Chapter II.
Reduced Models via Variational Approximation

There is a wide variety of models, or approximations, that are intermediate between the full time-dependent many-body quantum mechanics and classical mechanics. Most of them are based on a time-dependent variational principle, first used by Dirac (1930), which plays a similarly fundamental role for the time-dependent Schrödinger equation as the Rayleigh-Ritz variational principle for the Schrödinger eigenvalue problem. Indeed, several of the methods for the stationary problem, as for example the Hartree–Fock method, have a time-dependent analogue that comes about by the same choice of approximation manifold to which the variational principle is restricted. There are, however, different aspects that come into play in the time-dependent situation, both in the modeling/approximation aspects and in the numerical treatment of the reduced models.

We first give an abstract formulation and various interpretations of the time-dependent variational principle, and then turn to some basic examples that take us from the full molecular Schrödinger equation to classical molecular dynamics: the adiabatic or time-dependent Born–Oppenheimer approximation that eliminates the electronic degrees of freedom, the time-dependent self-consistent field approximation that separates the nuclei, and Gaussian wavepacket dynamics that parametrizes the single-particle wave functions. At the end of the chapter we address the theoretical question of approximation properties of variational approximations.

II.1 The Dirac–Frenkel Time-Dependent Variational Principle

In this section we give the abstract formulation of the time-dependent variational principle and discuss its structural properties.

II.1.1 Abstract Formulation

We consider an abstract Schrödinger equation on a complex Hilbert space $\mathcal{H}$ with inner product $\langle\cdot|\cdot\rangle$, with a Hamiltonian H that is a self-adjoint linear operator on $\mathcal{H}$,

$$\frac{d\psi}{dt} = \frac{1}{i\hbar} H\psi\,. \tag{1.1}$$

Let $\mathcal{M}$ be a submanifold of $\mathcal{H}$, and for $u \in \mathcal{M}$ denote by $T_u\mathcal{M}$ the tangent space at u, which consists of the derivatives of all differentiable paths on $\mathcal{M}$ passing through u. We think of $\mathcal{M}$ as an approximation manifold on which an approximate solution $u(t)$ to the solution $\psi(t)$ of (1.1) with initial data $u(0) = \psi(0) \in \mathcal{M}$ is sought. The function $t \mapsto u(t) \in \mathcal{M}$ is determined from the condition that at every time t, its derivative $du/dt\,(t)$, which lies in the tangent space $T_{u(t)}\mathcal{M}$, be such that the residual in the Schrödinger equation is orthogonal to the tangent space:

$$\frac{du}{dt} \in T_u\mathcal{M} \quad \text{such that} \quad \left\langle v \,\middle|\, \frac{du}{dt} - \frac{1}{i\hbar}Hu \right\rangle = 0 \quad \forall\, v \in T_u\mathcal{M}\,. \tag{1.2}$$

The tangent space $T_u\mathcal{M}$ is known to be a real-linear closed subspace of $\mathcal{H}$. We will always assume that in fact

$$T_u\mathcal{M} \text{ is a complex linear space,} \tag{1.3}$$

that is, with $v \in T_u\mathcal{M}$, also $iv \in T_u\mathcal{M}$. In this situation we get the same condition if we consider only the real part or the imaginary part of the inner product of (1.2). We will see, however, that these two cases lead to very different interpretations: as an orthogonal projection onto the tangent space in case of the real part, as a symplectic projection and as the Euler–Lagrange equations of an action functional in case of the imaginary part.

We remark that from a numerical analysis point of view, condition (1.2) can be seen as a Galerkin condition on the state-dependent approximation space $T_u\mathcal{M}$.

Historical Note. Dirac (1930) used condition (1.2) without further comment to derive the equations of motion of what is now known as the time-dependent Hartree–Fock method. Frenkel (1934), p. 253, gives the interpretation as an orthogonal projection and refers to the appendix of the Russian translation of Dirac's book as the origin of the argument. Some thirty years later, the Dirac–Frenkel reasoning was taken up again by McLachlan (1964) and enriched by further examples. Condition (1.2) is therefore often called the Dirac–Frenkel–McLachlan time-dependent variational principle in the chemical physics literature, see Heller (1976) and, e.g., Baer & Billing (2002). In theoretical and nuclear physics, the derivation from Dirac's quantum-mechanical action functional and with it the symplectic viewpoint has rather been emphasized; see Kerman & Koonin (1976), Rowe, Ryman & Rosensteel (1980), Kramer & Saraceno (1981) and, e.g., Feldmeier & Schnack (2000).

II.1.2 Interpretation as an Orthogonal Projection

Taking the real part in (1.2), we arrive at the minimum condition for the following linear approximation problem:

$$\frac{du}{dt} \text{ is chosen as that } w \in T_u\mathcal{M} \text{ for which } \left\| w - \frac{1}{i\hbar}Hu \right\| \text{ is minimal.} \tag{1.4}$$

(Note that $\|w + v - \frac{1}{i\hbar}Hu\|^2 = \|w - \frac{1}{i\hbar}Hu\|^2 + 2\,\mathrm{Re}\,\langle v\,,\, w - \frac{1}{i\hbar}Hu\rangle + \|v\|^2$.)

In other words, du/dt is the *orthogonal projection* of $\frac{1}{i\hbar}Hu$ onto the tangent space $T_u\mathcal{M}$. With the orthogonal projection operator onto $T_u\mathcal{M}$ denoted by $P(u)$, we can thus rewrite (1.2) as a differential equation on the manifold $\mathcal{M}$,

$$\frac{du}{dt} = P(u)\frac{1}{i\hbar}Hu\,, \tag{1.5}$$

which is *nonlinear* unless $\mathcal{M}$ is a linear subspace of $\mathcal{H}$. The (global or local in time) existence of a solution $u(t) \in D(H) \cap \mathcal{M}$ can be ascertained only with further specifications about the operator H and the manifold $\mathcal{M}$. In the following we make formal calculations which implicitly assume that a sufficiently regular solution $u(t)$ exists.

II.1.3 Interpretation as a Symplectic Projection

The real-bilinear form

$$\omega(\xi,\eta) = -2\hbar\,\mathrm{Im}\,\langle\xi\,|\,\eta\,\rangle\,, \qquad \xi,\eta\in\mathcal{H}\,,$$

is antisymmetric, and ω is called the canonical *symplectic two-form* on $\mathcal{H}$. Since $T_u\mathcal{M}$ is a complex linear space, for every $\varphi\in\mathcal{H}$ there exists a unique

$$w = P(u)\varphi \in T_u\mathcal{M} \quad \text{ such that } \quad \omega(v,w) = \omega(v,\varphi) \quad \forall\, v\in T_u\mathcal{M}\,.$$

This non-degeneracy of the two-form ω makes $\mathcal{M}$ a *symplectic submanifold* of $\mathcal{H}$, and $P(u)$ is the *symplectic projection* operator onto $T_u\mathcal{M}$. (Here $P(u)$ actually coincides with the orthogonal projection considered in the previous subsection.) Taking the imaginary part in condition (1.2) and multiplying with $-2\hbar$ yields

$$\omega\Big(v, \frac{du}{dt}\Big) = 2\,\mathrm{Re}\,\big\langle v\,\big|\,Hu\big\rangle \qquad \forall\, v\in T_u\mathcal{M}\,. \tag{1.6}$$

With the average of the Hamiltonian

$$H(u) = \langle u\,|\,H\,|\,u\rangle\,,$$

the right-hand side in (1.6) is recognized as the derivative $dH(u)v$ in the direction of v. Now, (1.6) rewritten as

$$\omega\Big(v, \frac{du}{dt}\Big) = dH(u)v \qquad \forall\, v\in T_u\mathcal{M}\,, \tag{1.7}$$

is a *Hamiltonian system* on the symplectic manifold $\mathcal{M}$ with the Hamilton function $H(u)$; see Marsden & Ratiu (1999), Chap. 5.4. Let us state and verify basic properties of this system.

Theorem 1.1. *The total energy $\langle H\rangle$ is conserved along solutions of the Hamiltonian system (1.7) on $\mathcal{M}$.*

Proof. We have (with $\dot{} = d/dt$)

$$\frac{d}{dt}\,\langle u \,|\, H \,|\, u\rangle = 2\,\mathrm{Re}\,\big\langle \dot{u}\,\big|\,Hu\big\rangle = \omega(\dot{u},\dot{u}) = 0$$

on using (1.6) with $v = \dot{u} \in T_u\mathcal{M}$ in the second equation. □

There is also the following important conservation property, which we first state briefly and then explain in detail.

Theorem 1.2. *The flow of the Hamiltonian system (1.7) is symplectic.*

This means that the symplectic two-form ω is preserved in the following sense: Let $u_0 \in \mathcal{M}$, and let $v_0 \in T_{u_0}\mathcal{M}$ be a tangent vector at u_0. Then there is a path $\gamma(\tau)$ on $\mathcal{M}$ with $\gamma(0) = u_0$ and $d\gamma/d\tau\,(0) = v_0$. Let $u(t) = u(t, u_0)$ be the solution of (1.7) with initial data u_0, and denote by

$$v(t) = \frac{d}{d\tau}\Big|_{\tau=0} u(t, \gamma(\tau)) \in T_{u(t)}\mathcal{M}$$

the tangent vector propagated along the solution $u(t, u_0)$ (note that $v(t)$ is the solution with initial data v_0 to the differential equation linearized at $u(t, u_0)$). Let $w(t)$ be another tangent vector propagated along the same solution, corresponding to an initial tangent vector w_0 at u_0. Then, the statement of Theorem 1.2 is that

$$\frac{d}{dt}\,\omega(v(t), w(t)) = 0\,. \tag{1.8}$$

Proof. By the bilinearity and antisymmetry of ω we have

$$\frac{d}{dt}\,\omega(v, w) = -\omega(w, \dot{v}) + \omega(v, \dot{w})\,.$$

Differentiating (1.6) with respect to the inital value, we obtain that this equals

$$\frac{d}{dt}\,\omega(v, w) = -2\,\mathrm{Re}\,\big\langle w\,\big|\,Hv\big\rangle + 2\,\mathrm{Re}\,\big\langle v\,\big|\,Hw\big\rangle = 0\,. \qquad \square$$

We will further discuss symplectic and Hamiltonian aspects in Section II.4.2 where we consider the non-canonical Hamiltonian structure of the equations of motion for parametrized wave functions.

II.1.4 Interpretation as an Action Principle

Taking the imaginary part in (1.2) also yields that every solution of (1.2) makes the *action functional*

$$S(u) = \int_{t_0}^{t_1} \Big\langle u(t)\,\Big|\, i\hbar\frac{du}{dt}(t) - Hu(t)\Big\rangle\, dt \tag{1.9}$$

stationary with respect to variations of paths on the manifold $\mathcal{M}$ with fixed endpoints, because by partial integration and the symmetry of H,

$$\begin{aligned}\delta S(u) &= \int_{t_0}^{t_1} \left(\left\langle \delta u(t) \,\middle|\, i\hbar \frac{du}{dt}(t) - Hu(t) \right\rangle + \left\langle u(t) \,\middle|\, i\hbar \frac{d\delta u}{dt}(t) - H\delta u(t) \right\rangle \right) dt \\ &= -2\hbar \int_{t_0}^{t_1} \operatorname{Im} \left\langle \delta u(t) \,\middle|\, \frac{du}{dt}(t) - \frac{1}{i\hbar} Hu(t) \right\rangle dt\,.\end{aligned}$$

The condition $\delta S = 0$ is the quantum-mechanical analogue of Hamilton's principle in classical mechanics. Also note that $S(u)$ is real if $\|u(t)\|^2 = \textit{Const.}$, as is seen by partial integration in (1.9).

II.1.5 Conservation Properties

We know from the Heisenberg equation (4.4) that the average $\langle A \rangle$ is conserved along solutions of the Schrödinger equation if A commutes with the Hamiltonian H. For variational approximations (1.2) there is the following criterion.

Theorem 1.3. *Let the self-adjoint operator A commute with the Hamiltonian H, $[A, H] = 0$. If*

$$Au \in T_u\mathcal{M} \qquad \forall\, u \in \mathcal{M} \cap D(A)\,, \tag{1.10}$$

then the average of A along variational approximations $u(t) \in \mathcal{M} \cap D(A)$ is conserved: $\langle u(t) \,|\, A \,|\, u(t) \rangle = Const.$

Proof. We have

$$\frac{d}{dt} \langle u \,|\, A \,|\, u \rangle = 2 \operatorname{Re} \langle Au \,|\, \dot{u} \rangle = 2 \operatorname{Re} \langle Au \,|\, \frac{1}{i\hbar}\, Hu \rangle = \langle u \,|\, \frac{1}{i\hbar}\, [A, H] \,|\, u \rangle = 0$$

on using (1.2) and (1.10) in the second equality. □

Choosing A as the identity operator, we obtain the following useful corollary.

Theorem 1.4. *The norm is conserved along variational approximations if $\mathcal{M}$ contains rays, that is, with $u \in \mathcal{M}$ also $\alpha u \in \mathcal{M}$ for every $\alpha > 0$.*

Proof. The stated condition implies $u \in T_u\mathcal{M}$ for $u \in \mathcal{M}$, and hence the result follows from Theorem 1.3. □

II.1.6 An A Posteriori Error Bound

A simple but sometimes useful general error bound for variational approximations is obtained in terms of the L^2 distance $\operatorname{dist}\left(\frac{1}{i\hbar} Hu, T_u\mathcal{M}\right)$ of $\frac{1}{i\hbar} Hu$ along the variational approximation $u(t)$ to the corresponding tangent space.

Theorem 1.5. *If $u(0) = \psi(0) \in \mathcal{M}$, then the error of the variational approximation is bounded by*

$$\|u(t) - \psi(t)\| \leq \int_0^t \text{dist}\left(\frac{1}{i\hbar} Hu(s), T_{u(s)}\mathcal{M}\right) ds\,. \tag{1.11}$$

Proof. We subtract (1.1) from (1.5), so that

$$\frac{d}{dt}(u-\psi) = \frac{1}{i\hbar} H(u-\psi) - P^\perp(u)\frac{1}{i\hbar} Hu \quad \text{with} \quad P^\perp(u) = I - P(u)\,.$$

Multiplying with $u - \psi$ and taking the real part gives

$$\begin{aligned}\|u-\psi\| \cdot \frac{d}{dt}\|u-\psi\| = \frac{1}{2}\frac{d}{dt}\|u-\psi\|^2 &= \text{Re}\left\langle u-\psi \,|\, \frac{d}{dt}(u-\psi)\right\rangle \\ &= \text{Re}\left\langle u-\psi \,|\, -P^\perp(u)\frac{1}{i\hbar}Hu\right\rangle \leq \|u-\psi\| \cdot \|P^\perp(u)\frac{1}{i\hbar}Hu\|\,.\end{aligned}$$

Dividing by $\|u - \psi\|$, integrating from 0 to t and noting

$$\text{dist}\left(\frac{1}{i\hbar}Hu,\, T_u\mathcal{M}\right) = \left\|P^\perp(u)\frac{1}{i\hbar}Hu\right\| = \left\|\frac{du}{dt} - \frac{1}{i\hbar}Hu\right\|$$

then yields the error bound (1.11). □

For the error in the average of an observable A along the variational approximation we note the bound

$$\big|\langle u \,|\, A \,|\, u\rangle - \langle \psi \,|\, A \,|\, \psi\rangle\big| = \big|\langle u-\psi \,|\, Au\rangle + \langle A\psi \,|\, u-\psi\rangle\big| \leq \|u-\psi\| \cdot \big(\|Au\| + \|A\psi\|\big).$$

II.2 Adiabatic / Born–Oppenheimer Approximation

In the following three sections we turn to basic examples of variational approximation, which take us in steps from the full molecular Schrödinger equation down to classical molecular dynamics. We begin with the adiabatic approximation that separates the motion of heavy nuclei and light electrons.

II.2.1 Electronic Schrödinger Equation

We return to the molecular Hamiltonian (I.5.6), viz.,

$$H_{\text{mol}} = T_N + T_e + V\,. \tag{2.1}$$

In a first step we ignore the contribution from the kinetic energy of the nuclei, T_N (vaguely motivated by the fact that $M_n \gg m$), and work with the electronic Hamiltonian

$$H_e(x) = T_e + V(x,\cdot)\,, \tag{2.2}$$

which acts on functions of the electronic coordinates y and depends parametrically on the nuclear coordinates x. We consider the electronic structure problem, the Schrödinger eigenvalue problem

$$H_e(x)\Phi(x,\cdot) = E(x)\Phi(x,\cdot)\,, \tag{2.3}$$

typically for the smallest eigenvalue, the ground state energy. Actually computing eigenvalues and eigenfunctions of the electronic Schrödinger equation is the primary concern of computatonal quantum chemistry; see, e.g., Szabo & Ostlund (1996), and Le Bris (2003) from a more mathematical viewpoint. Here we just suppose that this problem is solved in some satisfactory way.

We fix an eigenfunction $\Phi(x,\cdot)$ of $H_e(x)$ corresponding to the eigenvalue $E(x)$, and assume that $\Phi(x,y)$ is of unit L^2 norm as a function of y and depends smoothly on x. For fixed nuclear coordinates x, the solution of the *time-dependent electronic Schrödinger equation*

$$i\hbar\,\frac{\partial \Psi_e}{\partial t} = H_e(x)\Psi_e \tag{2.4}$$

with initial data $\psi_0(x)\Phi(x,\cdot)$ is given by

$$\Psi_e(x,y,t) = e^{-iE(x)t/\hbar}\psi_0(x)\cdot\Phi(x,y)\,.$$

II.2.2 Schrödinger Equation for the Nuclei on an Electronic Energy Surface

This motivates the *adiabatic approximation* to the molecular Schrödinger equation, which is the variational approximation on

$$\mathcal{M} = \{u \in L^2_{x,y} : u(x,y) = \psi(x)\,\Phi(x,y),\ \psi \in L^2_x\}\,. \tag{2.5}$$

Here $L^2_x = L^2(\mathbb{R}^{3N})$ denotes the Lebesgue space of square integrable functions depending only on the nuclear coordinates x, and $L^2_{x,y} = L^2(\mathbb{R}^{3N}\times\mathbb{R}^{3L})$ is the L^2 space of functions depending on both nuclear and electronic coordinates. Note that here $\mathcal{M}$ is a linear space so that $T_u\mathcal{M} = \mathcal{M}$ for all $u \in \mathcal{M}$. As we show below, the Dirac-Frenkel variational principle (1.2) then leads to a *Schrödinger equation for the nuclei* on the electronic energy surface E:

$$i\hbar\,\frac{\partial\psi}{\partial t} = H_N\psi \quad \text{with} \quad H_N = T_N + E + B_1 + B_2\,, \tag{2.6}$$

$$B_1 = \sum_{n=1}^{N}\frac{\hbar}{M_n}\,\mathrm{Im}\,\langle\nabla_{x_n}\Phi\,|\,\Phi\rangle_{L^2_y}\cdot p_n\,, \quad B_2 = \sum_{n=1}^{N}\frac{\hbar^2}{2M_n}\,\|\nabla_{x_n}\Phi\|^2_{L^2_y}\,,$$

with $p_n = -i\hbar\nabla_{x_n}$. The Hamiltonian H_N acts on functions of only the nuclear coordinates x, with the electronic eigenvalue E as a potential. The last two terms

B_1 and B_2 contain derivatives of the electronic wave function Φ with respect to the nuclear coordinates x. They are usually neglected in computations, first because they are expensive to compute or simply not available and second by the formal argument – to be taken with caution – that they carry the large masses M_n in the denominator and are of lower differentiation order than the kinetic energy term. The resulting simplified approximation with the Hamiltonian

$$H_{\mathrm{BO}} = T_N + E$$

is known as the *time-dependent Born–Oppenheimer approximation*. It describes the motion of the nuclei as driven by the potential energy surface E of the electrons. It underlies the vast majority of computations in molecular dynamics.

The term B_2 can indeed be safely neglected since it can be shown that this omission introduces an error that is of the same magnitude as the approximation error in the adiabatic approximation.

The term B_1, known as the *Berry connection*, vanishes for real eigenfunctions Φ and, more generally, it can be made to vanish by a gauge transformation $\Phi(x,y) \to e^{i\theta(x)}\Phi(x,y)$ with θ satisfying $\nabla_{x_n}\theta(x) = -\mathrm{Im}\,\langle \nabla_{x_n}\Phi \,|\, \Phi\rangle_{L^2_y}$. This transformation of Φ changes $\psi(x,t) \to e^{-i\theta(x)}\psi(x,t)$. Note that θ is uniquely determined up to a constant if Φ is indeed a smooth function of x on all of $\mathbb{R}^{3N}$, but is only locally uniquely determined if Φ is a differentiable function of x only on a domain that is not simply connected. In the latter case, B_1 can cause physical effects that are not retained in the model otherwise; see the extensive literature on Berry's phase, starting with Berry (1984) and Simon (1983).

Derivation of (2.6): We note that for $u(x,y) = \psi(x)\Phi(x,y)$ we have

$$T_N u = -\sum_{n=1}^{N} \frac{\hbar^2}{2M_n}\Big(\Delta_{x_n}\psi\cdot\Phi + 2\,\nabla_{x_n}\psi\cdot\nabla_{x_n}\Phi + \psi\cdot\Delta_{x_n}\Phi\Big),$$

and recall that $\|\Phi(x,\cdot)\|^2_{L^2_y} = 1$ for all x. We then obtain from (1.2) with $v(x,y) = \varphi(x)\Phi(x,y)$ for arbitrary $\varphi \in L^2_x$ that

$$\Big\langle \varphi \,\Big|\, i\hbar\frac{\partial\psi}{\partial t} - E\psi \quad + \quad \sum_{n=1}^{N}\frac{\hbar^2}{2M_n}\Big(\Delta_{x_n}\psi + 2\,\langle\nabla_{x_n}\Phi\,|\,\Phi\rangle_{L^2_y}\cdot\nabla_{x_n}\psi \\ - \langle\nabla_{x_n}\Phi\,|\,\nabla_{x_n}\Phi\rangle_{L^2_y}\,\psi\Big)\Big\rangle_{L^2_x} = 0\,.$$

On noting that $0 = \nabla_{x_n}\|\Phi\|^2_{L^2_y} = 2\,\mathrm{Re}\,\langle\nabla_{x_n}\Phi\,|\,\Phi\rangle_{L^2_y}$, we obtain (2.6). □

II.2.3 Semiclassical Scaling

One property to the success of the adiabatic approximation is the smallness of the mass ratio of electrons and nuclei,

$$\varepsilon^2 = \frac{m}{M} \ll 1 \tag{2.7}$$

with $M = \min_n M_n$. For ease of presentation we assume in the following that the masses of the nuclei are all equal: $M_n = M$ for all n. In atomic units ($\hbar = 1$, $m = 1$, $r = 1$, $e = 1$) and with the small parameter ε of (2.7), the molecular Hamiltonian then takes the form

$$H^\varepsilon_{\mathrm{mol}} = -\frac{\varepsilon^2}{2}\Delta_x + H_e(x) \quad \text{with} \quad H_e(x) = -\frac{1}{2}\Delta_y - V(x,\cdot)\,. \tag{2.8}$$

We are interested in solutions to the Schrödinger equation of bounded energy, and in particular of bounded kinetic energy

$$\langle \Psi \,|\, -\frac{\varepsilon^2}{2}\Delta_x \,|\, \Psi\rangle = \frac{1}{2}\,\|\varepsilon\nabla_x \Psi\|^2 = \mathcal{O}(1)\,.$$

For a wavepacket $e^{ip\cdot x}a(x)$ this condition corresponds to a momentum $p \sim \varepsilon^{-1}$ and hence to a velocity $v = p/M \sim \varepsilon$. Motion of the nuclei over a distance ~ 1 can thus be expected on a time scale ε^{-1}. We therefore rescale time

$$t \to t/\varepsilon\,,$$

so that with respect to the new time nuclear motion over distances ~ 1 can be expected to occur at time ~ 1. The molecular Schrödinger equation in the rescaled time then takes the form

$$i\varepsilon\frac{\partial \Psi}{\partial t} = H^\varepsilon_{\mathrm{mol}}\Psi\,. \tag{2.9}$$

The Schrödinger equation (2.6) for the nuclei becomes

$$i\varepsilon\frac{\partial \psi}{\partial t} = H^\varepsilon_N \psi \quad \text{with} \quad H^\varepsilon_N = -\frac{\varepsilon^2}{2}\Delta_x + E + \varepsilon B_1 + \varepsilon^2 B_2\,, \tag{2.10}$$
$$B_1 = \mathrm{Im}\,\langle \nabla_x \Phi \,|\, \Phi\rangle_{L^2_y}\cdot p\,, \qquad B_2 = \frac{1}{2}\,\|\nabla_x \Phi\|^2_{L^2_y}\,,$$

with $p = -i\varepsilon\nabla_x$. We are interested in solutions over times $t = \mathcal{O}(1)$.

II.2.4 Spectral Gap Condition

A small error of the adiabatic approximation will be seen to be caused by two properties: in addition to the smallness of the mass ratio $\varepsilon^2 = m/M$, we require a separation of the eigenvalue $E(x)$ from the remainder of the spectrum $\sigma(H_e(x))$ of the electronic Hamiltonian $H_e(x)$,

$$\mathrm{dist}\big(E(x),\, \sigma(H_e(x)) \setminus \{E(x)\}\big) \geq \delta > 0\,, \tag{2.11}$$

uniformly for all x in a region where the wavefunction remains approximately localized. We will give a result on the approximation error in the situation of a globally

well-separated single eigenvalue $E(x)$, where (2.11) is assumed to hold uniformly for all $x \in \mathbb{R}^{3N}$.

Remark. It is known that the adiabatic approximation generally breaks down near crossings of eigenvalues. A remedy then is to enlarge the approximation space by including several energy bands that are well separated from the remaining ones in the region of physical interest, e.g., using

$$\mathcal{M} = \{u \in L^2_{x,y} : u(x,y) = \psi_1(x)\Phi_1(x,y) + \psi_2(x)\Phi_2(x,y),\ \psi_1, \psi_2 \in L^2_x\}\ , \tag{2.12}$$

where $\Phi_1(x,\cdot)$, $\Phi_2(x,\cdot)$ span an invariant subspace of the electronic Hamiltonian $H_e(x)$. The variational approximation on $\mathcal{M}$ then leads to a system of coupled Schrödinger equations:

$$i\hbar\,\frac{\partial \psi}{\partial t} = T_N\psi + B_1\psi + B_2\psi + V\psi \qquad \text{for} \quad \psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \tag{2.13}$$

with the matrix-valued potential

$$V = \begin{pmatrix}V_{11} & V_{12}\\ V_{21} & V_{22}\end{pmatrix} \qquad \text{with} \quad V_{ij}(x) = \langle \Phi_i(x,\cdot)\,|\,H_e(x)\,|\,\Phi_j(x,\cdot)\rangle_{L^2_y} \tag{2.14}$$

and with the diagonal operators $B_j = \begin{pmatrix}B_j^1 & 0\\ 0 & B_j^2\end{pmatrix}$, where B_j^k are defined as B_j in (2.6) with Φ_k instead of Φ.

The non-adiabatic solution behaviour near eigenvalue crossings has attracted much attention in recent years; see, e.g., Baer & Billing (2002), Domcke, Yarkony & Köppel (2004), and Lasser & Teufel (2005).

II.2.5 Approximation Error

We derive an error bound of the adiabatic approximation that works for a modified Hamiltonian where the Coulomb interactions of the nuclei are mollified to smooth bounded potentials. We assume

$$\|\nabla_x V(x,y)\| \le C_V \qquad \text{for} \quad x \in \mathbb{R}^{3N},\ y \in \mathbb{R}^{3L} \tag{2.15}$$

and consider initial data on the approximation space $\mathcal{M}$ of (2.5),

$$\Psi_0(x,y) = \psi_0(x)\Phi(x,y) \qquad \text{with} \quad \|H_N^\varepsilon \psi_0\| \le C_0\ ,\ \|\psi_0\| = 1\ . \tag{2.16}$$

We consider the adiabatic approximation $u(t) = u(\cdot,\cdot,t)$, with initial data Ψ_0, determined by the time-dependent variational principle:

$$\frac{\partial u}{\partial t} \in \mathcal{M} \quad \text{such that} \quad \left\langle v \,\middle|\, \frac{\partial u}{\partial t} - \frac{1}{i\varepsilon} H^\varepsilon_{\mathrm{mol}} u \right\rangle = 0 \quad \forall\, v \in \mathcal{M}\ . \tag{2.17}$$

We know already that

$$u(x,y,t) = \psi(x,t)\Phi(x,y)\,,$$

where $\psi(x,t)$ is the solution of the nuclear Schrödinger equation (2.10) with initial data $\psi_0(x)$. This is compared with the exact solution $\Psi(t) = \Psi(\cdot,\cdot,t)$ of the molecular Schrödinger equation (2.9) with initial data $\Psi_0(x,y) = \psi_0(x)\Phi(x,y)$.

Theorem 2.1 (Space-Adiabatic Theorem, Teufel 2003). *Under the above conditions, the error of the adiabatic approximation is bounded by*

$$\|u(t) - \Psi(t)\| \le C\,(1+t)\,\varepsilon \qquad \textit{for} \quad t \ge 0\,,$$

where C is independent of ε and t but depends on the gap δ of (2.11) (uniform for $x \in \mathbb{R}^{3N}$), on bounds of partial derivatives with respect to x up to third order of the eigenfunctions Φ, and on the bounds C_V of (2.15) and C_0 of (2.16).

Teufel (2003) gives a more general result, including the case of higher-dimensional invariant subspaces as in (2.12), and a wealth of related theory. The result is also related to the time-adiabatic theorem of Born & Fock (1928) and Kato (1950), which states that in a quantum system with a slowly time-varying Hamiltonian a wave function that is an eigenfunction initially, approximately remains an eigenfunction of the Hamiltonian at any instant for long times.

Proof. We let $H = H^\varepsilon_{\mathrm{mol}}$ for brevity. With the orthogonal projection P onto $\mathcal{M}$, we reformulate (2.17) as

$$i\varepsilon\frac{\partial u}{\partial t} = Ku \qquad \text{with} \quad K = PHP\,.$$

We then have

$$u(t) = e^{-itK/\varepsilon}\Psi_0 = Pe^{-itK/\varepsilon}\Psi_0 \in \mathcal{M}\,, \qquad \Psi(t) = e^{-itH/\varepsilon}\Psi_0\,,$$

and by the variation-of-constants formula,

$$\begin{aligned} u(t) - \Psi(t) &= e^{-itK/\varepsilon}\Psi_0 - e^{-itH/\varepsilon}\Psi_0 \\ &= -\frac{1}{i\varepsilon}\int_0^t e^{-i(t-s)H/\varepsilon}(H-K)Pe^{-isK/\varepsilon}\Psi_0\,ds\,. \end{aligned}$$

We note that $(H-K)P = P^\perp HP$ (with $P^\perp = I - P$ the complementary orthogonal projection). The key idea is now to write $P^\perp HP$ essentially as a commutator with H, which becomes possible by the gap condition (2.11). Lemma 2.2 below tells us that $P^\perp HP = \varepsilon[H,G] + \varepsilon^2 R$ with operators G and R that are bounded independently of ε in appropriate norms as stated there. The remainder term $\varepsilon^2 R$ immediately gives an $\mathcal{O}(\varepsilon)$ bound on time intervals of length $\mathcal{O}(1)$ as desired. We then have

$$u(t) - \Psi(t) = ie^{-itH/\varepsilon}\int_0^t e^{isH/\varepsilon}\,[H,G]\,e^{-isH/\varepsilon}\cdot e^{isH/\varepsilon}\,e^{-isK/\varepsilon}\,\Psi_0\,ds + \mathcal{O}(\varepsilon)\,,$$

where we note that

$$e^{isH/\varepsilon}\,[H,G]e^{-isH/\varepsilon} = -i\varepsilon\,\frac{d}{ds}\Big(e^{isH/\varepsilon}\,G\,e^{-isH/\varepsilon}\Big)\,.$$

We now use partial integration and observe

$$\frac{d}{ds}\Big(e^{isH/\varepsilon}\,e^{-isK/\varepsilon}\Psi_0\Big) = \frac{i}{\varepsilon}e^{isH/\varepsilon}\,(H-K)P\,e^{-isK/\varepsilon}\Psi_0\,.$$

Expressing once again $(H-K)P = P^\perp HP$, we obtain

$$\begin{aligned} u(t)-\Psi(t) &= \varepsilon\,Ge^{-itK/\varepsilon}\,\Psi_0 - \varepsilon\,e^{-itH/\varepsilon}G\Psi_0 \\ &\quad - \int_0^t e^{-i(t-s)H/\varepsilon}\,GP^\perp HP\,e^{-isK/\varepsilon}\Psi_0\,ds + \mathcal{O}(\varepsilon)\,. \end{aligned}$$

The result now follows with the estimates of Lemmas 2.2 and 2.3. □

It remains to state and prove the two lemmas to which we referred in the above proof. They use scaled Sobolev norms of functions on $\mathbb{R}^{3N}$ or $\mathbb{R}^{3N}\times\mathbb{R}^{3L}$. The squares of these norms are defined by

$$\begin{aligned} \|\varphi\|_{1,\varepsilon}^2 &= \|\varepsilon\nabla_x\varphi\|^2 + \|\varphi\|^2\,, \\ \|\varphi\|_{2,\varepsilon}^2 &= \|\varepsilon^2\Delta_x\varphi\|^2 + \|\varphi\|^2\,, \end{aligned}$$

where the norm on the right-hand side is the L^2 norm (the L^2_x or $L^2_{x,y}$ norm, as appropriate).

Lemma 2.2. *The projected Hamiltonian $P^\perp HP$ can be written as*

$$P^\perp HP = \varepsilon[H,G] + \varepsilon^2 R \tag{2.18}$$

where the operators G and R are bounded by

$$\|G\Psi\| \le C_1\,\|\Psi\|_{1,\varepsilon}\,, \qquad \|R\Psi\| \le C_2\,\|\Psi\|_{2,\varepsilon} \tag{2.19}$$

for all $\Psi\in C_0^\infty(\mathbb{R}^{3N}\times\mathbb{R}^{3L})$. Moreover, $P^\perp HP$ is bounded by

$$\|P^\perp HP\Psi\|_{1,\varepsilon} \le C\varepsilon\,\|\Psi\|_{2,\varepsilon}\,. \tag{2.20}$$

Proof. In the following we write $\nabla=\nabla_x$ and $\Delta=\Delta_x$ for the gradient and Laplacian with respect to the nuclear coordinates x.

(a) We begin by computing $P^\perp HP$ for $H = -\frac{\varepsilon^2}{2}\Delta + H_e$. The orthogonal projection P onto $\mathcal{M}$ is fibered as

$$(P\Psi)(x) = P(x)\Psi(x,\cdot)\,,$$

where $P(x)$ is the L^2_y-orthogonal projection onto the span of the eigenfunction $\Phi(x,\cdot)$ of the electronic Hamiltonian $H_e(x)$. We have, for $\eta\in L^2_y$,

$$P(x)\eta = \langle \Phi(x,\cdot) \,|\, \eta \rangle \, \Phi(x,\cdot)\,,$$

with the inner product of L^2_y. Since $\Phi(x,\cdot)$ spans an invariant subspace of $H_e(x)$, we have $P^\perp(x) H_e(x) P(x) = 0$, and hence, for $\Psi \in L^2_{x,y}$,

$$P^\perp H P \Psi = -\frac{\varepsilon^2}{2}\, P^\perp \Delta (P\Psi) = -\varepsilon^2 P^\perp (\nabla P)\cdot \nabla \Psi - \frac{\varepsilon^2}{2}\, P^\perp (\Delta P)\Psi\,.$$

For the first term on the right-hand side we note, using $(\nabla P) P^\perp \Psi = \langle \nabla \Phi \,|\, P^\perp \Psi \rangle \Phi$ and $P^\perp \Phi = 0$,

$$Q := -P^\perp (\nabla P) = -P^\perp (\nabla P) P\,.$$

We thus obtain

$$P^\perp H P = \varepsilon\, Q \cdot \varepsilon \nabla + \varepsilon^2 R_0\,, \tag{2.21}$$

where $R_0(x) = -\frac{1}{2} P(x)^\perp (\Delta P)(x)$ is bounded on L^2_y uniformly in $x \in \mathbb{R}^{3N}$, provided that the eigenfunction Φ has bounded derivatives with respect to x. We also note that (2.21) implies the bound (2.20).

(b) We construct $F(x)$ such that

$$[H_e(x), F(x)] = Q(x)\,. \tag{2.22}$$

Writing H_e as an operator matrix with blocks corresponding to $\mathcal{M}$ and $\mathcal{M}^\perp$,

$$H_e = \begin{pmatrix} E & 0 \\ 0 & H_e^\perp \end{pmatrix} \qquad \text{with} \quad H_e^\perp = P^\perp H_e P^\perp\,,$$

we can rewrite (2.22) as

$$\left[\begin{pmatrix} E & 0 \\ 0 & H_e^\perp \end{pmatrix}, \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \right] = \begin{pmatrix} 0 & 0 \\ Q & 0 \end{pmatrix}$$

which is solved by setting $F_{11} = 0$, $F_{12} = 0$, $F_{22} = 0$ and determining $F_{21} = P^\perp F P$ from

$$H_e^\perp F_{21} - F_{21} E = Q\,.$$

By the spectral gap condition (2.11), this equation has a unique solution, and we thus obtain the solution to (2.22) as

$$F(x) = \big(H_e^\perp(x) - E(x)\big)^{-1} Q(x)\,.$$

This is bounded in L^2_y uniformly for $x \in \mathbb{R}^{3N}$ by the uniform gap condition, and so are $\nabla F(x)$ and $\Delta F(x)$.

(c) We next show that the commutator of $H = -\frac{\varepsilon^2}{2}\Delta + H_e$ with F is a small perturbation to $[H_e, F] = Q$. For this we note that

$$\big[-\frac{\varepsilon^2}{2}\Delta, F\big] = -\varepsilon \nabla F \cdot \varepsilon \nabla - \frac{\varepsilon^2}{2}\, \Delta F(x)\,,$$

so that

$$[H, F] = Q - \varepsilon R_1 \,, \tag{2.23}$$

where R_1 is bounded by $\|R_1\Psi\| \le c_1\|\Psi\|_{1,\varepsilon}$ for all Ψ.

(d) We set

$$G = F \cdot \varepsilon\nabla \tag{2.24}$$

and show that the commutator with H equals $Q \cdot \varepsilon\nabla$ up to a small perturbation. By (2.23) we have, using the Leibniz rule of the commutator,

$$\begin{aligned} [H, G] &= [H, F] \cdot \varepsilon\nabla + F \cdot [H, \varepsilon\nabla] \\ &= Q \cdot \varepsilon\nabla - \varepsilon R_1 \cdot \varepsilon\nabla - \varepsilon F \cdot \nabla V \,. \end{aligned}$$

For the term with the potential V we recall assumption (2.15), which bounds ∇V. The term $Q \cdot \varepsilon\nabla$ is the same as in (2.21), and hence we obtain the desired result (2.18) with $R = R_0 + R_1 \cdot \varepsilon\nabla + F \cdot \nabla V$. The bounds (2.19) are immediate from the construction of the operators G and R. □

We also need the following regularity result.

Lemma 2.3. *In the situation of Theorem 2.1, we have*

$$\|u(t)\|_{2,\varepsilon} \le C\left(\|H_N^\varepsilon \psi_0\| + 1\right) \quad \textit{for} \quad t \ge 0 \,.$$

Proof. We use the bounds, for $\psi\Phi \in \mathcal{M}$,

$$\|\psi\Phi\|_{2,\varepsilon} \le c\,\|\psi\|_{2,\varepsilon} \le C\left(\|H_N^\varepsilon \psi\| + \|\psi\|\right),$$

for which we omit the straightforward derivation. We have $u(t) = \left(e^{-itH_N^\varepsilon/\varepsilon}\psi_0\right)\Phi$, and the above inequality thus yields

$$\|u(t)\| \le C\big(\|H_N^\varepsilon e^{-itH_N^\varepsilon/\varepsilon}\psi_0\| + \|\psi_0\|\big) = C\big(\|H_N^\varepsilon \psi_0\| + 1\big)\,,$$

which is the stated bound. □

II.3 Separating the Particles: Self-Consistent Field Methods

The remaining high dimensionality requires further model reductions. The many-body wave function is approximated by appropriate linear combinations of tensor products of single-particle wave functions. The simplest case arises in approximating the dynamics of the nuclei by a single tensor product, which yields the *time-dependent Hartree method*. This model describes the motion of each particle driven by the mean field of the other particles.

Its antisymmetrized version, suitable for electron dynamics, is known as the *time-dependent Hartree–Fock method*. The equations of motion for the orbitals were

derived by Dirac (1930) in what is the historically first application of the time-dependent variational principle. This method is the time-dependent counterpart of the stationary Hartree–Fock method, which uses antisymmetrized products of orbitals to approximate eigenfunctions of the Schrödinger operator and is the basic approach to electronic structure computations; see, e.g., Szabo & Ostlund (1996).

Taking linear combinations of tensor products or their antisymmetrizations yields the *multi-configuration* time-dependent Hartree and Hartree–Fock methods, put forward by Meyer, Manthe & Cederbaum (1990). In this section we describe these various methods, derive the nonlinear equations of motion and discuss some of their properties.

The model reductions of this section can be viewed as *low-rank approximations* to the high-dimensional multi-particle wave function. Mostly independently of the developments in quantum mechanics, low-rank approximations to huge matrices and tensors have been widely used as computationally viable approximations in many other fields including, for example, information retrieval and option pricing. It seems, however, that using the time-dependent variational principle for low-rank approximations in areas outside quantum mechanics has been considered only recently (Koch & Lubich 2007b, Nonnenmacher & Lubich 2007, Jahnke & Huisinga 2007).

II.3.1 Time-Dependent Hartree Method (TDH)

We consider the Schrödinger equation for the nuclei obtained from the Born–Oppenheiner approximation,

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi\,, \qquad H = T + V \tag{3.1}$$

with kinetic energy $T = -\sum_{n=1}^{N} \frac{\hbar^2}{2M_N}\,\Delta_{x_n}$ and a potential $V(x_1,\dots,x_N)$. We assume that the domain $D(V)$ contains $D(T) = H^2(\mathbb{R}^{3N})$.

Hartree Products. We look for an approximation to the wave function of the tensor product form

$$\psi(x_1,\dots,x_N,t) \approx a(t)\,\varphi_1(x_1,t)\dots\varphi_N(x_N,t)$$

with a scalar phase factor $a(t)$ and with *single-particle functions* (or *molecular orbitals*) $\varphi_n(x_n,t)$. We thus consider the variational approximation (1.2) on the infinite-dimensional manifold

$$\mathcal{M} = \{u \in L^2(\mathbb{R}^{3N}) \,:\, u \neq 0,\ u = a\,\varphi_1\otimes\cdots\otimes\varphi_N\,,\ a\in\mathbb{C},\ \varphi_n \in L^2(\mathbb{R}^3)\} \tag{3.2}$$

(or instead we might consider tensor products of $3N$ functions in $L^2(\mathbb{R})$). The representation of $u \in \mathcal{M}$ as $u = a\,\varphi_1\otimes\cdots\otimes\varphi_N$ is not unique: for any choice of complex numbers $c_n \neq 0$, u remains unaltered under the transformation

$$\varphi_n \to c_n \varphi_n \,, \qquad a \to \frac{a}{c_1 \dots c_N} \,. \tag{3.3}$$

Tangent Functions. Although we do not have a unique representation of functions in the Hartree manifold $\mathcal{M}$, we can obtain a unique representation of tangent functions. This is what matters in derivng the equations of motion for the single-particle functions. Consider $u = a\,\varphi_1 \otimes \cdots \otimes \varphi_N$ with a of unit modulus and all φ_n of unit L^2 norm. Every tangent function $\dot{u} \in T_u\mathcal{M}$ (for the moment, $\dot{u}$ is just a symbol for any tangent function) is of the form

$$\dot{u} = \dot{a}\,\varphi_1 \otimes \cdots \otimes \varphi_N + a\,\dot{\varphi}_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_N + \cdots + a\,\varphi_1 \otimes \cdots \otimes \varphi_{N-1} \otimes \dot{\varphi}_N \tag{3.4}$$

where $\dot{a} \in \mathbb{C}$ and $\dot{\varphi}_n \in L^2$. These turn out to be uniquely determined by $\dot{u}$ and the fixed $a, \varphi_1, \dots, \varphi_N$ if we impose the *gauge condition*

$$\langle \varphi_n \,|\, \dot{\varphi}_n \rangle = 0 \,. \tag{3.5}$$

Indeed, taking the inner product of both sides of (3.4) with $u = a\,\varphi_1 \otimes \cdots \otimes \varphi_N$ and using (3.5) and $\|\varphi_n\| = 1$ and $a = 1/\overline{a}$, determines $\dot{a}$ as

$$\dot{a} = \langle u \,|\, \dot{u} \rangle\, a \,. \tag{3.6}$$

Taking the inner product with the function in which the nth factor φ_n in u is replaced by some L^2 function ϑ_n, viz., with $a\,\varphi_1 \otimes \cdots \otimes \vartheta_n \otimes \cdots \otimes \varphi_N \in T_u\mathcal{M}$, determines $\dot{\varphi}_n$ uniquely by the equation

$$\langle \vartheta_n \,|\, \dot{\varphi}_n \rangle + \overline{a}\dot{a} \langle \vartheta_n \,|\, \varphi_n \rangle = \langle a\,\varphi_1 \otimes \cdots \otimes \vartheta_n \otimes \cdots \otimes \varphi_N \,|\, \dot{u} \rangle \qquad \forall\, \vartheta_n \in L^2 \,. \tag{3.7}$$

Equations of Motion for the Single-Particle Functions. We now consider the variational approximation (1.2) on the Hartree manifold $\mathcal{M}$, viz.,

$$\left\langle v \,\middle|\, \frac{du}{dt} - \frac{1}{i\hbar} H u \right\rangle = 0 \qquad \forall\, v \in T_u\mathcal{M} \,. \tag{3.8}$$

Applying the above argument with $\dot{u} = du/dt \in T_u\mathcal{M}$ and using (3.8) to replace $\dot{u}$ by $\frac{1}{i\hbar} H u$ in (3.6) and (3.7), we obtain evolution equations for the factors in $u = a\,\varphi_1 \otimes \cdots \otimes \varphi_N$:

$$\begin{aligned} \frac{da}{dt} &= \left\langle u \,\middle|\, \frac{1}{i\hbar} H u \right\rangle a \\ \left\langle \vartheta_n \,\middle|\, \frac{\partial \varphi_n}{\partial t} \right\rangle &= \left\langle a\,\varphi_1 \otimes \cdots \otimes \vartheta_n \otimes \cdots \otimes \varphi_N \,\middle|\, \frac{1}{i\hbar} H u \right\rangle \\ &\quad - \left\langle u \,\middle|\, \frac{1}{i\hbar} H u \right\rangle \langle \vartheta_n \,|\, \varphi_n \rangle \qquad\qquad \forall\, \vartheta_n \in L^2 \,. \end{aligned} \tag{3.9}$$

With the total energy $E = \langle u \,|\, H \,|\, u \rangle$, which by Theorem 1.1 is constant in time, and with the *mean-field Hamiltonian* for the nth particle,

$$\langle H\rangle_n = \langle \psi_n \,|\, H \,|\, \psi_n\rangle \qquad \text{with} \quad \psi_n = \bigotimes_{j\neq n} \varphi_j \tag{3.10}$$

(the inner product on the right-hand side is over all variables except x_n), the equations of motion become the trivial linear constant-coefficient differential equation $i\hbar\, da/dt = Ea$ and

$$i\hbar \frac{\partial \varphi_n}{\partial t} = \langle H\rangle_n \varphi_n - E\varphi_n \,. \tag{3.11}$$

Multiplying with $\frac{1}{i\hbar}\varphi_n$ and noting

$$\frac{d}{dt}\,\|\varphi_n\|^2 = 2\,\mathrm{Re}\,\big\langle \varphi_n \,\big|\, \frac{\partial \varphi_n}{\partial t}\big\rangle = 0\,,$$

we see that φ_n indeed remains of unit norm, as was assumed in the derivation.

The last term $E\varphi_n$ in (3.11) can be dropped if we rescale $\varphi_j \to e^{-iEt/\hbar}\varphi_j$. For a Hamiltonian $H = T + V$ as in (3.1), we obtain for all $\vartheta_n \in L^2(\mathbb{R}^3)$ that are orthogonal to φ_n,

$$\Big\langle \varphi_1 \otimes \cdots \otimes \vartheta_n \otimes \cdots \otimes \varphi_N \,\Big|\, Tu\Big\rangle = \Big\langle \vartheta_n \,\Big|\, -\frac{h^2}{2M_n}\,\Delta_{x_n}\varphi_n\Big\rangle\,,$$

and hence for such ϑ_n we have by (3.9)

$$\Big\langle \vartheta_n \,\Big|\, i\hbar\,\frac{\partial \varphi_n}{\partial t} + \frac{\hbar^2}{2M_n}\,\Delta_{x_n}\varphi_n - \langle V\rangle_n \varphi_n\Big\rangle = 0\,,$$

where the mean-field potential $\langle V\rangle_n$ is defined in the same way as in (3.10) with V instead of H. It follows that the right-hand expression in the inner product is a multiple of φ_n. Since this term adds to $\dot u = du/dt$ in (3.4) only a scalar multiple of u and hence yields only a modified phase factor a in u, this term is ignored. Let us summarize the result obtained.

Theorem 3.1 (Time-Dependent Hartree Method). *For a Hamiltonian (3.1), the variational approximation (1.2) on the Hartree manifold (3.2), for initial data $u(x_1,\dots,x_N,0) = \varphi_1(x_1,0)\dots\varphi_N(x_N,0)$ with $\varphi_n(\cdot,0)$ of unit L^2 norm, is given as*

$$u(x_1,\dots,x_N,t) = a(t)\,\varphi_1(x_1,t)\dots\varphi_N(x_N,t)\,,$$

where $|a(t)| = 1$ and $\varphi_n(x_n,t)$ are solutions to the system of nonlinear partial differential equations

$$i\hbar\,\frac{\partial \varphi_n}{\partial t} = -\frac{\hbar^2}{2M_n}\,\Delta_{x_n}\varphi_n + \langle V\rangle_n\varphi_n\,. \tag{3.12}$$

This holds on time intervals $0 \le t \le \overline{t}$ on which a strong solution to this system exists, that is, for $\varphi_n \in C^1([0,\overline{t}], L^2(\mathbb{R}^3)) \cap C([0,\overline{t}], H^2(\mathbb{R}^3))$. □

Equations (3.12) look like usual Schrödinger equations, but since the mean-field potential $\langle V\rangle_n$ depends on the single-particle functions of the other particles, we obtain a coupled system of low-dimensional *nonlinear* partial differential equations.

A strong solution to (3.12) exists globally for all times $t \geq 0$ for example in the case of a smooth bounded potential with bounded derivatives. This is shown by Picard iteration in the Sobolev space $H^2(\mathbb{R}^3)^N$ on the integrated equation

$$\varphi_n(t) = e^{-itT_n/\hbar}\varphi_n(0) + \int_0^t e^{-i(t-s)T_n/\hbar}\,\langle V\rangle_n(s)\,\varphi_n(s)\,ds\ ,$$

where $T_n = -\frac{\hbar^2}{2M_n}\,\Delta_{x_n}$. By the same argument, the solution then has H^k regularity for arbitrary k whenever the initial data is in H^k.

Remark 3.2 (Principal Bundle Structure). On the Hartree manifold $\mathcal{M}$ of (3.2), $y = (a, \varphi_1, \ldots, \varphi_N)$ are not coordinates, but the underlying mathematical structure here and in the following subsections is that of a *principal bundle*, which is a familar concept in differential geometry that we now describe. There is a map $\chi : \mathcal{N} \to \mathcal{M}$ from a manifold $\mathcal{N}$ onto $\mathcal{M}$, so that every $u \in \mathcal{M}$ can be represented, though not uniquely, as

$$u = \chi(y) \qquad \text{for some } y \in \mathcal{N}.$$

(We have $\chi(y) = a\,\varphi_1 \otimes \cdots \otimes \varphi_n$ on the Hartree manifold.) The map χ is invariant under the action of a Lie group G on $\mathcal{N}$, which we denote by $\cdot : G \times \mathcal{N} \to \mathcal{N}$:

$$\chi(g \cdot y) = \chi(y) \qquad \forall\, g \in G\,,\ y \in \mathcal{N}\,.$$

In the Hartree method, the group is the componentwise multiplicative group $G = (\mathbb{C}^*)^N$ (with $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$), and the action is given by (3.3).

Moreover, there is a *gauge map* γ, which at every $y \in \mathcal{N}$ associates to a tangent vector $\dot y \in T_y\mathcal{N}$ an element $\gamma(y)\dot y$ in the Lie algebra $\mathfrak{g}$ of G ($\mathfrak{g}$ is the tangent space at the unit element of G). The linear map $\gamma(y) :\ T_y\mathcal{N} \to \mathfrak{g}$ is such that the extended derivative map, with $u = \chi(y)$,

$$T_y\mathcal{N} \to T_u\mathcal{M} \times \mathfrak{g}\ :\ \dot y \mapsto \big(d\chi(y)\dot y,\, \gamma(y)\dot y\big) \qquad \text{is an isomorphism.}$$

Hence, under the gauge condition $\gamma(y)\dot y = 0$ (or with any fixed element of $\mathfrak{g}$ instead of 0), $\dot y \in T_y\mathcal{N}$ is determined uniquely by y and $\dot u \in T_u\mathcal{M}$. In the Hartree method, a gauge map is given by $\gamma(y)\dot y = \big(\langle\varphi_n \,|\, \dot\varphi_n\rangle\big)_{n=1}^N \in \mathbb{C}^N$.

II.3.2 Time-Dependent Hartree–Fock Method (TDHF)

Slater Determinants. For a system of N identical fermions the wave function is antisymmetric (see Sect. I.5.2) and we wish to retain this property in the approximation. We therefore look for an approximate wave function in an antisymmetrized tensor product form, that is, as a *Slater determinant*

$$\psi(x_1,\dots,x_N,t) \approx a(t)\frac{1}{\sqrt{N!}}\,\det\bigl(\varphi_n(x_j,t)\bigr)_{n,j=1}^{N}$$

with a scalar phase factor $a(t)$ and with orbitals $\varphi_n(x,t)$ that are time-dependent functions of $x \in \mathbb{R}^3$. In the following we write the scaled determinant as the wedge product

$$\varphi_1 \wedge \cdots \wedge \varphi_n = \frac{1}{\sqrt{N!}} \sum_{\sigma\in S_N} \mathrm{sign}\,(\sigma)\, \varphi_{\sigma(1)} \otimes \cdots \otimes \varphi_{\sigma(N)}\,,$$

where the sum is over all permutations of $\{1,\dots,N\}$. We consider the variational approximation (1.2) on the manifold

$$\mathcal{M} = \{u \in L^2(\mathbb{R}^{3N}) \,:\, u \neq 0,\ u = a\,\varphi_1 \wedge \cdots \wedge \varphi_N\,,\ a \in \mathbb{C},\ \varphi_n \in L^2(\mathbb{R}^3)\}\,. \tag{3.13}$$

The representation of $u \in \mathcal{M}$ as $u = a\,\varphi_1 \wedge \cdots \wedge \varphi_N$ again is not unique: u remains unaltered under the transformation by any invertible $N \times N$ matrix, $A \in \mathrm{GL}(N)$, by

$$\begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_N \end{pmatrix} \to A \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_N \end{pmatrix}, \qquad a \to \frac{a}{\det(A)}\,.$$

We may therefore choose to work with orthonormal orbitals:

$$\langle \varphi_n \,|\, \varphi_j \rangle = \delta_{nj} \quad \text{ for all } n, j\,. \tag{3.14}$$

Tangent Functions. Consider $u = a\,\varphi_1 \wedge \cdots \wedge \varphi_N$ with a of unit modulus and with orthonormal orbitals φ_n. Every tangent function $\dot{u} \in T_u\mathcal{M}$ is of the form

$$\dot{u} = \dot{a}\,\varphi_1 \wedge \cdots \wedge \varphi_N + a\,\dot{\varphi}_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_N + \cdots + a\,\varphi_1 \wedge \cdots \wedge \varphi_{N-1} \wedge \dot{\varphi}_N \tag{3.15}$$

where $\dot{a} \in \mathbb{C}$ and $\dot{\varphi}_n \in L^2$. These turn out to be uniquely determined by $\dot{u}$ and the fixed $a, \varphi_1, \dots, \varphi_n$ if we impose the gauge condition

$$\langle \varphi_n \,|\, \dot{\varphi}_j \rangle = 0 \qquad \text{for all } n, j\,. \tag{3.16}$$

Indeed, taking the inner product of both sides of (3.4) with $u = a\,\varphi_1 \wedge \cdots \wedge \varphi_N$ and using (3.14) and (3.16) and $a = 1/\overline{a}$, determines $\dot{a}$ again as

$$\dot{a} = \langle u \,|\, \dot{u} \rangle\, a\,. \tag{3.17}$$

Taking the inner product with the function in which φ_n is replaced by some L^2 function ϑ_n, determines $\dot{\varphi}_n$ uniquely by the analogue of (3.7), where now simply the wedge product replaces the tensor product:

$$\langle \vartheta_n \,|\, \dot{\varphi}_n \rangle + \overline{a}\dot{a}\,\langle \vartheta_n \,|\, \varphi_n \rangle = \langle a\,\varphi_1 \wedge \cdots \wedge \vartheta_n \wedge \cdots \wedge \varphi_N \,|\, \dot{u} \rangle \quad \forall\, \vartheta_n \in L^2\,. \tag{3.18}$$

Equations of Motion for the Orbitals. The equations of motion for the orbitals in the variational approximation (1.2) on the Hartree–Fock manifold $\mathcal{M}$ in the weak form therefore still are of the same type as in (3.9), where just $\wedge$ formally replaces $\otimes$. With the constant total energy $E = \langle u \,|\, H \,|\, u \rangle$, we have

$$i\hbar \left\langle \vartheta_n \,\middle|\, \frac{\partial \varphi_n}{\partial t} \right\rangle = \left\langle a\, \varphi_1 \wedge \cdots \wedge \vartheta_n \wedge \cdots \wedge \varphi_N \,\middle|\, Hu \right\rangle - E \big\langle \vartheta_n \,\big|\, \varphi_n \big\rangle \quad \forall\, \vartheta_n \in L^2 \,. \tag{3.19}$$

To proceed further, we now consider a Hamiltonian composed of identical one- and two-body Hamiltonians:

$$H = \sum_{j=1}^{N} \left(-\frac{\hbar^2}{2m}\, \Delta_{x_j} + U_j \right) + \sum_{k<\ell} W_{k\ell} \equiv \sum_{j=1}^{N} S_j + \sum_{k<\ell} W_{k\ell} \tag{3.20}$$

with identical one-body potentials $U_j(x_1, \ldots, x_N) = U(x_j)$ and identical symmetric two-body potentials

$$W_{k\ell}(x_1, \ldots, x_N) = W(x_k, x_\ell) = W(x_\ell, x_k) \,.$$

The situation of primary interest is that of the electronic Schrödinger equation (2.4), where

$$W(x, y) = \frac{e^2}{|x - y|}$$

is the electron-electron Coulomb potential, and $U(x)$ describes the Coulomb interaction between an electron at $x \in \mathbb{R}^3$ and all nuclei at fixed positions.

We abbreviate the single-particle operator as $S = -\frac{\hbar^2}{2m}\, \Delta_x + U$, and write S_j when it is considered as an operator acting on the variable x_j of functions of $(x_1, \ldots, x_N)$.

We return to (3.19) and consider functions $\vartheta_n \in L^2(\mathbb{R}^3)$ that satisfy the orthogonality condition

$$\langle \vartheta_n \,|\, \varphi_j \rangle = 0 \quad \text{for all } \; n, j \,. \tag{3.21}$$

Using the definition of the wedge product and the orthogonality relations (3.14) and (3.21) we calculate

$$\langle \varphi_1 \wedge \cdots \wedge \vartheta_n \wedge \cdots \wedge \varphi_N \,|\, S_1 \,|\, \varphi_1 \wedge \cdots \wedge \varphi_N \rangle = \frac{1}{N} \, \langle \vartheta_n \,|\, S\varphi_n \rangle \,,$$

Since the same result is obtained for $S_2, \ldots, S_N$, we obtain

$$\left\langle \varphi_1 \wedge \cdots \wedge \vartheta_n \wedge \cdots \wedge \varphi_N \,\middle|\, \sum_{j=1}^{N} S_j \,\middle|\, \varphi_1 \wedge \cdots \wedge \varphi_N \right\rangle = \big\langle \vartheta_n \,\big|\, S\varphi_n \big\rangle \,. \tag{3.22}$$

For the two-body interaction we obtain similarly, using in addition the symmetry of W,

$$\begin{aligned}&\big\langle \varphi_1 \wedge \cdots \wedge \vartheta_n \wedge \cdots \wedge \varphi_N \,\big|\, W_{12} \,\big|\, \varphi_1 \wedge \cdots \wedge \varphi_N \big\rangle \\ &= \frac{2}{N(N-1)} \sum_{j \neq n} \Big(\langle \vartheta_n \otimes \varphi_j \,|\, W \,|\, \varphi_n \otimes \varphi_j \rangle - \langle \vartheta_n \otimes \varphi_j \,|\, W \,|\, \varphi_j \otimes \varphi_n \rangle \Big) ,\end{aligned}$$

and the same result for the other $W_{k\ell}$. Hence we have

$$\begin{aligned}\big\langle \varphi_1 \wedge \cdots \wedge \vartheta_n \wedge \cdots \wedge \varphi_N \,\big|\, \sum_{k<\ell} W_{k\ell} \,\big|\, \varphi_1 \wedge \cdots \wedge \varphi_N \big\rangle& \\ = \big\langle \vartheta_n \,\big|\, K_n \varphi_n - \sum_{j \neq n} X_{nj} \varphi_j \big\rangle& \qquad (3.23)\end{aligned}$$

with the Hartree potential K_n and the exchange potentials X_{nj} given as

$$K_n(x) = \sum_{j \neq n} \int_{\mathbb{R}^3} W(x,y)\, |\varphi_j(y)|^2 \, dy \qquad (3.24)$$

$$X_{nj}(x) = \int_{\mathbb{R}^3} W(x,y)\, \varphi_j^*(y)\, \varphi_n(y)\, dy \,. \qquad (3.25)$$

Substituting (3.22) and (3.23) into (3.19), we thus obtain, for all $\vartheta_n \in L^2(\mathbb{R}^3)$ satisfying the orthogonality relations (3.21),

$$\Big\langle \vartheta_n \,\Big|\, i\hbar \frac{\partial \varphi_n}{\partial t} - S\varphi_n - K_n \varphi_n + \sum_{j \neq n} X_{nj} \varphi_j \Big\rangle = 0 \,.$$

It follows that the right-hand expression in the inner product is in the linear span of $\varphi_1, \ldots, \varphi_N$. Since adding such a term to $\partial \varphi_n / \partial t$ adds to $\dot{u} = du/dt$ of (3.15) only a scalar multiple of u and hence changes only the scalar phase factor, this term is ignored and we set the right-hand expression in the inner product to zero. On multiplying with φ_j and interchanging n and j, we then further obtain

$$\frac{d}{dt} \langle \varphi_n \,|\, \varphi_j \rangle = \Big\langle \varphi_n \,\Big|\, \frac{\partial \varphi_j}{\partial t} \Big\rangle + \overline{\Big\langle \varphi_j \,\Big|\, \frac{\partial \varphi_n}{\partial t} \Big\rangle} = 0 \,,$$

so that the orthonormality relations (3.14) are preserved for all times. We summarize the result as follows.

Theorem 3.3 (Time-Dependent Hartree–Fock Method, Dirac 1930). *For a Hamiltonian (3.20), the variational approximation (1.2) on the Hartree–Fock manifold (3.13), for initial data* $u(x_1, \ldots, x_N, 0) = \frac{1}{\sqrt{N!}} \det \big(\varphi_n(x_j, 0)\big)_{n,j=1}^N$ *with* $\varphi_n(\cdot, 0)$ *satisfying the orthonormality relations (3.14), is given as*

$$u(x_1, \ldots, x_N, t) = a(t)\, \frac{1}{\sqrt{N!}} \det \big(\varphi_n(x_j, t)\big)_{n,j=1}^N \,,$$

where $|a(t)| = 1$ *and* $\varphi_n(x,t)$ *are solutions to the system of partial differential equations*

$$i\hbar \frac{\partial \varphi_n}{\partial t} = -\frac{\hbar^2}{2m} \Delta \varphi_n + U\varphi_n + K_n \varphi_n - \sum_{j \neq n} X_{nj} \varphi_j \,, \tag{3.26}$$

which is nonlinear through the Hartree and exchange potentials given by (3.24) and (3.25). This holds on time intervals $0 \leq t \leq \overline{t}$ *on which a strong solution to this system exists, that is, for* $\varphi_n \in C^1([0,\overline{t}], L^2(\mathbb{R}^3)) \cap C([0,\overline{t}], H^2(\mathbb{R}^3))$. *The orthonormality (3.14) of the orbitals is preserved on the whole time interval.* □

Comparing (3.26) with the Hartree equations (3.12), we note that the only, but essential difference is in the presence of the fermionic exchange terms $X_{nj}\varphi_j$.

Global existence of strong solutions to the equations of motion (3.26) in the case of Coulomb potentials is shown by Chadam & Glassey (1974). The line of their argument runs as follows: first it is shown by Picard iteration that solutions in H^2 exist *locally* in time, where the growth in the H^2 norm is exponential in terms of a bound of the H^1 norm of the solution. Since the H^1 norm of a strong solution can be bounded by the constant total energy $\langle H \rangle$, it follows that the H^1 norm remains in fact bounded for all times and the H^2 norm grows at worst exponentially.

Spin Orbitals. Electrons are distinguished by their spin which can take the two values up ($\uparrow$) and down ($\downarrow$). In a system with K electrons of spin up and $N-K$ electrons of spin down, the separable approximation with the correct antisymmetry properties is

$$u = a\,(\varphi_1 \wedge \cdots \wedge \varphi_K) \otimes (\varphi_{K+1} \wedge \cdots \wedge \varphi_N) \tag{3.27}$$

with $a \in \mathbb{C}$, $\varphi_n \in L^2(\mathbb{R}^3)$. The equations of motion for variational approximations of this type can be derived in the same way as above and turn out to be identical to (3.26) if the interpretation of inner products is modified as follows: we extend each orbital φ_n to a *spin orbital* $\widehat{\varphi}_n = (\varphi_n, s_n)$ with spin $s_n \in \{\uparrow, \downarrow\}$. For any observable A of orbitals we define

$$\langle \widehat{\varphi}_n \,|\, A \,|\, \widehat{\varphi}_j \rangle = \begin{cases} \langle \varphi_n \,|\, A \,|\, \varphi_j \rangle & \text{if } \; s_n = s_j \,, \\ 0 & \text{else.} \end{cases}$$

With this interpretation of all arising inner products, the equations of motion (3.26) remain valid for the spin orbitals $\widehat{\varphi}_n$, with non-vanishing exchange terms remaining only between spin orbitals of the same spin.

As opposed to the *unrestricted* Hartree-Fock approximation just described, the *restricted* Hartree-Fock method in the case of an even number N of electrons assumes an equal number $N/2$ of electrons with spin up and spin down with the spin orbitals $(\varphi_n, \uparrow)$ and $(\varphi_n, \downarrow)$ for $n = 1, \ldots, N/2$, that is, with the *same* spatial orbital φ_n for both spin up and spin down. The approximation to the wave function is thus chosen of the form

$$u = a\,(\varphi_1 \wedge \cdots \wedge \varphi_{N/2}) \otimes (\varphi_1 \wedge \cdots \wedge \varphi_{N/2}) \tag{3.28}$$

in the restricted Hartree-Fock method. For an initial state of this type, it is seen that this restricted form is preserved for all times in the equations of motion (3.26) of the unrestricted Hartree-Fock method with $N/2$ electrons of spin up and down each. Therefore half of the equations can be dropped in this case.

II.3.3 Multi-Configuration Methods (MCTDH, MCTDHF)

Multi-Configurations. We consider again the Schrödinger equation (3.1) for the nuclei that are supposed to be distinguishable by their different types or by their well-separated positions. It is to be expected, and has found ample confirmation in computations, that a better approximation to the wave function can be obtained by using a linear combination of tensor products instead of just a single tensor product, as is done in the time-dependent Hartree method of Section II.3.1. We therefore consider approximations

$$\begin{aligned}\psi(x_1,\dots,x_N,t) \quad &\approx \quad \sum_{(j_1,\dots,j_N)} a_{j_1,\dots,j_N}(t)\,\varphi_{j_1}^{(1)}(x_1,t)\cdots\varphi_{j_N}^{(d)}(x_N,t) \\ &\equiv \sum_J a_J(t)\,\Phi_J(x,t)\,. \end{aligned} \tag{3.29}$$

Here, the multi-indices $J=(j_1,\dots,j_N)$ vary for $j_n=1,\dots,d_n,\ n=1,\dots,N$, the $a_J(t)$ are complex coefficients depending only on t, and the single-particle functions $\varphi_{j_n}^{(n)}(x_n,t)$ depend on the coordinates $x_n\in\mathbb{R}^3$ of particle n and on time t. Alternatively, we might take Hartree products of $3N$ functions depending on $x_n\in\mathbb{R}$.

This is a model reduction analogous to low-rank approximation of matrices, where a large system matrix is replaced by a linear combination of rank-1 matrices $v\otimes w$, or to low-rank approximation of tensors by linear combinations of rank-1 tensors $v_1\otimes\cdots\otimes v_N$.

In the *multi-configuration time-dependent Hartree* (MCTDH) method proposed by Meyer, Manthe & Cederbaum (1990) and developed further as described by Beck, Jäckle, Worth & Meyer (2000), the Dirac–Frenkel time-dependent variational principle (1.2) is used to derive differential equations for the coefficients a_J and the single-particle functions $\varphi_{j_n}^{(n)}$. The MCTDH method determines approximations to the wave function that, for every time t, lie in the set

$$\overline{\mathcal{M}} = \Big\{ u\in L^2(\mathbb{R}^{3N}) : u=\sum_J a_J\,\varphi_{j_1}^{(1)}\otimes\cdots\otimes\varphi_{j_d}^{(d)} \ \text{ with } \ a_J\in\mathbb{C},\ \varphi_{j_n}^{(n)}\in L^2(\mathbb{R}^3)\Big\}$$

with multi-indices $J=(j_1,\dots,j_N)$ ranging over $j_n=1,\dots,d_n$. This set $\overline{\mathcal{M}}$ is not a manifold, but it contains a dense subset $\mathcal{M}$ that is a manifold and is characterized by a full-rank condition to be given below.

The representation of $u\in\overline{\mathcal{M}}$ by a coefficient tensor $A=(a_J)$ and single-particle functions $\Phi=\big(\varphi_{j_n}^{(n)}\big)$ clearly is not unique: the transformation

$$\begin{aligned}\varphi_{j_n}^{(n)} &\to \widehat{\varphi}_{j_n}^{(n)} = \sum_{k_n=1}^{d_n} S_{j_n,k_n}^{(n)}\varphi_{k_n}^{(n)}, \\ a_J &\to \widehat{a}_J = \sum_{i_1=1}^{d_1}\cdots\sum_{i_N=1}^{d_N} a_I (S^{(1)})^{-1}_{i_1,j_1}\cdots(S^{(N)})^{-1}_{i_N,j_N}\end{aligned}$$

yields the same function u for any choice of nonsingular matrices $S^{(1)}, \dots, S^{(N)}$. We may assume that the orbitals $\varphi_{j_n}^{(n)}$ corresponding to the same particle n are orthonormal:

$$\langle \varphi_{j_n}^{(n)} \,|\, \varphi_{k_n}^{(n)} \rangle = \delta_{j_n,k_n}, \qquad j_n, k_n = 1, \dots, d_n, \;\; n = 1, \dots, N\,. \tag{3.30}$$

Tangent Functions. Consider a differentiable path $(A(t), \Phi(t))$ of coefficients and single-particle functions representing a path $u(t)$ on $\mathcal{M}$. Then, the derivative $\dot{u}$ is of the form

$$\dot{u} = \sum_J \dot{a}_J \, \Phi_J + \sum_{k=1}^{d} \sum_{j_k=1}^{d_k} \dot{\varphi}_{j_k}^{(k)} \, \psi_{j_k}^{(k)} \tag{3.31}$$

with the Hartree products $\Phi_J = \bigotimes_{n=1}^N \varphi_{j_n}^{(n)}$ and with the *single-hole functions*

$$\begin{aligned} \psi_{j_n}^{(n)} &= \langle \varphi_{j_n}^{(n)} \,|\, u \rangle^{(n)} \\ &= \sum_{j_1=1}^{d_1} \cdots \sum_{j_{n-1}=1}^{d_{n-1}} \sum_{j_{n+1}=1}^{d_{n+1}} \cdots \sum_{j_N=1}^{d_N} a_{j_1,\dots,j_d} \bigotimes_{k \neq n} \varphi_{j_k}^{(k)} \end{aligned} \tag{3.32}$$

where the superscript (n) on the inner product indicates that the L^2 inner product is taken only with respect to the variable x_n, leaving a function depending on all the other variables x_k with $k \neq n$.

Conversely, the $\dot{a}_J$ are uniquely determined by $\dot{u}$ and (A, Φ) if we impose the orthogonality condition

$$\langle \varphi_{j_n}^{(n)} \,|\, \dot{\varphi}_{k_n}^{(n)} \rangle = 0\,, \qquad j_k, k_n = 1, \dots, d_n, \;\; n = 1, \dots, N\,, \tag{3.33}$$

which together with (3.30) implies

$$\dot{a}_J = \langle \Phi_J \,|\, \dot{u} \rangle\,. \tag{3.34}$$

Taking the inner product of (3.31) with $\psi_{i_n}^{(n)}$ then gives

$$\sum_{j_n=1}^{d_n} \rho_{i_n,j_n}^{(n)} \, \dot{\varphi}_{j_n}^{(n)} = \Big\langle \psi_{i_n}^{(n)} \,\Big|\, \dot{u} - \sum_J \dot{a}_J \, \Phi_J \Big\rangle^{(\neg n)} \tag{3.35}$$

with the hermitian, positive semi-definite *density matrices*

$$\rho^{(n)} = \big(\rho_{i_n,j_n}^{(n)}\big)_{i_n,j_n=1}^{d_n} \qquad \text{given by} \qquad \rho_{i_n,j_n}^{(n)} := \langle \psi_{i_n}^{(n)} | \psi_{j_n}^{(n)} \rangle. \tag{3.36}$$

The superscript $(\neg n)$ indicates that the L^2 inner product is taken over all variables except x_n, leaving a function depending on x_n. The orthonormality relations (3.30) allow us to express the entries of the density matrices in terms of the coefficients a_J:

$$\rho^{(n)}_{i_n,j_n} = \sum_{j_1=1}^{d_1} \cdots \sum_{j_{n-1}=1}^{d_{n-1}} \sum_{j_{n+1}=1}^{d_{n+1}} \cdots \sum_{j_N=1}^{d_N} \bar{a}_{j_1,\ldots,j_{n-1},i_n,j_{n+1},\ldots,j_N} a_{j_1,\ldots,j_N} \,. \quad (3.37)$$

The $\dot\varphi^{(n)}_{j_n}$ are thus uniquely determined from (3.35) under the *full-rank condition* that

$$\rho^{(n)} \text{ is an invertible matrix for each } n = 1, \ldots, N \,. \quad (3.38)$$

(In view of (3.37), a necessary condition for this property is $d_n \le \prod_{k\neq n} d_k$.)

The MCTDH manifold. With the above construction of the $\dot a_J$ and $\dot\varphi^{(n)}_{j_n}$, one can construct local charts on

$$\begin{aligned} \mathcal{M} \;=\; \Big\{ & u \in L^2(\mathbb{R}^{3N}) : \; u = \sum_J a_J \, \varphi^{(1)}_{j_1} \otimes \cdots \otimes \varphi^{(d)}_{j_d} \text{ with } a_J \in \mathbb{C} \text{ and} \\ & \varphi^{(n)}_{j_n} \in L^2(\mathbb{R}^3) \text{ satisfying the orthonormality condition (3.30)} \\ & \text{and the full-rank condition (3.38)} \Big\} \,, \end{aligned} \quad (3.39)$$

making this set an infinite-dimensional manifold, for which the tangent space at $u \in \mathcal{M}$ consists of the elements $\dot u$ of the form (3.31).

Equations of Motion for the Multi-Configuration Time-Dependent Hartree Method. The MCTDH method uses the time-dependent variational principle (1.2) on this approximation manifold $\mathcal{M}$. The equations of motion are thus obtained by substituting $\frac{1}{i\hbar} H u$ for $\dot u$ in (3.34) and (3.35), and so we have the following result.

Theorem 3.4 (MCTDH Method; Meyer, Manthe & Cederbaum 1990). *The variational approximation on the MCTDH manifold (3.39) is given by (3.29), where the coefficients and single-particle functions are solutions to the system of coupled ordinary and partial differential equations*

$$i\hbar \frac{da_J}{dt} \;=\; \sum_K \langle \Phi_J | H | \Phi_K \rangle \, a_K \,, \qquad \forall\, J = (j_1, \ldots, j_N) \,, \quad (3.40)$$

$$i\hbar \frac{\partial \varphi^{(n)}_{j_n}}{\partial t} \;=\; (I - P^{(n)}) \sum_{k_n=1}^{d_n} \sum_{l_n=1}^{d_n} (\rho^{(n)})^{-1}_{j_n,k_n} \langle \psi^{(n)}_{k_n} \,|\, H \,|\, \psi^{(n)}_{l_n} \rangle^{(\neg n)} \varphi^{(n)}_{l_n}, \quad (3.41)$$
$$j_n = 1, \ldots, d_n, \;\; n = 1, \ldots, N \,.$$

This holds on every time interval on which a strong solution to these equations exists. Here, the Hartree products Φ_J, the single-hole functions $\psi^{(n)}_{j_n}$, and the density matrices $\rho^{(n)}$ are defined in (3.29), (3.32), and (3.37), respectively. The superscript $(\neg n)$ indicates that the inner product is over all variables except x_n, and $P^{(n)}$ is the orthogonal projector onto the linear span of $\varphi^{(n)}_1, \ldots, \varphi^{(n)}_{d_n}$. □

We note that the projector $P^{(n)}$ is given as $P^{(n)}\vartheta = \sum_{j_n=1}^{d_n} \varphi_{j_n}^{(n)} \langle \varphi_{j_n}^{(n)} \mid \vartheta \rangle$, with the inner product over the variable x_n.

For a smooth bounded potential with bounded derivatives, it is shown by Koch & Lubich (2007a) that a strong solution $\varphi_{j_n}^n \in C^1([0,\overline{t}), L^2(\mathbb{R}^3)) \cap C([0,\overline{t}), H^2(\mathbb{R}^3))$ to the MCTDH equations exists either globally for all times or up to a time $\overline{t}$ where a density matrix $\rho^{(n)}$ becomes singular.

At a singularity of a density matrix $\rho^{(n)}$, the equations of motion break down. To avoid such problems in computations, the density matrices are usually regularized to $\rho^{(n)} + \mu I$ with a small parameter μ. Although such regularized solutions exist for all times, a near-singularity can still cause numerical problems, for example in the step size selection of a time integration method.

The MCTDH method has been used successfully for accurately computing the quantum dynamics of small molecules in a variety of chemical situations such as photodissociation and reactive scattering, for problems involving 6 to 24 nuclear degrees of freedom and one or several electronic states; see, e.g., Raab, Worth, Meyer & Cederbaum (1999).

The complexity of the method grows exponentially with the number of particles: there are d^N coefficients a_J if $d_n = d$ orbitals are taken for each particle. Several variants and extensions of the MCTDH method have been designed for the computational treatment of larger systems, such as the coupling with Gaussian wavepackets for secondary modes (Burghardt, Meyer & Cederbaum 1999) and the *hierarchical, cascadic* or *multilayer* versions of MCTDH (Beck, Jäckle, Worth & Meyer 2000, Wang & Thoss 2003) with which particular systems with up to 500 degrees of freedom have been treated.

Hierarchical MCTDH Method. Considering for simplicity a system with $N = 2^L$ particles, the binary cascadic MCTDH method determines an approximation to the wave function in the form

$$u = \sum_{j,k=1}^{d} a_{jk}\, \varphi_j^{(0)} \otimes \varphi_k^{(1)}$$

where, for a binary number $B = (b_1, \ldots, b_\ell)$ with $b_m \in \{0,1\}$ and $\ell < L$, we set recursively

$$\varphi_i^B = \sum_{j,k=1}^{d} a_{i,jk}^B\, \varphi_j^{(B,0)} \otimes \varphi_k^{(B,1)}\,,$$

and for $\ell = L$ we have the single-particle functions. The variational approximation u is thus built up from a binary tree, with the single-particle functions sitting at the end of the branches. This approach uses only $d^3 N$ instead of d^N coefficients.

The orthogonality relations (3.30) and (3.33) can now be imposed on each level: at the final level for the single-particle functions and at the other levels by

$$\langle \varphi_\ell^B \,|\, \varphi_i^B \rangle = \sum_{j,k=1}^{d} \overline{a_{\ell,jk}^B}\, a_{i,jk}^B = \delta_{i\ell}$$
$$\langle \varphi_\ell^B \,|\, \dot\varphi_i^B \rangle = \sum_{j,k=1}^{d} \overline{a_{\ell,jk}^B}\, \dot a_{i,jk}^B = 0\,.$$

The derivation of the equations of motion is then analogous to that of the MCTDH method given above, with recurrences climbing up and down the tree for the computation of the required inner products.

Multi-Configuration Time-Dependent Hartree-Fock Method (MCTDHF). For electron dynamics, a multi-configuration extension of the TDHF method is obtained by using the time-dependent variational principle for approximations of the form

$$u = \sum_{1\le j_1<\cdots<j_N\le K} c_{j_1\ldots j_N}\, \varphi_{j_1} \wedge \cdots \wedge \varphi_{j_N} \tag{3.42}$$

with $K > N$. The sum is over all $\binom{K}{N}$ subsets with N elements of $\{1,\ldots,K\}$. The equations of motion of the MCTDHF method for a Hamiltonian (3.20) are those of the MCTDH method with $\varphi_j^{(n)} = \varphi_j$ independent of n and with an antisymmetric tensor: for every multi-index $J = (j_1,\ldots,j_N)$ and permutation $\sigma \in S_N$ and with $\sigma(J) = (j_{\sigma(1)},\ldots,j_{\sigma(N)})$,

$$a_{\sigma(J)} = \mathrm{sign}(\sigma)\, a_J\,.$$

We refer to Zanghellini, Kitzler, Fabian, Brabec & Scrinzi (2003) and Koch, Kreuzer & Scrinzi (2006) for uses and properties of the MCTDHF approach.

No Theoretical Approximation Estimates. While the neighbouring sections close with theoretical results on the approximation error, apparently no such results are available for the methods considered in this section. One might hope that the multi-configuration methods converge to the exact wave function as the number of configurations is increased to infinity, but to date no such result exists. One obstacle to such a convergence result is the fact that the density matrices $\rho^{(n)}$ become more and more ill-conditioned as more nearly irrelevant configurations are included. Another difficulty lies in the time-dependent orbitals whose approximation properties are not under control. In Sect. II.6 we show, however, that for a *fixed* number of configurations, the variational approximation is quasi-optimal in the sense that its error – on sufficiently short time intervals – is bounded in terms of the error of the best approximation to the wave function by the given number of configurations.

Notwithstanding the deficiencies in theory, the methods considered in this section have proven their value in computations on realistic chemical and physical systems — the tool apparently works.

II.4 Parametrized Wave Functions: Gaussian Wave Packets

A further modelling or approximation step consists in replacing the wave function by a function that depends only on a finite number of real or complex parameters. The time-dependent variational principle then yields evolution equations for these parameters that retain a Hamiltonian character, albeit with a non-canonical Poisson bracket. Gaussian wave packets parametrized by position, momentum, complex width and phase are a prominent example. In the classical limit, their variational equations of motion for position and momentum yield the Newtonian equations of classical molecular dynamics.

II.4.1 Variational Gaussian Wave-Packet Dynamics

We consider a Schrödinger equation in semiclassical scaling, for $x \in \mathbb{R}^d$,

$$i\varepsilon \frac{\partial \psi}{\partial t} = H\psi\,, \qquad H = H^\varepsilon = -\frac{\varepsilon^2}{2M}\,\Delta + V\,, \tag{4.1}$$

with a small positive parameter $\varepsilon \ll 1$ (formally in place of $\hbar$, see Sect. II.2.3), a fixed mass parameter $M \sim 1$, and a potential V. The typical situation of (4.1) is the time-dependent Born-Oppenheimer approximation for the motion of nuclei.

As proposed by Heller (1976), the variational approximation of (4.1) can be done by complex Gaussians of the type

$$\psi(x,t) \approx u(x,t) = \exp\!\left(\frac{i}{\varepsilon}\Big((x-q(t))^T C(t)(x-q(t)) + p(t)\cdot(x-q(t)) + \zeta(t)\Big)\right), \tag{4.2}$$

where $q(t) \in \mathbb{R}^d$ is the position average and $p(t) \in \mathbb{R}^d$ is the momentum average of the wave packet. The matrix $C(t) \in \mathbb{C}^{d\times d}$ is a complex symmetric width matrix with positive definite imaginary part, possibly further restricted to a diagonal matrix or just a multiple of the identity, $c(t)I_d$ with complex $c(t)$. Finally, $\zeta(t) \in \mathbb{C}$ is a phase and normalization parameter.

The choice of Gaussians appears attractive because the exact wave function retains the form of a multidimensional Gaussian for all times in the case of a quadratic potential, even for a time-dependent quadratic potential. This useful fact follows from the observation that Hu then is in the tangent space at u, and therefore the variational approximation and the exact wave function coincide. For a narrow wave packet, of width $\sim \varepsilon^{1/2}$ in (4.2), a smooth potential appears locally approximately quadratic, and we may then expect good approximation by Gaussians, as will be made more precise in Sect. II.4.4 in an argument based on the error bound (1.11).

The equations of motion for the parameters read as follows (Heller 1976, Coalson & Karplus 1990): with $\langle A\rangle = \langle u \,|\, A \,|\, u\rangle$ denoting the average of an observable A in the Gaussian state u of unit L^2 norm, we have classically-looking equations for position and momentum, with the average of the gradient ∇V of the potential,

$$\begin{aligned} \dot{q} &= \frac{p}{M} \\ \dot{p} &= -\langle \nabla V \rangle . \end{aligned} \tag{4.3}$$

For the width matrix C and the complex phase ζ we have, with the Hessian $\nabla^2 V$ and with tr denoting the trace of a matrix,

$$\dot{C} = -\frac{2}{M} C^2 - \frac{1}{2} \langle \nabla^2 V \rangle \tag{4.4}$$

$$\dot{\zeta} = \frac{|p|^2}{2M} - \langle V \rangle + \frac{i\varepsilon}{M} \operatorname{tr} C + \frac{\varepsilon}{8} \left\langle \operatorname{tr} \left((\operatorname{Im} C)^{-1} \nabla^2 V \right) \right\rangle . \tag{4.5}$$

When C is restricted to diagonal matrices, then the diagonal part is to be taken on the right-hand side of the differential equation for C. When $C = cI$ is restricted to a multiple of the identity (spherical Gaussians), then the differential equation for c is obtained by taking the trace on both sides of the differential equation for C. If the width matrix is taken constant (frozen Gaussians, Heller 1981), then the equation for C is disregarded, and only the equations for position and momentum and phase remain.

As $\varepsilon \to 0$, the Gaussians (4.2) become narrower and increasingly concentrated at q, and we have $\langle \nabla V \rangle \to \nabla V(q)$ for a Gaussian of unit L^2 norm. Hence the equations for position q and momentum p become the

classical equations of motion in the limit $\varepsilon \to 0$.

The differential equations (4.3)–(4.5) are a regular perturbation to the equations for $\varepsilon = 0$: letting $\varepsilon \to 0$ gives a well-defined limit on the right-hand side. They are no longer a singularly perturbed system as (4.1) is. In contrast to the Gaussian wave packet, the time-dependent parameters are not highly oscillatory functions.

We shall give a derivation of the equations of motion (4.3)–(4.5) that highlights their mathematical structure as a non-canonical Hamiltonian system (or a Poisson system in another terminology). We first study the structure of the variational equations of motion in coordinates on an approximation manifold in a general setting and then return to the particular case of Gaussian wave packets. The presentation in this section essentially follows Faou & Lubich (2006).

II.4.2 Non-Canonical Hamilton Equations in Coordinates

Canonical Poisson structure of the Schrödinger equation. We split $\psi \in L^2(\mathbb{R}^d, \mathbb{C})$ into the real and imaginary parts $\psi = v + iw$. The functions v and w are thus functions in the *real* Hilbert space $L^2(\mathbb{R}^d, \mathbb{R})$. We denote the complex inner product by $\langle \cdot \,|\, \cdot \rangle$ and the real inner product by $(\cdot \,|\, \cdot)$.

As the Hamiltonian H is a real operator, the Schrödinger equation (4.1) can be written

$$\begin{aligned} \varepsilon \dot{v} &= Hw, \\ \varepsilon \dot{w} &= -Hv . \end{aligned} \tag{4.6}$$

With the canonical structure matrix

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and the Hamiltonian function

$$H(v,w) = \langle \psi \,|\, H \,|\, \psi \rangle = (v \,|\, Hv) + (w \,|\, Hw)$$

for $\psi = v + iw$ (we use the same symbol H as for the operator), this becomes the canonical Hamiltonian system

$$\begin{pmatrix} \dot v \\ \dot w \end{pmatrix} = \frac{1}{2\varepsilon}\, J^{-1} \nabla H(v,w)\,.$$

We note that the real multiplication with J corresponds to the complex multiplication with the imaginary unit i.

As in Theorem 1.2, the flow of this system preserves the canonical symplectic two-form

$$\omega(\xi,\eta) = (\xi \,|\, J\eta), \qquad \xi,\eta \in L^2(\mathbb{R}^d,\mathbb{R})^2. \tag{4.7}$$

The associated Poisson bracket is

$$\{F,G\}_{\rm can} = (\nabla F \,|\, J^{-1} \nabla G) \tag{4.8}$$

for functions $F, G : H^1(\mathbb{R}^d,\mathbb{R})^2 \to \mathbb{R}$.

Poisson structure of variational approximations. We consider a finite-dimensional submanifold $\mathcal{M}$ of the complex Hilbert space $L^2(\mathbb{R}^d,\mathbb{C})$ with the property (1.3), i.e., with $v \in T_u\mathcal{M}$ also $iv \in T_u\mathcal{M}$ at every $u \in \mathcal{M}$.

Taking the imaginary part in the Dirac–Frenkel time-dependent variational principle (1.2) on $\mathcal{M}$ yields, upon identifying $u = v+iw$ with the real pair $u = (v,w)^T$,

$$\big(\, \mu \,|\, 2\varepsilon J\dot u - \nabla_u H(u)\big) = 0 \quad \text{ for all } \quad \mu \in T_u\mathcal{M}\,. \tag{4.9}$$

We choose (local) coordinates on $\mathcal{M}$ so that we have a parametrization

$$u = \chi(y)$$

of $\mathcal{M}$, for y in an open subset of $\mathbb{R}^m$. We denote the derivative $X_{\mathbb{C}}(y) = d\chi(y) = V(y) + iW(y)$ or in the real setting as $X = \begin{pmatrix} V \\ W \end{pmatrix}$, which is of full rank for a coordinate map χ. We denote by X^T the adjoint of X with respect to the real inner product $(\cdot \,|\, \cdot)$. Since $\dot u = X(y)\dot y$ and the tangent vectors in $T_u\mathcal{M}$ are given as $\mu = X(y)\eta$ with arbitrary $\eta \in \mathbb{R}^m$, we obtain from (4.9) the differential equation in $\mathbb{R}^m$,

$$2\varepsilon X(y)^T J X(y)\, \dot y = X(y)^T \nabla_u H(\chi(y))\,. \tag{4.10}$$

With $X_{\mathbb{C}}^*$ denoting the adjoint of $X_{\mathbb{C}}$ with respect to the complex inner product $\langle \cdot \,|\, \cdot \rangle$, we note $X_{\mathbb{C}}^* X_{\mathbb{C}} = (V^TV + W^TW) + i(V^TW - W^TV) = X^TX - iX^TJX$ and hence

$$X^TJX = -\operatorname{Im} X_{\mathbb{C}}^* X_{\mathbb{C}}. \tag{4.11}$$

This skew-symmetric matrix is invertible, as the following lemma shows.

Lemma 4.1. *If $T_u\mathcal{M}$ is a complex linear space for every $u \in \mathcal{M}$, then*

$$X(y)^T J X(y) \text{ is invertible for all } y.$$

Proof. We fix $u = \chi(y) \in \mathcal{M}$ and omit the argument y in the following. Since $T_u\mathcal{M} = \text{Range}(X_{\mathbb{C}})$ is complex linear by assumption, there exists a real linear mapping $L : \mathbb{R}^m \to \mathbb{R}^m$ such that $iX_{\mathbb{C}}\eta = X_{\mathbb{C}}L\eta$ for all $\eta \in \mathbb{R}^m$. This implies

$$JX = XL \quad \text{and} \quad L^2 = -\mathrm{Id}$$

and hence $X^T J X = X^T X L$, which is invertible, since X is of full rank. □

We denote the inverse, which is again skew-symmetric, by

$$B(y) = \frac{1}{2\varepsilon}\left(X(y)^T J X(y)\right)^{-1}. \tag{4.12}$$

Introducing the Hamiltonian function on the manifold $\mathcal{M}$ in the coordinates y as

$$K(y) = H(\chi(y)), \tag{4.13}$$

we note $X(y)^T \nabla_u H(\chi(y)) = \nabla_y K(y)$ in (4.10). We then have the following result.

Theorem 4.2 (Variational Equations of Motion in Coordinates). *The differential equations of the variational approximation in coordinates read*

$$\dot{y} = B(y)\nabla_y K(y)\,. \tag{4.14}$$

This is a non-degenerate Poisson system, i.e., the structure matrix $B(y)$ is invertible and generates a bracket

$$\{F, G\}(y) = \nabla F(y)^T B(y) \nabla G(y) \tag{4.15}$$

on smooth real-valued functions F, G, which is antisymmetric ($\{G, F\} = -\{F, G\}$) and satisfies the Jacobi identity ($\{E, \{F, G\}\} + \{F, \{G, E\}\} + \{G, \{E, F\}\} = 0$) and the Leibniz rule ($\{E \cdot F, G\} = E \cdot \{F, G\} + F \cdot \{E, G\}$).

Proof. By (4.10) and the definitions of $B(y)$ and $K(y)$, we have (4.14). It remains to prove the properties of the bracket. Since ε plays no role here, we let $B(y) = \left(X(y)^T J X(y)\right)^{-1}$ (without the factor $\frac{1}{2\varepsilon}$) in this proof. For points $u \in \mathcal{M}$ we introduce the symplectic projector $\Pi(u)$ from the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^D, \mathbb{R})^2$ onto the tangent space $T_u\mathcal{M}$ as

$$\Pi(u) = X(y)B(y)X(y)^T J, \qquad u = \chi(y) \in \mathcal{M}\,,$$

From the induced decomposition $\mathcal{H} = \Pi(u)\mathcal{H} \oplus \big(I - \Pi(u)\big)\mathcal{H}$ we obtain, by the implicit function theorem, a corresponding splitting in a neighbourhood of the manifold $\mathcal{M}$ in $\mathcal{H}$,

$$\psi = u + v \quad \text{with } u \in \mathcal{M},\ \Pi(u)v = 0\,.$$

This permits us to extend functions F to a neighbourhood of $\mathcal{M}$ by setting

$$\widehat{F}(\psi) = F(y) \quad \text{for } \psi = u + v \text{ with } u = \chi(y),\ \Pi(u)v = 0\,.$$

We then have for the derivative $d\widehat{F}(u) = d\widehat{F}(u)\Pi(u)$ and hence for its adjoint, the gradient, $\nabla\widehat{F}(u) = \Pi(u)^T\nabla\widehat{F}(u)$. Moreover, $\nabla F(y) = X(y)^T\nabla\widehat{F}(u)$ for $u = \chi(y)$. For the canonical bracket this gives, at $u = \chi(y)$,

$$\begin{aligned}\{\widehat{F},\widehat{G}\}_{\rm can}(u) &= \nabla\widehat{F}(u)^T\Pi(u)J^{-1}\Pi(u)^T\nabla\widehat{G}(u)\\ &= \nabla F(y)^T B(y)\nabla G(y) = \{F,G\}(y)\,.\end{aligned}$$

Therefore the stated properties follow from the corresponding properties of the canonical bracket. □

More on Poisson systems can be found in Hairer, Lubich & Wanner (2006), Chap. VII.2, and Marsden & Ratiu (1999), Chap. 8.5. In particular, the flow map $\phi_t : y(0) \mapsto y(t)$ is a *Poisson map*, that is, it preserves the Poisson bracket as

$$\{F\circ\phi_t,\, G\circ\phi_t\} = \{F,G\}\circ\phi_t \qquad \forall\, F, G\,.$$

The property of being a Poisson map in the coordinates can be translated to be an equivalent formulation of the symplecticity of the flow on the manifold $\mathcal{M}$ as stated by Theorem 1.2.

II.4.3 Poisson Structure of Gaussian Wave-Packet Dynamics

The variational Gaussian wavepacket dynamics (4.3)–(4.5) is obtained by choosing the manifold $\mathcal{M}$ as consisting of complex Gaussians (4.2). For ease of presentation we give the derivation for spherical Gaussians, where $C = cI_d$ with a complex $c = \alpha + i\beta$ with $\beta > 0$, and I_d is the d-dimensional identity. We write the complex phase as $\zeta = \gamma + i\delta$. We then have the approximation manifold

$$\mathcal{M} = \{u = \chi(y) \in L^2(\mathbb{R}^d) : y = (p, q, \alpha, \beta, \gamma, \delta) \in \mathbb{R}^{2d+4} \text{ with } \beta > 0\} \quad (4.16)$$

with

$$\Big(\chi(y)\Big)(x) = \exp\Big(\frac{i}{\varepsilon}\big((\alpha+i\beta)\,|x-q|^2 + p\cdot(x-q) + \gamma + i\delta\big)\Big). \qquad (4.17)$$

The tangent space $T_u\mathcal{M} \subset L^2(\mathbb{R}^d)$ at a given point $u = \chi(y) \in \mathcal{M}$ is $(2d+4)$-dimensional and is made of the elements of $L^2(\mathbb{R}^d)$ written as

$$\frac{i}{\varepsilon}\Big((A+iB)\,|x-q|^2 + (P - 2(\alpha+i\beta)Q)\cdot(x-q) - p\cdot Q + C + iD\Big)u \quad (4.18)$$

with arbitrary $(P, Q, A, B, C, D)^T \in \mathbb{R}^{2d+4}$. The tangent space $T_u\mathcal{M}$ is indeed complex linear (note $\beta > 0$). Moreover, we have $u \in T_u\mathcal{M}$, and hence Theorem 1.4 shows the preservation of the squared L^2 norm of $u = \chi(y)$, which is given by

$$N(y) = \|\chi(y)\|^2 = \exp\left(-\frac{2\delta}{\varepsilon}\right)\left(\frac{\pi\varepsilon}{2\beta}\right)^{d/2}. \tag{4.19}$$

We then we have the following result.

Theorem 4.3 (Gaussian Wave-Packet Dynamics as a Poisson System). *The variational approximation on the Gaussian wave-packet manifold $\mathcal{M}$ of (4.16)–(4.17) yields the Poisson system*

$$\dot{y} = B(y)\nabla_y K(y) \tag{4.20}$$

where, for $y = (p, q, \alpha, \beta, \gamma, \delta) \in \mathbb{R}^{2d+4}$ with $\beta > 0$,

$$B(y) = \frac{1}{N(y)}\begin{pmatrix} 0 & -I_d & 0 & 0 & -p & 0 \\ I_d & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4\beta^2}{\varepsilon d} & 0 & -\beta \\ 0 & 0 & -\frac{4\beta^2}{\varepsilon d} & 0 & \beta & 0 \\ p^T & 0 & 0 & -\beta & 0 & \frac{d+2}{4}\varepsilon \\ 0 & 0 & \beta & 0 & -\frac{d+2}{4}\varepsilon & 0 \end{pmatrix} \tag{4.21}$$

defines a Poisson structure, and for $u = \chi(y)$,

$$K(y) = \langle\, u \,|\, H \,|\, u \,\rangle = K_T(y) + K_V(y) \tag{4.22}$$

is the total energy, with kinetic and potential parts

$$K_T(y) = N(y)\left(\frac{|p|^2}{2M} + \frac{\varepsilon d}{2M}\,\frac{\alpha^2+\beta^2}{\beta}\right) = \left\langle\, u \,\middle|\, -\frac{\varepsilon^2}{2M}\,\Delta \,\middle|\, u \,\right\rangle$$

and

$$K_V(y) = \int_{\mathbb{R}^d} V(x)\exp\Big(-\frac{2}{\varepsilon}\big(\beta|x-q|^2+\delta\big)\Big)\,dx = \langle\, u \,|\, V \,|\, u \,\rangle.$$

Both $K(y)$ and $N(y)$ are conserved quantities of the system.

Proof. By (4.17), the derivative $X_{\mathbb{C}}(y) = d\chi(y) = \big(\frac{\partial u}{\partial p}, \frac{\partial u}{\partial q}, \frac{\partial u}{\partial \alpha}, \frac{\partial u}{\partial \beta}, \frac{\partial u}{\partial \gamma}, \frac{\partial u}{\partial \delta}\big)$ for $u = \chi(y)$ is written

$$X_{\mathbb{C}}(y) = \frac{i}{\varepsilon}\left(x-q\,,\; -2(\alpha+i\beta)(x-q)-p\,,\; |x-q|^2\,,\; i|x-q|^2\,,\; 1\,,\; i\right)u\,.$$

Calculating the Gaussian integrals, we obtain from (4.11) that

$$2\varepsilon X^T(y) J X(y) = N(y) \begin{pmatrix} 0 & I_d & 0 & 0 & 0 & 0 \\ -I_d & 0 & 0 & \frac{dp}{2\beta} & 0 & \frac{2p}{\varepsilon} \\ 0 & 0 & 0 & -\frac{\varepsilon d(d+2)}{8\beta^2} & 0 & -\frac{d}{2\beta} \\ 0 & -\frac{dp^T}{2\beta} & \frac{\varepsilon d(d+2)}{8\beta^2} & 0 & \frac{d}{2\beta} & 0 \\ 0 & 0 & 0 & -\frac{d}{2\beta} & 0 & -\frac{2}{\varepsilon} \\ 0 & -\frac{2p^T}{\varepsilon} & \frac{d}{2\beta} & 0 & \frac{2}{\varepsilon} & 0 \end{pmatrix}.$$

The inverse of this matrix can be computed explicitly to give the above matrix $B(y)$. Theorem 4.2 then yields the Poisson system, and Theorems 1.1 and 1.4 give the conservation of energy and norm. □

II.4.4 Approximation Error

From the error bound (1.11) we derive the following result, which is closely related to a result by Hagedorn (1980) on non-variational Gaussian wave packets.

Theorem 4.4 (Error Bound for Variational Gaussian Wave Packets). *Consider the variational multidimensional Gaussian wave packet approximation (4.3)–(4.5). Assume that the smallest eigenvalue of the width matrix* $\operatorname{Im} C(t)$ *is bounded from below by a constant* $\rho > 0$*. Assume that the potential* V *is three-times continuously differentiable with a bounded third derivative. Then, the error between the Gaussian wave packet* $u(t)$ *and the exact wave function* $\psi(t)$ *with Gaussian initial data* $\psi(0) = u(0)$ *is bounded in the* L^2 *norm by*

$$\|u(t) - \psi(t)\| \le c\, t\, \varepsilon^{1/2},$$

where c *depends only on* ρ *and the bound of* $\partial^3 V$*.*

Proof. In view of the error bound of Theorem 1.5, we estimate the distance of $\frac{1}{i\varepsilon} H u(t)$ to the tangent space $T_{u(t)}\mathcal{M}$. We split the potential into the quadratic Taylor polynomial at the current position $q(t)$ and the non-quadratic remainder,

$$V = Q_{q(t)} + R_{q(t)},$$

where we note $|R_q(x)| \le \frac{1}{3!} B_3 |x-q|^3$ with a bound B_3 of $\partial^3 V$. Since both Δu and $Q_q u$ are in the tangent space $T_u\mathcal{M}$ given by (4.18), we have

$$\operatorname{dist}\left(\frac{1}{i\varepsilon} H u, T_u\mathcal{M}\right) = \operatorname{dist}\left(\frac{1}{i\varepsilon} R_q u, T_u\mathcal{M}\right) \le \left\| \frac{1}{\varepsilon} R_q u \right\|.$$

With the above bound for R_q and the condition on the width matrix we obtain, for a Gaussian state u of unit L^2 norm,

$$\|R_q u\| \le c_1 \left(\varepsilon^{-d/2} \int_{\mathbb{R}^d} e^{-2\rho|x-q|^2/\varepsilon} |x-q|^6 \, dx \right)^{1/2} \le c\, \varepsilon^{3/2},$$

and hence the result follows with Theorem 1.5. □

As is clear from the proof, the global boundedness of $\partial^3 V$ can be weakened to a bound in a neighbourhood of the positions $q(t)$ and exponential growth outside this region.

We remark that an analogous result does not hold for Gaussian wave packets where the width matrix is restricted to a diagonal matrix.

Though the above result is asymptotically comforting, it must be noted that for realistic values of $\varepsilon \approx 10^{-2}$, a result with a predicted error of $\varepsilon^{1/2}$ cannot necessarily be considered accurate. We will turn to more accurate semiclassical methods briefly in the next section and in more detail in Chapter V.

II.5 Mixed Models, Quantum-Classical Models

There are numerous possibilities for extensions and combinations of the models described in the foregoing sections. The reader may invent his own favourite extension and combination and check out if it has not yet been tried out in the literature. For example, within an MCTDH framework, for some parts of the system the single-particle functions might be chosen as Gaussians, while they are left of a general form for other particles (Burghardt, Meyer & Cederbaum, 1999). Considering the Gaussians of frozen width in such a model and passing to the classical limit $\varepsilon \to 0$ in the equations of motions for positions and momenta then yields equations of motion where most particles are described classically while some are treated quantum-mechanically. For example, this is desired for studying proton transfer in a critical region of a molecule, or more generally for describing a quantum subsystem in a classical bath.

II.5.1 Mean-Field Quantum-Classical Model

Among the various possible mixed quantum-classical models, we now describe the conceptually simplest one which has found widespread use in computations, in spite of its known flaws. Consider a system of light and heavy particles (e.g., protons and the other, heavier nuclei in a molecule), where one would like to describe the light particles quantum-mechanically and the heavy particles classically. Let x and y denote the position coordinates of heavy and light particles, respectively. We consider the Schrödinger equation with the Hamiltonian $H = -\frac{\varepsilon^2}{2}\Delta_x - \frac{1}{2}\Delta_y + V(x,y)$, where ε^2 is the mass ratio as in Section II.2.3. We start from a time-dependent Hartree approximation to the full wave function $\Psi(x,y,t)$

$$\Psi(x,y,t) \approx \gamma(x,t)\,\psi(y,t)\,,$$

where we restrict $\gamma(x,t)$ further to take the form of a frozen Gaussian at variable position $q(t)$ and with variable momentum $p(t)$. When we write down the equations of motion for the corresponding variational approximation and let the width of the Gaussians tend to zero, so that averages over x are replaced by evaluations at the

position $q(t)$, then we obtain the following coupled system of classical and quantum equations where the classical particles are driven by the mean-field potential of the quantum particles, the wave function of which is determined by a Schrödinger equation with a potential evaluated at the current classical position:

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\nabla_q \langle \psi \,|\, V(q,\cdot) \,|\, \psi \rangle \\ i\varepsilon \frac{\partial \psi}{\partial t} &= -\frac{1}{2}\Delta\psi + V(q,\cdot)\psi \,. \end{aligned} \tag{5.1}$$

While this appears as an attractive model at first sight, its mean-field character is flawed. The problem becomes clear by the following argument: Suppose we start with an initial wave function

$$\Psi(x,y,0) = \alpha_1 \gamma_1^0(x)\Phi_1(x,y) + \alpha_2 \gamma_2^0(x)\Phi_2(x,y)\,,$$

where $\Phi_j(x,\cdot)$ are eigenfunctions of $H_e(x) = -\frac{1}{2}\Delta + V(x,\cdot)$ to well-separated eigenvalues $E_j(x)$, of unit L^2_y norm, and γ_j^0 are complex Gaussians of width $\sim \varepsilon^{1/2}$ and unit L^2_x norm. The coefficients should satisfy $|\alpha_1|^2 + |\alpha_2|^2 = 1$ so that Ψ is of unit $L^2_{x,y}$ norm. We then know from Theorems 2.1 and 4.4 that for times $t \sim 1$ the exact wave function $\Psi(x,y,t)$ is approximately, up to an error of order $\varepsilon^{1/2}$,

$$\Psi(x,y,t) \approx \alpha_1 \gamma_1(x,t)\Phi_1(x,y) + \alpha_2 \gamma_2(x,t)\Phi_2(x,y)\,,$$

where $\gamma_j(x,t)$ is a Gaussian located at a position $q_j(t)$ that follows classical equations of motion

$$\dot{q}_j = p_j, \quad \dot{p}_j = -\nabla_q E_j(q_j)\,. \tag{5.2}$$

On the other hand, in (5.1) we have by the time-adiabatic theorem mentioned after Theorem 2.1 that for times $t \sim 1$,

$$\psi(y,t) \approx e^{i\phi_1(t)/\varepsilon}\alpha_1 \Phi_1(q(t),y) + e^{i\phi_2(t)/\varepsilon}\alpha_2 \Phi_2(q(t),y)\,,$$

so that

$$\langle \psi \,|\, H_e(q) \,|\, \psi \rangle \approx |\alpha_1|^2 E_1(q) + |\alpha_2|^2 E_2(q)$$

and hence the classical motion in (5.1) is approximately determined by

$$\dot{q} = p\,, \quad \dot{p} = -\nabla_q \big(|\alpha_1|^2 E_1(q) + |\alpha_2|^2 E_2(q)\big)\,, \tag{5.3}$$

with a potential that is a convex linear combination of the potentials in (5.2). Unless the potentials E_j happen to be quadratic, not even the average position $\alpha_1 q_1 + \alpha_2 q_2$ is described correctly by the equations for q. The equations (5.1) are asymptotically correct, however, if we start from a pure eigenstate (where $\alpha_1 = 1$, $\alpha_2 = 0$).

This example illustrates that even very plausible-looking models must be considered with care and assessed critically by analysis and (numerical and physical) experiment.

For an asymptotic analysis of the above mixed quantum-classical model we refer to Bornemann & Schütte (1999). The quantum-mechanical part can be further restricted, assuming for example $\psi(y,t)$ in the form of a Slater determinant, thus combining classical motion and the time-dependent Hartree-Fock method. Global existence of solutions for such a model has been studied by Cancès & Le Bris (1999).

II.5.2 Quantum Dressed Classical Mechanics

Even if the approximation by a Gaussian wave packet is too rough, it can nevertheless be reused in a correction scheme, which is once more based on the time-dependent variational principle. We briefly describe such an approach due to Billing (2003). Let $q(t)$, $p(t)$ be defined by Gaussian wave packet dynamics with a diagonal width matrix with entries $c_n(t)$, possibly further simplified by using the classical equations of motion for q and p and a similar simplification in the differential equations for the widths, replacing averages by point evaluations. We search for an approximation to the wave function of the form

$$\psi(x_1,\ldots,x_N,t) \approx \sum_J a_J(t)\,\gamma_{j_1}^{(1)}(x_1,t)\cdot\ldots\cdot\gamma_{j_N}^{(N)}(x_N,t)\,,$$

where the sum is over a set of multi-indices $J=(j_1,\ldots,j_N)$ and the functions $\gamma_j^{(n)}$ are shifted and scaled Gauss-Hermite basis functions defined by (we assume all x_n one-dimensional for simplicity)

$$\begin{aligned}\gamma_j^{(n)}(x_n,t) \;&=\; \exp\!\left(\frac{i}{\varepsilon}\Big(c_n(t)\,(x_n-q_n(t))^2+p_n(t)\,(x_n-q_n(t))\Big)\right)\cdot\\ &\qquad H_j\!\left(\sqrt{\frac{2\,\mathrm{Im}\,c_n(t)}{\varepsilon}}\,(x_n-q_n(t))\right)\end{aligned}$$

with Hermite polynomials H_j and the known Gaussian parameters $q_n(t)$, $p_n(t)$, and $c_n(t)$. The unknown coefficients $a_J(t)$ are determined by differential equations obtained from the variational principle on the *time-dependent* approximation manifold (here actually a linear space)

$$\mathcal{M}_t=\Big\{u : u(x_1,\ldots,x_N)=\sum_J a_J\,\gamma_{j_1}^{(1)}(x_1,t)\cdot\ldots\cdot\gamma_{j_N}^{(N)}(x_N,t),\ a_J\in\mathbb{C}\Big\}\,,$$

at every instant t as previously in (1.2), except that now du/dt is not sought for in the tangent space of $\mathcal{M}_t$, but as the derivative of a path $u(t)\in\mathcal{M}_t$.

This approach leads to a method which adapts the location and width of the Hermite basis functions to Gaussian wave packets that follow classical trajectories. We will consider in more detail a somewhat related, but computationally favourable approach in Chap. V.

II.5.3 Swarms of Gaussians

In a conceptually similar approach, frozen Gaussians $\gamma_k(x,t)$ first evolve independently according to the classical equations of motion for position and momentum and with the phase given by the action integral $\int_0^t \big(\frac{1}{2}|p_k|^2 - \langle V\rangle_{\gamma_k}\big)ds$, as proposed by Heller (1981). This approximation is then improved upon by taking a linear combination

$$\psi(x,t) \approx \sum_k a_k(t)\,\gamma_k(x,t)\,,$$

where the coefficients $a_k(t)$ are determined by the time-dependent variational principle:

$$\Big\langle \sum_j b_j\gamma_j \,\Big|\, \sum_k (\dot a_k\gamma_k + a_k\dot\gamma_k) - \frac{1}{i\varepsilon}\,H\sum_k a_k\gamma_k \Big\rangle = 0 \qquad \forall\, b = (b_j)\,.$$

This yields a linear system of differential equations for $a = (a_k)$,

$$M\dot a = \frac{1}{i\varepsilon}\,Ka - La$$

with the matrices $M = \big(\langle\gamma_j\,|\,\gamma_k\rangle\big)$, $L = \big(\langle\gamma_j\,|\,\dot\gamma_k\rangle\big)$, $K = \big(\langle\gamma_j\,|\,H\,|\,\gamma_k\rangle\big)$. While the L^2 norm of the approximation is conserved, the total energy and symplecticity are *not* conserved by applying the variational principle on a time-dependent approximation space as is done here, in contrast to the case of a time-independent approximation manifold as studied in Sect. II.1.3.

The above approach was mentioned by Heller (1981) and has been carried further by Ben-Nun & Martinez (1998, 2000) together with criteria when to create, or "spawn" new basis functions. It is related in spirit to particle methods in fluid dynamics; see, e.g., Monaghan (1992) and Yserentant (1997).

II.6 Quasi-Optimality of Variational Approximations

In this theoretical section we consider variational approximation on a manifold $\mathcal{M}$ and study the following question: In case the true wave function remains close to the manifold, does the time-dependent variational principle then provide a good approximation? Stated differently: Can the error of the variational approximation be bounded in terms of the error of the best approximation to the wave function on $\mathcal{M}$?

This is a familiar question in other areas of numerical analysis; cf. Céa's lemma on the optimality of Galerkin approximations of elliptic boundary value problems as stated, e.g., in Ciarlet (1991), p. 113. A positive answer to this question separates the problems of approximability of the wave function on the chosen manifold, which often is a modeling hypothesis, and the quality of the time-dependent variational principle for obtaining approximate wave functions.

Following Lubich (2005), we give a conditionally positive answer under assumptions that include, for example, the time-dependent Hartree method and its multi-configuration versions.

Assumptions. We consider the Schrödinger equation (1.1) on a Hilbert space $\mathcal{H}$, with $\hbar = 1$ in the following, and the variational approximation given by the Dirac-Frenkel principle (1.2) on the manifold $\mathcal{M}$. The Hamiltonian H is split as

$$H = A + B \tag{6.1}$$

with self-adjoint linear operators A and B where A corresponds to the separable part: $u \in \mathcal{M}$ implies $e^{-itA}u \in \mathcal{M}$ for all t. This is satisfied if and only if

$$Au \in T_u\mathcal{M} \qquad \text{for all} \quad u \in \mathcal{M} \cap D(A). \tag{6.2}$$

We assume that the non-separable remainder B is bounded:

$$\|B\varphi\| \le \beta \, \|\varphi\| \tag{6.3}$$

for all $\varphi \in \mathcal{H}$. About the approximation manifold $\mathcal{M}$ we assume the condition (1.3) of complex linear tangent spaces $T_u\mathcal{M}$, and a condition that is satisfied if $\mathcal{M}$ contains rays (cf. Theorem 1.4):

$$u \in T_u\mathcal{M} \qquad \text{for all} \quad u \in \mathcal{M} \, , \tag{6.4}$$

A bound of the curvature of $\mathcal{M}$ is formulated in terms of the orthogonal projectors $P(u) : \mathcal{H} \to T_u\mathcal{M}$ and $P^\perp(u) = I - P(u)$:

$$\begin{aligned} \| \, (P(u) - P(v))\varphi \, \| &\le \kappa \, \|u - v\| \cdot \|\varphi\| && (6.5) \\ \| \, P^\perp(v)(u - v) \, \| &\le \kappa \, \|u - v\|^2 && (6.6) \end{aligned}$$

for all $u, v \in \mathcal{M}$ and $\varphi \in \mathcal{H}$. We assume that $P(u(t))\varphi$ is a continuously differentiable function of t in $\mathcal{H}$ for every continuously differentiable path $u(t)$ on $\mathcal{M}$ and $\varphi \in \mathcal{H}$.

The initial data $\psi(0)$ is assumed to be on $\mathcal{M}$ and of unit norm. We consider a time interval on which the solution $\psi(t)$ to (1.1) remains near $\mathcal{M}$, in the sense that

$$\operatorname{dist}(\psi(t), \mathcal{M}) \le \frac{1}{2\kappa} \qquad \text{for} \quad 0 \le t \le \overline{t} \, . \tag{6.7}$$

Both the exact wave function $\psi(t)$ and the variational approximation $u(t)$ of (1.2) are required to be in the domain of H for $0 \le t \le \overline{t}$, with a bound

$$\|H\psi(t)\| \le \mu \, , \quad \|Hu(t)\| \le \mu \quad \text{and} \quad \|Au(t)\| \le \mu \, . \tag{6.8}$$

Further we consider the distance bound $\delta \le \mu$ given by

$$\operatorname{dist}\,(H\psi(t), T_{v(t)}\mathcal{M}) \le \delta \, , \quad \operatorname{dist}\,(Hu(t), T_{u(t)}\mathcal{M}) \le \delta \, , \tag{6.9}$$

where $v(t) \in \mathcal{M}$ is the nearest point to $\psi(t)$ on $\mathcal{M}$:

$$\|v(t) - \psi(t)\| = \operatorname{dist}(\psi(t), \mathcal{M}) .$$

Discussion of the assumptions. In all the examples of this chapter, A might be chosen as the kinetic energy operator T, though this might not always be the optimal choice. A critical assumption is the boundedness of the non-separable remainder B. It is a reasonable assumption in the Schrödinger equation of the nuclei and its Hartree and Gaussian wave packet approximations (and their multiconfiguration versions). The condition is not satisfied, however, in the time-dependent Hartree-Fock method for the electronic Schrödinger equation where the Coulomb potentials are non-separable and unbounded. We refer to Lubich (2005) for a corresponding result in the Coulomb case.

We have assumed the splitting (6.1) independent of time for ease of presentation, though the result would directly extend to the situation of a time-dependent splitting $H = A(t) + B(t)$. For example, in the (multi-configuration) Hartree method we might choose $A(t) = T + V_1 + \cdots + V_N$ with the mean-field potentials V_n, so that $B(t)$ becomes the difference between the given potential and the sum of the mean-field potentials. This can be expected to give more favourable error bounds than a time-independent splitting into kinetic energy and potential. Similarly, in Gaussian wave packets we might split into $A(t) = T + Q(t)$ with the local quadratic approximation $Q(t)$ to the potential, and the non-quadratic remainder $B(t)$.

Condition (6.4) is satisfied for all the examples in this chapter. Conditions (6.5) and (6.6) encode curvature information of $\mathcal{M}$ in a form that is suitable for our analysis. Condition (6.7) ensures that $\psi(t)$ has a unique nearest point on $\mathcal{M}$. The regularity assumption (6.8) for $\psi(t)$ is satisfied if the initial value has such regularity. The regularity (6.8) of the approximate solution $u(t)$ needs to be ascertained, but is known to hold, e.g., for the (multiconfiguration) time-dependent Hartree method when the Schrödinger equation for the nuclei has a smooth bounded potential.

The following result bounds the error of the variational approximation in terms of the best-approximation error.

Theorem 6.1 (Quasi-Optimality of Variational Approximations). *Under conditions (6.1)–(6.9), the error of the variational approximation is bounded by*

$$\|u(t) - \psi(t)\| \le d(t) + Ce^{Kt} \int_0^t d(s)\, ds \qquad \text{with} \quad d(t) = \operatorname{dist}(\psi(t), \mathcal{M}) \tag{6.10}$$

and with $K = 2\kappa\delta$ and $C = \beta + 3\kappa\mu$, for $0 \le t \le \overline{t}$.

Though the bound (6.10) can be pessimistic in a concrete situation, it does identify sources that can make the variational approximation deviate far from optimality even if the best-approximation error $d(t)$ is small: large curvature of the approximation manifold (κ), a large effective non-separable potential in the Hamiltonian (β, δ), lack of regularity in the exact or approximate solution (μ, δ), and long time intervals (t).

Proof. The proof compares the differential equation for $u(t)$ with the equation satisfied by the best approximation $v(t) \in \mathcal{M}$ with $\|v(t) - \psi(t)\| = d(t)$.

(a) The function $v(t)$ is implicitly characterized by the condition (omitting the obvious argument t in the sequel)

$$P(v)\,(v - \psi) = 0\,. \tag{6.11}$$

Under condition (6.7), the implicit function theorem can be used to show that this equation has a unique solution in the ball of radius $1/(2\kappa)$ around ψ, which depends continuously differentiable on t. We derive a differential equation for $v(t)$ by differentiating (6.11) with respect to t ($\dot{} = d/dt$):

$$0 = P(v)(\dot{v} - \dot{\psi}) + \big(P'(v) \cdot (v - \psi)\big)\dot{v} \tag{6.12}$$

with $(P'(v)\cdot\varphi)\dot{v} = (d/dt)P(v(t))\varphi$ for $\varphi \in \mathcal{H}$. Since $\dot{v} \in T_v\mathcal{M}$, we have $P(v)\dot{v} = \dot{v}$, and the equation becomes

$$\Big(I + P'(v) \cdot (v - \psi)\Big)\dot{v} = P(v)\dot{\psi}\,. \tag{6.13}$$

By (6.5) and (6.7) we have

$$\|P'(v) \cdot (v - \psi)\| \le \kappa\,\|v - \psi\| \le \frac{1}{2}\,,$$

so that the operator in (6.13) is invertible and

$$\dot{v} = P(v)\dot{\psi} + r(v, \psi) \qquad \text{with} \quad \|r(v, \psi)\| \le 2\kappa\mu\,\|v - \psi\|\,. \tag{6.14}$$

Here we have used the bound (6.8), $\|\dot{\psi}\| = \|H\psi\| \le \mu$. Inserting (1.1) in (6.14), the equation can be written as

$$\dot{v} = P(v)\frac{1}{i}Hv - P(v)\frac{1}{i}H(v - \psi) + r(v, \psi)\,. \tag{6.15}$$

We will compare this differential equation with Equation (1.5) for $u(t)$, viz.,

$$\dot{u} = P(u)\frac{1}{i}Hu\,. \tag{6.16}$$

In the following we tacitly assume $v(t) \in D(H) = D(A)$. If v does not have this regularity, then the proof would proceed by replacing v by a regularized family (v_ε) with $v_\varepsilon(t) \in D(H)$ and $v_\varepsilon \to v$ in $C^1([0, \overline{t}], \mathcal{H})$ as $\varepsilon \to 0$. Applying the arguments below to v_ε and letting $\varepsilon \to 0$ in the final estimate then gives the result.

(b) We form the difference of (6.16) and (6.15), take the inner product with $u - v$ and consider the real part. We then have

$$\|u - v\| \cdot \frac{d}{dt}\,\|u - v\| = \frac{1}{2}\,\frac{d}{dt}\,\|u - v\|^2 = \mathrm{Re}\,\langle u - v \,|\, \dot{u} - \dot{v}\rangle = I + II + III$$

with

$$\begin{aligned} I &= -\operatorname{Re}\langle u-v \,|\, P(u)iHu - P(v)iHv\rangle \\ II &= -\operatorname{Re}\langle u-v \,|\, P(v)iH(v-\psi)\rangle \\ III &= -\operatorname{Re}\langle u-v \,|\, r(v,\psi)\rangle\,. \end{aligned}$$

(c) Using the self-adjointness of $H = A + B$ and condition (6.2), which implies $P^\perp(v)iAv = 0$, we write

$$\begin{aligned} I &= \operatorname{Re}\langle u-v \,|\, P^\perp(u)iHu - P^\perp(v)iHv\rangle \\ &= \operatorname{Re}\langle u-v \,|\, P^\perp(u)iHu\rangle - \operatorname{Re}\langle u-v \,|\, P^\perp(v)iBv\rangle\,. \end{aligned}$$

To treat the expression II, we split

$$II = -\operatorname{Re}\langle u-v \,|\, P(v)iA(v-\psi)\rangle - \operatorname{Re}\langle u-v \,|\, P(v)iB(v-\psi)\rangle\,.$$

It is in the first term that condition (6.4) is used. This condition implies $P(v)v = v$ and hence, by (6.11),

$$v = P(v)\psi, \qquad v-\psi = P^\perp(v)(v-\psi) = -P^\perp(v)\psi\,.$$

It follows that

$$\langle v \,|\, P(v)iA(v-\psi)\rangle = -\langle v \,|\, P(v)iAP^\perp(v)\psi\rangle = \langle P^\perp(v)iAv \,|\, \psi\rangle = 0\,,$$

since $P^\perp(v)iAv = 0$ by (6.2). Similarly, (6.2) implies

$$\langle u \,|\, iAP^\perp(u)(v-\psi)\rangle = 0\,.$$

These equations yield

$$\begin{aligned} &\langle u-v \,|\, P(v)iA(v-\psi)\rangle \\ &\quad = \langle u \,|\, iA(v-\psi)\rangle - \langle u-v \,|\, P^\perp(v)iA(v-\psi)\rangle \\ &\quad = -\langle u \,|\, iA(P^\perp(u)-P^\perp(v))(v-\psi)\rangle + \langle u-v \,|\, P^\perp(v)iA\psi\rangle \\ &\quad = -\langle iAu \,|\, (P(u)-P(v))(v-\psi)\rangle + \langle P^\perp(v)(u-v) \,|\, P^\perp(v)iH\psi\rangle \\ &\qquad -\langle u-v \,|\, P^\perp(v)iB\psi\rangle\,. \end{aligned}$$

We then arrive at the basic equation of the proof,

$$\begin{aligned} I+II &= \operatorname{Re}\langle P^\perp(u)(u-v) \,|\, P^\perp(u)iHu\rangle \\ &\quad - \operatorname{Re}\langle u-v \,|\, iB(v-\psi)\rangle \\ &\quad + \operatorname{Re}\langle iAu \,|\, (P(u)-P(v))(v-\psi)\rangle \\ &\quad - \operatorname{Re}\langle P^\perp(v)(u-v) \,|\, P^\perp(v)iH\psi\rangle\,. \end{aligned}$$

With (6.3)–(6.9) we thus obtain

$$\begin{aligned} |I+II| &\le \kappa\,\|u-v\|^2\cdot\delta + \|u-v\|\cdot\beta\,\|v-\psi\| \\ &\quad + \mu\cdot\kappa\,\|u-v\|\cdot\|v-\psi\| + \kappa\,\|u-v\|^2\cdot\delta \\ &= 2\kappa\delta\,\|u-v\|^2 + (\beta+\kappa\mu)\,\|u-v\|\cdot\|v-\psi\|\,. \end{aligned}$$

(d) Together with (6.14) for bounding III, this estimate gives

$$\frac{d}{dt}\,\|u-v\| \le K\|u-v\| + C\|v-\psi\|$$

with $K = 2\kappa\delta$ and $C = \beta + 3\kappa\mu$. The Gronwall inequality then implies

$$\|u(t)-v(t)\| \le Ce^{Kt}\int_0^t \|v(s)-\psi(s)\|\,ds\,, \tag{6.17}$$

and the triangle inequality for $u-\psi = (u-v)+(v-\psi)$ together with $d = \|v-\psi\|$ yield the result. □

Bibliography

M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965. *[III.1]*

M. Baer & G.D. Billing, eds., *The Role of Degenerate States in Chemistry*, Advances in Chemical Physics, Vol. 124, Wiley, 2002. *[II.1], [II.2]*

M.H. Beck, A. Jäckle, G.A. Worth & H.-D. Meyer, *The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets*, Phys. Reports 324 (2000), 1–105. *[II.3]*

M. Ben-Nun & T.J. Martinez, *Nonadiabatic molecular dynamics: validation of the multiple spawning method for a multidimensional problem*, J. Chem. Phys. 108 (1998), 7244–7257. *[II.5]*

M. Ben-Nun & T.J. Martinez, *A multiple spawning approach to tunneling dynamics*, J. Chem. Phys. 112 (2000), 6113–6121. *[II.5]*

M.V. Berry, *Quantal phase factors accompanying adiabatic changes*, Proc, Royal Soc. London. Series A, Math. Phys. Sci. 392 (1984), 45–57. *[II.2]*

G.D. Billing, *The Quantum Classical Theory*, Oxford Univ. Press, 2003. *[II.5]*

S. Brandt & H.D. Dahmen, *The Picture Book of Quantum Mechanics*, 3rd ed., Springer, New York, 2001. *[I.0]*

M. Born, *Quantenmechanik der Stoßvorgänge*, Z. Phys. 38 (1926), 803–827. *[I.1]*

M. Born & V. Fock, *Beweis des Adiabatensatzes*, Z. Phys. 51 (1928), 165–180. *[II.2]*

F. A. Bornemann and C. Schütte, *On the singular limit of the quantum-classical molecular dynamics model*, SIAM J. Appl. Math. 59 (1999), 1208–1224. *[II.5]*

L. de Broglie, *Recherche sur la théorie des quanta*, Thèse, Paris, 1924. http://tel.archives-ouvertes.fr/tel-00006807 *[I.2]*

I. Burghardt, H.-D. Meyer, & L. S. Cederbaum. *Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method*, J. Chem. Phys. 111 (1999), 2927–2939. *[II.3], [II.5]*

E. Cancès & C. Le Bris, *On the time-dependent HartreeFock equations coupled with a classical nuclear dynamics*, Math. Mod. Meth. Appl. Sci 9 (1999), 963–990. *[II.5]*

J.M. Chadam & R.T. Glassey, *Global existence of solutions to the Cauchy problem for time-dependent Hartree equations*, J. Math. Phys. 16 (1975), 1122–1130. *[II.3]*

P.G. Ciarlet, *Introduction to the Finite Element Method*, in Handbook of Numerical Analysis, Vol. II (P.G. Ciarlet and J.L. Lions, eds.), North-Holland, Amsterdam, 1991, 17–352. *[II.6]*

R.D. Coalson & M. Karplus, *Multidimensional variational Gaussian wavepackets dynamics with application to photodissociation spectroscopy*, J. Chem. Phys. 93 (1990), 3919–3930. *[II.4]*

C. Cohen-Tannoudji, B. Diu & F. Laloë, *Quantum Mechanics*, Vols. I & II, Wiley, New York, 1977. *[I.0]*

P.A.M. Dirac, *Note on exchange phenomena in the Thomas atom*, Proc. Cambridge Phil. Soc. 26 (1930) 376–385. *[.0], [II.0], [II.1], [II.3]*

P.A.M. Dirac, *The Principles of Quantum Mechanics*, Clarendon Press, Oxford, 1930. *[I.0]*

W. Domcke, D.R. Yarkony & H. Köppel (eds.), *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy*, World Scientific, 2004. *[II.2]*

P. Ehrenfest, *Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik*, Z. Phys. 45 (1927), 455–457. *[I.4]*

A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, Ann. Phys. 17 (1905), 132–148. *[I.2]*

E. Faou & V. Gradinaru, *Gauss-Hermite wavepacket dynamics: convergence of the spectral and pseudo-spectral approximation*, Report, 2007. *[III.1]*

E. Faou & C. Lubich, *A Poisson integrator for Gaussian wavepacket dynamics*, Comput. Vis. Sci. 9 (2006), 45–55. *[II.4]*

H. Feldmeier & J. Schnack, *Molecular dynamics for fermions*, Rev. Modern Phys. 72 (2000), 655–688. *[II.1]*

J. Frenkel, *Wave Mechanics, Advanced General Theory*, Clarendon Press, Oxford, 1934. *[II.1]*

W. Gautschi, *Numerical Analysis, An Introduction*, Birkhäuser, Basel, 1997. *[III.1]*

S.J. Gustafson & I.M. Sigal, *Mathematical Concepts of Quantum Mechanics*, Springer, Berlin, 2003. *[I.0], [I.3]*

G.A. Hagedorn, *Semiclassical quantum mechanics. I. The $\hbar \to 0$ limit for coherent states*, Comm. Math. Phys. 71 (1980), 77–93.

E. Hairer, C. Lubich & G. Wanner, *Geometric numerical integration illustrated by the Störmer–Verlet method*, Acta Numerica 12 (2003), 399–450. *[I.3]*

E. Hairer, C. Lubich & G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd edition. Springer, Berlin, 2006. *[II.4]*

W. Heisenberg, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Z. Phys. 33 (1925), 879–893. *[I.4]*

W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Z. Phys. 43 (1927), 172–198. *[I.4]*

E.J. Heller, *Time dependent variational approach to semiclassical dynamics*, J. Chem. Phys. 64 (1976), 63–73. *[II.1], [II.4]*

E.J. Heller, *Frozen Gaussians: A very simple semiclassical approximation*, J. Chem. Phys. 75 (1981), 2923–2931. *[II.4], [II.5]*

P.D. Hislop & I.M. Sigal, *Introduction to Spectral Theory. With Applications to Schrdinger Operators*, Springer, Berlin, 1996. *[I.3]*

T. Jahnke & W. Huisinga, *A dynamical low-rank approach to the chemical master equation*, Report, FU Berlin, 2007.

T. Kato, *On the adiabatic theorem of quantum mechanics*, Phys. Soc. Jap. 5 (1950), 435–439. *[II.2]*

T. Kato, *Fundamental properties of Hamiltonian operators of Schrödinger type*, Trans. Amer. Math. Soc. 70 (1951), 195–211. *[I.5]*

T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Springer, Berlin, 1980. *[I.3]*

Y. Katznelson, *An Introduction to Harmonic Analysis*, Dover Publ., New York, 1976. *[I.2]*

A.K. Kerman & S.E. Koonin, *Hamiltonian formulation of time-dependent variational principles for the many-body system*, Ann. Phys. 100 (1976), 332–358. *[II.1]*

O. Koch, W. Kreuzer & A. Scrinzi. *Approximation of the time-dependent electronic Schrödinger equation by MCTDHF*, Appl. Math. Comput. 173 (2006), 960–976. *[II.3]*

O. Koch & C. Lubich, *Regularity of the multi-configuration time-dependent Hartree approximation in quantum molecular dynamics*, M2AN 41 (2007), 315–332. *[II.3]*

O. Koch & C. Lubich, *Dynamical low rank approximation*, SIAM J. Matrix Anal. Appl. 29 (2007), 434–454. *[II.3]*

P. Kramer & M. Saraceno, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, Lecture Notes in Physics 140, Springer, Berlin, 1981. *[II.1]*

C. Lasser & S. Teufel, *Propagation through conical crossings: an asymptotic semigroup*, Comm. Pure Appl. Math. 58 (2005), 1188–1230. *[II.2]*

C. Le Bris (ed.), *Handbook of Numerical Analysis, Special Volume Computational Chemistry*, Elsevier, Amsterdam, 2003. *[II.2]*

C. Lubich, *On variational approximations in quantum molecular dynamics*, Math. Comp. 74 (2005), 765–779. *[II.6]*

J.E. Marsden & T.S. Ratiu, *Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems*, Second edition, Texts in Applied Mathematics 17, Springer-Verlag, New York, 1999. *[II.1], [II.4]*

A.D. McLachlan, *A variational solution of the time-dependent Schrodinger equation*, Mol. Phys. 8 (1964), 39–44. *[II.1]*

A. Messiah, *Quantum Mechanics*, Dover Publ., New York, 1999 (reprint of the two-volume edition published by Wiley, 1961–1962). *[I.0]*

H.-D. Meyer, U. Manthe & L. S. Cederbaum, *The multi-configurational time-dependent Hartree approach*, Chem. Phys. Lett. 165 (1990), 73–78. *[II.3]*

J.J. Monaghan, *Smoothed particle hydrodynamics*, Annu. Rev. Astron. Astrophys. 30 (1992), 543–574. *[II.5]*

J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932, Reprint 1996. *[I.0], [I.3]*

A. Nonnenmacher & C. Lubich, *Dynamical low-rank approximation: applications and numerical experiments*, Math. Comput. Simulation, to appear. *[II.3]*

A. Raab, G.A. Worth, H.-D. Meyer & L.S. Cederbaum, *Molecular dynamics of pyrazine after excitation to the S_2 electronic state using a realistic 24-mode model Hamiltonian*, J. Chem. Phys. 110 (1999), 936–946. *[II.3]*

M. Reed & B. Simon, *Methods of Modern Mathematical Physics. I: Functional Analysis*, Academic Press, New York, 1972. *[I.3]*

M. Reed & B. Simon, *Methods of Modern Mathematical Physics. II: Fourier Analysis, Self-Adjointness*, Academic Press, New York, 1975. *[I.2], [I.3], [I.5]*

D.J. Rowe, A. Ryman & G. Rosensteel, *Many-body quantum mechanics as a symplectic dynamical system*, Phys. Rev. A 22 (1980), 2362–2373. *[II.1]*

E. Schrödinger, *Quantisierung als Eigenwertproblem (Vierte Mitteilung)*, Ann. Phys. 81 (1926), 109–139. *[I.1]*

B. Simon, *Holonomy, the quantum adiabatic theorem, and Berry's phase*, Phys. Rev. Letters 51 (1983), 2167–2170. *[II.2]*

M.H. Stone, *On One-Parameter Unitary Groups in Hilbert Space*, Ann. Math., 2nd Ser., 33 (1932), 643–648. *[I.3]*

A. Szabo & N.S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Revised ed., Dover Publ., New York, 1996. *[II.2], [II.3]*

D.J. Tannor, *Introduction to Quantum Mechanics: A Time-Dependent Perspective*, University Science Books, Sausalito, CA USA, 2007. *[I.0]*

S. Teufel, *Adiabatic Perturbation Theory in Quantum Dynamics*, Lecture Notes in Mathematics, 1821. Springer-Verlag, Berlin, 2003. *[II.2]*

B. Thaller, *Visual Quantum Mechanics*, Springer, New York, 2000. *[I.0], [III.1]*

H.F. Trotter, *On the product of semigroups of operators*, Proc. Amer. Math. Soc. 10 (1959), 545–551. *[I.3]*

H. Wang & M. Thoss, *Multilayer formulation of the multiconfiguration time-dependent Hartree theory*, J. Chem. Phys. 119 (2003), 1289–1299. *[II.3]*

H. Yserentant, *A new class of particle methods*, Numer. Math. 76 (1997), 87–109. *[II.5]*

J. Zanghellini, M. Kitzler, C. Fabian, T. Brabec & A. Scrinzi, *An MCTDHF approach to multielectron dynamics in laser fields*, Laser Phys. 13 (2003), 1064–1068. *[II.3]*