *Phys. Rev.* **135** A685–97
[14] des Cloizeaux J 1964 *Phys. Rev.* **135** A698–707
[15] Nenciu G 1983 *Commun. Math. Phys.* **91** 81–5
[16] Nenciu A and Nenciu G 1993 *Phys. Rev.* B **47** 10112–5
[17] Panati G 2007 *Ann. Henri Poincaré* **8** 995–1011
[18] Wu X, Dieguez O, Rabe K M and Vanderbilt D 2006 *Phys. Rev. Lett.* **97** 107602
[19] Bhattacharjee J and Waghmare U V 2005 *Phys. Rev.* B **71** 045106
[20] Bruno-Alfonso A and Nacbar D R 2007 *Phys. Rev.* B **75** 115428
[21] Kohn W and Onffroy J 1973 *Phys. Rev.* B **8** 2485–95
[22] Rehr J J and Kohn W 1973 *Phys. Rev.* B **10** 448–55
[23] Geller M R and Kohn W 1993 *Phys. Rev.* B **48** 14085–8
[24] Niu Q 1991 *Mod. Phys. Lett.* B **14, 15** 923–31
[25] Kivelson S 1982 *Phys. Rev.* B **26** 4269–74
[26] Nenciu A and Nenciu G 1998 *Commun. Math. Phys.* **190** 541–8
[27] Reed M and Simon B 1975 *Methods of Modern Mathematical Physics: II. Fourier Analysis. Self-adjointness* (New York: Academic)
[28] Avron J E 1979 *J. Phys. A: Math. Gen.* **12** 2393–8
[29] Nenciu A and Nenciu G 1981 *J. Phys. A: Math. Gen.* **14** 2817–27