APPROXIMATION OF SEMICLASSICAL EXPECTATION VALUES BY SYMPLECTIC GAUSSIAN WAVE PACKET DYNAMICS

TOMOKI OHSAWA

ABSTRACT. This paper concerns an approximation of the expectation values of the position and momentum of the solution to the semiclassical Schrödinger equation with a Gaussian as the initial condition. Of particular interest is the approximation obtained by our symplectic/Hamiltonian formulation of the Gaussian wave packet dynamics that introduces a correction term to the conventional formulation using the classical Hamiltonian system by Hagedorn and others. The main result is a proof that our formulation gives a higher-order approximation than the classical formulation does to the expectation value dynamics under certain conditions on the potential function. Specifically, as the semiclassical parameter ε approaches 0, our dynamics gives an $O(\varepsilon^{3/2})$ approximation of the expectation value dynamics whereas the classical one gives an $O(\varepsilon)$ approximation.

1. INTRODUCTION

1.1. Semiclassical Schrödinger Equation and Gaussian Wave Packet.

Consider the following initial value problem of the semiclassical Schrödinger equation on $\mathbb{R}^d$:

$$\mathrm{i}\varepsilon\frac{\partial}{\partial t}\psi(t,x) = \hat{H}\psi(t,x), \qquad \hat{H} = \frac{\hat{p}^2}{2} + V(x), \tag{1a}$$

$$\psi(0,x) = \phi_0(q(0), p(0), Q(0), P(0), S(0); x), \tag{1b}$$

where $\varepsilon > 0$ is the semiclassical parameter, $\hat{p} := -\mathrm{i}\varepsilon\partial/\partial x$ is the momentum operator, and ϕ_0 is the Gaussian wave function

$$\phi_0(q, p, Q, P, S; x) := \frac{(\det Q)^{-1/2}}{(\pi\varepsilon)^{d/4}} \exp\left\{\frac{\mathrm{i}}{\varepsilon}\left(\frac{1}{2}(x-q)^T PQ^{-1}(x-q) + p\cdot(x-q) + S\right)\right\}. \tag{2}$$

The parameters (q,p) live in the cotangent bundle $T^*\mathbb{R}^d \cong \mathbb{R}^d \times \mathbb{R}^d$, whereas $Q, P \in \mathsf{M}_d(\mathbb{C})$ (the set of $d \times d$ complex matrices) satisfy

$$Q^T P - P^T Q = 0 \quad \text{and} \quad Q^* P - P^* Q = 2\mathrm{i}I,$$

and $S \in \mathbb{R}$ is a phase factor. It is worth noting that the imaginary part of PQ^{-1} is given by $(QQ^*)^{-1}$; see, e.g., Lubich [17, Lemma V.1.1].

The seminal works by Hagedorn [6, 9] (see also Heller [12, 13, 14], Robert [31], Combescure and Robert [2]) showed that one may approximate the solution $\psi(t,x)$ of the above initial value problem (1) in the semiclassical limit $\varepsilon \to 0$ by the time-dependent Gaussian wave packet

$$\phi_0(t,x) := \phi_0(q(t), p(t), Q(t), P(t), S(t); x) \tag{3}$$

---

whose parameters evolve in time according to the ordinary differential equations

$$\dot{q} = p, \qquad \dot{p} = -DV(q), \qquad \dot{Q} = P, \qquad \dot{P} = -D^2V(q)\,Q,$$

$$\dot{S} = \frac{p^2}{2} - V(q), \tag{4}$$

where DV and D^2V stand for the gradient and Hessian of the potential V. Note that the equations for (q, p) are the *classical* Hamiltonian system. Specifically, Hagedorn [6, 9] proved, under certain conditions on the potential V, that the error in terms of the L^2-norm $\|\cdot\|$ in $L^2(\mathbb{R}^d)$ is $O(\varepsilon^{1/2})$ in the sense that there exists a function $\mathscr{C}(t)$ such that

$$\|\psi(t, x) - \phi_0(t, x)\| \le \mathscr{C}(t)\,\varepsilon^{1/2}. \tag{5}$$

1.2. **Symplectic Gaussian Wave Packet Dynamics.** In a series of works [23, 25, 26], we proposed the following symplectic/Hamiltonian alternative to the evolution equation (4):

$$\dot{q} = p, \qquad \dot{p} = -\partial_q\Big(V(q) + \varepsilon\,V^{(1)}(q, Q)\Big), \qquad \dot{Q} = P, \qquad \dot{P} = -D^2V(q)\,Q, \tag{6a}$$

$$\dot{S} = \frac{p^2}{2} - V(q), \tag{6b}$$

where ∂_q is a shorthand for $\partial/\partial q$ and

$$V^{(1)}(q, Q) := \frac{1}{4}\,\mathrm{tr}\,\big(QQ^*D^2V(q)\big). \tag{7}$$

The only difference from (4) of Hagedorn is the $O(\varepsilon)$ correction term in the potential. The correction term renders the coupled system (6a) for (q, p, Q, P) a Hamiltonian system on $T^*\mathbb{R}^d \times \mathsf{M}_d(\mathbb{C}) \times \mathsf{M}_d(\mathbb{C})$ with a natural symplectic structure and the following Hamiltonian [23]:

$$H^\varepsilon(q, p, Q, P) := \frac{p^2}{2} + V(q) + \frac{\varepsilon}{4}\big(\mathrm{tr}(P^*P) + \mathrm{tr}\,\big(QQ^*\,D^2V(q)\big)\big). \tag{8}$$

Note also that (6b) is decoupled and $t \mapsto S(t)$ is obtained by a quadrature using the solution to (6a). In what follows, the time dependent functions $t \mapsto (q(t), p(t), Q(t), P(t), S(t))$ refer to the solution to (6) with the initial condition $(q(0), p(0), Q(0), P(0), S(0))$ at $t = 0$, unless otherwise stated.

This is in contrast to (4), which is Hamiltonian in the decoupled classical dynamics of (q, p) in the classical phase space $T^*\mathbb{R}^d$ but not as a system for (q, p, Q, P). We also note in passing that Watson et al. [33] obtained (6) from a different perspective, and also that a correction term of the above form was proposed earlier by Pattanayak and Schieve [27] for the one-dimensional case, and also by Prezhdo and Pereverzev [29, 30] and Prezhdo [28] in a different manner. Our formulation gave a symplectic-geometric account of the variational formulation of Faou and Lubich [4] (see also Lubich [17, Section II.4]), and yielded the correction term as a result of an asymptotic expansion of the resulting potential term [23, 25].

1.3. **Main Result.** The main question we would like to address is how the $O(\varepsilon)$ correction term in (6) contributes to the accuracy of the approximation by the Gaussian wave packet dynamics. We are particularly interested in approximating the dynamics of the expectation values of the position and momentum operators

$$\hat{z} := (\hat{x}, \hat{p}) = \left(\hat{x}, -\mathrm{i}\varepsilon\frac{\partial}{\partial x}\right),$$

that is, the dynamics defined as

$$t \mapsto \langle \hat{z} \rangle(t) = \langle \psi(t, \cdot), \hat{z}\, \psi(t, \cdot) \rangle, \tag{9}$$

where $\langle \cdot, \cdot \rangle$ is the standard (right-linear) inner product on $L^2(\mathbb{R}^d)$, and $t \mapsto \psi(t, \cdot)$ is the solution of the initial value problem (1).

Numerical experiments [26] suggest that our dynamics (6) gives a better approximation than the classical dynamics (4) does to the exact expectation value dynamics. We note that the comparisons were made with respect to the expectation value dynamics obtained by Egorov's method [2, 15, 16] or the Initial Value Representation (IVR) method [19–21, 32], which is known to give an $O(\varepsilon^2)$ approximation to the exact dynamics (9); see, e.g., Egorov [3], Bouzouina and Robert [1], and Zworski [34, Chapter 11].

Our main result gives a rigorous account of this observation:

Theorem 1.1. *Suppose that $V \in C^4(\mathbb{R}^d)$ is bounded from below, i.e., $C_1 \leq V(x)$ for some $C_1 \in \mathbb{R}$ for any $x \in \mathbb{R}^d$, and also that $x \mapsto D_{ij}^2 V(x)$, i.e., the (i,j)-component of the Hessian $D^2 V(x)$, is bounded for any $i, j \in \{1, \ldots, d\}$. Let $t \mapsto (z(t), Q(t), P(t))$ with $z(t) := (q(t), p(t))$ be the solution to (6a), $t \mapsto z^0(t) = (q^0(t), p^0(t))$ be that to the classical Hamiltonian system in (4) with the initial condition $z^0(0) = z(0)$, and let $t \mapsto \langle \hat{z} \rangle(t)$ be the exact expectation value dynamics (9). Then, for any $i \in \{1, \ldots, 2d\}$, $z_i(t) - \langle \hat{z}_i \rangle(t) = O(\varepsilon^{3/2})$ in the sense that there exists a function $\mathscr{C}_i(t)$ such that*

$$|z_i(t) - \langle \hat{z}_i \rangle(t)| \leq \mathscr{C}_i(t)\, \varepsilon^{3/2},$$

whereas $z_i^0(t) - \langle \hat{z}_i \rangle(t) = O(\varepsilon)$ in the same sense.

Several remarks are in order. This paper is *not* about improving the accuracy of an ansatz for the wave function itself. In fact, replacing (4) by (6) does *not* improve the estimate (5) in terms of ε, as we shall show in Corollary 3.5 (with $\mathbf{n} = 0$). The focus of the paper is rather on improving the approximation of the expectation value dynamics *without having any additional time evolution equations other than (4) nor assuming any ansatz other than the Gaussian (3)*: We achieve it by simply introducing a correction term to (4).

We also would like to stress that, as shown in [23, 25], the derivation of (6) also involves only the Gaussian (2). In other words, the $O(\varepsilon)$ correction term does *not* come from any higher-order ansatz as one might expect. This is in contrast to assuming a higher-order ansatz than just the Gaussian (3) as is done in [7–11] and [2, Theorem 24 on p. 109]. One can certainly improve the accuracy of the ansatz that way, but needs additional evolution equations in addition to those for (q, p, Q, P, S).

Note also that our approximation involves only a single initial value problem of (6) as opposed to averaging solutions over numerous initial conditions like the Egorov/IVR method mentioned above.

Throughout the paper, we will carry out asymptotic analysis as $\varepsilon \to 0$ of time-dependent functions, and will employ the same notation used in the statement of the above theorem for brevity. Specifically, when we write $f(t, \varepsilon) = O(\varepsilon^r)$ for some time-dependent function f with some $r \in \mathbb{R}$, it means that there exists a function $\mathscr{C}(t)$ such that $|f(t, \varepsilon)| \leq \mathscr{C}(t)\, \varepsilon^r$ as $\varepsilon \to 0$.

1.4. **Approximation of Other Observables.** One also naturally wonders whether Theorem 1.1 extends to the expectation values of general observables as well. We defer this question to future work. However, it is easy to see that the result holds for the Hamiltonian with an even better approximation: As we mentioned above, our dynamics (6) is a Hamiltonian system with the Hamiltonian H^ε given in (8). But then this Hamiltonian is an $O(\varepsilon^2)$ approximation to the expectation value of the Hamiltonian operator $\hat{H}$ from (1a) with respect to the Gaussian:

$$\left\langle \phi_0(q,p,Q,P,S), \hat{H}\phi_0(q,p,Q,P,S) \right\rangle = H^\varepsilon(q,p,Q,P) + O(\varepsilon^2).$$

This follows from Laplace's method (see, e.g., Miller [18, Section 3.7]) applied to the integral on the left. Now, note that $t \mapsto \langle \hat{H} \rangle(t) := \langle \psi(t), \hat{H}\psi(t) \rangle$ along the exact solution to (1) and $t \mapsto H^\varepsilon(q(t), p(t), Q(t), P(t))$ along (6) are both constant. So their difference is constant at the initial value—where $\psi(0)$ is the initial Gaussian (1b):

$$\langle \hat{H} \rangle(t) - H^\varepsilon(q(t), p(t), Q(t), P(t)) = \langle \hat{H} \rangle(0) - H^\varepsilon(q(0), p(0), Q(0), P(0)) = O(\varepsilon^2).$$

On the other hand, with the classical system (4) and the classical Hamiltonian $H^0(q,p) := p^2/2 + V(q)$,

$$\langle \hat{H} \rangle(t) - H^0(q^0(t), p^0(t)) = \langle \hat{H} \rangle(0) - H^0(q(0), p(0)) = O(\varepsilon),$$

because $H^\varepsilon(q,p,Q,P) = H^0(q,p) + O(\varepsilon)$.

A similar argument works for, e.g., the angular momentum $J(q,p) := q \diamond p$—where $q \diamond p$ denotes the $d \times d$ skew-symmetric matrix defined by $(q \diamond p)_{ij} := q_j p_i - q_i p_j$—when the potential V has $\mathsf{SO}(d)$-symmetry by approximating the expectation value $\langle \hat{J} \rangle$ by the semiclassical angular momentum [22, 23]:

$$J^\varepsilon(q,p,Q,P) = q \diamond p + \frac{\varepsilon}{2}\operatorname{Re}(PQ^* - QP^*),$$

because both $\langle \hat{J} \rangle$ and J^ε are invariants.

1.5. **Outline.** We prove Theorem 1.1 in the rest of the paper. The main part of the proof is in Section 4, whereas Sections 2 and 3 are devoted to some lemmas and propositions needed in Section 4. Therefore, the reader might want to first skim through Section 4 to have an overview of the proof.

Much of what we do is a detailed analysis of the error $\mathcal{Z}_0(t;x) := \psi(t,x) - \phi_0(t,x)$, i.e., the difference between the exact solution to (1) and the Gaussian wave packet (3). In fact, the difference between the exact expectation value $\langle \hat{x} \rangle(t)$ of the position and the position variable $q(t)$ in (6a) is, dropping the spatial variables x for brevity,

$$
\begin{aligned}
\langle \hat{x} \rangle(t) - q(t) &= \langle \psi(t), (\hat{x} - q(t))\psi(t) \rangle \\
&= \langle \phi_0(t), (\hat{x} - q(t))\phi_0(t) \rangle + \langle \phi_0(t), (\hat{x} - q(t))\mathcal{Z}_0(t) \rangle + \langle \mathcal{Z}_0(t), (\hat{x} - q(t))\phi_0(t) \rangle \\
&\quad + \langle \mathcal{Z}_0(t), (\hat{x} - q(t))\mathcal{Z}_0(t) \rangle \\
&= 2\operatorname{Re} \langle \mathcal{Z}_0(t), (\hat{x} - q(t))\phi_0(t) \rangle + \langle \mathcal{Z}_0(t), (\hat{x} - q(t))\mathcal{Z}_0(t) \rangle.
\end{aligned}
$$

Therefore, our analysis boils down to estimates of the above two terms involving the error $\mathcal{Z}_0$. Those lemmas and propositions in Sections 2 and 3 mainly concern those key properties of $\mathcal{Z}_0$ that are pertinent to our analysis.

2. Hagedorn Wave Packets

2.1. **Overview.** We first give a brief review of the Hagedorn wave packets [6–9] (see also [2] and [24]), and then derive the evolution equation satisfied by the Gaussian wave packet (3) where the parameters evolve in time according to our equations (6). The evolution equation resembles the Schrödinger equation (1a) but differs by a residual term. We then prove several key properties of the residual term. Later, in Section 3, we will find an expression for the error $\mathcal{Z}_0(t, x)$ in terms of the residual term analyzed here.

2.2. **The Hagedorn Wave Packets.** Following Hagedorn [9], let us define the lowering operator

$$\mathscr{A}(q, p, Q, P) := -\frac{\mathrm{i}}{\sqrt{2\varepsilon}} \big(P^T(\hat{x} - q) - Q^T(\hat{p} - p) \big)$$

as well as its adjoint or the raising operator

$$\mathscr{A}^*(q, p, Q, P) := \frac{\mathrm{i}}{\sqrt{2\varepsilon}} (P^*(\hat{x} - q) - Q^*(\hat{p} - p)). \tag{10}$$

We refer to the lowering and raising operators collectively as the *ladder operators*. Note that both are operators on the Schwartz space $\mathscr{S}(\mathbb{R}^d)$. It is straightforward to see that they satisfy the following relationship for any $j, k \in \{1, \ldots, d\}$:

$$[\mathscr{A}_j(q, p, Q, P), \mathscr{A}_k^*(q, p, Q, P)] = \delta_{jk}. \tag{11}$$

It also turns out to be convenient to write the position and momentum operators in terms of the ladder operators as follows:

$$\hat{x} - q = \sqrt{\frac{\varepsilon}{2}} \big(\overline{Q} \mathscr{A}(q, p, Q, P) + Q \mathscr{A}^*(q, p, Q, P) \big), \tag{12}$$

$$\hat{p} - p = \sqrt{\frac{\varepsilon}{2}} \big(\overline{P} \mathscr{A}(q, p, Q, P) + P \mathscr{A}^*(q, p, Q, P) \big). \tag{13}$$

Let $\mathbb{N}_0$ be the set of integers greater than or equal to zero. Then one can generate a set of functions $\{\phi_{\mathbf{n}}(q, p, Q, P, S)\}_{\mathbf{n} \in \mathbb{N}_0^d}$ by recursively defining, for any multi-index $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}_0^d$ and $j \in \{1, \ldots, d\}$,

$$\phi_{\mathbf{n}+\mathbf{e}_j}(q, p, Q, P, S; x) := \frac{1}{\sqrt{n_j + 1}} \mathscr{A}_j^*(q, p, Q, P) \phi_{\mathbf{n}}(q, p, Q, P, S; x), \tag{14}$$

where $\mathbf{e}_j$ is the unit vector in $\mathbb{R}^d$ whose j-th entry is 1; then they also satisfy

$$\phi_{\mathbf{n}-\mathbf{e}_j}(q, p, Q, P, S; x) := \frac{1}{\sqrt{n_j}} \mathscr{A}_j(q, p, Q, P) \phi_{\mathbf{n}}(q, p, Q, P, S; x).$$

Hagedorn [6, 7, 8, 9] showed that the set $\{\phi_{\mathbf{n}}(q, p, Q, P, S)\}_{\mathbf{n} \in \mathbb{N}_0^d}$ then forms an orthonormal basis for $L^2(\mathbb{R}^d)$, where ϕ_0 is the "ground state" here in the sense that

$$\mathscr{A}(q, p, Q, P) \phi_0(q, p, Q, P, S; x) = 0. \tag{15}$$

In fact, one may think of them as a generalization of the Hermite functions, and also can find a unitary operator on $L^2(\mathbb{R}^d)$ that relates each element of the Hagedorn wave packet with the Hermite function of the same index [24]. Because of this correspondence, we refer to $\phi_{\mathbf{n}}$ as an $|\mathbf{n}|$-th *excited state* with $|\mathbf{n}| := \sum_{i=1}^{d} n_i$ for any multi-index $\mathbf{n} \in \mathbb{N}_0^d$.

We note in passing that Hagedorn [9] actually constructed an orthonormal basis $\{\varphi_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}_0^d}$ without the phase factor starting with the Gaussian

$$\varphi_0(q,p,Q,P;x) := \frac{(\det Q)^{-1/2}}{(\pi\varepsilon)^{d/4}}\exp\left\{\frac{\mathrm{i}}{\varepsilon}\left(\frac{1}{2}(x-q)^T PQ^{-1}(x-q) + p\cdot(x-q)\right)\right\}$$

instead of ϕ_0 from (2). It is just a matter of convenience that we use the basis $\{\phi_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}_0^d}$ with the phase factor instead of $\{\varphi_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}_0^d}$.

2.3. **Evolution Equation of the Gaussian Wave Packet.** Hagedorn [6, 7, 8, 9] and Hagedorn and Joye [10, 11] have proved error estimates of various approximations to the solution to the Schrödinger equation (1a) constructed by taking linear combinations of the Hagedorn wave packets. In their works, each wave packet evolves in time according to (4), and one of the key ideas of these estimates is to find the Schrödinger-type equation satisfied by those wave packets and identify the residual term that accounts for the difference from the Schrödinger equation (1a).

Following their approach, we would like to first find the Schrödinger-type equation satisfied by the time-dependent Gaussian wave packet (3). The resulting residual term slightly differs from Hagedorn's because of the $O(\varepsilon)$ correction term:

Lemma 2.1. *Consider the Gaussian $\phi_0(t,x)$ from (3) whose time-dependent parameters satisfy (6). Then it satisfies the Schrödinger-type equation*

$$\mathrm{i}\varepsilon\frac{\partial}{\partial t}\phi_0(t,x) = \hat{H}\phi_0(t,x) + \varepsilon^{3/2}\zeta_0(t,x), \tag{16}$$

where we defined the residual term

$$\zeta_0(t,x) := \alpha(q(t),Q(t);x)\,\phi_0(t,x) \tag{17}$$

with

$$\alpha(q,Q;x) := \varepsilon^{-1/2}\partial_q V^{(1)}(q,Q)\cdot(x-q) + \varepsilon^{-3/2}\left(\sum_{k=0}^{2}\frac{1}{k!}D^k V(q)\cdot(x-q)^k - V(x)\right). \tag{18}$$

Furthermore, we may split α as

$$\alpha(q,Q;x) = \alpha^{(0)}(q,Q;x) + \varepsilon^{1/2}\alpha^{(1)}(q;x) \tag{19a}$$

with

$$\begin{aligned}
\alpha^{(0)}(q,Q;x) &:= \varepsilon^{-1/2}\partial_q V^{(1)}(q,Q)\cdot(x-q) - \frac{\varepsilon^{-3/2}}{3!}D^3 V(q)\cdot(x-q)^3 \\
&= \partial_q V^{(1)}(q,Q)\cdot\xi - \frac{1}{6}D^3 V(q)\cdot\xi^3, \tag{19b}
\end{aligned}$$

$$\begin{aligned}
\alpha^{(1)}(q;x) &:= \varepsilon^{-2}\left(\sum_{k=0}^{3}\frac{1}{k!}D^k V(q)\cdot(x-q)^k - V(x)\right) \\
&= \frac{\varepsilon^{-2}}{4!}D^4 V(\sigma_1(x,q))\cdot(x-q)^4 \\
&= \frac{1}{4!}D^4 V(\sigma_1(x,q))\cdot\xi^4. \tag{19c}
\end{aligned}$$

Note that we set $\xi := \varepsilon^{-1/2}(\hat{x} - q)$ and $D^0 V := V$, and that $\sigma_1(x, q)$ is a point in the segment joining x and q; we also used the shorthand ξ^m with $m \in \mathbb{N}_0$ for the m-tensor defined as $\xi^m_{i_1 \dots i_m} := \xi_{i_1} \cdots \xi_{i_m}$ as well as

$$D^m V(q) \cdot \xi^m := D^m_{i_1 \dots i_m} V(q) \xi^m_{i_1 \dots i_m}$$

with Einstein's summation convention on repeated indices.

Remark 2.2. Those terms with $D^3 V$ in $\alpha^{(0)}$ and $\alpha^{(1)}$ cancel with each other in α, but as we shall see below, splitting the terms in α in this manner is crucial for us as we shall see in the next subsection.

Proof of Lemma 2.1. It follows from tedious but straightforward calculations: Dropping the time variable t for brevity in the calculations, we have, using (6),

$$
\begin{aligned}
\varepsilon^{3/2} \alpha(q(t), Q(t); x) &= \left(\mathrm{i}\varepsilon \frac{\partial}{\partial t} \phi_0(t, x) - \hat{H} \phi_0(t, x) \right) \Big/ \phi_0(t, x) \\
&= (\dot{q} - p)^T P Q^{-1}(x - q) - \frac{1}{2}(x - q)^T \left(\dot{P} Q^{-1} - P Q^{-1} \dot{Q} Q^{-1} + (PQ^{-1})^2 \right)(x - q) \\
&\quad - \dot{p} \cdot (x - q) + p \cdot (\dot{q} - p) - V(x) + V(q) \\
&\quad + \left(\frac{p^2}{2} - V(q) - \dot{S} \right) - \frac{\mathrm{i}}{2}\varepsilon\left(\mathrm{tr}(Q^{-1}\dot{Q}) - \mathrm{tr}(PQ^{-1}) \right) \\
&= \varepsilon\, \partial_q V^{(1)}(q, Q) \cdot (x - q) \\
&\quad + V(q) + DV(q) \cdot (x - q) + \frac{1}{2}D^2 V(q) \cdot (x - q)^2 - V(x),
\end{aligned}
$$

which gives (18). We may then split α as follows:

$$
\begin{aligned}
\alpha(q, Q; x) &= \varepsilon^{-1/2}\, \partial_q V^{(1)}(q, Q) \cdot (x - q) - \frac{\varepsilon^{-3/2}}{3!} D^3 V(q) \cdot (x - q)^3 \\
&\quad + \varepsilon^{-3/2}\left(\sum_{k=0}^{3} \frac{1}{k!} D^k V(q) \cdot (x - q)^k - V(x) \right) \\
&= \alpha^{(0)}(q, Q; x) + \varepsilon^{1/2}\alpha^{(1)}(q; x).
\end{aligned}
$$

The second expression for $\alpha^{(1)}$ follows from Taylor's Theorem because V is of class C^4. $\qquad\square$

2.4. Properties of the Residual Term.

Let us prove some key properties of the residual term ζ_0 as lemmas. These lemmas show why we split α into $\alpha^{(0)}$ and $\alpha^{(1)}$ as shown in (19).

Lemma 2.3. *Under the assumptions on the potential V from Theorem 1.1, we have the following estimates for the residual term ζ_0 defined in (17):*

(i) $\left\| \alpha^{(0)}(q(t), Q(t); \cdot)\phi_0(t, \cdot) \right\| = O(1)$;

(ii) $\left\| \alpha^{(1)}(q(t); \cdot)\phi_0(t, \cdot) \right\| = O(1)$;

(iii) $\left\| \zeta_0(t, \cdot) \right\| = O(1)$.

(iv) $\left\| \hat{\xi}_i(t)\, \zeta_0(t, \cdot) \right\| = O(1)$ *for any $i \in \{1, \dots, d\}$ with $\hat{\xi}(t) := \varepsilon^{-1/2}(\hat{x} - q(t))$.*

(v) $\left\| \hat{\eta}_i(t)\, \zeta_0(t, \cdot) \right\| = O(1)$ *for any $i \in \{1, \dots, d\}$ with $\hat{\eta}(t) := \varepsilon^{-1/2}(\hat{p} - p(t))$.*

Proof. To prove (i), notice that it is the square root of the integral with respect to x of

$$|\phi_0(t, x)|^2 = \frac{|\det Q(t)|^{-1}}{(\pi\varepsilon)^{d/2}} \exp\left(-\frac{1}{\varepsilon}(x - q(t))^T (Q(t)Q(t)^*)^{-1}(x - q(t))\right)$$

multiplied by a polynomial of $\varepsilon^{-1/2}(x - q(t))$. It is straightforward to see that, by performing the integral using the change of variables from x to $\xi := \varepsilon^{-1/2}(x - q(t))$, the integral does not depend on ε. More specifically, for any $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{N}_0^d$,

$$\left\| \xi_1^{m_1} \ldots \xi_d^{m_d} \phi_0(t, \cdot) \right\| = O(1),$$

and hence it follows that $\left\| \alpha^{(0)}(q(t), Q(t); \cdot)\phi_0(t, \cdot) \right\| = O(1)$; see, e.g., Hagedorn [9, Eq. (3.30)] for an equivalent statement.

For (ii), we mimic the proof of Theorem 2.9 of Hagedorn [9]. Take an arbitrarily small $r > 0$ and let $\bar{\mathbb{B}}_r(q(t)) \subset \mathbb{R}^d$ be the closed ball with radius r centered at $q(t)$. Then,

$$\alpha^{(1)}(q(t); x) = \mathbf{1}_{\bar{\mathbb{B}}_r(q(t))}(x)\, \alpha^{(1)}(q(t); x) + \mathbf{1}_{\bar{\mathbb{B}}_r(q(t))^{\mathrm{c}}}(x)\, \alpha^{(1)}(q(t); x),$$

where $\mathbf{1}_A$ stands for the characteristic function for a subset $A \subset \mathbb{R}^d$. Since V is of class C^4, there exists a 4-tensor-valued time-dependent function $F_r(t)$ such that, for any $x \in \bar{\mathbb{B}}_r(q(t))$,

$$\left| \alpha^{(1)}(q(t); x) \right| = \left| \frac{1}{4!} D^4 V(\sigma_1(x, q(t))) \cdot \xi^4 \right| \le \left| F_r(t) \cdot \xi^4 \right|.$$

On the other hand, there exist $C_2(t) > 0$ and a polynomial $\mathcal{P}(x)$ such that, for any $x \in \bar{\mathbb{B}}_r(q(t))^{\mathrm{c}}$,

$$\left| \alpha^{(1)}(q(t); x) \right| \le \varepsilon^{-2} C_2(t)\mathcal{P}(x)$$

because the boundedness assumption on the Hessian $D^2 V$ implies that the potential V is dominated by a quadratic function on $\bar{\mathbb{B}}_r(q(t))^{\mathrm{c}}$, and also the rest of $\alpha^{(1)}(q(t); x)$ is cubic in x. Therefore,

$$\left| \alpha^{(1)}(q(t); x)\phi_0(t, x) \right| = \mathbf{1}_{\bar{\mathbb{B}}_r(q(t))}(x) \left| \alpha^{(1)}(q(t); x)\phi_0(t, x) \right| + \mathbf{1}_{\bar{\mathbb{B}}_r(q(t))^{\mathrm{c}}}(x) \left| \alpha^{(1)}(q(t); x)\phi_0(t, x) \right|$$

$$\le \mathbf{1}_{\bar{\mathbb{B}}_r(q(t))}(x) \left| F_r(t) \cdot \xi^4 \phi_0(t, x) \right| + \varepsilon^{-2} C_2(t)\mathbf{1}_{\bar{\mathbb{B}}_r(q(t))^{\mathrm{c}}}(x)\, \mathcal{P}(x) \left| \phi_0(t, x) \right|.$$

However, the norm of the first term is $O(1)$ following the same argument as in (i), whereas the norm of the second term is $o(\varepsilon^r)$ for any real r effectively canceling ε^{-2} in the coefficient. Hence $\left\| \alpha^{(1)}(q(t); \cdot)\phi_0(t, \cdot) \right\| = O(1)$.

The estimate in (iii) follows easily from (i) and (ii):

$$\|\zeta_0(t, \cdot)\| \le \left\| \alpha^{(0)}(q(t), Q(t); \cdot)\phi_0(t, \cdot) \right\| + \varepsilon^{1/2}\left\| \alpha^{(1)}(q(t); \cdot)\phi_0(t, \cdot) \right\| = O(1).$$

The estimate in (iv) holds similarly because the above estimates do not change upon multiplying $\phi_0(t, x)$ by $\hat{\xi}_i(t) := \varepsilon^{-1/2}(\hat{x} - q(t))_i$.

The estimate in (v) holds because straightforward calculations (see Appendix A.1) show that

$$\hat{\eta}_i(t)\zeta_0(t, x) = \beta_i(q(t), Q(t); x)\phi_0(t, x) + (P(t)Q(t)^{-1})_{ij}\xi_j(t)\zeta_0(t, x), \tag{20}$$

where $\beta_i(q, Q; x) := \beta_i^{(0)}(q, Q; x) + \varepsilon^{1/2}\beta_i^{(1)}(q; x)$ with

$$\beta_i^{(0)}(q, Q; x) := -\mathrm{i}\left(\partial_{q_i} V^{(1)}(q, Q) - \frac{1}{2}D_{ijk}^3 V(q)\xi_{jk}^2 \right),$$

$$\beta_i^{(1)}(q; x) := -\mathrm{i}\,\varepsilon^{-3/2}\left(D_i V(q) + D_{ij}^2 V(q)(x - q)_j + \frac{1}{2}D_{ijk}^3 V(q)(x - q)_{jk}^2 - D_i V(x) \right). \tag{21}$$

For the first term on the right-hand side of (20), notice the similarity between β_i and α; so we can obtain the estimate of the first term essentially the same way we did for ζ_0, i.e., its norm is $O(1)$. We also know from (iv) that the norm of the second term in (20) is $O(1)$ as well. $\qquad\square$

Furthermore, the first part $\alpha^{(0)}\phi_0$ of the residual term ζ_0 satisfies the following orthogonality property that later turns out to be crucial:

Lemma 2.4. *For any multi-index $\mathbf{k} \in \mathbb{N}_0^d$ with $0 \le |\mathbf{k}| \le 2$, we have*

$$\Big\langle \alpha^{(0)}(q, Q; \cdot)\, \phi_0(q, p, Q, P, S; \cdot),\, \phi_{\mathbf{k}}(q, p, Q, P, S; \cdot) \Big\rangle = 0.$$

More specifically, $\alpha^{(0)}(q, Q; \cdot)\, \phi_0(q, p, Q, P, S; \cdot)$ is a linear combination of the third excited states

$$\Big\{ \phi_{\mathbf{n}}(q, p, Q, P, S; \cdot) \mid \mathbf{n} \in \mathbb{N}_0^d \text{ with } |\mathbf{n}| = 3 \Big\}.$$

Proof. Substituting (7) into the expression (19b) for $\alpha^{(0)}$, we have, suppressing the variables in $\alpha^{(0)}$ and ϕ_0 for brevity,

$$\alpha^{(0)}\phi_0 = \left(\frac{\varepsilon^{-1/2}}{4} Q_{jl}\overline{Q}_{kl} D_{ikj}^3 V(q)(x-q)_i - \frac{\varepsilon^{-3/2}}{3!} D_{ijk}^3 V(q)(x-q)_{ijk}^3 \right)\phi_0.$$

Using (12) and noting (15), the first term becomes

$$\frac{\varepsilon^{-1/2}}{4} Q_{jl}\overline{Q}_{kl} D_{ikj}^3 V(q)(x-q)_i \phi_0 = \frac{1}{4\sqrt{2}} D_{ikj}^3 V(q) Q_{jl}\overline{Q}_{kl} Q_{in} \mathscr{A}_n^* \phi_0$$

$$= \frac{1}{4\sqrt{2}} \mathcal{T}_{ijk} \overline{Q}_{il} Q_{jl} Q_{kn} \phi_{\mathbf{e}_n}, \tag{22}$$

where we used the shorthand $\mathcal{T}_{ijk} := D_{ijk}^3 V(q)$ and its symmetry with respect to permutations of the indices as well as (14). On the other hand, after similar but more tedious calculations (see Appendix A.2), the second term becomes

$$-\frac{\varepsilon^{-3/2}}{3!} D_{ijk}^3 V(q)(x-q)_{ijk}^3 \phi_0 = -\frac{1}{4\sqrt{3}} \mathcal{T}_{ijk} Q_{il} Q_{jm} Q_{kn} \phi_{\mathbf{e}_l+\mathbf{e}_m+\mathbf{e}_n} - \frac{1}{4\sqrt{2}} \mathcal{T}_{ijk} \overline{Q}_{il} Q_{jl} Q_{kn} \phi_{\mathbf{e}_n}. \tag{23}$$

As a result,

$$\alpha^{(0)}\phi_0 = -\frac{1}{4\sqrt{3}} \mathcal{T}_{ijk} Q_{il} Q_{jm} Q_{kn} \phi_{\mathbf{e}_l+\mathbf{e}_m+\mathbf{e}_n},$$

which is a linear combination of the third excited states. $\qquad\square$

Notice the role played by the first term in $\alpha^{(0)}$ in canceling the term with the first excited states, and recall that this term (22) came from the $O(\varepsilon)$ correction term in the potential in our dynamics (6a).

Remark 2.5. What if one uses the classical Hamiltonian system for $t \mapsto (q(t), p(t))$ as in (4) of Hagedorn? Then the function α in the residual term ζ_0 becomes

$$\alpha(q, Q; x) = \varepsilon^{-3/2}\left(\sum_{k=0}^{2} \frac{1}{k!} D^k V(q) \cdot (x-q)^k - V(x) \right)$$

$$= -\varepsilon^{-3/2}\left(\frac{1}{3!} D^3 V(q) \cdot (x-q)^3 + \frac{1}{4!} D^4 V(\sigma_1(x, q)) \cdot (x-q)^4 \right), \tag{24}$$

where $\sigma_1(x, q)$ is defined in Lemma 2.1. The absence of the term coming from the correction term indicates that there is no cancellation of those terms involving the first excited states. Indeed, as we shall see in Remark 4.1 of Section 4.2, this residual term does not enjoy the same property as ours does; it turns out to be detrimental in the error estimate.

3. Time Evolution of the Hagedorn Wave Packets

3.1. **Overview.** While the main focus of the paper is the Gaussian (2) and its associated equations (6) for the parameters, it turns out that the proof of the main result requires some analysis on the time evolution of some other Hagedorn wave packets as well. Therefore, in this section, we derive the Schrödinger-type evolution equation for the Hagedorn wave packets as opposed to just the Gaussian.

3.2. **Evolution Equation of the Hagedorn Wave Packets.** Lemma 2.1 applies only to the Gaussian wave packet ϕ_0. However, it turns out that Lemma 2.1 generalizes to $\phi_{\mathbf{n}}$ with any $\mathbf{n} \in \mathbb{N}_0^d$:

Proposition 3.1. *Let us define, for* $\mathbf{n} \in \mathbb{N}_0^d$,

$$\phi_{\mathbf{n}}(t, x) := \phi_{\mathbf{n}}(q(t), p(t), Q(t), P(t), S(t); x),$$

where $t \mapsto (q(t), p(t), Q(t), P(t), S(t))$ *satisfies* (6). *Then* $\phi_{\mathbf{n}}(t, x)$ *satisfies the Schrödinger-type equation*

$$\mathrm{i}\varepsilon \frac{\partial}{\partial t} \phi_{\mathbf{n}}(t, x) = \hat{H} \phi_{\mathbf{n}}(t, x) + \varepsilon^{3/2} \zeta_{\mathbf{n}}(t, x), \tag{25}$$

where

$$\zeta_{\mathbf{n}}(t, x) := \alpha(q(t), Q(t); x) \phi_{\mathbf{n}}(t, x) \tag{26}$$

with the same α *defined in* (18).

This result follows easily from the following lemma regarding the time evolution of the raising operator:

Lemma 3.2. *Suppose that* $t \mapsto (q(t), p(t), Q(t), P(t), S(t))$ *satisfies* (6) *and let us write*

$$\mathscr{A}^*(t) := \mathscr{A}^*(q(t), p(t), Q(t), P(t)).$$

Then its time evolution is governed by

$$\mathrm{i}\varepsilon \frac{d}{dt} \mathscr{A}^*(t) + \left[\mathscr{A}^*(t), \hat{H} \right] = \frac{\varepsilon^2}{\sqrt{2}} Q^*(t) \partial_x \alpha(q(t), Q(t); x)$$

as operators with domain $\mathscr{S}(\mathbb{R}^d)$.

Proof. It is straightforward to see, in view of (6) and (10), that

$$\mathrm{i}\varepsilon \frac{d}{dt} \mathscr{A}^*(t) = \sqrt{\frac{\varepsilon}{2}} \Big(-\dot{P}^*(t)(x - q(t)) + P^*(t)\dot{q}(t) + \dot{Q}^*(t)(\hat{p} - p(t)) - Q^*(t)\dot{p}(t) \Big)$$

$$= \sqrt{\frac{\varepsilon}{2}} \Big(Q^*(t) D^2 V(q(t))(x - q(t)) + P^*(t)\hat{p} + Q^*(t) \Big(DV(q(t)) + \varepsilon \, \partial_q V^{(1)}(q(t), Q(t)) \Big) \Big),$$

whereas, for $j \in \{1, \ldots, d\}$,

$$\left[\mathscr{A}_j^*(q, p, Q, P), \hat{H}\right] = \frac{\mathrm{i}}{\sqrt{2\varepsilon}}\left(\frac{1}{2}P_{jk}^*[x_k, \hat{p}^2] - Q_{jk}^*[\hat{p}_k, V(x)]\right)$$

$$= -\sqrt{\frac{\varepsilon}{2}}\left(P_{jk}^*\hat{p}_k + Q_{jk}^*\frac{\partial V}{\partial x_k}(x)\right)$$

and so

$$\left[\mathscr{A}^*(t), \hat{H}\right] = -\sqrt{\frac{\varepsilon}{2}}(P^*(t)\hat{p} + Q^*(t)DV(x)).$$

Therefore,

$$\mathrm{i}\varepsilon\frac{d}{dt}\mathscr{A}^*(t) + \left[\mathscr{A}^*(t), \hat{H}\right]$$

$$= \sqrt{\frac{\varepsilon}{2}}\,Q^*(t)\Big(D^2V(q(t))(x - q(t)) + DV(q(t)) + \varepsilon\,\partial_q V^{(1)}(q(t), Q(t)) - DV(x)\Big)$$

$$= \sqrt{\frac{\varepsilon}{2}}\,Q^*(t)\frac{\partial}{\partial x}\Big(\varepsilon\,\partial_q V^{(1)}(q(t), Q(t)) \cdot (x - q(t))$$

$$+ V(q(t)) + DV(q(t)) \cdot (x - q(t)) + \frac{1}{2}D^2V(q(t)) \cdot (x - q(t))^2 - V(x)\Big)$$

$$= \frac{\varepsilon^2}{\sqrt{2}}\,Q^*(t)\partial_x\alpha(q(t), Q(t); x). \qquad \square$$

Proof of Proposition 3.1. By induction on $\mathbf{n} \in \mathbb{N}_0^d$. Lemma 2.1 shows that the assertion holds for $\mathbf{n} = 0$. Let us suppose that the assertion holds for $\mathbf{n} \in \mathbb{N}_0^d$ and show that it holds for $\mathbf{n} + \mathbf{e}_j$ for any $j \in \{1, \ldots, d\}$. Using (14) and the above lemma,

$$\sqrt{n_j + 1}\left(\mathrm{i}\varepsilon\frac{\partial}{\partial t}\phi_{\mathbf{n}+\mathbf{e}_j}(t, x) - \hat{H}\phi_{\mathbf{n}+\mathbf{e}_j}(t, x)\right)$$

$$= \mathrm{i}\varepsilon\frac{\partial}{\partial t}\big(\mathscr{A}_j^*(t)\,\phi_{\mathbf{n}}(t, x)\big) - \hat{H}\mathscr{A}_j^*(t)\,\phi_{\mathbf{n}}(t, x)$$

$$= \left(\mathrm{i}\varepsilon\frac{d}{dt}\mathscr{A}_j^*(t) + \left[\mathscr{A}_j^*(t), \hat{H}\right]\right)\phi_{\mathbf{n}}(t, x) + \mathscr{A}_j^*(t)\left(\mathrm{i}\varepsilon\frac{\partial}{\partial t}\phi_{\mathbf{n}}(t, x) - \hat{H}\phi_{\mathbf{n}}(t, x)\right)$$

$$= \frac{\varepsilon^2}{\sqrt{2}}Q_{jk}^*(t)\frac{\partial}{\partial x_k}\alpha(q(t), Q(t); x)\phi_{\mathbf{n}}(t, x) + \varepsilon^{3/2}\mathscr{A}_j^*(t)(\alpha(q(t), Q(t); x)\phi_{\mathbf{n}}(t, x)).$$

However, since every term in $\mathscr{A}^*$ (see (10)) except the one with $\hat{p}$ is a multiplication operator,

$$\mathscr{A}^*(q, p, Q, P)(\alpha(q, Q; x)\phi_{\mathbf{n}}(q, p, Q, P, S; x))$$

$$= \alpha(q, Q; x)\mathscr{A}^*(q, p, Q, P)\phi_{\mathbf{n}}(q, p, Q, P, S; x)$$

$$- \frac{\mathrm{i}}{\sqrt{2\varepsilon}}Q^*\hat{p}(\alpha(q, Q; x))\phi_{\mathbf{n}}(q, p, Q, P, S; x)$$

$$= \sqrt{n_j + 1}\,\alpha(q, Q; x)\phi_{\mathbf{n}+\mathbf{e}_j}(q, p, Q, P, S; x) - \sqrt{\frac{\varepsilon}{2}}Q^*\partial_x\alpha(q, Q; x)\phi_{\mathbf{n}}(q, p, Q, P, S; x).$$

Therefore, we obtain

$$\mathrm{i}\varepsilon\frac{\partial}{\partial t}\phi_{\mathbf{n}+\mathbf{e}_j}(t, x) - \hat{H}\phi_{\mathbf{n}+\mathbf{e}_j}(t, x) = \varepsilon^{3/2}\alpha(q(t), Q(t); x)\phi_{\mathbf{n}+\mathbf{e}_j}(t, x). \qquad \square$$

3.3. **Errors in Wave Functions.** Let us first note that, in what follows, we will suppress the spatial variables x for brevity. We also note that the assumption that the potential V is bounded from below guarantees that there exists a self-adjoint extension of the Schrödinger operator $\hat{H}$ so that the unitary operators of the form $e^{-i\hat{H}t/\varepsilon}$ with $t \in \mathbb{R}$ would make sense.

Now we would like to compare the exact solution $t \mapsto e^{-i\hat{H}(t-s)/\varepsilon}\phi_{\mathbf{n}}(s)$ of the Schrödinger equation (1a) and the wave packet $t \mapsto \phi_{\mathbf{n}}(t)$, both with the initial wave function being $\phi_{\mathbf{n}}(s)$ with any $\mathbf{n} \in \mathbb{N}_0^d$ at time $s \in \mathbb{R}$. To that end, let us define the difference between them (i.e., error in wave functions): For any $\mathbf{n} \in \mathbb{N}_0^d$, and any $s, t \in \mathbb{R}$ (for which both $\phi_{\mathbf{n}}(s)$ and $\phi_{\mathbf{n}}(t)$ are defined),

$$\mathcal{Z}_{\mathbf{n}}(t,s) := e^{-i\hat{H}(t-s)/\varepsilon}\phi_{\mathbf{n}}(s) - \phi_{\mathbf{n}}(t), \qquad \mathcal{Z}_{\mathbf{n}}(t) := \mathcal{Z}_{\mathbf{n}}(t,0) = e^{-i\hat{H}t/\varepsilon}\phi_{\mathbf{n}}(0) - \phi_{\mathbf{n}}(t). \tag{27}$$

The following lemma is critical in finding an estimate of these errors:

Lemma 3.3. *For any* $\mathbf{n} \in \mathbb{N}_0^d$, $\|\zeta_{\mathbf{n}}(t)\| = O(1)$.

Proof. The proof is almost identical to that of Lemma 2.3. In fact, one can show that, for any $\mathbf{m}, \mathbf{n} \in \mathbb{N}_0^d$,

$$\left\|\xi_1^{m_1}\dots\xi_d^{m_d}\,\phi_{\mathbf{n}}(t)\right\| = O(1)$$

because $\phi_{\mathbf{n}}$ is ϕ_0 multiplied by an $|\mathbf{n}|$-th order polynomial of $\xi = \varepsilon^{-1/2}(x - q(t))$; see also Hagedorn [9, Eq. (3.30)]. It implies that those arguments with ϕ_0 from Lemma 2.3 still apply upon replacing ϕ_0 by $\phi_{\mathbf{n}}$. Hence it follows that $\left\|\alpha^{(0)}(q(t), Q(t))\phi_{\mathbf{n}}(t)\right\| = O(1)$ as well as that $\left\|\alpha^{(1)}(q(t))\phi_{\mathbf{n}}(t)\right\| = O(1)$ as well. $\qquad\square$

As a result, we have an expression and an estimate for $\mathcal{Z}_{\mathbf{n}}$ as follows:

Proposition 3.4. *The errors defined in* (27) *can be written in terms of the residual term* $\zeta_{\mathbf{n}}$ *from* (26) *as follows: For any* $\mathbf{n} \in \mathbb{N}_0^d$ *and any* $s, t \in \mathbb{R}$ *for which* $\phi_{\mathbf{n}}(s)$ *and* $\phi_{\mathbf{n}}(t)$ *are defined,*

$$\mathcal{Z}_{\mathbf{n}}(t,s) := e^{-i\hat{H}(t-s)/\varepsilon}\phi_{\mathbf{n}}(s) - \phi_{\mathbf{n}}(t) = i\,\varepsilon^{1/2}\int_s^t e^{-i\hat{H}(t-\tau)/\varepsilon}\zeta_{\mathbf{n}}(\tau)\,d\tau, \tag{28}$$

and hence $\|\mathcal{Z}_{\mathbf{n}}(t,s)\| = O(\varepsilon^{1/2})$ *in the sense that there exists some function* $\mathscr{C}$ *such that* $\|\mathcal{Z}_{\mathbf{n}}(t,s)\| \leq \mathscr{C}(t,s)\varepsilon^{1/2}$.

Proof. This is essentially the same as the proof of Hagedorn [9, Lemma 2.8], but we briefly reproduce it here for completeness. Using the Schrödinger-type equation (25) satisfied by $\tau \mapsto \phi_{\mathbf{n}}(\tau)$, we have

$$\frac{\partial}{\partial \tau}\left(e^{-i\hat{H}(t-\tau)/\varepsilon}\phi_{\mathbf{n}}(\tau)\right) = -i\,\varepsilon^{1/2}e^{-i\hat{H}(t-\tau)/\varepsilon}\zeta_{\mathbf{n}}(\tau).$$

Integrating both sides with respect to τ over the time interval between s and t yields

$$\phi_{\mathbf{n}}(t) - e^{-i\hat{H}(t-s)/\varepsilon}\phi_{\mathbf{n}}(s) = -i\,\varepsilon^{1/2}\int_s^t e^{-i\hat{H}(t-\tau)/\varepsilon}\zeta_{\mathbf{n}}(\tau)\,d\tau.$$

The left hand side is $-\mathcal{Z}_{\mathbf{n}}(t,s)$, and so (28) follows. The estimate in norm follows by taking the norm of both sides of (28):

$$\|\mathcal{Z}_{\mathbf{n}}(t,s)\| \leq \varepsilon^{1/2}\int_s^t \|\zeta_{\mathbf{n}}(\tau)\|\,d\tau = O(\varepsilon^{1/2}).$$

due to the unitarity of $e^{-i\hat{H}(t-s)/\varepsilon}$ as well as Lemma 3.3. $\qquad\square$

Particularly, setting $s = 0$, we have the following:

Corollary 3.5. *Let* $t \mapsto \psi(t)$ *be the solution to the Schrödinger equation* (1a) *with the initial condition* $\psi(0) = \phi_{\mathbf{n}}(0)$ *with* $\mathbf{n} \in \mathbb{N}_0^d$. *Then*

$$\mathcal{Z}_{\mathbf{n}}(t) = \psi(t) - \phi_{\mathbf{n}}(t) = \mathrm{i}\,\varepsilon^{1/2} \int_0^t e^{-\mathrm{i}\hat{H}(t-s)/\varepsilon} \zeta_{\mathbf{n}}(s)\, ds, \tag{29}$$

and hence $\|\psi(t) - \phi_{\mathbf{n}}(t)\| = O(\varepsilon^{1/2})$.

The above result reproduces those estimates obtained by Hagedorn [6, 9] using our equations (6), and also indicates that using (6) in place of Hagedorn's (4) does not improve the errors in wave function in terms of L^2-norm—at least not with the above method of estimation. The reason why there is still a difference in the error estimates of the observables as stated in Theorem 1.1 is that our estimates involve a more detailed analysis of the residual term ζ_0 as opposed to just having an L^2-norm estimate of it.

4. Proof of Main Result

4.1. **Error Terms in Observables.** Let $t \mapsto \psi(t)$ be the exact solution of the initial value problem (1) of the Schrödinger equation. From the definition of $\mathcal{Z}_0$ in (27) with $\mathbf{n} = 0$, we have $\psi(t) = \phi_0(t) + \mathcal{Z}_0(t)$, and so, as we have shown in Section 1.5,

$$\langle \hat{x} \rangle(t) - q(t) = 2\operatorname{Re} \langle \mathcal{Z}_0(t), (\hat{x} - q(t))\phi_0(t)\rangle + \langle \mathcal{Z}_0(t), (\hat{x} - q(t))\mathcal{Z}_0(t)\rangle,$$

and similarly,

$$\langle \hat{p} \rangle(t) - p(t) = 2\operatorname{Re} \langle \mathcal{Z}_0(t), (\hat{p} - p(t))\phi_0(t)\rangle + \langle \mathcal{Z}_0(t), (\hat{p} - p(t))\mathcal{Z}_0(t)\rangle.$$

In the remaining subsections, we finish the proof of Theorem 1.1 by showing that the two terms on the right-hand side of each of the above equations are both $O(\varepsilon^{3/2})$.

4.2. **Estimates for First Error Term.** First we see that, using the expression (29) for $\mathcal{Z}_0$ and Fubini's Theorem,

$$\langle \mathcal{Z}_0(t), (\hat{x} - q(t))\phi_0(t)\rangle = -\mathrm{i}\,\varepsilon^{1/2} \int_0^t \left\langle e^{-\mathrm{i}\hat{H}(t-s)/\varepsilon} \zeta_0(s), (\hat{x} - q(t))\phi_0(t)\right\rangle ds. \tag{30}$$

However, we can rewrite the inner product inside the integral as follows using the relationship (12) between the operator $\hat{x} - q$ and the ladder operators: For any $i \in \{1, \ldots, d\}$,

$$\left\langle e^{-\mathrm{i}\hat{H}(t-s)/\varepsilon} \zeta_0(s), (\hat{x} - q(t))_i\, \phi_0(t)\right\rangle = \sqrt{\frac{\varepsilon}{2}} \left\langle e^{-\mathrm{i}\hat{H}(t-s)/\varepsilon} \zeta_0(s), \left(\overline{Q}_{ij}(t)\mathscr{A}_j(t) + Q_{ij}(t)\mathscr{A}_j^*(t)\right)\phi_0(t)\right\rangle$$

$$= \sqrt{\frac{\varepsilon}{2}}\, Q_{ij}(t) \left\langle e^{-\mathrm{i}\hat{H}(t-s)/\varepsilon} \zeta_0(s), \phi_{\mathbf{e}_j}(t)\right\rangle$$

$$= \sqrt{\frac{\varepsilon}{2}}\, Q_{ij}(t) \left\langle \zeta_0(s), e^{-\mathrm{i}\hat{H}(s-t)/\varepsilon} \phi_{\mathbf{e}_j}(t)\right\rangle$$

$$= \sqrt{\frac{\varepsilon}{2}}\, Q_{ij}(t) \left(\langle \zeta_0(s), \phi_{\mathbf{e}_j}(s)\rangle + \langle \zeta_0(s), \mathcal{Z}_{\mathbf{e}_j}(s, t)\rangle\right)$$

where $\mathscr{A}^*(t)$ is defined in Lemma 3.2 and similarly for $\mathscr{A}(t)$; we used (14) for the second equality, and (28) (s and t swapped) with $\mathbf{n} = \mathbf{e}_j$ for the last equality, i.e., for any $j \in \{1, \dots, d\}$, $e^{-i\hat{H}(s-t)/\varepsilon}\phi_{\mathbf{e}_j}(t) = \phi_{\mathbf{e}_j}(s) + \mathcal{Z}_{\mathbf{e}_j}(s, t)$.

Let us evaluate the above two terms: First, recalling the formulas (17) and (19) for ζ_0 and exploiting the orthogonality in Lemma 2.4, we have, for any $j \in \{1, \dots, d\}$,

$$
\begin{aligned}
\left\langle \zeta_0(s), \phi_{\mathbf{e}_j}(s) \right\rangle &= \left\langle \alpha^{(0)}(q(s), Q(s))\phi_0(s), \phi_{\mathbf{e}_j}(s) \right\rangle + \varepsilon^{1/2} \left\langle \alpha^{(1)}(q(s))\phi_0(s), \phi_{\mathbf{e}_j}(s) \right\rangle \\
&= \varepsilon^{1/2} \left\langle \alpha^{(1)}(q(s))\phi_0(s), \phi_{\mathbf{e}_j}(s) \right\rangle,
\end{aligned}
$$

and thus by the Cauchy–Schwarz inequality and Lemma 2.3 (ii),

$$
\left| \left\langle \zeta_0(s), \phi_{\mathbf{e}_j}(s) \right\rangle \right| \leq \varepsilon^{1/2} \left\| \alpha^{(1)}(q(s))\phi_0(s) \right\| \left\| \phi_{\mathbf{e}_j}(s) \right\| = O(\varepsilon^{1/2}).
$$

On the other hand, again by the Cauchy–Schwarz inequality, Lemma 2.3 (iii), and Proposition 3.4, we have, for any $j \in \{1, \dots, d\}$,

$$
\left| \left\langle \zeta_0(s), \mathcal{Z}_{\mathbf{e}_j}(s, t) \right\rangle \right| \leq \left\| \zeta_0(s) \right\| \left\| \mathcal{Z}_{\mathbf{e}_j}(s, t) \right\| = O(\varepsilon^{1/2}).
$$

Hence we see that

$$
\left| \left\langle e^{-i\hat{H}(t-s)/\varepsilon}\zeta_0(s), (\hat{x} - q(t))_i \phi_0(t) \right\rangle \right| = O(\varepsilon),
$$

and therefore (30) yields, for any $i \in \{1, \dots, d\}$,

$$
\left| \left\langle \mathcal{Z}_0(t), (\hat{x} - q(t))_i \phi_0(t) \right\rangle \right| \leq \varepsilon^{1/2} \int_0^t \left| \left\langle e^{-i\hat{H}(t-s)/\varepsilon}\zeta_0(s), (\hat{x} - q(t))\phi_0(t) \right\rangle \right| ds = O(\varepsilon^{3/2}).
$$

Using the relationship (13) between the operator $\hat{p} - p$ and the ladder operators, we can proceed in the same way to obtain

$$
\left| \left\langle \mathcal{Z}_0(t), (\hat{p} - p(t))_i \phi_0(t) \right\rangle \right| = O(\varepsilon^{3/2})
$$

for any $i \in \{1, \dots, d\}$ as well.

Remark 4.1. What if one uses the classical Hamiltonian system for (q, p) as in (4) of Hagedorn? As discussed in Remark 2.5, we have α as shown in (24). Then, as shown in Appendix A.3, we have

$$
\left\langle \zeta_0(s), \phi_{\mathbf{e}_j}(s) \right\rangle = O(1)
$$

in this case as opposed to $O(\varepsilon^{1/2})$. Indeed the leading $O(1)$ term (see (33) in Appendix A.3) is exactly the term canceled due to the first term in (18) coming from the correction term in our case. This underscores the importance of the correction term alluded in Remark 2.5. As a result, the above estimates become

$$
\left| \left\langle \mathcal{Z}_0(t), (\hat{x} - q(t))_i \phi_0(t) \right\rangle \right| = O(\varepsilon), \qquad \left| \left\langle \mathcal{Z}_0(t), (\hat{p} - p(t))_i \phi_0(t) \right\rangle \right| = O(\varepsilon).
$$

as opposed to $O(\varepsilon^{3/2})$.

4.3. **Estimates for Second Error Term.** It now remains to show that, for any $i \in \{1, \ldots, d\}$,

$$|\langle \mathcal{Z}_0(t), (\hat{x} - q(t))_i \mathcal{Z}_0(t) \rangle| = O(\varepsilon^{3/2}), \qquad |\langle \mathcal{Z}_0(t), (\hat{p} - p(t))_i \mathcal{Z}_0(t) \rangle| = O(\varepsilon^{3/2}).$$

From (1a) and (16), we see that

$$\dot{\mathcal{Z}}_0(t) = \dot{\psi}(t) - \dot{\phi}_0(t) = -\frac{\mathrm{i}}{\varepsilon} \hat{H} \mathcal{Z}_0(t) + \mathrm{i} \varepsilon^{1/2} \zeta_0(t).$$

Therefore,

$$\frac{d}{dt} ((\hat{x} - q(t)) \mathcal{Z}_0(t)) = -\dot{q}(t) \mathcal{Z}_0(t) + (\hat{x} - q(t)) \dot{\mathcal{Z}}_0(t)$$

$$= -p(t) \mathcal{Z}_0(t) - \frac{\mathrm{i}}{\varepsilon} (\hat{x} - q(t)) \hat{H} \mathcal{Z}_0(t) + \mathrm{i} \varepsilon^{1/2} (\hat{x} - q(t)) \zeta_0(t)$$

$$= -p(t) \mathcal{Z}_0(t) - \frac{\mathrm{i}}{\varepsilon} \left(\left[\hat{x}, \hat{H} \right] + \hat{H}(\hat{x} - q(t)) \right) \mathcal{Z}_0(t) + \mathrm{i} \varepsilon^{1/2} (\hat{x} - q(t)) \zeta_0(t)$$

$$= (\hat{p} - p(t)) \mathcal{Z}_0(t) - \frac{\mathrm{i}}{\varepsilon} \hat{H}(\hat{x} - q(t)) \mathcal{Z}_0(t) + \mathrm{i} \varepsilon \hat{\xi}(t) \zeta_0(t),$$

where the last equality follows from

$$\left[\hat{x}, \hat{H} \right] = \left[\hat{x}, \frac{\hat{p}^2}{2} \right] = \mathrm{i} \varepsilon \hat{p},$$

and also setting $\hat{\xi}(t) := \varepsilon^{-1/2} (\hat{x} - q(t))$. Applying $e^{\mathrm{i} \hat{H} t / \varepsilon}$ to both sides, we have

$$e^{\mathrm{i} \hat{H} t / \varepsilon} \frac{d}{dt} ((\hat{x} - q(t)) \mathcal{Z}_0(t)) + e^{\mathrm{i} \hat{H} t / \varepsilon} \frac{\mathrm{i}}{\varepsilon} \hat{H}(\hat{x} - q(t)) \mathcal{Z}_0(t) = e^{\mathrm{i} \hat{H} t / \varepsilon} (\hat{p} - p(t)) \mathcal{Z}_0(t) + \mathrm{i} \varepsilon \, e^{\mathrm{i} \hat{H} t / \varepsilon} \hat{\xi}(t) \zeta_0(t)$$

or

$$\frac{d}{dt} \left(e^{\mathrm{i} \hat{H} t / \varepsilon} (\hat{x} - q(t)) \mathcal{Z}_0(t) \right) = e^{\mathrm{i} \hat{H} t / \varepsilon} (\hat{p} - p(t)) \mathcal{Z}_0(t) + \mathrm{i} \varepsilon \, e^{\mathrm{i} \hat{H} t / \varepsilon} \hat{\xi}(t) \zeta_0(t).$$

Integrating both sides on the interval $[0, t]$ and using $\mathcal{Z}_0(0) = 0$, we have

$$e^{\mathrm{i} \hat{H} t / \varepsilon} (\hat{x} - q(t)) \mathcal{Z}_0(t) = \int_0^t e^{\mathrm{i} \hat{H} s / \varepsilon} (\hat{p} - p(s)) \mathcal{Z}_0(s) \, ds + \mathrm{i} \varepsilon \int_0^t e^{\mathrm{i} \hat{H} s / \varepsilon} \hat{\xi}(s) \zeta_0(s) \, ds.$$

Taking the L^2-norm of the i-th component of both sides with $i \in \{1, \ldots, d\}$,

$$\|(\hat{x} - q(t))_i \mathcal{Z}_0(t)\| \leq \int_0^t \|(\hat{p} - p(s))_i \mathcal{Z}_0(s)\| \, ds + \varepsilon \int_0^t \left\| \hat{\xi}_i(s) \zeta_0(s) \right\| ds$$

$$= \int_0^t \|(\hat{p} - p(s))_i \mathcal{Z}_0(s)\| \, ds + O(\varepsilon), \tag{31}$$

where we used the estimate $\left\| \hat{\xi}_i(s) \zeta_0(s) \right\| = O(1)$ from Lemma 2.3 (iv).

Similarly,

$$
\begin{aligned}
\frac{d}{dt}\big((\hat{p} - p(t))\mathcal{Z}_0(t)\big) &= -\dot{p}(t)\,\mathcal{Z}_0(t) + (\hat{p} - p(t))\dot{\mathcal{Z}}_0(t) \\
&= \Big(DV(q(t)) + \varepsilon\,\partial_q V^{(1)}(q(t), Q(t))\Big)\mathcal{Z}_0(t) - \frac{\mathrm{i}}{\varepsilon}(\hat{p} - p(t))\hat{H}\mathcal{Z}_0(t) \\
&\quad + \mathrm{i}\,\varepsilon^{1/2}(\hat{p} - p(t))\zeta_0(t) \\
&= \Big(DV(q(t)) + \varepsilon\,\partial_q V^{(1)}(q(t), Q(t))\Big)\mathcal{Z}_0(t) - \frac{\mathrm{i}}{\varepsilon}\Big(\big[\hat{p}, \hat{H}\big] + \hat{H}(\hat{p} - p(t))\Big)\mathcal{Z}_0(t) \\
&\quad + \mathrm{i}\,\varepsilon^{1/2}(\hat{p} - p(t))\zeta_0(t) \\
&= \Big(DV(q(t)) - DV(x) + \varepsilon\,\partial_q V^{(1)}(q(t), Q(t))\Big)\mathcal{Z}_0(t) - \frac{\mathrm{i}}{\varepsilon}\hat{H}(\hat{p} - p(t))\mathcal{Z}_0(t) \\
&\quad + \mathrm{i}\,\varepsilon^{1/2}(\hat{p} - p(t))\zeta_0(t) \\
&= \Big(-D^2 V(\sigma_2(x, q(t)))(\hat{x} - q(t)) + \varepsilon\,\partial_q V^{(1)}(q(t), Q(t))\Big)\mathcal{Z}_0(t) \\
&\quad - \frac{\mathrm{i}}{\varepsilon}\hat{H}(\hat{p} - p(t))\mathcal{Z}_0(t) + \mathrm{i}\,\varepsilon\,\hat{\eta}(t)\zeta_0(t),
\end{aligned}
$$

where the second last equality follows from

$$
\big[\hat{p}, \hat{H}\big] = [\hat{p}, V(x)] = -\mathrm{i}\,\varepsilon\,DV(x),
$$

and $\sigma_2(x, q(t))$ is a point in the segment joining x and $q(t)$ in $\mathbb{R}^d$; we also set $\hat{\eta}(t) := \varepsilon^{-1/2}(\hat{p} - p(t))$. Applying $e^{\mathrm{i}\hat{H}t/\varepsilon}$ to both sides,

$$
\begin{aligned}
\frac{d}{dt}&\Big(e^{\mathrm{i}\hat{H}t/\varepsilon}(\hat{p} - p(t))\mathcal{Z}_0(t)\Big) \\
&= e^{\mathrm{i}\hat{H}t/\varepsilon}\Big(-D^2 V(\sigma_2(x, q(t)))(\hat{x} - q(t)) + \varepsilon\,\partial_q V^{(1)}(q(t), Q(t))\Big)\mathcal{Z}_0(t) + \mathrm{i}\,\varepsilon\,e^{\mathrm{i}\hat{H}t/\varepsilon}\hat{\eta}(t)\zeta_0(t).
\end{aligned}
$$

Integrating both sides on $[0, t]$, we have

$$
\begin{aligned}
e^{\mathrm{i}\hat{H}t/\varepsilon}(\hat{p} - p(t))\mathcal{Z}_0(t) &= \int_0^t e^{\mathrm{i}\hat{H}s/\varepsilon}\Big(-D^2 V(\sigma_2(x, q(s)))(\hat{x} - q(s)) + \varepsilon\,\partial_q V^{(1)}(q(s), Q(s))\Big)\mathcal{Z}_0(s)\,ds \\
&\quad + \mathrm{i}\,\varepsilon \int_0^t e^{\mathrm{i}\hat{H}s/\varepsilon}\hat{\eta}(s)\zeta_0(s)\,ds.
\end{aligned}
$$

Taking the L^2-norm of the i-th component of both sides for any $i \in \{1, \ldots, d\}$,

$$
\begin{aligned}
\|(\hat{p} - p(t))_i \mathcal{Z}_0(t)\| &\le \int_0^t \left\| \sum_{j=1}^d D_{ij}^2 V(\sigma_2(x, q(s)))(\hat{x} - q(s))_j \mathcal{Z}_0(s) \right\| ds \\
&\quad + \varepsilon \int_0^t \Big(\big|\partial_{q_i} V^{(1)}(q(s), Q(s))\big|\,\|\mathcal{Z}_0(s)\| + \|\hat{\eta}_i(s)\zeta_0(s)\| \Big)\,ds \\
&\le \int_0^t \sum_{j=1}^d \big\|D_{ij}^2 V(\sigma_2(x, q(s)))(\hat{x} - q(s))_j \mathcal{Z}_0(s)\big\|\,ds + O(\varepsilon) \\
&\le C_3 \int_0^t \sum_{j=1}^d \|(\hat{x} - q(s))_j \mathcal{Z}_0(s)\|\,ds + O(\varepsilon), \tag{32}
\end{aligned}
$$

where we used the following bound of the second derivative of V

$$C_3 := \max_{1 \leq i,j \leq d} \sup_{x \in \mathbb{R}^d} \left| D_{ij}^2 V(x) \right|$$

as well as the following: $\|\mathcal{Z}_0(s)\| = O(\varepsilon^{1/2})$ from Corollary 3.5 and $\|\hat{\eta}_i(s)\zeta_0(s)\| = O(1)$ from Lemma 2.3 (v).

Now, let us set

$$f(t) := \sum_{i=1}^{d} \left(\|(\hat{x} - q(t))_i \mathcal{Z}_0(t)\| + \|(\hat{p} - p(t))_i \mathcal{Z}_0(t)\| \right).$$

Then, using (31) and (32), we have

$$f(t) \leq \int_0^t \sum_{i=1}^{d} \|(\hat{p} - p(s))_i \mathcal{Z}_0(s)\| \, ds + d \, C_3 \int_0^t \sum_{j=1}^{d} \|(\hat{x} - q(s))_j \mathcal{Z}_0(s)\| \, ds + O(\varepsilon)$$

$$\leq C_4 \int_0^t \sum_{i=1}^{d} \left(\|(\hat{p} - p(s))_i \mathcal{Z}_0(s)\| + \|(\hat{x} - q(s))_i \mathcal{Z}_0(s)\| \right) ds + O(\varepsilon)$$

$$= C_4 \int_0^t f(s) \, ds + O(\varepsilon),$$

where we defined $C_4 := \max\{1, d \, C_3\}$. Therefore, by Gronwall's inequality [5], we obtain

$$f(t) \leq O(\varepsilon) \exp(C_4 t),$$

that is, $f(t) = O(\varepsilon)$, and so we have, for any $i \in \{1, \dots, d\}$,

$$\|(\hat{x} - q(t))_i \mathcal{Z}_0(t)\| = O(\varepsilon), \qquad \|(\hat{p} - p(t))_i \mathcal{Z}_0(t)\| = O(\varepsilon).$$

As a result, by the Cauchy–Schwarz inequality, we obtain

$$|\langle \mathcal{Z}_0(t), (\hat{x} - q(t))_i \mathcal{Z}_0(t) \rangle| \leq \|\mathcal{Z}_0(t)\| \, \|(\hat{x} - q(t))_i \mathcal{Z}_0(t)\| = O(\varepsilon^{3/2}),$$

and similarly

$$|\langle \mathcal{Z}_0(t), (\hat{p} - p(t))_i \mathcal{Z}_0(t) \rangle| = O(\varepsilon^{3/2})$$

as well.

Therefore, we conclude that, for any $i \in \{1, \dots, d\}$,

$$\langle \hat{x}_i \rangle(t) - q_i(t) = O(\varepsilon^{3/2}), \qquad \langle \hat{p}_i \rangle(t) - p_i(t) = O(\varepsilon^{3/2}).$$

Remark 4.2. If one uses the classical Hamiltonian system for (q, p) as in (4) of Hagedorn, then the estimate of the second error term proceeds similarly and hence it is still $O(\varepsilon^{3/2})$—the difference is the expression of the residual term ζ_0 as well as the absence of the term with $V^{(1)}$. These do not affect the estimate of the second error term—it is still $O(\varepsilon^{3/2})$. However, as discussed in Remark 4.1, the estimate of the first error term now becomes $O(\varepsilon)$ and hence the total error is $O(\varepsilon)$.

Appendix A. Additional Details

A.1. **Details of Proof of Lemma 2.3 (v).** We suppress the time dependence of ζ_0, ϕ_0, and (q, p, Q, P) for brevity here. First we have

$$(\hat{p}_i - p_i)\zeta_0(x) = (\hat{p}_i \alpha(q, Q; x))\phi_0(x) + \alpha(q, Q; x)(\hat{p}_i \phi_0(x)) - p_i \zeta_0(x).$$

However,

$$\begin{aligned}
\hat{p}_i \alpha(q, Q; x) &= -\mathrm{i}\,\varepsilon \frac{\partial}{\partial x_i} \alpha(q, Q; x) \\
&= \varepsilon^{1/2} \beta_i(q, Q; x) \\
&= \varepsilon^{1/2} \Big(\beta_i^{(0)}(q, Q; x) + \varepsilon^{1/2} \beta_i^{(1)}(q; x)\Big),
\end{aligned}$$

where we defined

$$\beta_i(q, Q; x) := \beta_i^{(0)}(q, Q; x) + \varepsilon^{1/2} \beta_i^{(1)}(q; x),$$

$$\beta_i^{(0)}(q, Q; x) := -\mathrm{i}\,\varepsilon^{1/2} \frac{\partial}{\partial x_i} \alpha^{(0)}(q, Q; x), \qquad \beta_i^{(1)}(q; x) := -\mathrm{i}\,\varepsilon^{1/2} \frac{\partial}{\partial x_i} \alpha^{(1)}(q; x),$$

which yield the expressions in (21). On the other hand, using the expression (2) for ϕ_0,

$$\hat{p}\phi_0(x) = \big(PQ^{-1}(x - q) + p\big)\phi_0(x).$$

Therefore,

$$(\hat{p}_i - p_i)\zeta_0(x) = \Big(\varepsilon^{1/2}\beta_i(q, Q; x) + \alpha(q, Q; x)(PQ^{-1})_{ij}(x - q)_j\Big)\phi_0(x),$$

and thus

$$\hat{\eta}_i \zeta_0(x) = \big(\beta_i(q, Q; x) + \alpha(q, Q; x)(PQ^{-1})_{ij}\xi_j\big)\phi_0(x),$$

which gives (20).

A.2. **Details of Proof of Lemma 2.4.** Let us show the detailed derivation of (23). We first have, using (12),

$$\frac{\varepsilon^{-3/2}}{3!} D_{ijk}^3 V(q)(x-q)_{ijk}^3 \phi_0 = \frac{1}{12\sqrt{2}} \mathcal{T}_{ijk}(\overline{Q}_{il}\mathscr{A}_l + Q_{il}\mathscr{A}_l^*)(\overline{Q}_{jm}\mathscr{A}_m + Q_{jm}\mathscr{A}_m^*)(\overline{Q}_{kn}\mathscr{A}_n + Q_{kn}\mathscr{A}_n^*)\phi_0.$$

Notice that, applying lowering operator(s) twice and a raising operator once—regardless of the order—to ϕ_0 results in zero, and the same goes with lowering operator(s) thrice as well. Therefore, we have

$$\begin{aligned}
(\overline{Q}_{il}\mathscr{A}_l &+ Q_{il}\mathscr{A}_l^*)(\overline{Q}_{jm}\mathscr{A}_m + Q_{jm}\mathscr{A}_m^*)(\overline{Q}_{kn}\mathscr{A}_n + Q_{kn}\mathscr{A}_n^*)\phi_0 \\
&= Q_{il}Q_{jm}Q_{kn}\mathscr{A}_l^*\mathscr{A}_m^*\mathscr{A}_n^*\phi_0 \\
&\quad + \big(\overline{Q}_{il}Q_{jm}Q_{kn}\mathscr{A}_l\mathscr{A}_m^*\mathscr{A}_n^* + Q_{il}\overline{Q}_{jm}Q_{kn}\mathscr{A}_l^*\mathscr{A}_m\mathscr{A}_n^* + Q_{il}Q_{jm}\overline{Q}_{kn}\mathscr{A}_l^*\mathscr{A}_m^*\mathscr{A}_n\big)\phi_0 \\
&= Q_{il}Q_{jm}Q_{kn}\mathscr{A}_l^*\mathscr{A}_m^*\mathscr{A}_n^*\phi_0 + \big(\overline{Q}_{il}Q_{jm}Q_{kn}\mathscr{A}_l\mathscr{A}_m^*\mathscr{A}_n^* + Q_{il}\overline{Q}_{jm}Q_{kn}\mathscr{A}_l^*\mathscr{A}_m\mathscr{A}_n^*\big)\phi_0,
\end{aligned}$$

where we used (15).

However, we may use (11) and (15) to simplify the second and last terms as follows:

$$\mathscr{A}_l\mathscr{A}_m^*\mathscr{A}_n^*\phi_0 = (\delta_{lm} + \mathscr{A}_m^*\mathscr{A}_l)\mathscr{A}_n^*\phi_0$$
$$= \delta_{lm}\mathscr{A}_n^*\phi_0 + \mathscr{A}_m^*(\delta_{ln} + \mathscr{A}_n^*\mathscr{A}_l)\phi_0$$
$$= (\delta_{lm}\mathscr{A}_n^* + \delta_{ln}\mathscr{A}_m^*)\phi_0,$$

and

$$\mathscr{A}_l^*\mathscr{A}_m\mathscr{A}_n^*\phi_0 = \mathscr{A}_l^*(\delta_{mn} + \mathscr{A}_n^*\mathscr{A}_m)\phi_0$$
$$= \delta_{mn}\mathscr{A}_l^*\phi_0.$$

Therefore,

$$(\overline{Q}_{il}\mathscr{A}_l + Q_{il}\mathscr{A}_l^*)(\overline{Q}_{jm}\mathscr{A}_m + Q_{jm}\mathscr{A}_m^*)(\overline{Q}_{kn}\mathscr{A}_n + Q_{kn}\mathscr{A}_n^*)\phi_0$$
$$= Q_{il}Q_{jm}Q_{kn}\mathscr{A}_l^*\mathscr{A}_m^*\mathscr{A}_n^*\phi_0 + \left(\overline{Q}_{il}Q_{jl}Q_{kn}\mathscr{A}_n^* + \overline{Q}_{il}Q_{jm}Q_{kl}\mathscr{A}_m^* + Q_{il}\overline{Q}_{jm}Q_{km}\mathscr{A}_l^*\right)\phi_0,$$

and so

$$\frac{\varepsilon^{-3/2}}{3!}D^3_{ijk}V(q)(x-q)^3_{ijk}\phi_0 = \frac{1}{12\sqrt{2}}\mathcal{T}_{ijk}Q_{il}Q_{jm}Q_{kn}\mathscr{A}_l^*\mathscr{A}_m^*\mathscr{A}_n^*\phi_0$$
$$+ \frac{1}{12\sqrt{2}}\mathcal{T}_{ijk}\left(\overline{Q}_{il}Q_{jl}Q_{kn}\mathscr{A}_n^* + \overline{Q}_{il}Q_{jm}Q_{kl}\mathscr{A}_m^* + Q_{il}\overline{Q}_{jm}Q_{km}\mathscr{A}_l^*\right)\phi_0$$
$$= \frac{1}{4\sqrt{3}}\mathcal{T}_{ijk}Q_{il}Q_{jm}Q_{kn}\phi_{\mathbf{e}_l+\mathbf{e}_m+\mathbf{e}_n} + \frac{1}{4\sqrt{2}}\mathcal{T}_{ijk}\overline{Q}_{il}Q_{jl}Q_{kn}\phi_{\mathbf{e}_n},$$

where we used (14) to rewrite $\mathscr{A}_l^*\mathscr{A}_m^*\mathscr{A}_n^*\phi_0$ as $\sqrt{6}\,\phi_{\mathbf{e}_l+\mathbf{e}_m+\mathbf{e}_n}$ as well as the permutation symmetry of $\mathcal{T}$ in its indices. Hence we obtain (23).

A.3. **Details on Remark 4.1.** We drop the time and spatial dependence for brevity here. Just as we have done in the above subsection, rewriting the cubic term in (24) for α using (12), we obtain

$$\zeta_0 = \alpha(q,Q)\,\phi_0 = -\frac{1}{4\sqrt{3}}\mathcal{T}_{abc}Q_{al}Q_{bm}Q_{cn}\phi_{\mathbf{e}_l+\mathbf{e}_m+\mathbf{e}_n} - \frac{1}{4\sqrt{2}}\mathcal{T}_{abc}\overline{Q}_{al}Q_{bl}Q_{cn}\phi_{\mathbf{e}_n}$$
$$- \varepsilon^{1/2}\frac{1}{4!}D^4V(\sigma_1(x,q))\cdot\xi^4\phi_0,$$

where $\mathcal{T}_{abc} := D^3_{abc}V(q)$ and $\xi := \varepsilon^{-1/2}(x-q)$. Therefore,

$$\langle\zeta_0,\phi_{\mathbf{e}_j}\rangle = -\frac{1}{4\sqrt{3}}\mathcal{T}_{abc}\overline{Q}_{al}\overline{Q}_{bm}\overline{Q}_{cn}\langle\phi_{\mathbf{e}_l+\mathbf{e}_m+\mathbf{e}_n},\phi_{\mathbf{e}_j}\rangle - \frac{1}{4\sqrt{2}}\mathcal{T}_{abc}Q_{al}\overline{Q}_{bl}\overline{Q}_{cn}\langle\phi_{\mathbf{e}_n},\phi_{\mathbf{e}_j}\rangle$$
$$- \varepsilon^{1/2}\frac{1}{4!}\langle D^4V(\sigma_1(x,q))\cdot\xi^4\phi_0,\phi_{\mathbf{e}_j}\rangle$$
$$= -\frac{1}{4\sqrt{2}}\mathcal{T}_{abc}Q_{al}\overline{Q}_{bl}\overline{Q}_{cj} + O(\varepsilon^{1/2})$$
$$= -\frac{1}{\sqrt{2}}Q_{jc}^*\partial_{q_c}V^{(1)}(q,Q) + O(\varepsilon^{1/2}), \tag{33}$$

because we can obtain the estimate

$$\frac{1}{4!}\langle D^4V(\sigma_1(x,q))\cdot\xi^4\phi_0,\phi_{\mathbf{e}_j}\rangle = O(1)$$

just as we did in the proof of Lemma 2.3 (ii). We also used the expression (7) for $V^{(1)}$ in the last equality. As a result, we have $\langle\zeta_0,\phi_{\mathbf{e}_j}\rangle = O(1)$.

References

[1] A. Bouzouina and D. Robert. Uniform semiclassical estimates for the propagation of quantum observables. *Duke Math. J.*, 111(2):223–252, 2002.

[2] M. Combescure and D. Robert. *Coherent States and Applications in Mathematical Physics*. Springer, 2012.

[3] Y. V. Egorov. The canonical transformations of pseudodifferential operators. *Uspekhi Mat. Nauk*, 24(5(149)):235–236, 1969.

[4] E. Faou and C. Lubich. A Poisson integrator for Gaussian wavepacket dynamics. *Computing and Visualization in Science*, 9(2):45–55, 2006.

[5] T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. 20(4):292–296, 1919.

[6] G. A. Hagedorn. Semiclassical quantum mechanics. I. The $\hbar \to 0$ limit for coherent states. *Communications in Mathematical Physics*, 71(1):77–93, 1980.

[7] G. A. Hagedorn. Semiclassical quantum mechanics. III. the large order asymptotics and more general states. *Annals of Physics*, 135(1):58–70, 1981.

[8] G. A. Hagedorn. Semiclassical quantum mechanics, IV: large order asymptotics and more general states in more than one dimension. *Annales de l'institut Henri Poincaré (A) Physique théorique*, 42(4):363–374, 1985.

[9] G. A. Hagedorn. Raising and lowering operators for semiclassical wave packets. *Annals of Physics*, 269(1):77–104, 1998.

[10] G. A. Hagedorn and A. Joye. Semiclassical dynamics with exponentially small error estimates. *Communications in Mathematical Physics*, 207(2):439–465, 1999.

[11] G. A. Hagedorn and A. Joye. Exponentially accurate semiclassical dynamics: Propagation, localization, ehrenfest times, scattering, and more general states. *Annales Henri Poincaré*, 1(5):837–883, 2000.

[12] E. J. Heller. Time-dependent approach to semiclassical dynamics. *Journal of Chemical Physics*, 62(4):1544–1555, 1975.

[13] E. J. Heller. Classical S-matrix limit of wave packet dynamics. *Journal of Chemical Physics*, 65(11):4979–4989, 1976.

[14] E. J. Heller. Wavepacket dynamics and quantum chaology. In M. Giannoni, A. Voros, and J. Zinn-Justin, editors, *Chaos and quantum physics*, pages 547–663. North-Holland, 1991.

[15] C. Lasser and S. Röblitz. Computing expectation values for molecular quantum dynamics. *SIAM Journal on Scientific Computing*, 32(3):1465–1483, 2010.

[16] C. Lasser and C. Lubich. Computing quantum dynamics in the semiclassical regime. *Acta Numerica*, 29:229–401, 2020.

[17] C. Lubich. *From quantum to classical molecular dynamics: reduced models and numerical analysis*. European Mathematical Society, Zürich, Switzerland, 2008.

[18] P. D. Miller. *Applied Asymptotic Analysis*. American Mathematical Society, Providence, R.I., 2006.

[19] W. H. Miller. Classical S matrix: Numerical application to inelastic collisions. *The Journal of Chemical Physics*, 53(9):3578–3587, 1970.

[20] W. H. Miller. Quantum mechanical transition state theory and a new semiclassical model for reaction rate constants. *The Journal of Chemical Physics*, 61(5):1823–1834, 1974.

[21] W. H. Miller. The semiclassical initial value representation: A potentially practical way for adding quantum effects to classical molecular dynamics simulations. *The Journal of Physical Chemistry A*, 105(13):2942–2955, 2001.

[22] T. Ohsawa. Symmetry and conservation laws in semiclassical wave packet dynamics. *Journal of Mathematical Physics*, 56(3):032103, 2015.

[23] T. Ohsawa. The Siegel upper half space is a Marsden–Weinstein quotient: Symplectic reduction and Gaussian wave packets. *Letters in Mathematical Physics*, 105(9):1301–1320, 2015.

[24] T. Ohsawa. The Hagedorn–Hermite correspondence. *Journal of Fourier Analysis and Applications*, 25(4):1513–1552, 2019.

[25] T. Ohsawa and M. Leok. Symplectic semiclassical wave packet dynamics. *Journal of Physics A: Mathematical and Theoretical*, 46(40):405201, 2013.

[26] T. Ohsawa and C. Tronci. Geometry and dynamics of Gaussian wave packets and their Wigner transforms. *Journal of Mathematical Physics*, 58(9):092105, 2017.

[27] A. K. Pattanayak and W. C. Schieve. Gaussian wave-packet dynamics: Semiquantal and semiclassical phase-space formalism. *Physical Review E*, 50(5):3601–3615, 1994.

[28] O. V. Prezhdo. Quantized Hamilton dynamics. *Theoretical Chemistry Accounts*, 116(1-3): 206–218, 2006.

[29] O. V. Prezhdo and Y. V. Pereverzev. Quantized Hamilton dynamics. *Journal of Chemical Physics*, 113(16):6557–6565, 2000.

[30] O. V. Prezhdo and Y. V. Pereverzev. Quantized Hamilton dynamics for a general potential. *Journal of Chemical Physics*, 116(11):4450–4461, 2002.

[31] D. Robert. Propagation of coherent states in quantum mechanics and applications. In X. Wang, editor, *Partial differential equations and applications*, volume 15 of *Séminaires et Congrès*, pages 181–252. Société Mathématique de France, 2007.

[32] H. Wang, X. Sun, and W. H. Miller. Semiclassical approximations for the calculation of thermal rate constants for chemical reactions in complex molecular systems. *The Journal of Chemical Physics*, 108(23):9726–9736, 1998.

[33] A. B. Watson, J. Lu, and M. I. Weinstein. Wavepackets in inhomogeneous periodic media: Effective particle-field dynamics and Berry curvature. *Journal of Mathematical Physics*, 58(2): 021503, 2017.

[34] M. Zworski. *Semiclassical Analysis*. American Mathematical Society, Providence, R.I., 2012.

Department of Mathematical Sciences, The University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080-3021

Email address: `tomoki@utdallas.edu`