Adiabatic Decoupling and Time-Dependent Born-Oppenheimer Theory

Herbert Spohn and Stefan Teufel

Zentrum Mathematik and Physik Department,
Technische Universität München,
80290 München, Germany
email: spohn@ma.tum.de, teufel@ma.tum.de

July 9, 2001

Abstract

We reconsider the time-dependent Born-Oppenheimer theory with the goal to carefully separate between the adiabatic decoupling of a given group of energy bands from their orthogonal subspace and the semiclassics within the energy bands. Band crossings are allowed and our results are local in the sense that they hold up to the first time when a band crossing is encountered. The adiabatic decoupling leads to an effective Schrödinger equation for the nuclei, including contributions from the Berry connection.

1 Introduction

Molecules consist of light electrons, mass m_e, and heavy nuclei, mass M which depends on the type of nucleus. Born and Oppenheimer [3] wanted to explain some general features of molecular spectra and realized that, since the ratio m_e/M is small, it could be used as an expansion parameter for the energy levels of the molecular Hamiltonian. The time-*independent* Born-Oppenheimer theory has been put on firm mathematical grounds by Combes, Duclos, and Seiler [5], Hagedorn [8], and more recently in [16].

With the development of tailored state preparation and ultra precise time resolution there is a growing interest in understanding and controlling the dynamics of molecules, which requires an analysis of the solutions to the time-*dependent* Schrödinger equation, again exploiting that m_e/M is small. The molecular Hamiltonian is of the form

$$H = \frac{\hbar^2}{2m_\mathrm{e}}\Big(-i\nabla_x - A_\mathrm{ext}(x)\Big)^2 + \frac{\hbar^2}{2M}\Big(-i\nabla_X + A_\mathrm{ext}(X)\Big)^2 + V_\mathrm{e}(x) + V_\mathrm{en}(X,x) + V_\mathrm{n}(X)\,. \quad (1)$$

For notational simplicity we ignore spin degrees of freedom and assume that all nuclei have the same mass. We have k electrons with positions $\{x_1,\dots,x_k\} = x$ and l nuclei with positions $\{X_1,\dots,X_l\} = X$. The first and second term of H are the kinetic energies of the electrons and of the nuclei, respectively. An external magnetic field is included through the vector potential A_ext. Electrons and nuclei interact via the static Coulomb potential. Therefore V_e is the electronic, V_n the nucleonic repulsion, and V_en the attraction between electrons and nuclei. V_e and V_n may also contain an external electrostatic potential.

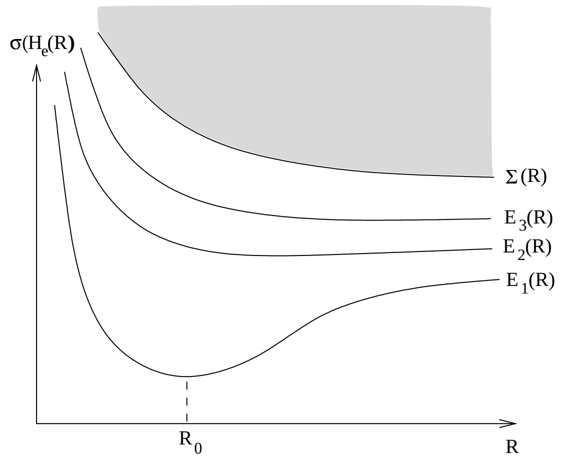

Figure 1: The schematic spectrum of $H_{\mathrm{e}}(R)$ for a diatomic molecule as a function of the separation R of the two nuclei.

In atomic units ($m_{\mathrm{e}} = \hbar = 1$) the Hamiltonian (1) can be written more concisely as

$$H = \frac{m_{\mathrm{e}}}{M} \frac{1}{2} \Big(-i\nabla_X + A_{\mathrm{ext}}(X) \Big)^2 + H_{\mathrm{e}}(X)\,, \tag{2}$$

emphasizing that the nuclear kinetic energy will be treated as a "small perturbation". $H_{\mathrm{e}}(X)$ is the electronic Hamiltonian for given position X of the nuclei,

$$H_{\mathrm{e}}(X) = \frac{1}{2} \Big(-i\nabla_x - A_{\mathrm{ext}}(x) \Big)^2 + V_{\mathrm{e}}(x) + V_{\mathrm{en}}(X, x) + V_{\mathrm{n}}(X)\,. \tag{3}$$

$H_{\mathrm{e}}(X)$ is a self-adjoint operator on the electronic Hilbert space $L^2(\mathbb{R}^{3k})$ restricted to its antisymmetric subspace. Later on we will need some smoothness of $H_{\mathrm{e}}(X)$, which can be established easily if the electrons are treated as point-like and the nuclei have an extended, rigid charge distribution.

Generically $H_{\mathrm{e}}(X)$ has, possibly degenerate, eigenvalues $E_1(X) < E_2(X) < \ldots$ which terminate at the continuum edge $\Sigma(X)$. Thereby one obtains the band structure as plotted schematically in Figure 1. The discrete bands $E_j(X)$ may cross and possibly merge into the continuous spectrum as indicated in Figure 2.

Comparing kinetic energies, we find for the speeds $|v_{\mathrm{n}}| \approx (m_{\mathrm{e}}/M)^{1/2}|v_{\mathrm{e}}|$, which means that on the atomic scale the nuclei move very slowly. If we regard $X(t)$ as a given nucleonic trajectory, then $H_{\mathrm{e}}(X(t))$ is a Hamiltonian with slow time variation and the time-adiabatic theorem [15, 14, 1] can be applied [2]. For us X are quantum mechanical degrees of freedom. The Hamiltonian H of (2) is time-independent and we can only exploit that the nucleonic Laplacian carries a small prefactor. To distinguish, we refer to our situation as space-adiabatic. Since the nuclei move very slowly, their dynamics must be followed over sufficiently long times. From the speed ratio we conclude that these times are of order $(m_{\mathrm{e}}/M)^{1/2}$ in atomic units. To simplify notation we define

$$\varepsilon = \sqrt{\frac{m_{\mathrm{e}}}{M}} \tag{4}$$

as the small dimensionless parameter. Then

$$H^{\varepsilon} = \varepsilon^2 \frac{1}{2} \Big(-i\nabla_X + A_{\mathrm{ext}}(X) \Big)^2 + H_{\mathrm{e}}(X)\,, \tag{5}$$

and we want to study the solutions of the time-dependent Schrödinger equation

$$i\varepsilon\frac{\partial\psi}{\partial t} = H^\varepsilon\psi \tag{6}$$

in the limit of small ε.

The crude physical picture underlying the analysis of (6) is that the nuclei behave semiclassically because of their large mass and that the electrons rapidly adjust to the slow nucleonic motion. Thus, in fact, the time-dependent Born-Oppenheimer approximation involves two limits. If the electrons are initially in the eigenstate $\chi_j(X_0)$ of the j-th band with energy $E_j(X_0)$, where X_0 is the approximate initial configuration of the nuclei, then the j-th band is adiabatically protected provided there is an energy gap separating it from the rest of the spectrum. Thus at later times, up to small error, the electronic wave function is still in the subspace corresponding to the j-th band. But this implies that the nuclei are governed by the Born-Oppenheimer Hamiltonian

$$H^\varepsilon_{\mathrm{BO}} = \varepsilon^2\frac{1}{2}\Big(-i\nabla_X + A_{\mathrm{ext}}(X)\Big)^2 + E_j(X)\,. \tag{7}$$

Since $\varepsilon \ll 1$, $H^\varepsilon_{\mathrm{BO}}$ can be analyzed through semiclassical methods where to leading order the contributions come from the classical flow Φ^t corresponding to the classical Hamiltonian $H^{\mathrm{cl}}_{\mathrm{BO}} = \frac{1}{2}p^2 + E_j(q)$ on nucleonic phase space.

In general, $E_j(X)$ may touch another band as X varies. To allow for such band crossings we introduce the region $\Lambda \subset \mathbb{R}^n$, $n = 3l$, in nucleonic configuration space, such that E_j restricted to Λ does not cross or touch any other energy band. The classical flow Φ^t then has $\Lambda \times \mathbb{R}^n$ as phase space and is defined only up to the time when it first hits the boundary $\partial\Lambda \times \mathbb{R}^n$. Up to that time (7) still correctly describes the quantum evolution. To follow the tunneling through a band crossing other methods have to be used [11, 7], in particular, the codimension of the crossing is of relevance.

The mathematical investigation of the time-dependent Born-Oppenheimer theory was initiated and carried out in great detail by Hagedorn. In his pioneering work [9] he constructs approximate solutions to (6) of the form $\phi_{q(t),p(t)} \otimes \chi_j(q(t))$, where $\phi_{q(t),p(t)}$ is a coherent state carried along the classical flow, $(q(t),p(t)) = \Phi^t(q_0,p_0)$. The difference to the true solution with the same initial condition is of order $\sqrt{\varepsilon}$ in the L^2-norm over times of order ε^{-1} in atomic units and the approximation holds until the first hitting time of $\partial\Lambda \times \mathbb{R}^n$. In a recent work Hagedorn and Joye [10] construct solutions to (6) satisfying exponentially small error estimates. In Hagedorn's approach the "adiabatic and semiclassical limits are being taken simultaneously, and they are coupled [10]".

In our paper we carefully separate the space-adiabatic and the semiclassical limit. One immediate benefit is the generalization of the first order analysis of Hagedorn from coherent states to arbitrary wave functions.

Let us explain our result for the space-adiabatic part in more detail. We assume that there is some region $\Lambda \subset \mathbb{R}^n$ in the nucleonic configuration space, such that some subset $\sigma_*(X)$ of $\sigma(H_{\mathrm{e}}(X))$ is separated from the remainder of the spectrum by a gap for all $X \in \Lambda$, i.e.

$$\mathrm{dist}(\sigma_*(X), \sigma(H_{\mathrm{e}}(X)) \setminus \sigma_*(X)) \geq d > 0 \quad \text{for all } X \in \Lambda\,.$$

Λ could be punctured by small balls (for $n = 2$) because of band crossings. Λ could also terminate because the point spectrum merges in the continuum, which physically means that the molecule loses an electron through ionization. Let $P_*(X)$ be the spectral projection of $H_{\mathrm{e}}(X)$ associated with $\sigma_*(X)$ and $P_* = \int_\Lambda^\oplus dX\, P_*(X)$. We will establish that the unitary time evolution $e^{-iH^\varepsilon t/\varepsilon}$

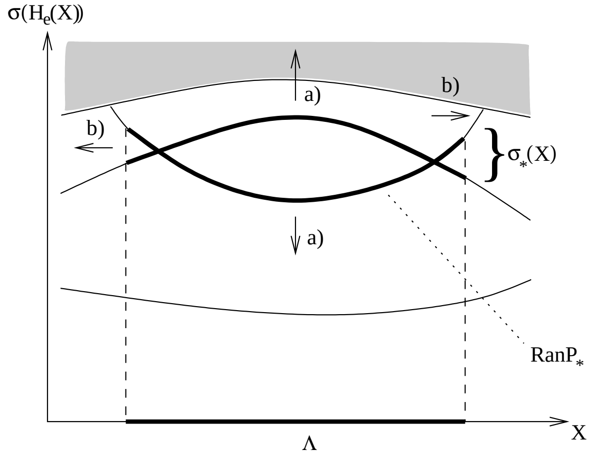

Figure 2: The wave function can leave RanP_* in two different ways. Either by transitions to other bands (a) or through the boundary of Λ (b).

agrees on RanP_* with the diagonal evolution $e^{-iH^\varepsilon_{\mathrm{diag}}t/\varepsilon}$ generated by $H^\varepsilon_{\mathrm{diag}} := P_* H^\varepsilon P_*$ up to errors of order ε as long as the leaking through the boundary of Λ is sufficiently small.

To complete the analysis one has to control the flow of the wave function through $\partial\Lambda$. One possibility is to simply avoid the problem by assuming that $\Lambda = \mathbb{R}^n$, hence $\partial\Lambda = \emptyset$. We will refer to this case as a *globally isolated* band. Of course, the set $\{(X, y) \in \mathbb{R}^n \times \mathbb{R} : y \in \sigma_*(X)\}$ may contain arbitrary band crossings. As one of our main results, we prove that the subspace RanP_* is adiabatically protected. In particular for the purpose of studying band crossings the full molecular Hamiltonian may be replaced by a simplified model with two bands only.

In general one has $\partial\Lambda \neq \emptyset$, to which we refer as a *locally isolated* band. To estimate the flow out of Λ the only technique available seems to be semiclassical analysis. But this requires a control over the semiclassical evolution, for which one needs, at present, that $\{(X, y) \in \Lambda \times \mathbb{R} : y \in \sigma_*(X)\}$ contains no band crossings. Then $\{(X, y) \in \Lambda \times \mathbb{R} : y \in \sigma_*(X)\} = \cup_j \{(X, y) \in \Lambda \times \mathbb{R} : y = E_j(X)\}$ is the disjoint union of possibly degenerate energy bands $E_j(X)$. We will prove that each band separately is adiabatically protected.

In the special case where $\sigma_*(X) = E_j(X)$ is a nondegenerate eigenvalue for $X \in \Lambda$, $e^{-iH^\varepsilon_{\mathrm{diag}}t/\varepsilon}$ is well approximated through $e^{-iH^\varepsilon_{\mathrm{BO}}t/\varepsilon}$ on $L^2(\mathbb{R}^n)$. Since $H^\varepsilon_{\mathrm{BO}}$ is a standard semiclassical operator, one can easily control the X-support of the wave function and therefore prove a result for rather general $\Lambda \subset \mathbb{R}^n$, for the details see Theorem 4. Roughly speaking, it says that if ϕ_t is a solution of the effective Schrödinger equation for the nuclei

$$i\varepsilon \frac{\partial \phi_t}{\partial t} = H^\varepsilon_{\mathrm{BO}} \phi_t \,, \tag{8}$$

with supp$\phi_0 \subset \Lambda$, then, modulo an error of order ε,

$$\psi_t := \phi_t(X)\chi_j(X, x)$$

is a solution of the full Schrödinger equation (6) with initial condition $\psi_0(X, x) = \phi_0(X)\chi_j(X, x)$ as long as ϕ_t is supported in Λ up to L^2-mass of order ε. This maximal time span can be computed using the classical flow Φ^t.

As first observed by Mead and Truhlar [19], in general $H^\varepsilon_{\mathrm{BO}}$ acquires as a first order correction an additional vector potential $A_{\mathrm{geo}}(X) = -i\langle\chi_j(X), \nabla_X\chi_j(X)\rangle$ and (7) has to be replaced by

$$H^\varepsilon_{\mathrm{BO}} = \varepsilon^2\frac{1}{2}\Big(-i\nabla_X + A_{\mathrm{ext}}(X) + A_{\mathrm{geo}}(X)\Big)^2 + E_j(X)\,. \tag{9}$$

Multiplying $\chi_j(X)$ with a smooth X-dependent phase factor induces a gauge transformation for A_{geo}, which implies that the physical predictions based on (9) do not change, as it should be. As noticed in [19], if Λ is not contractible, then A_{geo} cannot be removed through a gauge transformation and (9) and (7) describe different physics. Berry realized that geometric phases appear whenever the Hamiltonian has slowly changing parameters. Therefore $A_{\mathrm{geo}}(X)$ is referred to as Berry connection, cf. [22] for an instructive collection of reprints. In fact, the motion of nuclei as governed by the Born-Oppenheimer Hamiltonian (9) is one of the paradigmatic examples for geometric phases.

If $\sigma_*(X) = E(X)$ is k-fold degenerate, not much of the above analysis changes. $H^\varepsilon_{\mathrm{BO}}$ becomes matrix-valued and acts on $L^2(\mathbb{R}^n)^{\oplus k}$, i.e.

$$H^\varepsilon_{\mathrm{BO}} = \left(\frac{\varepsilon^2}{2}\Big(-i\nabla_X + A_{\mathrm{ext}}(X)\Big)^2 + E_j(X)\right)\mathbf{1}_{k\times k}+\frac{\varepsilon}{2}\Big((-i\varepsilon\nabla_X)\cdot A_{\mathrm{geo}}(X)+A_{\mathrm{geo}}(X)\cdot(-i\varepsilon\nabla_X)\Big)\,.$$

The connection $A_{\mathrm{geo}}(X)$ contains in general also off-diagonal terms and matrix-valued semiclassics must be applied. However, since the only nondiagonal term is in the subprincipal symbol, the leading order semiclassical analysis reduces to the scalar case and, in particular, agrees with the nondegenerate band case. We do not carry out the straightforward extension of Theorem 4 below to the degenerate band case, because the technicalities of matrix-valued semiclassics would obscure the simple ideas behind our analysis.

In their recent work [18] Martinez and Sordoni independently study the time-dependent Born-Oppenheimer approximation as based on techniques developed by Nenciu and Sordoni [20]. They consider the case of a globally isolated band for a Hamiltonian of the form (1) with smooth V and $A_{\mathrm{ext}} = 0$. They succeed in proving the adiabatic decoupling to any order in ε for subspaces P^ε_* which are ε-close to the unperturbed subspaces P_* considered by us. With this result, in principle, higher order corrections to the effective Hamiltonian (7) could be computed.

The paper is organized as follows. Section 2 contains the precise formulation of the results. Section 3 gives a short discussion of the semiclassical limit of $H^\varepsilon_{\mathrm{BO}}$ and on how such results extend to the full molecular system. Proofs are provided in Section 4. In spirit they rely on techniques developed in [23] in the context of the semiclassical limit for dressed electron states. In practice the Born-Oppenheimer approximation requires several novel constructions, since the "perturbation" $-\frac{\varepsilon^2}{2}\Delta$ increases quadratically.

Our results can be formulated and proved in a more general framework dealing with, possibly time-dependent, perturbations of fibered operators. Also the gap condition can be removed by using arguments similar to those developed by Avron and Elgart in [1]. The general operator theoretical results will appear elsewhere [24].

2 Main results

The specific form (3) of the electronic part of the Hamiltonian will be of no importance in the following. Thus we only assume that

$$H_{\mathrm{e}} = \int^\oplus_{\mathbb{R}^n} dX\, H_{\mathrm{e}}(X)\,,\qquad H_{\mathrm{e}}(X) = H_{\mathrm{e}0} + H_{\mathrm{e}1}(X)\,,$$

where H_{e0} is self-adjoint on some dense domain $\mathcal{D} \subset \mathcal{H}_e$ and bounded from below and $H_{e1}(X) \in \mathcal{L}(\mathcal{H}_e)$ is a continuous family of self-adjoint operators, bounded uniformly for $X \in \mathbb{R}^n$. Thus H_e is self-adjoint on $D(H_e) = L^2(\mathbb{R}^n) \otimes \mathcal{D} \subset \mathcal{H} := L^2(\mathbb{R}^n) \otimes \mathcal{H}_e$ and bounded from below. For the definition of $L^2(\mathbb{R}^n) \otimes \mathcal{D}$ we equip $\mathcal{D}$ with the graph-norm $\|\cdot\|_{H_{e0}}$, i.e., for $\psi \in \mathcal{D}$, $\|\psi\|_{H_{e0}} = \|H_{e0}\psi\| + \|\psi\|$.

Let $A_{\mathrm{ext}} \in C^1_{\mathrm{b}}(\mathbb{R}^n, \mathbb{R}^n)$, where for any open set $\Omega \subset \mathbb{R}^m$, $m \in \mathbb{N}$, $C^k_{\mathrm{b}}(\Omega)$ denotes the set of functions $f \in C^k(\Omega)$ such that for each multi-index α with $|\alpha| \leq k$ there exists a $C_\alpha < \infty$ with

$$\sup_{x\in\Omega} |\partial^\alpha f(x)| \leq C_\alpha \,.$$

Then $\frac{\varepsilon^2}{2}\Big(-i\nabla_X + A_{\mathrm{ext}}(X)\Big)^2$ is self-adjoint on $W^2(\mathbb{R}^n)$, the second Sobolev space, since $-i\nabla_X$ is infinitesimally operator bounded with respect to $-\Delta_X$. It follows that

$$H^\varepsilon = \frac{\varepsilon^2}{2}\Big(-i\nabla_X + A_{\mathrm{ext}}(X)\Big)^2 \otimes \mathbf{1} + H_e \tag{10}$$

self-adjoint on $D(H^\varepsilon) = W^2(\mathbb{R}^n) \otimes \mathcal{H}_e \cap D(H_e)$.

For $X \in \Lambda$, $\Lambda \subset \mathbb{R}^n$ open, we require in addition some regularity for $H_e(X)$ as a function of X:

H$_k$ $\quad H_{e1}(\cdot) \in C^k_{\mathrm{b}}(\Lambda, \mathcal{L}(\mathcal{H}_e))$.

The exact value of k will depend on whether $\Lambda = \mathbb{R}^n$ or $\Lambda \subset \mathbb{R}^n$. For the type of Hamiltonian considered in the introduction, cf. (1), all the above conditions including Condition **H**$_k$ are easily checked and put constraints only on the smoothness of the external potentials and on the smoothness and the decay of the charge distribution of the nuclei. For point nuclei **H**$_k$ fails and a suitable substitute would require a generalization of the Hunziker distortion method of [16].

We will be interested in subsets of $\{(X,s) \in \Lambda \times \mathbb{R} : s \in \sigma(H_e(X)\}$ which are isolated from the rest of the spectrum in the following sense.

S *For $X \in \Lambda$, let $\sigma_*(X) \subset \sigma(H_e(X))$ be such that there are functions $f_\pm \in C_{\mathrm{b}}(\Lambda, \mathbb{R})$ and a constant $d > 0$ with*

$$[f_-(X) + d, f_+(X) - d] \cap \sigma_*(X) = \sigma_*(X)$$

and

$$[f_-(X), f_+(X)] \cap (\sigma(H_e(X) \setminus \sigma_*(X)) = \emptyset \,.$$

We set $P_* = \int_\Lambda^\oplus dX\, P_*(X)$, where $P_*(X) = 1\!\!1_{\sigma_*(X)}(H_e(X))$ is the spectral projection of $H_e(X)$ with respect to $\sigma_*(X)$. As explained in the introduction we have to distinguish two cases.

(i) Globally isolated bands

We assume $\Lambda = \mathbb{R}^n$ and let

$$H^\varepsilon_{\mathrm{diag}} := P_*\, H^\varepsilon\, P_* + P_*^\perp\, H^\varepsilon\, P_*^\perp \,. \tag{11}$$

Since we aim at a uniform result for the adiabatic theorem, we introduce the Sobolev spaces $W^{1,\varepsilon}(\mathbb{R}^n)$ and $W^{2,\varepsilon}(\mathbb{R}^n)$ with respect to the ε-scaled gradient, i.e.

$$W^{1,\varepsilon}(\mathbb{R}^n) := \big\{\phi \in L^2(\mathbb{R}^n) : \|\phi\|_{W^{1,\varepsilon}} := \|\varepsilon\,|\nabla\phi|\,\| + \|\phi\| < \infty\big\}$$

and

$$W^{2,\varepsilon}(\mathbb{R}^n) := \left\{\phi \in L^2(\mathbb{R}^n) : \|\phi\|_{W^{2,\varepsilon}} := \|\varepsilon^2\Delta\phi\| + \|\phi\| < \infty\right\} .$$

Alternatively we will project on finite total energies and define $\mathcal{E}(H^\varepsilon) := 1\!\!1_{(-\infty,\mathcal{E}]}(H^\varepsilon)$ as the projection on total energies smaller than $\mathcal{E}$.

Theorem 1. *Assume* $\mathbf{H}_3$ *and* $\mathbf{S}$ *for* $\Lambda = \mathbb{R}^n$*. Then* $H^\varepsilon_{\mathrm{diag}}$ *is self-adjoint on the domain of* H^ε*. There are constants* $C, \widetilde{C} < \infty$ *such that for all* $t \in \mathbb{R}$

$$\left\| e^{-iH^\varepsilon t/\varepsilon} - e^{-iH^\varepsilon_{\mathrm{diag}} t/\varepsilon} \right\|_{\mathcal{L}(W^{2,\varepsilon}\otimes\mathcal{H}_{\mathrm{e}},\mathcal{H})} \leq \varepsilon\, C\,(1+|t|)^3 \tag{12}$$

and for all $\mathcal{E} \in \mathbb{R}$

$$\left\| \left(e^{-iH^\varepsilon t/\varepsilon} - e^{-iH^\varepsilon_{\mathrm{diag}} t/\varepsilon} \right) \mathcal{E}(H^\varepsilon) \right\|_{\mathcal{L}(\mathcal{H})} \leq \varepsilon\, \widetilde{C}\,(1+|\mathcal{E}|)\,(1+|t|)\,. \tag{13}$$

$\mathcal{L}(W^{2,\varepsilon}\otimes\mathcal{H}_{\mathrm{e}},\mathcal{H})$ *denotes the space of bounded linear operators from* $W^{2,\varepsilon}\otimes\mathcal{H}_{\mathrm{e}}$ *to* $\mathcal{H}$ *equipped with the operator norm.*

This result should be understood as an adiabatic theorem for the subspaces $\mathrm{Ran}P_*$ and $\mathrm{Ran}P_*^\perp$, which are *not* spectral subspaces. Let us point out one immediate application of Theorem 1. The behavior near band crossings is usually investigated using simplified models involving only two energy bands and ignoring the rest of the spectrum, cf. [11, 7]. Theorem 1 shows that this strategy is indeed justified modulo errors of order ε.

(ii) Locally isolated bands

$\sigma_*(X) = E(X)$ is a nondegenerate eigenvalue for all $X \in \Lambda$. Λ may now be any open subset of $\mathbb{R}^n$ and for such a Λ we assume $\mathbf{H}_\infty$ and $\mathbf{S}$. We also assume that Λ is connected. Otherwise one could treat each connected component separately.

It is easy to see that, given $\mathbf{H}_\infty$ and $\mathbf{S}$, the family of projections $P_*(\cdot) \in C^\infty_{\mathrm{b}}(\Lambda, \mathcal{L}(\mathcal{H}_{\mathrm{e}}))$. However, in order to "map" the dynamics from $\mathrm{Ran}P_*$ to $L^2(\Lambda)$ we need in addition a smooth version $\chi(\cdot) \in C^\infty_{\mathrm{b}}(\Lambda, \mathcal{H}_{\mathrm{e}})$ of the normalized eigenvector of $H_{\mathrm{e}}(X)$ with eigenvalue $E(X)$. In other words we require the complex line bundle over Λ defined by P_* to be trivial. This always holds for contractible Λ, but, as discussed below, also for some relevant examples where Λ is not contractible.

Given a smooth version of $\chi(X)$ with $\|\chi(X)\| = 1$, one has $\mathrm{Re}\langle\chi(X), \nabla_X\chi(X)\rangle = 0$, but, in general, $\mathrm{Im}\langle\chi(X), \nabla_X\chi(X)\rangle \neq 0$. In the following we distinguish two cases: Either it is possible to achieve $\mathrm{Im}\langle\widetilde{\chi}(X), \nabla_X\widetilde{\chi}(X)\rangle = 0$ by a smooth gauge transformation $\chi(X) \to \widetilde{\chi}(X) = e^{i\theta(X)}\chi(X)$ or not. In the latter case

$$A_{\mathrm{geo}}(X) := -i\langle\chi(X), \nabla_X\chi(X)\rangle$$

is the gauge potential of a connection on the trivial complex line bundle over Λ, the Berry connection, and has to be taken into account in the definition of the effective operator

$$H^\varepsilon_{\mathrm{BO}} := \frac{\varepsilon^2}{2}\Big(-i\nabla_X + A_{\mathrm{ext}}(X) + A_{\mathrm{geo}}(X)\Big)^2 + E(X) \tag{14}$$

with domain $W^2(\mathbb{R}^n)$. Thus A_{geo} acts as an additional external magnetic vector potential. Although A_{ext} and A_{geo} appear in $H^\varepsilon_{\mathrm{BO}}$ with an ε in front only, and therefore are not retained in the semiclassical limit to leading order, they do contribute to the solution of the Schrödinger equation for times of order ε^{-1}. If the full Hamiltonian is real in position representation, as it is the case for the Hamiltonians considered in the introduction whenever $A_{\rm ext}=0$, then $\chi(X)$ can be chosen real-valued. If, in addition, Λ is contractible, the existence of a smooth version of $\chi(X)$ with $\mathrm{Im}\langle\chi(X),\nabla_X\chi(X)\rangle=0$ follows.

To define $H^\varepsilon_{\rm BO}$ on $L^2(\mathbb{R}^n)$ through (14), the functions $E(X)$ and $A_{\rm geo}(X)$, which are a priori defined on Λ only, must be continued to functions on $\mathbb{R}^n$. Hence we arbitrarily extend $E(X)$ and $A_{\rm geo}(X)$ to functions in $C^\infty_{\rm b}(\mathbb{R}^n)$ by modifying them, if necessary, on $\Lambda\setminus(\Lambda-\delta/5)$ (cf. (17)) for some $\delta>0$. The parameter δ will be fixed in the formulation of Theorem 4 and will appear in several places. It controls how close the states are allowed to come to $\partial\Lambda$.

The generic example for the Berry phase is a band crossings of codimension 2 (cf. [22, 11, 7]). If $E(X)$ is an isolated energy band except for a codimension 2 crossing, then $\Lambda=\mathbb{R}^n\setminus\{\text{closed neighborhood of the crossing}\}$ is no longer contractible, but the line bundle is still trivial. Although the underlying Hamiltonian is real, the Berry connection cannot be gauged away. Within the time-independent Born-Oppenheimer approximation Herrin and Howland [12] study a model with a nontrivial eigenvector bundle.

With the fixed choice for $\chi(X)$ we have

$$\mathrm{Ran}\,P_*=\left\{\int_\Lambda^\oplus dX\,\phi(X)\chi(X);\;\phi\in L^2(\Lambda)\right\}\subset\mathcal{H}\,. \tag{15}$$

Thus there is a natural identification $\mathcal{U}:\mathrm{Ran}P_*\to L^2(\mathbb{R}^n)$ connecting the relevant subspace on which the full quantum evolution takes place and the Hilbert space $L^2(\mathbb{R}^n)$ on which the effective Born-Oppenheimer evolution is defined. According to (15), we set

$$\mathcal{U}(\phi\chi)=\phi\,,\qquad\text{i.e.}\quad(\mathcal{U}P_*\psi)(X)=\langle\,\chi(X)\,,\,(P_*\psi)(X)\,\rangle_{\mathcal{H}_{\rm e}}\,.$$

Its adjoint $\mathcal{U}^*:L^2(\mathbb{R}^n)\to\mathrm{Ran}P_*$ is given by

$$\mathcal{U}^*\phi=\int_\Lambda^\oplus dX\,\phi(X)\chi(X)\,.$$

Clearly $\mathcal{U}$ is an isometry and $\mathcal{U}^*\mathcal{U}=\mathbf{1}$ on $\mathrm{Ran}P_*$. But $\mathcal{U}$ is not surjective and thus not unitary.

By construction, $e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}$ is a good approximation to the true dynamics only as long as the wave function of the nuclei is supported in Λ modulo errors of order ε. Since $H^\varepsilon_{\rm BO}$ is a standard semiclassical operator, the X-support of solutions of (8) can be calculated approximately from the classical dynamics generated by its principal symbol $H_{\rm cl}(q,p)=\frac{1}{2}p^2+E(q)$ on phase space $Z:=\mathbb{R}^n\times\mathbb{R}^n$,

$$\frac{d}{dt}q=p\,,\qquad\frac{d}{dt}p=-\nabla E(q)\,. \tag{16}$$

The solution flow to (16) exists for all times and will be denoted by Φ^t.

In order to make these notions more precise, we need to introduce some notation. The Weyl quantization of $a\in C^\infty_{\rm b}(Z)$ is the linear operator

$$\left(a^{\mathrm{W},\varepsilon}\phi\right)(X)=(2\pi)^{-n}\int_{\mathbb{R}^n}dY\,dk\,a\left(\frac{X+Y}{2},\varepsilon\,k\right)e^{-i(X-Y)\cdot k}\phi(Y)\,,$$

as acting on Schwartz functions. $a^{\mathrm{W},\varepsilon}$ extends to $\mathcal{L}(L^2(\mathbb{R}^n))$ with operator norm bounded uniformly in ε (cf., e.g., Theorem 7.11 in [6]). The wave functions with phase space support in a compact set $\Gamma\subset Z$ do not form a closed subspace of $L^2(\mathbb{R}^n)$. Hence we cannot project on this set. In order to define approximate projections, let for $\Gamma\subset\mathbb{R}^m$, $m\in\mathbb{N}$, and for $\alpha>0$

$$\Gamma-\alpha:=\left\{z\in\Gamma:\;\inf_{w\in\mathbb{R}^m\setminus\Gamma}|w-z|\geq\alpha\right\}\,. \tag{17}$$

Definition 2. *An* approximate characteristic function $\mathbb{1}_{(\Gamma,\alpha)} \in C_{\mathrm{b}}^\infty(\mathbb{R}^m)$ *of a set* $\Gamma \subset \mathbb{R}^m$ *with margin* α *is defined by the requirement that* $\mathbb{1}_{(\Gamma,\alpha)}|_{\Gamma-\alpha} = 1$ *and* $\mathbb{1}_{(\Gamma,\alpha)}|_{\mathbb{R}^m\setminus\Gamma} = 0$.

If $\mathbb{1}_{(\Gamma,\alpha)}$ *is an approximate characteristic function on phase space* Z*, then the corresponding* approximate projection *is defined as its Weyl quantization* $\mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}$*. We will say that functions in* $\mathrm{Ran}\mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}$ *have phase space support in* Γ*.*

For $\Gamma \subset Z$ we will use the abbreviations

$$\begin{aligned} \Gamma_q &:= \{q \in \mathbb{R}^n : (q,p) \in \Gamma \text{ for some } p \in \mathbb{R}^n\}\,, \\ \Gamma_p &:= \{p \in \mathbb{R}^n : (q,p) \in \Gamma \text{ for some } q \in \mathbb{R}^n\}\,. \end{aligned}$$

Let the phase space support Γ of the initial wave function be such that $\Gamma_q \subset \Lambda - \delta$. Then the maximal time interval for which the X-support of the wave function of the nuclei stays in Λ up to errors of order ε can be written as

$$I^\delta_{\max}(\Gamma,\Lambda) := [T^\delta_-(\Gamma,\Lambda), T^\delta_+(\Gamma,\Lambda)]\,,$$

where the "first hitting times" $T_\pm$ are defined by the classical dynamics through

$$T^\delta_+(\Gamma,\Lambda) := \sup\left\{t \geq 0 : (\Phi^s(\Gamma))_q \subseteq \Lambda - \delta \ \forall s \in [0,t]\right\}$$

and $T^\delta_-(\Gamma,\Lambda)$ analogously for negative times. This are just the first times for a particle starting in Γ to hit the boundary of $\Lambda-\delta$ when dragged along the classical flow Φ^t.

The following proposition, which is an immediate consequence of Egorov's Theorem [4, 21], shows that for times in $I^\delta_{\max}(\Gamma,\Lambda)$ the support of the wave function of the nuclei stays indeed in $\Lambda - \delta$, up to errors of order ε uniformly on $\mathrm{Ran}\mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}$ for any approximate projection $\mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}$.

Proposition 3. *Let* $\Gamma \subset Z$ *be such that* $\Gamma_q \subset \Lambda - \delta$ *and let* $\mathbb{1}_{\Lambda-\delta}$ *denote multiplication with the characteristic function of* $\Lambda - \delta$ *on* $L^2(\mathbb{R}^n)$*. For any approximate projection* $\mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}$ *and any bounded interval* $I \subseteq I^\delta_{\max}(\Gamma,\Lambda)$ *there is a constant* $C < \infty$ *such that for all* $t \in I$

$$\left\| (\mathbf{1} - \mathbb{1}_{\Lambda-\delta})\, e^{-iH^\varepsilon_{\mathrm{BO}}t/\varepsilon}\, \mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)} \right\|_{\mathcal{L}(L^2(\mathbb{R}^n))} \leq C\,\varepsilon\,.$$

An approximate projection on Γ in $\mathcal{H}$ is defined as $P^\alpha_\Gamma := \mathcal{U}^*\, \mathbb{1}_{(\Lambda,\delta)}\, \mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}\, \mathcal{U}\, P_*$, where $\mathbb{1}^{\mathrm{W},\varepsilon}_{(\Gamma,\alpha)}$ is an approximate projection on Γ according to Definition 2 and $\mathbb{1}_{(\Lambda,\delta)}$ is an approximate characteristic function for Λ. Using the latter instead of the sharp cutoff from $\mathcal{U}^*$ makes $\mathrm{Ran}P^\alpha_\Gamma$ a bounded set in $W^{2,\varepsilon} \otimes \mathcal{H}_{\mathrm{e}}$ whenever Γ_p is a bounded set.

Theorem 4. *Assume* $\mathbf{H}_\infty$ *and* $\mathbf{S}$ *with dim(Ran*$P_*(X)$*)* $= 1$ *for some open* $\Lambda \subseteq \mathbb{R}^n$*. Let* $\Gamma \subset Z$ *be such that* $\Gamma_q \subset \Lambda - \delta$ *for some* $\delta > 0$ *and* Γ_p *bounded. For any approximate projection* P^α_Γ *and any bounded interval* $I \subseteq I^\delta_{\max}(\Gamma,\Lambda)$ *there is a constant* $C < \infty$ *such that for all* $t \in I$

$$\left\| \left(e^{-iH^\varepsilon t/\varepsilon} - \mathcal{U}^*\, e^{-iH^\varepsilon_{\mathrm{BO}}t/\varepsilon}\,\mathcal{U}\right) P^\alpha_\Gamma \right\|_{\mathcal{L}(\mathcal{H})} \leq C\varepsilon\,. \tag{18}$$

Theorem 4 establishes that the electrons adiabatically follow the motion of the nuclei up to errors of order ε as long as the leaking through the boundary of Λ is small. The semiclassics was used only to control such a leaking uniformly. However, for $H^\varepsilon_{\mathrm{BO}}$ the limit $\varepsilon \to 0$ is a semiclassical limit and, as discussed in the following section, beyond the mere support of the wave function more detailed information is available.

3 Semiclassics for a single band

The semiclassical limit of Equation (8) with a Hamiltonian of the form (14) is well understood and there is a variety of different approaches. For example one can construct approximate solutions $\phi_{q(t)}$ of (8) which are localized along a classical trajectory $q(t)$, i.e. along a solution of (16). Then it follows from Theorem 4 that $\phi_{q(t)}\chi$ is a solution of the full Schrödinger equation, (6), up to an error of order ε as long as $q(t) \in \Lambda - \delta$. Roughly speaking, this coincides with the result of Hagedorn [9]. In applications the assumption that the wave function of the nuclei is well described by a coherent state seems to be rather restrictive and a more general approach to the semiclassical analysis of a Schrödinger equation of the form (8) is to consider the distributions of semiclassical observables, i.e. of operators obtained as Weyl quantization $a^{\mathrm{W},\varepsilon}$ of classical phase space functions $a : Z \to \mathbb{R}$.

Consider a general initial wave function $\phi^\varepsilon \in L^2(\mathbb{R}^n)$, such that ϕ^ε corresponds to a probability measure $\rho_{\mathrm{cl}}(dq\,dp)$ on phase space in the sense that for all semiclassical observables with symbols $a \in C^\infty_{\mathrm{b}}(Z)$

$$\lim_{\varepsilon\to 0}\left| \langle \phi^\varepsilon,\, a^{\mathrm{W},\varepsilon}\, \phi^\varepsilon\rangle - \int_Z a(q,p)\,\rho_{\mathrm{cl}}(dq\,dp)\right| = 0\,. \tag{19}$$

The definition is equivalent to saying that the Wigner transform of ϕ^ε converges to ρ_{cl} weakly on test functions in $C^\infty_{\mathrm{b}}(Z)$ [17]. An immediate application of Egorov's theorem yields

$$\lim_{\varepsilon\to 0}\left| \langle \phi^\varepsilon,\, e^{iH^\varepsilon_{\mathrm{BO}}t/\varepsilon}\, a^{\mathrm{W},\varepsilon}\, e^{-iH^\varepsilon_{\mathrm{BO}}t/\varepsilon}\, \phi^\varepsilon\rangle - \int_Z (a\circ\Phi^t)(q,p)\,\rho_{\mathrm{cl}}(dq\,dp)\right| = 0 \tag{20}$$

uniformly on bounded intervals in time, where we recall that Φ^t is the flow generated by (16). In (20) one can of course shift the time evolution from the observables to the states on both sides and write instead

$$\lim_{\varepsilon\to 0}\left| \langle \phi^\varepsilon_t,\, a^{\mathrm{W},\varepsilon}\, \phi^\varepsilon_t\rangle - \int_Z a(q,p)\,\rho_{\mathrm{cl}}(dq\,dp,t)\right| = 0\,. \tag{21}$$

Here $\phi^\varepsilon_t = e^{-iH_{\mathrm{BO}}t/\varepsilon}\phi^\varepsilon$ and $\rho_{\mathrm{cl}}(dq\,dp,t) = (\rho_{\mathrm{cl}}\circ\Phi^{-t})(dq\,dp)$ is the initial distribution $\rho_{\mathrm{cl}}(dq\,dp)$ transported along the classical flow. Thus with respect to certain type of experiments the system described by the wave function ϕ^ε_t behaves like a classical system.

For a molecular system the object of real interest is the left hand side of (21) with ϕ^ε_t replaced by the solution ψ^ε_t of the full Schrödinger equation and $a^{\mathrm{W},\varepsilon} =: a^\varepsilon_{\mathrm{BO}}$ as acting on $L^2(\mathbb{R}^n)$ replaced by $a^{\mathrm{W},\varepsilon}\otimes\mathbf{1}$ as acting on $\mathcal{H}$. In order to compare the expectations of $a^\varepsilon_{\mathrm{BO}}$ with the expectations of $a^{\mathrm{W},\varepsilon}\otimes\mathbf{1}$, we need the following proposition.

Proposition 5. *In addition to the assumptions of Theorem 4 let $a \in C^\infty_{\mathrm{b}}(Z)$ with*

$$\int d\xi \sup_{x\in\mathbb{R}^n} |\xi|\,|\widehat{a}^{(2)}(x,\xi)| < \infty\,, \tag{22}$$

where $\widehat{\ }^{(2)}$ denotes Fourier transformation in the second argument. Then there is a constant $C<\infty$ such that

$$\big\| \big(a^{\mathrm{W},\varepsilon}\otimes\mathbf{1} - \mathcal{U}^*\, a^{\mathrm{W},\varepsilon}\,\mathcal{U}\big)\, 1\!\!1_{\Lambda-\delta}P_*\big\| \le C\,\varepsilon\,.$$

For the proof of Proposition 5 see the end of Section 4.2. With its help we obtain the semiclassical limit for the nuclei as governed by the full Hamiltonian.

Corollary 6. *Let Γ and I be as in Theorem 4. Let $\psi^\varepsilon \in \mathcal{H}$ such that (19) is satisfied for $\phi^\varepsilon := \mathcal{U}P_*\psi^\varepsilon$ for some $\rho_{\rm cl}$ with $\mathrm{supp}\rho_{\rm cl} \subset \Gamma - \alpha$. Let $\psi_t^\varepsilon = e^{-iH^\varepsilon t/\varepsilon}\psi^\varepsilon$ then for all $a \in C_{\rm b}^\infty(Z)$ which satisfy (22)*

$$\lim_{\varepsilon\to 0}\left|\langle\psi_t^\varepsilon,\,(a^{\mathrm{W},\varepsilon}\otimes\mathbf{1})\,\psi_t^\varepsilon\rangle - \int_Z a(q,p)\,\rho_{\rm cl}(dq\,dp,t)\right| = 0 \tag{23}$$

uniformly for $t \in I$.

Translated to the language of Wigner measures Corollary 6 states the following. Let us define the marginal Wigner transform for the nuclei as

$$W_{\rm nuc}^\varepsilon(\psi_t^\varepsilon)(q,p) := (2\pi)^{-n}\int_{\mathbb{R}^n} dX\, e^{iX\cdot p}\,\langle\psi_t^{\varepsilon *}(q+\varepsilon X/2), \psi_t^\varepsilon(q-\varepsilon X/2)\rangle_{\mathcal{H}_{\rm e}}\,.$$

Then, whenever $W_{\rm nuc}^\varepsilon(P_*\psi_0^\varepsilon)(q,p)\,dq\,dp$ converges weakly to some probability measure $\rho_{\rm cl}(dq\,dp)$, $W_{\rm nuc}^\varepsilon(P_*\psi_t^\varepsilon)(q,p)\,dq\,dp$ converges weakly to $(\rho_{\rm cl}\circ\Phi^{-t})(dq\,dp)$.

Corollary 6 follows by applying first Proposition 5 and then Theorem 4 to the left hand side in the difference (23), where we note that $\lim_{\varepsilon\to 0}\|(1-P_\Gamma^\alpha)\psi^\varepsilon\| = 0$ and thus also $\lim_{\varepsilon\to 0}\|(1-P_{\Lambda-\delta'})\psi_t^\varepsilon\| = 0$ for any $\delta' < \delta$. This yields the left hand side of (20) and thus (23).

We mention some standard examples of initial wave functions ϕ^ε of the nuclei which approximate certain classical distributions. The initial wave function for the full system is, as before, recovered as $\psi^\varepsilon = \mathcal{U}^*\phi^\varepsilon = \phi^\varepsilon(X)\chi(X)$. In these examples one regains some control on the rate of convergence with respect to ε which was lost in (19).

(i) Wave packets tracking a classical trajectory.

For $\phi \in L^2(\mathbb{R}^n)$ let

$$\phi_{q_0,p_0}^\varepsilon(X) = \varepsilon^{-\frac{n}{4}}\,e^{-i\frac{p_0\cdot(X-q_0)}{\varepsilon}}\,\phi\Big(\frac{X-q_0}{\sqrt{\varepsilon}}\Big)\,.$$

Then $|\phi_{q_0,p_0}^\varepsilon(X)|^2$ is sharply peaked at q_0 for ε small and its ε-scaled Fourier transform is sharply peaked at p_0. Thus one expects that the corresponding classical distribution is given by $\delta(q-q_0)\delta(p-p_0)\,dq\,dp$. As was shown, e.g. in [23], this is indeed true for $\phi \in L^2(\mathbb{R}^n)$ such that $\phi, |x|\phi, \widehat{\phi}, |p|\widehat{\phi} \in L^1(\mathbb{R}^n)$. Then Corollary 6 holds with (23) replaced by

$$\begin{aligned}&\big|\langle\psi_t^\varepsilon,\,(a^{\mathrm{W},\varepsilon}\otimes\mathbf{1})\,\psi_t^\varepsilon\rangle - a(q(t),p(t))\big| \\ &\quad= O(\sqrt{\varepsilon})\,\Big(\|\phi\|_{L^2}^2 + \|\phi\|_{L^1}\,\||p|\widehat{\phi}\|_{L^1} + \||x|\phi\|_{L^1}\,\|\widehat{\phi}\|_{L^1}\Big)\,,\end{aligned} \tag{24}$$

where $(q(t),p(t))$ is the solution of the classical dynamics with initial condition (q_0,p_0). (24) generalizes Hagedorn's first order result in [9] to a larger class of localized wave functions.

(ii) Either sharp momentum or sharp position.

For $\phi \in L^2(\mathbb{R}^n)$ let

$$\widehat{\phi}_{p_0}^\varepsilon(p) = \widehat{\phi}(p - p_0/\varepsilon)\,,$$

where $\widehat{\ }$ denotes the ε-scaled Fourier transformation, then the corresponding classical distribution is $\rho_{\rm cl}(dq\,dp) = \delta(p-p_0)|\phi(q)|^2\,dq\,dp$. Note that the absolute value of ϕ does not depend on ε in that case. Equivalently one defines

$$\phi_{q_0}^\varepsilon(X) = \varepsilon^{-\frac{n}{2}}\phi\Big(\frac{X-q_0}{\varepsilon}\Big)$$

and obtains $\rho_{\rm cl}(dq\,dp) = \delta(q-q_0)|\widehat{\phi}(p)|^2\,dq\,dp$. In both cases one finds that the difference in (23) is bounded a constant times either $\varepsilon(\|\phi\|^2_{L^2} + \|\phi\|_{L^1}\||p|\widehat{\phi}\|_{L^1})$ for $\phi^\varepsilon_{p_0}$ or $\varepsilon(\|\phi\|^2_{L^2} + \||x|\phi\|_{L^1}\|\widehat{\phi}\|_{L^1})$ for $\phi^\varepsilon_{q_0}$.

(iii) WKB wave functions.

For $f \in L^2(\mathbb{R}^n)$ and $S \in C^1(\mathbb{R}^n)$ both real valued let

$$\phi^\varepsilon(X) = f(X)\,e^{i\frac{S(X)}{\varepsilon}}\,,$$

then $\rho_{\rm cl}(dq\,dp) = f^2(q)\,\delta(p - \nabla S(q))\,dq\,dp$. In this case one expects that (23) is bounded as $\sqrt{\varepsilon}$, which has been shown in [23] for a smaller set of test functions.

4 Proofs

4.1 Globally isolated bands

We collect some immediate consequences of $\mathbf{H}_3$ and $\mathbf{S}$. Using the Riesz formula

$$P_*(X) = -\frac{1}{2\pi i}\oint_{\gamma(X)} d\lambda\, R_\lambda(H_{\rm e}(X))\,, \tag{25}$$

with $\gamma(X)$ a smooth curve in the complex plain circling $\sigma_*(X)$ only and $R_\lambda(H_{\rm e}(X)) = (H_{\rm e}(X) - \lambda)^{-1}$, one easily shows that $P_*(\cdot) \in C^2_{\rm b}(\mathbb{R}^n, \mathcal{L}(\mathcal{H}_{\rm e}))$. Assumption $\mathbf{S}$ enters at this point, since it allows to chose $\gamma(X)$ locally independent of X. Hence, when taking derivatives with respect to X in (25), one only needs to differentiate the integrand. In particular one finds that

$$\begin{aligned}&P_*^\perp(X)(\nabla_X P_*)(X)P_*(X) = \\ &\quad\frac{1}{2\pi i}\oint_{\gamma(X)} d\lambda\, R_\lambda(H_{\rm e}(X))\,P_*^\perp(X)\,(\nabla_X H_{\rm e})(X)\,R_\lambda(H_{\rm e}(X))\,P_*(X)\,.\end{aligned} \tag{26}$$

Since $P_*(X)(\nabla_X P_*)(X)P_*(X) = P_*^\perp(X)(\nabla_X P_*)(X)P_*^\perp(X) = 0$, which follows from $(\nabla_X P_*)(X) = (\nabla_X P_*^2)(X) = (\nabla_X P_*)(X)P_*(X) + P_*(X)(\nabla_X P_*)(X)$, we have that

$$(\nabla_X P_*)(X) = P_*^\perp(X)(\nabla_X P_*)(X)P_*(X) + \text{ adjoint}\,. \tag{27}$$

In (27) and in the following "+ adjoint" means that the adjoint operator of the first term in a sum is added.

Starting with (12), we find, at the moment formally, that

$$\begin{aligned}e^{-iH^\varepsilon_{\rm diag}t/\varepsilon} - e^{-iH^\varepsilon t/\varepsilon} &= e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\left(\mathbf{1} - e^{iH^\varepsilon_{\rm diag}t/\varepsilon}\,e^{-iH^\varepsilon t/\varepsilon}\right) = \\ &= i\,e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\int_0^{t/\varepsilon} ds\, e^{iH^\varepsilon_{\rm diag}s}\,(H^\varepsilon - H^\varepsilon_{\rm diag})e^{-iH^\varepsilon s}\,,\end{aligned} \tag{28}$$

where

$$\begin{aligned}H^\varepsilon - H^\varepsilon_{\rm diag} &= P_*^\perp\,H^\varepsilon\,P_* + \text{ adjoint} \\ &= P_*^\perp\left[\frac{\varepsilon^2}{2}\Big(-i\nabla_X + A_{\rm ext}(X)\Big)^2, P_*\right]P_* + \text{ adjoint}\,.\end{aligned} \tag{29}$$

Let $D_A := -i\nabla_X + A_{\rm ext}(X)$. Then the commutator is easily calculated as

$$\begin{aligned}\left[\frac{\varepsilon^2}{2}(D_A\otimes\mathbf{1})^2, P_*\right] &= -i\,\varepsilon\,(\nabla_X P_*)\,\cdot\,(\varepsilon D_A\otimes\mathbf{1}) + O(\varepsilon^2) &(30)\\ &= -\varepsilon\,(\nabla_X P_*)\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1}) + O(\varepsilon^2)\,, &(31)\end{aligned}$$

where $O(\varepsilon^2)$ holds in the norm of $\mathcal{L}(\mathcal{H},\mathcal{H})$ as $\varepsilon\to 0$. For (30) and (31) it was used that $A_{\rm ext}(X)$ and $P_*(X)$ are both differentiable with bounded derivatives and that $A_{\rm ext}(X)$ commutes with P_*.

Before we can continue, we need to justify (28) by showing that $H^\varepsilon_{\rm diag}$ is self-adjoint on $D(H^\varepsilon)$. To see this, note that $-i\varepsilon\nabla_X$ is bounded with respect to $\varepsilon^2\Delta_X$ with relative bound 0 and that for $\psi\in D(H^\varepsilon)$

$$\begin{aligned}\|(\varepsilon^2\Delta_X\otimes\mathbf{1})\,\psi\| &\le c_1\,\left(\|(\varepsilon^2 D_A^2\otimes\mathbf{1})\,\psi\| + \|\psi\|\right)\\ &\le c_2\,\left(\|(\varepsilon^2 D_A^2\otimes\mathbf{1} + \mathbf{1}\otimes H_0)\,\psi\| + \|\psi\|\right)\\ &\le c_3\,\left(\|H^\varepsilon\,\psi\| + \|\psi\|\right)\,, &(32)\end{aligned}$$

where we used that $H_{\rm e0}$ is bounded from below and that $H_{\rm e1}$ is bounded. Hence $H^\varepsilon - H^\varepsilon_{\rm diag}$ is infinitesimally operator bounded with respect to H^ε, consequently $H^\varepsilon_{\rm diag}$ is self-adjoint on $D(H^\varepsilon)$ and thus (28) holds on $D(H^\varepsilon)$.

(29) and (31) in (28) give

$$\begin{aligned}&P_*^\perp\left(e^{-iH^\varepsilon_{\rm diag}t/\varepsilon} - e^{-iH^\varepsilon t/\varepsilon}\right) = &(33)\\ &= -i\varepsilon\, e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\int_0^{t/\varepsilon} ds\, e^{iH^\varepsilon_{\rm diag}s}\,P_*^\perp\,(\nabla_X P_*)\,P_*\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1})\,e^{-iH^\varepsilon s} + O(\varepsilon)|t|\,,\end{aligned}$$

where we used that the term of order $O(\varepsilon^2)$ in (31) yields a term of order $O(\varepsilon)|t|$ after integration, since all other expressions in the integrand are bounded uniformly in time and the domain of integration grows like t/ε. In (33) and in the following we omit the adjoint term from (29) and thus consider the difference of the groups projected on ${\rm Ran}P_*^\perp$ only. The argument for the difference projected on ${\rm Ran}P_*$ goes through analogously by taking adjoints at the appropriate places.

Now $\varepsilon(\nabla_X P_*)\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1})$ is only $O(\varepsilon)$ in the norm of $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$ and thus, according to the naive argument, only $O(1)|t|$ after integration. As in [13] and [23] we proceed by writing $(\nabla_X P_*)\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1})$ as the commutator of a bounded operator B with H^ε modulo terms of order $O(\varepsilon)$. This is in analogy to the proof of the time-adiabatic theorem [15] and allows one to write the first order part of the integrand in (33) as the time derivative of a bounded operator and, as a consequence, to do the integration without losing one order in ε.

In view of (26) we define

$$\widetilde{B}(X) := \frac{1}{2\pi i}\oint_{\gamma(X)} d\lambda\, R_\lambda(H_{\rm e}(X))^2\,P_*^\perp(X)\,(\nabla_X H_{\rm e})(X)\,R_\lambda(H_{\rm e}(X))\,P_*(X)\,. \tag{34}$$

An easy calculation shows that

$$\left[H_{\rm e},\widetilde{B}\right] = -\,P_*^\perp\,(\nabla_X P_*)\,P_*\,. \tag{35}$$

By assumption $\partial_{X_j}H_{\rm e}(X)\in C^2(\mathbb{R}^n,\mathcal{L}(\mathcal{H}_{\rm e}))$, $j=1,\dots,n$, hence $\widetilde{B}_j(X)\in C^2(\mathbb{R}^n,\mathcal{L}(\mathcal{H}_{\rm e}))$ and thus

$$\left[\frac{\varepsilon^2}{2}D_A^2\otimes\mathbf{1},\widetilde{B}\right] = -\varepsilon\,(\nabla_X\widetilde{B})\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1}) + O(\varepsilon^2) = O(\varepsilon) \tag{36}$$

in the norm of $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$. (35) and (36) combined yield that

$$\left[H^\varepsilon,\widetilde{B}\right] = -\,P_*^\perp\,(\nabla_X P_*)\,P_* + O(\varepsilon)$$

with $O(\varepsilon)$ in the norm of $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$. Since $\nabla_X H_{\rm e}\in\mathcal{L}(\mathcal{H})$, a short calculation shows that $[H^\varepsilon,\varepsilon\nabla_X\otimes\mathbf{1}] = O(\varepsilon)$ in $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$. Hence we define

$$B := \widetilde{B}\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1})$$

and obtain

$$[H^\varepsilon,B] = -\,P_*^\perp\,(\nabla_X P_*)\,P_*\,\cdot\,(\varepsilon\nabla_X\otimes\mathbf{1}) + O(\varepsilon)$$

with $O(\varepsilon)$ in the norm of $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$. Let

$$B(s) = e^{iH^\varepsilon s}\,B\,e^{-iH^\varepsilon s}$$

then

$$-i\frac{d}{ds}\,B(s) = e^{iH^\varepsilon s}\,[H^\varepsilon,B]\,e^{-iH^\varepsilon s}\,.$$

Continuing (33), we have

$$\begin{aligned}
&P_*^\perp\left(e^{-iH^\varepsilon_{\rm diag}t/\varepsilon} - e^{-iH^\varepsilon t/\varepsilon}\right) = \\
&= \; i\,\varepsilon\,e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\int_0^{t/\varepsilon} ds\,e^{iH^\varepsilon_{\rm diag}s}\,[H^\varepsilon,B]\;e^{-iH^\varepsilon s} + O(\varepsilon)(|t|+|t|^2) \\
&= \; \varepsilon\,e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\int_0^{t/\varepsilon} ds\,e^{iH^\varepsilon_{\rm diag}s}\,e^{-iH^\varepsilon s}\left(\frac{d}{ds}B(s)\right) + O(\varepsilon)(|t|+|t|^2)\,,
\end{aligned} \tag{37}$$

where $O(\varepsilon)$ holds now in the norm of $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$. The additional factor of $|t|$ in (37) comes from the fact that

$$\|e^{-iH^\varepsilon s}\|_{\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e})} \le c\,(1+\varepsilon\,|s|) \tag{38}$$

for some constant $c<\infty$, i.e. the scaled momentum of the nuclei may grow in time. Using $\|A_{\rm ext}\|_\infty = C < \infty$ and

$$\|[(\varepsilon D_A\otimes\mathbf{1}),H^\varepsilon\,]\|_{\mathcal{L}(\mathcal{H})} \le \widetilde{C}\,\varepsilon\,,$$

(38) follows from

$$\begin{aligned}
\left\|(-i\varepsilon\nabla_X\otimes\mathbf{1})\,e^{-iH^\varepsilon s}\,\psi\right\| &\le \left\|(\varepsilon D_A\otimes\mathbf{1})\,e^{-iH^\varepsilon s}\,\psi\right\| + \left\|(\varepsilon A_{\rm ext}\otimes\mathbf{1})\,e^{-iH^\varepsilon s}\,\psi\right\| \\
&\le \; \|(\varepsilon D_A\otimes\mathbf{1})\,\psi\| + \left\|\left[(\varepsilon D_A\otimes\mathbf{1}),\,e^{-iH^\varepsilon s}\right]\,\psi\right\| + C\,\|\psi\| \\
&\le \; \|(-i\varepsilon\nabla_X\otimes\mathbf{1})\,\psi\| \;+\; \widetilde{C}\,\varepsilon\,|s|\,\|\psi\| \;+\; 2\,C\,\|\psi\|
\end{aligned}$$

for $\psi\in W^1\otimes\mathcal{H}_{\rm e}$.

Finally, continuing (37), integration by parts yields

$$\begin{aligned}
&P_*^\perp\left(e^{-iH^\varepsilon_{\rm diag}t/\varepsilon} - e^{-iH^\varepsilon t/\varepsilon}\right) = \\
&= \; \varepsilon\,e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\int_0^{t/\varepsilon} ds\,e^{iH^\varepsilon_{\rm diag}s}\,e^{-iH^\varepsilon s}\left(\frac{d}{ds}B(s)\right) + O(\varepsilon)(|t|+|t|^2) \\
&= \; \varepsilon\left(B\,e^{-iH^\varepsilon t/\varepsilon} - e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\,B\right) \\
&\quad + i\,\varepsilon\,e^{-iH^\varepsilon_{\rm diag}t/\varepsilon}\int_0^{t/\varepsilon} ds\,e^{iH^\varepsilon_{\rm diag}s}\,(H^\varepsilon - H^\varepsilon_{\rm diag})B\,e^{-iH^\varepsilon s} + O(\varepsilon)(|t|+|t|^2) \\
&= \; O(\varepsilon)(1+|t|)^3\,,
\end{aligned} \tag{39}$$

where $O(\varepsilon)$ holds in the norm of $\mathcal{L}(W^{2,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$. For the last equality we used that B is bounded in $\mathcal{L}(W^{2,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$ as well as in $\mathcal{L}(W^{2,\varepsilon}\otimes\mathcal{H}_{\rm e},W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e})$ uniformly with respect to ε, $H^\varepsilon - H^\varepsilon_{\rm diag}$ is $O(\varepsilon)$ in $\mathcal{L}(W^{1,\varepsilon}\otimes\mathcal{H}_{\rm e},\mathcal{H})$, as we saw in (29) and (31), and

$$\left\|e^{-iH^\varepsilon s}\right\|_{\mathcal{L}(W^{2,\varepsilon}\otimes\mathcal{H}_{\rm e})} \leq c\,(1+\varepsilon\,|s|)^2 \tag{40}$$

for some constant $c<\infty$. (40) follows from arguments similar to those used in the proof of (38).

We are left to prove (13). This follows from exactly the same proof using that $\mathcal{E}(H^\varepsilon)$ commutes with $e^{-iH^\varepsilon s}$ and that, according to (32),

$$\|(\varepsilon^2\Delta_X\otimes\mathbf{1})\,\mathcal{E}(H^\varepsilon)\,\psi\| \leq c_3\,\left(\|H^\varepsilon\,\mathcal{E}(H^\varepsilon)\,\psi\| + \|\psi\|\right) \leq c_4\,(|\mathcal{E}|+1)\,\|\psi\|\,.$$

4.2 Locally isolated bands

To prove Theorem 4 we proceed along the same lines as in the previous section, with the one modification that we use Proposition 3 to control the flux out of $\partial\Lambda$. However, one cannot use $P_* = \int_\Lambda^\oplus dX\,P_*(X)$ to define $H^\varepsilon_{\rm diag}$ anymore, because the functions in its range would not be in the range of H^ε and some smoothing in the cutoff is needed. For $i\in\{0,1,2,3\}$ let $1\!\!1_i = 1\!\!1_{(\Lambda-\frac{4-i}{5}\delta,\frac{1}{5}\delta)}$ be approximate characteristic functions according to Definition 2. Then the smoothed projections are defined with $P_i(X) = 1\!\!1_i(X)\,P_*(X)$ as $P_i = \int^\oplus dX\,P_i(X)$. In the following it will be used that for $i<j$ we have $P_iP_j = P_jP_i = P_i$ and hence $(1-P_j)P_i = P_i(1-P_j) = 0$.

Proposition 3 yields

$$\left(e^{-iH^\varepsilon t/\varepsilon} - \mathcal{U}^*\,e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}\,\mathcal{U}\right)\,P^\alpha_\Gamma \;=\; \left(e^{-iH^\varepsilon t/\varepsilon} - P_1\,\mathcal{U}^*\,e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}\,\mathcal{U}\right)\,P^\alpha_\Gamma + O(\varepsilon)\,. \tag{41}$$

We make also use of the fact that the phase space support of the initial wave function lies in Γ and has thus bounded energy with respect to $H_{\rm cl}$. Let $E := \sup_{z\in\Gamma}H_{\rm cl}(z) < \infty$, let $1\!\!1_{((-\infty,E+\alpha),\alpha)}$ be a smooth characteristic function on $\mathbb{R}$ and let $\mathcal{E} := \left(1\!\!1_{((-\infty,E+\alpha),\alpha)}(H_{\rm cl}(\cdot))\right)^{\rm W,\varepsilon}$. Then standard results from semiclassical analysis imply the following relations.

Proposition 7.

(a) $1\!\!1^{\rm W,\varepsilon}_{(\Gamma,\alpha)} = \mathcal{E}\,1\!\!1^{\rm W,\varepsilon}_{(\Gamma,\alpha)} + O(\varepsilon)$;

(b) $e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}\,\mathcal{E} = \mathcal{E}\,e^{-iH^\varepsilon_{\rm BO}t/\varepsilon} + O(\varepsilon)$ *uniformly for* $t\in I$;

(c) $[\,H^\varepsilon_{\rm BO},\,\mathcal{E}\,] = O(\varepsilon^2)$;

(d) $\mathcal{E}\in\mathcal{L}(L^2(\mathbb{R}^n),W^{2,\varepsilon})$.

In (a)–(c) $O(\varepsilon)$ resp. $O(\varepsilon^2)$ hold in the norm of $\mathcal{L}(L^2(\mathbb{R}^n))$.

Proposition (7) (a), (c) and (d) are direct consequences of the product rule for pseudo-differential operators (see, e.g., [21, 6]) and (b) is again Egorov's Theorem.

Using Proposition 7 (a) and (b) we continue (41) and obtain

$$\left(e^{-iH^\varepsilon t/\varepsilon} - P_1\,\mathcal{U}^*\,e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}\,\mathcal{U}\right)\,P^\alpha_\Gamma \;=\; \left(e^{-iH^\varepsilon t/\varepsilon} - P_1\,\mathcal{U}^*\,\mathcal{E}\,e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}\,\mathcal{U}\right)\,P^\alpha_\Gamma + O(\varepsilon)\,. \tag{42}$$

We proceed as in the globally isolated band case and write

$$
\begin{aligned}
&\left(e^{-iH^\varepsilon t/\varepsilon} - P_1\,\mathcal{U}^*\,\mathcal{E}\,e^{-iH^\varepsilon_{\rm BO}t/\varepsilon}\,\mathcal{U}\right)\,P^\alpha_\Gamma\\
&= \; -ie^{-iH^\varepsilon t/\varepsilon}\int_0^{t/\varepsilon} ds\, e^{iH^\varepsilon s}\,(H^\varepsilon\,P_1\,\mathcal{U}^*\,\mathcal{E} \;-\; P_1\,\mathcal{U}^*\,\mathcal{E}\,H^\varepsilon_{\rm BO})e^{-iH^\varepsilon_{\rm BO}s}\,\mathcal{U}\,P^\alpha_\Gamma\\
&= \; -ie^{-iH^\varepsilon t/\varepsilon}\int_0^{t/\varepsilon} ds\, e^{iH^\varepsilon s}\,(H^\varepsilon \;-\; H^\varepsilon_{\rm diag})\,P_1\,\mathcal{U}^*\,\mathcal{E}e^{-iH^\varepsilon_{\rm BO}s}\,\mathcal{U}\,P^\alpha_\Gamma \qquad (43)\\
&\quad -ie^{-iH^\varepsilon t/\varepsilon}\int_0^{t/\varepsilon} ds\, e^{iH^\varepsilon s}\,(H^\varepsilon_{\rm diag}\,P_1\,\mathcal{U}^*\,\mathcal{E} - \; P_1\,\mathcal{U}^*\,\mathcal{E}\,H^\varepsilon_{\rm BO})e^{-iH^\varepsilon_{\rm BO}s}\,\mathcal{U}\,P^\alpha_\Gamma\,, \qquad (44)
\end{aligned}
$$

where

$$H^\varepsilon_{\rm diag} := P_3\,H^\varepsilon\,P_3\,.$$

One can now show that (43) is bounded in norm by a constant times $\varepsilon(1+|t|)$ using exactly the same sequence of arguments as in the proof in the previous section. One must only keep track of the "hierarchy" of smoothed projections, e.g., instead of (29) one has

$$(H^\varepsilon \;-\; H^\varepsilon_{\rm diag})P_1 = (1-P_3)\left[-\frac{\varepsilon^2}{2}\Delta_X\otimes\mathbf{1}, P_2\right]P_1 \;+\; O(\varepsilon^2).$$

The adjoint part drops out completely, because this time only the difference on the band, i.e. on $\mathrm{Ran}P_1$, is of interest. Note also that the smoothed projections P_i are bounded operators on the respective scaled Sobolev spaces and thus, according to Proposition 7 (d), all estimates hold in the norm of $\mathcal{L}(\mathcal{H})$.

It remains to show that also (44) is $O(\varepsilon)$. First note that, according to Proposition 7 (c), commuting $\mathcal{E}$ and $H^\varepsilon_{\rm BO}$ yields an error of order $O(\varepsilon^2)$ in the integrand and thus an error of order $O(\varepsilon)$ after integration. For $\phi\in W^2$ we compute

$$
\begin{aligned}
(H^\varepsilon_{\rm diag}\,P_1\,\mathcal{U}^*\phi)(X) \;&=\; \mathbb{1}_1(X)\,E(X)\,\phi(X)\chi(X) + \mathbb{1}_1(X)\left(\frac{\varepsilon^2}{2}(-i\nabla_X + A_{\rm ext})^2\,\phi\right)(X)\,\chi(X)\\
&\quad +\varepsilon\,\mathbb{1}_1(X)\,(-i\varepsilon\nabla\phi)\,(X)\cdot(-i\langle\chi(X),\nabla_X\chi(X)\rangle_{\mathcal{H}_{\rm e}})\;\chi(X)\\
&\quad -i\,\varepsilon\,(\nabla\mathbb{1}_1)(X)\cdot(-i\varepsilon\nabla\phi)\,(X)\,\chi(X) \;+\; O(\varepsilon^2)\,. \qquad (45)
\end{aligned}
$$

On the other hand, again for $\phi\in W^2$,

$$
\begin{aligned}
(P_1\,\mathcal{U}^*\,H^\varepsilon_{\rm BO}\,\phi)(X) \;&=\; \mathbb{1}_1(X)\,E(X)\,\phi(X)\chi(X) + \mathbb{1}_1(X)\left(\frac{\varepsilon^2}{2}(-i\nabla_X + A_{\rm ext})^2\,\phi\right)(X)\,\chi(X)\\
&\quad +\varepsilon\,\mathbb{1}_1(X)\,(-i\varepsilon\nabla\phi)\,(X)\cdot A_{\rm geo}(X)\,\chi(X) \;+\; O(\varepsilon^2)\,. \qquad (46)
\end{aligned}
$$

Hence

$$H^\varepsilon_{\rm diag}\,P_1\,\mathcal{U}^*\mathcal{E} - P_1\,\mathcal{U}^*\,H^\varepsilon_{\rm BO}\,\mathcal{E} = -\varepsilon\,\mathcal{U}^*\,(\nabla\mathbb{1}_1)\cdot\varepsilon\nabla_X\,\mathcal{E} + O(\varepsilon^2)$$

Thus the norm of (44) is, up to an error of order $O(\varepsilon)$, bounded by the norm of

$$\varepsilon\,\mathcal{U}^*\int_0^{t/\varepsilon} ds\,(\nabla\mathbb{1}_1)\cdot\varepsilon\nabla_X\,\mathcal{E}\,e^{-iH^\varepsilon_{\rm BO}s}\,\mathcal{U}\,P^\alpha_\Gamma\,. \qquad (47)$$

$(\nabla\mathbb{1}_1)\cdot\varepsilon\nabla_X\,\mathcal{E}$ is a bounded operator and we can apply Proposition 3 in the integrand of (47) once more, this time however with the smoothed projection P_0, and obtain

$$(47) \;=\; \varepsilon\,\mathcal{U}^*\int_0^{t/\varepsilon} ds\,(\nabla\mathbb{1}_1)\cdot\varepsilon\nabla_X\,\mathcal{E}\,\mathbb{1}_0\,e^{-iH^\varepsilon_{\rm BO}s}\,\mathcal{U}\,P^\alpha_\Gamma \;+ O(\varepsilon) = O(\varepsilon). \qquad (48)$$

The last equality in (48) follows from the fact that $[\varepsilon\nabla_X \mathcal{E}, 1\!\!1_0] = O(\varepsilon)$ and that $(\nabla 1\!\!1_1)$ and $1\!\!1_0$ are disjointly supported.

Proof of Proposition 5. For the following calculations we continue $\chi(\cdot) \in C_{\mathrm{b}}^\infty(\Lambda, \mathcal{H}_{\mathrm{e}})$ arbitrarily to a function $\chi(\cdot) \in C_{\mathrm{b}}^\infty(\mathbb{R}^n, \mathcal{H}_{\mathrm{e}})$ by possibly modifying it on $\Lambda \setminus (\Lambda - \delta/2)$. For ϕ in a dense subset of $L^2(\Lambda - \delta)$ and $X \in \Lambda - \delta/2$, by making the substitutions $\widetilde{k} = \varepsilon k$ and $\widetilde{Y} = (Y - X)/\varepsilon$ and using Taylor expansion with rest, we have:

$$
\begin{aligned}
\big((a^{\mathrm{W},\varepsilon} \otimes \mathbf{1})\, \phi\chi\big)(X) = (2\pi)^{-n} & \int dY\, dk\, a\left(\frac{X+Y}{2}, \varepsilon k\right) e^{-i(X-Y)\cdot k}\, \phi(Y)\chi(Y) \\
= \;& (2\pi)^{-n} \int d\widetilde{Y}\, \widehat{a}^{(2)}\left(X + \frac{\varepsilon}{2}\widetilde{Y}, -\widetilde{Y}\right) \phi(X + \varepsilon\widetilde{Y})\, \chi(X) \\
& + \varepsilon\, (2\pi)^{-n} \int d\widetilde{Y}\, \widehat{a}^{(2)}\left(X + \frac{\varepsilon}{2}\widetilde{Y}, -\widetilde{Y}\right) \phi(X + \varepsilon\widetilde{Y})\, \widetilde{Y} \cdot (\nabla_X \chi)(f(X, \varepsilon\widetilde{Y})) \\
= \;& \big(\mathcal{U}^*\, a^{\mathrm{W},\varepsilon}\, \mathcal{U}\, \phi\chi\big)(X) \,+\, R^\varepsilon\,.
\end{aligned}
\tag{49}
$$

From (49) we conclude that

$$
\left\| (1\!\!1_{\Lambda-\delta/2}(\cdot) \otimes \mathbf{1})\, \left(a^{\mathrm{W},\varepsilon} \otimes \mathbf{1} \,-\, \mathcal{U}^*\, a^{\mathrm{W},\varepsilon}\, \mathcal{U}\right) P_{\Lambda-\delta} \right\| \leq \|R^\varepsilon\|\,. \tag{50}
$$

Since

$$
\begin{aligned}
& \left\| (\mathbf{1} - 1\!\!1_{\Lambda-\delta/2}(\cdot) \otimes \mathbf{1})\, \left(a^{\mathrm{W},\varepsilon} \otimes \mathbf{1} \,-\, \mathcal{U}^*\, a^{\mathrm{W},\varepsilon}\, \mathcal{U}\right) P_{\Lambda-\delta} \right\| \\
& \quad = \; \left\| (\mathbf{1} - 1\!\!1_{\Lambda-\delta/2}(\cdot) \otimes \mathbf{1})\, \left(a^{\mathrm{W},\varepsilon} \otimes \mathbf{1} \,-\, \mathcal{U}^*\, a^{\mathrm{W},\varepsilon}\, \mathcal{U}\right) (1\!\!1_{\Lambda-\delta}(\cdot) \otimes \mathbf{1}) P_{\Lambda-\delta} \right\| = O(\varepsilon^n)
\end{aligned}
$$

for arbitrary n, Proposition 5 follows by showing that R^ε is of order ε:

$$
\begin{aligned}
\|R^\varepsilon\| \;\leq\;& \varepsilon\, (2\pi)^{-n} \int d\widetilde{Y}\, \left\| \widehat{a}^{(2)}\left(\cdot + \frac{\varepsilon}{2}\widetilde{Y}, -\widetilde{Y}\right) \phi(\cdot + \varepsilon\widetilde{Y})\, \widetilde{Y} \cdot (\nabla_X \chi)(f(\cdot, \varepsilon\widetilde{Y})) \right\|_{\mathcal{H}} \\
\leq\;& \varepsilon\, (2\pi)^{-n} \sup_{X \in \mathbb{R}^n} \|(\nabla_X \chi)(X)\|_{\mathcal{H}_{\mathrm{e}}} \int d\widetilde{Y}\, \left\| \widehat{a}^{(2)}\left(\cdot + \frac{\varepsilon}{2}\widetilde{Y}, -\widetilde{Y}\right) |\widetilde{Y}|\, \phi(\cdot + \varepsilon\widetilde{Y}) \right\|_{L^2(\mathbb{R}^n)} \\
\leq\;& \varepsilon\, C\, \|\phi\|_{L^2(\mathbb{R}^n)} \int d\widetilde{Y} \sup_{X \in \mathbb{R}^n} |\widetilde{Y}|\, |\widehat{a}^{(2)}(X, \widetilde{Y})| \\
=\;& \varepsilon\, \widetilde{C}\, \|\phi\chi\|_{\mathcal{H}}\,.
\end{aligned}
$$

Acknowledgment: We are grateful to André Martinez and Gheorghe Nenciu for explaining to us their work in great detail. S. T. would like to thank George Hagedorn for stimulating discussions and, in particular, for helpful advice on questions concerning the Berry connection and Markus Klein and Ruedi Seiler for explaining their treatment of Coulomb singularities. We thank Caroline Lasser and Gianluca Panati for careful reading of the manuscript and the referee for pointing out Reference [12].

References

[1] J. E. Avron and A. Elgart. *Adiabatic theorems without a gap condition*, Commun. Math. Phys. **203**, 445–463 (1999).

[2] F. Bornemann and C. Schütte. *On the singular limit of the quantum-classical molecular dynamics model*, SIAM J. Appl. Math. **59**, 1208-1224 (1999).

[3] M. Born and R. Oppenheimer. *Zur Quantentheorie der Molekeln*, Ann. Phys. (Leipzig) **84**, 457–484 (1927).

[4] A. Bouzouina and D. Robert. *Uniform semi-classical estimates for the propagation of Heisenberg observables*, Math. Phys. Preprint Archive mp_arc 99-409 (1999).

[5] J.-M. Combes, P. Duclos, R. Seiler. *The Born-Oppenheimer approximation*, in: Rigorous Atomic and Molecular Physics (eds. G. Velo, A. Wightman), New York, Plenum, 185–212 (1981).

[6] M. Dimassi and J. Sjöstrand. *Spectral Asymptotics in the Semi-Classical Limit*, London Mathematical Society Lecture Note Series 268, Cambridge University Press (1999).

[7] C. Fermanian Kammerer and P. Gérard. *A Landau-Zener formula for two-scaled Wigner measures*, preprint (2001).

[8] G.A. Hagedorn. *High order corrections to the time-independent Born-Oppenheimer approximation I: smooth potentials*, Ann. Inst. H. Poincaré Sect. A **47**, 1–19 (1987).

[9] G.A. Hagedorn. *A time dependent Born-Oppenheimer approximation*, Comm. Math. Phys. **77**, 1–19 (1980).

[10] G.A. Hagedorn and A. Joye. *A time-dependent Born-Oppenheimer approximation with exponentially small error estimates*, Math. Phys. Preprint Archive mp_arc 00-209 (2000).

[11] G.A. Hagedorn. *Molecular Propagation Through Electronic Eigenvalue Crossings*, Memoirs Amer. Math. Soc. **536** (1994).

[12] J. Herrin and J. S. Howland. *The Born-Oppenheimer approximation: straight-up and with a twist*, Rev. Math. Phys. **9**, 467–488 (1997).

[13] F. Hövermann, H. Spohn, S. Teufel. *Semiclassical limit for the Schrödinger equation with a short scale periodic potential*, Commun. Math. Phys. **215**, 609–629 (2001).

[14] A. Joye and C.-E. Pfister. *Quantum adiabatic evolution*, in: On Three Levels (eds. M. Fannes, C. Maes, A. Verbeure), Plenum, New York, 139–148 (1994).

[15] T. Kato. *On the adiabatic theorem of quantum mechanics*, Phys. Soc. Jap. **5**, 435–439 (1958).

[16] M. Klein, A. Martinez, R. Seiler, X.P. Wang. *On the Born-Oppenheimer expansion for polyatomic molecules*, Commun. Math. Phys. **143**, 607–639 (1992).

[17] P. L. Lions and T. Paul. *Sur les mesures de Wigner*, Revista Mathematica Iberoamericana **9**, 553–618 (1993).

[18] A. Martinez and V. Sordoni. *On the time-dependent Born-Oppenheimer approximation with smooth potential*, Math. Phys. Preprint Archive mp_arc 01-37 (2001).

[19] C.A. Mead and D.G. Truhlar. *On the determination of Born-Oppenheimer nuclear motion wave functions including complications due to conical intersections and identical nuclei*, J. Chem. Phys. **70**, 2284–2296 (1979).

[20] G. Nenciu and V. Sordoni. *Semiclassical limit for multistate Klein-Gordon systems: almost invariant subspaces and scattering theory*, Math. Phys. Preprint Archive mp_arc 01-36 (2001).

[21] D. Robert. *Autour de l'Approximation Semi-Classique*, Progress in Mathematics, Volume 68, Birkhäuser (1987).

[22] A. Shapere and F. Wilczek (Eds.). *Geometric Phases in Physics*, World Scientific, Singapore (1989).

[23] S. Teufel and H. Spohn. *Semi-classical motion of dressed electrons*, Preprint ArXiv.org math-ph/0010009, to appear in Rev. Math. Phys. (2001).

[24] S. Teufel. *Adiabatic decoupling for perturbations of fibered Hamiltonians*, in preparation.