COMPUTATIONAL ERROR ESTIMATES FOR BORN-OPPENHEIMER MOLECULAR DYNAMICS WITH NEARLY CROSSING POTENTIAL SURFACES

CHRISTIAN BAYER, HÅKON HOEL, ASHRAFUL KADIR, PETR PLECHÁČ, MATTIAS SANDBERG, AND ANDERS SZEPESSY

ABSTRACT. The difference of the values of observables for the time-independent Schrödinger equation, with matrix valued potentials, and the values of observables for ab initio Born-Oppenheimer molecular dynamics, of the ground state, depends on the probability to be in excited states and the electron/nuclei mass ratio. The paper first proves an error estimate (depending on the electron/nuclei mass ratio and the probability to be in excited states) for this difference of microcanonical observables, assuming that molecular dynamics space-time averages converge, with a rate related to the maximal Lyapunov exponent. The error estimate is uniform in the number of particles and the analysis does not assume a uniform lower bound on the spectral gap of the electron operator and consequently the probability to be in excited states can be large. A numerical method to determine the probability to be in excited states is then presented, based on Ehrenfest molecular dynamics and stability analysis of a perturbed eigenvalue problem.

CONTENTS

1. Motivation for error estimates in ab initio molecular dynamics
1.1. The Schrödinger and molecular dynamics models with an outline of the main results
2. Molecular dynamics approximation of Schrödinger observables for matrix-valued potentials
3. Determining the probability to be in excited states
3.1. Construction of WKB-solutions
3.2. The Landau-Zener model, transition probabilities and Ehrenfest dynamics
3.3. An estimate of the probability to be in excited states for matrix-valued potentials
4. Numerical examples
4.1. Model 1: A one dimensional problem
4.2. Model 2: A two dimensional problem
5. Regularity and Weyl calculus
6. Appendix: Fourier integral WKB states including caustics
6.1. A preparatory example with the simplest caustic
6.2. A general Fourier integral ansatz
7. Appendix: The stationary phase expansion
8. Appendix: An alternative motivation for assumption (2.17)
Acknowledgment
References

1. Motivation for error estimates in ab initio molecular dynamics

Molecular dynamics is a computational method to study molecular systems in materials science, chemistry and molecular biology. The simulations are used, for example, in designing and understanding new materials or for determining biochemical reactions in drug design [14].

The wide popularity of molecular dynamics simulations relies on the fact that in many cases it agrees very well with experiments. Indeed when we have experimental data it is often possible to verify the correctness of the method by comparing with experiments at certain parameter regimes. However, if we want the simulation to predict something that has no comparing experiment, we need a mathematical estimate of the accuracy of the computation. In the case of molecular systems with few particles, such studies are made by directly solving the Schrödinger equation. A fundamental and still open question

---

Key words and phrases. Landau-Zener probability, Egorov’s theorem, Born-Oppenheimer approximation, WKB expansion, microcanonical ensemble, Ehrenfest dynamics, ab initio molecular dynamics.
in classical molecular dynamics simulations is how to verify the accuracy computationally, i.e., when the solution of the Schrödinger equation is not a computational alternative.

The aim of this paper is to determine quantitative error estimates for molecular dynamics, including the case with nearly crossing electron potential surfaces that can yield large errors, without solving the Schrödinger equation but by combining mathematical stability analysis of eigenvalue problems with quantitative numerical Ehrenfest molecular dynamics computations of perturbations. Having molecular dynamics error estimates opens, for instance, the possibility of systematically evaluating which density functionals or empirical force fields are good approximations and under what conditions the approximation properties hold. Computations with such error estimates could also give improved understanding of when quantum effects are important and when they are not, in particular in cases when the Schrödinger equation is too computationally complex to solve.

The first step to check the accuracy of a molecular dynamics simulation is to know what to compare with. Here we compare the value of any observable \( g : \mathbb{R}^{3N} \to \mathbb{R} \), of nuclei positions \( X \), for the time-independent Schrödinger eigenvalue equation, \( \hat{H} \Phi = E \Phi \), with the corresponding molecular dynamics observable, defined by a Hamiltonian \( \hat{H} \), the wave function \( \Phi : \mathbb{R}^{3N} \to \mathbb{C}^J \) and the eigenvalue \( E \in \mathbb{R} \).

The approximation error we study is therefore the microcanonical setting

\[
\left( 1.1 \right) \quad \frac{\int_{\mathbb{R}^{3N}} \langle \Phi(X), g(X) \Phi(X) \rangle \, dX}{\int_{\mathbb{R}^{3N}} \langle \Phi(X), \Phi(X) \rangle \, dX} = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} g(X_t) \, dt,
\]

based on a molecular dynamics path \( X_t \), with total energy equal to the Schrödinger eigenvalue \( E \). The inner product \( \langle \cdot, \cdot \rangle \) is defined in \( \mathbb{C}^J \), where \( \Phi \in \mathbb{C}^J \) corresponds to \( J \) discrete (electron) states. The observable can for instance be the local potential energy, used in [50] to determine phase-field partial differential equations from molecular dynamics simulations. For the sake of simplicity we first explain the basic idea for observables depending only on position variables \( X \). The general case of position and momentum dependent observables is treated in Section 2 using the Weyl quantization. The time-independent Schrödinger equation has a remarkable property of accurately predicting experiments in combination with no unknown data, thereby forming the foundation of computational chemistry. However, the drawback is the high dimensional solution space for nuclei-electron systems with several particles, restricting numerical solution to small molecules. In this paper we study the time-independent setting of the Schrödinger equation as the reference. The proposed approach has the advantage of not requiring any initial data as input, on the other hand, an assumption on convergence rates of time averages of molecular dynamics observables is needed.

The second step to check the accuracy is to derive error estimates. We have three types of error: time discretization error, sampling error and modeling error. The time discretization error comes from approximating the differential equation for molecular dynamics with a numerical method, based on replacing time derivatives with difference quotients and time steps \( \Delta t \). The sampling error is due to truncating the infinite \( T \) and using a finite value of \( T \) in the integral in [1.1]. The modeling error (also called coarse-graining error) originates from eliminating the electrons in the Schrödinger nuclei-electron system and replacing the nuclei dynamics with their classical paths; this approximation error was first analyzed by Born and Oppenheimer in their seminal paper [4].

The time discretization and truncation error components are in some sense simpler to handle by comparing simulations with different choices of \( \Delta t \) and \( T \), although it can, of course, be difficult to know that the behavior does not change with even smaller \( \Delta t \) and larger \( T \) due to metastability, see [30, 7]. The modeling error is more difficult to check since a direct approach would require to solve the Schrödinger equation. Currently the Schrödinger partial differential equation can only be solved with few particles, therefore it is not an option to solve the Schrödinger equation in general. The reason for using molecular dynamics is precisely to avoid solving the Schrödinger equation. Consequently, the modeling error requires mathematical error analysis. Egorov’s theorem, cf. [39, 5], provides such error estimates and is used also here. However, in the literature there seems to be no error analysis that is precise, simple and constructive enough so that a molecular dynamics simulation can use it in practice to assess the modeling error also in the case when the electron operator has eigenvalues (i.e., potential surfaces) which are not well separated and may cause large modeling error. For instance, crossing eigenvalue surfaces can form so called conical intersections, which provide the mechanism for many chemical reactions, see [47, 51]. If the excited electron energy levels are well separated from the ground state energy, molecular dynamics based on the ground state energy is a good approximation, as first analyzed by Born and Oppenheimer [4]. On the other hand, in a quantum system with two electron energy levels that are not well separated, the dynamic transition probability from one state to the other state for a moving particle can be substantial, as determined by Landau and Zener [53] and generalized to a similar
dynamic case in two space dimensions in [27]. We denote the transition probability to go from the ground state to excited states for a time-dependent problem "dynamic transition probability", as in the Landau-Zener model, to distinguish it from the "probability to be in excited states" for the time-independent Schrödinger eigenvalue problem.

Our alternative error analysis presented here relates computable dynamic transition probabilities to the probability to be in excited states for the time-independent Schrödinger equation and is developed with the aim to give a different point of view that could help to construct algorithms that estimate the modeling error in molecular dynamics computations. Our analysis differs from previous ones by using, in three main steps:

- analyzing the time-independent Schrödinger equation as the reference model, including excited electron states with near crossing potential surfaces and the accuracy of observables as a function of the probability to be in excited states (in Section 2),
- studying stochastic molecular dynamics, constrained to the manifold of constant energy, based on small stochastic perturbations of the standard Born-Oppenheimer Hamiltonian dynamics,
- using stability analysis of a perturbed eigenvalue problem to estimate the probability to be in excited electron states, based on perturbations related to (Landau-Zener like) dynamic transition probabilities (in Section 3),
- applying Ehrenfest molecular dynamics to computationally estimate the dynamic transition probabilities (in Section 4).

The estimation method is tested on one and two dimensional problems but to conclude how useful the method is for realistic chemistry problems will require more work.

The next section introduces the Schrödinger and molecular dynamics models and ends with a more detailed outline.

1.1. The Schrödinger and molecular dynamics models with an outline of the main results.

In deriving the approximation of the solutions to the full Schrödinger equation, the heavy particles are often treated within classical mechanics, i.e., by defining the evolution of their positions and momenta by equations of motions of classical mechanics. Therefore we denote by $X_1 : [0, \infty) \rightarrow \mathbb{R}^{3N}$ the time-dependent function of the nuclei positions, with time derivatives denoted by

$$\dot{X}_i = \frac{dX_i}{dt}, \quad \ddot{X}_i = \frac{d^2X_i}{dt^2}. $$

We denote the Euclidean scalar product on $\mathbb{R}^{3N}$ by

$$X \cdot Y = \sum_{n=1}^{3N} X^n_i Y^n_i .$$

Furthermore, we use the notation $\nabla \psi(X) = (\nabla X_1 \psi(X), \ldots, \nabla X_N \psi(X))$, and as customary $\nabla X_n \psi = (\partial X_1 \psi, \partial X_2 \psi, \partial X_3 \psi)$. The notation $\psi(X) = \mathcal{O}(M^{-\alpha})$ is also used for complex valued functions, meaning that $|\psi(X)| = \mathcal{O}(M^{-\alpha})$ holds uniformly in $X$ for $\psi(X) \in \mathbb{C}^J$.

The time-independent Schrödinger equation

$$(1.2) \quad \hat{H}(X) \Phi(X) = E \Phi(X)$$

models many-body (nuclei-electron) quantum systems and is obtained from minimization of the energy in the solution space of wave functions, see [12, 11, 2, 16, 29, 0]. It is an eigenvalue problem for the energy $E \in \mathbb{R}$ of the system, described by the Hamiltonian operator $\hat{H}(X)$

$$(1.3) \quad \hat{H}(X) = V(X) - \frac{1}{2} M^{-1} I \sum_{n=1}^{N} \Delta X_n ,$$

where $I$ denotes the $J \times J$ identity matrix, and by the wave functions, $\Phi(X)$, belonging to a set of permissible electron states which we assume, for simplicity, to be finite, $\Phi(X) \in \mathbb{C}^J$. We use a normalized solution $\int_{\mathbb{R}^{3N}} |\Phi(X)|^2 dX = 1$ and without loss of generality we assume that all nuclei have the same mass $M$. If this is not the case, we can introduce new coordinates $M^{1/2} \hat{X}_k = \tilde{X}_k$, which transform the Hamiltonian to the form we want $V(M^{1/2} M^{-1/2} \hat{X}) - (2M) \int \sum_{k=1}^{N} \Delta \hat{X}_k \hat{X}$. In computational chemistry, the electron operator Hamiltonian $\hat{V}$, is independent of $M$ and it is precisely determined by the sum of the kinetic energy of electrons and the Coulomb interaction between nuclei and electrons. The wave function is then in a subspace of $L^2(\mathbb{R}^{3N} dX dx)$, with electron coordinates $x$, see [29]. By introducing a finite dimensional basis $\{\phi_i\}_{i=1}^{N}$ of the electron solution space, minimization of the energy $\int_{\mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} \Phi(X,x)^* V(X) \Phi(X,x) dX dx = 1$ leads to the
eigenvalue problem \([1.2]\) for the solution \(\Phi(X,x) = \sum_{j=1}^{J} \Phi_j(X) \phi_j(x)\), with the \(J \times J\) matrix \(V := S^{-1} \tilde{V}\) defined by the matrix components \(S_{ij} := \int_{\mathbb{R}^3} \phi_i(x) \phi_j(x)^* dx\) and \(\tilde{V}_{ij} := \int_{\mathbb{R}^3} V(X,x) \phi_i(x) \phi_j(x)^* dx\). To be able to compute with several electrons, i.e. \(n/3 \gg 1\), the electron solution space is often simplified in the form of Hartree-Fock or density functional approximations, which also lead to eigenvalue problems which now are non linear, see \([29]\). We assume that the electron operator \(V(X)\) is linear and self-adjoint, in a finite dimensional complex valued Euclidian space \(\mathbb{C}^J\) (for simplicity), with the usual inner product \(\langle \cdot, \cdot \rangle\), and acts as a matrix multiplication. An essential feature of the partial differential equation \([1.2]\) is the high computational complexity of finding the solution in a subset of the Sobolev space \(H^1(\mathbb{R}^{3N})\).

The mass of the nuclei, \(M\), are much greater than one (electron mass).

In contrast to the Schrödinger equation, a molecular dynamics model of \(N\) nuclei \(\tilde{X} : [0,\infty) \to \mathbb{R}^{3N}\), with a given potential \(V_p : \mathbb{R}^{3N} \to \mathbb{R}\), can be computationally studied for large \(N\) by solving the ordinary differential equations \(M \ddot{X}_r/dt^2 = -\nabla V_p(\tilde{X}_r)\) in the fast time scale. We will use the slow time scale \(t = M^{-1/2}\tau\) with positions \(X_t = \tilde{X}_\tau\) and scaled momenta \(P_t := \tilde{P}_\tau\) satisfying
\[
X_t = -\nabla V_p(X_t) \tag{1.4}
\]
In the slow time scale the nuclei move \(O(1)\) in unit time, since \(\dot{P}_t = -\nabla V_p(X_t)\). This computational and conceptual simplification motivates the study to determine the potential and its implied accuracy compared with the the Schrödinger equation, as started already in the 1920’s with the Born-Oppenheimer approximation \([4]\). The purpose of our work is to contribute to the current understanding of such derivations by showing convergence rates under new assumptions. The precise aim in this paper is to estimate the error
\[
\frac{\int_{\mathbb{R}^{3N}} \langle \Phi(X), g(X)\Phi(X) \rangle dX}{\int_{\mathbb{R}^{3N}} \langle \Phi(X), \Phi(X) \rangle dX} - \lim_{T \to \infty} \frac{1}{T} \int_0^T g(X_t) dt \tag{1.5}
\]
for scalar smooth observables \(g\) of the time-independent Schrödinger equation \([1.2]\) approximated by the corresponding molecular dynamics observable \(\lim_{T \to \infty} \frac{1}{T} \int_0^T g(X_t) dt\), which is computationally cheaper to evaluate in the case with more than one nuclei; Section \([2]\) includes also the general case with position and momentum dependent observables.

The main step to relate the Schrödinger wave function and the molecular dynamics solution is the so-called zero-order Born-Oppenheimer approximation, where \(X_t\) solves the classical \textit{ab initio} molecular dynamics \([4]\) with the potential \(V_p : \mathbb{R}^{3N} \to \mathbb{R}\) determined as an eigenvalue of the electron Hamiltonian \(V(X)\) for a given nuclei position \(X\). That is, \(V_p(X) = \lambda_0(X)\) and
\[
V(X)\Psi_0(X) = \lambda_0(X)\Psi_0(X) \tag{1.6}
\]
for a normalized electron eigenvector \(\Psi_0(X) \in \mathbb{C}^J\) (here the ground state). The initial data \((X_0,P_0)\) is chosen to have its energy equal to the eigenvalue \(E\), i.e. \(|P_0|^2/2 + \lambda_0(X_0)\). The Born-Oppenheimer expansion \([4]\) is an approximation of the solution of the time-independent Schrödinger equation which is shown in \([17,25]\) to solve the time-independent Schrödinger equation approximately when the electron operator \(V\) have eigenvalues \(\lambda_j(X)\) that are well separated, satisfying
\[
\min_X (\lambda_1(X) - \lambda_0(X)) > c \quad \text{for a positive constant } c \text{ independent of } M \tag{1.7}
\]
This expansion, analyzed by the methods of multiple scales, pseudo-differential operators and spectral analysis, e.g., in \([16,17,25,13]\), can be used to study the approximation error \([1.5]\). The work \([16,17,25,31]\) prove asymptotic expansions for the eigenfunction. In the literature one can also find precise statements on the error for the setting of the time-dependent Schrödinger equation, e.g., in \([31,5,35,26]\), which in some sense is easier since the stability issue is more subtle in the eigenvalue setting. The aim of our work is to present a method to analyze this stability of the eigenvalue problem, using the dynamic transition probability estimated from Ehrenfest dynamics simulations as quantitative input, without assuming an \(M\)-uniform lower bound \(c\) on the spectral gap.

We present error estimates for molecular dynamics approximation of time-independent observables in quantum mechanics, valid also when the eigenvalues \(\lambda_0\) of the electron operator \(V\) may nearly intersect and thereby increase the probability of being in excited states. For the Schrödinger solution \(\Phi\) to \([1.2]\), with the electron ground state \(\Psi_0\) satisfying \([1.6]\), the probability \(p_{ex}\) to be in excited states is
\[
p_{ex} := \int_{\mathbb{R}^{3N}} \langle \Phi^+(X), \Phi^+(X) \rangle dX \tag{1.8}
\]
where \(\Phi^+(X) := \Phi(X) - \langle \Psi_0(X), \Phi(X) \rangle \Psi_0(X)\). The derivation of the approximation error is divided into three steps which defines the outline of the paper:
I. The first step, in Section 2, proves that observables based on the time-independent Schrödinger equation can be approximated by stochastic molecular dynamics on the constant energy manifold. The convergence rate depends on the nuclei/electron mass ratio $M$ and the probability $p_{\text{ex}}$ to be in excited states. The proofs use Egorov’s theorem and assume that the expected value of space-time averages of the molecular dynamics observable converge in distributional sense with a rate related to the maximal Lyapunov exponent.

II. The second step, in Section 3, presents a numerical method to determine the probability to be in excited states from stability analysis of a perturbed eigenvalue problem. The perturbation is determined from a time-dependent excitation problem related to the Landau-Zener model (with $p_{\text{d}}$ denoting the dynamic transition probability to go to the excited state). Resonances of the eigenvalue problem then determines the larger probability $p_{\text{ex}}$ (to be in excited states for the Schrödinger eigenvalue problem) from the perturbations $p_{\text{d}}$ in the dynamic problem.

III. The final step, in Section 4, on numerical results, illustrate that the dynamic transition probability $p_{\text{d}}$ can be determined from Ehrenfest molecular dynamics simulations applied to two simple model problems.

Section 5 proves a lemma on the regularity of the expected value of the molecular dynamics solution and establishes in three other lemmas estimates on remainder terms in Moyal expansions used in the theorems. Section 6 formulates WKB solutions in the case with caustics.

The main ingredient in step I is Egorov’s theorem which has been used extensively to study semiclassical limits both in the time-dependent and time-independent case, cf. 39 8 14. We derive two estimates: Theorem 2.3 proves an estimate of a weighted difference of the Schrödinger and molecular dynamics observables and Theorem 2.2 derives an estimate of the difference of the Schrödinger and molecular dynamics observables, without the weight. The idea to assume ergodic classical dynamics to prove convergence of observables, in 14 28, initiated the activity on quantum ergodicity 52 and related results on semiclassical limits for scalar potentials, 20, and matrix potentials with distinct eigenvalues 8. The work 14 includes a recent review of semiclassical limit results in the case of well separated electron eigenvalue surfaces. In this case with well separated electron eigenvalues, satisfying 1.7, the probability to be in the excited state is proved to be small $p_{\text{ex}} = \mathcal{O}(M^{-1})$. We want to include the case when molecular dynamics simulations become inaccurate in practice. For an actual molecular dynamics simulation the mass $M$ is given and the electron potential $V$ does not depend in $M$, so a positive spectral gap $\gamma$ does not depend on $M$. However, if the mass is not large enough the molecular dynamics approximation may be inaccurate. Our analysis is asymptotic for large $M$ and to include the situation with a small gap for a certain mass we therefore consider a model when the spectral gap

\begin{align}
\min_X (\lambda_1(X) - \lambda_0(X)) =: \delta_M > 0,
\end{align}

can be close to zero, i.e. when it is not uniformly bounded from below with respect to $M$, and then $p_{\text{ex}}$ can be large as seen e.g. in the Landau-Zener probability 3.10 and Figure 4. There are four main new ingredients in our theorems compared to previous studies. It is well known that the standard proof of Egorov’s theorem, in the time-dependent case, combined with an assumption on the molecular dynamics convergence rate towards its ergodic limit yields an error estimate for the molecular dynamics observables as compared to the Schrödinger observables. The precise form of the assumption on the convergence to the ergodic limit is important for the conclusion of the approximation error compared to Schrödinger observables: (1) for instance, assuming an algebraic convergence rate $\mathcal{O}(T^{-1/2})$, for molecular dynamics time-averages of length $T$, would only give logarithmic error estimates $\mathcal{O}(1/\log^{1/2} M)$ (also in the case with a uniform spectral gap); (2) it is difficult to verify ergodicity for Hamiltonian dynamics, both theoretically and numerically. The first new ingredient here is to use stochastic dynamics constrained on the constant energy surface instead of Hamiltonian dynamics. Introducing small stochastic perturbations on the constant energy manifold have the advantage to guarantee ergodicity with respect to the microcanonical ensemble and to provide observables with higher regularity. Our second ingredient is to use a more precise convergence rate assumption based on space-time averages, which can be tested numerically (see Figure 10) and yields a better error estimate $\mathcal{O}(M^{-\gamma})$, for a positive $\gamma$. The third ingredient is to prove uniform convergence rates in the number of particles $N$ by careful use of localized mollifiers. Previous convergence proofs require for instance, for a constant $C$, the quantity $\max_{z \in \mathbb{R}^D, |z| \leq CN} |\partial_\mu g(z)| M^{-1}$ to be small but typically this is a large number for systems with many particles unless the mass ratio $M$ is very large and increasing with the number of particles $N$, see 20 and Theorem 15.4 in 25. The fourth ingredient is that we include non uniformly spectral gap bounds 1.9. Theorems 2.1 and 2.2 prove the error estimate in step I in terms of the probability $p_{\text{ex}}$ to be in excited states, which is not a typical a priori information. Consequently the computational
approximation of $p_{ex}$ in steps II and III is an important complement. Although to evaluate the practical value of the suggested method to estimate the computational error in molecular dynamics these two steps need to be tested on more realistic chemistry problems.

The alternative to consider molecular dynamics at constant temperature using the canonical ensemble is important in practice and it also has the advantage of proved ergodic limits. The canonical ensemble is in a sense the average of the microcanonical ensemble over all energies weighted with the Gibbs density. Therefore the results here are useful for ongoing analysis in the constant temperature setting.

2. Molecular dynamics approximation of Schrödinger observables for matrix-valued potentials

For every mass $M \geq 1$, we consider $L^2(\mathbb{R}^{3N})$ wave functions $\Phi : \mathbb{R}^{3N} \to \mathbb{R}^J$ and eigenvalues $E \in \mathbb{R}$, solving the Schrödinger eigenvalue equation

$$\hat{H} \Phi = E \Phi,$$

where

$$\hat{H} := -\frac{1}{2M} \Delta x + V(x).$$

Here, the symmetric smooth matrix valued potential $V : \mathbb{R}^{3N} \to \mathbb{R}^{J^2}$ has (electron) eigenvalues $\{\lambda_n\}_{n=0}^{J-1}$ ordered increasingly, with normalized eigenvectors $\{\Psi_n\}_{n=0}^{J-1}$ satisfying

$$V(x) \Psi_n(x) = \lambda_n(x) \Psi_n(x)$$

and the spectral gap condition [1,9]. The eigenvalues $\{\lambda_n, n \geq 1\}$ may be degenerate but $\lambda_0$ is not. Both $\Phi$ and $E$ depend on the mass $M$ and we will only consider those eigenfunctions where the corresponding eigenvalue is bounded, i.e., the $\Phi$ for which $E = O(1)$ as $M \to \infty$. That is, we assume that for every $M \geq 1$ there are solutions $\Phi = \Phi_E \in L^2(\mathbb{R}^{3N})$ with eigenvalue $E = O(1)$ to (2.1), $M \to \infty$. On the other hand, the matrix $V$ does not depend on $M$ explicitly and consequently also the eigenvectors and eigenvalues $\Psi_n$ and $\lambda_n$ do not depend on $M$ explicitly. However to include the case of a small spectral gap, we assume that the positive spectral gap, $\delta$, in (1.9) is not uniform in $M$ and we study estimates depending on the two parameters $M$ and $\delta$. To handle ergodicity for the Born-Oppenheimer dynamics, we will below assume that the smallest eigenvalue $\lambda_0(X)$ is smooth and tends to infinity as $|X| \to \infty$.

This assumption in fact also implies that the spectrum of $\hat{H}$ is discrete, see [9].

We consider a given smooth scalar observable $g : \mathbb{R}^{6N} \to \mathbb{R}$ in the Schwartz space constructed as

$$g = \hat{g} * \phi_\eta,$$

where for some $\eta > 4M^{-1/2}$ the mollifier $\phi_\eta : \mathbb{R}^{6N} \to \mathbb{R}$ is defined by $\phi_\eta(z) := (2\pi \eta)^{-3N} e^{-|z|^2/(2\eta)}$ and $\hat{g} : \mathbb{R}^{6N} \to \mathbb{R}$ is smooth and compactly supported. We assume that each solution $\Phi = \Phi_E$ to (2.1) is normalized, i.e. $\int_{\mathbb{R}^{3N}} \langle \Phi(X), \Phi(X) \rangle dX = 1$, and define for each $\Phi$ the observable

$$g_\Phi := \int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{g} \Phi(X) \rangle dX,$$

based on the Weyl quantization, see [32],

$$\hat{g} \Phi(X) = \mbox{Op}[g] \Phi(X) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} e^{iM^{1/2}(X-Y) \cdot P} g \left( \frac{X+Y}{2}, P \right) \Phi(Y) dY dP,$$

using the inner products

$$\langle w, v \rangle := \sum_{j=1}^J w_j^* v_j \quad \text{for } w, v \in \mathbb{C}^J, \quad X \cdot P := \sum_{n=1}^{3N} X^n P^n \quad \text{for } X, P \in \mathbb{R}^{3N}.$$

We note that in the case $g(X, P) = g(X)$ is a function of $X$ only, then the Weyl quantization acts as a multiplication operator, i.e. $\hat{g} \Phi(X) = g(X) \Phi(X)$. We also define the Wigner transform

$$W(X, P) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} e^{iM^{1/2} Y \cdot P} \left( \Phi \left( X + \frac{Y}{2} \right), \Phi \left( X - \frac{Y}{2} \right) \right) dY,$$

and

$$W_{jk}(X, P) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} e^{iM^{1/2} Y \cdot P} \Phi_j^* \left( X + \frac{Y}{2} \right) \Phi_k \left( X - \frac{Y}{2} \right) dY,$$
and note that Φ = Φ_E is different for different eigenvalues E. The change of variables
\[ X' = X + \frac{Y}{2}, \quad Y' = X - \frac{Y}{2} \]
shows that for any \( g : \mathbb{R}^{6N} \to \mathbb{R} \) and \( A : \mathbb{R}^{6N} \to \mathbb{R}^{J^2} \)
\[ \int_{\mathbb{R}^{6N}} (\Phi(X), \dot{g}(X))dX = \int_{\mathbb{R}^{6N}} g(X, P)W(X, P)dXdP, \]
\[ \int_{\mathbb{R}^{6N}} \langle \Phi(X), \dot{A}(X) \rangle dX = \int_{\mathbb{R}^{6N}} \sum_{j=1}^{J} \sum_{k=1}^{J} A_{jk}(X, P)W_{jk}(X, P)dXdP. \]

Our estimates of remainder terms use the corresponding function \( W(s) \) defined for \( s \in [0, 1] \) by
\[ W(s)(X, P) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{2N}} e^{iM^{1/2}Y \cdot P} \langle \Phi(X + sY), \Phi(X - (1-s)Y) \rangle dY, \]
which is the "Wigner"-function corresponding to the alternative quantization
\[ (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{2N}} \langle \Phi(X), g(X + s(Y - X), P)\Phi(Y) \rangle e^{iM^{1/2}(X-Y) \cdot P}dYdXdP \]
\[ = \int_{\mathbb{R}^{6N}} g(X, P)W(s)(X, P)dXdP. \]

We will approximate the Schrödinger observable by dynamics related to ab initio molecular dynamics
\[ \dot{X}_t = P_t, \]
\[ \dot{P}_t = -\nabla_X \lambda_0(X_t), \]
based on the Hamiltonian
\[ H_0(X, P) := \frac{|P|^2}{2} + \lambda_0(X). \]
We assume that \( \lambda_0 \) is smooth and coercive in the sense that
\[ \lim_{|X| \to +\infty} \lambda_0(X) = +\infty. \]

In order to obtain precise estimates of the approximation error which are uniform in \( N \) we consider a mollified Hamiltonian
\[ H_\eta(X, P) := \frac{|P|^2}{2} + \lambda_\eta(X), \]
\[ \lambda_\eta(X) := \lambda_0 * \phi_\eta(X), \]
where \( \phi_\eta(X) = (2\pi \eta)^{-3N/2}e^{-|X|^2/(2\eta)} \) is a standard mollifier on the small scale \( \sqrt{\eta} \). To obtain good approximation of observables compared to the Schrödinger case we use \( \eta = 4M^{-1/2} \). To ensure ergodic molecular dynamics and still approximate the Born-Oppenheimer dynamics we will compare the Schrödinger observable to the observable for stochastic dynamics restricted to the manifold
\[ \Sigma_E := \{ (X, P) \in \mathbb{R}^{6N} \mid H_\eta(X, P) = E \}, \]
for the constant energy \( E \) equal to the Schrödinger eigenvalue, with a small stochastic perturbation of \[\Sigma_E.\]

Let \( Z = (X, P) \in \mathbb{R}^{6N} \) denote the phase space variable, then one example of such stochastic Stratonovich dynamics on \( \Sigma_E \) takes the form
\[ dz_t = J\nabla H_\eta(Z_t)dt + \sqrt{2\epsilon\tau} \mathbb{P}(Z_t) d\tilde{W}_t, \]
where \( \epsilon < M^{-1/2} \) is a small positive parameter, \( \tilde{W} \) is the standard Wiener process in \( \mathbb{R}^{6N} \) with independent components, the skew symmetric matrix \( J \) is defined by \( J\nabla Z H_\eta(X, P) = (P, -\nabla_X \lambda_\eta(X)) \), the parameter \( \tau \sim 1 \) is positive, and the projection \( \mathbb{P}(Z) \) onto the tangent space of \( \Sigma_E \) at \( Z \in \Sigma_E \) reads
\[ \mathbb{P}(Z) := \text{Id} - \hat{n}(Z) \otimes \hat{n}(Z) \]
with the normal \( \hat{n}(Z) := \nabla H_\eta(Z)/|\nabla H_\eta(Z)| \) to \( \Sigma_E \). We assume that \( \nabla H_\eta|_{\Sigma_E} \neq 0 \) for all \( \epsilon \in \mathbb{R} \). Since also \( H_\eta \) is assumed to be smooth and coercive, the set \( \Sigma_E \) is a smooth compact codimension one manifold in \( \mathbb{R}^{6N} \) for every \( \epsilon \in \mathbb{R} \). The dynamics \[\Sigma_E \text{ is the projection of the Ito differential equation} \]
\[ dz_t = (J\nabla H_\eta(z_t) - \epsilon \nabla H_\eta(z_t))dt + \sqrt{2\epsilon\tau} d\tilde{W}_t \]
to \( \Sigma_E \) and \( z_t \) has a unique equilibrium measure, which is the Gibbs measure
\[ e^{-H_\eta(z)}/\tau dz / \int_{\mathbb{R}^{6N}} e^{-H_\eta(z)}/\tau dz. \]
The diffusion parameter \( \epsilon \) can, for instance, also be a semi-positive definite constant diagonal matrix, so that \( (2.11) \) includes the Langevin equation. The work \cite{12} proves that the projected dynamics \( (2.10) \) is also ergodic and the equilibrium measure is the microcanonical measure

\[
dv(Z) := \left( \int_{\Sigma_E} \frac{d\Sigma}{|\nabla H_q(z)|} \right)^{-1} \frac{d\Sigma(Z)}{|\nabla H_q(Z)|},
\]

i.e. \( \lim_{T \to \infty} T^{-1} \int_0^T g(Z_t)dt = \int_{\Sigma_E} g(z)dv(z) \) where \( d\Sigma \) is the surface measure on \( \Sigma_E \) induced by the Lebesgue measure in \( \mathbb{R}^{6N} \). Alternative ergodic dynamics on \( \Sigma_E \) sampling the microcanonical measure are presented in \cite{12}.

The Ito dynamics corresponding to \( (2.10) \) reads, see \cite{12} and \cite{8},

\[
dZ_t = \mathbb{P}(Z_t)(J \nabla H_q(Z_t)) dt - \epsilon \nabla \kappa(Z_t)\hat{n}(Z_t)dt + \sqrt{2\epsilon} \mathbb{P}(Z_t)d\tilde{W}_t,
\]

where \( \kappa(z) := \text{div} \hat{n}(z) \) is the mean curvature at \( z \) on \( \Sigma_E \). The Kolmogorov equation for the expected value \( u(t,z) := \mathbb{E}[g(Z_T) \mid Z_t = z] \), for \( t \leq T \), of paths \( Z \) satisfying \( (2.10) \) becomes

\[
\partial_t u(t,z) + (J \nabla H_q(z) - \epsilon \nabla \kappa(z)\hat{n}(z)) \cdot \nabla_z u(t,z) + \epsilon \tau \text{trace}(\mathbb{P}(z)\nabla^2 u(t,z)) = 0 \quad t < T,
\]

\[
\tau = \left[ L_\omega(t,z) \right] = g,
\]

where \( \nabla^2 u(t,Z) \) denotes the Hessian with respect to \( z \), so that

\[
\text{trace}(\mathbb{P}(z)\nabla^2 u(t,z)) = \sum_{i,j} \mathbb{P}_{ij}(z) \partial_{i,z} \partial_{j,z} u(t,z) \text{ and } \mathbb{P}_{ij} = \delta_{ij} - \hat{n}_i \hat{n}_j.
\]

This Kolmogorov equation has an intrinsic definition based only on the coordinates on \( \Sigma_E \), see \cite{12}. The expected value \( u \) is smooth and the compact support of \( g \) implies that \( u \) decays rapidly as \( |Z| \to \infty \), so that \( u \) is in the Schwartz class. The work \cite{12} also proves an exponential convergence rate to equilibrium: let the initial data \( g \) be smooth in a neighborhood of \( \Sigma_E \) so that \( (2.11) \) includes the Langevin equation. The work \cite{12} proves that the projected dynamics \( (2.10) \) depends on the observable \( \kappa \) depending on the initial energy \( E \).

We write \( S_t(Z) := (Z_t \mid Z_s = Z) \), for \( t > s \) be the stochastic flow of the dynamics \( (2.10) \), i.e. the solution \( Z_t \) that starts in \( Z \) at time \( s \), for a realization \( \tilde{W}(\cdot,\omega) \) of the Wiener process. Then, the expected value \( u(0,Z) = \mathbb{E}[g \circ S_t(Z)] \). This function satisfies the exponential growth

\[
\sup_{Z \in \mathbb{R}^{6N}} \left( \sum_{|\beta| \leq n} |\partial^\beta_{Z_t} \mathbb{E}[g \circ S_t(Z)]|^2 \right)^{1/2} \leq e^{\tilde{C}(t+1)} \delta^{\text{min}(0,−n+1)}, \quad n \leq 4,
\]

where \( \tilde{C} \) is independent of \( M, \delta \) and \( \epsilon \), as derived in Lemma \ref{lemma:5.1} under assumption \ref{assumption:5.3} including weak near crossing of eigenvalues.

Define the molecular dynamics microcanonical observable

\[
g_{\text{MD}}(Z_0) := \lim_{T \to \infty} \frac{1}{T} \int_0^T g(S_t(Z_0))dt = \int_{\Sigma_E} g(z)dv(z),
\]

depending on the initial energy \( E = H_q(Z_0) \). Our main assumption on the dynamics is that we assume that for some \( \epsilon < M^{-1/2} \) there is a positive constant \( \gamma \), independent of \( \epsilon \) and \( M \), such that

\[
\int_{\mathbb{R}^N} \frac{2}{T} \int_{T/2}^T \mathbb{E}[g \circ S_s(Z_0) - g_{\text{MD}}(Z_0)]^2 ds W(Z_0)dZ_0 \leq e^{-\gamma T} + O(M^{-1}), \quad \text{for } T \leq \log M/(\tilde{C} + \gamma).
\]

The constant \( \tilde{C} \) depends on the observable \( g \) so that an observable related to a macroscopic quantity may have better convergence rate in the case of a large system.
From (2.14), we know that there is a constant $\gamma \sim \epsilon$ satisfying (2.17). Assumption (2.17) says that the decay with respect to time $T$ of

$$\frac{2}{T} \int_0^{T/2} \left( u(t, Z_0) - \int_{\mathbb{R}^N} u(T, \cdot) d\nu \right) dt$$

integrated over the Wigner distribution is exponential with the decay rate $\gamma \sim 1$ (for some $\epsilon < M^{-1}$ and $T < \log M/(\bar{C} + \gamma)$) plus possibly an additional constant small term $M^{-1}$ independent of $t$. The ergodic assumption

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T g \circ S_0(Z_0) dt = \int_{\mathbb{R}^N} g(z) d\nu(z),$$

for the Hamiltonian dynamics with $\epsilon = 0$ is used in [33, 28, 20] and to prove convergence of Schrödinger observables. This ergodic assumption is difficult to verify numerically, since a numerical approximation of the dynamics will always perturb the dynamics and it is theoretically unclear how this perturbation effects the dynamics over infinite time, see [49]. To find a method to prove ergodicity for general Hamiltonian systems also remains a challenge, cf. [49]. An advantage with assumption (2.17) is that it can be tested numerically for some initial points $Z_0$, see Figure 11 since the time discretization error can be made small compared to the bound $O(M^{-1})$, by taking a small time step $\Delta t < M^{-1}$, and the simulation time $T < \log M/\bar{C}$ is finite. Not all dynamics satisfy (2.17). For instance there exists billiard dynamics in “stadium” domains in $\mathbb{R}^2$ that is proven to be ergodic for almost all initial data but is not ergodic for some data with corresponding concentrated Schrödinger eigenfunctions, see [19]. Small perturbations of this data with non ergodic dynamics will need very long time to reach asymptotically ergodic behavior and consequently also the stochastic regularization above is likely to have $\gamma$ in (2.17) depending on $\epsilon$ so that our assumption would not hold.

Having introduced the necessary notation and terminology, we are now ready to present an estimate for the weighted difference between the Schrödinger and stochastic molecular dynamics observables that is expressed in terms of the spectral gap, the mass and the excitation probability. Theorem 2.2 presented in the end of this section, then estimates the difference $g_s - g_{MD}$ of the observables without the weight.

**Theorem 2.1.** For every $M \gg 1$ consider the Schrödinger eigenvalue problem (2.1)–(2.2) with a solution $\Phi \in L^2(\mathbb{R}^{3N})$, for a corresponding eigenvalue $E = O(1)$, and assume:

(i) the observable $g$ satisfies (2.3) and the potential $V$ is smooth with

$$\sup_{X \in \mathbb{R}^{3N}} \sum_{1 \leq |\alpha| \leq 3} |\partial^\alpha_X V(X)|^2$$

bounded, uniformly in $N$, where $| \cdot |$ is the matrix $2$-norm in $\mathbb{C}^J$,

(ii) the molecular dynamics limit $g_{MD}$ in (2.16) satisfies the weak convergence rate (2.17),

(iii) the electron eigenvector $\Psi_0$ and eigenvalue $\lambda_0$ are smooth and have bounded derivatives satisfying, for some $\delta = \delta_M > 0$ defined in (1.9) (which may depend on $M$) and any $|\beta| \leq 2$, and for any $|\alpha| \leq 4$,

$$|\partial^\alpha_X \Psi_0(X)| = O(\delta^{-|\beta|}),$$

$$|\partial^\alpha_X \lambda_0(X)| = O(\delta^{-|\alpha|+1}),$$

(iv) the Hamiltonian $H_q$ is smooth, satisfies the coercivity (2.9) and $\nabla H_q|_{X_0} \neq 0$ for all $e \in \mathbb{R}$, and the exponential bound (2.15) holds with the constant $\bar{C}$ uniformly bounded in $N$.

Let the probability to be in excited states be denoted by $p_{ex} = p_{ex}(\Phi)$, as defined in (1.8), and define the Wigner transform $W = W(\Phi)$ by (2.5). Then, for $M^{-1/2} \leq \epsilon = O(1)$, the molecular dynamics observable $g_{MD}$ approximates the Schrödinger observable $g_s$, defined in (2.4), with the weighted error estimate

$$(2.18) \quad \int_{\mathbb{R}^N} \left( g_s - g_{MD}(Z_0) \right) W(Z_0) dZ_0 \leq C\left( e^{\epsilon T} (M^{-1/2} + C_N \epsilon^{-1} + \delta^{-2} + p_{ex}^{1/2}) + e^{-\gamma T} \right),$$

for any $T > 0$, and by choosing $T$ such that $\bar{C} = (M^{-1/2} \delta^{-2} + p_{ex}^{1/2}), \eta = 4M^{-1/2}$ and $\epsilon$ sufficiently small

$$(2.19) \quad \int_{\mathbb{R}^N} (g_s - g_{MD}(Z_0)) W(Z_0) dZ_0 \leq C(M^{-1/2} \delta^{-2} + p_{ex}^{1/2}) \nu_{\mathbb{R}^N},$$

for a constant $C$, independent of $M \gg 1$, $\eta$, $\bar{C}$, $\gamma$, $\delta$ and $M$.

To prove this theorem, we modify the proof of Egorov’s theorem [39] to include stochastic dynamics and the assumption (2.17) on the space-time convergence rate and to replace the uniform spectral gap bound for matrix valued perturbation potentials $V(X)$, in [3] [14], with Assumption (iii). Assumption (iii) with $\delta > 0$ excludes a conical intersection but avoided crossings are allowed with a spectral gap in [19] that is not uniformly bounded from below as a function of $M$. The Landau-Zener probability (3.10)
in Section 3.2 (combined with 3.12 and 3.42) illustrate for instance that $p_{\text{ex}}$ can be close to one for an avoided crossing, with a small spectral gap $\min_X (\lambda_1(X) - \lambda_0(X)) = \delta = M^{-1/4}$. That $p_{\text{ex}}^{1/2}$ can be of the same order as $M^{-1} \delta^{-4}$ in the estimate (2.18) motivates the study of the approximation error as a function of $p_{\text{ex}}$ and $M$. To have $p_{\text{ex}}$ large seems to require smaller gaps $\delta \lesssim M^{1/4}$, see Section 3. It would be desirable to have analytic estimates of $p_{\text{ex}}$, since it is determined by $M$ and $V$. Such estimates are known, $p_{\text{ex}} = o(1)$ as $M \to \infty$, for the case when spectral gap $\delta > c > 0$ is uniformly bounded from below while in the case with no uniform spectral gap the required smoothness of projection to the ground state is assumed, see [48]. The lack of analytic estimates of $p_{\text{ex}}$ is our motivation for using $p_{\text{ex}}$ in addition to $M$ as parameter in the Theorems and presenting the computational method to determine $p_{\text{ex}}$ in this work.

We note that stochastic dynamics has a regularizing effect since $\mathbb{E}[g \circ S_t]$, with $\epsilon > 0$, is in general more regular than $g \circ S_t$, with $\epsilon = 0$. In fact, we do not explicitly need the stochastic flow $S$: the proof only uses the deterministic value $u(t, \cdot) = \mathbb{E}[g \circ ST_t(\cdot)]$ and its flow $u(t, \cdot) =: S_T u(T, \cdot)$, for $t < T$, with the generator $\partial_t u(t, \cdot) = -Lu(t, \cdot)$. The exponential bound in assumption (v) is proved in Lemma 5.1 under an assumption that allows weak near crossing of eigenvalues away from regions where $P$ vanishes.

**Proof.** Consider the solution operator $e^{iM^{1/2}t\hat{H}}$ of the time-dependent Schrödinger equation

$$iM^{-1/2}\partial_t \Psi(t, X) = \hat{H} \Psi(t, X)$$

defined by $\Psi(t, \cdot) = e^{-iM^{1/2}t\hat{H}} \Psi(0, \cdot)$. We will compare the observable

$$g_S = \int_{\mathbb{R}^N} \langle e^{-iM^{1/2}t\hat{H}} \Phi(X), \bar{g} e^{-iM^{1/2}t\hat{H}} \Phi(X) \rangle dX$$

$$= \int_{\mathbb{R}^N} \langle \Phi(X), e^{iM^{1/2}t\hat{H}} \bar{g} e^{-iM^{1/2}t\hat{H}} \Phi(X) \rangle dX$$

to the expected value

$$\frac{2}{T} \int_{T/2}^T \int_{\mathbb{R}^N} \mathbb{E}[g \circ S_t(Z_0)] W(Z_0) dZ_0 dt = \frac{2}{T} \int_{T/2}^T \int_{\mathbb{R}^N} \langle \Phi(X_0), \mathbb{E}[g \circ S_t] \Phi(X_0) \rangle dX_0 dt.$$ 

By construction we have that $S_t(Z_0)$ is in the compact manifold $\Sigma_{H_0}(Z_0)$. The coercivity (2.9) of $\lambda_0$ and the decay of $g$ in phase space imply therefore that $g(S_t(Z_0))$ decays sufficiently fast for large $|Z_0|$. The integrals above are therefore well defined. Since $V$ is smooth, elliptic regularity implies that the solution $\Phi$ to the Schrödinger eigenvalue problem also is smooth. Let $\bar{u}$ be the solution to

$$\partial_t \bar{u}(t, z) - L \bar{u}(t, z) = 0 \quad t > 0$$

$$\bar{u}(0) = \bar{g},$$

(2.20)
then $\mathbb{E}[\tilde{g} \circ S_t(Z_0)] = \tilde{u}(t, Z_0)$ and $g = \tilde{g} * \phi_\eta$. We have, using the commutator $[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}$ and that $\hat{H}\Phi = E\Phi$,

\[(2.21)\]

\[
\int_{\mathbb{R}^N} \tilde{u}(t, \cdot) * \phi_\eta(Z)W(Z) dZ - g_\eta = \int_{\mathbb{R}^N} \langle \Phi(X), \tilde{u}(t, \cdot) * \phi_\eta - e^{iM^{1/2}t\hat{H}} \tilde{u}(0, \cdot) * \phi_\eta e^{-iM^{1/2}t\hat{H}} \rangle \Phi(X) dX
\]

\[
= \int_{\mathbb{R}^N} \langle \Phi(X), \int_0^t \frac{\partial}{\partial s} \left( e^{iM^{1/2}(t-s)\hat{H}} \tilde{u}(s, \cdot) * \phi_\eta e^{-iM^{1/2}(t-s)\hat{H}} \right) ds \Phi(X) dX
\]

\[
= \int_{\mathbb{R}^N} \langle \Phi(X), \int_0^t e^{iM^{1/2}(t-s)\hat{H}} \left( \frac{\partial}{\partial s} \text{Op}[\tilde{u}(s, \cdot) * \phi_\eta] - iM^{1/2} \left[ \hat{H}, \tilde{u}(s, \cdot) * \phi_\eta \right] \right) e^{-iM^{1/2}(t-s)\hat{H}} \Phi(X) dX ds
\]

\[
= \int_0^t \int_{\mathbb{R}^N} \langle \Phi(X), \text{Op}\left[ \left( \frac{(P \cdot \nabla X - \nabla \lambda_0 \cdot \nabla P) \tilde{u}(s, \cdot) * \phi_\eta}{H} \right) \right] e^{-iM^{1/2}(t-s)\hat{H}} \Phi(X) dX dS
\]

In the fourth equality we have used (2.20) to conclude that
\[
\frac{\partial}{\partial s} \text{Op}[\tilde{u}(s, \cdot) * \phi_\eta(Z)] = \text{Op}\left( \frac{\partial}{\partial s} \tilde{u}(s, \cdot) * \phi_\eta(Z) \right)
\]

\[
= \text{Op}\left( [L \tilde{u}(s, \cdot)] * \phi_\eta(Z) \right)
\]

\[
= \text{Op}\left( \left( J \nabla_Z H_0 - \frac{\epsilon \tau K}{\nabla_Z H_0} \nabla Z \Phi(\eta) \right) \tilde{u}(s, \cdot) * \phi_\eta \right)
\]

\[
= \text{Op}\left( [J \nabla_Z H_0 - \nabla Z \tilde{u}(\cdot)] * \phi_\eta(Z) \right) + \mathcal{O}(C_N \epsilon^\delta - e^{\epsilon C t})
\]

Here the estimate
\[(2.22)\]

\[
\left| \int_{\mathbb{R}^N} \langle \Phi(X), \text{Op}\left( \frac{-\epsilon \tau K}{\nabla_Z H_0} \nabla Z + \epsilon \tau \text{trace}(P \nabla_Z^2) \right) \tilde{u}(s, \cdot) * \phi_\eta \right| dX \right| \leq C_N \epsilon \delta - e^{\epsilon C t}
\]

follows by Lemma 5.4 and we note that the term is negligible small by choosing $\epsilon$ sufficiently small, although the constant $C_N$ typically is large as it depends on order $N$ derivatives of the symbol.

To estimate terms in the error representation (2.21) is now the remaining four steps in the proof.

1. The Moyal expansion for the commutator of two Weyl operators in \[\text{[25]}\] and Lemma 5.5 is stated for scalars and here we apply Lemma 5.5 to each matrix component $j$ of $\hat{H}$ to obtain

\[
\text{[23]}\]

\[
\left( \text{Op}[\nabla P H_k(X, P) \cdot \nabla X - \nabla X H_k(X, P) \cdot \nabla P] \left( \tilde{u}(s, \cdot) * \phi_\eta \right)(X, P) + \tilde{R}_M \right) \Phi_k(X)
\]

for the matrix components

\[
H_k(X, P) = \frac{|P|^2}{2} \delta_{jk} + V_{jk}(X)
\]

where the remainder takes the form, see \[\text{[5]}\] and Section 5

\[(2.24)\]

\[
R_M := \text{Op}\left[ \sum_{n=1}^m 2M^{-n} (2r)^{-n} \sum_{|\alpha| = 2n+1} \frac{(-1)^{|\alpha|}}{\alpha!} \partial_\alpha^2 V_{jk}(X) \left( \partial_\alpha^2 \tilde{u}(s, \cdot) * \phi_\eta \right)(X, P) + M^{-(m+1)} r_m \right]
\]

and $r_m$ is smooth. We will use $\tilde{R}_M$ for $m = 1$. Lemma 5.5 in Section 5 shows that

\[(2.25)\]

\[
\int_{\mathbb{R}^N} \langle \Phi, \tilde{R}_M \Phi \rangle = \mathcal{O}(\epsilon C_T \delta^{-2} M^{-1}).
\]
2. The normalization property \(1 = \int_{R^N} W(Z) dZ\) implies that
\[ g_S = \int_{R^N} g_S W(Z) dZ \]
and we obtain by (2.21) and (2.23)
\[
\int_{R^N} \left( g_S - \frac{2}{T} \int_{T/2}^T \bar{u}(t, \cdot) \ast \phi_\eta(Z) dt \right) W(Z) dZ \\
= \frac{2}{T} \int_{T/2}^T \left( g_S - \int_{R^N} \bar{u}(t, \cdot) \ast \phi_\eta(Z) \right) W(Z) dZ dt \\
= \frac{2}{T} \int_{T/2}^T \int_{R^N} (\Phi(X), Op\left( \left( \nabla \lambda_\eta(\cdot) I - \nabla X V(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right) \Phi(X)) dX ds dt \\
= \frac{2}{T} \int_{T/2}^T \int_{R^N} (\Phi(X), (\Omega(\epsilon) + \hat{R}_M) \Phi(X)) dX ds dt,
\]
where we use the bounds (2.25) and (2.22) in last term.

3. We have the desired quantity \(\int_{R^N} (g_S - g_{MD}) W dX dP\) in the left hand side above by adding and subtracting the molecular dynamics observable \(g_{MD}\), to the second term in the left hand side, and using its convergence rate \(e^{-\gamma T} + M^{-1}\) from (2.17) as follows
\[
\int_{R^N} \frac{2}{T} \int_{T/2}^T \bar{u}(t, \cdot) \ast \phi_\eta(Z) dt W(Z) dZ = \int_{R^N} g_{MD}(Z) W(Z) dZ \\
+ \int_{R^N} \frac{2}{T} \int_{T/2}^T \left( \bar{u}(t, \cdot) \ast \phi_\eta(Z) - g_{MD}(Z) \right) dt W(Z) dZ \\
= \int_{R^N} g_{MD}(Z) W(Z) dZ + O(e^{-\gamma T} + M^{-1}),
\]
which yields the last term in the error bound (2.18).

4. It remains to estimate the first term in the right hand side of (2.26). The rule for the composition of Weyl quantizations, cf. [32] and Lemma 5.5 in Section 5
\[
(2.27) \quad \hat{A} \hat{B} = \hat{A} \hat{B} - i M^{-1/2} Op \left\{ A, B \right\} + O(M^{-1}),
\]
for smooth scalar functions \(A, B : \mathbb{R}^6 \rightarrow \mathbb{R}\). The first term in the right hand side of (2.26) can be written
\[
\int_{R^N} \left( \Phi(X), Op\left( \left( \nabla \lambda_\eta(\cdot) I - \nabla X V(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right) \Phi(X) \right) dX
\]
and to estimate this term we use (1.8) and write \(\Phi = \Phi^\perp + \Pi_0 \Phi\), where \(\|\Phi^\perp\|_{L^2(dX)} = \frac{1}{\rho_{cs}}\) and \(\Pi_0 w := \langle w, \Psi_0 \rangle \Psi_0\), for \(w \in \mathbb{R}^J\), defines the projection onto the ground state. The integral terms including the factor \(\Phi^\perp\) is bounded using Lemma 5.5
\[
\int_{R^N} \left( \Phi^\perp(X), Op\left[ - \left( \nabla \lambda_\eta(\cdot) I - \nabla X V(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right] \Phi(X) \right) dX
\]
\[
= \frac{1}{\rho_{cs}} \int_{R^N} \left( \Phi^\perp(X), Op\left[ - \left( \nabla \lambda_\eta(\cdot) I - \nabla X V(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right] \Phi(X) \right) dX = O(e^{\frac{1}{2} \rho_{cs}}).
\]
The main term related to the consistency of the approximation can be written as follows
\[
Op\left[ - \left( \nabla \lambda_\eta(\cdot) I - \nabla X V(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right]
\]
\[
= Op\left[ - \left( \nabla \lambda_\eta(X) I - \nabla X V(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right]
\]
\[
+ Op\left[ - \left( \nabla \lambda_\eta(\cdot) - \nabla \lambda_\eta(X) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right],
\]
where Lemma 5.3 shows that the last term is bounded by
\[
\sup_{(X, P) \in R^N} \left\| \left( \nabla \lambda_\eta(X) - \nabla \lambda_\eta(\cdot) \right) \cdot \nabla P \bar{u}(s, \cdot) \ast \phi_\eta(X, P) \right\| + O(\eta e^{C_s \delta^{-2}}) = O(e^{C_s \eta \delta^{-2}})
\]
since the symbol satisfies

\[
\int_{\mathbb{R}^N} (\nabla \lambda_\eta(X) - \nabla \lambda_\eta(X - X')) \cdot \nabla P \tilde{u}(s, X - X', P - P') \phi_\eta(X', P')dX'dP'
\]

\[
= \int_{\mathbb{R}^N} \sum_{j=1}^{3N} \partial_X \nabla \lambda_\eta(X) \cdot \nabla P \tilde{u}(s, X - X', P - P') \int_{\mathbb{R}^N} X_j' \phi_\eta(X', P')dX'dP'
\]

\[
+ \int_{\mathbb{R}^N} \sum_{j,k=1}^{3N} \int_0^1 (1 - t) \left( \partial_X \nabla \lambda_\eta(X - tX') \cdot \nabla P \tilde{u}(s, X - tX', P - P') \right) X_j' X_k' \phi_\eta(X', P')dX'dP'dt
\]

\[
\leq \| \sum_{j,k=1}^{3N} \int_0^1 (1 - t) \left( |\partial_X \nabla \lambda_\eta(X - tX') \cdot \nabla P \tilde{u}(s, X - tX', P - P')| 
\right. 
\]

\[
+ 2|\partial_X \nabla \lambda_\eta(X - tX') \cdot \partial_X \nabla P \tilde{u}(s, X - tX', P - P')| 
\]

\[
+ |\partial_X \nabla \lambda_\eta(X - tX') \cdot \partial_X \nabla P \tilde{u}(s, X - tX', P - P')| 
\]

\[\times \sup_{j,k} \| X_j' X_k' \phi_\eta(X', P') \|_{L^\infty} dt \cdot \]

The remaining term satisfies, using [2.27] and the notation \( \tilde{V}(X) := V(X) - \lambda_\eta(X) \),

\[
\int_{\mathbb{R}^N} \left( \Pi_0 \Phi \cdot \text{Op} \left[ \nabla_X \tilde{V}(X) \cdot \nabla P \tilde{u}(s, \cdot) \cdot \phi_\eta(X, P) \right] \right) \Pi_0 \Phi dX
\]

\[
= \int_{\mathbb{R}^N} \left( \Phi \cdot \Pi_0 \text{Op} \left[ \nabla_X \tilde{V}(X) \cdot \nabla P \tilde{u}(s, \cdot) \cdot \phi_\eta(X, P) \right] \right) \Pi_0 \Phi dX
\]

\[
= \int_{\mathbb{R}^N} \left( \Phi \cdot \Pi_0 \frac{1}{2} \left( \nabla_X \tilde{V}(X) \cdot \text{Op} \left[ \nabla P \tilde{u}(s, \cdot) \cdot \phi_\eta(X, P) \right] \right) \right) \Pi_0 \Phi dX + O(e^{C'r M^{-1} \delta^{-4}})
\]

\[
(2.28)
\]

\[
+ \int_{\mathbb{R}^N} \left( \Phi \cdot \Pi_0 \nabla_X \tilde{V}(X) \cdot \text{Op} \left[ \nabla P \tilde{u}(s, \cdot) \cdot \phi_\eta(X, P) \right] \right) \Pi_0 \Phi dX
\]

\[
+ \int_{\mathbb{R}^N} \left( \Phi \cdot \Pi_0 \nabla_X \tilde{V}(X) \cdot \text{Op} \left[ \nabla P \tilde{u}(s, \cdot) \cdot \phi_\eta(X, P) \right] \right) \Pi_0 \Phi dX
\]

\[
+ \int_{\mathbb{R}^N} \left( \Phi \cdot \Pi_0 \nabla_X \tilde{V}(X) \cdot \text{Op} \left[ \nabla P \tilde{u}(s, \cdot) \cdot \phi_\eta(X, P) \right] \right) \Pi_0 \Phi dX
\]

where the first term is \( O(\eta) \) since \( \tilde{V} = (V - \lambda_\eta) + \lambda_\eta - \lambda_\eta \) and the gradient of the ground state condition \((V - \lambda_\eta) \Psi_0 = 0 \) implies

\[
\langle \Psi_0, \nabla (V - \lambda_\eta) \Psi_0 \rangle = -\langle \Psi_0, (V - \lambda_\eta) \nabla \Psi_0 \rangle = -\langle (V - \lambda_\eta) \Psi_0, \nabla \Psi_0 \rangle = 0,
\]

and the second term is estimated by Lemma 5.5. Therefore the right hand side in (2.26) yields the contribution \( O(e^{C'r M^{-1} \delta^{-4}}) \) to the error bound [2.18]. \( \square \)

The next result shows that the difference of the observables can be estimated without the weight \( W \).

**Theorem 2.2.** Let \( \bar{E} := \int_{\mathbb{R}^N} H_\eta(Z) W(Z) dZ \) and assume that

(i) the assumptions in Theorem 2.7 hold, and
(ii) the function \( \tilde{g}_{MD} : \mathbb{R} \to \mathbb{R} \) is defined by

\[
g_{MD}(Z) =: \tilde{g}_{MD}(H_\eta(Z)), \quad \text{for } Z \in \mathbb{R}^{6N},
\]

then

\[
\int_{\mathbb{R}^{6N}} g_{MD}(Z)W(Z)dZ = \tilde{g}_{MD}(E) + \mathcal{O}(M^{-1}\delta^{-4} + p_{ex}^{1/2})
\]

and

\[
g_S - \tilde{g}_{MD}(E) = \mathcal{O}((M^{-1/2}\delta^{-2} + p_{ex}^{1/2})^{\frac{1}{(2\gamma+1)}}),
\]

where

\[
\dot{E} = E + \mathcal{O}(p_{ex} + \eta).
\]

**Proof.** The function \( \tilde{g}_{MD}(E) = \int_{\Sigma_E} g(z)d\nu(z)/\int_{\Sigma_E} d\nu(z) \) can be written as the convolution

\[
(2.29) \quad \tilde{g}_{MD}(E) = \int_{\Sigma_E} g(z)d\nu(z) = \int_{\mathbb{R}^{6N}} \int_{\Sigma_E} \tilde{g}(z-z')d\nu(z)\phi_\eta(z')dz',
\]

and as

\[
\tilde{g}_{MD}(E) = \int_{\mathbb{R}^{6N}} g(Z)\delta(H_\eta(Z) - E)dZ
\]

\[
\int_{\mathbb{R}^{6N}} \delta(H_\eta(Z) - E)dZ.
\]

Using the assumption that \( |\nabla_Z H_\eta(Z)|_{\Sigma_E} \) does not vanish, the function \( \tilde{g}_{MD} \) is smooth since we have for \( Z = (X, P) \)

\[
\frac{d}{dE} \int_{\mathbb{R}^{6N}} g(Z)\delta(H_\eta(Z) - E)dZ = -\int_{\mathbb{R}^{6N}} g(Z)\delta'(H_\eta(Z) - E)dZ
\]

\[
= -\int_{\mathbb{R}^{6N}} \frac{g(Z)}{|\nabla H_\eta(Z)|} \nabla H_\eta(Z) \cdot \nabla H_\eta(Z)\delta'(H_\eta(Z) - E)dZ
\]

\[
= -\int_{\mathbb{R}^{6N}} \frac{g(Z)}{|\nabla H_\eta(Z)|} \cdot \nabla_{\xi}(H_\eta(Z) - E)dZ
\]

\[
= \int_{\mathbb{R}^{6N}} \text{div}\left(\frac{g(Z)\nabla H_\eta(Z)}{|\nabla H_\eta(Z)|^2}\right)\delta(H_\eta(Z) - E)dZ
\]

and similarly for higher order derivatives.

Taylor expansion implies

\[
\tilde{g}_{MD}(\dot{E}) = \tilde{g}_{MD}(\dot{E}) + \tilde{g}_{MD}'(\dot{E})(\dot{E} - E) + \frac{1}{2} \tilde{g}_{MD}''(\xi)(\dot{E} - E)^2,
\]

for some \( \xi = \xi(\dot{E}) \) between \( \dot{E} \) and \( \dot{E} \) satisfying \( \tilde{g}_{MD}''(\xi(\dot{E})) := 2 \int_0^1 \tilde{g}_{MD}'(s\dot{E} + (1-s)E)(1-s)ds, \) and integration with \( \dot{E} = H_\eta(Z) \) yields the moment relation

\[
(2.30) \quad \int_{\mathbb{R}^{6N}} g_{MD}(Z)W(Z)dZ
\]

\[
= \int_{\mathbb{R}^{6N}} \tilde{g}_{MD}(H_\eta(Z))W(Z)dZ
\]

\[
= \tilde{g}_{MD}(E) \int_{\mathbb{R}^{6N}} W(Z)dZ
\]

\[
+ \tilde{g}_{MD}'(E) \int_{\mathbb{R}^{6N}} (H_\eta(Z) - \dot{E})W(Z)dZ
\]

\[
+ \frac{1}{2} \int_{\mathbb{R}^{6N}} \tilde{g}_{MD}''(\xi(H_\eta(Z)))(H_\eta(Z) - \dot{E})^2W(Z)dZ.
\]
The definition of $\tilde{E}$ shows the error estimate

\begin{equation}
\tilde{E} = \int_{R^N} H_0(Z) W(Z) dZ
= \int_{R^2} \langle \Phi(X), \hat{H}_0 \Phi(X) \rangle dX
= \int_{R^2} \langle \Phi(X), \hat{H} \Phi(X) \rangle dX - \int_{R^2} \langle \Phi(X), (V - \lambda_0) \Phi(X) \rangle dX
= E + O(p_{ex}) + O(\eta),
\end{equation}

since $\hat{H}_0 = \hat{H} - \hat{V} = \hat{H} - (V - \lambda_0) - (\lambda_0 - \lambda_0)$.

The rule for composition of Weyl quantizations, in (5.23) in Lemma 5.5 shows that

$$\hat{A} \hat{B} = \hat{A} \hat{B} - iM^{-1/2} \text{Op}[\{A, B\}] + O(M^{-1} \delta^{-4}),$$

for $A = H_0 - \tilde{E}$ and $B(Z) = \tilde{g}_MD''(\xi(Z))$, which implies

$$\text{Op}[\tilde{g}_MD''(\xi(Z))] = \text{Op}[\hat{H}_0 - \tilde{E}] \text{Op}[\tilde{g}_MD''(\xi(Z))] \text{Op}[H_0 - \tilde{E}] + O(M^{-1} \delta^{-4}),$$

since we have

$$\{\tilde{g}_MD''(\xi(Z)), H_0 - \tilde{E}\} = 0$$

$$\{H_0 - \tilde{E}, H_0 - \tilde{E}\} = 0.$$
where the amplitude function $\phi : \mathbb{R}^{3N} \to \mathbb{C}$ is complex valued, the phase $\theta : \mathbb{R}^{3N} \to \mathbb{R}$ is real valued in the classically allowed region $\{ X \in \mathbb{R}^{3N} \mid \lambda_0(X) \leq E \}$ (and purely imaginary elsewhere, see Remark 3.1), and the factor $M^{1/2}$ is introduced in order to have well-defined limits of $\phi$ and $\theta$ as $M \to \infty$. Note that it is trivially always possible to find functions $\phi$ and $\theta$ satisfying (3.1), even in the sense of a true equality. Of course, the ansatz only makes sense if $\phi$ and $\theta$ do not have strong oscillations for large $M$. The standard WKB-construction [33, 34] is based on a series expansion in powers of $M^{1/2}$, which solves the Schrödinger equation with arbitrarily high accuracy. Instead of an asymptotic solution, we introduce an actual local solution based on a time-dependent Schrödinger transport equation. This transport equation reduces to the formulation in [33].

3.1.1. A WKB-solution. The WKB-function (3.1) satisfies the Schrödinger equation (1.2) provided that

$$0 = (\hat{H} - E)\phi e^{iM^{1/2}\theta(X)} = \left( \frac{1}{2} |\nabla \theta|^2 + V - E \right)\phi - \frac{1}{2M} \Delta \phi - \frac{i}{M^{1/2}} (\nabla \phi \cdot \nabla \theta + \frac{1}{2} \phi \Delta \theta) e^{iM^{1/2}\theta(X)}.$$  \hspace{1cm} (3.2)

We shall see that only eigensolutions $\Phi$ that correspond to dynamics without caustics correspond to such a single WKB-mode, as for instance when the eigenvalue $E$ is inside an electron eigenvalue gap. Solutions in the presence of caustics use a Fourier integral of such WKB-modes, see [33] and Appendix 6. The purpose of the phase function $\theta$ is to generate an accurate approximation in the limit as $M \to \infty$. A possible definition, see [33], is the Hamilton-Jacobi equation (3.3)

$$\frac{1}{2} |\nabla \theta|^2 = E - \lambda_0.$$  \hspace{1cm} (3.3)

The solution to the Hamilton-Jacobi eikonal equation can be constructed locally from the associated Hamiltonian system

$$\dot{X}_t = P_t,$$

$$\dot{P}_t = -\nabla \lambda_0(X_t),$$  \hspace{1cm} (3.4)

through the characteristics path $(X_t, P_t)$ satisfying $\nabla \theta(X_t) = P_t$. For the time being, we fix some $P_0 \in \mathbb{R}^{3N}$ and vary $X_0$ over a hyperplane (denoted by $I$) in $\mathbb{R}^{3N}$ orthogonal to $P_0$, see Figure 1. Hence, the trajectories $X_t$ can (locally) cover the space. Later on, in Section 3.3 we shall argue that a superposition of such solutions actually allows us to reconstruct the true (local) solution of the Schrödinger equation. In particular, we will solve the system (3.4) not just once, but for all different $X_0$.

![Figure 1](image_url) 

**Figure 1.** Paths $X_t$ with velocity $\dot{X}_t = P_t$ going out from the plane with fixed normal $P_0 = \text{constant}$. 

The amplitude function $\phi$ can be determined by requiring the ansatz (3.1) to be a solution to (3.2), which gives

$$
0 = (\hat{H} - E)\phi e^{iM^{1/2}\theta(X)}
$$

$$
= \left( \frac{1}{2} |\nabla \theta|^2 + \lambda_0 - E \right) \phi
$$

$$
- \frac{1}{2M} \Delta \phi + (V - \lambda_0)\phi - \frac{i}{M^{1/2}} (\nabla \phi \cdot \nabla \theta + \frac{1}{2} \phi \Delta \theta) e^{iM^{1/2}\theta(X)}.
$$

Thus by using (3.3) we have

$$
(3.5) \quad - \frac{1}{2M} \Delta \phi + (V - \lambda_0)\phi - \frac{i}{M^{1/2}} (\nabla \phi \cdot \nabla \theta + \frac{1}{2} \phi \Delta \theta) = 0.
$$

The usual method for determining $\phi$ from this so-called transport equation uses an asymptotic expansion $\phi \simeq \sum_{k=0}^{K} M^{-k/2} \phi_k$, see [17, 25]. An alternative is to write it as a Schrödinger equation, as in [33]: we apply the characteristics in (3.4) to write

$$
\frac{d}{dt} \phi(X_t) = \nabla \phi \cdot \dot{X}_t = \nabla \phi \cdot \nabla \theta,
$$

and define the weight function $G$ by

$$
(3.6) \quad \frac{d}{dt} \log G_t = \frac{1}{2} \Delta \theta(X_t),
$$

and the variable $\psi_t := \phi(X_t) G_t$. Then the transport equation (3.5) becomes a Schrödinger-type equation

$$
(3.7) \quad iM^{-1/2} \psi_t = (V - \lambda_0)\psi_t - \frac{G_t}{2M} \Delta X \left( \frac{\psi_t}{G_t} \right) =: \hat{V} \psi_t.
$$

As (3.7) is (formally) a time-dependent Schrödinger equation, we need to impose initial conditions. Once more, we refer to Section 3.3 for a detailed discussion of this subtle issue.

The last step in the construction of the WKB system $(\theta, \phi)$ is to patch together the different trajectories obtained in (3.3)–(3.7). A priori, $X_t, P_t, G_t$ and $\psi_t$ are functions of $(t, X_0)$, for $t \geq 0$ and $X_0 \in I$, where $I$ was a hyperplane in $\mathbb{R}^{3N}$. By uniqueness of solutions to (3.4), the map $(t, X_0) \mapsto X_t$ is locally injective. Considering $-P_0$ in addition to $P_0$, we obtain that $(t, X_0) \mapsto X_t$ is locally invertible. Hence, we may also interpret the functions $P = P(t, X_0), G = G(t, X_0), \psi = \psi(t, X_0)$ as functions of the space variable alone. Abusing notation, we set $\nabla \theta(X) = P(t, X_0)$, $G(X) = G(t, X_0)$, $\psi(X) = \psi(t, X_0)$ for $X = X(t, X_0)$ in a neighborhood of $I$.

In conclusion, equations (3.3)–(3.7) determine the WKB-ansatz (3.1) to be a local solution to the Schrödinger equation (1.2) in the following sense. Assume that the Hamilton-Jacobi equation $\frac{\nabla \theta(X)}{2} + \lambda_0(X) = E$ has a $C^2$ solution $\theta : \mathcal{U} \to \mathbb{R}$ in a domain $\mathcal{U} \subseteq \mathbb{R}^{3N}$. Let $\dot{X}_t = \nabla \theta(X_t)$ and $P_t = \nabla \theta(X_t)$, then $(X_t, P_t)$ solves the Hamiltonian system (3.4), for $t \in [0, t_*)$ such that $X_t \in \mathcal{U}$. Then

$$
(3.8) \quad \Phi(X_t) = G^{-1}(X_t) \psi(X_t) e^{iM^{1/2}\theta(X_t)}
$$

solves (1.2) in $\mathcal{U}$, since both (3.3) and (3.5) are satisfied and consequently also (3.2). It is well known that Hamilton-Jacobi equations in general do not have global $C^2$ solutions, due to $X$-paths that collide and generate blow up in $\partial_X \theta(X)$. However if the domain is small enough and the data on the boundary is compatible (in the sense that $H_p(X, \nabla \theta(X)) = E$ on the boundary), non-characteristic (in the sense that the normal derivative $\partial_n \theta(X) \neq 0$ on the boundary) and $\lambda_0$ is smooth, then the converse property holds, i.e. the characteristics generate a local solution to the Hamilton-Jacobi equation, see [11]. Maslov’s method to find a global asymptotic solution by patching together local solutions is described in [35].

**Remark 3.1.** In the classically forbidden region $\{X \in \mathbb{R}^{3N} | E < \lambda_0(X)\}$ we can apply the same WKB method by replacing the phase function $\theta$ by the imaginary phase $i\theta$. The Eikonal equation becomes $-|P|^2/2 + \lambda_0(X) = E$, which then has a solution, and the transport equation becomes real valued

$$
M^{-1/2} \psi_t = -(V - \lambda_0)\psi_t + \frac{G_t}{2M} \Delta X \left( \frac{\psi_t}{G_t} \right).
$$

The matrix $V - \lambda_0$ is positive semi-definite, so that $\psi$ remains bounded for bounded time and approaches the ground state $\Psi_0$. In the classical region we instead use the oscillatory behavior to conclude that $\psi$ tends to the ground state, see [11].
3.2. The Landau-Zener model, transition probabilities and Ehrenfest dynamics. Notice that the parameter $t$ in (3.3)–(3.7) is just a numerical parameter, which, prima facie, has no connection with physical time. Indeed, we are working in a time-independent setting after all. On the other hand, if $t$ is interpreted as time, (3.7) corresponds to a time dependent Schrödinger equation, and we may ask for dynamic transition probabilities of the time-dependent model, denoted by $p_d(t) = p_d(X_t)$ and rigorously defined in (3.12). It turns out below that the time-dependent transition probabilities are easier to analyze than the time-independent ones, $p_{ex}$, at least under some simplifying assumptions. The link between the time-dependent transition probabilities $p_d$ and the time-independent $p_{ex}$ is not trivial and will be explored in detail in Section 3.3.

To have some intuition, it might be helpful to think of $p_d(t)$ as a “local” excitation probability around $X_t$, with the idea that $p_{ex}$, in turn, is given as a time/space average of $p_d$.

We start our discussion by looking at a simple special case of (3.7), the Landau-Zener model, for which the first results on transition probabilities with crossing or nearly crossing electron potentials were obtained. It is given by

\begin{equation}
M^{-1/2} \hat{\psi}_t = \begin{bmatrix} P_0 t & \delta \\ \delta & -P_0 t \end{bmatrix} \phi_t
\end{equation}

with a wave function $\phi : \mathbb{R} \to \mathbb{C}^2$, constant positive parameters $(M, P_0, \delta)$ and initial data $\lim_{t \to -\infty} \phi(t) = (1, 0)$. The transition probability

\begin{equation}
p_{\text{LZ}} := e^{-\pi \delta^2 M^{1/2}/P_0},
\end{equation}

is the so-called Landau-Zener probability, determined using Weber functions in (53) and illustrated in Figure 3. In this particular model, we note that $p_{\text{LZ}} = \lim_{t \to -\infty} |\phi_2(t)|^2$.

In the context of (3.4)–(3.7), the Landau-Zener model can be seen as a special case of

\begin{align*}
\dot{X}_t &= P_1, \\
\dot{P}_1 &= -\nabla \lambda_0(X_t), \\
iM^{-1/2} \hat{\psi}_t &= \begin{bmatrix} (V(X_t) - \lambda_0(X_t)) \psi_t - (2M)^{-1} \Delta \psi/G \\
\end{bmatrix} = \hat{V}(X_t)
\end{align*}

which by the WKB-method (3.8) determines a Schrödinger WKB-solution $\Phi_Q$ locally and hence the transition probability

\begin{equation}
p_d(Q, X_t) := \langle \psi_t, \psi_t \rangle - |\langle \psi_t, \Psi_0(X_t) \rangle|^2,
\end{equation}

using the initialization $\psi(X_t) = \Psi_0(X_t)$ for $X_t \in I$, where the inflow domain is given by

\begin{equation}
I = \{ X : (X - Y) \cdot P = 0 \}
\end{equation}

for a given point $Q := (Y, P) \in \mathbb{R}^{2N}$ and the ground state $\Psi_0$ is defined in (1.6). By neglecting the small term $-G/(2M) \Delta \psi/G$ in the transport equation (3.11) and using the simplification with the potential $\lambda_0(X) = 0$ and

\begin{equation}
\hat{V}(X) = \begin{bmatrix} X & \delta \\
\delta & -X \end{bmatrix},
\end{equation}

we obtain the Landau-Zener model. The Landau Zener model was constructed to model and explain the dynamic transitions from the ground state to an excited state when the electron potential surfaces cross or nearly cross (with a minimal distance $\delta$) and the eigenstates change rapidly (near $X = 0$ for small values of $\delta$): the eigenvalues of $\hat{V}$ in (3.14) are $\lambda_{\pm}(X) = \pm \sqrt{X^2 + \delta^2}$ and the eigenvectors

\begin{equation}
\Psi_{\pm} = \frac{1}{\sqrt{\delta^2 + (\lambda_{\pm}(X) - X)^2}} \begin{bmatrix} \delta \\
\lambda_{\pm}(X) - X \end{bmatrix}.
\end{equation}

Remark 3.2. In the one dimensional avoided crossing case (3.14) the electron eigenvectors satisfy $\Psi_{\pm}(X) = f(X/\delta)$ for a smooth function $f : \mathbb{R} \to \mathbb{R}^2$ with

\begin{align*}
\lim_{X \to \infty} \Psi_{\pm}(X) &= \begin{bmatrix} 1 \\
0 \end{bmatrix}, \\
\lim_{X \to -\infty} \Psi_{\pm}(X) &= \begin{bmatrix} 0 \\
1 \end{bmatrix}.
\end{align*}
We see that
\[ \| \partial_X \Psi_+ \|_{L^\infty(\mathbb{R})} = O(\delta^{-1}), \]
for \( \delta \neq 0 \). The other eigenvector \( \Psi_- \) satisfies the same bounds. Therefore assumption (iii) in Theorem 2.1 holds for this avoided crossing with any \( \delta \neq 0 \).

3.2.1. Ehrenfest dynamics. As far as computations are concerned, the system (3.11) is still very demanding due to the Laplacian on the right hand side. A further simplification mainly introduced for computational reasons (and used in the computational examples of this paper) is the so-called Ehrenfest dynamics
\[
\begin{align*}
\dot{X}_t &= P_t, \\
\dot{P}_t &= -\nabla \lambda_0(X_t) - \frac{\langle \psi_t, \nabla \hat{V}(X_t) \psi_t \rangle}{\langle \psi, \psi \rangle},
\end{align*}
\]
which is an approximate WKB solution, by neglecting the small term \(-G/(2M)\Delta(\psi/G)\) in the transport equation and replacing the Eikonal equation with \( |P|^2/2 + \lambda_0(X) + \langle \psi, \hat{V}(X) \psi \rangle = E \), as in \( \| \). If we write the wave function in its real and imaginary parts \( \psi = \psi^r + i\psi^i \), and use initial data that satisfy \( \langle \psi_0, \psi_0 \rangle = 2M^{-1/2} \), then the Ehrenfest dynamics (3.15) is a Hamiltonian system with the Hamiltonian
\[
H_E = |P|^2/2 + \lambda_0(X) + M^{1/2}\langle \psi, \hat{V}(X) \psi \rangle/2,
\]
using the primal variables \( X \) and \( \psi^r \), the dual variables \( P \) and \( \psi^i \). If the probability, \( \langle \psi^r, \psi^r \rangle \), to be in the excited state is small (using the projection in \( L^2 \)) the Ehrenfest Hamiltonian is \( O(\langle \psi^r, \psi^r \rangle) \) close to the Born-Oppenheimer Hamiltonian \( |P|^2/2 + \lambda_0(X) \).

3.3. An estimate of the probability to be in excited states for matrix-valued potentials. This section presents a formal stability study of a perturbed eigenvalue problem that provides an approximation for the probability to be in excited states, \( p_{\text{ex}} \). The two ingredients are first to determine the perturbation as a dynamic transition problem, related to the Landau-Zener model, and thereafter to use the stability analysis of a matrix eigenvalue problem to identify the probability to be in the excited state with the squared norm of the change of the eigenvector. The dynamic transition problem can be approximated numerically using Ehrenfest dynamics, as described in Sections 3.2 and 3.3.1. Our formulation of the perturbed eigenvalue problem is related to the transformation from local WKB solutions, which relate to the dynamic transition probability \( p_d \), to a global solution of the Schrödinger equation. In dimension one we view a local WKB solution as a solution to the left of the hyperplane \( I \) (which is a point in dimension one) and another WKB solution to the right of the hyperplane. The continuity condition to match these two WKB solutions to the right and left to one global differentiable solution forms the eigenvalue problem we will study. The eigenvalue problem is consequently not formulated directly for the unbounded Schrödinger operator. It is instead an eigenvalue problem with bounded solution operators obtained from WKB solutions. A reason we use this form of the perturbation analysis is that small dynamic transition probabilities \( p_d \) from Landau-Zener like dynamics can be viewed as small regular perturbations while the corresponding perturbations of the potential in the Schrödinger equation is not small and not regular. The construction is extended to several dimensions in Section 3.3.3.

3.3.1. The perturbed eigenvalue problem in one space dimension. Consider first the scalar Schrödinger eigenvalue problem in one space dimension. Assume we have solutions \( \Phi \) in the domain \( X > 0 \) and \( \Phi \) in \( X < 0 \), with right and left limits \( \lim_{X \to 0^+} \Phi(X) =: \Phi^r \), \( \lim_{X \to 0^-} \Phi(X) =: \Phi^l \), and \( \lim_{X \to 0^+} \Phi'(X) =: \Phi'^r \), \( \lim_{X \to 0^-} \Phi'(X) =: \Phi'^l \). In the one dimensional case, the WKB method for an eigenvalue problem typically gives a caustic to the left, with phase \( \theta_r \), and a caustic to the right, with phase \( \theta_r \), (see Section 3.1, 6.3, and 7.1) so that
\[
\begin{align*}
\Phi(X) &\simeq \phi_r(X) \cos \left( M^{1/2} \theta_r(X) - \pi/4 \right), \text{ for } X > 0, \\
\Phi(X) &\simeq \phi_L(X) \cos \left( M^{1/2} \theta_L(X) - \pi/4 \right), \text{ for } -X > 0,
\end{align*}
\]
for smooth functions \( \theta_r \) and \( \phi_r \). The continuity condition in order to have a global solution
\[
\begin{align*}
\Phi^r &= \Phi^l, \\
\Phi'^r &= \Phi'^l,
\end{align*}
\]
can with the notation $\Phi'' =: R\Phi'$ (and similarly for the limit to the left) be written

\[(3.18) \quad \begin{bmatrix} 1 & -1 \\ Rr & -Rl \end{bmatrix} \begin{bmatrix} \Phi' \\ \Phi' \end{bmatrix} = 0\]

which implies that the $2 \times 2$ matrix

\[
\begin{bmatrix} 1 & -1 \\ Rr & -Rl \end{bmatrix}
\]

must be singular, i.e., $Rr = Rl$. For the WKB solutions the derivative satisfies

\[
Rr = -\frac{M_1}{2} \theta'(0) \tan \left( \frac{M_1}{2} \theta(0) - \frac{\pi}{4} \right) + O(M_0),
\]

\[
Rl = -\frac{M_1}{2} \theta'(0) \tan \left( \frac{M_1}{2} \theta(0) - \frac{\pi}{4} \right) + O(M_0).
\]

The two eigenvalues of the matrix

\[
\begin{bmatrix} 1 & -1 \\ Rr & -Rl \end{bmatrix}
\]

satisfy

\[(3.19) \quad \mu = \frac{1 - Rl}{2} \pm \sqrt{\left(\frac{1 - Rl}{2}\right)^2 + Rl - Rr}.
\]

Consider the two decoupled scalar Schrödinger eigenvalue problems

\[
\begin{align*}
&\left(-\frac{1}{2M} \Delta + \lambda_1(X)\right) \Phi_1(X) = E \Phi_1(X), \quad \text{and} \\
&\left(-\frac{1}{2M} \Delta + \lambda_2(X)\right) \Phi_2(X) = E \Phi_2(X).
\end{align*}
\]

The continuity condition corresponding to (3.18) for (3.20) will be used as our unperturbed eigenvalue problem, and in one space dimension this continuity condition becomes the unperturbed problem

\[(3.21) \quad \begin{bmatrix} 1 & -1 \\ Rr_1 & -Rl_1 \end{bmatrix} \begin{bmatrix} 0 \\ \Phi_1' \end{bmatrix} = 0.
\]

We consider a special case of the Schrödinger eigenvalue equation (1.2) in one space dimension where the potential is diagonal outside the interval $[-a, 0]$, i.e., $V(X) = \begin{bmatrix} \lambda_1(X) & 0 \\ 0 & \lambda_2(X) \end{bmatrix}$, $X \not\in [-a, 0]$, and view it as a perturbation of the diagonal Schrödinger equation (3.20), as illustrated in Figure 2. To have a diagonal potential in the unperturbed case is not necessary. What is important is that the electron eigenvalues are well separated so that the transition from ground state to excited states is negligible.

**Figure 2.** The eigenvalues $\lambda_1$ and $\lambda_2$ of the potential $V$ in the unperturbed (left) and perturbed (right) cases as a function of position $X$. For $X \not\in [-a, 0]$ the potential $V(X)$ is diagonal in both the unperturbed and perturbed cases while for $X \in [-a, 0]$ the potential $V(X)$ is diagonal in the unperturbed case and non diagonal in the perturbed case.

The WKB-method (3.8) shows that

\[(3.22) \quad \Phi^\pm(X) := \psi_\pm(X) G_\pm^{-1}(X) e^{\pm i M^{1/2} \theta(X)}\]
are local solutions to the Schrödinger equation (1.2), with positive and negative wave speeds \( \pm \theta'(X) \). When the path \( X_t \) passes through a domain where the electron eigenvalues are close, the wave function \( \psi_+ \) (or \( \psi_- \)) determined by the transport equation (3.7) yields the probability \( p_d \) to transit from the ground state to an excited state, as in the Landau-Zener model (3.9). For a WKB-solution (3.22), with \( \Phi = \Phi^+ \) (or \( \Phi = \Phi^- \)), we can define the transition operator \( \hat{S}^+ \) by

\[
\Phi^+(0) = \hat{S}^+ \Phi^+(-a) = \begin{bmatrix} \hat{S}_{11} & \hat{S}_{12} \\ \hat{S}_{21} & \hat{S}_{22} \end{bmatrix} \begin{bmatrix} \Phi_1^+(a) \\ \Phi_2^+(a) \end{bmatrix},
\]

(and similarly for \( \Phi^- \)) by solving the eikonal and transport equations (3.11) with data given on \( X = -a \). The idea is that the point \( X = -a \) and the momentum \( P \) in this point the two first equations determine the path \( (X(t), P(t)) \). The third Schrödinger like transport equation in (3.11) then yields a linear solution operator along this path determining \( \psi \) at \( X = -a \). Let us now determine \( \hat{S} \) more precisely. We want to determine the global effect of a small transist probability \( |\hat{S}_{21}|^2 \) viewed as the perturbation to the eigenvalue problem (3.21) with \( \hat{S}^+ \Phi^+(-a) + \hat{S}^- \Phi^-(-a) \) replacing \( \Phi' \).

To simplify the perturbation analysis we factorize the transition

\[
\begin{bmatrix} \hat{S}_{11} & \hat{S}_{12} \\ \hat{S}_{21} & \hat{S}_{22} \end{bmatrix} = \begin{bmatrix} 1 & \hat{S}_{12} \\ \hat{S}_{21} & 1 \end{bmatrix} \begin{bmatrix} \hat{S}_{11} & 0 \\ 0 & \hat{S}_{22} \end{bmatrix}
\]

into a transition between the states \( S := \begin{bmatrix} 1 & \hat{S}_{12} \\ \hat{S}_{21} & 1 \end{bmatrix} \) and a diagonal matrix \( \begin{bmatrix} \hat{S}_{11} & 0 \\ 0 & \hat{S}_{22} \end{bmatrix} \), which does not contribute to the transition between states. Since we are only interested in estimates of the small components \( \hat{S}_{12} = S_{21} S_{22} \) and \( \hat{S}_{21} = S_{11} S_{21} \), we consider only perturbations generated by the transition matrix \( S \). We note that the transition element \( \hat{S}_{21} \) by (3.23) can be written as the 2-component of \( \Phi(0) \) if \( \Phi(-a) = (1, 0) \). Therefore (3.22) shows that the transition element

\[
S_{21}^+ = \frac{\hat{S}_{21}^+}{S_{21}^-} = \frac{G_+(0)e^{i\alpha_{1/2}g(0)}\psi_2(0)}{G_-(0)e^{i\alpha_{1/2}g(0)}\psi_1(0)} = \frac{\psi_2(0)}{\psi_1(0)} \quad \text{where } |\psi_1(0)| = 1
\]

can be determined by the WKB-amplitude function \( \psi_+ \) and similarly we have the transition element \( S_{21}^- = \psi_2(0)/\psi_1(0) \). The dynamic transition probability \( p_d \), defined by Ehrenfest dynamics in (3.12), measures in this case the amplitude squared in the excited state at \( X = 0 \) for a wave starting in \( X = -a \) in the ground state. Consequently we have, for \( X_t = 0 \) and \( Q \) corresponding to the + wave speed in (3.22), that

\[
p_d(Q, 0) = \langle \psi_+(0), \psi_+(0) \rangle - |\langle \psi_+(0), \Psi_0(0) \rangle|^2 = |\psi_1(0)|^2 + |\psi_2(0)|^2 - |\psi_1(0)|^2 = |\psi_2(0)|^2
\]

so that by (3.24) \( S_{21}^\pm = \mathcal{O}(p_d^{1/2}) \), as \( p_d \to 0 \) (namely as the spectral gap becomes large). We assume for simplicity that the transition matrix \( S = S^\pm \) is translation invariant, so that \( \Phi^\pm (0) = S^\pm \Phi^\pm (-a) \). The total perturbation becomes the sum \( S\Phi := S^+ \Phi^+ + S^- \Phi^- \) of the perturbations of the two WKB-solutions \( \Phi^\pm \), for the splitting \( \Phi = \Phi^+ + \Phi^- \). Hence, we consider perturbations

\[
\Phi^\pm(0) = S^\pm \Phi^\pm(-a) = \begin{bmatrix} 1 - \mathcal{O}(p_d) & \mathcal{O}(p_d^{1/2}) \\ \mathcal{O}(p_d^{1/2}) & 1 - \mathcal{O}(p_d) \end{bmatrix} \begin{bmatrix} \Phi^+_1(-a) \\ \Phi^-_2(-a) \end{bmatrix}
\]

with the transition from the state 1 to state 2 determined by the matrix component \( S_{21}^\pm \), which is of the (small) order \( p_d^{1/2} \), where \( p_d \), defined in (3.12), is related to the Landau-Zener probability. A reason to decompose the solution into WKB solutions is that their transition is determined by the amplitude functions \( \psi_\pm \) with approximately conserved norm \( |\psi_\pm(t)| \). The approximate conservation follows from the conservation \( \frac{d}{dt} |\psi_\pm(t)|^2 = 0 \) for the Ehrenfest dynamics (3.15) and the fact that the transport equation (3.11), determining \( \psi_\pm \), becomes the Ehrenfest dynamics in the limit \( M \to \infty \).

The perturbed eigenvalue condition becomes

\[
(A + \beta)r_\beta = 0,
\]
For sufficiently small matrices $\beta$ and $r$, we want to determine the change in the eigenvector $|r_\beta - r_0|^2$ (which measures the probability to be in the excited state) from the perturbed eigenvalue problem (3.25). For this we use the result in the following lemma, whose derivation follows [15].

**Lemma 3.3.** Assume that the matrix $A \in \mathbb{C}^{n \times n}$ has $n$ distinct eigenvalues, $\mu_j$, $0 \leq j \leq n - 1$. Let $l_j$ and $r_j$ be left and right eigenvectors of $A$, i.e. $l_j A = \mu_j l_j$, $A r_j = \mu_j r_j$, satisfying

$$l_j \cdot r_k = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases}$$

For sufficiently small matrices $\beta \in \mathbb{C}^{n \times n}$ the eigenvectors and eigenvalues of the perturbation matrix $A + \beta$ are differentiable functions of $\beta$. If $r_\beta$ and $\mu_\beta$ denote the eigenvector and eigenvalue to $A + \beta$ that equal the eigenpair $(r_0, \mu_0)$ to the matrix $A$ for $\beta = 0$, we have that

$$l_j \cdot (r_\beta - r_0) = -\frac{l_j \cdot \beta r_0}{\mu_j - \mu_0} + o(|\beta|).$$

and

$$\mu_\beta - \mu_0 = l_0 \cdot \beta r_0 + o(|\beta|).$$

**Proof.** That the eigenvectors and eigenvalues of $A + \beta$ are differentiable functions of $\beta$ in a neighborhood of $\beta = 0$ follows directly by differentiation of the relations

$$(A + \beta)r = \mu r, \quad l(A + \beta) = \mu l.$$

Let $B := |\beta|^{-1}\beta$, where $|\beta|$ is a matrix norm, e.g., the Euclidean operator norm. Let $r(\gamma)$ and $\mu(\gamma)$ be the perturbed normalized eigenvector and eigenvalue corresponding to $r_0$ and $\mu_0$ given by $r(0) = r_0$ and $\mu(0) = \mu_0$, and

$$(A + \gamma B)r(\gamma) = \mu(\gamma)r(\gamma).$$

Differentiation of (3.29) at $\gamma = 0$ gives

$$(A \gamma B)r(\gamma) = \mu(\gamma)r(\gamma).$$

Using that $r(0) = r_0$ and $\mu(0) = \mu_0$ and multiplying by $l_j$ from the left, we get

$$(\mu_j - \mu_0)l_j \cdot r(0) = l_j \cdot Ar(0) - \mu_0 l_j \cdot r(0) = -l_j \cdot Br_0 + \mu(0)l_j \cdot r_0 = -l_j \cdot Br_0, \quad \text{for } j \neq 0,$$

by the orthogonality between left and right eigenvectors. The Taylor expansion $r(|\beta|) = r(0) + r(0)|\beta| + o(|\beta|)$ gives (3.27).

To prove (3.28) we take the scalar product with $l_0$ from the left in (3.30), which gives $\mu'(0) = l_0 \cdot Br_0$. The Taylor expansion $\mu(|\beta|) = \mu(0) + \mu'(0)|\beta| + o(|\beta|)$ gives (3.28). □
The combination of (3.27) in Lemma 3.3 and the assumption that \( r_β - r_0 = \mathcal{O}(1) \), as \( |β| \to 0^+ \), also when \( |μ_j - μ_0| \) is small, implies that

\[
\ell_j \cdot (r_β - r_0) = \begin{cases} 
- \ell_j \cdot βr_0 + o(|β|), & \text{if } |μ_j - μ_0|^{-1} = o(|β|^{-1}), \\
\mathcal{O}(1), & \text{otherwise},
\end{cases}
\]

(3.31)

\[
|ℓ_j| \cdot βr_0 + o(|β|)
\]

[ and by the definition \( (μ_j - μ_0)^2 := μ_j - μ_0 + c|ℓ_j \cdot βr_0| \text{sign}(μ_j - μ_0) \]

\[
= - \frac{ℓ_j \cdot βr_0}{(μ_j - μ_0)^2} + o(|β|),
\]

for some positive constant \( c \), which determines the perturbation orthogonal to \( r_0 \). We will also use that this representation holds separately for each perturbation \( S^+ \) and \( S^- \). We do not determine the perturbation in the \( r_0 \) direction, which does not contribute to the transition between states. We denote by \( r_β^+ - r_0 \) the projection of the perturbation \( r_β - r_0 \) on the hyperplane orthogonal to \( r_0 \). Since this plane is spanned by the left eigenvectors \( ℓ_1, \ldots, ℓ_{n-1} \), we have by (3.27) in Lemma 3.3 that

\[
r_β^+ - r_0 = - \sum_{i_j, r_0 \neq 0} \frac{l_j \cdot βr_0}{μ_j - μ_0} + O(|β|).
\]

Returning to the perturbed problem (3.25), we recall from equation (3.19) that the unperturbed eigenvalues are given by

\[
\begin{bmatrix}
μ_0 \\
μ_1 \\
μ_2
\end{bmatrix} =
\begin{bmatrix}
\frac{1-R^t_1}{2} & \frac{1-R^t_1}{2} & \frac{1-R^t_1}{2} \\
\frac{1-R^t_1}{2} & \frac{1-R^t_1}{2} & \frac{1-R^t_1}{2} \\
\frac{1-R^t_1}{2} & \frac{1-R^t_1}{2} & \frac{1-R^t_1}{2}
\end{bmatrix}
\]

\frac{1-R^t_2}{2} + \frac{1-R^t_2}{2} + \frac{1-R^t_2}{2},

and the corresponding unperturbed eigenvectors are

\[
\ell_j = \begin{cases} 
\frac{(1-j)(1-j)}{R^t_2} & \text{when } j = 0, 1, \\
\frac{(1-j)(1-j)}{R^t_2} & \text{when } j = 2, 3.
\end{cases}
\]

(3.32)

\[
μ_β = μ_0 = l_0 \cdot βr_0 + o(|β|) = (1 - S_{11}) \frac{1 - (1 - μ_0)^2}{2} \frac{R^t_2}{R^t_1} + o(\sqrt{p_d}) = o(\sqrt{p_d}).
\]

To have the perturbed eigenvalue equal to zero means that we start with an unperturbed eigenvalue \( μ_0 = o(\sqrt{p_d}) \) in order to obtain the perturbed eigenvalue \( μ_β = 0 \), satisfying \((A + β)r_β = 0\). Using the unperturbed eigenvector \( r_0 = (1, 1 + o(\sqrt{p_d}), 0, 0) \), with \( μ_0 = o(\sqrt{p_d}) \), yields a perturbation in the \( Φ_2 \) component from zero (in the third component of \( r_0 \) to)

\[
S_{21} \sum_{j=2} \frac{(μ_j - μ_0)^2}{2} \frac{R^t_2}{R^t_2} (1 - (1 - μ_j)^2 \frac{1}{R^t_1}) + o(\sqrt{p_d}) = O(\sqrt{p_d})
\]

and its squared absolute value measures the probability to be in the excited state, i.e. the probability density \( |Φ_2|^2 \), which integrated yields \( p_{ex} \) defined in (1.8).

**Remark 3.4.** Having a zero eigenvalue of \( A_1 \) in (3.26) means that the scalar Schrödinger eigenvalue problem

\[
( - \frac{1}{2M} \Delta + λ_1(X))Φ_1(X) = EΦ_1(X)
\]

has a solution in the whole domain. We see from (3.33) that when also \( A_2 \) has a zero eigenvalue \( μ_2 = 0 \) we obtain a resonance in the sense that a small perturbation \( β \) (where \( |S_{21}| = O(|β|) = O(p_d^{1/2}) \)) yields a large change in the eigenvector \( r_β = Φ \), meaning that the probability to be in the excited state (which by (3.33) is of the order \( |β|^2/|μ_2|^2 \)) can be of order one even if the perturbation \( β \) is tiny. Resonance
means that the two scalar Schrödinger eigenvalue problems have a coinciding eigenvalue $E$, see Figure 3. An approximation of this probability is

\[(3.34) \quad |\beta|^2/|\mu|^2 \approx p_d/(|\mu - \mu_0|^2 + c^2p_d).\]

### 3.3.2. Numerical test of the perturbation analysis

A method to numerically test the validity of the perturbation analysis leading to the estimate (3.33), for the one dimensional model, is to compare the probability to be in excited states $p_{ex}$ in (1.8) (determined from numerical approximation of the Schrödinger equation (4.3)) with the expression (3.34) derived from (3.33) in the perturbation analysis. If $p_{ex}$ shows some similarity with (3.34) the perturbation analysis is in some sense justified. The difference $|\mu - \mu_0|$, of the eigenvalues in (3.34), can be obtained from numerical approximation of a difference $E_{2,k'} - E_{1,k}$, of the eigenvalues of the decoupled system (5.20), $(\lambda_i(X) - \frac{1}{2}M^{-1}\frac{\partial^2}{\partial X^2})\Phi_{i,k}(X) = E_{i,k}\Phi_{i,k}(X)$, $i = 1, 2$, using the same centered difference approximation as in (4.3). Let $E'_i = \{E_{i,k} | k = 1, 2, 3, \ldots\}$ be the set of all $E_i$ eigenvalues, for $i = 1, 2$. By the approximation

\[|\mu_2 - \mu_0|^2 \approx C'|E_{1,k} - E_{2,k'}|^2 \approx C^{-1}|E_{1,k} - E_{2,k'}|^2/|E_{1,k} - E_{1,k+1}|^2,\]

where $k$ and $k'$ is chosen such that $E_{1,k}$ and $E_{2,k'}$ are the closest to $E$ and using two constants $C$ and $C' = C^{-1}/|E_{1,k} - E_{1,k+1}|^2$, we obtain the following estimate of the probability to be in the excited state

\[(3.35) \quad \hat{p}_{ex} := \frac{p_d}{C^{-1}|E'_1 - E'_2|^2/|E'_1 - E'_m|^2 + p_d},\]

with $C$ being a constant, $E$ being an eigenvalue of the discrete two-state Schrödinger equation corresponding to (3.35) and

\[E'_1 := \min\{E'_{1} - E'_{1}|, \quad E'_m := \min\{E'_{1} - E'_{m}|, \quad E'_2 := \min\{E'_{2} - E'_{1}|, \quad p_d := \exp(-\pi\delta^2\sqrt{M}/\sqrt{2(E - \lambda_-(0))}).\]

### Conclusions

Figures 3 and 3 show that the estimation of the probabilities, $\hat{p}_{ex}$, to be in the excited state, obtained using the formula (3.35), and the numerically computed probabilities, $p_{ex}$, to be in the excited state, obtained using the formula (4.4) from the solution of the discrete one dimensional Schrödinger eigenvalue problem (4.3), have similar qualitative behavior in the following two aspects: (1) the minimal value of $p_{ex}$ and $\hat{p}_{ex}$ are similar, and (2) when $E$ is close to a resonance, seen for $\hat{p}_{ex}$ in (3.35) as $|E'_1 - E'_2| \ll |E'_1 - E'_m|$, both $p_{ex}$ and $\hat{p}_{ex}$ become larger. By also adjusting the constant $C$ in (3.35) the behavior can also be quantitatively similar, see Figure 3 and 3. That is the perturbation analysis seems to capture this resonance phenomenon, at least qualitatively.

### 3.3.3. The perturbed eigenvalue problem in the multi-dimensional case

The main idea to extend the one dimensional perturbed eigenvalue problem to multiple dimensions is to write the Schrödinger wave function as a phase-space integral of highly oscillatory functions using the Fourier-Bros-Iagolnitzer transform (FBI-transform). This integral can be viewed as an integral over all hyperplanes $I$ in $\mathbb{R}^N$ defined in (3.13). Therefore each FBI-mode can be initial data for a WKB solution (3.11) to the right of a hyperplane $I$ in $\mathbb{R}^N$ and a WKB solution to the left of the hyperplane. The condition to have a differentiable WKB solution across the hyperplane then uses the continuity condition (3.25), as the one dimensional setting, for each point in the hyperplane. The direction of the oscillations in the FBI-mode generates through the WKB-method a molecular dynamics path and an amplitude function. The amplitude function solves, along the path, a transport equation that is well approximated by Ehrenfest dynamics (3.15) and can consequently be computed similarly as the molecular dynamics paths but with smaller time steps, see Sections 3.2 and 3. The integral of WKB-solutions over all hyperplanes contribute through their amplitude functions to the transition, $S$. We will now explain this extension to multiple dimensions more precisely and then in the next section show how to computationally approximate the probability to be in excited states $p_{ex}$.

We will use the FBI-transform. We could use the standard semi-classical Fourier transform as the initial data for the WKB-method but the FBI-transform has the advantage of giving the high frequency content locally in $X$, i.e. microlocally, which yields a more accurate WKB Ansatz. The important property of the FBI transform

\[(3.36) \quad T\varphi(Y, P) := (2^{1/2})^{-3N/4}M^{3N/8} \int_{\mathbb{R}^{3N}} e^{iM^{1/2}(Y - X)P - |Y - X|^2M^{1/2}/2} \varphi(X) dX = \alpha_M\]
Figure 3. Probabilities to be in the excited state, $p_{ex}$, using numerical approximation of (1.8), precisely defined in (4.4), and the estimations, $\hat{p}_{ex}$, using the formula (3.35), obtained by solving the discrete one-dimensional Schrödinger eigenvalue problem (4.3) with $M = 1000$, $\delta = \sqrt{\frac{3}{5}}M^{-0.2} \approx 0.1946$, $C = 0.09$, and mesh size $h = 0.0001$. 

(a) The probability $p_{ex}$

(b) The probability $p_{ex}$ (left panel) to be in an excited state compared to the estimate $\hat{p}_{ex}$ (right panel)

(c) The probability $p_{ex}$ (‘o’) to be in an excited state compared to the estimate $\hat{p}_{ex}$ (‘*’)
\[ \delta = \left( \frac{3}{5} \right)^{1/2} M^{-1/5}, \ C = 0.09 \]

\( E - \) eigenvalue

\( p \)

\( \hat{p}_{ex} \)

(a)

\[ \delta = M^{-1/5}, \ C = 0.09 \]

\( E - \) eigenvalue

\( p \)

(b)

\[ \delta = \left( \frac{9}{5} \right)^{1/2} M^{-1/5}, \ C = 5 \]

\( E - \) eigenvalue

\( p \)

(c)

\[ \delta = \left( \frac{9}{5} \right)^{1/2} M^{-1/5}, \ C = 1 \]

\( E - \) eigenvalue

\( p \)

(d)

Figure 4. Probabilities to be in the excited state, \( p_{ex} \), using numerical approximation of (1.8), precisely defined in (4.4), and the estimations, \( \hat{p}_{ex} \), using the formula (3.35), obtained by solving the discrete one-dimensional Schrödinger eigenvalue problem (4.3) with \( M = 20000 \).

is the identity

(3.37)  \[ \varphi = T^* T \varphi \]

for \( \varphi \in L^2(\mathbb{R}^{3N}) \), where the adjoint operator \( T^* \) is defined by

\[ T^* \phi(X) = \alpha_M \int_{\mathbb{R}^N} e^{-iM^{1/2} (Y-X) \cdot P - |Y-X|^2 M^{1/2}/2} \phi(Y,P) dY dP \]

e.g. for all Schwartz functions \( \phi \) on \( \mathbb{R}^{6N} \), see [25]. Therefore the integral representation

\[ \Phi(X) = T^* T \Phi(X) \]

\[ = \alpha_M \int_{\mathbb{R}^N} e^{-iM^{1/2} (Y-X) \cdot P - |Y-X|^2 M^{1/2}/2} T \Phi(Y,P) dY dP \]

yields suitable boundary data on hyperplanes (e.g. \( X_1 = 0 \)) for the WKB-method. We will, for each point \( (Y,P) =: Q \) (in the classically allowed region), use a WKB-function (3.1). In the case of caustics this single WKB-function is replaced by a finite sum of WKB-functions based on the Legendre transform \( \theta^*(P) \) of \( \theta(X) \), as explained in Section 6 and 7

(3.38)  \[ \Phi_Q(X) \sim \sum_{\{P^* : (P^* , \nabla \theta^*(P^*) = X\}} \phi_Q(X;P^*) e^{iM^{1/2} \theta^*(X;P^*)} \]

This WKB solution \( \Phi_Q \) solves the Schrödinger eigenvalue problem (1.2) in the two domains to the left and right of the hyperplane, with the boundary condition constructed so that the WKB-mode \( \Phi_Q(X) \) is equal to the FBI-mode

\[ \alpha_M e^{-iM^{1/2} (Y-X) \cdot P - |Y-X|^2 M^{1/2}/2} T \Phi(Y,P) \]
for $X - Y$ in the hyperplane $L_P := \{ q \in \mathbb{R}^N \mid q \cdot P = 0 \}$ orthogonal to $P$:

$$q \cdot \nabla \theta_0(X) = q \cdot P = 0, \quad \text{for } q \in L_P \text{ and } X - Y \in L_P$$  \hspace{1cm} (3.39)

$$\phi_Q(X; P^*) = \alpha_M T \Phi(Q) e^{-(|X-Y|^2 / M^{1/2})}, \quad \text{for } X - Y \in L_P.$$  \hspace{1cm} (3.40)

The construction of the WKB-solutions in the case of caustics uses asymptotic convergence as $M \to \infty$, described by the asymptotically equal sign $\sim$, see [54, 33] and (7.1).

The obtained decomposition

$$\Phi(X) = \int_{\mathbb{R}^N} \Phi_Q(X) dQ$$  \hspace{1cm} (3.40)

determines, for each $Q$, the corresponding operators $R_{c,Q}^{r,t}$. The total perturbation comes from transitions from all $\phi_Q$ along the paths: for each $Q = (Y, P)$ we let the function $\phi_Q(X)$ solve the transport equation (3.7) along the WKB-path $\{ X_\tau \}_{\tau = 0}^\infty$ and initialize $\phi_Q(X_\tau)$ to the ground state $\Psi_0(X_\tau)$ each time $X_\tau$ is in the hyperplane $(X_\tau - Y) \cdot P = 0$, generated by $Q$. The perturbed wave function is then as in (3.33) determined by

$$\langle S(Q, X), \Psi_1(X) \rangle = \frac{\langle \phi_Q(X), \Psi_1(X) \rangle}{\langle \phi_Q(X), \Psi_0(X) \rangle},$$

which depends on $Q$, by relating $\psi_{1+}$ in (3.24) with $\langle \phi_Q, \Psi_0 \rangle$ and $\psi_{2+}$ with $\langle \phi_Q, \Psi_1 \rangle$ in the case of two states. Integrating over all $Q$ yields the total perturbation, as in (3.26),

$$\int_{\mathbb{R}^N} S(Q, X) \Phi_Q(X) dQ$$

at the point $X$. The corresponding perturbed eigenvalue problem (3.25) for the wave function $\Phi_Q$ implies, as in (3.33), the following larger change in the excited state

$$\Phi^\perp(X) = \Phi(X) - \langle \Psi_0(X), \Phi(X) \rangle \Psi_0(X)$$

$$= \int_{\mathbb{R}^N} S_{21}(Q, X) \Phi_Q(X) \sum_{j=2,3} \frac{1 - (1 - \mu_j)^2 R_{c,j}^Q}{\pi^2 \mu_j} dQ$$

$$+ O(\int_{\mathbb{R}^N} S_{21}(Q, X) \Phi_Q(X) dQ)^2)$$

at the point $X$, where $\mu_j$ and $R_{c,j}^Q$ depend on $Q$. The index 1 in $S$ now corresponds to the component in the ground state and the index 2 to be in an excited state, i.e.,

$$\Phi(X) = \gamma_1(X) \Psi_0(X) + \gamma_2(X) \Phi^\perp(X),$$

and $\gamma_j(X) \in \mathbb{C}$ with the orthogonality conditions

$$\langle \Psi_0, \Phi^\perp \rangle = 0, \quad \langle \Psi_0, \Psi_0 \rangle = 1, \quad \langle \Psi^\perp, \Psi^\perp \rangle = 1.$$
negligible small, we have
\[
A[(\Phi^+ (X), \Phi^+ (X))] \leq CA \int_{\mathbb{R}^N} |S_{21}(Q, X)\Phi_Q(X)|^2 dQ
\]
\[
\leq CA \int_{\mathbb{R}^N} \frac{|S_{21}(Q, X)\Phi_Q(X)|^2}{|\mu_2 - \mu_0|^2} dQ
\]
\[
= C \int_{\mathbb{R}^N} \int_{\mathbb{R}} |S_{21}(Q, X)||\Phi_Q(X)|^2 \frac{1}{|c^2|S_{21}(Q, X)|^2} dQ
\]
\[
= C \int_{\mathbb{R}^N} |S_{21}(Q, X)||\Phi(X)|^2 \frac{\mu_2}{c^2|S_{21}(Q, X)|^2} dQ
\]
\[
\leq C \int_{\mathbb{R}^N} |S_{21}(Q, X)||\Phi(X)|^2 \|\Phi\|_{L^\infty(\mathbb{R})} dQ
\]
\[
= O \left( \int_{\mathbb{R}^N} |S_{21}(Q, X)||\Phi(X)|^2 dQ \right).
\]

We conclude that in a model of local ensemble averages of eigenvalues \( E \), assuming that for each \( X \) the eigenvalue difference \( |\mu_2 - \mu_0| \) has a bounded density, then the average probability to be in the excited state has the bound
\[
A[p_{ex}] = O(\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |S_{21}(Q, X)||\Phi_Q(X)|^2 dQ dX),
\]
which, by (3.41) can be written as the average of the square root of the dynamic transition probability \( p_d \)
\[
A[p_{ex}] = O(\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} p_d^{1/2}(Q, X)|\Phi(X)|^2 dQ dX).
\]

3.3.4. Ergodic computation of probabilities to be in excited states. The FBI-transform satisfies
\[
\|T\Phi\|_{L^2(\mathbb{R}^N)} = \|\Phi\|_{L^2(\mathbb{R}^N)} = 1
\]
and \( T\Phi \) concentrates on the phase-space set \( H_0(Y, P) = |P|^2/2 + \lambda_0(Y) = E \) in the limit as \( M \to \infty \) and \( p_{ex} \to 0+ \), see [25]. When the dynamics is ergodic, the phase-space measure is in addition uniform on the set \( E < H_0(Q) < E + \gamma \) as \( \gamma \to 0+ \), cf. [37]. The WKB functions \( \Phi_Q(X) \) behave similarly as \( T\Phi \) locally, since their initial conditions are given by \( \Phi \) and they solve the Shrödinger equation. We may write \( S_{21}(Q, X) = S_{21}(Q; X, P) \) where \( P \) is the momentum for the path at the position \( X \) that started in the plane given by \( Q = (Y', P') \), since this \( P \) is a function of \( Q \) and \( X \). Therefore we approximate \( p_{ex} \) by the following time average of \( p_d^{1/2} \)
\[
p_{ex} \approx \lim_{\gamma \to 0+} \int_{\mathbb{R}^N} \int_{E < H_0(Q) < E + \gamma} |S_{21}(Q, X, P)|^2 dQ dX
\]
\[
= \lim_{\gamma \to 0+} \int_{\mathbb{R}^N} \int_{E < H_0(Q) < E + \gamma} |S_{21}(X, P; Q)|^2 dQ dX
\]
\[
= \lim_{T \to \infty} \int_{\mathbb{R}^N} \int_{0}^{T} \left( \frac{|S_{21}(X, P; X_t, P_t)|}{T} \right) dt dX,
\]
where we used that the transition probability \( S_{21}(Q; X, P) = S_{21}(X, P; Q) \) is symmetric by reversing time \( t \) to \( -t \). In fact, along a given path \( \{X_t\}_{t=0}^{T} \) also the momentum \( \mathcal{P}' \), at position \( Y' \), is determined by the position \( Y' \) and \( X_t \) so that \( p_d^{1/2}(Y', X_t) = S_{21}(Y', P'; X_t, P_t) \) is well defined. The relation (3.43) means that we sample square roots, \( p_d^{1/2}(X, X_t) \), of transition probabilities along the molecular dynamics path, normalized for each passage through a hyperplane, and then take the average of all hyperplanes.
In our computations we sample hyperplanes by means of the phase space trajectory of ground state Born-Oppenheimer molecular dynamics, cf. Algorithm [1] and Figure [5].
Algorithm 1 Approximations of the probabilities, \( \hat{p}_{ex} \), to be in excited states using molecular dynamics in 2D based on (3.43)

**Input:** Energy \( E \); potential functions \( V \); mass \( M \); time \( T \); initial position \( X_0 \), initial momentum \( P_0 \).

**Output:** Approximated probability \( \hat{p}_{ex} \) to be in excited states.

1. Sampling of initialization times \( \{T_n\}_n \) and hyperplane coordinates \( \{\overline{X}_{T_n}, \overline{P}_{T_n}\}_n \):
   - Set \( t \leftarrow 0, n \leftarrow 0, T_n \leftarrow 0 \) and \( (X_{T_0}, P_{T_0}) \leftarrow (X_0, P_0) \) and define the hyperplane \( \overline{P}_{T_0} := \{X \in \mathbb{R}^2 \mid X \cdot \overline{P}_{T_0} = 0\} \).
   - while \( T_n < T \) do
     - Simulate the ground state Born-Oppenheimer molecular dynamics for \( (X_t, P_t) \) until \( X_t \) crosses the plane \( \overline{P}_{T_n} \).
     - At the crossing time, set \( n \leftarrow n + 1, T_n \leftarrow t \) and \( (X_{T_n}, P_{T_n}) \leftarrow (X_t, P_t) \) and define the hyperplane \( \overline{P}_{T_n} := \{X \in \mathbb{R}^2 \mid X \cdot \overline{P}_{T_n} = 0\} \).
   - end while

2. Solve the Ehrenfest molecular dynamics using the Störmer-Verlet method with the entries in \( \{T_k, \overline{X}_{T_k}, \overline{P}_{T_k}\}_{k=0}^n \) as input parameters and compute \( \hat{p}_{ex} \):
   - \( \text{temp} \leftarrow 0 \)
   - for \( k = 0 \) to \( n - 1 \) do
     - \( t_1 \leftarrow T_k \)
     - \( t_2 \leftarrow T_{k+1} \)
     - \( X_{t_1} \leftarrow \overline{X}_{T_k} \)
     - \( P_{t_1} \leftarrow \overline{P}_{T_k} \)
     - \( \psi_{t_1} = \Psi_-(X_{t_1}) \)
     - \( (X_t, P_t, \psi_t) \leftarrow \) Ehrenfest dynamics path obtained from time \( t_1 \) to \( t_2 \) with initial data \( X_{t_1}, P_{t_1}, \psi_{t_1} \)
     - \( \text{temp} \leftarrow \text{temp} + \int_{t_1}^{t_2} |\langle \psi_t, \Psi_+(X_t) \rangle| dt \)
   - end for
   - \( \hat{p}_{ex} \leftarrow \text{temp}/T_n \)

---

**Figure 5.** Illustration of the sampling of hyperplanes and initialization times described in part 1 of Algorithm 1. Left figure: the 2D Born-Oppenheimer molecular dynamics trajectory \( X_t \) (blue line) with initial conditions \( X_{T_0} = (-2, 0.5) \) and \( X_{T_0} = P_{T_0} \) is simulated until the time \( t = T_1 \) when it crosses the hyperplane \( P_{T_0} \) (dashed line). Right figure: a new hyperplane \( P_{T_1} \) (dash-dotted line) is constructed and the trajectory of \( X_t \) is simulated further until the time \( t = T_2 \) when it crosses that hyperplane. The sampling procedure is iterated until the final time is reached.
4. Numerical examples

The purpose of this section is to present two simple model problems, where the Schrödinger eigenvalue solution can be studied computationally and compared to the molecular dynamics approximation. In the following subsections we

- show that the dynamic transition probability $p_d$ in (3.12) can be determined by numerical solution of Ehrenfest dynamics,
- verify assumption (2.17) on exponential convergence rate in finite time,
- compare numerical approximations of observables from the Schrödinger equation and Born-Oppenheimer molecular dynamics,
- compare approximations of the probability to be in excited states $p_{ex}$ from the Schrödinger equation with the molecular dynamics approximation $\hat{p}_{ex}$ from Algorithm 1.

4.1. Model 1: A one dimensional problem. We consider the one dimensional, time-independent Schrödinger equation (1.2) with the heavy-particle coordinate $X \in \mathbb{R}$, two electron states $J = 2$ and the Hamiltonian operator, $\hat{H}(X)$, defined by

\begin{equation}
\hat{H}(X) := V(X) - \frac{1}{2} M^{-1} \frac{\partial^2}{\partial X^2},
\end{equation}

with the potential operator $V$ defined by the matrix

\begin{equation}
V(X) := \begin{bmatrix} X + r(X) & \delta \\ \delta & -X + r(X) \end{bmatrix},
\end{equation}

where the parameter $\delta$ is a non-negative constant and the function $r : \mathbb{R} \to \mathbb{R}$ is given by

\begin{equation}
\begin{aligned}
r(X) := \begin{cases}
(a_l - X)^2 & \text{if } X < a_l, \\
(X - a_r)^2 & \text{if } X > a_r, \\
0 & \text{otherwise}.
\end{cases}
\end{aligned}
\end{equation}

For each $X$ the potential $V$ defines the eigenvalue problem $V(X)\Psi_{\pm}(X) = \lambda_{\pm}(X)\Psi_{\pm}(X)$, with the two eigenvalues $\lambda_{\pm}(X) \in \mathbb{R}$ and two eigenvectors $\Psi_{\pm}(X) \in \mathbb{R}^2$. The eigenvalues are given by $\lambda_{\pm}(X) = r(X) \pm \sqrt{X^2 + \delta^2}$, and $\Psi_{-}(X)$ and $\Psi_{+}(X)$ respectively denote the ground and the excited state vector.

Choosing $\delta = 0$ gives a conical intersection at $X = 0$, and a positive value of $\delta$ gives a minimum gap $2\delta$ between $\lambda_{-}(X)$ and $\lambda_{+}(X)$ at $X = 0$. A small value of $\delta$ corresponds to a large probability to be in excited states and a large value of $\delta$ corresponds to a small probability to be in excited states. Figure 6 illustrates examples of small and large spectral gaps between $\lambda_{+}$ and $\lambda_{-}$ with two different values of $\delta$, respectively.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{delta.png}
\caption{Eigenvalues of the potential matrix (4.2) for small and large values of the parameter $\delta$ corresponding to the small and large minimum gaps between the eigenvalues, respectively.}
\end{figure}

4.1.1. The 1D discretized Schrödinger equation. We use a central difference method to solve the one dimensional two-state Schrödinger equation (1.2) in the domain $X \in (-2\pi, 2\pi)$ with $(a_l, a_r) = (-2, 3)$, mesh-size $h = 4\pi/[10M^{3/4}]$ and partition $X^j = -2\pi + jh$, $j = 0, 1, 2, \ldots, [10M^{3/4}]$. The approximation
of the eigenvalue problem with the eigenvector components \( \Phi_j^i \simeq \Phi(X^i) \) and corresponding eigenvalue \( E_h \in \mathbb{R} \) becomes

\[
(4.3) \quad -\frac{1}{2M} \Phi_{j-1}^i - 2\Phi_j^i + \Phi_{j+1}^i + V(X^i)\Phi_j^i = E_h \Phi_j^i, \quad \text{for } j = 1, 2, \ldots, [10M^{3/4}] - 1,
\]

using homogeneous Dirichlet boundary conditions \( \Phi_h^0 = \Phi_h^{10M^{3/4}} = 0 \).

The approximation of the probability \( p_{ex} \) to be in the excited state is given by

\[
(4.4) \quad p_{ex} = \frac{\sum_{j=0}^{[10M^{3/4}]} |\langle \Phi_j^h, \Psi_+(X^j) \rangle|^2}{\sum_{j=0}^{[10M^{3/4}]} |\langle \Phi_j^h, \Phi_j^i \rangle|^2}.
\]

4.1.2. Approximation of the dynamic transition probability \( p_d \). We show in this computational example that the dynamic transition probability, \( p_d \), defined in (3.12), can be obtained from Ehrenfest molecular dynamics simulations (3.15) using the following formula

\[
(4.5) \quad p_d(t) = \left| \frac{\langle \Psi_+(X_t), \psi_t \rangle}{\langle \Psi_+(X_t), \Psi_+(X_t) \rangle} \right|^2,
\]

with \( t \in \mathbb{R} \) denoting the time. We observe that numerical experiments illustrate that the transition probability \( p_d(t) \) approximates, as time tends to infinity, the Landau-Zener probability \( p_{LZ} \) as defined in (3.10), given the Landau-Zener model (3.9), see Figure 8.

We approximate the solution of the transport equation (3.7) by the Ehrenfest molecular dynamics simulations (3.15) with \( \lambda_0 = \lambda_- \) (equal to the smallest eigenvalue of \( V \) in (4.2)) in the Hamiltonian (3.6). For this computational example, we choose \( E = 1 \), \( (a, a_c) = (-2, 2) \) and the initial conditions \( t = 0 \), \( X_0 = -4.0 \), \( \psi_0 = \Psi_-(X_0) \), and \( P_0 = \sqrt{2(E - \lambda_-(X_0))} \). We approximate the solution of the molecular dynamics using the Störmer-Verlet method, [18], based on the two symplectic Euler steps,

\[
(4.6) \quad \begin{align*}
\psi_{k+\frac{1}{2}}^i &= \psi_k^i - \frac{\Delta t}{2} \sqrt{M} \dot{V}(X_k) \psi_k^i, \\
P_{k+\frac{1}{2}} &= P_k - \frac{\Delta t}{2} \left( \lambda_-(X_k) + \langle \psi_k^i, \ddot{V}(X_k) \psi_k^i \rangle + \langle \psi_{k+\frac{1}{2}}^i, \ddot{V}(X_k) \psi_{k+\frac{1}{2}}^i \rangle \right), \\
X_{k+1} &= X_k + \Delta t P_{k+\frac{1}{2}}, \\
\psi_{k+1}^i &= \psi_{k+\frac{1}{2}}^i + \frac{\Delta t}{2} \sqrt{M} \dot{V}(X_{k+1}) \psi_{k+\frac{1}{2}}^i, \\
P_{k+1} &= P_{k+\frac{1}{2}} - \frac{\Delta t}{2} \left( \lambda_-(X_{k+1}) + \langle \psi_{k+1}^i, \ddot{V}(X_{k+1}) \psi_{k+1}^i \rangle + \langle \psi_{k+\frac{1}{2}}^i, \ddot{V}(X_{k+1}) \psi_{k+\frac{1}{2}}^i \rangle \right).
\end{align*}
\]

Figure 7. The potential energy function, \( t \mapsto \langle \psi_t, V(X_t) \psi_t \rangle \), approximated from the Ehrenfest molecular dynamics simulations (3.15), deviates away from the eigenvalues after the avoided crossing, if the gap is small and the mass is sufficiently small.
Conclusions. Figure 7 illustrates that the potential energy function \( t \rightarrow \langle \psi_t, V(X_t)\psi_t \rangle \), obtained from the Ehrenfest molecular dynamics simulations is equal to the ground state energy, \( \lambda_-(X_t) \), until the dynamics reaches the avoided conical intersection. Then it deviates from the ground state energy and continues in between the ground and excited state energies. Figure 8 shows that the transition probability, \( p_d(t) \), obtained using the formula (4.5) based on Ehrenfest molecular dynamics simulations, remains zero until it reaches the avoided conical intersection and then it oscillates near the avoided conical intersection region and approaches \( p_{LZ} \) asymptotically as \( t \to \infty \).

4.2. Model 2: A two dimensional problem. We consider the two dimensional, time-independent Schrödinger equation (1.2) with the heavy-particle coordinate \( X = (X_1, X_2) \in \mathbb{R}^2 \) and two electron states \( J = 2 \). The Hamiltonian is

\[
\hat{H}(X) = V(X) - \frac{1}{2} M^{-1} \Delta X,
\]

Figure 8. Estimation of the Landau-Zener probability, \( p_{LZ} \), from the Ehrenfest molecular dynamics simulations (3.15). The estimation, \( p_d(t) \), overshoots near the avoided crossing and eventually stabilizes around \( p_{LZ} \). The overshoot period and the relative magnitude are more prominent for the cases of smaller Landau-Zener probabilities.
with

\[ V(X) := \left( \frac{1}{2}(X_1^2 + \alpha X_2^2) + \beta \sin(X_1 X_2) \right) \eta + \eta \begin{bmatrix} v_1(X) & v_2(X) \\ v_2(X) & -v_1(X) \end{bmatrix}, \]

where \( I \) is the \( 2 \times 2 \) identity matrix, the functions \( v_1, v_2 : \mathbb{R}^2 \to \mathbb{R} \) are given by (A) or (B) below and \( \alpha = \sqrt{2}, \beta = 2 \) and \( \eta = 1/2 \). As in Example 1, we have for each \( X \) the eigenvalue problem,

\[ V(X) \Psi_\pm(X) = \lambda_\pm(X) \Psi_\pm(X), \]

with the eigenvalues \( \lambda_\pm(X) = \eta \sqrt{v_1(X)^2+(v_2(X))^2} \) and the ground state and excited state eigenvectors \( \Psi_-(X) \) and \( \Psi_+(X) \), respectively. We choose the energy, \( E = 1.5 \). In this example we study the following two cases of interactions between the two potential surfaces \( \lambda_+ \) and \( \lambda_- \):

(A) A line intersection. We choose \( v_1(X) = \arctan(X_1/\eta) \) and \( v_2(X) = \delta/\eta \), where \( \delta \) is a non-negative constant that defines the minimum distance between the potential surfaces. In this example, the potential surfaces intersect each other (for \( \delta = 0 \)) or have the minimum distance between each other at the line \( X_1 = 0 \). Here, choosing a small value for \( \delta \) corresponds to a large probability to be in excited states and a large value for \( \delta \) will give a small probability to be in excited states. The parameter \( \delta \) in this two dimensional example is analogous to the parameter \( \delta \) in the onedimensional example given in Section 4.1

(B) A conical intersection. We choose \( v_1(X) = \arctan((X_1-a_1)/\eta) \) and \( v_2(X) = \arctan((X_2-a_2)/\eta) \) with \( a = (a_1, a_2) \in \mathbb{R}^2 \) being a chosen point in the two dimensional space. If \( a \) is chosen such that \( \lambda_\pm(a) \) are smaller than the energy, \( E \), we have a conical intersection between the potential surfaces at the point \( a \) in the classically allowed region \( R := \{ X : \lambda_-(X) \leq E \} \), otherwise we have a positive gap between the potential surfaces in the domain \( R \). Here, choosing \( a \) in the origin gives a larger absolute momentum, \( |P| \), at the conical intersection whereas choosing \( a \) far from the origin will yield a smaller \( |P| \), which yield larger and smaller probabilities to be in excited states, respectively. Figure 9 shows level curves of the eigenvalues, along with example molecular dynamics paths, for varying conical intersection points.

4.2.1. The 2D discretized Schrödinger equation. We use a standard finite difference method to discretize the two dimensional Schrödinger eigenvalue problem \( \{12\} \) with mesh-size \( h \) in both the \( X_1 \) and \( X_2 \) directions in the computational domain \( \Omega = [-4,4]^2 \). The unknown eigenvalue components \( \Phi_{h}^{j,k} \simeq \Phi(X_{j,k}) \), with the nodal point \( X_{j,k} = (-4+jh,-4+kh) \) for \( j,k = 1,2,\ldots,8/h-1 \), solve the discrete eigenvalue problem

\[ -\frac{1}{2M} \left( \frac{\Phi_{h}^{j-1,k} - 2\Phi_{h}^{j,k} + \Phi_{h}^{j+1,k}}{h^2} + \frac{\Phi_{h}^{j,k-1} - 2\Phi_{h}^{j,k} + \Phi_{h}^{j,k+1}}{h^2} \right) + V(X_{j,k})\Phi_{h}^{j,k} = E_{h}\Phi_{h}^{j,k} \]

with homogeneous Dirichlet boundary conditions and \( h = 1/(4\sqrt{M}) \). We use the solution of the discrete Schrödinger eigenvalue problem to determine the approximate Schrödinger observables

\[ g_{s} = \frac{\sum_{j,k} g(X_{j,k})p(X_{j,k})}{\sum_{j,k} p(X_{j,k})}, \]

with \( g(X) : \mathbb{R}^2 \to \mathbb{R} \) and \( p(X) = |\Phi_h(X)|^2 \).

The value of the molecular dynamics observable, \( g_{\text{MD}} \), is given by

\[ g_{\text{MD}} := \lim_{\delta \to 0+} \frac{\int_{E<H_0(X,P)<E+\delta} g(X,P)dXdP}{\int_{E<H_0(X,P)<E+\delta} dXdP}, \]

which, for the two dimensional case, can be written\(^1\) as

\[ g_{\text{MD}} = \frac{\int_{H_0(X,0)\leq E} g(X)dX}{\int_{H_0(X,0)\leq E} dX}, \]

when \( g \) depends only on \( X \).

\(^1\)For a fixed \( X \in \mathbb{R}^d \), \( P \)-integration using spherical coordinates in the shell \( E < |P|^2/2 + \lambda_0(X) < E + \delta \) yields the \( P \)-dependence \( [|P|^2]^{\frac{E-\lambda_0(X)/2}{\sqrt{E-\lambda_0(X)}}} \), which differentiated with respect to \( \delta \) gives \( |P| \) for \( d = 2 \) and the additional factor \( (E-\lambda_0(X))^{d/2-1} \) for \( d \neq 2 \).
The numerical method is given by the steps \( \bar{X} \) based on numerical approximation of the ergodic projected stochastic dynamics \( \bar{X} \).

Verification of the ergodic rate condition

Figure 10 determines the convergence rate \( \gamma \) in assumption (2.17) based on numerical approximation of the ergodic projected stochastic dynamics (2.10). The numerical method is given by the steps \( \bar{Z}(n\Delta t) \rightarrow \bar{Z}((n+1)\Delta t) \)

\[
(X, P) = \bar{Z}(n\Delta t) \\
X_* = X + P\Delta t \\
P_* = P - \nabla \lambda(X_*)\Delta t \\
Z^1 = (X_*, P_*) \\
Z^2 = Z^1 + \sqrt{\tau}(Z^1)\Delta W \\
Z^3 = Z^2 + \mu \nabla H(Z^2)/|\nabla H(Z^2)|
\]

where \( \mu \in \mathbb{R} \) is chosen so that \( H(Z^2 + \mu \nabla H(Z^2)/|\nabla H(Z^2)|) = E \). Here \( \Delta W = \sqrt{\Delta t}(\xi_1, \xi_2, \xi_3, \xi_4) \) and \( \xi_i \) are all independent taking values \( \pm 1 \) with probability \( 1/2 \). The temperature parameter is \( \tau = 1/2 \). This scheme with phase-space path \( \bar{Z}(n\Delta t) = (\bar{X}(n\Delta t), \bar{P}(n\Delta t)) \) \( n = 0, 1, 2 \ldots \) satisfies the convergence assumptions in (12) so that \( \mathbb{E}[g(\bar{X}(T), \bar{P}(T)) - g(X(T), P(T))] = O(\Delta t) \), uniformly in \( T \). Figure 10 shows that in this case \( 2\gamma > 0.2 \). The maximal Lyapunov constant is determined numerically to \( \hat{C} = 0.35 \) and the convergence rate exponent \( \gamma/(\hat{C} + \gamma) \) in Theorem 2.1 becomes 0.1/0.45 \( \approx 0.2 \).

Convergence of observables in 2D. Figures 11 and 12 show that the Schrödinger observables, \( g_S \), are close to the ergodic molecular dynamics observables, \( g_{\text{MD}} \). We can compare the numerical results in Figures 11 and 12 with Theorem 2.2 for the case when the probabilities to be in the excited state, \( p_{ex}\), are small, as in Figure 13 for \( a = (5, 0) \) in the conical intersection case. We see in Figures 11 and 12 that

\( \text{Related to examples (4.15) and (4.7) in [12].} \)
Figure 10. The figure shows the function $T \mapsto T^{-1} \log \left( T^{-1} \int_0^T (\mathbb{E}[g \circ S_{t0}(Z_0)] - g_{MD}(Z_0)) dt \right)$, which approximates the decay rate $2\gamma$ in assumption (2.17), for the projected dynamics (2.10) with the potential $\lambda(X) = X_1^2/2 + X_2^2/\sqrt{2} + 2\sin(X_1)X_2$ and energy $E = H(Z_0) = 1.5$. The empirical mean is $\mathbb{E}[g \circ S_{t0}(Z_0)] := \sum_{n=1}^N \sin(X_1(t;n)X_2(t;n))/N$ and the number of stochastic paths are $N = 2 \times 10^5$. The initial point is $Z_0 = (0, 0, \sqrt{1.5}, \sqrt{1.5})$ and $(X(t;n), P(t;n))$ is the phase space point for sample path $n$ at time $t$ determined by the numerical method (4.11) for time steps $\Delta t = 0.01$. The blue solid curve is for diffusion coefficient $\epsilon = 10^{-2}$ and the green dashed-dot for $\epsilon = 0.25$. Here the approximation $g_{MD}(Z_0) = -0.4388$ is determined by an accurate Monte Carlo simulation in the microcanonical ensemble.

<table>
<thead>
<tr>
<th>$g = \sin^2(X_1X_2)$</th>
<th>$g = \lambda_-(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_\infty$</td>
<td>-0.3079</td>
</tr>
<tr>
<td>$e_W$</td>
<td>-0.2744</td>
</tr>
<tr>
<td>$e_1$</td>
<td>-0.2671</td>
</tr>
<tr>
<td>$e_2$</td>
<td>-0.2890</td>
</tr>
</tbody>
</table>

Table 1. Numerical results for the order of convergence of the observables in 2D as explained in Section 4.2.3. The table shows $\varrho$ where the order of convergence is $O(M^{\varrho})$. The convergence rate is computed as $\varrho = (\log e(M_1) - \log e(M_2))/\log M_1 - \log M_2$, with $e(M_i)$ being the norm of error corresponding to the mass $M_i$, where $M_1 = 200$, and $M_2 = 3200$. We use the norms $e_\infty = \max(|(g_\text{ref} - g_{MD})(E_i)|$, $e_W = |\sum_i (g_\text{ref} - g_{MD})(E_i)|/N_E$, $e_1 = \sqrt{\sum_i ((g_\text{ref} - g_{MD})(E_i))^2}/N_E$, and $e_2 = \sqrt{\sum_i ((g_\text{ref} - g_{MD})(E_i))^2}/N_E$, with $N_E$ being the number of eigenvalues $E_i$.

the standard deviation of $|g_\text{ref} - g_{MD}|$ for $M = 3200$ is roughly 1/2 times of the standard deviation for $M = 200$, i.e. the numerical results agree roughly with an $O(M^{-1/4})$ estimate in the theorem, see Table 1.

4.2.4. Approximation of the probability to be in excited states. We approximate the probability to be in the excited state, $p_{ex}$, from the Schrödinger equation by the approximation, $\bar{p}_{ex}$, from Algorithm 1, using Ehrenfest molecular dynamics simulations based on the Störmer-Verlet method as in (4.6). We choose $t_\text{0} = 0$, $p_\text{0} = 1, \sqrt{2(E - \lambda_-(X_0)) - 1}$, $\phi_\text{0} = \Psi_-(X_0)$, $X_\text{0} = [2, -0.5]$ (line intersection) and $X_\text{0} = [-2, 0.5]$ (conical intersection).

Conclusions. Figure 14 shows the probabilities, $p_{ex}$, to be in the excited state determined from the discrete Schrödinger eigenvalue problem (4.8) based on the formula (4.4). We note that for large gaps, in the right column, the probabilities to be in the excited state, $p_{ex}$, are either small or close to one, as expected, while for small gaps, in the left column, $p_{ex}$ takes values in the whole interval $[0,1]$. We also observe the lack of the structure present in the analogous Figure 4 for the one dimensional case, due
Figure 11. Schrödinger observables, \( g_S \), as a function of Schrödinger eigenvalues, \( E_h \), for mass \( M = 200 \), and the conical intersection case with \( a = (5, 0) \) outside the classically allowed region, compared with the corresponding molecular dynamics observables, \( g_{MD} \), which we compute using Monte Carlo integrations based on the formula (4.10). The plots show that the solutions obtained with the mesh size \( h = 1/(4\sqrt{M}) \) and \( h = 1/(16\sqrt{M}) \) are comparable, whereas the solution obtained with the mesh size \( h = 1/(2\sqrt{M}) \) appears less accurate.

Figure 13 shows \( p_{ex} \) determined from the discrete Schrödinger eigenvalue problem and from molecular dynamics simulations for both line and conical intersection cases computed by Algorithm 1. The shapes of the respective methods’ solutions are similar, especially when the probability to be in the exited state, \( p_{ex} \), is small. Figure 14 presents an empirical probability density of inter-arrival hitting times of hyperplanes, cf. Algorithm 1. The empirical density is well approximated by a heavy tailed Burr XII probability density, which indicates that long inter-arrival hitting times are non-negligible. In Figure 16 we observe an expected decay of \( O(T^{-1/2}) \) for the relative error \( |\tilde{p}_{ex}(T) - \tilde{p}_{ex}|/\tilde{p}_{ex} \) when studying the conical intersection problem with \( a = (5, 0) \). In the more unstable parameter setting of \( a = (0, 0) \), however, the error decay is slower with an observed rate of convergence slightly slower than the expected \(-1/2\). This might be due to a longer correlation length in the molecular dynamics and hyperplane sampling of Algorithm 1 when the parameter values of \( a \) are close to \((0, 0)\).

Remark 4.1. [Extension to higher dimension] The formation of a conical intersection for a \( 2 \times 2 \) symmetric matrix \( V \) requires the eigenvalues to be equal, which means that the off diagonal element is zero and the diagonal elements are equal at an intersection point. To have a conical intersection point is therefore generic in dimension two and in higher dimensions the intersection is typically a co-dimension two set. An implementation of the algorithm for computing \( \hat{p}_{ex} \) in higher dimensions and for potentials used in chemistry is future work. The relevant geometry in the multi-dimensional case is the distance to this co-dimension two set, which makes the link to two dimensional case studied here.
Figure 12. Schrödinger observables, $g_S$, as a function of Schrödinger eigenvalues, $E_h$, for mass $M = 3200$, and the conical intersection case with $a = (5,0)$ outside the classically allowed region, compared with the corresponding molecular dynamics observables, $g_{MD}$, which we compute using Monte Carlo integrations based on the formula (4.10). The plots show that the solutions obtained with the mesh size $h = 1/(4\sqrt{M})$ and $h = 1/(8\sqrt{M})$ are comparable, whereas the solutions obtained with the mesh size $h = 1/(2\sqrt{M})$ appear less accurate.
Figure 13. Plots showing $p_{ex}$ estimated from the solution of the discrete Schrödinger eigenvalue problems (left column) and the corresponding $\hat{p}_{ex}$ estimated by the molecular dynamics Algorithm 1 (right column) for the line intersection cases (first row) and conical intersection cases (second row). The Schrödinger solution is computed by averaging the solutions for 20 eigenvalues around $E = 1.5$ for the line intersection and by averaging the solutions for 500 eigenvalues around $E = 1.5$ for the conical intersection.

Figure 14. Normalized histogram of the inter-arrival initialization times obtained by Algorithm 1 when applied to the conical intersection problem with $a = (5, 0)$ and $M = 12800$, and, for comparison, the probability density function of the Burr XII distribution with parameter values $\alpha = 1.65$, $c = 3.6$ and $k = 0.5$. We do not know why this distribution fits reasonably well to the data.
Figure 15. Probabilities, \( p_{ex} \), to be in the excited state as a function of the eigenvalues, \( E_h \), computed from the solution of the two dimensional Schrödinger equations and based on the formula (4.4). The first row corresponds to the line intersection for two different values of \( \delta \); the second row corresponds to the conical intersection case for the intersection point \( a = (1.5, 0) \) inside and \( a = (4.5, 0) \) outside the classically allowed region. We note that the plots contain the probabilities to be in the excited state, \( p_{ex} \), for 20 and 1000 eigenvalues for line and conical intersection cases, respectively.
Figure 16. Convergence plots for $\hat{p}_{ex}$ obtained for the hyperplane sampling and molecular dynamics simulation methods in Algorithm 1 applied to the conical intersection problem.
5. Regularity and Weyl calculus

This section establishes four lemmas where the last three are used in Theorem 2.1 and 2.2. The first lemma proves the bound \( \frac{2}{15} \) on derivatives of the solution \( u(t, Z) = \mathbb{E}[g \circ S_T(Z)] \) to \( (2.13) \), which verifies that assumption (v) in Theorem 2.1 holds for a case with a near crossing of eigenvalues. The second lemma derives \( L^2 \)-bounds on Weyl quantized operators obtained from mollified symbols, the third and fourth lemma estimate remainder terms related to the Weyl quantization and establishes the Moyal expansion for the setting used in the proofs of Theorem 2.1 and Theorem 2.2.

We will use equation \((2.12)\) written as the Ito equation

\[
dZ_t = a(Z_t)dt + b(Z_t)d\hat{W}_t
\]

for the molecular dynamics path \( Z : [0, T] \times \Omega \to \mathbb{R}^{6N} \), where \( \Omega \) is the sample space and assume that the Jacobians \( a' \) and \( b' \) and the higher order derivatives \( \partial^2_a u(Z) \in \mathbb{R}^{(6N)^{(\alpha+1)}} \), with \( \alpha = n \) for \( n = 1, 2, 3, \ldots \), satisfy the \( \ell^2 \) bounds

\[
\|a'\|_{2,\infty} + \|b'\|_{2,\infty} = \mathcal{O}(\delta^{-1}), \\
\|\partial^2_a a\|_{2,\infty} + \|\partial^2_a b\|_{2,\infty} = \mathcal{O}(\delta^{-|\alpha|})
\]

uniformly in \( N \) (but not in \( |\alpha| \)). Here we use the \( \ell^2 \) (Frobenius) norm: assume that \( |\alpha| = n \) then

\[
\|a(Z_t)\|_{2,\infty} := \sup_{\omega \in \Omega} \|a(Z_t(\omega))\|_2.
\]

**Lemma 5.1.** Assume \((5.2)\) and that the near crossing is weak, with non vanishing velocity \( P \) across the near crossing, namely that

\[
\max_{X \in \mathbb{R}^{3N}} \|\lambda^\alpha_n(X)\|_2 \leq c_3 \delta^{-1}, \\
\int_0^T \|\lambda^\alpha_n(X_t)\|_{2,\infty} dt \leq C_T,
\]

\[
\int_0^T (\mathbb{E}[\|\partial^2_X \lambda^\alpha_n(X_t)\|_{2^n}^2])^{1/(2^n)} dt \leq C \alpha T \delta^{-|\alpha|},
\]

for a constant \( c_3 \leq 1/\log \delta^{-1} \) and constants \( C_\alpha \) independent of \( \delta, M, N \) and \( \epsilon \). Then for each multi index \( \beta \) there is a constant \( \tilde{C} \), independent of \( M, \delta, N \) and \( \epsilon \), such that

\[
\sup_{Z_0 \in \mathbb{R}^{6N}} \|\partial^\beta_Z \mathbb{E}[g \circ S_T(Z_0) | Z_0]\|_2 \leq e^{\tilde{C}(T+1)} \delta \min(0, -|\beta|+1).
\]

**Proof.** Remark \(5.2\) below motives the assumptions. First we study solutions \( w : [0, T] \times \Omega \to \mathbb{R}^{6N} \) to the linearized Ito equation

\[
dw_t = a'(Z_t)w_t dt + b'(Z_t)w_t d\hat{W}_t + \alpha_t dt + \beta_t d\hat{W}_t
\]

where \( Z_t \) solves \((5.1) = (2.12)\) and the stochastic processes \( \alpha : [0, T] \times \Omega \to \mathbb{R}^{6N} \) and \( \beta : [0, T] \times \Omega \to \mathbb{R}^{6N} \times \mathbb{R}^{6N} \) satisfy for each \( n \in \mathbb{N} \ \setminus \ 0 \) and some positive number \( m \) the bound

\[
\int_0^t \mathbb{E}[\|\beta_s\|_{2^n}^2 | Z_0] ds = O(e^{C_1 t}), \\
\int_0^t (\mathbb{E}[\|\alpha_s\|_{2^n}^2 | Z_0])^{1/(2^n)} ds = O(e^{C_1 \delta^{-m}}).
\]

Ito’s formula implies

\[
d\mathbb{E}[\|w_t\|_2^2 | Z_0] = 2\mathbb{E}[w_t \cdot dw_t | Z_0] + \mathbb{E}[dw_t \cdot dw_t | Z_0] \\
\leq 2\mathbb{E}[a'(Z_t)_{|z_2} w_t \cdot |Z_0] dt + \mathbb{E}[b'(Z_t)_{|z_2} w_t \cdot |Z_0] dt \\
+ 2\mathbb{E}[\|w_t\|_2 \alpha_t | Z_0] dt + \mathbb{E}[\|\beta_t\|_2 | Z_0] dt + 2\mathbb{E}[\|\beta_t\|_2 | Z_0] dt + \mathbb{E}[\|a'(Z_t)_{|z_2} \|_2 + \|b'(Z_t)_{|z_2} \|_2 | Z_0] dt
\]

where by assumption \((5.2)\) we have \( 2\|a'_s\|_{2,\infty} + \|b'_s\|_{2,\infty} = \mathcal{O}(\delta^{-1} + \epsilon \delta^{-2}) = \mathcal{O}(\delta^{-1}) \) and by \((5.3)\) there holds \( \int_0^t 2\|a'_s\|_{2,\infty} + \|b'_s\|_{2,\infty} ds = O(t) \). Here \( \|b'_s\|_{2,\infty} = \sum_i \|b'_i\|_{2,\infty} \). We use \( \epsilon \ll \delta \), with the arbitrarily...
small diffusion constant $\epsilon$ in (2.10), so that the contribution from $b'$ and $\beta$ are negligible compared to those from $a'$ and $\alpha$. Gronwall’s inequality shows that

$$
\mathbb{E}[||w_t||^2_2 \mid Z_0] \leq (\mathbb{E}[||w_0||^2_2 \mid Z_0] + \int_0^t \mathbb{E}[2||w_s||_2\|\alpha_s\|_2 + \|\beta_s\|_2^2 \mid Z_0] ds) e^{Ct} ||\nabla Z_t||^2_{2,\infty} ds
$$

which by (5.5) implies

$$
\sup_{0<s<t} \mathbb{E}[||w_s||^2_2 \mid Z_0] \leq e^{Ct} \left(\mathbb{E}[||w_0||^2_2 \mid Z_0] + \int_0^t 2\mathbb{E}[||\alpha_s||^2_2 \mid Z_0] + \mathbb{E}[||\beta_s||^2_2 \mid Z_0] ds\right) e^{Ct}
$$

provided this holds initially for $t = 0$. Analogously for $n > 1$, we obtain by Hölder’s inequality

$$
d\mathbb{E}[||w_t||^{2n}_2 \mid Z_0] = n\mathbb{E}[||w_t||^{2(n-1)}_2 w_t \cdot dw_t + n(n-1)||w_t||^{2(n-2)}_2 \left<\frac{w_t \cdot dw_t}{2}\right> + n||w_t||^{2(n-1)}_2 dw_t \cdot dw_t \mid Z_0]
$$

so that

$$
\sup_{0<s<t} \mathbb{E}[||w_s||^{2n}_2 \mid Z_0] \leq e^{Ct} \left(\mathbb{E}[||w_0||^{2n}_2 \mid Z_0] + \int_0^t 2\mathbb{E}[||\alpha_t||^{2n}_2 \mid Z_0] ds + O(e^{Ct})\right)
$$

provided this holds initially for $t = 0$.

The first variation $Z'_t = \partial Z_t / \partial Z_0$ satisfies

$$
dZ'_t = a'(Z_t) Z'_t dt + b'(Z_t) Z'_t d\tilde{W}_t,
$$

so that (5.7), for $m = 0$, implies

$$
\mathbb{E}[||Z'_t||^{2n}_2 \mid Z_0] = O(e^{Ct}).
$$

The second variation satisfies

$$
dZ''_t = a''(Z_t) Z''_t dt + b''(Z_t) Z''_t d\tilde{W}_t + a'''(Z_t) Z'_t Z'_t dt + b'''(Z_t) Z'_t Z'_t d\tilde{W}_t.
$$

We have for $\alpha = a'' Z' Z'$ that

$$
||a'' Z' Z'||^2_2 = \sum_{ijk} a''_{iqr} Z'_i Z'_j Z''_{qk}^2
$$

and similarly for $\beta$, which combined with Ito’s formula and Gronwall’s inequality as above imply

$$
\sqrt{\mathbb{E}[||Z''_t||^2_2 \mid Z_0]} = O(\delta^{-1} e^{Ct})
$$

using (5.3), that (5.8) yields $\mathbb{E}[||Z''_t||^2_2 \mid Z_0]^{1/4} = O(e^{Ct})$ and

$$
\int_0^t \sqrt{\mathbb{E}[||a'' Z' Z''_t||^2_2 \mid Z_0]} ds \leq \int_0^t \left(\mathbb{E}[||a'' Z' Z''_t||^2_2 \mid Z_0]^{1/4} (\mathbb{E}[||Z''_t||^2_2 \mid Z_0])^{1/4}\right) ds \leq e^{Ct} O(\delta^{-1}).
$$

Also the higher variations satisfy

$$
dZ^{(m)}_t = a'(Z_t) Z^{(m)}_t dt + b'(Z_t) Z^{(m)}_t d\tilde{W}_t + \alpha_t dt + \beta_t d\tilde{W}_t
$$
Remark 5.2. which by (5.10) proves (5.4).

Similarly, an assumption (5.11) provided the near crossing of potential surfaces is located away from the part of \( \Sigma_E \) where the stochastic processes \( \alpha \) and \( \beta \) have the bound (5.5), since following (5.9) we have
\[
\|a'' Z'' Z'' \|_2^2 \leq \|a'' \|_2^2 \|Z'' \|_2^2, \\
\|a'' Z'' Z'' \|_2^2 \leq \|a'' \|_2^2 \|Z'' \|_2^2,
\]
and similarly for the higher order derivatives. Therefore (5.7) and (5.8) yield (5.10)
\[
(E [\|Z_t^{(m)} \|_2^2 | Z_0])^{1/(2n)} = O(e^{Ct \delta(1-m)}).
\]

We have by Jensens inequality, for \( |\beta| = n \),
\[
\sup_{Z_0} \|E [\partial_{Z_0}^2 g(Z_t) | Z_0] \|_2 \leq \sup_{Z_0} E [\|\partial_{Z_0}^2 g(Z_t) \|_2 | Z_0] \\
\leq \sup_{Z_0} \sqrt{E [\|\partial_{Z_0}^2 g(Z_t) \|_2^2 | Z_0]} = \begin{cases} \\
\sup_{Z_0} \sqrt{E [\|g_t Z_t'' \|_2^2 | Z_0]}, & n = 1 \\
\sup_{Z_0} \sqrt{E [\|g_t Z_t'' + g_t Z_t Z_t'' \|_2^2 | Z_0]}, & n = 2 \\
\sup_{Z_0} \sqrt{E [\|g_t Z_t'' + 2 g_t Z_t Z_t'' + g_t Z_t Z_t'' Z_t'' \|_2^2 | Z_0]}, & n = 3
\end{cases}
\]
which by (5.10) proves (5.4). \( \square \)

Remark 5.2. We motivate assumption (5.3) by the behavior for the example of the avoided conical intersection eigenvalues \( \lambda_0(X) = \pm c \sqrt{X^2 + \delta} \) for \( X \in \mathbb{R}^2 \) or \( X \in \mathbb{R}^4 \) discussed in Section 4.2 and (3.14). Although \( \max_{X \in K} \|\lambda_0''(X)\|_2 = O(\delta^{-1}) \), we assume the near crossing is weak, i.e. that \( \|\lambda_0''(X)\|_2 \leq c_3 (\text{dist}(X, a)^2 + \delta^2) \) for some constant \( c_3 \leq 1/\log \delta^{-1} \) and a codimension 2 set \( a \subset \mathbb{R}^{3N} \). Then the exponent in Gronwall’s inequality has the bound
\[
\int_0^T \|\lambda_0''(X_s)\| ds \leq C(T - t) + c_3 \int_0^1 \frac{ds}{C s^2 + \delta^2} \leq C(T - t) + c_3^{-1} c_3 \log \delta^{-1}
\]
provided the near crossing of potential surfaces is located away from the part of \( \Sigma_E \) where \( |P| = 0 \). Here we assume that the velocity through the avoided crossing domain is bounded from below by \( c > 0 \). Similarly, an assumption \( \|\partial_{X}^2 \lambda_0''(X)\|_2 \leq C (\text{dist}(X, a)^2 + \delta^2)^{-1/2} \) satisfies the last estimate in (5.3).

We will use the s-quantization and the notation
\[
\int_{\mathbb{R}^{3N}} \langle \Phi(X), O_{\hat{\phi}}(h) \Phi(X) \rangle dX = \int_{\mathbb{R}^{1/2}} h(Z) \cdot W^{(s)}(Z) dZ
\]
for any smooth symbol \( h \) and the Wigner functions \( W = W^{(1/2)} \) and \( W^{(s)} \) defined in (2.6).

Lemma 5.3. Assume that \( h : \mathbb{R}^{3N} \to \mathbb{C} \) and \( f : \mathbb{R}^{3N} \to \mathbb{C} \) can be written as \( h = \hat{h} \ast \phi_\eta \) and \( f = \hat{f} \ast \phi_\eta \),
where \( \phi_\eta(z) := (2\pi \eta)^{-3N} e^{-z^2/(2\eta)} \), and that \( h \) and \( f \) have continuous second derivatives with polynomial growth, i.e. for some \( n \in \mathbb{Z} \) and \( m \in \mathbb{Z} \) and \( |\alpha| \leq 2 \) there is a constant \( C \) such that \( \partial_\alpha^2 h(z) \leq C(1 + |z|^2)^m \) and \( \partial_\alpha^2 f(z) \leq C(1 + |z|^2)^n \) uniformly in \( z \), then for every \( s \in [0, 1] \)
\[
| \int_{\mathbb{R}^{3N}} \langle \Phi(X), O_{\hat{\phi}}(h(z) \phi_\eta) \Phi(X) \rangle dX | = \sup_{z \in \mathbb{R}^{3N}} | h(z) \phi_\eta(z) | + O(\eta),
\]
uniformly in \( N \).

Proof. The proof to estimate the \( L^2(\mathbb{R}^{3N}) \) operator norm for a symbol \( r \) has three steps: to determine a representation of the solution of \( \int_{\mathbb{R}^{3N}} \langle \Phi(X), O^* r(\hat{\phi}) \Phi(X) \rangle dX \) using the FBI transform \( T \Phi \), to calculate a Fourier multiplier applied to \( r \) using the representation and in the third step to verify that the product of two FBI transformed functions is the convolution of the Wigner function with a Gaussian.

1. We will use that the components of the Wigner function \( W = W^{(1/2)} \) convolved with the Gaussian \( \phi_{M^{-1/2}}(X) = (2\pi M^{-1/2})^{-3N} e^{-|X|^2/(2M^{1/2})} \) is the product of the FBI transforms (3.36) of the components \( \Phi_i \) and \( \Phi_j \)
\[
(5.11) \ W_{ij} \ast \phi_{M^{-1/2}}(X, P) = T \Phi_i(X, P) T \Phi_j(X, P)
\]
as verified in Step 3 below. Consequently, the diagonal entries are non negative
\[
W_{ii} \ast \phi_{M^{-1/2}}(X, P) = |T \Phi_i(X, P)|^2 \geq 0.
\]

An s-quantization remainder \( r^{s*} \) can be related to a remainder \( r = \text{Op}[h] \) in the Wigner quantization (with \( s = 1/2 \)) by
\[
\int_{\mathbb{R}^{3N}} r(X, P) \cdot W^{(s*)}(X, P) dX dP = \int_{\mathbb{R}^{3N}} r^{s*}(X, P) \cdot W^{(1/2)}(X, P) dX dP
\]
where for $s_\ast \in [0,1]$
\[ r^{s_\ast}(X,P) = e^{-iM^{-1/2}(s_\ast - \frac{1}{2})\nabla_N \cdot \nabla}_R X, P, \]
which is proved in [22] Remark 2.7.3 and related to the expansion of the exponential in [5.18].

We will introduce the convolution with the Gaussian in the estimate of the remainder using the Taylor expansion of the exponential
\[ \int_{\mathbb{R}^N} (\Phi, r^{s_\ast} \Phi) dX = \int_{\mathbb{R}^N} r^{s_\ast}(X,P) \cdot W^{(1/2)}(X,P) dXdP \]
\[ = \int_{\mathbb{R}^N} \mathcal{F}(r^{s_\ast})(\xi,\xi_P) \cdot \mathcal{F}W^{(1/2)}(\xi,\xi_P) d\xi d\xi_P \]
\[ = \int_{\mathbb{R}^N} (\mathcal{F}^{\phi_{M-1/2}})(\xi,\xi_P) \cdot \mathcal{F}r^{s_\ast}(\xi,\xi_P) \mathcal{F}W^{(1/2)}(\xi,\xi_P) d\xi d\xi_P \]
\[ = \int_{\mathbb{R}^N} r^{\ast}(X,P) \cdot W^{(1/2)} \ast \phi_{M-1/2}(X,P) dXdP, \]
where the function $r^{\ast}$ is defined by $r^{\ast} = \mathcal{F}^{-1}\{(\mathcal{F}^{\phi_{M-1/2}})(\xi,\xi_P)\}^{-1} \mathcal{F}r^{s_\ast}(\xi,\xi_P)$. Here we use the notation $r \cdot W := \sum_{ij} r_{ij}W_{ij}$. The tensor product property [5.11] of $W \ast \phi_{M-1/2}$ implies the matrix norm estimate
\[ \int_{\mathbb{R}^N} (\Phi(X), r^{\ast} \Phi(X)) dX \leq \int_{\mathbb{R}^N} r^{\ast}(X,P) \cdot W^{(1/2)} \ast \phi_{M-1/2}(X,P) dXdP \]
\[ = \int_{\mathbb{R}^N} (T\Phi(X,P), r^{\ast}(X,P)T\Phi(X,P)) dXdP \]
\[ \leq \sup_{(X,P) \in \mathbb{R}^N} \|r^{\ast}(X,P)\|_2 |T\Phi(X,P)|^2 dXdP \]
\[ \leq \sup_{(X,P) \in \mathbb{R}^N} \|r^{\ast}(X,P)\|_2 \int_{\mathbb{R}^N} (T^{*}T\Phi, \Phi) dX \]
\[ = \sup_{(X,P) \in \mathbb{R}^N} \|r^{\ast}(X,P)\|_2 \int_{\mathbb{R}^N} \Phi(X) \Phi(X) dX \]
\[ = \sup_{(X,P) \in \mathbb{R}^N} \|r^{\ast}(X,P)\|_2 \int_{\mathbb{R}^N} \Phi(X) \Phi(X) dX = \int_{\mathbb{R}^N} \Phi(X) \Phi(X) dX = 1, \]
using the $L^2$ identity $T^{*}T = I$ in $[3.37]$ for the FBI transform and the $l^2$ matrix norm $\|r^{\ast}(X,P)\|_2$ (or the less sharp Frobenius norm).

2. We have $\mathcal{F}r^{\ast} = e^{i|x|^2M^{-1/2}/2} e^{i\omega_x \cdot \omega_p}M^{-1/2} \mathcal{F}r(\omega)$. The goal is to absorb this exponentially growing pre-factor $e^{i|x|^2M^{-1/2}/2} e^{i\omega_x \cdot \omega_p}M^{-1/2}$ in $\mathcal{F}r(z)$.

Consider first the case $r = r^{\ast} \ast \phi_{\eta}$, then $\mathcal{F}r^{\ast} = e^{i|x|^2M^{-1/2}/2} e^{i\omega_x \cdot \omega_p}M^{-1/2} \mathcal{F}\phi_{\eta}(\omega) \mathcal{F}r(\omega)$. We have $\mathcal{F}\phi_{\eta}(\omega) = e^{-|\omega|^2\eta^2/2}$ so that
\[ e^{i|x|^2M^{-1/2}/2} e^{-|\omega|^2\eta^2/2} = e^{-|\omega|^2\nu^2/2} = \mathcal{F}\phi_{\nu}(\omega) \]
where $\nu = \eta - M^{-1/2} > 0$. The inverse Fourier transform of the pre factor $e^{i\omega_x \cdot \omega_p}M^{-1/2} \mathcal{F}\phi_{\nu}(\omega)$ is
\[ (2\pi)^{-6N} \int_{\mathbb{R}^N} e^{-\nu|\omega_x|^2/2 - \nu|\omega_p|^2/2 + i\omega_x \cdot \omega_p M^{-1/2} + i\omega_x \cdot x + i\omega_p \cdot p} d\omega_x dp \]
\[ = e^{-\frac{(\nu x + \frac{1}{2})^2}{2(\nu + i\frac{1}{2})^2}} e^{-i\nu x M^{-1/2} + \frac{1}{\nu} + \frac{1}{2} \int_{\mathbb{R}^N} e^{-\nu|\omega_x|^2/2 + \nu|\omega_p|^2/2 + i\omega_x \cdot \omega_p M^{-1/2} + i\omega_x \cdot x + i\omega_p \cdot p} d\omega_x dp \]
\[ = \phi_{\nu,s}, \]
so that $r^{\ast} = r^{\ast} \ast \phi_{\nu,s}$. The estimate (5.12) implies then in the case $r = r^{\ast} \ast \phi_{\eta}$
\[ \int_{\mathbb{R}^N} (\Phi(X), r^{\ast} \Phi(X)) dX \leq \sup_{z \in \mathbb{R}^N} \|r^{\ast}(z)\|_2 \]
\[ = \sup_{z \in \mathbb{R}^N} \|r \ast \phi_{\nu,s}(z)\|_2. \]
If $r$ is uniformly bounded in $C^2(\mathbb{R}^N)$ we have by the two moments $\int_{\mathbb{R}^N} \phi_{\eta}(z) dz = 1$ and $\int_{\mathbb{R}^N} z\phi_{\eta}(z) dz = 0$ that
\[ r - \tilde{r} = O(\eta) \]
and by the properties \( \int_{\mathbb{R}^N} \phi_{v,s}(z)dz = 1 \) and \( \int_{\mathbb{R}^N} z\phi_{v,s}(z)dz = 0 \) we obtain
\[
\tilde{r} \ast \phi_{v,s} - \tilde{r} = O(\eta)
\]
so that
\[
(5.13) \quad \int_{\mathbb{R}^N} \langle \Phi(X), \tilde{r}^\text{as} \Phi(X) \rangle dX = \sup_{z \in \mathbb{R}^N} |r(z)| + O(\nu).
\]
In the case \( r = (h \ast \phi_\eta)(\tilde{f} \ast \phi_\eta) \) we have similarly
\[
(2\pi)^{-12N} \int_{\mathbb{R}^{12N}} \mathcal{F}h(\omega) \mathcal{F}\tilde{f}(\omega - \omega') e^{-|\omega|^2/2} e^{i\omega \cdot \omega} e^{i\omega \cdot \omega} e^{M^{-1/2} + |\omega|^2/(2M^{-1/2})} e^{i\omega \cdot z} d\omega' d\omega
\]
\[
= (2\pi)^{-12N} \int_{\mathbb{R}^{12N}} \mathcal{F}h(\omega) \mathcal{F}\tilde{f}(\omega - \omega') e^{-|\omega|^2/2} e^{i\omega \cdot \omega} e^{M^{-1/2} + |\omega|^2/(2M^{-1/2})} e^{i\omega \cdot z} d\omega' d\omega
\]
\[
= (2\pi)^{-9N} \eta^{-3N} \int_{\mathbb{R}^{12N}} \mathcal{F}^{-1} \{ \mathcal{F}h \mathcal{F}(\tilde{f}(\omega - \cdot)) \} (z') e^{-|z'|^2/(4\eta)} + i\omega \cdot z' e^{i\omega \cdot z} e^{M^{-1/2} - |\omega|^2/(2M^{-1/2})} e^{i\omega \cdot z'} d\omega dz'
\]
\[
= (2\pi)^{-3N} (2\pi\nu)^{-3N} (1 + \frac{s^2}{\nu^2M})^{-3N/2} \times
\]
\[
\times \int_{\mathbb{R}^{12N}} \tilde{h}(z' - z'') \tilde{f}(-z'') e^{-|z'|^2/(4\eta)} e^{-|z-z''+z'/2|^2/(4\nu(s))} e^{(\frac{(z'' - z) + (z' - z'')}{2})^{2} + (\frac{z'' - z}{2})^{2} + M^{-1/2}} d\omega dz
\]
\[
= \sigma_{N,M} \int_{\mathbb{R}^{12N}} \tilde{h}(z) \tilde{f}(w) e^{-|w-v|^2/(4\eta)} e^{-|z+w+v/2|^2/(4\nu(s))} e^{(\frac{(z + w + v)}{2})^{2} + (\frac{z + w + v}{2})^{2} + M^{-1/2}} d\omega dw
\]
\[
= \sigma_{N,M} \int_{\mathbb{R}^{12N}} \tilde{h}(z) \tilde{f}(z - \nu) e^{-|w-v|^2/(4\eta)} e^{-|z-w-v/2|^2/(4\nu(s))} e^{(\frac{z + w - v}{2})^{2} + (\frac{z + w - v}{2})^{2} + M^{-1/2}} d\omega dw
\]
\[
=: Q(h\tilde{f})(z).
\]
We have by the Fourier transform
\[
\sigma_{N,M} \int_{\mathbb{R}^{12N}} e^{-|w-v|^2/(4\eta)} e^{-|z-w-v/2|^2/(4\nu(s))} e^{(\frac{(z + w - v)}{2})^{2} + (\frac{z + w - v}{2})^{2} + M^{-1/2}} d\omega dw = 1
\]
\[
\int_{\mathbb{R}^{12N}} we^{-|w-v|^2/(4\eta)} e^{-|z-w-v/2|^2/(4\nu(s))} e^{(\frac{(z + w - v)}{2})^{2} + (\frac{z + w - v}{2})^{2} + M^{-1/2}} d\omega dw = 0
\]
\[
\int_{\mathbb{R}^{12N}} we^{-|w-v|^2/(4\eta)} e^{-|z-w-v/2|^2/(4\nu(s))} e^{(\frac{(z + w - v)}{2})^{2} + (\frac{z + w - v}{2})^{2} + M^{-1/2}} d\omega dw = 0
\]
so that
\[
|Q(h\tilde{f})(z) - \tilde{h}(z)\tilde{f}(z)| = |Q(h\tilde{f})(z) - \tilde{h}(z)\tilde{f}(z)| \leq C(||\tilde{h}(z) - \tilde{h}(\sqrt{2}z)||^2_{L^2(\mathbb{R}^N)} + ||\tilde{f}(z)(1 + |z|^2)||^2_{L^2(\mathbb{R}^N)})\nu(s)
\]
Similarly using the moments \( \int_{\mathbb{R}^N} \phi_{v,s}(z)dz = 1 \) and \( \int_{\mathbb{R}^N} z\phi_{v,s}(z)dz = 0 \) and the second derivatives \( ||\tilde{h}||^2_{L^2(\mathbb{R}^N)} + ||\tilde{f}||^2_{L^2(\mathbb{R}^N)} \leq C \) we obtain
\[
\tilde{h}(z)f(z) - \tilde{h}(z)\tilde{f}(z) = O(\eta),
\]
and we conclude that \( |Q(h\tilde{f})(z) - h(z)f(z)| = O(\eta) \) which combined with (5.12), as in [5.13], proves the theorem.

3. This step verifies (5.11) following the proof of Proposition 1 in [21]. We have
\[
W_{ij} \ast \phi_{M^{-1/2}}(X,P)
\]
\[
= (\pi M^{1/2} \times 2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} e^{iM^{1/2}Y \cdot P} \Phi_i(X' - \frac{Y}{2}) \Phi^*_j(X' + \frac{Y}{2}) e^{-\frac{1}{2} ||X' - X||^2 + \frac{1}{2} ||P - P'||^2} dY dX' dP'
\]
\[
= (\pi M^{1/2})^{-3N/2} (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} e^{iM^{1/2}Y \cdot P} \Phi_i(X' - \frac{Y}{2}) \Phi^*_j(X' + \frac{Y}{2}) e^{-\frac{1}{2} ||X' - X||^2 + \frac{1}{2} ||Y||^2} M^{1/2} dY dX'.
\]
The change of variables \( X' - \frac{Y}{2} = v \) and \( X' + \frac{Y}{2} = w \) implies
\[
M^{1/2}(||X - X'||^2 + \frac{1}{4} ||Y||^2) = \frac{M^{1/2}}{2} (||v - X||^2 + ||w - X||^2)
\]
and we conclude that the Wigner function and the FBI transform have the relation

\[
W_{ij} \ast \phi_{M^{-1/2}}(X, P) = (2^{1/3} \pi)^{-3N/4} M^{3N/8} \int_{\mathbb{R}^{3N}} e^{iM^{1/2}(X-v)} P e^{-|X-v|^2M^{1/2}} \Phi_i(v) dv \\
\times (2^{1/3} \pi)^{-3N/4} M^{3N/8} \int_{\mathbb{R}^{3N}} e^{-iM^{1/2}(X-w)} P e^{-|X-w|^2M^{1/2}} \Phi_j(w) dw \\
= T \Phi_i(X, P)(T \Phi_j(X, P))^*.
\]

\[\Box\]

**Lemma 5.4.** Assume that \( r : \mathbb{R}^{6N} \to \mathbb{C} \) is smooth and for each \( n \in \mathbb{N} \) and \( |\alpha| \leq n \) there is a constant \( C_n \) such that \( \sup_{x \in \mathbb{R}^{6N}} |\partial^\alpha r(x, P)| dP \leq C_n \), then there is a constant \( C \) such that

\[
\int_{\mathbb{R}^{3N}} \langle \Phi(X), Op^*[r]\Phi(X) \rangle dX \leq C \sup_{n \leq 3N+3} C_n.
\]

If for \( |\alpha| \leq n \) there is a constant \( C_n \) such that \( \sup_{x \in \mathbb{R}^{6N}} |\partial^\alpha r(z)| \leq C_n \delta^{\min(0,-n+1)} \), then there is a constant \( C \) such that

\[
\int_{\mathbb{R}^{3N}} \langle \Phi(X), Op^*[r]\Phi(X) \rangle dX \leq C (\sup_{z \in \mathbb{R}^{6N}} |r(z)| + M^{-1/4} \delta^{-1/2} \sup_{n \leq C_n} C_n).
\]

**Proof.** A proof if this is given e.g. in [54] (Theorem 4.21) and presented here for completeness.

The \( s \)-symbol \( r \) has the integral kernel

\[
K(X, Y) = \left( \frac{1}{2\pi M^{-1/2}} \right)^{3N} \int_{\mathbb{R}^{3N}} e^{iM^{1/2}(X-Y)} P r((1-s)X + sY, P) dP
\]

so that

\[
Op^*[r]\Phi(X) = \int_{\mathbb{R}^{3N}} K(X, Y)\Phi(Y) dY
\]

and

\[
\|Op^*[r]\Phi(X)\|^2_{L^2(\mathbb{R}^{3N})} = \int_{\mathbb{R}^{3N}} \mathcal{K}(Z, Y) K(Z, X) \Phi(Y) \Phi(X) dX dY dZ
\]

\[
\leq \frac{1}{2} \int_{\mathbb{R}^{3N}} |K(Z, Y)||K(Z, X)||\Phi(Y)||\Phi(X)||^2 dX dY dZ.
\]

Here

\[
\int_{\mathbb{R}^{3N}} |K(Z, Y)||K(Z, X)||\Phi(Y)||\Phi(X)||^2 dX dY dZ
\]

\[
\leq \sup_{Y \in \mathbb{R}^{3N}} \left( \int_{\mathbb{R}^{3N}} |K(Z, Y)||K(Z, X)||dX dZ \right) \int_{\mathbb{R}^{3N}} |\Phi(Y)||^2 dY
\]

\[
\leq \sup_{Y \in \mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} |K(Z, Y)||dZ \sup_{Z \in \mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} |K(Z, X)||dX
\]

and similarly for the term with \( Y \) and \( X \) replaced. We have

\[
K(X, Y) = (M^{1/2})^{-3N} \mathcal{F}^{-1} \{r((1-s)X + sY, \cdot)\}((X - Y)M^{1/2})
\]

and

\[
\int_{\mathbb{R}^{3N}} |K(X, Y)||dX = \int_{\mathbb{R}^{3N}} \mathcal{F}^{-1} \{r((1-s)X'M^{-1/2} + sY, \cdot)\}(X')|dX'
\]

\[
= \int_{\mathbb{R}^{3N}} (1 + |X'|^{3N+1}) \mathcal{F}^{-1} \{r((1-s)X'M^{-1/2} + sY, \cdot)\}(X') \frac{1}{1 + |X'|^{3N+1}} dX'
\]

\[
= \int_{\mathbb{R}^{3N}} \mathcal{F}^{-1} \{(1 + \Delta^{(3N+1)/2})r((1-s)X'M^{-1/2} + sY, \cdot)\}(X') \frac{1}{1 + |X'|^{3N+1}} dX'
\]

\[
\leq \|\mathcal{F}^{-1} \{(1 + \Delta^{(3N+1)/2})r((1-s)X'M^{-1/2} + sY, \cdot)\}(X')\|_{L^\infty} \int_{\mathbb{R}^{3N}} \frac{1}{1 + |X'|^{3N+1}} dX'
\]

\[
\leq C \sup_{Y'} \int_{\mathbb{R}^{3N}} \{(1 + \Delta^{(3N+1)/2})r(Y', P)\}|dP.
\]

where \( \Delta \) means the Laplacian with respect to the second variable in \( r \). The function \( r \) satisfies

\[
\sup_{Y' \in \mathbb{R}^{3N}} |\partial^\alpha r(Y', P)| dP \leq C_n \delta^{-|\alpha|} \text{ Theorem 5.1 in [54] applied to the symbol } r \text{ proves (5.14).}
\]

In the case \( \sup_{z} |\partial^\alpha r(z)| \leq C_n \delta^{-|\alpha|} \text{ Theorem 5.1 in [54] applied to the symbol } r \text{ proves (5.15).} \quad \Box\]
Lemma 5.5. The composition $\hat{C} = \hat{A}\hat{B}$ of two Fourier integral operators, with smooth symbols $A(X)$ and $B(X, P)$ in the Schwartz space, has the Weyl symbol

$$C(X, P) = e^{\frac{i}{\hbar}M^{-1/2}\sum_k (-\partial_{x_k} \partial_{p_k}) A(X)}B(X', P')\bigg|_{(X, P) = (X', P')}$$

(5.18)

$$= \sum_{n=0}^{\infty} \frac{(iM^{-1/2})^n}{2^n n!} (\nabla_X \cdot \nabla_{P'} - \nabla_X \cdot \nabla_{P})^n A(X)B(X', P')\bigg|_{(X, P) = (X', P')} + M^{-3/2}r_2.$$ 

The remainder $r_2$ is smooth and if $B(X_0, P_0) = \mathbb{E}[\tilde{g} \circ S_t(\cdot) \mid X_0, P_0] * \phi_0(X_0, P_0)$, and $A(X_0) = \tilde{A} * \phi_0(X_0)$ the remainder satisfies

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_2 \rangle dX = O(e^{\hbar t \delta^{-2}}).$$

If $A \in C^\infty(\mathbb{R}^{6N})$ and $B \in C^\infty(\mathbb{R}^{6N})$ are in the Schwartz space there holds

$$C(X, P) = e^{\frac{i}{\hbar}M^{-1/2}\sum_k (-\partial_{x_k} \partial_{p_k}) A(X)}B(X', P')\bigg|_{(X, P) = (X', P')}$$

(5.19)

$$= \sum_{n=0}^{\infty} \frac{(iM^{-1/2})^n}{2^n n!} (\nabla_X \cdot \nabla_{P'} - \nabla_X \cdot \nabla_{P})^n A(X)B(X', P')\bigg|_{(X, P) = (X', P')} + M^{-3/2}r_2,$$

where for $B(X_0, P_0) = \mathbb{E}[\tilde{g} \circ S_t(\cdot) \mid X_0, P_0] * \phi_0(X_0, P_0)$ and $A(X_0, P_0) = \tilde{A} * \phi_0(X_0, P_0)$

(5.20)

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_m \rangle dX = O(e^{\hbar t \delta^{-m}}),$$

for $A(X) = \nabla V(X) - \nabla \lambda_\eta * \phi_\nu(X)$ and $B(X, P_0) = \partial_P \mathbb{E}[\tilde{g} \circ S_t(\cdot) \mid X_0, P_0] * \phi_\eta(X_0, P_0)$

(5.21)

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_1 \rangle dX = O(e^{\hbar t \delta^{-4}}),$$

and for $A(Z) = H_\eta(Z)$ and $B(Z) = 2 \int_0^1 g_{MD}(sH_\eta(Z) + (1-s)E)(1-s)ds$

(5.22)

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_1 \rangle dX = O(e^{\hbar t \delta^{-4}}).$$

Proof. The proof has five steps. The first step uses the definition of the Weyl quantization to define an integral kernel for the product $\hat{C} = \hat{A}\hat{B}$ following Hörmander’s work [22]. The next step formulates the Moyal expansion and identifies the remainder as an average of $s$- quantizations (2.7), which proves (5.18). The third step writes the $s$-quantized remainder as a Wigner quantization (with $s = 1/2$) and estimates the remainder, using that the product of two FBI transformed functions is the convolution of the Wigner function with a Gaussian. Step 4 proves (5.20) and (5.21). Step 5 proves (5.23).

1. To verify (5.18) we start with the definition of the Weyl operator

$$\hat{C}\Phi(X) = \left(2\pi M^{-1/2}\right)^{-3N} \int_{\mathbb{R}^{3N}} C\left(\frac{X + Y}{2}, P\right) e^{i M^{1/2}(X - Y).P} \Phi(Y) dY dP.$$ 

The $L^2$ inner product

$$\int_{\mathbb{R}^{3N}} \hat{C}\Phi(X) \Psi^*(X) dX = \gamma \int_{\mathbb{R}^{3N}} C\left(\frac{X + Y}{2}, P\right) e^{i M^{1/2}(X - Y).P} \Phi(Y) \Psi^*(X) dP dY dX$$

defines the kernel

(5.23)

$$C_K(X, Y) := \gamma \int_{\mathbb{R}^{3N}} C\left(\frac{X + Y}{2}, P\right) e^{i M^{1/2}(X - Y).P} dP,$$

and the inverse Fourier transform implies

$$C(U, P) = \int_{\mathbb{R}^{3N}} C_K(U + \frac{Z}{2}, U - \frac{Z}{2} e^{-i M^{1/2}Z}.P) dZ.$$ 

Our examples have $A$ independent of $P$, i.e. $A(X, P) = A(X)$, and the kernel of the composition $\hat{A}\hat{B}$ becomes a multiplication

$$C_K(X, Y) = \gamma \int_{\mathbb{R}^{3N}} A(X) B\left(\frac{X + Y}{2}, P\right) e^{i M^{1/2}(X - Y).P} dP.$$
The definition \( \tilde{C} = \tilde{A} \tilde{B} \) and \((5.25)\) yields

\[
C(U, P) = (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} A(U + \frac{Z}{2}) B(U, P + P') e^{iM^{1/2}P'R} dZ dP' \\
= (\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} A(U + Z') B(U, P + P') e^{2iM^{1/2}P'R} dZ' dP'.
\]

The remainder term can be written

\[
(5.26)
\]

The remainder is based on a Taylor expansion of the exponential function as follows

\[
\text{2. We use in this step the Fourier transform } F_P \text{ in the } P\text{-direction defined for } f(X_0, P_0) \in \mathbb{R} \text{ by}
\]

\[
F_P[f](X_0, Y) := \int_{\mathbb{R}^{3N}} f(X_0, \xi) e^{-i\xi Y} d\xi.
\]

The remainder can be written on a Taylor expansion of the exponential function as follows

\[
(5.26)
\]

The remainder term can be written

\[
(\frac{1}{2\pi})^{6N} \int_{\mathbb{R}^{3N}} \mathcal{F} f(\xi X, \xi P) e^{-\frac{i}{2}M^{-1/2}\xi X} \xi d\xi X d\xi P
\]

\[
= (\frac{1}{2\pi})^{6N} (\pi M^{-1/2})^{3N} \int_{\mathbb{R}^{3N}} \mathcal{F} f(\xi X, \xi P) \left( \sum_{n=0}^{m} \frac{(-\xi X \cdot \xi P)^n}{2M^{1/2}} \frac{1}{n!} \right)
\]

\[
+ \left( \frac{1}{2M^{1/2}} \right)^{m+1} \int_{0}^{1} \int_{\mathbb{R}^{3N}} F((-i\nabla_X \cdot \nabla P)^{m+1} f) (\xi X, \xi P) e^{-\frac{i}{2}M^{-1/2}\xi X} \xi d\xi X d\xi P d\xi P
\]

\[
= \left( \frac{1}{2M^{1/2}} \right)^{m+1} \int_{0}^{1} \int_{\mathbb{R}^{3N}} F_P (-i\nabla_X \cdot \nabla P)^{m+1} f(X + \cdot, P + \cdot) \left( \frac{s\xi P}{2M^{1/2}} \xi P \right) \left( \frac{1}{2\pi} \right)^{3N} \xi d\xi P d\xi P d\xi P
\]

where \( r_m \) is smooth since \( f \) is a Schwartz function. The change of variables \( P + P' = P'' \) and \((5.24)\) shows that kernel of the remainder becomes

\[
r_m, K(X, Y) = \gamma \int_{0}^{1} \int_{\mathbb{R}^{3N}} (-i\nabla_X \cdot \nabla P)^{m+1} f \left( \frac{X + Y}{2} + \frac{s\xi P}{2M^{1/2}} \xi P \right)
\]

\[
\times e^{iM^{1/2}(X-Y) \cdot P'} e^{i(P'' - P) \cdot \xi P} \left( \frac{1}{2\pi} \right)^{3N} \xi d\xi P d\xi P d\xi P
\]

\[
= \gamma \int_{0}^{1} \int_{\mathbb{R}^{3N}} (-i\nabla X \cdot \nabla P)^{m+1} f \left( \frac{X + Y}{2} + \frac{s(X-Y)}{2M^{1/2}} \right)\xi d\xi P d\xi P d\xi P
\]

\[
= \gamma \int_{0}^{1} \int_{\mathbb{R}^{3N}} F_P (\xi P)^{m+1} f \left( \frac{X + Y}{2} + \frac{s(X-Y)}{2M^{1/2}(X-Y)} \right) \left( \frac{1}{m} \right) \xi d\xi P d\xi P d\xi P
\]
We note that the kernel \( A^s_K(X, Y) \) for any symbols \( A \) in the \( s_* \)-quantization \(^{2.7}\) is

\[
A^s_K(X, Y) = \gamma \int_{\mathbb{R}^{3N}} A(X + s_*(Y - X), P) e^{iM^{1/2}(X - Y) \cdot P} dP
\]

\[
= \gamma \mathcal{F}_P \{ A \} \{ X + s_*(Y - X), M^{1/2}(X - Y) \} = \gamma \mathcal{F}_P \{ A \} \{ \frac{X + Y}{2} + \frac{s}{2}(Y - X), M^{1/2}(X - Y) \}
\]

for \( s_* = (1 - s)/2 \). Consequently we have

\[
(5.27) \quad \int_{\mathbb{R}^{3N}} (\Phi, \hat{r}_m \Phi) dX = \int_0^1 \int_{\mathbb{R}^{6N}} (\nabla_X \cdot \nabla_P)^{m+1} f(X, P) W^{(s_*)}(X, P) dXdP \left( \frac{1 - s}{m!} \right)^m ds
\]

and the next step shows the final bound

\[
(5.28) \quad \int_{\mathbb{R}^{3N}} (\Phi, \hat{r}_2 \Phi) dX = O(e^{C_T \delta^{-2}}).
\]

3. It remains to estimate the remainder for \( r = r_2 \) in \(^{5.28}\). Lemma \(^{5.3}\) and Lemma \(^{5.1}\) imply that we have

\[
\sup_{(X, P) \in \mathbb{R}^{6N}} ||r^*(X, P)||_2 \leq C \sup_{(X, P) \in \mathbb{R}^{6N}} ||r_2(X, P)||_\infty + O(\eta) = O(e^{C_T \delta^{-2}}),
\]

uniformly in \( N \), which proves \(^{5.28}\).

4. The general case \( \hat{C} = \hat{A} \hat{B} \) of two Fourier integral operators, with smooth symbols \( A, B \in C^\infty(\mathbb{R}^6) \) in the Schwartz space yields

\[
C(U, P) = \gamma^2 \int_{\mathbb{R}^{12N}} A(U + Z, P) B(U + Z, P) e^{iM^{1/2}F} dP dP' dZ dT
\]

where

\[
F := (U - Z + \frac{T}{2}) \cdot P'' + (Z - U + \frac{T}{2}) \cdot P' - T \cdot P
\]

\[
= (U - Z + \frac{T}{2}) \cdot (P'' - P) + (Z - U + \frac{T}{2}) \cdot (P' - P).
\]

The change of variables \( (P'' - P, P' - P, (Z - U + \frac{T}{2})/2, (Z - U - \frac{T}{2})/2) \) replacing \( (P'', P', Z, T) \) has the Jacobian \( 2^{6N} \) and implies

\[
C(U, P) = (\pi M^{-1/2})^{-6N} \int_{\mathbb{R}^{12N}} A(U + Z, P + P'') B(U + T, P + P'') e^{iM^{1/2}(P' - T \cdot P'')} dZ dP dP'.
\]

The same expansion as in \(^{5.26}\) shows that

\[
C(X, P) = e^{i \frac{2M^{1/2}}{\pi} \sum_k (\partial_X \partial_{P'} - \partial_X \partial_{P''}) A(X, P) B(X', P')} \bigg|_{(X, P) = (X', P')}
\]

\[
= \sum_{n=0}^{m} \frac{(iM^{-1/2})^n}{2^n n!} \left( \nabla_P' \cdot \nabla_X' - \nabla_X \cdot \nabla_P' \right)^n A(X, P) B(X', P') \bigg|_{(X, P) = (X', P')} + M^{-(m+1)/2} r_m
\]

based on the representation

\[
\int_{\mathbb{R}^N} f(X + X', P + P', X + X'', P + P'') e^{2iM^{1/2}(X' - X'' \cdot P')} dX' dP' dX'' dP''
\]

\[
= \left( \frac{\pi}{M^{1/2}} \right)^{3N} \sum_{n=0}^{m} \frac{1}{2^n n!} \left( i(\nabla_X' \cdot \nabla_{P'} - \nabla_X \cdot \nabla_{P''}) \right)^{n+1} f(X + X', P + P', X + X'', P + P'') \bigg|_{X' = X'' = 0}
\]

\[
= \left( \frac{\pi}{M^{1/2}} \right)^{3N} \sum_{n=0}^{m} \frac{1}{2^n n!} (1 - s)^m dP' dP'' \frac{d\xi_P}{(2\pi)^{6N} m!} \delta_P \delta_{P''} ds.
\]

\[
r_m := \int_0^1 \int_{\mathbb{R}^{12N}} (i(\nabla_X' \cdot \nabla_{P'} - \nabla_X \cdot \nabla_{P''}) \right)^{n+1} f(X + \frac{s\xi_P}{2M^{1/2}}, P + P', X - \frac{s\xi_{P'}}{2M^{1/2}}, P + P'')
\]

\[
\times e^{i(P'' \cdot \xi_{P'} - P' \cdot \xi_{P'')} (1 - s)^m} dP' d\xi_P dP'' ds.
\]
The kernel for \( m = 1 \) becomes
\[
\begin{align*}
\rho_{1,k}(X,Y) & = \gamma \int_0^1 \mathcal{F}_{p'} p'' \left\{ \left| i (\nabla X'' \cdot \nabla p'' - \nabla X' \cdot \nabla p') \right|^2 \right\} \\
& \quad \times \left( \frac{X+Y}{2} - \frac{s(X-Y)}{2M^{1/2}}, M^{1/2}(X-Y), \frac{X+Y}{2} - \frac{s(X-Y)}{2M^{1/2}}, M^{1/2}(X-Y) \right) \left( 1-s \right) ds . \end{align*}
\]
and as in (5.27) we obtain
\[
\int_{\mathbb{R}^N} \langle \Phi, \hat{r}_1 \Phi \rangle dX = O(e^{CT} \delta^{-4}) .
\]
\[\square\]

6. Appendix: Fourier integral WKB states including caustics

6.1. A preparatory example with the simplest caustic. As an example of a caustic, we study first the simplest example of a fold caustic based on the Airy function \( A : \mathbb{R} \to \mathbb{R} \) which solves
\[
\partial_{xx} A(x) + x A(x) = 0 .
\]
The scaled Airy function
\[
u(x) = C A(M^{1/3} x)
\]
solves the Schrödinger equation
\[
- \frac{1}{M} \partial_{xx} u(x) + x u(x) = 0 ,
\]
for any constant \( C \). In our context an important property of the Airy function is the fact that it is the inverse Fourier transform of the function
\[
\hat{A}(p) = \sqrt{\frac{2}{\pi}} e^{ip^3/3} ,
\]
i.e.,
\[
A(x) = \frac{1}{\pi} \int_{\mathbb{R}} e^{ixp + p^3/3} \, dp .
\]
In the next section, we will consider a general Schrödinger equation and determine a WKB Fourier integral corresponding to (6.3) for the Airy function; as an introduction to the general case we show how the derive (6.3): by taking the Fourier transform of the ordinary differential equation (6.1)
\[
0 = \int_{\mathbb{R}} (-\partial_{xx} + x) \Lambda(x) e^{-ixp} \, dx = \left( p^2 + i \partial_p \right) \hat{A}(p) ,
\]
we obtain an ordinary differential equation for the Fourier transform \( \Lambda(p) \) with the solution \( \Lambda(p) = C e^{ip^3} \), for any constant \( C \). Then, by differentiation, it is clear that the scaled Airy function \( u \) solves (6.2). Furthermore, the stationary phase method, cf. Section 7, shows that to the leading order \( u \) is approximated by
\[
u(x) \simeq C \left( -x M^{1/3} \right)^{-1/4} \cos \left( M^{1/2}(-x)^{3/2} - \pi/4 \right) , \quad \text{for } x < 0,
\]
and \( u(x) \simeq 0 \) to any order (i.e., \( O(M^{-K}) \) for any positive \( K \)) when \( x > 0 \). The behaviour of the Airy function is illustrated in Figure 17.

![Airy function graph](image)

**Figure 17.** The Airy function.

### 6.1.1. Molecular dynamics for the Airy function.

The eikonal equation corresponding to (6.2) is

\[
p^2 + x = 0
\]

with solutions for \( x \leq 0 \), which leads to the phase

\[
p = \theta'(x) = \pm (-x)^{1/2}, \quad \theta(x) = \mp \frac{2}{3}(-x)^{3/2}.
\]

We have the Legendre transform

\[
\theta^*(p) = \min_y (yp - \theta(y))
\]

with the solution \( p = \theta'(y) \). Consequently, we obtain by (6.6) that \( x = y \) and

\[
\theta^*(p) = xp - \theta(x) = -p^2p + \frac{2}{3}p^3 = -\frac{p^3}{3}.
\]

We note that this solution is also obtained from the eikonal equation

\[
p^2 + \partial_p \theta^*(p) = 0,
\]

which is solved by

\[
\theta^*(p) = -\frac{p^3}{3}.
\]

Thus we recover the relation for the Legendre transform

\[
-xp + \theta^*(p) = -\theta(x).
\]

### 6.2. A general Fourier integral ansatz.

In order to treat a more general case with a caustic of the dimension \( d \) we use the Fourier integral ansatz

\[
\Phi(X) = \int_{\mathbb{R}^d} \phi(X)e^{-iM^{1/2}\Theta(\hat{X},\hat{P})} d\hat{P}
\]

where

\[
\phi(X) \in \mathbb{C}^d,
\]

\[
X = (\hat{X},\hat{X}) \in \mathbb{R}^{3N}, \quad P = (\hat{P},\hat{P}) \in \mathbb{R}^{3N}
\]

\[
\hat{X} \cdot \hat{P} = \sum_{j=1}^{d} \hat{X}^j \hat{P}^j, \quad \hat{X} \cdot \hat{P} = \sum_{j=d+1}^{N} \hat{X}^j \hat{P}^j
\]

\[
\Theta(\hat{X},\hat{X},\hat{P}) = \hat{X} \cdot \hat{P} - \theta^*(\hat{X},\hat{P}),
\]

based on the Legendre transform relationships (cf. [11])

\[
\theta^*(\hat{X},\hat{P}) = \min_\hat{X} \left( \hat{X} \cdot \hat{P} - \theta(\hat{X},\hat{X}) \right),
\]

\[
\theta(\hat{X},\hat{X}) = \min_\hat{P} \left( \hat{X} \cdot \hat{P} - \theta^*(\hat{X},\hat{P}) \right).
\]

If the function \( \theta^*(\hat{X},\hat{P}) \) is not defined for all \( \hat{P} \in \mathbb{R}^d \), but only for \( \hat{P} \in \mathcal{U} \subset \mathbb{R}^d \) we replace the integral over \( \mathbb{R}^d \) by integration over \( \mathcal{U} \) using a smooth cut-off function \( \chi(\hat{P}) \). The cut-off function is zero outside
\( \mathcal{U} \) and equal to one in a large part of the interior of \( \mathcal{U} \). The ansatz (6.7) is inspired by Maslov’s work [33], although it is not the same since our amplitude function \( \phi \) depends on \((\hat{X}, \hat{X})\) but not on \( \hat{P} \).

6.2.1. Making the ansatz for a Schrödinger solution. In this section we construct a solution to the Schrödinger equation from the ansatz (6.7). The constructed solution will be an actual solution and not only an asymptotic solution as in [33]. We consider first the case when the integration is over \( \mathbb{R}^d \) and then conclude in the Remark 6.2 that the cut-off function \( \chi(\hat{P}) \) can be included in all integrals without changing the property of the Fourier integral ansatz being a solution in the \( \hat{X} \)-domain where \( \hat{X} = \nabla_{\rho} \theta^\ast(\hat{X}, \hat{P}) \) for some \( \hat{P} \) satisfying \( \chi(\hat{P}) = 1 \).

The requirement to be a solution means that there should hold

\[
0 = (\hat{H} - E) \Phi
\]

\[
= \int_{\mathbb{R}^d} \left( \frac{1}{2} |\nabla X \theta^\ast(\hat{X}, \hat{P})|^2 + \frac{1}{2} |\hat{P}|^2 + V_0(X) - E \right) \phi(X)e^{-iM^{1/2} \Theta(\hat{X}, \hat{P})} d\hat{P}
\]

Comparing this expression to the previously discussed case of a single WKB-mode we see that the zero order term is now \( \Delta \hat{X} \theta^\ast \) instead of \( \Delta \hat{X} \theta \) and that we have \( -\nabla_{\hat{X}} \phi \cdot \hat{P} \) instead of \( \nabla_{\hat{X}} \phi \cdot \nabla \hat{X} \theta \). However, the main difference is that the first integral is not zero (only the leading order term of its stationary phase expansion is zero, cf. (7.1)). Therefore, the first integral contributes to the second integral. The goal is now to determine a function \( F(\hat{X}, \hat{X}, \hat{P}) \) satisfying

\[
\int_{\mathbb{R}^d} \left( \frac{1}{2} |\nabla X \theta^\ast|^2 + \frac{1}{2} |\hat{P}|^2 + V_0(X) - E \right) e^{-iM^{1/2} \Theta(\hat{X}, \hat{P})} d\hat{P}
\]

\[
= iM^{-1/2} \int_{\mathbb{R}^d} F(\hat{X}, \hat{X}, \hat{P}) e^{-iM^{1/2} \Theta(\hat{X}, \hat{P})} d\hat{P} ,
\]

and verify that it is bounded.

**Lemma 6.1.** There holds \( F = F_0 + F_1 \) where

\[
F_0 = \frac{1}{2} \sum_{i,j} \partial_{\hat{X}^i \hat{X}^j} V_0 \left( \nabla_{\rho} \theta^\ast(\hat{P}) \right) \partial_{\hat{P}^i \hat{P}^j} \theta^\ast(\hat{P}) ,
\]

\[
F_1 = iM^{-1/2} \int_0^1 \int_{\mathbb{R}^d} \sum_{i,j,k} t(1-t) \partial_{\hat{P}^i} \left[ \partial_{\hat{X}^i \hat{X}^j} V_0 \left( \nabla_{\rho} \theta^\ast(\hat{P}) \right) + s t \delta \theta^\ast(\hat{P}) \right] \partial_{\hat{P}^j} \nabla_{\rho} \theta^\ast(\hat{P}) \] \( dt ds .
\]

**Proof.** The function \( \theta^\ast(\hat{X}, \hat{P}) \) is defined as a solution to the Hamilton-Jacobi (eikonal) equation

\[
\frac{1}{2} |\nabla \theta^\ast(\hat{X}, \hat{P})|^2 + \frac{1}{2} |\hat{P}|^2 + V_0 \left( \hat{X}, \nabla_{\rho} \theta^\ast(\hat{X}, \hat{P}) \right) - E = 0
\]

for all \((\hat{X}, \hat{P})\). Consequently, the integral on the left hand side of (6.9) is

\[
\int_{\mathbb{R}^d} \left( V_0(\hat{X}, \hat{X}) - V_0(\hat{X}, \nabla_{\rho} \theta^\ast(\hat{X}, \hat{P})) \right) e^{-iM^{1/2}(\hat{X}, \hat{P}) - \Theta(\hat{X}, \hat{P})} d\hat{P} .
\]

Let \( \hat{P}_0(\hat{X}) \) be any solution to the stationary phase equation \( \hat{X} = \nabla_{\rho} \theta^\ast(\hat{X}, \hat{P}_0) \) and introduce the notation

\[
\Theta'(\hat{X}, \hat{X}, \hat{P}) := \nabla_{\rho} \theta^\ast(\hat{X}, \hat{P}_0) \cdot \hat{P} - \theta^\ast(\hat{X}, \hat{P}) .
\]
Then by writing a difference as \( V(y_1) - V(y_2) = \int_0^1 \partial_y V(y_2 + t(y_1 - y_2)) dt \cdot (y_1 - y_2) \), identifying a derivative \( \partial_{\hat{p}} \) and integrating by parts the integral can be written

\[
\int_{\mathbb{R}^d} \left( V_0(\hat{X}, \nabla_\rho \theta^*(\hat{X}, \hat{P}_0)) - V_0(\hat{X}, \nabla_\rho \theta^*(\hat{X}, \hat{P})) \right) e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P}
\]

\[
= \int_0^1 \int_{\mathbb{R}^d} \sum_i \partial_{\hat{X}_i} V_0(\nabla_\rho \theta^*(\hat{P})) + t \left[ \nabla_\rho \theta^*(\hat{P}_0) - \nabla_\rho \theta^*(\hat{P}) \right] \times
\]

\[
\times (\partial_{\hat{p}} \theta^*(\hat{P}_0) - \partial_{\hat{p}} \theta^*(\hat{P})) \, e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P} \, dt
\]

\[
= -iM^{-1/2} \int_0^1 \int_{\mathbb{R}^d} \sum_i \partial_{\hat{X}_i} V_0(\nabla_\rho \theta^*(\hat{P})) + t \left[ \nabla_\rho \theta^*(\hat{P}_0) - \nabla_\rho \theta^*(\hat{P}) \right] \partial_{\hat{p}} e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P} \, dt dt.
\]

Therefore the leading order term in \( F =: F_0 + F_1 \) is

\[
F_0 := \int_0^1 \sum_{i,j} (1 - t) \partial_{\hat{X}_i \hat{X}_j} V_0(\nabla_\rho \theta^*(\hat{P})) \, \partial_{\hat{p}_i \hat{p}_j} \theta^*(\hat{P}) \, dt
\]

\[
= \frac{1}{2} \sum_{i,j} \partial_{\hat{X}_i \hat{X}_j} V_0(\nabla_\rho \theta^*(\hat{P})) \, \partial_{\hat{p}_i \hat{p}_j} \theta^*(\hat{P}).
\]

Denoting \( \delta \theta^*(\hat{P}) := \nabla_\rho \theta^*(\hat{P}_0) - \nabla_\rho \theta^*(\hat{P}) \) the remainder becomes

\[
-i M^{-1/2} \int_0^1 \int_{\mathbb{R}^d} \left[ \partial_{\hat{X}_i \hat{X}_j} V_0(\nabla_\rho \theta^*(\hat{P}) - \partial_{\hat{X}_i \hat{X}_j} V_0(\nabla_\rho \theta^*(\hat{P}) + t \delta \theta^*(\hat{P})) \right] \times
\]

\[
\times (1 - t) \partial_{\hat{p}_i \hat{p}_j} \theta^*(\hat{P}) e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P} \, dt
\]

\[
= M^{-1/2} \int_0^1 \int_{\mathbb{R}^d} \int_0^1 \sum_{i,j,k} t(1 - t) \partial_{\hat{X}_i \hat{X}_j \hat{X}_k} V_0(\nabla_\rho \theta^*(\hat{P}) + s t \delta \theta^*(\hat{P})) \partial_{\hat{p}_i \hat{p}_j \hat{p}_k} \theta^*(\hat{P})
\]

\[
\times (\partial_{\hat{p}_a} \theta^*(\hat{P}_0) - \partial_{\hat{p}_a} \theta^*(\hat{P})) e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P} \, dt \, ds
\]

\[
= -\frac{1}{M} \int_0^1 \int_{\mathbb{R}^d} \int_0^1 \sum_{i,j,k} t(1 - t) \partial_{\hat{p}_a} \left[ \partial_{\hat{X}_i \hat{X}_j \hat{X}_k} V_0(\nabla_\rho \theta^*(\hat{P}) + s t \delta \theta^*(\hat{P})) \partial_{\hat{p}_i \hat{p}_j \hat{p}_k} \nabla_\rho \theta^*(\hat{P}) \right]
\]

\[
\times e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P} \, dt \, ds,
\]

hence the function \( F_1 \) is purely imaginary and small

\[
F_1 = M^{-1/2} \int_0^1 \int_{\mathbb{R}^d} \int_0^1 \sum_{i,j,k} t(1 - t) \partial_{\hat{p}_a} \left[ \partial_{\hat{X}_i \hat{X}_j \hat{X}_k} V_0(\nabla_\rho \theta^*(\hat{P}) + s t \delta \theta^*(\hat{P})) \partial_{\hat{p}_i \hat{p}_j \hat{p}_k} \nabla_\rho \theta^*(\hat{P}) \right] \, dt \, ds,
\]

and

(6.11) \[ 2Re F = \sum_{i,j} \partial_{\hat{X}_j \hat{X}_j} V_0(\nabla_\rho \theta^*(\hat{P})) \, \partial_{\hat{p}_i \hat{p}_j} \theta^*(\hat{P}). \]

\( \square \)

The eikonal equation (6.10) and the requirement that \( (\hat{H} - E)\Phi = 0 \) in (6.8) then imply that

(6.12) \[ 0 = \int_{\mathbb{R}^d} \left[ i M^{-1/2} \left( \nabla_\chi \phi \cdot \nabla_\chi \theta^* - \nabla_\chi \phi \cdot \hat{P} + \frac{1}{2} \phi (\Delta_\chi \theta^* - 2F(X, \hat{P})) \right) \right.
\]

\[ - (V - V_0) \phi + \frac{1}{2M} \Delta_\chi \phi \left] e^{-iM^{1/2} \theta^*(X, X, P)} \, d\hat{P}. \right. \]
The Hamilton-Jacobi eikonal equation (6.10), in the primal variable \((\hat{X}, \hat{P})\) with the corresponding dual variable \((\hat{P}, \hat{X})\), can be solved locally by the characteristics

\[
\begin{align*}
\dot{\hat{X}} &= \hat{P} \\
\dot{\hat{P}} &= -\nabla_{\hat{X}} V_0(\hat{X}, \hat{X}) \\
\dot{\hat{X}} &= -\hat{P} \\
\dot{\hat{P}} &= \nabla_{\hat{X}} V_0(\hat{X}, \hat{X}) ,
\end{align*}
\]

(6.13)

using the definition

\[
\nabla_{\hat{X}} \theta^*(\hat{X}, \hat{P}) = \hat{P} \\
\nabla_{\hat{P}} \theta^*(\hat{X}, \hat{P}) = \hat{X} .
\]

The characteristics give

\[
\frac{d}{dt}\phi = \nabla_{\hat{X}} \phi \cdot \nabla_{\hat{X}} \theta^* - \nabla_{\hat{P}} \phi \cdot \dot{\hat{P}} ,
\]

so that the Schrödinger transport equation becomes, as in (3.7),

\[
i M^{-1/2} \left( \phi + \phi \frac{\hat{G}}{G} \right) = (V - V_0)\phi - \frac{1}{2M} \Delta_{\hat{X}} \phi
\]

(6.14)

and for \(\psi = G\phi\)

\[
i M^{-1/2} \dot{\psi} = (V - V_0)\psi - \frac{G}{2M} \Delta_{\hat{X}} \psi
\]

(6.15)

with the complex valued weight function \(G\) defined by

\[
\frac{d}{dt}\log G_t = \frac{1}{2} \Delta_{\hat{X}} \theta^*(\hat{X}_t, \hat{P}_t) - F(\hat{X}_t, \hat{P}_t) .
\]

This transport equation is of the same form as the transport equation for a single WKB-mode, with a modification of the weight function \(G\).

Differentiation of the second equation in the Hamiltonian system (6.13) implies that the first variation \(\partial \hat{P}_t / \partial \hat{X}_0\) satisfies

\[
\frac{d}{dt} \left( \frac{\partial \hat{P}_t}{\partial \hat{X}_0} \right) = \sum_{j,k} \partial_X V_0(\hat{X}, \hat{X}_t) \partial_{P_j, P_k} \theta^*(\hat{P}) \partial_X \frac{\partial \hat{P}_t}{\partial \hat{X}_0} ,
\]

which by the Liouville formula (6.18) and the equality

\[
2\text{Re} F = \sum_{i,j} \partial_{\hat{X}_i} V_0 \partial_{P_j, P_k} \theta^* = \text{Tr} \left( \sum_j \partial_{\hat{X}_i} V_0 \partial_{P_j, P_k} \theta^* \right)
\]

in (6.11) yields the relation,

\[
e^{-2 \int_0^t \text{Re} F \, dt} = C \left| \det \frac{\partial \hat{P}_t}{\partial \hat{X}_0} \right| ,
\]

(6.17)

for the constant \(C := | \det \frac{\partial X_0}{\partial \hat{P}_0} |\).

**Remark 6.2.** The conclusion in this section holds also when all integrals over \(d\hat{P}\) in \(\mathbb{R}^d\) are replaced by integrals with the measure \(\chi(\hat{P}) \, d\hat{P}\). Then there holds \(2\text{Re} F = \sum_{i,j} \partial_{\hat{X}_i} \hat{V} \partial_{P_j, P_k} \theta^* \).

### 6.2.2. Liouville’s formula. Here we state Liouville’s formula

\[
\frac{G_0}{G_t} = e^{-\int_0^t \text{Tr} (\nabla P(X)) \, dt} \left| \det \frac{\partial (X_0)}{\partial (X_t)} \right| ,
\]

(6.18)

given in [33].
7. Appendix: The stationary phase expansion

Consider the phase function \( \dot{X} \cdot \dot{P} - \theta^*(\dot{X}, \dot{P}) \) and let \( \dot{P}_0(\dot{X}) \) be any solution to the stationary phase equation \( \dot{X} = \nabla_P \theta^*(\dot{X}, \dot{P}_0) \). We rewrite the phase function

\[
\dot{X} \cdot \dot{P} - \theta^*(\dot{X}, \dot{P}) = \dot{X} \cdot \dot{P}_0 - \theta^*(\dot{X}, \dot{P}_0) + (\dot{P} - \dot{P}_0) \cdot \int_0^1 (1 - t) \partial_{PP} \theta^*(\dot{P}_0 + t[\dot{P} - \dot{P}_0]) \, dt \, [\dot{P} - \dot{P}_0].
\]

The relation

\[
\frac{1}{2} Y \cdot \partial_{PP} \theta(\dot{P}_0) Y = (\dot{P} - \dot{P}_0) \cdot \int_0^1 (1 - t) \partial_{PP} \theta (\dot{P}_0 + t[\dot{P} - \dot{P}_0]) \, dt \, [\dot{P} - \dot{P}_0]
\]

defines the function \( Y(\dot{P}) \), and its inverse \( \dot{P}(Y) \), so that the phase is a quadratic function in \( Y \). The stationary phase expansion of an integral takes the form, see [54],

\[
\int_{\mathbb{R}^d} w(\dot{P}) e^{-iM^{1/2}(\dot{X} \cdot \dot{P} - \theta^*(\dot{X}, \dot{P}))} \, d\dot{P} \\
\sim \sum_{(\dot{P}_0, \nabla_P \theta^*(\dot{P}_0) = \dot{X})} (2\pi M^{-1/2})^{d/2} \left| \det \left( \frac{\partial (\dot{P}_0)}{\partial (\dot{X})} \right) \right|^{1/2} e^{i\frac{2\pi}{M} \text{sgn}(\partial_{PP} \theta^*(\dot{P}_0))} e^{-iM^{1/2} \theta(\dot{X}, \dot{P}_0)} \\
\times \sum_{k=0}^{\infty} \frac{M^{-k/2}}{k!} \left( \sum_{l,j} i (\partial_{PP}, \theta^*)^{-1}(\dot{P}_0) \partial_{YY} \right) \left( \left. w(\dot{P}(Y)) \left| \det \left( \frac{\partial (\dot{P})}{\partial (Y)} \right) \right| \right)_{Y=0}.
\]

8. Appendix: An alternative motivation for assumption (2.17)

The double average (2.17) in time and phase-space is in Monte Carlo methods approximated by sampling several paths with random initial points \((X_0, P_0)\). For a single path one expects \( O(T^{-1/2}) \) convergence rate, in the case of bounded correlation times. Many paths correspond in some sense to longer averaging time for a single path and one may ask how long. This answer is related to how the dynamics maps an initial phase-space neighborhood \( D \) to \( S_{T/2}(D) \). The exponential \( e^{CT} \) growth of the first variation \(|\partial_P g \circ S_T(X_0, P_0)|\), for \(|\alpha| = 1\), implies that an initial domain \( D \), on the constant energy hyper-surface \( F := \{(X, P) | H_0(X, P) = \text{constant}\} \), is stretched but mapped to the bounded energy surface \( F \). If the dynamics is ergodic the sampling density becomes uniform on this energy hyper-surface, asymptotically in time. Paths that are very close are not acting as independent samples for finite \( T \). The assumption (2.17) means that the exponential growth of the first variation measures how small the initial distance, \( e^{-CT/2} \), of the paths can be in order to lead to approximately independent paths after time \( T/2 \). The paths that are initially \( e^{-CT/2} \) close will after time \( T/2 \) have a distance of order one. The remaining \( T/2 \) time this \( O(e^{CT/2}) \) number of paths would act as approximately independent samples (over time \( T/2 \)) as tested in Figure 18 and would give the expected decay \( e^{-CT/4} \), if the Lyapunov exponent is positive in only one direction, \( d = 1 \). If \( d \) Lyapunov exponents are similar to the maximal, we expect the decay \( e^{-CTd/4} \).

Figure 18 tests ergodicity for Born-Oppenheimer molecular dynamics simulations with the potential \( \lambda_n(X) = X_1^2/2 + X_2^2/\sqrt{2} + 2 \sin(X_1 X_2) \) and the double average over time and \( N \) initial points \( \{(X_0[n], P_0[n]) \sim (0, 0, \sqrt{2}(E - \lambda_n(0)) \cos(1.2 + n\Delta v), \sqrt{2}(E - \lambda_n(0)) \sin(1.2 + n\Delta v))\}_{n=1}^{\infty} \) in (2.17). The vertical axis shows the sampling error \( \sum_{n=1}^{N} \frac{1}{\sqrt{T/2}} \int_{T/2}^T (g(X_t) - g_{\text{MD}}(X_0[n], P_0[n])) dt/N \) and the horizontal axis shows the number of samples \( N \) for the symplectic Euler method. The convergence rate \( O(N^{-1/2}) \) for \( \Delta v = 10^{-6} \) (left figure) and \( \Delta v = 10^{-10} \) (middle figure) indicate independent sample paths while the closer paths with \( \Delta v = 10^{-14} \) in (right figure) are not acting independently since the convergence is slower. The maximal Lyapunov exponent is in this case computed to be roughly 0.35 (the others are \( -0.35, 0, 0 \)), so that the expansion of the path distance from time zero to time 70 becomes \( e^{0.35 \times 70} \approx 10^{10} \) and the approximate \( N^{-1/2} \) convergence rate for \( \Delta v \geq 10^{-10} \) indicates that \( \gamma \approx 1/4 \) as suggested in the paragraph following (2.17). Here \( T = 140, \Delta t = 0.01, E = 1.5, g(X) = \sin(X_1 X_2) \). The ergodic limit \( g_{\text{MD}} \) is approximated in two different ways: by a single symplectic Euler path to \(-0.4389 \) and by the Monte Carlo samples [4,10] to \(-0.4388 \). The dashed line shows \( N^{-1/2} \).
Acknowledgment

The research was partially supported by the U.S. National Science Foundation under the grant NSF-DMS-0813893, U.S. DOD-ARO Grant Award W911NF-14-1-0247, Swedish Research Council grants 621-2010-5647 and 621-2014-4776 and the Swedish e-Science Research Center. P.P. also thanks KTH and Nordita for their hospitality during his visit when the presented research was initiated.

References


