Semi-classical motion of dressed electrons

Stefan Teufel and Herbert Spohn

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

October 5, 2000

Abstract

We consider an electron coupled to the quantized radiation field and subject to a slowly varying electrostatic potential. We establish that over sufficiently long times radiation effects are negligible and the dressed electron is governed by an effective one-particle Hamiltonian. In the proof only a few generic properties of the full Pauli-Fierz Hamiltonian $H_{\rm PF}$ enter. Most importantly, $H_{\rm PF}$ must have an isolated ground state band for $|p| < p_{\rm c} \leq \infty$ with p the total momentum and $p_{\rm c}$ indicating that the ground state band may terminate. This structure demands a local approximation theorem, in the sense that the one-particle approximation holds until the semi-classical dynamics violates $|p| < p_{\rm c}$. Within this framework we prove an abstract Hilbert space theorem which uses no additional information on the Hamiltonian away from the band of interest. Our result is applicable to other time-dependent semi-classical problems. We discuss semi-classical distributions for the effective one-particle dynamics and show how they can be translated to the full dynamics by our results.

1 Introduction

Electrons, protons, and other elementary charged particles move in essence along classical orbits provided the external potentials have a slow variation. E.g. in accelerators it is safe to compute the orbits by means of classical mechanics. As a somewhat crude physical picture one imagines that the electron dressed with its photon cloud at any given instant t is in its state of lowest energy consistent with the current total momentum $p(t)$. During the time span δt the external forces change the momentum of the electron followed by a rapid adjustment of the photon cloud resulting in the new total momentum $p(t+\delta t)$. In this picture two physical mechanisms are interlaced. Since the forces are weak, the acceleration is small and radiative losses can be ignored. The effective dynamics of the dressed electron is conservative within a good approximation. In addition the total momentum $p(t)$ has a slow variation and can be regarded as a semiclassical variable with respect to the effective Hamiltonian of the ground state band (lowest energy shell). The goal of our paper is to establish the validity of the physical picture.

To set the scene, let us introduce the quantum Hamiltonian under consideration. In fact, as will be explained below, our technique is fairly general and uses only a few generic properties of the Hamiltonian. Still, it is useful to have a specific example in mind. We consider a free electron coupled to the quantized Maxwell field. The Hilbert space of states for the electron is $\mathcal{H}_{\rm el} = L^2(\mathbb{R}^3)$ and its time evolution is governed by the Hamiltonian $-\frac{1}{2m}\Delta$. m is the mass of the electron and we have set $\hbar = 1$. For the photon field we introduce the Fock space $\mathcal{F}_{\rm b} = \mathcal{H}_{\rm f}$ over the one-particle space $L^2(\mathbb{R}^3)\otimes\mathbb{C}^2$, i.e. $\mathcal{F}_{\rm b} = \bigoplus_{n=0}^{\infty} S_n(L^2(\mathbb{R}^3)\otimes\mathbb{C}^2)^{\otimes n}$ and S_n the symmetrizer. Thus a state $\phi\in\mathcal{F}_{\rm b}$ is a sequence of vectors $\{\phi^{(0)},\phi^{(1)},\ldots\}$ with $\phi^{(n)}\in S_n(L^2(\mathbb{R}^3)\otimes\mathbb{C}^2)^{\otimes n}$ such that $\|\phi\|^2 = \sum_{n=0}^{\infty}\|\phi^{(n)}\|^2 < \infty$. On $\mathcal{F}_{\rm b}$ we define the two component Bose field with annihilation operators $a(k,\lambda)$, where $k\in\mathbb{R}^3$ stands for the wave number and $\lambda = 1,2$ for the helicity of the photon. The fields satisfy the canonical commutation relations $[a(k,\lambda),a^*(k',\lambda')] = \delta(k-k')\delta_{\lambda\lambda'}$, $[a(k,\lambda),a(k',\lambda')] = 0 = [a^*(k,\lambda),a^*(k',\lambda')]$. The Hamiltonian of the free photon field is the given by

$$H_{\rm f} = \sum_{\lambda=1,2}\int d^3k\,\omega(k)a^*(k,\lambda)a(k,\lambda) \tag{1}$$

with dispersion relation $\omega(k) = |k|$, the velocity of light $c = 1$. The electron and the photon field are minimally coupled through the transverse vector potential $A(x)$. To assure transversality we introduce the standard dreibein $e_1(k)$, $e_2(k)$, $k/|k|$. Then

$$A(x) = (2\pi)^{-3/2}\sum_{\lambda=1,2}\int d^3k\,\frac{1}{\sqrt{2\omega(k)}}e_\lambda(k)\left(e^{ik\cdot x}a(k,\lambda)+e^{-ik\cdot x}a^*(k,\lambda)\right)\,. \tag{2}$$

A is an operator valued distribution. To make it an unbounded operator we smoothen over the form factor ρ as

$$A_\rho(x) = \int d^3x'\,\rho(x-x')A(x') \tag{3}$$

and assume that ρ is radial, smooth, of rapid decrease, and normalized as $\int d^3x\,\rho(x) = 1$. In the corresponding classical Hamiltonian ρ would be the rigid charge distribution. With all these preparations we can introduce the Pauli-Fierz operator of a free electron as

$$H_0 = \frac{1}{2m}\big(-i\nabla\otimes\mathbf{1} - eA_\rho(x)\big)^2 + \mathbf{1}\otimes H_{\mathrm{f}} \tag{4}$$

acting on $\mathcal{H} = \mathcal{H}_{\mathrm{el}}\otimes\mathcal{H}_{\mathrm{f}}$. e is the charge of the electron. In (4) x denotes the position operator of the electron on $L^2(\mathbb{R}^3)$. H_0 is invariant under translations jointly of the electron and the photons. Thus the total momentum

$$p = p_{\mathrm{el}}\otimes\mathbf{1} + \mathbf{1}\otimes p_{\mathrm{f}}\,, \qquad p_{\mathrm{f}} = \sum_{\lambda=1,2}\int d^3k \quad k\,a^*(k,\lambda)a(k,\lambda)\,, \tag{5}$$

is conserved, $[H_0,p]=0$. This can be seen more directly by rewriting (4) in momentum representation as

$$H_0 = \frac{1}{2m}\big(p_{\mathrm{el}}\otimes\mathbf{1} - eA_\rho(i\nabla_{p_{\mathrm{el}}})\big)^2 + \mathbf{1}\otimes H_{\mathrm{f}} \tag{6}$$

and then going to a representation in which p is diagonal. This is achieved through the unitary transformation T defined by

$$(T\psi)^{(n)}(p,k_1,\dots,k_n) = \psi^{(n)}(p-\sum_{i=1}^n k_i,k_1,\dots,k_n)\,, \tag{7}$$

where $\psi^{(n)}$ is the n-particle sector component of ψ in electron momentum representation. The transformed Hamiltonian $T^{-1}H_0T$, again denoted by H_0, takes the form

$$H_0 = \frac{1}{2m}\big(p - p_{\mathrm{f}} - eA_\rho(0)\big)^2 + H_{\mathrm{f}}\,. \tag{8}$$

We decompose $\mathcal{H}$ and H_0 on the spectrum of p,

$$\mathcal{H} = \int_{\mathbb{R}^3}^{\oplus} d^3p\,\mathcal{H}_p\,, \qquad H_0 = \int_{\mathbb{R}^3}^{\oplus} d^3p\,H_0(p)\,. \tag{9}$$

The spaces $\mathcal{H}_p$ are isomorphic to $\mathcal{H}_{\mathrm{f}}$ in a natural sense and will be identified in the following. $H_0(p)$ is just H_0 acting on $\mathcal{H}_{\mathrm{f}}$ for a given value of p.

Physically one expects to have a dressed electron state for given momentum p, at least if $|p|\le p_{\mathrm{c}}$. In our semi-relativistic model states with $|p|\ge p_{\mathrm{c}}$ decay through Cherenkov radiation to lower momentum states. Thus provisionally we assume that there exists a p_{c} such that for every $p\in\Lambda_{\mathrm{g}} = \{p : |p|\le p_{\mathrm{c}}\}$ $H_0(p)$ has a unique ground state with energy $E(p)$, i.e.

$$H_0(p)\psi_0(p) = E(p)\psi_0(p) \tag{10}$$

has a solution $\psi_0(p) \in \mathcal{H}_{\rm f}$ which is unique up to scalar multiples and $E(p) = \inf \mathrm{spec}(H_0(p))$. We will discuss in Section 3 under which additional conditions on $\widehat{\rho}$ our assumptions can be verified.

$E(p)$ is the ground state band energy. The corresponding projection is $P_{\rm g} = \int_{\Lambda_{\rm g}}^{\oplus} d^3p\, P_0(p)$ where $P_0(p)$ denotes the orthogonal projection onto the one-dimensional subspace spanned by $\psi_0(p)$. The states in $\mathrm{Ran} P_{\rm g}$ are called dressed electron states. More explicitly

$$\mathrm{Ran}\, P_{\rm g} = \left\{ \int_{\Lambda_{\rm g}}^{\oplus} d^3p\, \phi(p)\psi_0(p) \in \mathcal{H};\ \phi \in L^2(\Lambda_{\rm g}) \right\} . \tag{11}$$

On $\mathrm{Ran} P_{\rm g}$ we have

$$e^{-iH_0 t}\psi = \int_{\Lambda_{\rm g}} d^3p\, \left(e^{-iE(p)t}\phi(p) \right) \psi_0(p)\,. \tag{12}$$

Thus, if initially in $\mathrm{Ran} P_{\rm g}$, the dressed electron propagates like a single quantum particle with dispersion relation $E(p)$, which is generated through the interaction with the photons. In particular, it stays in the dressed electron subspace for all times.

As already remarked, the motion of an electron is modified through external electro-magnetic potentials. In general, they have a slow variation on the scale set by the Compton wave length. Thus we add to H_0 the external potential $V(\varepsilon x)$ (and possibly also an external vector potential $A_{\rm ex}(\varepsilon x)$). ε is a small dimensionless parameter which controls the variation of V. The external forces break the translation invariance of H_0 and the total momentum is no longer conserved. A state ψ initially in $\mathrm{Ran} P_{\rm g}$ will no longer remain so under the time evolution generated by the full Hamiltonian H, which reflects that an accelerated charge looses energy through radiation. Since the external forces are weak of order ε, we can expect radiation losses to be negligible. More precisely, the acceleration is of order ε and by Larmor's formula the energy radiated over the time span τ is $\varepsilon^2\tau$. Thus the relevant time scale is of order ε^{-1}. On that time scale the radiation loss is of order ε, whereas the cumulative effect of the forces is of order 1 and the electron moves on the scale set by the potential V. If the initial $\psi \in \mathrm{Ran} P_{\rm g}$, the dressed electron should still be governed by an effective one-particle Hamiltonian, which is obtained from the dispersion relation $E(p)$ through the Peierls substitution

$$H_1 = E\left(p - eA_{\rm ex}(\varepsilon x)\right) + V(\varepsilon x)\,, \tag{13}$$

and

$$e^{-iHt}\psi = \int_{\Lambda_{\rm g}} d^3p\, \left(e^{-iH_1 t}\phi(p) \right) \psi_0(p) + O(\varepsilon) \tag{14}$$

for $0 \le t \le \varepsilon^{-1}T$ with some suitable macroscopic time T. We will establish (14) under some additional assumptions on H_0 and for $A_{\rm ex} = 0$.

In (14) it is crucial to assume that $\psi \in \mathrm{Ran} P_{\rm g}$. For a general $\psi \in \mathcal{H}$ one expects that, in essence on a time scale of order 1, ψ splits into outgoing radiation and a piece which is approximately in $\mathrm{Ran}P_{\mathrm{g}}$. This second piece is then governed by the approximation (14). To prove that this really happens is a problem of scattering theory for free electrons and outside the scope of our present investigation. To our knowledge this is an unsolved and very challenging problem.

There are several difficulties with the picture proposed in (14). Firstly the Pauli-Fierz Hamiltonian is infrared divergent. For $p \neq 0$, the photon cloud has an infinite number of photons, though of finite total energy. The physical ground state at fixed $p \neq 0$ does not lie in Fock space. Even if we introduce a suitable infrared cut-off by assuming that $\widehat{\rho}(k) \to 0$ as $k \to 0$, with $\widehat{\rho}$ the Fourier transform of ρ, $E(p)$ is not separated from the rest of the spectrum of $H_0(p)$. At present, we have no technique available to control (14) in case $H_0(p)$ has no spectral gap. To overcome both difficulties we are forced to give the photons a small mass, which means to set $\omega(k) = (m_{\mathrm{ph}}^2 + k^2)^{1/2}$. We emphasize that the radiation losses are not affected by this cut-off.

As a second difficulty, which has been accounted for already, we observe that $E(p)$ will cease to exist beyond a certain critical value p_{c}, i.e. for $|p| \geq p_{\mathrm{c}}$. This is most easily understood by considering the uncoupled Hamiltonian $H_0(p)$ at $e = 0$. It has absolutely continuous spectrum and the only eigenvalue $E(p) = \frac{1}{2m_{\mathrm{el}}}p^2$. This eigenvalue is isolated and below the continuum edge for $|p| < m_{\mathrm{el}}$ (m_{ph} small) and is embedded in the continuum for $|p| > m_{\mathrm{el}}$. For small coupling the embedded eigenvalue should dissolve and $E(p)$ exists only for $|p| < p_{\mathrm{c}} \cong m_{\mathrm{el}}$.

We could avoid the bounded extension of $E(p)$ through a suitable modification of the boson dispersion $\omega(k)$ and/or the electron dispersion $p^2/2m$, cf. Section 2.3. However, the termination of bands is a fairly generic phenomenon. Thus, viewed in a more general setting, we have a classical Hamiltonian

$$H_{\mathrm{cl}}(q,p) = E(p) + V(q) \tag{15}$$

corresponding to (13) on phase space $\Gamma = \mathbb{R}^3 \times \Lambda_{\mathrm{g}}$. The solution flow does not exist globally in time and for given initial conditions (q,p) there is a first time T (including $T = \infty$) when the solution trajectory hits the boundary of Γ. This means we have to control the approximation (14) up to the time when substantial parts of the wave packet "leave" the allowed phase space. Beyond that time new physical phenomena appear not accounted for by (14). We believe that one of our main achievements is to develop a technical machinery which allows for such local approximations.

It turns out that the validity of (14) relies only on a few rather general facts. Therefore we decided to prove (14) as an abstract operator problem. The basic assumptions are the decomposition (9) and a non-degenerate energy band of possibly finite extension separated by a gap from the rest of the spectrum. We mention three widely studied physical systems which posses this abstract mathematical structure, but are, at first sight, very different from the dressed electron.
(i) For an electron moving in a periodic crystal potential, p becomes the quasi-momentum and the energy band is one particular Bloch band. Bloch bands are

discrete but may cross. The small parameter arises, as for Pauli-Fierz, through a, on the scale of the lattice spacing, slowly varying external potential. In [7] we studied the semi-classical motion in periodic potentials for Bloch bands that are isolated over the whole Brillouin zone. While we used a similar approach, the present paper is a substantial improvement in two respects. The approximation is local, no isolated bands have to be assumed, and we consistently avoid using any information on $H(p)$ away from the band of interest.
(ii) Electrons are lighter than nuclei by a factor of $2 \cdot 10^3$–$5 \cdot 10^5$, which is the starting point of the Born-Oppenheimer approximation for molecular dynamics. p is now replaced by $\vec{R}$, the collection of nucleonic coordinates. $H(\vec{R})$ is the electron Hamiltonian for fixed nuclei and the band structure arises from eigenvalues of $H(\vec{R})$. They may cross or dive into the continuum. If we include the kinetic energy of the nuclei, the small parameter becomes $\varepsilon = (m_{\mathrm{el}}/m_{\mathrm{nucleus}})^{1/2}$ and, except for the interchange of p and $\vec{R}$, the Born-Oppenheimer approximation fits our framework. Since in the following only bounded potentials will be covered—the kinetic energy is p^2 and plays the role of the potential—we postpone a discussion of the time-dependent Born-Oppenheimer approximation to a separate paper. For an analysis from the point of view of wave packet dynamics we refer to [5].
(iii) The Dirac equation for a single particle has the electron and the positron band. One studies the motion of an electron, say, under slowly varying external potentials [1], [13]. As novel feature, the bands are doubly degenerate. Presumably this would also be the case if we include the electron spin in (8) as

$$H = \frac{1}{2m}\big(\sigma \cdot (p - p_{\mathrm{f}} - eA_\rho)\big)^2 + H_{\mathrm{f}}\,.$$

In the semi-classical limit the internal degrees of freedom (degeneracy) remain quantum mechanical and one has to approximate by matrix valued classical mechanics.

Our paper is organized as follows. The abstract setting will be explained in Section 2. In this framework, almost by necessity, the theorems are stated as uniform convergence of certain unitary groups and of the corresponding time dependent semi-classical observables in the Heisenberg picture. The reader may worry, as we did, whether such convergence results imply the semi-classical approximation of quantities of physical interest, like position and momentum distributions. In fact, they do under very mild assumptions on the initial wave function. We could not find a coverage of sufficient generality in the literature and therefore discuss semi-classical distributions in Section 6. Our main theorems are stated in Section 2 with proofs given in Sections 3 to 5.

2 General setting and main results

2.1 General setting

As the "momentum space" M we take either $\mathbb{R}^d$, $d \in \mathbb{N}$, in which case dp will denote Lebesgue measure on $\mathbb{R}^d$, or a flat d-torus. In the latter case we take

dp to denote normalized Lebesgue measure on the torus. In the following $\mathcal{H}_{\mathrm{f}}$ denotes any separable Hilbert space, although the notation should remind one of the Hilbert space for the Bose field in case of our main application.

Let H_0 be a self-adjoint operator on $D(H_0) \subset \mathcal{H} = L^2(M) \otimes \mathcal{H}_{\mathrm{f}}$ that can be decomposed on M as

$$H_0 = \int_M^{\oplus} dp\, H_0(p)\,, \tag{16}$$

where $\{H_0(p), p \in M\}$ is a family of self-adjoint operators with a common domain $\mathcal{D}_0 \subset \mathcal{H}_{\mathrm{f}}$. We assume that the map $p \mapsto H_0(p)$ is differentiable in the sense that for all $p \in M$ and $j = 1, \ldots, d$ the limit

$$\big(\partial_{p_j} H_0(p)\big)(H_0(p) - i)^{-1} := \lim_{h \to 0} \frac{H_0(p + he_j) - H_0(p)}{h}(H_0(p) - i)^{-1}$$

exists in the norm of bounded operators on $\mathcal{H}_{\mathrm{f}}$. This defines, in particular, $(\nabla_p H_0)(p)$ as a self-adjoint operator on $\mathcal{D}_0^d$.

For p in some compact and convex $\Lambda_{\mathrm{g}} \subseteq M$ with non-empty interior let $H_0(p)$ have an isolated non-degenerate eigenvalue $E(p)$, i.e.

$$H_0(p)\psi_0(p) = E(p)\psi_0(p)$$

with $\psi_0(p) \in \mathcal{H}_{\mathrm{f}}$ and $\mathrm{dist}\big(E(p), \sigma(H_0(p)) \setminus E(p)\big) > 0$. We assume that $E(\cdot) \in C^\infty(\Lambda_{\mathrm{g}}, \mathrm{I\!R})$. As before, $P_0(p)$ denotes the rank-one projection onto $\psi_0(p)$ and P_{g} is defined as the orthogonal projection on $\mathrm{Ran}P_{\mathrm{g}}$ given by (11). On Λ_{g} we require $H_0(\cdot)$ to be twice continuously differentiable in the same sense as above, i.e. for all $p \in \Lambda_{\mathrm{g}}$ and $j, k \in 1, \ldots, d$ the limit

$$\big(\partial_{p_k}\partial_{p_j} H_0(p)\big)(H_0(p) - i)^{-1} := \lim_{h \to 0} \frac{(\partial_{p_j} H_0)(p + he_k) - (\partial_{p_j} H_0)(p)}{h}(H_0(p) - i)^{-1}$$

exists in the norm of bounded operators on $\mathcal{H}_{\mathrm{f}}$ and depends continuously on p in the same topology.

Let the potential $V : \mathrm{I\!R}^d \to \mathrm{I\!R}$ be such that $V(x) = \int dk\, e^{ik\cdot x}\, \widehat{V}(k)$, where $\int dk\, |k|^n |\widehat{V}(k)| < \infty$ for all $n \in \mathrm{I\!N}_0$. This guarantees, in particular, that V and all its partial derivatives are in $C^\infty(\mathrm{I\!R}^d)$ and vanish at infinity. $V^\varepsilon := V(i\varepsilon\nabla_p)$ is a bounded self-adjoint operator on $L^2(M)$ and the full Hamiltonian

$$H = H_0 + V^\varepsilon \otimes \mathbf{1}_{\mathcal{H}_{\mathrm{f}}} \tag{17}$$

is self-adjoint on $D(H_0)$.

We define the corresponding one-particle Hamiltonian on $L^2(M)$ by

$$H_1 = E(p) + V^\varepsilon\,. \tag{18}$$

Note that $E(p)$ is a priori only defined for $p \in \Lambda_{\mathrm{g}}$. To make things simple, we continue $E(p)$ to a smooth and compactly supported but otherwise arbitrary function on M. Since we will be interested only in the behavior for $p \in \Lambda_{\mathrm{g}}$, we use this global function to define H_1. The corresponding unitary groups are denoted by $U(t) = e^{-iHt}$ and $U_1(t) = e^{-iH_1 t}$.

Our goal is to show that, in a suitable sense,

$$\big(U\,(t/\varepsilon) - U_1\,(t/\varepsilon)\,\big)P_{\rm g} = O(\varepsilon) \tag{19}$$

for macroscopic times $t < T < \infty$, where T depends only on V and the initial momenta.

But before we can state the precise result, we have to make sense of H_1 acting on $\mathrm{Ran}P_{\rm g}$. According to (11) we define the map $\mathcal{U} : \mathrm{Ran}P_{\rm g} \to L^2(M)$

$$\mathcal{U}(\phi\psi_0) = \phi\,. \tag{20}$$

Its adjoint $\mathcal{U}^* : L^2(M) \to \mathrm{Ran}P_{\rm g}$ is given by

$$\mathcal{U}^*\phi = \int^{\oplus} dp\, 1\!\!1_{\Lambda_{\rm g}}(p)\phi(p)\psi_0(p)\,, \tag{21}$$

where here and in the following $1\!\!1_A$ denotes the characteristic function of a set A. Clearly $\mathcal{U}$ is an isometry and $\mathcal{U}^*\mathcal{U} = \mathbf{1}$ on $\mathrm{Ran}P_{\rm g}$.

The effective dynamics cease to make sense, once the momentum of the particle leaves $\Lambda_{\rm g}$. If that is not excluded by energy conservation, the comparison (19) can hold only up to some finite time, which can be determined from the classical dynamics generated by

$$H_{\rm cl}(x,p) = E(p) + V(x)$$

on phase space $\mathbb{R}^d \times M$. Let $\Lambda_{\rm i} \subset M$, the "set of initial momenta", and $\Lambda_{\rm m} \subset M$ some "maximal set of allowed momenta", both be compact. For $\Lambda \subset M$ compact and $\delta \geq 0$ we let

$$\Lambda + \delta := \left\{ p \in M : \inf_{k\in\Lambda} |p-k| \leq \delta \right\}$$

be the corresponding δ-enlarged set, which is again compact. With $\Phi_p^t : \mathbb{R}^d \times M \to M$ denoting the momentum component of the classical flow, we define for $\Lambda_{\rm i} + \delta \subset \Lambda_{\rm m}$

$$T^\delta_{\rm m}(\Lambda_{\rm i},\Lambda_{\rm m}) := \sup_{t\geq 0} \left\{ t \,:\, \mathrm{supp}\left(\big(1\!\!1_{\Lambda_{\rm i}} \circ \Phi_p^{-s}\big)(x,\cdot)\right) + \delta \subseteq \Lambda_{\rm m}\ \forall\, s \in [0,t], x \in \mathbb{R}^d \right\}$$

as the maximal time for which the momentum of any (i.e. with any starting position) classical particle with initial momentum in $\Lambda_{\rm i}$ stays within a δ-margin inside of $\Lambda_{\rm m}$.

The following lemma shows that the classical bound on the momentum is respected also by the quantum dynamics in the limit $\varepsilon \to 0$. Let $P_{\rm i} = 1\!\!1_{\Lambda_{\rm i}}(p)$ and $P_{\rm m} = 1\!\!1_{\Lambda_{\rm m}}(p)$.

Lemma 1. Let $\Lambda_{\rm i}$, $\Lambda_{\rm m} \subset M$ be both compact and $\Lambda_{\rm i} + \delta \subset \Lambda_{\rm m}$ for some $\delta > 0$. For any $T < \infty$ with $T \leq T^\delta_{\rm m}(\Lambda_{\rm i},\Lambda_{\rm m})$ there is a $C < \infty$ such that for all $t \in [0,T]$

$$\left\| (1-P_{\rm m})\; U_1\left(\frac{t}{\varepsilon}\right)\, P_{\rm i} \right\|_{\mathcal{L}(L^2(M))} \leq C\,\varepsilon^2\,. \tag{22}$$

For the proof of Lemma 1 see Section 3.

2.2 Main results

In the following we will always assume that $\Lambda_{\rm i} \subset \Lambda_{\rm g}$ and we can therefore identify $P_{\rm i}$ with $\mathcal{U}^* P_{\rm i} \mathcal{U} P_{\rm g}$ to keep notation simple. I.e., $P_{\rm i}$ projects on "dressed electron states" with momenta in $\Lambda_{\rm i}$.

Theorem 2. Let $\Lambda_{\rm i}$ be compact with $\Lambda_{\rm i} + \delta \subset \Lambda_{\rm g}$ for some $\delta > 0$. For any $T < \infty$ with $T \leq T_{\rm m}^{\delta}(\Lambda_{\rm i}, \Lambda_{\rm g})$ there is a $C < \infty$ such that for all $t \in [0, T]$

$$\left\| \left(U\left(\frac{t}{\varepsilon}\right) - \mathcal{U}^* U_1 \left(\frac{t}{\varepsilon}\right) \mathcal{U} \right) P_{\rm i} \right\|_{\mathcal{L}(\mathcal{H})} \leq C\,\varepsilon\,. \tag{23}$$

This means, in the case of our main application, that dressed electron states with initial momenta in $\Lambda_{\rm i}$ evolve according to the effective dynamics without radiation losses for times of order ε^{-1} as long as the momenta of the corresponding classical orbits stay inside $\Lambda_{\rm g}$. Note, in particular, that the macroscopic time $T_{\rm m}^{\delta}$ depends only on $\Lambda_{\rm i}$, but not on ε.

The equivalence between the full dynamics and the effective one-particle dynamics on the isolated energy band extends to the level of semi-classical or macroscopic observables. On a suitable domain in $\mathcal{H}$ the macroscopic (Heisenberg) position operator is given by

$$x^{\varepsilon}(t) = U\left(-t/\varepsilon\right)\,\left(\varepsilon i \nabla_p \otimes \mathbf{1}\right)\, U\left(t/\varepsilon\right)$$

and its effective one-particle counterpart on $H^1(\mathbb{R}^d)$ by

$$x_1^{\varepsilon}(t) = U_1\left(-t/\varepsilon\right)\,\varepsilon i \nabla_p\, U_1\left(t/\varepsilon\right)\,.$$

Theorem 3. Let $\Lambda_{\rm i}$ be compact with $\Lambda_{\rm i} + \delta \subset \Lambda_{\rm g}$ for some $\delta > 0$. For every $T < \infty$ with $T \leq T_{\rm m}^{\delta}(\Lambda_{\rm i}, \Lambda_{\rm g})$ there is a $C < \infty$ such that for all $t \in [0, T]$

$$\left\| \left(x^{\varepsilon}(t) - \mathcal{U}^* x_1^{\varepsilon}(t) \mathcal{U} \right) P_{\rm i} \right\|_{\mathcal{L}(\mathcal{H})} < C\,\varepsilon\,. \tag{24}$$

Theorem 3 is not a trivial corollary of Theorem 2 for several reasons. It is a statement about unbounded operators, whose difference, however, is bounded. More importantly, the difference of the position operators does not vanish even at time $t = 0$ because $[x^{\varepsilon}(0), P_{\rm g}] \neq 0$.

More generally, consider classical symbols $a \in C^{\infty}(\mathbb{R}^d \times M)$ such that for all multi-indices α, β there exists $C_{\alpha,\beta} < \infty$ with

$$\sup_{x,p} \left| \partial_x^{\alpha} \partial_p^{\beta} a(x,p) \right| \leq C_{\alpha,\beta}\,.$$

We use the notation of [2] and denote this set of symbols by $S^0_0(1)$. For $a \in S^0_0(1)$ Weyl quantization leads to an operator $a^{\mathrm{W},\varepsilon} \in \mathcal{L}(L^2(M))$ given by

$$\left(a^{\mathrm{W},\varepsilon}\phi\right)(p) = (2\pi)^{-d}\int dx\, dk\; a\left(\varepsilon x, \frac{p+k}{2}\right)\, e^{-i(p-k)\cdot x}\,\phi(k)\,, \tag{25}$$

with operator norm that is bounded uniformly in ε (cf. Section 4 for details).

For $a \in S^0_0(1)$ let

$$a^\varepsilon(t) = U\left(-t/\varepsilon\right)\,\left(a^{\mathrm{W},\varepsilon}\otimes \mathbf{1}\right)\,U\left(t/\varepsilon\right) \tag{26}$$

and

$$a^\varepsilon_1(t) = U_1\left(-t/\varepsilon\right)\,a^{\mathrm{W},\varepsilon}\,U_1\left(t/\varepsilon\right)\,. \tag{27}$$

Theorem 4. Let Λ_{i} be compact with $\Lambda_{\mathrm{i}}+\delta\subset\Lambda_{\mathrm{g}}$ for some $\delta>0$ and $a\in S^0_0(1)\cap C_\infty(\mathbb{R}^d\times M)$. For every $T<\infty$ with $T\le T^\delta_{\mathrm{m}}(\Lambda_{\mathrm{i}},\Lambda_{\mathrm{g}})$ there is a $C<\infty$ such that for all $t\in[0,T]$

$$\left\|\left(a^\varepsilon(t)-\mathcal{U}^* a^\varepsilon_1(t)\,\mathcal{U}\right)P_{\mathrm{i}}\right\|_{\mathcal{L}(\mathcal{H})} < C\,\varepsilon\,. \tag{28}$$

C_∞ denotes the set of continuous functions vanishing at infinity. Presumably Theorem 4 holds as well for $a\in S^0_0(1)$, however, the assumption $a\in C_\infty$ allows for a simpler approximation argument in the proof.

2.3 The massive Nelson and Pauli-Fierz model

It is of interest to see whether our abstract assumptions are in fact satisfied for physically, at least semi-realistic models. The best studied case is the Nelson model [9], where the coupling is to the position of the particle and the Maxwell field is replaced by a scalar field. Switching immediately to the total momentum representation, cf. (8), the Nelson Hamiltonian reads

$$H_{\mathrm{N}}(p) = \frac{1}{2}(p-p_{\mathrm{f}})^2 + H_{\mathrm{f}} + (2\pi)^{-d/2}\int dk\,\frac{\widehat{\rho}(k)}{\sqrt{2\omega(k)}}\left(a(k)+a^*(-k)\right)\,, \tag{29}$$

where instead of (1), (5), $a(k)$, $a^*(k)$ is a one-component Bose field over $\mathbb{R}^d$. Again, p is the total momentum and regarded as a parameter. If $\int dk\,|\widehat{\rho}|^2(\omega^{-3}+\omega^{-1})<\infty$, then $H_{\mathrm{N}}(p)$ is bounded from below and self-adjoint with domain $D(H_{\mathrm{f}})$.

According to a result of Fröhlich [4], under this condition with $\omega(k)=(m^2_{\mathrm{ph}}+k^2)^{1/2}$, $m_{\mathrm{ph}}>0$ and in dimension $d=3$, H_{N} has an isolated, nondegenerate ground state band for $|p|<p_{\mathrm{c}}$ with $p_{\mathrm{c}}>\sqrt{3}-1$. If in (29) we replace the electron dispersion by $E_0(p)=(m^2_{\mathrm{el}}+p^2)^{1/2}$, $m_{\mathrm{el}}>0$, still keeping $\omega(k)=(m^2_{\mathrm{ph}}+k^2)^{1/2}$, $m_{\mathrm{ph}}>0$, then at zero coupling the ground state lies strictly below the continuum edge for all p, which means $p_\mathrm{c} = \infty$ for $\widehat{\rho} = 0$. As proved in [4] $p_\mathrm{c} = \infty$ persists to arbitrary coupling strength.

A larger class of bosonic dispersion relations is studied in [14]. For the particular case $\omega(k) = \omega_0 > 0$ and dimensions $d = 1, 2$ one has $p_\mathrm{c} = \infty$, whereas for $d = 3$ and small coupling indeed $p_\mathrm{c} < \infty$ [8].

Fröhlich encloses the Bose field in a periodic box and proves that as the box size goes to infinity, there is still spectrum of $H_\mathrm{N}(p)$ strictly below the continuum edge, provided $|p| < p_\mathrm{c}$ which has to be estimated through a separate argument. Thus $H_\mathrm{N}(p)$ has a ground state for $|p| < p_\mathrm{c}$. To establish uniqueness he uses that $e^{-tH_\mathrm{N}(p)}$ in Fock space representation is positivity improving, except for a small error which goes to zero as $t \to \infty$. From analytic perturbation theory one concludes then that the ground state band is isolated and analytic in p, provided $|p| < p_\mathrm{c}$. Hence, the massive Nelson model satisfies the assumptions underlying Theorem 2, 3 and 4.

For the Pauli-Fierz model (8) the existence part of the proof goes through without essential changes. Unfortunately, the positivity part fails and no argument is known to ensure the uniqueness of the ground state for $|p| < p_\mathrm{c}$. Since for $e = 0$ and small p the ground state band is isolated, it must remain so for small e by perturbation theory, however, with the undesirable feature that smallness depends now on m_ph. Even for the massive Pauli-Fierz model a general proof of uniqueness would be most welcome.

3 Preliminaries from semi-classical analysis

Although Theorem 2 compares a full quantum evolution to an effective quantum evolution and says nothing about semi-classics, its proof makes use of Lemma 1. To prove Lemma 1 we need some tools from semi-classical analysis, which are introduced and applied to the one-particle Hamiltonian H_1 in this section. Furthermore, in Section 6 we will consider semi-classical distributions based on the notions and results discussed here.

Since E and V are both smooth bounded functions, we can apply standard results of semi-classics to the Hamiltonian

$$H_1(p, i\varepsilon\nabla_p) = E(p) + V(i\varepsilon\nabla_p)$$

acting on $L^2(M)$, where the roles of momentum and position are exchanged and the role of $\hbar$ is taken by ε. In the following we will simply ignore this difference and leave the necessary changes as compared to the standard case to the reader. Note, however, that the change in sign, $i\nabla$ instead of $-i\nabla$, is canceled by the fact that $p' = q$ and $q' = -p$ is the canonical transformation interchanging q and p.

We will consider classical symbols $a \in S^0_0(1)$. As was stated in Section 2, Weyl quantization of functions in $S^0_0(1)$ leads to bounded operators. The following result sharpens this statement.

Proposition 5. [Calderon-Vaillancourt] There is a $C < \infty$ and a finite $n \in \mathbb{N}$

such that for all $a \in S_0^0(1)$ and $\varepsilon \in [0,1]$

$$\left\|a^{\mathrm{W},\varepsilon}\right\|_{\mathcal{L}(L^2(M))} \leq C \sup_{|\alpha|\leq n, |\beta|\leq n, (x,p)\in \mathbb{R}^d\times M} \left|\partial_x^\alpha \partial_p^\beta\, a(x,p)\right| . \tag{30}$$

For a proof see [2], Theorem 7.11. The statement there is slightly different, but their proof implies our Proposition 5. Note, in particular, that the constants n and C in Proposition 5 depend on the dimension d of configuration space.

The so called product rule, presented in the following Proposition 6, is at the basis of all semi-classical analysis we will apply. Given Proposition 5, however, its proof is mainly a matter of calculation. For details see [2], Proposition 7.7 and Theorem 7.9.

Proposition 6. [Product rule] Let $a, b \in S_0^0(1)$. Then for all $n \in \mathbb{N}_0$ there is a $d_n < \infty$ such that

$$\left\|a^{\mathrm{W},\varepsilon} b^{\mathrm{W},\varepsilon} - \Big(\sum_{k=0}^{n} \varepsilon^k\, c_k\Big)^{\mathrm{W},\varepsilon}\right\|_{\mathcal{L}(L^2(M))} \leq d_n\, \varepsilon^{n+1} , \tag{31}$$

with

$$c_k(x,p) = \left(\frac{i}{2}\right)^k \sum_{|\alpha|+|\beta|=k} \frac{(-1)^{|\beta|}}{|\alpha|!|\beta|!} \Big((\partial_x^\beta \partial_p^\alpha a)(\partial_p^\beta \partial_x^\alpha b)\Big)(x,p) . \tag{32}$$

We state two simple facts that are immediate consequences of the product rule as

Corollary 7. Let $a, b \in S_0^0(1)$. Then

(i)

$$\left\|a^{\mathrm{W},\varepsilon} b^{\mathrm{W},\varepsilon} - (ab)^{\mathrm{W},\varepsilon}\right\|_{\mathcal{L}(L^2(M))} = O(\varepsilon) .$$

(ii) If, in addition, $\mathrm{supp}(a) \cap \mathrm{supp}(b) = \emptyset$, then for any $n \in \mathbb{N}$

$$\left\|a^{\mathrm{W},\varepsilon} b^{\mathrm{W},\varepsilon}\right\|_{\mathcal{L}(L^2(M))} = O(\varepsilon^n) .$$

Proof. (i) is just Proposition 6 for $n = 0$. (ii) holds since in this case $c_k = 0$ in (31) for all $k \in \mathbb{N}_0$.

□

Here and in the following $O(\varepsilon^n)$ means that an expression, or its appropriate norm, is bounded by a constant times ε^n for sufficiently small ε.

The crucial ingredient to our semi-classical analysis is the following first-order version of a Theorem going back to Egorov [3] (cf. also [12], Théorème IV-10), which is also a direct consequence of the product rule.

Proposition 8. [Egorov's Theorem] Let $a \in S_0^0(1)$ and $0 \le T < \infty$. There is a $C < \infty$ such that for all $t \in [-T, T]$

$$\left\| U_1\left(-\frac{t}{\varepsilon}\right) a^{\mathrm{W},\varepsilon} U_1\left(\frac{t}{\varepsilon}\right) - \left(a \circ \Phi^t\right)^{\mathrm{W},\varepsilon} \right\|_{\mathcal{L}(L^2(M))} \le C\varepsilon^2 .$$

Proof. First note that for $a \in S_0^0(1)$ we have that $a \circ \Phi^t \in S_0^0(1)$ for all finite t, since the Hamiltonian vector field is smooth, uniformly bounded and all its partial derivatives are uniformly bounded. Moreover, $a \circ \Phi^t$ and all its partial derivatives are each bounded uniformly for $t \in [-T, T]$. Therefore $\|(a \circ \Phi^t)^{\mathrm{W},\varepsilon}\|_{\mathcal{L}(L^2(M))}$ is bounded uniformly for $t \in [-T, T]$ by Proposition 5.

Writing

$$\begin{aligned} & U_1\left(-\frac{t}{\varepsilon}\right) a^{\mathrm{W},\varepsilon} U_1\left(\frac{t}{\varepsilon}\right) - \left(a \circ \Phi^t\right)^{\mathrm{W},\varepsilon} = \\ & = \int_0^t ds \, \frac{d}{ds} \left(U_1\left(-\frac{s}{\varepsilon}\right) (a \circ \Phi^{t-s})^{\mathrm{W},\varepsilon} U_1\left(\frac{s}{\varepsilon}\right)\right) , \end{aligned} \tag{33}$$

one is led to consider

$$\begin{aligned} & \frac{d}{ds} U_1\left(-\frac{s}{\varepsilon}\right) (a \circ \Phi^{t-s})^{\mathrm{W},\varepsilon} U_1\left(\frac{s}{\varepsilon}\right) = \\ & = U_1\left(-\frac{s}{\varepsilon}\right) \left(\frac{i}{\varepsilon} \left[H_{\mathrm{cl}}^{\mathrm{W},\varepsilon}, (a \circ \Phi^{t-s})^{\mathrm{W},\varepsilon} \right] - \left\{ H_{\mathrm{cl}}, (a \circ \Phi^{t-s}) \right\}^{\mathrm{W},\varepsilon} \right) U_1\left(\frac{s}{\varepsilon}\right) , \end{aligned} \tag{34}$$

where $\{\cdot,\cdot\}$ denotes the Poisson bracket and we used that $H_1 = H_{\mathrm{cl}}^{\mathrm{W},\varepsilon}$. From the product rule one computes that for arbitrary $a, b \in S_0^0(1)$

$$\frac{i}{\varepsilon}\left[a^{\mathrm{W},\varepsilon}, b^{\mathrm{W},\varepsilon}\right] = \frac{i}{\varepsilon}\left(a^{\mathrm{W},\varepsilon} b^{\mathrm{W},\varepsilon} - b^{\mathrm{W},\varepsilon} a^{\mathrm{W},\varepsilon}\right) = \left\{a, b\right\}^{\mathrm{W},\varepsilon} + O(\varepsilon^2) , \tag{35}$$

which implies that (34) is $O(\varepsilon^2)$ for fixed $t-s$. However, one can easily convince oneself that this bound is uniform for $t - s$ in any bounded interval, since the semi-norm used in Proposition 5 is bounded uniformly for the c_2-terms appearing in a derivation of Equation (35) using the product rule.

□

To have a natural way for extending functions of x or p alone to functions on phase space, we introduce the projections $\pi_x : \mathbb{R}^d \times M \to \mathbb{R}^d$ and $\pi_p : \mathbb{R}^d \times M \to M$ as $\pi_x(x, p) = x$ and $\pi_p(x, p) = p$. We are now ready to prove Lemma 1.

Proof (of Lemma 1). In order to regularize $\mathbb{1}_{\Lambda_{\mathrm{i}}}$ and $\mathbb{1}_{\Lambda_{\mathrm{m}}}$ we pick some $f_{\mathrm{i}} \in C_0^\infty(M)$ and $f_{\mathrm{m}} \in C^\infty(M)$ such that $f_{\mathrm{i}}\big|_{\Lambda_{\mathrm{i}}} = 1$, $f_{\mathrm{m}}\big|_{M\setminus\Lambda_{\mathrm{m}}} = 1$ and $\mathrm{supp}(f_{\mathrm{i}} \circ \Phi_p^{-t}) \cap \mathrm{supp}(f_{\mathrm{m}} \circ \pi_p) = \emptyset$ for all $t \in [0, T_m^\delta]$. Egorov's Theorem implies that there is a $C < \infty$ such that for $t \in [0, T_m^\delta]$

$$\left\| U_1\left(\frac{t}{\varepsilon}\right) (f_{\mathrm{i}} \circ \pi_p)^{\mathrm{W},\varepsilon} U_1\left(-\frac{t}{\varepsilon}\right) - (f_{\mathrm{i}} \circ \Phi_p^{-t})^{\mathrm{W},\varepsilon} \right\|_{\mathcal{L}(L^2)} \le C\varepsilon^2 .$$

Since, by construction, $\mathrm{supp}(f_{\mathrm{i}} \circ \Phi_p^{-t}) \cap \mathrm{supp}(f_{\mathrm{m}} \circ \pi_p) = \emptyset$, we have $(f_{\mathrm{m}} \circ \pi_p)^{\mathrm{W},\varepsilon}(f_{\mathrm{i}} \circ \Phi_p^{-t})^{\mathrm{W},\varepsilon} = O(\varepsilon^2)$ by Corollary 7. This and the fact that $P_{\mathrm{i}} = (f_{\mathrm{i}} \circ \pi_p)^{\mathrm{W},\varepsilon} P_{\mathrm{i}}$ and that $(1 - P_{\mathrm{m}}) = (1 - P_{\mathrm{m}})(f_{\mathrm{m}} \circ \pi_p)^{\mathrm{W},\varepsilon}$ implies (22).

□

4 Convergence of the unitary groups

We prove Theorem 2. Let $\Lambda_{\mathrm{i}} + \delta \subset \Lambda_{\mathrm{g}}$. In the following $T < \infty$ with $T \leq T_{\mathrm{m}}^{\delta}(\Lambda_{\mathrm{i}}, \Lambda_{\mathrm{g}})$ will be fixed once and for all and we will always assume that $0 \leq t \leq T$.

For reasons that will become clear during the proof we have to introduce further sets in momentum space. Let

$$\Lambda_{\mathrm{m}} = \bigcup_{t \in [0,T], x \in \mathbb{R}^d} \mathrm{supp}\left(\mathbb{1}_{\Lambda_{\mathrm{i}}} \circ \Phi_p^{-t}(x, \cdot)\right)$$

be the subset in momentum space that is reached by the classical dynamics starting from Λ_{i} in our relevant time span. By construction we have that $\Lambda_{\mathrm{m}} + \delta \subseteq \Lambda_{\mathrm{g}}$ and we define $\Lambda_{\mathrm{m}1} := \Lambda_{\mathrm{m}} + \delta/4$, $\Lambda_{\mathrm{m}2} := \Lambda_{\mathrm{m}} + \delta/2$ and $\Lambda_{\mathrm{m}3} := \Lambda_{\mathrm{m}} + 3\delta/4$. The respective projections will be denoted by $P_{\mathrm{m}1}$, $P_{\mathrm{m}2}$ and $P_{\mathrm{m}3}$. The following Lemma ensures that V^ε maps wave functions supported in $\Lambda_{\mathrm{m}1}$ (resp. in $\Lambda_{\mathrm{m}2}$, $\Lambda_{\mathrm{m}3}$) to wave functions supported in $\Lambda_{\mathrm{m}2}$ (resp. in $\Lambda_{\mathrm{m}3}$, Λ_{g}) up to errors of arbitrary order in ε. We will use this result in the following implicitly many times.

Lemma 9. Let $\widehat{V} \in L^1(\mathbb{R}^d)$ such that $\int dk\, |k|^n |\widehat{V}(k)| < \infty$ for all $n \in \mathbb{N}$. Then for each compact $\Lambda \subset \mathbb{R}^d$ and all $m \in \mathbb{N}$ and $\delta > 0$ there is a $C_{\delta,m}$ such that

$$\|(V^\varepsilon - P_{\Lambda+\delta} V^\varepsilon)\, P_\Lambda\|_{\mathcal{L}(L^2(M))} \leq C_{\delta,m} \varepsilon^m \,.$$

Proof. Let $\phi \in \mathrm{Ran} P_\Lambda$, i.e. $\mathrm{supp}\phi \subset \Lambda$, and note that

$$\begin{aligned}(V^\varepsilon \phi)(p) &= \int dk\, \widehat{V}(k) \phi(p - \varepsilon k) \\ &= \int_{|k| \leq \varepsilon^{-1/2}} dk\, \widehat{V}(k) \phi(p - \varepsilon k) + \int_{|k| > \varepsilon^{-1/2}} dk\, \widehat{V}(k) \phi(p - \varepsilon k) \,. \end{aligned} \tag{36}$$

The first term in (36) is supported in $\Lambda + \delta$ for ε sufficiently small. For the second term note that

$$\int_{|k| > \varepsilon^{-1/2}} dk\, \widehat{V}(k) \phi(p - \varepsilon k) = \int_{|k'| > \varepsilon^{1/2}} dk'\, \varepsilon^{-d} \widehat{V}(k'/\varepsilon) \phi(p - k')$$

amounts to convolution with the function $\mathbb{1}_{|k'|>\varepsilon^{1/2}}\varepsilon^{-d}\widehat{V}(k'/\varepsilon)$ and therefore, by Young's inequality, the $\mathcal{L}(L^2)$-norm of the corresponding map is bounded by

$$\int\limits_{|k'|>\varepsilon^{1/2}} dk'\,\varepsilon^{-d}|\widehat{V}(k'/\varepsilon)| = \int\limits_{|k|>\varepsilon^{-1/2}} dk\,|k|^{-2m}|k|^{2m}|\widehat{V}(k)| \leq C_m\varepsilon^m\,.$$

□

We turn to the proof of (23). Using Lemma 1 together with $\mathcal{U}\mathcal{U}^* = \mathrm{Id}$ on $\mathrm{Ran}P_{\mathrm{m2}}$ we get

$$\begin{aligned}
&\left(U\left(\frac{t}{\varepsilon}\right) - \mathcal{U}^*U_1\left(\frac{t}{\varepsilon}\right)\mathcal{U}\right)P_{\mathrm{i}} = \\
&= -iU\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,U(-s)\left(H\mathcal{U}^* - \mathcal{U}^*H_1\right)U_1(s)\mathcal{U}\,P_{\mathrm{i}} \\
&= -iU\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,U(-s)\left(H - \mathcal{U}^*H_1\mathcal{U}\right)P_{\mathrm{m2}}\mathcal{U}^*U_1(s)\mathcal{U}\,P_{\mathrm{i}} + O(\varepsilon)\,.
\end{aligned} \tag{37}$$

For the last equality note also that a factor of order $O(\varepsilon^2)$ in the otherwise uniformly bounded integrand leads to the integral being $O(\varepsilon)$. We would be done by the same argument, if we could show that $H-\mathcal{U}^*H_1\mathcal{U}$ acting on $\mathrm{Ran}P_{\mathrm{m2}}$ is $O(\varepsilon^2)$. However, the first order term does not vanish, as we will see, and we have to treat the integral more carefully.

In order to separate the leading order term we write

$$H - \mathcal{U}^*H_1\mathcal{U} = (H - H_{\mathrm{diag}}) + (H_{\mathrm{diag}} - \mathcal{U}^*H_1\mathcal{U})\ ,$$

where

$$H_{\mathrm{diag}} = H_0 + P_{\mathrm{g}}V^\varepsilon P_{\mathrm{g}} + P_{\mathrm{g}}^\perp V^\varepsilon P_{\mathrm{g}}^\perp \tag{38}$$

and $P_{\mathrm{g}}^\perp := \mathbf{1}_{\mathcal{H}} - P_{\mathrm{g}}$.

We will treat the easy part first and show in Lemma 11 that the difference $H_{\mathrm{diag}} - \mathcal{U}^*H_1\mathcal{U}$ vanishes sufficiently fast on $\mathrm{Ran}P_{\mathrm{m2}}$.

However, to do that we need to show that $P_0(p)$ is twice continuously differentiable as a function of p and, for later purposes, we give an explicit expression for $\nabla_p P_0(p)$.

Lemma 10. $P_0(\cdot) \in C^2(\Lambda_{\mathrm{g}}, \mathcal{L}(\mathcal{H}_{\mathrm{f}}))$ and for $p \in \Lambda_{\mathrm{g}}$ one has

$$\big(\nabla_p P_0\big)(p) = -\big(R_E Q_0\big)(p)\big(\nabla_p H_0\big)(p)P_0(p) - P_0(p)\big(\nabla_p H_0\big)(p)\big(R_E Q_0\big)(p)\,, \tag{39}$$

where $\big(R_E Q_0\big)(p) = \big(H_0(p) - E(p)\big)^{-1}Q_0(p)$ and $Q_0(p) = \mathbf{1}_{\mathcal{H}_{\mathrm{f}}} - P_0(p)$.

Note that $(R_E Q_0)(p)$ is bounded since $E(p)$ is an isolated eigenvalue for all $p \in \Lambda_{\mathrm{g}}$ and $Q_0(p)$ projects on the orthogonal complement of the corresponding eigenvector.

Proof. For $p \in \Lambda_{\mathrm{g}}$ we can express $P_0(p)$ as a contour integral:

$$P_0(p) = -\frac{1}{2\pi i} \oint_{c(p)} d\lambda \, R_\lambda(H_0(p)) \, ,$$

where $c(p)$ is a smooth curve in the complex plane circling the isolated eigenvalue $E(p)$ only and $R_\lambda(H_0(p)) = (H_0(p) - \lambda)^{-1}$. Multiplying

$$\begin{aligned} 0 &= \nabla_p \mathbf{1} = \nabla_p (H_0(p) - \lambda) R_\lambda(H_0(p)) \\ &= (\nabla_p H_0(p)) R_\lambda(H_0(p)) + (H_0(p) - \lambda)(\nabla_p R_\lambda(H_0(p))) \end{aligned}$$

with R_λ from the left we get

$$\nabla_p R_\lambda(H_0(p)) = -R_\lambda(H_0(p))(\nabla_p H_0(p)) R_\lambda(H_0(p)) \, ,$$

which is bounded uniformly for $\lambda \in c(p)$ by our assumptions on the family $H_0(p)$. Hence

$$\nabla_p P_0(p) = -\frac{1}{2\pi i} \oint_{c(p)} d\lambda \, \nabla_p R_\lambda(H_0(p)) \, .$$

Collecting the above information we obtain

$$\begin{aligned} Q_0(p)\nabla_p P_0(p) &= Q_0(p)\nabla_p P_0(p)(P_0(p) + Q_0(p)) \\ &= \frac{1}{2\pi i} \oint_{c(p)} d\lambda \, Q_0(p) R_\lambda(H_0(p))(\nabla_p H_0(p)) R_\lambda(H_0(p)) P_0(p) \\ &= \frac{1}{2\pi i} \oint_{c(p)} d\lambda \, R_\lambda(H_0(p)) Q_0(p)(\nabla_p H_0(p)) \frac{1}{E(p) - \lambda} P_0(p) \\ &= -R_E(H_0(p)) Q_0(p) \left(\nabla_p H_0(p)\right) P_0(p) \, , \end{aligned}$$

where the $Q_0(p)\nabla_p P_0(p) Q_0(p)$ vanishes, since in this case the integrand is an analytic function inside of $c(p)$. Therefore $Q_0(p)\nabla_p P_0(p) = Q_0(p)\nabla_p P_0(p) P_0(p)$ and, using self adjointness of $\nabla_p P_0(p)$, it follows that

$$P_0(p)\nabla_p P_0(p) = P_0(p)\nabla_p P_0(p) Q_0(p) = \left(-R_E(H_0(p)) Q_0(p) \left(\nabla_p H_0(p)\right) P_0(p)\right)^* .$$

This yields (39) and $P_0(\cdot) \in C^2(\Lambda_{\mathrm{g}}, \mathcal{L}(\mathcal{H}_{\mathrm{f}}))$ by applying the same argument again to (39).

□

Lemma 11. For ε sufficiently small there is a $C < \infty$ such that

$$\left\| \left(H_{\mathrm{diag}} - \mathcal{U}^* H_1 \mathcal{U}\right) P_{\mathrm{m2}} \right\| \leq C\varepsilon^2 \, . \tag{40}$$

Proof. Note that Lemma 9 implies that

$$\|P_{\mathrm{g}}^{\perp} H_{\mathrm{diag}} P_{\mathrm{m2}}\| \leq C\varepsilon^2 \, ,$$

and we have by construction that $P_{\mathrm{g}}^{\perp}\mathcal{U}^* H_1 \mathcal{U} P_{\mathrm{m}2} = 0$. Hence it suffices to consider the difference projected onto $\mathrm{Ran}P_{\mathrm{g}}$. On $\mathrm{Ran}P_{\mathrm{m}2}$ we have

$$\begin{aligned} P_{\mathrm{g}}(H_{\mathrm{diag}}\phi\psi_0)(p) \;&= \\ & E(p)\phi(p)\psi_0(p) + \mathbb{1}_{\Lambda_{\mathrm{g}}}(p)\int dk\,\widehat{V}(k)\phi(p-\varepsilon k)\,\langle\psi_0(p),\psi_0(p-\varepsilon k)\rangle_{\mathcal{H}_{\mathrm{f}}}\,\psi_0(p) \end{aligned}$$

and

$$(\mathcal{U}^* H_1 \mathcal{U}\phi\psi_0)(p) = E(p)\phi(p)\psi_0(p) + \mathbb{1}_{\Lambda_{\mathrm{g}}}(p)\int dk\,\widehat{V}(k)\phi(p-\varepsilon k)\,\psi_0(p)\,.$$

Hence

$$\begin{aligned} & P_{\mathrm{g}}\left(H_{\mathrm{diag}} - \mathcal{U}^* H_1 \mathcal{U}\right)(\phi\psi_0)(p) = \\ & \quad = \; \mathbb{1}_{\Lambda_{\mathrm{g}}}(p)\int dk\,\widehat{V}(k)\phi(p-\varepsilon k)\,\left(\langle\psi_0(p),\psi_0(p-\varepsilon k)\rangle_{\mathcal{H}_{\mathrm{f}}} - 1\right)\,\psi_0(p)\,. \end{aligned} \tag{41}$$

We will show that there is a constant C such that

$$|\langle\psi_0(p),\psi_0(p-\varepsilon k)\rangle_{\mathcal{H}_{\mathrm{f}}} - 1| \leq C|k|^2\varepsilon^2 \tag{42}$$

for Lebesgue almost all k and $p, p-\varepsilon k \in \Lambda_{\mathrm{g}}$. Therefore

$$\begin{aligned} \left\| \left(H_{\mathrm{diag}} - \mathcal{U}^* H_1 \mathcal{U}\right)\phi\psi_0 \right\|_{\mathcal{H}} \;&\leq\; C\varepsilon^2 \int dk\,|\widehat{V}(k)|\,|k|^2 \left\|\phi(\cdot - \varepsilon k)\psi_0(\cdot)\right\|_{\mathcal{H}} \\ &\leq\; C\varepsilon^2 \|\phi\|_{L^2(M)} \int dk\,|\widehat{V}(k)|\,|k|^2 \\ &=\; C\varepsilon^2 \|\phi\psi_0\|_{\mathcal{H}} \int dk\,|\widehat{V}(k)|\,|k|^2\,. \end{aligned}$$

To see (42) note that one can chose $\psi_0(p)$ such that the map $p \mapsto \psi_0(p) \in \mathcal{H}_{\mathrm{f}}$ is twice continuously differentiable Lebesgue almost everywhere. This is because the projections $P_0(p)$ are twice continuously differentiable and hence one can locally define $\psi_0(p) = P_0(p)\psi_0(p_n)/\|P_0(p)\psi_0(p_n)\|$. Now one can cover Λ_{g} by finitely many open disjoint sets U_n such that $\Lambda_{\mathrm{g}} \setminus \cup_n U_n$ is a set of Lebesgue measure zero and $\psi_0(p)$ can be defined on the closure of each U_n in the way described above. One obtains a family $\psi_0(p)$ of eigenfunctions which are twice continuously differentiable except at the boundaries between the sets, where we pick $\psi_0(p)$ with arbitrary phase. Wherever $\psi_0(p)$ is twice continuously differentiable, Taylor expansion yields $\psi_0(p-\varepsilon k) = \psi_0(p) - \varepsilon k\cdot\nabla_p\psi_0(p) + \frac{1}{2}\varepsilon^2 k\cdot H(\psi_0)(p')k$, where $H(\psi_0)$ denotes the Hessian and $\frac{1}{2}\varepsilon^2 k\cdot H(\psi_0)(p')k$ is the Lagrangian remainder. In view of $\langle\psi_0(p),\nabla_p\psi_0(p)\rangle_{\mathcal{H}_{\mathrm{f}}} = 0$, which follows from comparing (39) with

$$(\nabla_p P_0\psi)(p) = \langle\psi_0(p),\psi\rangle_{\mathcal{H}_{\mathrm{f}}}\nabla_p\psi_0(p) + \langle\nabla_p\psi_0(p),\psi\rangle_{\mathcal{H}_{\mathrm{f}}}\psi_0(p)\,,$$

we obtain

$$|\langle\psi_0(p),\psi_0(p-\varepsilon k)\rangle_{\mathcal{H}_{\mathrm{f}}} - 1| \leq C(p)|k|^2\varepsilon^2\,.$$

Here $C(p) = \frac{1}{2}\sum_{i,j} |\langle\psi_0(p'), \partial_{p_i}\partial_{p_j}\psi_0(p')\rangle|$. However, $C(p)$ is bounded uniformly in p, since $\psi_0(p)$ is twice continuously differentiable, by construction, on each compact $\overline{U}_n$.

$\square$

Remark 12. In the case $M = \mathbb{R}^d$ one easyly gets $\psi_0 \in C^2(\Lambda_{\rm g}, \mathcal{H}_{\rm f})$. However, we want to include the case where $\Lambda_{\rm g} = M$ is a torus. Then the construction of a phase function making $\psi_0 \in C^2(\Lambda_{\rm g}, \mathcal{H}_{\rm f})$ becomes highly nontrivial (cf., e.g., [10]). In the proof of Lemma 11 that was not needed and therefore we relied on "local diffferentiability" only.

With the help of Lemma 11 and (37), we have

$$\begin{aligned}
&\left(U\left(\frac{t}{\varepsilon}\right) - \mathcal{U}^* U_1\left(\frac{t}{\varepsilon}\right)\mathcal{U}\right) P_{\rm i} = \\
&= \; -iU\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\, U(-s)\,(H - H_{\rm diag})\, P_{\rm m2}\,\mathcal{U}^* U_1(s)\mathcal{U}\, P_{\rm i} + O(\varepsilon)\,.
\end{aligned}$$

In the following lemma we isolate the leading order term in $(H - H_{\rm diag})P_{\rm m2}$.

Lemma 13.

$$(H - H_{\rm diag})P_{\rm m2} = -\varepsilon Q_{\rm g}(\nabla_p P_{\rm g})P_{\rm m3}\cdot F^\varepsilon P_{\rm m2} + O(\varepsilon^2)\,, \tag{43}$$

where $(\nabla_p P_{\rm g}) = \int_{\Lambda_{\rm g}}^{\oplus} dp\, \nabla_p P_0(p)$, $Q_{\rm g} = \int_{\Lambda_{\rm g}}^{\oplus} dp\, Q_0(p)$ and

$$(F^\varepsilon\psi)(p) := -i\int dk\, k\,\widehat{V}(k)\psi(p - \varepsilon k)\,.$$

Proof. Using Lemma 9 we obtain

$$\begin{aligned}
(H - H_{\rm diag})P_{\rm m2}\psi &= P_{\rm g}^{\perp} V^\varepsilon P_{\rm m2}\psi \\
&= \int_{\Lambda_{\rm m3}}^{\oplus} dp\,\left(Q_0(p)\int dk\,\widehat{V}(k)P_0(p-\varepsilon k)(P_{\rm m2}\psi)(p-\varepsilon k)\right) + O(\varepsilon^2)\,.
\end{aligned}$$

Since $P_0 : \Lambda_{\rm g} \to \mathcal{L}(\mathcal{H}_{\rm f})$ is twice continuously differentiable, we have that

$$P_0(p - \varepsilon k) = P_0(p) - \varepsilon k\cdot(\nabla_p P_0)(p) + \varepsilon^2\, k\cdot H(P_0)(p'(p,\varepsilon k))\cdot k\,, \tag{44}$$

where the last term is the Lagrangian remainder with H denoting the Hessian. Hence, for $p \in \Lambda_{\rm g}$,

$$\begin{aligned}
&\int dk\,\widehat{V}(k)P_0(p-\varepsilon k)(P_{\rm m2}\psi)(p-\varepsilon k) \\
&= \int dk\,\widehat{V}(k)\big(P_0(p) - \varepsilon k\cdot(\nabla_p P_0)(p)\big)(P_{\rm m2}\psi)(p-\varepsilon k) \qquad (45) \\
&\quad + \varepsilon^2\int dk\,\widehat{V}(k)\,k\cdot H(P_0)(p'(p,\varepsilon k))\cdot k\,(P_{\rm m2}\psi)(p-\varepsilon k)\,. \qquad (46)
\end{aligned}$$

Since

$$\left\| \mathbb{1}_{\Lambda_{\mathrm{m3}}}(\cdot) \int dk\, \widehat{V}(k)\, k \cdot H(P_0)(p'(\cdot, \varepsilon k)) \cdot k\, (P_{\mathrm{m2}}\psi)(\cdot - \varepsilon k) \right\|_{\mathcal{H}} \tag{47}$$

$$\leq \int dk\, \widehat{V}(k) \left\| \mathbb{1}_{\Lambda_{\mathrm{m3}}}(\cdot)\, k \cdot H(P_0)(p'(\cdot, \varepsilon k)) \cdot k\, (P_{\mathrm{m2}}\psi)(\cdot - \varepsilon k) \right\|_{\mathcal{H}} \tag{48}$$

$$\leq \sup_{p \in \Lambda_{\mathrm{m2}}} \|H(P_0)(p)\| \int dk\, |\widehat{V}(k)|\, |k|^2\, \|\mathbb{1}_{\Lambda_{\mathrm{m3}}}(\cdot)\, (P_{\mathrm{m2}}\psi)(\cdot - \varepsilon k)\|_{\mathcal{H}}$$

$$\leq C\, \|\psi\|_{\mathcal{H}} \int dk\, |\widehat{V}(k)| k^2 \,,$$

(46) is $O(\varepsilon^2)$ in $\mathcal{L}(\mathcal{H}_{\mathrm{f}})$ and multiplying (45) with $Q_0(p)$ from the left establishes (43).

□

Including Lemma 13 we are left with

$$\left(U\left(\frac{t}{\varepsilon}\right) - \mathcal{U}^* U_1 \left(\frac{t}{\varepsilon}\right) \mathcal{U} \right) P_{\mathrm{i}} =$$
$$= i\varepsilon U\left(\frac{t}{\varepsilon}\right) \int_0^{\frac{t}{\varepsilon}} ds\, U(-s) Q_{\mathrm{g}}(\nabla_p P_{\mathrm{g}}) P_{\mathrm{m3}} \cdot F^\varepsilon P_{\mathrm{m2}} \mathcal{U}^* U_1(s) \mathcal{U}\, P_{\mathrm{i}} + O(\varepsilon)\,.$$

To exploit the time averaging we write $Q_{\mathrm{g}}(\nabla_p P_{\mathrm{g}})$ as a time derivative, at least in approximation. Let

$$B(p) := R_{E(p)}(H_0(p))^2 Q_0(p) (\nabla_p H_0)(p) P_0(p)\,.$$

Lemma 14.

$$Q_{\mathrm{g}}(\nabla_p P_{\mathrm{g}}) P_{\mathrm{m3}} = [B, H_0] P_{\mathrm{m3}} \tag{49}$$
$$= [B, H] P_{\mathrm{m3}} + O(\varepsilon) \tag{50}$$

and

$$[F^\varepsilon, H] P_{\mathrm{m2}} = O(\varepsilon) \tag{51}$$

in $\mathcal{L}(\mathcal{H})$.

Proof. Recalling (39), we have for $p \in \Lambda_{\mathrm{g}}$ that

$$Q_0(p)(\nabla_p P_0)(p) P_0(p) = -R_{E(p)}(H_0(p)) Q_0(p) (\nabla_p H_0)(p) P_0(p)$$

and thus (49) follows from direct computation,

$$B(p) H_0(p) - H_0(p) B(p) =$$
$$= -(H_0(p) - E(p)) R_{E(p)}(H_0(p))^2 Q_0(p) (\nabla_p H_0)(p) P_0(p)$$
$$= -R_{E(p)}(H_0(p)) Q_0(p) (\nabla_p H_0)(p) P_0(p)$$
$$= Q_0(p)(\nabla_p P_0)(p) P_0(p)\,.$$

For (50) we need to show that $[B,V^\varepsilon]P_{\mathrm{m3}} = O(\varepsilon)$:

$$[B,V^\varepsilon]P_{\mathrm{m3}} = [R_E(H_0)^2 Q_{\mathrm{g}}(\nabla_p H_0)P_{\mathrm{g}}, V^\varepsilon]P_{\mathrm{m3}}\,.$$

It follows from the proof of Lemma 13 and using Lemma 9 that

$$\begin{aligned}
[P_{\mathrm{g}},V^\varepsilon]P_{\mathrm{m3}} &= P_{\mathrm{g}}V^\varepsilon P_{\mathrm{m3}} - P_{\mathrm{g}}V^\varepsilon P_{\mathrm{g}}P_{\mathrm{m3}} - Q_{\mathrm{g}}V^\varepsilon P_{\mathrm{g}}P_{\mathrm{m3}} + O(\varepsilon)\\
&= -Q_{\mathrm{g}}V^\varepsilon P_{\mathrm{g}}P_{\mathrm{m3}} + O(\varepsilon) = O(\varepsilon)\,.
\end{aligned}$$

Hence

$$[B,V^\varepsilon]P_{\mathrm{m3}} = [R_E(H_0)^2 Q_{\mathrm{g}}(\nabla_p H_0), V^\varepsilon]P_{\mathrm{g}}P_{\mathrm{m3}} + O(\varepsilon)\,.$$

Using again Lemma 9 and the fact that also $\nabla_p H_0$ acts on p-fibers, we observe that

$$\begin{aligned}
&\big([\nabla_p H_0, V^\varepsilon]P_{\mathrm{m3}}\psi\big)(p) = \mathbb{1}_{\Lambda_{\mathrm{g}}}(p)\big([\nabla_p H_0, V^\varepsilon]P_{\mathrm{m3}}\psi\big)(p) + O(\varepsilon)\\
&= \mathbb{1}_{\Lambda_{\mathrm{g}}}(p)\int dk\,\widehat{V}(k)\big((\nabla_p H_0)(p) - (\nabla_p H_0)(p-\varepsilon k)\big)(P_{\mathrm{m3}}\psi)(p-\varepsilon k) + O(\varepsilon)\\
&= \varepsilon\mathbb{1}_{\Lambda_{\mathrm{g}}}(p)\int dk\,\widehat{V}(k)\,k\cdot H(H_0)(p'(k))(P_{\mathrm{m3}}\psi)(p-\varepsilon k) + O(\varepsilon) \qquad (52)
\end{aligned}$$

where $H(H_0)(p'(k))$ denotes again the Hessian evaluated at the appropriate point $p'(k)$. By assumption we have that $H(H_0)(p)$ is bounded uniformly for $p\in\Lambda_{\mathrm{g}}$ in the graph norm, which is equivalent to the $\mathcal{H}$-norm on $\mathrm{span}\{\psi_0(p), p\in\Lambda_{\mathrm{g}}\}$. Thus the above expression is $O(\varepsilon)$ as $\varepsilon\to 0$ (cf. estimate (47)). Now

$$[B,V^\varepsilon]P_{\mathrm{m3}} = [R_E(H_0)^2 Q_{\mathrm{g}}, V^\varepsilon](\nabla_p H_0)P_{\mathrm{m3}} + O(\varepsilon)\,.$$

Since $R_{E(p)}(H_0(p))^2 Q_0(p)$ is bounded and differentiable with respect to p, one can show by the same line of arguments as in the case of $[P_{\mathrm{g}},V^\varepsilon]P_{\mathrm{m2}}$ (cf. proof of Lemma 13) that

$$[R_E(H_0)^2 Q_{\mathrm{g}}, V^\varepsilon](\nabla_p H_0)P_{\mathrm{m3}} = O(\varepsilon)\,.$$

Thus we showed (50). For (51) first observe that $[F^\varepsilon, V^\varepsilon] = 0$. $[F^\varepsilon, H_0] = O(\varepsilon)$ follows from an estimate analogous to (52).

□

Let $A = B\cdot P_{\mathrm{m3}}F^\varepsilon$. Then Lemma 14 yields

$$Q_{\mathrm{g}}(\nabla_p P_{\mathrm{g}})\cdot P_{\mathrm{m3}}F^\varepsilon P_{\mathrm{m2}} = [A,H]P_{\mathrm{m2}} + O(\varepsilon)$$

in $\mathcal{L}(\mathcal{H})$, where $[P_{\mathrm{m3}}F^\varepsilon, H]P_2 = O(\varepsilon)$ follows immediately from (51) Let $A(t) = U(-t)AU(t)$, then

$$\begin{aligned}
&\left(U\left(\frac{t}{\varepsilon}\right) - \mathcal{U}^* U_1\left(\frac{t}{\varepsilon}\right)\mathcal{U}\right)P_{\mathrm{i}} =\\
&= i\varepsilon U\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,U(-s)[A,H]U(s)U(-s)P_{\mathrm{m2}}\,\mathcal{U}^*U_1(s)\mathcal{U}\,P_{\mathrm{i}} + O(\varepsilon)
\end{aligned}$$

$$
\begin{aligned}
&= -\varepsilon U\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,\left(\frac{d}{ds}A(s)\right) U(-s)P_{\mathrm{m}2}\,\mathcal{U}^*U_1(s)\mathcal{U}\,P_{\mathrm{i}} + O(\varepsilon)\\
&= -\varepsilon U\left(\frac{t}{\varepsilon}\right)\Big[A(s)U(-s)P_{\mathrm{m}2}\,\mathcal{U}^*U_1(s)\mathcal{U}\Big]_0^{\frac{t}{\varepsilon}} P_{\mathrm{i}} + O(\varepsilon)\\
&\quad + \varepsilon U\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,A(s)\frac{d}{ds}\Big(U(-s)P_{\mathrm{m}2}\,\mathcal{U}^*U_1(s)\mathcal{U}\Big)\,P_{\mathrm{i}}\\
&= i\varepsilon U\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,A(s)U(-s)\Big(HP_{\mathrm{m}2}\mathcal{U}^* - P_{\mathrm{m}2}\mathcal{U}^*H_1\Big)\,U_1(s)\mathcal{U}\,P_{\mathrm{i}} + O(\varepsilon)\\
&= i\varepsilon U\left(\frac{t}{\varepsilon}\right)\int_0^{\frac{t}{\varepsilon}} ds\,A(s)U(-s)\Big(H - \mathcal{U}^*H_1\mathcal{U}\Big)P_{\mathrm{m}1}\,\mathcal{U}^*U_1(s)\mathcal{U}\,P_{\mathrm{i}} + O(\varepsilon)\,.
\end{aligned}
$$

For the last equality we used again Lemma 1 and the fact that Lemma 9 guarantees $P_{\mathrm{m}2}\mathcal{U}^*H_1\mathcal{U}P_{\mathrm{m}1} = \mathcal{U}^*H_1\mathcal{U}P_{\mathrm{m}1} + O(\varepsilon^2)$. Finally also the last integral is $O(\varepsilon)$, since we showed already that $(H - \mathcal{U}^*H_1\mathcal{U})P_{\mathrm{m}1} = O(\varepsilon)$ and the proof of Theorem 2 is completed.

5 Convergence of the macroscopic observables

In this section we prove Theorem 3 and Theorem 4. We start by showing

Lemma 15. Let $g \in C^\infty(M)$ such that $g\circ\pi_p \in S^0_0(1)$ and let the assumptions of Theorem 3 be satisfied. Then

$$
\left\|\Big(U(-t/\varepsilon)\,(g(p)\otimes \mathbf{1})\,U(t/\varepsilon) - \mathcal{U}^*U_1(-t/\varepsilon)g(p)U_1(t/\varepsilon)\mathcal{U}\Big)P_{\mathrm{i}}\right\| \le C\,\varepsilon\,, \qquad (53)
$$

where $g(p) := (g\circ\pi_p)^{\mathrm{W},\varepsilon}$ denotes the operator of multiplication with the function g on $L^2(M)$.

Proof. Let $\Lambda_{\mathrm{i},\gamma} = \Lambda_{\mathrm{i}} + \gamma$ and $P_{\mathrm{i},\gamma}$ the corresponding projection. Let γ be sufficiently small to ensure that $\mathrm{supp}\big((\mathbb{1}_{\Lambda_{\mathrm{i}},\gamma}\circ\Phi_p^{-t})(x,\cdot)\big) + \delta/2 \subset \Lambda_{\mathrm{g}}$ for all $t \in [0, T^\delta_{\mathrm{m}}]$ and $x \in \mathbb{R}^d$. We abbreviate $g(t) := U(-t/\varepsilon)(g(p)\otimes\mathbf{1})U(t/\varepsilon)$ and $g_1(t) := U_1(-t/\varepsilon)g(p)U_1(t/\varepsilon)$ and split (53) into two parts:

$$
\begin{aligned}
\left\|\Big(g(t) - \mathcal{U}^*g_1(t)\mathcal{U}\Big)P_{\mathrm{i}}\right\| \;\le\; &\left\|P_{\mathrm{i},\gamma}\Big(g(t) - \mathcal{U}^*g_1(t)\mathcal{U}\Big)P_{\mathrm{i}}\right\| &(54)\\
&+\left\|P^{\perp}_{\mathrm{i},\gamma}\Big(g(t) - \mathcal{U}^*g_1(t)\mathcal{U}\Big)P_{\mathrm{i}}\right\|. &(55)
\end{aligned}
$$

Using $\mathcal{U}\mathcal{U}^*P_{\mathrm{g}} = P_{\mathrm{g}}$ on $L^2(M)$ we find that (54) can be estimated as

$$
\begin{aligned}
&\left\|P_{\mathrm{i},\gamma}\Big(g(t) - \mathcal{U}^*g_1(t)\mathcal{U}\Big)P_{\mathrm{i}}\right\|\\
&\overset{\text{Lemma 1}}{\le} \left\|P_{\mathrm{i},\gamma}\Big(U(-t/\varepsilon)\,\mathcal{U}^* - \mathcal{U}^*U_1(-t/\varepsilon)\Big)\mathcal{U}\mathcal{U}^*P_{\mathrm{g}}\,g(p)\,U_1(t/\varepsilon)\mathcal{U}P_{\mathrm{i}}\right\| + O(\varepsilon^2)\\
&\qquad + \left\|P_{\mathrm{i},\gamma}U(-t/\varepsilon)\,(g(p)\otimes\mathbf{1})\Big(U(t/\varepsilon) - \mathcal{U}^*U_1(t/\varepsilon)\mathcal{U}\Big)P_{\mathrm{i}}\right\|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \quad \left\| P_{\mathrm{i},\gamma}\Big(U\,(-t/\varepsilon) - \mathcal{U}^* U_1\,(-t/\varepsilon)\,\mathcal{U}\Big)\right\| \left\| P_{\mathrm{g}}\, g(p)\, U_1\,(t/\varepsilon)\,\mathcal{U} P_{\mathrm{i}}\right\| + O(\varepsilon^2)\\
&\qquad + \left\| P_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\ (g(p)\otimes \mathbf{1})\, P_{\mathrm{g}}\right\| \left\| \Big(U\,(t/\varepsilon) - \mathcal{U}^* U_1\,(t/\varepsilon)\mathcal{U}\Big) P_{\mathrm{i}}\right\|\\
&\overset{\text{Theorem 2}}{=} \quad O(\varepsilon)\,.
\end{aligned}
$$

It remains to show that in (55) both terms are of order $O(\varepsilon)$ separately. By construction we can pick f_{i}, $f^{\perp}_{\mathrm{i},\gamma} \in C^\infty(M)$ such that $f_{\mathrm{i}}\big|_{\Lambda_{\mathrm{i}}} = 1$, $f^{\perp}_{\mathrm{i},\gamma}\big|_{M\setminus\Lambda_{\mathrm{i},\gamma}} = 1$ and $\mathrm{supp} f_{\mathrm{i}} \bigcap \mathrm{supp} f^{\perp}_{\mathrm{i},\gamma} = \emptyset$. Using Egorov's Theorem and Corollary 7 (i) we get

$$
\begin{aligned}
g_1(t)\,(f_{\mathrm{i}}\circ\pi_p)^{\mathrm{W},\varepsilon} &= (g\circ\Phi^t_p)^{\mathrm{W},\varepsilon}\,(f_{\mathrm{i}}\circ\pi_p)^{\mathrm{W},\varepsilon} + O(\varepsilon^2)\\
&= ((g\circ\Phi^t_p)\,(f_{\mathrm{i}}\circ\pi_p))^{\mathrm{W},\varepsilon} + O(\varepsilon)
\end{aligned}
$$

and therefore, with Corollary 7 (ii), that

$$
\begin{aligned}
\left\|P^{\perp}_{\mathrm{i},\gamma}\mathcal{U}^*\, g_1(t)\mathcal{U}\, P_{\mathrm{i}}\right\| &\leq \left\|(f^{\perp}_{\mathrm{i},\gamma}\circ\pi_p)^{\mathrm{W},\varepsilon}\, g_1(t)\,(f_{\mathrm{i}}\circ\pi_p)^{\mathrm{W},\varepsilon}\,\mathcal{U}\, P_{\mathrm{i}}\right\|\\
&= \left\|(f^{\perp}_{\mathrm{i},\gamma}\circ\pi_p)^{\mathrm{W},\varepsilon}\ \big((g\circ\Phi^t_p)\,(f_{\mathrm{i}}\circ\pi_p)\big)^{\mathrm{W},\varepsilon}\ \mathcal{U}\, P_{\mathrm{i}}\right\| + O(\varepsilon)\\
&= O(\varepsilon)\,,
\end{aligned}
$$

since $f^{\perp}_{\mathrm{i},\gamma}\circ\pi_p$ and $(g\circ\Phi^t_p)\,(f_{\mathrm{i}}\circ\pi_p)$ are disjointly supported. Finally we compute that

$$
\begin{aligned}
\left\|P^{\perp}_{\mathrm{i},\gamma} g(t)\, P_{\mathrm{i}}\right\| &= \left\|P^{\perp}_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\ (g(p)\otimes\mathbf{1})\, U\,(t/\varepsilon)\, P_{\mathrm{i}}\right\|\\
&\overset{\text{Theorem 2}}{=} \left\|P^{\perp}_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\ \mathcal{U}^*\ g(p) U_1\,(t/\varepsilon)\mathcal{U} P_{\mathrm{i}}\right\| + O(\varepsilon)\\
&\overset{\text{Egorov}}{=} \left\|P^{\perp}_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\mathcal{U}^* U_1\,(t/\varepsilon)\,(g\circ\Phi^t_{\mathrm{p}})^{\mathrm{W},\varepsilon}(f_{\mathrm{i}}\circ\pi_p)^{\mathrm{W},\varepsilon}\mathcal{U} P_{\mathrm{i}}\right\| + O(\varepsilon)\\
&\overset{\text{Corollary 7 (i)}}{=} \left\|P^{\perp}_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\mathcal{U}^* U_1\,(t/\varepsilon)\,(f_{\mathrm{i}}\circ\pi_p)^{\mathrm{W},\varepsilon}(g\circ\Phi^t_{\mathrm{p}})^{\mathrm{W},\varepsilon}\mathcal{U} P_{\mathrm{i}}\right\| + O(\varepsilon)\\
&\leq \left\|P^{\perp}_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\mathcal{U}^* U_1\,(t/\varepsilon)\mathcal{U} P_{\mathrm{i},\gamma}\,\mathcal{U}^*(g\circ\Phi^t_{\mathrm{p}})^{\mathrm{W},\varepsilon}\mathcal{U} P_{\mathrm{i}}\right\| + O(\varepsilon)\\
&\overset{\text{Theorem 2}}{=} \left\|P^{\perp}_{\mathrm{i},\gamma} U\,(-t/\varepsilon)\, U\,(t/\varepsilon)\, P_{\mathrm{i},\gamma}(g\circ\Phi^t_{\mathrm{p}})^{\mathrm{W},\varepsilon}\mathcal{U} P_{\mathrm{i}}\right\| + O(\varepsilon)\\
&= O(\varepsilon)\,.
\end{aligned}
$$

For the third equality we inserted $U_1(t/\varepsilon)U_1(-t/\varepsilon) = \mathbf{1}$ in order to apply Egorov's Theorem on $g_1(t)$.

□

We now come to the proof of (24). We have (at the moment only formally) that

$$
x^\varepsilon(t) = i\,\varepsilon\nabla_p\otimes\mathbf{1}\ +\ \int_0^t ds\, U(-s/\varepsilon)\,[i\nabla_p\otimes\mathbf{1}, H]U(s/\varepsilon)\,.
$$

On $H^1(\mathrm{I\!R}^d)\otimes\mathcal{H}_{\mathrm{f}}$ this gives

$$
x^\varepsilon(t)P_{\mathrm{i}} \quad = \quad i\,\varepsilon\nabla_p\otimes\mathbf{1} + \int_0^t ds\, U(-s/\varepsilon)\,[i\nabla_p\otimes\mathbf{1}, H]U(s/\varepsilon)\, P_{\mathrm{i}}
$$

$$
\begin{aligned}
&= i\,\varepsilon\nabla_p \otimes \mathbf{1} + \int_0^t ds\, U(-s/\varepsilon)\,(\nabla_p H_0)\, P_{\mathrm{g}} U(s/\varepsilon)\, P_{\mathrm{i}} \\
&\qquad + \int_0^t ds\, U(-s/\varepsilon)\,(\nabla_p H_0)\, P_{\mathrm{g}}^{\perp} U(s/\varepsilon)\, P_{\mathrm{i}} \\
&= i\,\varepsilon\nabla_p \otimes \mathbf{1} + \int_0^t ds\, U(-s/\varepsilon)\,\left(i\nabla E(p) \otimes \mathbf{1}\right) U(s/\varepsilon)\, P_{\mathrm{i}} \\
&\qquad + O(\varepsilon)\,, \qquad (56)
\end{aligned}
$$

where we prove in a moment that this is well defined on $H^1(\mathbb{R}^d) \otimes \mathcal{H}_{\mathrm{f}}$ and that the last equality in (56) holds.

On the other hand, again on $H^1(\mathbb{R}^d) \otimes \mathcal{H}_{\mathrm{f}}$,

$$
\mathcal{U}^*\, x_1^{\varepsilon}(t)\, \mathcal{U}\, P_{\mathrm{i}} = \mathcal{U}^*\, i\,\varepsilon\nabla_p \mathcal{U}\, P_{\mathrm{i}} \;+\; \int_0^t ds\, \mathcal{U}^*\, U_1(-s/\varepsilon)\, i\nabla E(p)\, U_1(s/\varepsilon)\, \mathcal{U}\;P_{\mathrm{i}}\,.
$$

Since $\nabla E \in C_0^\infty(M)$ we can apply Lemma 15 and finish by showing that

$$
\left\| \left(i\,\varepsilon\nabla_p \otimes \mathbf{1} - \mathcal{U}^*\, i\,\varepsilon\nabla_p \mathcal{U} \right) P_{\mathrm{i}} \right\| = O(\varepsilon)\,. \qquad (57)
$$

Let $\psi = \phi(p)\psi_0(p)$, then

$$
\begin{aligned}
\left\| \left(i\,\varepsilon\nabla_p \otimes \mathbf{1} - \mathcal{U}^*\, i\,\varepsilon\nabla_p \mathcal{U} \right) \phi\psi_0 \right\|_{\mathcal{H}} &= \varepsilon \left\| \phi\nabla_p\psi_0 \right\|_{\mathcal{H}} \\
&\leq \varepsilon\, C \|\phi\|_{L^2(M)} = \varepsilon\, C \|\psi\|_{\mathcal{H}}\,,
\end{aligned}
$$

which proves (57). To complete the proof of Theorem 3 we are left to show the last equality of (56), i.e. that

$$
(\nabla_p H_0)\, P_{\mathrm{g}}^{\perp} U(s/\varepsilon)\, P_{\mathrm{i}} = O(\varepsilon)
$$

uniformly for $s \in [0, T]$. If $(\nabla_p H_0)$ was a bounded operator, Lemma 1 translated to the full dynamics via Theorem 2 would take care of that. Recall, however, that by assumption $(\nabla_p H_0)(H_0 - i)^{-1}$ is bounded and hence it suffices to show

$$
H_0\, P_{\mathrm{g}}^{\perp} U(s/\varepsilon)\, P_{\mathrm{i}} = O(\varepsilon)\,.
$$

We calculate

$$
\begin{aligned}
H_0\, P_{\mathrm{g}}^{\perp} U(s/\varepsilon)\, P_{\mathrm{i}} &= P_{\mathrm{g}}^{\perp}\, H_0\, U(s/\varepsilon)\, P_{\mathrm{i}} \\
&\overset{(*)}{=} P_{\mathrm{g}}^{\perp}\, H\, U(s/\varepsilon)\, P_{\mathrm{i}} + O(\varepsilon) \\
&= P_{\mathrm{g}}^{\perp}\, U(s/\varepsilon)\, H\, P_{\mathrm{i}} + O(\varepsilon) \\
&\overset{\text{Lemma 9}}{=} P_{\mathrm{g}}^{\perp}\, U(s/\varepsilon)\, P_{\mathrm{i},\gamma} H\, P_{\mathrm{i}} + O(\varepsilon) \\
&\overset{\text{Lemma 1}}{=} O(\varepsilon)\,, \qquad (58)
\end{aligned}
$$

where (*) follows from Lemma 1 together with $[P_{\mathrm{g}}^{\perp}, V^{\varepsilon}]P_{\mathrm{m}} = O(\varepsilon)$ (cf. proof of Lemma 14) and $P_{\mathrm{i},\gamma}$ is defined as in the proof of Lemma 15.

Finally we sketch the proof of Theorem 4. First we conclude from Theorem 3 that in the special case of $a = f \circ \pi_x \in S_0^0(1) \cap C_\infty(\mathbb{R}^d \times M)$, for some $f \in C^\infty(\mathbb{R}^d)$, (28) holds. To see this, note that by the functional calculus for self-adjoint operators it suffices to show

$$\|(f(x^\varepsilon(t)) - f(\mathcal{U}^*\, x_1^\varepsilon(t)\,\mathcal{U}))\, P_1\| \leq C\,\varepsilon\,. \tag{59}$$

However, (59) follows from a standard approximation argument using the general Stone-Weierstraß Theorem and norm-convergence of the resolvents (cf., e.g., Theorem VIII.20 in [11]).

Next we observe that for $a(x,p) = f(x)g(p)$ Lemma 1 and Corollary 7 imply (28) in a rather straightforward way.

Finally one can approximate general $a \in S_0^0(1) \cap C_\infty(\mathbb{R}^d \times M)$ by products referring again to the general Stone-Weierstraß Theorem.

6 Semi-classical distributions

Theorem 2 establishes that, restricted to the ground state band, the full unitary group $U(t)$ and the approximate one-particle unitary $U_1(t)$ are uniformly close to each other and Theorems 3 and 4 lift this assertion to semi-classical observables. Experimentally measured are empirical statistics of suitable observables, like position and momentum, and we still have the task to investigate in what sense they are approximated through the time evolution $U_1(t)$ for $\varepsilon \ll 1$.

For small ε, H_1 itself is a semi-classical Hamiltonian, which means that empirical distributions can be determined through the classical flow Φ^t generated by H_1. Somewhat crudely the scheme is as follows: one choses an initial wave function ψ^ε, which may or may not depend on ε, such that for small ε it determines the measure $\rho_{\mathrm{cl}}(dx\,dp)$ on phase space. We evolve ψ^ε as $\psi_t^\varepsilon = e^{-iH_1 t/\varepsilon}\psi^\varepsilon$ and $\rho_{\mathrm{cl}}(dx\,dp, t) = (\rho_{\mathrm{cl}} \circ \Phi^{-t})(dx\,dp)$. Then the empirical distributions computed from ψ_t^ε agree with those of $\rho_{\mathrm{cl}}(t)$ up to errors of order ε, i.e. quantum distributions are well approximated by their classical counterpart. In our context the true statistics must be compared with $U(t/\varepsilon)\psi^\varepsilon$ and we have to make sure that the approximations of Theorem 4 are so sharp that we can draw the same conclusions for $U(t/\varepsilon)\psi^\varepsilon$.

There are various "schools" which differ in what initial ψ's are regarded as physically natural.

(i) *wave packet dynamics.* The initial wave function is well localized in macroscopic position and momentum, i.e. $\rho_{\mathrm{cl}}(dx\,dp) = \delta(x - x_0)\delta(p - p_0)\,dx\,dp$. Then the wave packet follows the classical orbit , which only reflects that $\rho_{\mathrm{cl}}(dx dp, t) = \delta(x - x_t)\delta(p - p_t)\,dx\,dp$.

(ii) *microscopic wave function independent of* ε. On the macrscopic scale the position is localized, but there is momentum spread. Therefore $\rho_{\mathrm{cl}}(dx\,dp) = \delta(x)|\widehat{\psi}(p)|^2\,dx\,dp$. Such a choice is appropriate immediately after a scattering event. Then ψ is still localized at the scatterer but has considerable momentum spread.

(iii) *WKB.* The wave function is taken to be build up from local plane waves, which means it has the form $\psi^\varepsilon(x) = \varepsilon^{d/2} f(\varepsilon x) e^{iS(\varepsilon x)/\varepsilon}$ on the microscopic scale. ψ^ε is spread over macroscopic distance, but at any given point it has a sharp momentum. ψ^ε yields the phase space measure $\rho_{\mathrm{cl}}(dx\,dp) = |f(x)|^2 \delta(p - \nabla S(x))\,dx\,dp$.

As will be discussed, all three classes of initial wave functions, and in principle more, can be handled under rather mild regularity assumptions in a unified fashion. The approximation from $U(t/\varepsilon)$ to $U_1(t/\varepsilon)$ is covered by Theorems 2 and 4 and the semi-classical distributions for $U_1(t)$ are a consequence of standard results of semi-classical analysis as presented, for example, in [2]. Since we did not find a discussion of sufficient generality in the literature, we explain the arguments in some detail.

Note that in the following we only treat the case $M = \mathbb{R}^d$. If M is a torus, position and (quasi)-momentum are not related by Fourier transformation, but by a Bloch-Floquet transformation. Apart from this difference the analysis goes through analogously.

6.1 Wave packets following a classical trajectory

The conceptually simplest way for a quantum particle to behave classically is to have a well localized wave function that follows a classical trajectory. Hence we consider initial wave functions with sharply peaked momentum and, at the same time, sharply peaked macroscopic position. Let the initial wave function be

$$\phi_{x_0,p_0}(x) = \varepsilon^{d/4}\, e^{ix\cdot p_0}\, \phi(\sqrt{\varepsilon}(x - x_0/\varepsilon))$$

for some $\phi \in L^2(\mathbb{R}^d)$, i.e., a wave function that is peaked on the macroscopic scale and centered at x_0, but spread out on the microscopic scale. Its Fourier transform is given by

$$\widehat{\phi}_{x_0,p_0}(p) = \varepsilon^{-d/4}\, e^{-i\frac{x_0\cdot(p-p_0)}{\varepsilon}}\, \widehat{\phi}(\frac{p-p_0}{\sqrt{\varepsilon}})\,,$$

which becomes sharply peaked at p_0 for ε small. There is no difficulty to include also asymmetric scaling with weights $\varepsilon^{1-\alpha}$ and ε^α, $0 < \alpha < 1$, and the choice $\alpha = 1/2$ was made to simplify presentation. Under the time evolution generated by

$$H_1 = E(-i\nabla_x) + V(\varepsilon x)$$

it moves along the corresponding classical trajectory starting at (x_0, p_0) following the classical flow Φ^t generated by

$$H_{\mathrm{cl}} = E(p) + V(x)\,.$$

To be consistent with the previous chapters, we continue to work in momentum representation.

Proposition 16. Let $a \in S^0_0(1)$, $\widehat{\phi} \in H^1(\mathbb{R}^d) \cap D(p)$ and $T < \infty$. Then there is a $C < \infty$ such that for $t \in [0,T]$

$$\Big| \langle \widehat{\phi}_{x_0,p_0}, a_1^\varepsilon(t)\, \widehat{\phi}_{x_0,p_0} \rangle - \big(a \circ \Phi^t\big)(x_0,p_0) \Big| \le C\sqrt{\varepsilon}\, \|\widehat{\phi}\| \big(\|\widehat{\phi}\| + \|p\widehat{\phi}\| + \|\nabla_p \widehat{\phi}\| \big)\,. \tag{60}$$

Using Theorem 4, this translates immediately to the full dynamics.

Corollary 17. Let the assumptions of Theorem 4 be satisfied. Then there is a $C < \infty$ such that for $\psi \in \big((H^1(\mathbb{R}^d) \cap D(p)) \otimes \mathcal{H}_\mathrm{f} \big) \cap \mathrm{Ran} P_\mathrm{i}$ and $t \in [0,T]$

$$\begin{aligned} &\big| \langle \psi_{x_0,p_0}, a^\varepsilon(t)\, \psi_{x_0,p_0} \rangle - \big(a \circ \Phi^t\big)(x_0,p_0) \big| \\ &\qquad \le C\sqrt{\varepsilon}\, \|\psi\| \big(\|\psi\| + \|(p \otimes \mathbf{1})\psi\| + \|(\nabla_p \otimes \mathbf{1})\psi\| \big)\,. \end{aligned}$$

Hence, when initially, on the macroscopic scale, position and momentum are both sharply defined, the wave packet follows the classical orbit without spreading even for macroscopic times. Such a situation occurs for example in particle accelerators, where one can indeed calculate the particle trajectories based solely on classical dynamics in good approximation.

Proof [of Proposition 16.] Referring to Egorov's theorem we have to compute

$$\langle \widehat{\phi}_{x_0,p_0}, (a \circ \Phi^t)^{\mathrm{W},\varepsilon}\, \widehat{\phi}_{x_0,p_0} \rangle\,.$$

One can replace $(a \circ \Phi^t)^{\mathrm{W},\varepsilon}$ by the so called standard quantization of $a \circ \Phi^t$ defined by

$$\big((a \circ \Phi^t)^{\mathrm{S},\varepsilon}\, \widehat{\phi} \big)(p) := (2\pi)^{-d/2} \int dx\, \big(a \circ \Phi^t\big)(\varepsilon x, p)\; e^{-ip\cdot x}\, \phi(x) \tag{61}$$

where the error is of order ε uniformly in $t \in [0,T]$ (cf. Chapter 7 in [2]). For the standard quantization the result becomes a simple calculation:

$$\begin{aligned} &\langle \widehat{\phi}_{x_0,p_0}, (a \circ \Phi^t)^{\mathrm{S},\varepsilon}\, \widehat{\phi}_{x_0,p_0} \rangle \; = \\ &\quad = (2\pi)^{-d/2} \int dx\, dp\, e^{i\frac{x_0\cdot(p-p_0)}{\varepsilon}}\, \widehat{\phi}^* \left(\frac{p-p_0}{\sqrt{\varepsilon}} \right) (a \circ \Phi^t)(\varepsilon x, p)\, e^{-ip\cdot x}\, e^{ix\cdot p_0} \;\times \\ &\qquad\qquad \phi\left(\sqrt{\varepsilon} \left(x - \frac{x_0}{\varepsilon} \right) \right) \\ &\quad = (2\pi)^{-d/2} \int dx\, dp\, \widehat{\phi}^* \left(\frac{p}{\sqrt{\varepsilon}} \right) (a \circ \Phi^t)(\varepsilon x + x_0, p + p_0)\, e^{-ip\cdot x}\, \phi\big(\sqrt{\varepsilon}\,(x)\big) \\ &\quad = (2\pi)^{-d/2} \int d\overline{x}\, d\overline{p}\, \widehat{\phi}^*(\overline{p}) (a \circ \Phi^t)(\sqrt{\varepsilon}\overline{x} + x_0, \sqrt{\varepsilon}\overline{p} + p_0)\, e^{-i\overline{p}\cdot\overline{x}}\, \phi(\overline{x}) \\ &\quad = (a \circ \Phi^t)(x_0,p_0) \; + \; O(\sqrt{\varepsilon})\,, \end{aligned}$$

where the last equality follows from a Taylor expansion of $a\circ\Phi^t$ around (x_0,p_0) together with the assumption that $\phi\in H^1(\mathbb{R}^d)\cap D(p)$. Note also that $(\nabla_x(a\circ\Phi^t))\in S^0_0(1)$ and $(\nabla_p(a\circ\Phi^t))\in S^0_0(1)$ and hence $(\nabla_x(a\circ\Phi^t))^{\mathrm{S},\varepsilon}$, $(\nabla_p(a\circ\Phi^t))^{\mathrm{S},\varepsilon}\in\mathcal{L}(L^2(\mathbb{R}^d))$ (cf. Chapter 7 in [2])

□

6.2 Initial wave function with momentum spread

Generally wave functions are not of the special form described in the previous subsection. Typically they arise from microscopic interactions and thus "live" on the microscopic scale and do not depend on ε. But if the shape of the initial wave function does not depend on ε it will effectively look like a delta function on the macroscopic scale at $t=0$. More precisely, let $\phi\in L^2(\mathbb{R}^d)$, then the scaled position distribution is $\varepsilon^{-d}|\phi(x/\varepsilon)|^2$ which converges to $\delta(x)$ as a measure. However, the scaled momentum distribution is still $|\widehat{\phi}(p)|^2$, since in the quotient x/t the ε's cancel and thus the initially peaked wave function will spread if evolved to times of order ε^{-1}. More generally we consider as initial wave function

$$\phi_{x_0}(x)=\phi(x-x_0/\varepsilon)\,,$$

i.e. we move ϕ to the macroscopic initial position x_0.

Then it is natural to chose

$$\rho_{\mathrm{cl}}(dx\,dp)=\delta(x-x_0)|\widehat{\phi}_{x_0}(p)|^2\,dx\,dp$$

as the corresponding classical phase space distribution at $t=0$ and evolve it according to the classical flow to

$$\rho_{\mathrm{cl}}(dx\,dp,t)=\left(\rho\circ\Phi^{-t}\right)(dx\,dp)\,.$$

Let us now, as the simplest example, compare the quantum mechanical position distribution

$$\rho_\varepsilon(dx,t)=\varepsilon^{-d}|\phi_t(x/\varepsilon)|^2\,dx\,,\qquad \phi_t=e^{-iH_1t/\varepsilon}\phi_{x_0}\,,$$

with the classical one,

$$\rho^x_{\mathrm{cl}}(dx,t)=\int\rho_{\mathrm{cl}}(dx\,dp,t)\,,$$

in the limit $\varepsilon\to0$. As a first step we calculate, with $f\in C_0^\infty$ a test function,

$$\begin{aligned}\int\rho_\varepsilon(dx,t)\,f(x) &= \langle\phi^\varepsilon_t,f\,\phi^\varepsilon_t\rangle=\langle\phi_{x_0},f(x^\varepsilon(t))\,\phi_{x_0}\rangle\\ &= \langle\widehat{\phi}_{x_0},(f\circ\Phi^t_x)^{\mathrm{W},\varepsilon}\,\widehat{\phi}_{x_0}\rangle+O(\varepsilon^2)\,. \qquad (62)\end{aligned}$$

For the last equality we used Egorov's Theorem.

To proceed we need to know how the Weyl quantization of a time evolved classical observable acts on microscopic wave functions. We postpone the proof of the following Proposition to the end of the section.

Proposition 18. Let $a \in S^0_0(1)$. Then for each $T < \infty$ there is a $C < \infty$ such that for $\phi \in D(x)$ and $t \in [0,T]$

$$\left\| \Big(\big(a \circ \Phi^t\big)^{\mathrm{W},\varepsilon} - \big(a \circ \Phi^t\big)(x_0, \cdot) \Big) \,\widehat{\phi}_{x_0} \right\| \leq C\,\varepsilon\, \|x\phi\|\,, \tag{63}$$

where $\big(a \circ \Phi^t\big)(x_0,\cdot)$ denotes the operator of multiplication with the function $\big(a \circ \Phi^t\big)(x_0,p)$ in momentum representation.

Thus (62) becomes

$$\langle \widehat{\phi}_{x_0}, (f \circ \Phi^t_x)^{\mathrm{W},\varepsilon}\, \widehat{\phi}_{x_0}\rangle = \int dp\, |\widehat{\phi}_{x_0}(p)|^2 \,\big(f \circ \Phi^t_x\big)(x_0,p) + O(\varepsilon)\,. \tag{64}$$

However, the right hand side of (64) is exactly what we were aiming for:

$$\begin{aligned} \int dp\, |\widehat{\phi}_{x_0}(p)|^2 \,\big(f \circ \Phi^t_x\big)(x_0,p) &= \int \rho_{\mathrm{cl}}(dx\,dp, 0)\big(f \circ \Phi^t_x\big)(x,p) \\ &= \int \rho_{\mathrm{cl}}(dx\,dp, t) f(x) \\ &= \int \rho^x_{\mathrm{cl}}(dx, t) f(x)\,. \end{aligned}$$

In summary, one obtains that

$$\lim_{\varepsilon\to 0} \rho_\varepsilon(dx,t) = \rho^x_{\mathrm{cl}}(dx,t) \tag{65}$$

weakly as measures. Note that this means that the position distribution of the quantum particle converges to the classical one, although the wave function is *not* following the classical orbit, but spreading. Such a situation occurs for example in a scattering experiment. After the quantum particle is scattered off the target its wave function usually has a large momentum spread and thus it also spreads in position space. However, during the process of detection it is subject to relatively weak potentials and can be treated like a classical particle, at least on the level of statistics.

Such a result holds in fact not only for the position distribution but for all semi-classical observables. For general observables $a \in S^0_0(1)$ we get analogously

$$\langle \widehat{\phi}_{x_0}, a^\varepsilon_1(t)\, \widehat{\phi}_{x_0}\rangle = \int \rho_{\mathrm{cl}}(dx\,dp, t) a(x,p) \; + \; O(\varepsilon)\,. \tag{66}$$

Now Theorem 4 allows us to immediately translate this result to the full quantum dynamics.

Corollary 19. Let the assumptions of Theorem 4 be satisfied. Then there is a $C < \infty$ such that for $\psi \in (H^1(\mathbb{R}^d) \otimes \mathcal{H}_\mathrm{f}) \cap \mathrm{Ran} P_\mathrm{i}$ and $t \in [0, T]$

$$\left| \langle \psi_{x_0}, a^\varepsilon(t)\, \psi_{x_0} \rangle - \int \rho_\mathrm{cl}(dx\, dp, t) a(x,p) \right| < C\varepsilon\, \|\psi\| (\|\psi\| + \|(\nabla_p \otimes \mathbf{1})\psi\|)\,.$$

Before we close the section with the proof of Lemma 18, let us shortly comment on an approach "dual" to Weyl quantization of classical observables. One can translate the above result to the level of distributions, i.e. generalize (65), by looking at the scaled Wigner distribution. For $\phi \in L^2(\mathbb{R}^d)$ it is defined through

$$W^\varepsilon(dx\, dp, t) = \int d\xi\, \varepsilon^{-d}\, \phi^*_{x_0}(x/\varepsilon - \xi/2, t)\, \phi_{x_0}(x/\varepsilon + \xi/2, t)\, e^{ip\cdot\xi}\, dx\, dp\,.$$

The scaled Wigner distribution yields expectations of Weyl quantized operators through

$$\langle \widehat{\phi}_{x_0},\, a_1^\varepsilon(t)\, \widehat{\phi}_{x_0} \rangle = \int W^\varepsilon(dx\, dp, t)\, a(x,p)\,.$$

This together with (66) implies

$$\lim_{\varepsilon\to 0} W^\varepsilon(dx\, dp, t) = \rho_\mathrm{cl}(dx\, dp, t)$$

weakly as measures.

On the level of the full quantum dynamics we introduce the reduced Wigner distribution for $\psi \in L^2(\mathbb{R}^d) \otimes \mathcal{H}_\mathrm{f}$ as

$$W^\varepsilon_\mathrm{rd}(dx\, dp, t) = \int d\xi\, \varepsilon^{-d} \left\langle \widehat{\psi}^*_{x_0}(x/\varepsilon - \xi/2, t), \widehat{\psi}_{x_0}(x/\varepsilon + \xi/2, t) \right\rangle_{\mathcal{H}_\mathrm{f}} e^{ip\cdot\xi}\, dx\, dp\,,$$

where here $\widehat{\ }$ stands for Fourier transformation in the first argument, i.e. for $\mathcal{F} \otimes \mathbf{1}$. Corollary 19 translates into

$$\lim_{\varepsilon\to 0} W^\varepsilon_\mathrm{rd}(dx\, dp, t) = \rho_\mathrm{cl}(dx\, dp, t)$$

weakly as measures.

Proof [of Proposition 18]. Again we replace $(a\circ\Phi^t)^{\mathrm{W},\varepsilon}$ by the standard quantization of $a \circ \Phi^t$ and get

$$\begin{aligned} \big((a\circ\Phi^t)^{\mathrm{S},\varepsilon}\, \widehat{\phi}_{x_0}\big)(p) &:= (2\pi)^{-d/2} \int dx\, \big(a\circ\Phi^t\big)(\varepsilon x, p)\; e^{-ip\cdot x}\, \phi(x - x_0/\varepsilon) \\ &= (2\pi)^{-d/2} \int dx\, \big(a\circ\Phi^t\big)(\varepsilon x + x_0, p)\; e^{-ip\cdot x}\, e^{-ip\cdot x_0/\varepsilon}\, \phi(x)\,. \end{aligned} \tag{67}$$

Taylor expansion yields

$$(a\circ\Phi^t)(\varepsilon x + x_0, p) = (a\circ\Phi^t)(x_0, p) + \varepsilon x \cdot (\nabla_x (a \circ \Phi^t))(y(x), p)\,.$$

We insert this into (67) and conclude as in the proof of Proposition 16 that the term proportional to ε is bounded in norm by our assumption that $\phi \in D(x)$. This proves (63).

□

6.3 Initial wave function of WKB form

In the previous case of an initially localized wave function the different momentum components travel at different velocities and therefore such a wave function spreads on the macroscopic scale. After some time one expects the wave function to have locally well defined momentum as long as no interference occurs, i.e. it should be of WKB type. On the microscopic scale a WKB wave function has the form

$$\phi(x) = \varepsilon^{d/2} f(\varepsilon x) e^{i\frac{S(\varepsilon x)}{\varepsilon}}\,,$$

with f and S real valued. Hence it locally looks like a plane wave with momentum $\nabla S(\varepsilon x)$ and amplitude $f(\varepsilon x)$.

Time-dependent WKB approximation is concerned with showing that the time evolution of such a wave function is in first order given by

$$\left(e^{-iH_1 t/\varepsilon}\phi\right)(x) \approx f_t(\varepsilon x) e^{i\frac{S_t(\varepsilon x)}{\varepsilon}}\,,$$

where S_t is the solution of the classical Hamilton-Jacobi equation with initial condition S and f_t is the solution of the classical continuity equation

$$\partial_t f + \mathrm{div}\nabla S_t = 0\,.$$

The corresponding classical phase space distribution is therefore

$$\rho_{\mathrm{cl}}(dx\,dp) = f^2(x)\delta(p - \nabla S(x))\,dx\,dp\,. \tag{68}$$

The main drawback of the WKB approximation is that it works only as long as no caustics are reached, or, put differently, as long as no interference between different parts of the WKB wave function happens.

However, if we focus again on distributions of semi-classical observables on phase space no difficulty arises, only $\rho_{\mathrm{cl}}(t)$ is no longer of the particular form (68).

Proposition 20. Let $f \in C^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, $S \in C^\infty(\mathbb{R}^d)$ and

$$\phi(x) = \varepsilon^{d/2} f(\varepsilon x) e^{i\frac{S(\varepsilon x)}{\varepsilon}}$$

be normalized in $L^2(\mathbb{R}^d)$. Let $a \in C_0^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ and $T < \infty$. Then there is a $C < \infty$ such that for all $t \in [0, T]$

$$\left| \langle \widehat{\phi}, a_1^\varepsilon(t)\widehat{\phi}\rangle - \int \rho_{\mathrm{cl}}(dx\,dp, t) a(x,p) \right| \leq C\sqrt{\varepsilon}\,, \tag{69}$$

where $\rho_{\mathrm{cl}}(dx\,dp, t) = (\rho_{\mathrm{cl}} \circ \Phi^{-t})(dx\,dp)$ and $\rho_{\mathrm{cl}}(dx\,dp)$ was defined in (68).

As in the preceding cases we can again translate Proposition 20 to the full dynamics using Theorem 4, but we omit the corresponding statement this time.
Proof. We apply Egorov and switch to standard quantization:

$$\langle \widehat{\phi}, a_1^\varepsilon(t)\widehat{\phi}\rangle = \langle \widehat{\phi}, (a \circ \Phi^t)^{\mathrm{W},\varepsilon}\widehat{\phi}\rangle + O(\varepsilon) = \langle \widehat{\phi}, (a \circ \Phi^t)^{\mathrm{S},\varepsilon}\widehat{\phi}\rangle + O(\varepsilon)\,.$$

We calculate for $a \in C_0^\infty$

$$\begin{aligned}
(\widehat{a^{\mathrm{S},\varepsilon}\widehat{\phi}})(\tfrac{y}{\varepsilon}) &= (2\pi)^{-d}\varepsilon^{d/2}\int dx\, dp\, a(\varepsilon x,p)f(\varepsilon x)e^{-ip\cdot(x-y/\varepsilon)}e^{iS(\varepsilon x)/\varepsilon}\\
&= (2\pi)^{-d}\varepsilon^{-d/2}\int dx\, dp\, a(x,p)f(x)e^{i(S(x)-p\cdot(x-y))/\varepsilon}\\
&= (2\pi)^{-d}\varepsilon^{-d/2}\int dz\, dp\, a(z+y,p)f(z+y)e^{i(S(z+y)-p\cdot z)/\varepsilon}\,.
\end{aligned}$$

Stationary phase method (cf. Theorem 7.7.7 in [6] with $k = n+1$) yields that

$$\Big|\int dz\, dp\, a(z+y,p)f(z+y)e^{i(S(z+y)-p\cdot z)/\varepsilon} - (2\pi\varepsilon)^d a(y,\nabla S(y))f(y)e^{\frac{i}{\varepsilon}S(y)}\Big| \leq C\,\varepsilon^{d+\frac{1}{2}}\,,$$

where the constant is uniform in y and depends on a via a sum of sup-norms of finite many partial derivatives. Hence

$$((\widehat{a\circ\Phi^t)^{\mathrm{S},\varepsilon}\widehat{\phi}})(\tfrac{y}{\varepsilon}) = \varepsilon^{d/2}(a\circ\Phi^t)(y,\nabla S(y))f(y)e^{\frac{i}{\varepsilon}S(y)} + \varepsilon^{d/2}O(\sqrt{\varepsilon})$$

and

$$\langle\widehat{\phi},(a\circ\Phi^t)^{\mathrm{S},\varepsilon}\widehat{\phi}\rangle = \int dy\,(a\circ\Phi^t)(y,\nabla S(y))f^2(y) + O(\sqrt{\varepsilon})\,,$$

where we used that $f \in L^1(\mathbb{R}^d)$.

□

References

[1] J. Bolte and S. Keppeler, *A semiclassical approach to the Dirac equation*, Annals Phys. **274**, 125–162 (1999).

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

[3] Y.V. Egorov, *On canonical transformations of pseudodifferential operators*, Uspehi Math. Nauk **25**, 235–236 (1969).

[4] J. Fröhlich, *Existence of Dressed One Electron States in a Class of Persistent Models*, Fortschritte der Physik **22**, 159–198 (1974).

[5] G.A. Hagedorn, *A Time Dependent Born-Oppenheimer Approximation*, Comm. Math. Phys. **77**, 1–19 (1980).

[6] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Grundlehren der mathematischen Wissenschaften **256**, Springer (1983).

[7] F. Hövermann, H. Spohn and S. Teufel, *Semiclassical limit for the Schrödinger equation with a short scale periodic potential*, Comm. Math. Phys. (2000).

[8] R. Minlos, *Lower branch of the spectrum of a fermion interacting with a bosonic gas (polaron)*, Theor. Math. Phys. **92**, 869–877 (1993).

[9] E. Nelson, *Interaction of Nonrelativistic Particles with a Quantized Scalar Field*, Journal of Mathematical Physics, Vol. **5**, No. 9, 1190–1197 (1964).

[10] G. Nenciu, *Dynamics of band electrons in electric and magnetic fields: rigorous justification of the effective Hamiltonians*, Reviews of Modern Physics, Vol. **63**, No. 1, 91–127 (1991).

[11] M. Reed and B. Simon, *Methods of Modern Mathematical Physics 1, Functional Analysis*, Academic Press, San Diego (1972).

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

[13] H. Spohn, *Semiclassical Limit of the Dirac Equation and Spin Precession*, Annals Phys. **282**, 420–431 (2000).

[14] H. Spohn, *The polaron at large total momentum*, J. Phys. A **21**, 1199–1211 (1987).