LANGEVIN MOLECULAR DYNAMICS DERIVED FROM EHRENFEST DYNAMICS

ANDERS SZEPESSY

ABSTRACT. Stochastic Langevin molecular dynamics for nuclei is derived from quantum classical molecular dynamics, also called Ehrenfest dynamics, at positive temperature, assuming that the initial data for the electrons is stochastically perturbed from the ground state, with a large spectral gap, and that the ratio, $M$, of nuclei and electron mass tends to infinity. The Ehrenfest dynamics is approximated by the Langevin dynamics with accuracy $o(M^{-1/2})$ on bounded time intervals, which makes the $\mathcal{O}(M^{-1/2})$ small friction and fluctuation terms visible. The initial electron probability distribution is a Gibbs density derived from the Liouville equilibrium solution generated by the nuclei acting as a heat bath for the electrons in the Ehrenfest Hamiltonian system. The diffusion and friction coefficients in the Langevin equation satisfy the Einstein’s fluctuation-dissipation relation. The dissipative friction mechanism comes from the evolution of the electron ground state, due to slow dynamics of the nuclei. The stochastic perturbation of the ground state can also generate a temperature dependent contribution to the drift, depending on the spectral gap of the electron eigenvalues.

CONTENTS

1. Introduction to Ab Initio Molecular Dynamics and the Main Result 2
   1.1. Perturbation of the Born-Oppenheimer Approximation 5
   1.2. Stochastic Electron Initial Data 6
   1.3. The Main Result 7
2. The Force on Nuclei 8
3. Approximation Error of Observables 9
   3.1. The Dynamics 10
   3.2. An Error Representation 11
4. Estimates of the Friction, Fluctuation and Coupling 12
   4.1. The Friction Term 12
   4.2. The Fluctuation-Dissipation Property 13
   4.3. Couplings 14
   4.4. The Three Error Terms 16
4.5. The Born-Oppenheimer approximation 17
5. The Constrained Stochastic Initial Data 18
   5.1. The Constraint $\|\Psi\| = 1$ 18
   5.2. Entropy and the Standard Canonical Density Distribution 19
References 19

2000 Mathematics Subject Classification. Primary: 82C31; Secondary: 60H10, 82C10.

Key words and phrases. Langevin equation, Ehrenfest dynamics, quantum classical molecular dynamics, heat bath, fluctuation-dissipation, Born-Oppenheimer, Brownian particle, ab initio molecular dynamics.
1. Introduction to Ab Initio Molecular Dynamics and the Main Result

One method to simulate molecular motion is to use quantum classical molecular dynamics (QCMD), also called Ehrenfest dynamics, with the Hamiltonian

$$H(X) = -\frac{1}{2} \sum_{j=1}^{J} \Delta x_j + \sum_{1 \leq k < j \leq J} \frac{1}{|x_k - x_j|} - \sum_{n=1}^{N} \sum_{j=1}^{J} \frac{Z_n}{|x_j - X_n|} + \sum_{1 \leq n < m \leq N} \frac{Z_n Z_m}{|X_n - X_m|}$$

and the energy

$$\langle \Psi, H(X) \Psi \rangle := \int_{\mathbb{R}^{3J}} \Psi(x_1, \ldots, x_J)^* H(X) \Psi(x_1, \ldots, x_J) dx_1 \ldots dx_J$$

where the nuclear positions $X_n : [0, \infty) \to \mathbb{R}^3$, $n = 1, \ldots, N$ and the electron wave function $\Psi : [0, \infty) \times \mathbb{R}^{3J} \to \mathbb{C}$ solve the Ehrenfest dynamics system

$$M \dot{X}_n^t = -\langle \Psi^t, \partial_{X_n} H(X^t) \Psi^t \rangle,$$

$$i \partial_t \Psi = H(X^t) \Psi,$$

see [24, 27]. The Hamiltonian is composed of the kinetic energy of the electrons, the electron-electron repulsion, the electron-nuclei attraction, and the repulsion of nuclei (with charge $Z_n$), in the Hartree atomic units where the electron mass, electron charge, reduced Planck constant, and the Coulomb force constant $(4\pi\epsilon_0)^{-1}$ all are one. The mass of the nuclei, which are much greater than one (electron mass), are the diagonal elements in the diagonal matrix $M$. The Ehrenfest dynamics (1.2,1.3) is a Hamiltonian system with the Hamiltonian

$$M^{-1}|p|^2 + \frac{1}{2} \langle \Psi, H \Psi \rangle =: H_E$$

in the real variable $\langle (X, \dot{X}^r), (p, \dot{p}^i) \rangle = \langle (r_E, p_E) \rangle$, where $2p = M \dot{X}$ and $\Psi = \Psi^r + i \Psi^i$. Therefore, the Ehrenfest dynamics conserves this energy $M^{-1}|p|^2 + 2^{-1} \langle \Psi, H \Psi \rangle$. The superscript $t$ denotes the time variable $X^t := X(t)$, in the fast electron dynamics time scale.

The Ehrenfest dynamics can be derived from the time-independent Schrödinger equation

$$\left( H - \frac{1}{2M} \sum_{n=1}^{N} \Delta_{X_n} \right) \Phi = E \Phi,$$

see [24], and from the time-dependent Schrödinger equation, cf. [2, 23]. The wave function depends the electron and nuclei coordinates and on discrete spin states

$$\Phi(x^1, \sigma_1, \ldots, x^J, \sigma_J, X_1, \Sigma_1, \ldots, X_N, \Sigma_N),$$

which effect the solutions space: each electron spin $\sigma_j$ can take two different values and each nucleus can be in a finite set of spin states $\Sigma_n$, the Pauli exclusion principle restricts the solutions space to wave functions satisfying the antisymmetry/symmetry

$$\Phi(\ldots, x_j, \sigma_j, \ldots, x_k, \sigma_k, \ldots) = -\Phi(\ldots, x_k, \sigma_k, \ldots, x_j, \sigma_j, \ldots)$$

for any $1 \leq j, k \leq J$

and for any pair of nuclei $n$ and $m$, with $A$ nucleons and the same number of protons and neutrons,

$$\Phi(\ldots, X^m, \Sigma_m, \ldots, X^n, \Sigma_n, \ldots) = (-1)^A \Phi(\ldots, X^n, \Sigma_n, \ldots, X^m, \Sigma_m, \ldots)$$

cf. [4]. We simplify the notation by writing $\Phi(x, X)$ instead of the more complete (1.5), since the Hamiltonian $H$ does not depend on the spin of each particle. We use periodicity in the $X$-domain and set $T := \mathbb{R}/(L\mathbb{Z})$.
for a cell size $L$. The Ehrenfest solution is shown, in [24], to approximate the full Schrödinger solution with error bounded by $O(M^{-1})$, for the micro canonical ensemble and for equilibrium observables in the canonical ensemble, in the case without caustics and with electron eigenvalues having a large spectral gap. The work [26] describes with examples some of the weaknesses and strengths of the Ehrenfest approximation.

The Ehrenfest dynamics can be further coarse-grained by assuming the electron wave function is in its ground state – this so called Born-Oppenheimer approximation leads to $O(M^{-1})$ accurate approximation of observables of the time-independent Schrödinger equation, for the micro canonical ensemble in the case of a spectral gap and no caustics, see [24]. The work [24] also shows that Schrödinger equilibrium observables in the canonical ensemble, not including correlation in time can be $O(M^{-1})$ accurately approximated by Langevin dynamics with the Born-Oppenheimer drift and any friction coefficient proportional to the diffusion coefficient.

The purpose here is to study observables including time-correlation in a bounded time interval to precisely determine the friction/diffusion coefficient for the Langevin dynamics: Theorem 1.1 shows that Ehrenfest dynamics is approximated by the Langevin dynamics for observables including time-correlation, in the canonical ensemble when the initial data for the electrons is a temperature dependent stochastic perturbation of its ground state; the accuracy is $o(M^{-1/2})$ on bounded time intervals (in the slow time scale of nuclei motion), which makes the $O(M^{-1/2})$ small friction and fluctuation terms visible. The stochastic perturbation of the ground state can also generate a temperature dependent correction to the drift, depending on the spectral gap of the electron eigenvalues. The next section explains the stochastic setting and the Born-Oppenheimer approximation. The main idea in the work is the formulation of a classical Gibbs equilibrium distribution of eigenstates, motivated by nuclei acting as heat bath for the electrons in the quantum Ehrenfest and Schrödinger Hamiltonian systems. The error analysis uses the residual in the Kolmogorov equation of the Langevin stochastics, evaluated along the Ehrenfest dynamics on sufficiently large time steps, to have the fluctuations behaving as Brownian motion increments; here a critical idea is to use that the Ehrenfest solutions are the eigenfunctions of the logarithm of the solution operator. The results are based on the Hamiltonian $-2^{-1} \sum_j \Delta x_j + H_I$ with any potential $H_I$ that is smooth in $X$, e.g. a regularized version of the Coulomb potential.

The particles with coordinates $x_j$ in (1.2)-(1.3), can also be interpreted as a heat bath of lighter particles consisting of both nuclei and electrons, i.e. not necessarily only of electrons, so that the Langevin equation (1.15) also describes approximately the dynamics $X$ of heavy so called Brownian particles. This section ends with some background on deriving the Langevin equation for heavy particles in a heat bath. Theorem 1.1 is relevant for the central problem in statistical mechanics to show that Hamiltonian dynamics of heavy particles, coupled to a heat bath of many lighter particles with random initial data, can be approximately described by Langevin’s equation, as motivated by the pioneering work [6, 20] and continued with more precise heat bath models, based on harmonic interactions, in [12, 11, 28]. More recently these models of a heavy, so called Brownian, particle coupled to a heat bath are also used for numerical analysis studies related to coarse-graining in molecular dynamics and weak convergence analysis [25, 19, 15], for strong convergence analysis [14], and for computational studies on nonlinear heat bath models [8, 18]. Langevin’s equation has also been derived from a heavy particle colliding with an ideal gas heat bath, where the initial light particle positions are modeled by a Poisson point process and initial particle velocities are independent Maxwell distributed; the heavy particle collides elastically with the ideal gas particles and moves uniformly in between, see [9, 8].

The main inspiration of this paper is [28], where Zwanzig derives a generalized Langevin equation for a (Brownian) particle coupled to a heat bath particle system through a harmonic interaction potential with a Debye heat bath frequency distribution and an ad hoc coupling; similar assumptions on deterministic or stochastic frequency distributions and coupling are used in the above cited references. This work extends the ideas in [28] by using the ab initio Ehrenfest dynamics (1.2) for nuclei and (1.3) for electrons (or (1.2) for heavy nuclei coupled to a heat bath of lighter nuclei and all electrons modeled by a Schrödinger equation (1.3)). Other differences are that:
The residual of the Langevin Kolmogorov equation along the Ehrenfest dynamics is used to derive an error estimate, instead of using an explicit solution, and

the slow nuclei dynamics compared to the fast electron dynamics is exploited, in integration by parts of highly oscillatory integrands based on the Ehrenfest dynamics solution operator, to find a proper Langevin equation, avoiding the integral coupling in the generalized Langevin equation.

The two central ideas in deriving Langevin dynamics from coupling to a heat bath – to find the friction mechanism in the heavy particle coupling to the dynamics of the lighter particles and to find the diffusion from fluctuations in initial data of the light particles – were already the basis in [12, 11] and [28].

1.1. Perturbation of the Born-Oppenheimer Approximation. At low temperature we expect that the fast electron dynamics, compared to the slower nuclei in the Ehrenfest dynamics, yields an electron wave solution that is almost in its ground state \( \Psi_0 \), which solves the electron eigenvalue problem

\[
H \Psi_0 = \lambda_0 \Psi_0,
\]

and is normalized \((\Psi_0, \Psi_0) = 1\); here \( \lambda_0 = \lambda_0(X^t) \) is the smallest eigenvalue of \( H = H(X^t) \) in \( L^2(dx) \). The function

\[
\hat{\Psi} := \exp \left( -i \int_0^t \lambda_0(X^s) ds \right) \Psi_0(t, x; X^t)
\]
satisfies

\[
i \partial_t \hat{\Psi} - H \hat{\Psi} = i e^{-i \int_0^t \lambda_0(X^s) ds} \dot{\Psi}_0(t, x; X^t),
\]

so that if the nuclei do not move, the wave function \( \hat{\Psi} \) solves the time dependent Schrödinger equation; and if they move slowly, i.e. the \( L^2 \)-norm \( \|\Psi_0\| \) is small, the function \( \hat{\Psi} \) is an approximate solution to the Ehrenfest dynamics. The approximation \((X^t, \hat{\Psi})\), with \( \hat{\Psi} \) replacing \( \Psi \) in (1.2), is called Born-Oppenheimer molecular dynamics [23, 14] and it approximates observables of the time-independent Schrödinger equation for the electron-nuclei system with accuracy \( O(M^{-1}) \), when there is a spectral gap and no caustics, see [24]. The aim of this work is to study molecular dynamics when the electron states \( \hat{\Psi} \) are randomly perturbed from the ground state, with a Gibbs distribution at positive temperature. To study the perturbation of the ground state we first make the transformation

\[
\hat{\Psi}(t, x) = e^{-i \int_0^t \lambda_0(X^s) ds} \Psi(t, x)
\]

which implies that \( \Psi \) solves the Schrödinger equation

\[
i \partial_t \Psi^t = \left( H(X^t) - \lambda_0(X^t) \right) \Psi^t,
\]

with the translated Hamiltonian \( \tilde{H} \), and then we make the Ansatz

\[
\Psi = \Psi_0 + \tilde{\Psi}
\]

and expect \( \tilde{\Psi} : [0, \infty) \times \mathbb{R}^3J \to \mathbb{C} \) to be small. The Ansatz implies that \( \tilde{\Psi} \) solves

\[
i \partial_t \tilde{\Psi} = \tilde{H}(X^t) \tilde{\Psi} - i \partial_t \Psi_0,
\]

which yields the solution representation

\[
\tilde{\Psi}^t = S_{t,0} \tilde{\Psi}_0 - i \int_0^t S_{t,s} \tilde{\Psi}_{s}^s ds,
\]

with the solution operator \( S \) defined by

\[
S_{t,s} \varphi^s := \varphi^t
\]

for a solution

\[
i \partial_t \varphi^t = \tilde{H} \varphi^t \quad t > s.
\]
The first term in the representation depends only on the initial data and the second term depends only on the residual $\dot{\Psi}_0$. This splitting, inserted into the equation $\mathbb{L}_{\tau}^2$ for the nuclei, eliminates formally the electrons and generates fluctuations, from stochastic initial data $\Psi^0$, and friction, through the coupling to $\dot{\Psi}_0 = \partial_X \Psi_0 \dot{X}$.

1.2. Stochastic Electron Initial Data. The next step, to determine the stochastic initial data for $\Psi$, requires some additional modeling. Inspired by the study of a classical heat bath of harmonic oscillators in [28], we will sample the initial data for $\Psi$ randomly from a probability density given by an equilibrium solution $f$ (i.e., $\partial_t f = 0$) of the Liouville equation $\partial_t f + \partial_{p_E} H_E \partial_{r_E} f - \partial_{r_E} H_E \partial_{p_E} f = 0$, to the Ehrenfest dynamics $\mathbb{L}^2$ and $\mathbb{L}_f$. There are many such equilibrium solutions but there is only one where the momenta $p_j$ are independent and that is the standard canonical ensemble for Hamiltonian dynamics, i.e. $f(r_E, p_E) = c \exp(-H_E/T)$, where the positive parameter $T$ is the temperature, in units of the Boltzmann constant, and $c$ is a normalization constant. There are other equilibrium solutions, e.g. $f = h(H_E)$ for differentiable functions $h$. To find a unique equilibrium solution to sample $\Psi$ from, we may think of the nuclei as the heat bath for the electrons. The equilibrium distribution of the nuclei seems simpler to understand than the equilibrium of electrons: in a statistical mechanics view, the probability of finding the heat bath with the energy $H_E := M^{-1} |p|^2 + 2^{-1} \langle \Psi, H(X) \Psi \rangle$ (for given $\Psi$) is proportional to the Gibbs-Boltzmann factor $\exp(-H_E/T)$, cf. [10], [17]; an alternative motivation, based on the conclusion of this work (in a somewhat circular argument), is that the nuclei can be approximated by classical Langevin dynamics which has its unique invariant density $\exp(-M^{-1} |p|^2/2 - \lambda_0(X))/T$ as the marginal distribution of $\exp(-H_E/T)$ when integrating over all the electron states $\Psi$. Since the energy is conserved for the Ehrenfest dynamics – now viewed with the electrons as the primary systems coupled to the heat bath of nuclei – the probability of finding the electrons in a certain configuration $\Psi$ is the same as finding the nuclei in a state with energy $H_E$, which is proportional to $\exp(-H_E/T)$. This conclusion, that the probability to have an electron wave function $\Psi$ is proportional to $\exp(-H_E/T) d\Psi^t d\Psi^t$ is our motivation to sample the initial data for $\Psi$ from the conditioned density generated by $\exp(-\langle \Psi, H(X) \Psi \rangle/(2T))$: since we seek data for the electrons, we use the probability distribution for $\Psi$ conditioned on $(X, p)$.

We compare in Section 5.2 our model of initial data with a more standard model of initial data, having given probabilities to be in mixed states. It turns out that Einstein’s fluctuation-dissipation result does not hold for the canonical distribution, which is a motivation for our focus on the initial data sampled from the Liouville equilibrium solution above, where the Einstein fluctuation-dissipation results holds. To sample from such an equilibrium density is standard in classical Hamiltonian statistical mechanics but it seems non standard for Ehrenfest quantum dynamics. The Gibbs density for the Ehrenfest dynamics approximates $O(M^{-1})$ accurately the Gibbs density for the Hamiltonian dynamics of the time-independent Schrödinger equation derived in [23], in the case no caustics are present.

1.2.1. Slow Nuclear Dynamics. The large mass, $M \gg 1$, in the dynamics of the nuclei

\[
\begin{align*}
\frac{dX^t}{dt} &= u^t, \\
M \frac{du^t}{dt} &= -\langle \Psi^t, \partial_X H(\Psi^t) \rangle
\end{align*}
\]

implies that position increments $|X^t - X^0| = O(M^{-1/2})$ are asymptotically negligible for bounded time intervals $t = O(1)$. It is therefore necessary to follow the nuclei for long time $t > M^{1/2}$ to see non trivial dynamics and then it is convenient to change the time scale by introducing $t = \sqrt{M} \tau =: \dot{\tau}$, $X^\tau$ and $v^\tau := \sqrt{M} \dot{w}^\tau$ to obtain

\[
\begin{align*}
\frac{dX^\tau}{d\tau} &= v^\tau, \\
M \frac{dv^\tau}{d\tau} &= -\langle \Psi^\tau, \partial_X H(X^\tau) \Psi^\tau \rangle,
\end{align*}
\]

which means that we want to understand electron wave solutions $\tilde{\Psi}^\tau$ to $\mathbb{L}_3$ up to fast time of order $t = \dot{\tau} \simeq M^{1/2}$.
To describe the stochastic initial data, we use a spectral decomposition defined in the next section.

1.2.2. The Initial Data and the Energy Cut-off. Reference [24] shows that observables of the Ehrenfest dynamics approximates observables of the time-independent Schrödinger equation with error $O(M^{-1})$. The solution of the time-independent Schrödinger equation can therefore be characterized by data for the Ehrenfest dynamics. We choose these stochastic initial data sampled from the Gibbs-Boltzmann equilibrium density as follows.

Introduce the normalized electron eigenvectors $\{\tilde{\psi}_j(X) : j = 0,1,2,\ldots, J\}$ of the Hamiltonian $\tilde{H}(X) : H^1(T^{3J}) \to H^{-1}(T^{3J})$ with its eigenvalues $\{\tilde{\lambda}_j(X) : j = 0,1,2,\ldots, J\}$ in increasing order satisfying $\tilde{H}(X)\tilde{\psi}_j(X) = \tilde{\lambda}_j(X)\tilde{\psi}_j(X)$, for given nuclei position $X$. We let the initial data for the initial electron wave function $\Psi^0$ in the Ehrenfest dynamics [18] be a stochastic perturbation (1.9) of the ground state with initial data

$$
\Psi^0 = \sum_{j=0}^{J} \gamma_j \tilde{\psi}_j(X^0)
$$

where the complex coefficients $\gamma_j = \gamma_j^\alpha + i\gamma_j^\beta$, for $j = 1,\ldots, J$, are all independent, with independent normally distributed real and imaginary parts of mean zero and variance $T/\tilde{\lambda}_j(X^0)$ and $\gamma_0 := (1 - \sum_{j=1}^{J} |\gamma_j|^2)^{1/2}$. We consider the case when there is a large spectral gap around the ground state energy – more precisely $\tilde{\lambda}_{\gamma_0}\psi_0(X^0)$. Although the perturbation $\sum_{n>0} \gamma_n \tilde{\psi}_n$ is small its expected energy

$$
E[\sum_{n,m>0}\gamma_n^*\gamma_m\langle \tilde{\psi}_n, \tilde{H}\tilde{\psi}_m \rangle] = E[\sum_{n>0} |\gamma_n|^2 \langle \tilde{\psi}_n, \tilde{H}\tilde{\psi}_n \rangle] = T\tilde{J}
$$

is large and we will see that the small perturbation in fact influences the dynamics to order $O(M^{-1/2})$. The cut-off state $\tilde{J}$ is chosen such that

$$
\tilde{J} \to \infty
$$

(slowly) as $M \to \infty$, to include all eigenstates in the limit. Section 5 presents motivations of the normal distribution from the canonical Gibbs-Boltzmann distribution.

1.3. The Main Result. Let $W^\tau$ denote the standard Brownian process (at time $\tau$) in $\mathbb{R}^{3N}$ with independent components. To simplify the notation we assume that all nuclei have the same mass $M \gg 1 = $ electron mass; this can easily be extended to varying nuclear masses much larger than the electron mass. We apply the notation $\psi(x,X) = O(M^{-\alpha})$ also for complex valued functions, meaning that $|\psi(x,X)| = O(M^{-\alpha})$ holds uniformly in $x$ and $X$.

We will use the following spectral gap condition for the electron eigenvalues

$$
\sum_{n\geq 1} \frac{\lambda_n^{-1}}{n} = o(1) \quad \text{and} \quad \sum_{n=0}^{\infty} j\langle \tilde{\psi}_n, \partial_X \tilde{H}\tilde{\psi}_n \rangle = o(1)
$$

We also assume that the electron eigenvalues do not cross, i.e. $|\lambda_n - \lambda_m| > c$ for a positive constant $c$ along solution paths and $n \neq m$. The conditions are motivated and used as follows:

The first condition will imply that the Ehrenfest wave function is a small perturbation of the ground state, the temperature is small compared to the large gap $\tilde{\lambda}_1$ and the integral kernel for the friction and diffusion is well approximated by a point mass in (1.2). Since the eigenvalues may grow slowly for a Schrödinger problem, the first condition typically requires $\tilde{J} \to \infty$ slowly, i.e. the number of exited states $\tilde{J}$ is small compared to the gaps: the reason to have $\tilde{J} \to \infty$ is to establish $\sum_{n=0}^{\infty} j\langle \tilde{\psi}_n, \partial_X \tilde{H}\tilde{\psi}_n \rangle = o(1)$ in (1.3), which has a high convergence rate, since $\partial_X \tilde{H}$ is smooth and $\tilde{\psi}_n$ becomes highly oscillatory, compatible with $\tilde{J}$ tending to infinity slowly.
The second condition means that the electron eigenvalues are almost parallel to the ground state eigenvalue, as a function of $X$, which implies that the ground state Born-Oppenheimer drift is a good approximation for all states. When this condition is relaxed in (1.18), the drift receives a temperature dependent correction. We have $\partial_X \lambda_n = \langle \tilde{\psi}_n, \partial_X \tilde{H} \tilde{\psi}_n \rangle$ which decays fast with growing eigenvalues, since $\partial_X \tilde{H}$ is smooth and $\tilde{\psi}_n$ become highly oscillatory as $n \to \infty$.

Let $K$ be a positive parameter; the stochastic Langevin dynamics
\[
d\tilde{X}^\tau = \tilde{p}^\tau d\tau \\
d\tilde{p}^\tau = -\partial_X \lambda_0(\tilde{X}^\tau) d\tau - M^{-1/2} K \tilde{p}^\tau d\tau + \sqrt{2T M^{-1/2}} dW^\tau
\]
has the unique invariant probability density
\[
e^{-\langle \rho \partial \rho/2 + \lambda_0(X) \rangle / T} d\rho dX \\
\int_{\mathbb{R}^N} e^{-(\rho \partial \rho/2 + \lambda_0(X)) / T} d\rho dX
\]
cf. [5]. The work [24] shows that Langevin dynamics approximates equilibrium observables of the Schrödinger equation and Ehrenfest dynamics with accuracy $o(M^{-1/2})$ provided (1.14) holds, in the case where the observable does not depend on time and it is enough to determine an integral with respect to the invariant measure. Here we study Langevin dynamics approximating Ehrenfest dynamics on a bounded time interval. This determines the otherwise undetermined friction and diffusion matrix $K$, and makes it possible to determine observables that depend on the correlation of the positions at different time in the canonical ensemble $o(M^{-1/2})$ accurately.

We use the notation $\tilde{\tau} = M^{1/2} \tau$ for the transformation from slow to fast time, the eigenvalue $\lambda_0$ denotes the ground state electron energy (1.6) of $H$ with normalized ground state $\Psi_0$ and $\tilde{\lambda}_j$ are the translated eigenvalues of $H - \lambda_0$ defined in (1.1). Define also the rank one dissipation and fluctuation matrix
\[
K(X) := 2\langle \partial_X \Psi_0(X), \partial_X \Psi_0(X) \rangle, \\
K^{1/2}(X) = \sqrt{2}\langle \partial_X \Psi_0^{-1}(\partial_X \Psi_0(X), \partial_X \Psi_0(X)) \rangle.
\]

**Theorem 1.1.** Assume that the spectral gap condition (1.14) holds, the temperature is bounded $T = O(1)$, and the electron initial data is given by (1.12), then the Itô Langevin dynamics
\[
(1.15) \\
\tilde{d} \tilde{X}^\tau = -\partial_X \lambda_0(\tilde{X}^\tau) d\tau - M^{-1/2} K(\tilde{X}^\tau) \tilde{X}^\tau d\tau + \sqrt{2T M^{-1/2}} K^{1/2}(\tilde{X}^\tau) dW^\tau,
\]
approximates Ehrenfest dynamics with accuracy
\[
(1.16) \\
\left| \mathbb{E}[g(X^T, p^T) - g(\tilde{X}^\tau, \tilde{X}^\tau)] \right| = o(M^{-1/2}),
\]
for any bounded function $g : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R}$, provided the Langevin value function
\[
(1.17) \\
u(X, p, \sigma) := \mathbb{E}[g(\tilde{X}^\tau, \tilde{X}^\tau) \mid \tilde{X}^\sigma = X, \tilde{X}^\sigma = p]
\]
has bounded derivatives of order one to four
\[
D = \partial_\tau, \partial_X, \partial_p, \ldots, \partial_{pppp}.
\]

The approximation result uses a non interacting particle, with given velocity equal to one and position coordinate $X_0 = \tau$, that acts as the time coordinate, so that e.g. transport coefficients as diffusion can be studied.

The small dissipation term $M^{-1/2} K \tilde{p}$ and small diffusion term are visible in the convergence rate $o(M^{-1/2})$. The approximation of Schrödinger observables in equilibrium, in [24], is based on long time and determines also the diffusion matrix to be equal to the dissipation matrix. That is, the Langevin dynamics (1.15) satisfies the *Einstein’s fluctuation-dissipation* result: the square of the diffusion coefficient is the dissipation coefficient times twice the temperature.
The combination of the equilibrium result from [24], that Langevin dynamics approximates the equilibrium observables of the Ehrenfest and Schrödinger equation, and the approximation of Ehrenfest correlation observables over bounded time intervals $O(1)$ here shows that this Langevin dynamics is an accurate approximation of Ehrenfest dynamics for both short and long time. We cannot from this conclude that Langevin dynamics is also a good approximation of the Schrödinger equation for short time correlations but conjecture this is true, considering the qualitative similarity of the Ehrenfest dynamics and the Schrödinger dynamics as formulated in [24], where $\hat{H}$ is essentially replaced by $\hat{H} - \sum_n(2M)^{-1}\Delta X_n$.

**Theorem 1.2.** If we instead of the gap condition (1.14) assume the weaker condition

$$
\sum_{n=1}^{\infty} \lambda_n^{-1} = o(1) \quad \text{and} \quad \sum_{n=1}^{\infty} |\partial X, \lambda_n| \lambda_n^{-1} = o(M^{-1/2}),
$$

$$
\sum_{n>0} (|\partial X, \lambda_n| \lambda_n^{-1})^2 = o(M^{-1/2}) \quad \text{and} \quad \sum_{n>0} |\partial X, \lambda_n| \lambda_n^{-1} = O(1),
$$

then the result in Theorem 1.1 holds, provided the drift $\partial_X \lambda_0$ is replaced by

$$
\partial_X (\lambda_0(X) + \frac{T}{2} \text{trace}^\perp \log \hat{H}(X)),
$$

where $\text{trace}^\perp$ is the trace in the orthogonal complement of $\Psi_0$.

2. The Force on Nuclei

It is convenient to split the force on nuclei in (1.2) into two parts, using the definition of $\Psi$ in (1.7),

$$
\langle \Psi, \partial_X \hat{H}(X)\Psi \rangle = \langle \Psi, \partial_X \hat{H}(X)\Psi \rangle \\
= \langle \Psi, \partial_X \lambda_0(X)\Psi \rangle + \langle \Psi, \partial_X \hat{H}(X) - \lambda_0(X)\Psi \rangle \\
= \partial_X \lambda_0(X) + \langle \Psi, \partial_X \hat{H}(X)\Psi \rangle.
$$

With the Ansatz $\Psi = \Psi_0 + \tilde{\Psi}$, the second term in the nuclear force becomes

$$
\langle \Psi, \partial_X \hat{H}(X)\Psi \rangle = \langle \Psi_0, \partial_X \hat{H}(X)\Psi_0 \rangle \\
+ \langle \tilde{\Psi}, \partial_X \hat{H}(X)\Psi_0 \rangle + \langle \Psi_0, \partial_X \hat{H}(X)\tilde{\Psi} \rangle \\
+ \langle \tilde{\Psi}, \partial_X \hat{H}(X)\tilde{\Psi} \rangle.
$$

Use $\hat{H}\Psi_0 = 0$ and consequently

$$
\partial_X \hat{H}\Psi_0 + \hat{H}\partial_X \Psi_0 = 0
$$

to obtain for the first term

$$
\langle \Psi_0, \partial_X \hat{H}(X)\Psi_0 \rangle = -\langle \Psi_0, \hat{H}(X)\partial_X \Psi_0 \rangle \\
= -\langle \hat{H}\Psi_0, \partial_X \Psi_0 \rangle = 0.
$$

Let $\Re$ denote the real part. The second terms are

$$
\langle \tilde{\Psi}, \partial_X \hat{H}(X)\Psi_0 \rangle + \langle \Psi_0, \partial_X \hat{H}(X)\tilde{\Psi} \rangle = 2\Re\langle \tilde{\Psi}, \partial_X \hat{H}(X)\Psi_0 \rangle \\
= -2\Re\langle \tilde{\Psi}, \hat{H}(X)\partial_X \Psi_0 \rangle
$$

which, by the dynamics (1.9), generates a fluctuation term

$$
2\Re\langle S_t, 0, \hat{H}(X^t)\partial_X \Psi_0 \rangle
$$

and a friction term

$$
-2\Re\int_0^t \langle S_{t-s} \partial_X \Psi_0, \dot{X}^s, \hat{H}(X^t)\partial_X \Psi_0 \rangle ds.
$$

In Section 4.2 we show that the third term, $\langle \tilde{\Psi}, \partial_X \hat{H}\tilde{\Psi} \rangle$, is negligible small for spectral gap condition (1.14) and yields the correction $T \text{trace}^\perp \hat{H}/2$ for the weaker gap condition (1.18). The asymptotic estimates of forces are explained more precisely in Sections 4.1 and 4.2.
Remark 2.1 (Quantum initial data). Assume, instead of (1.12), that $\Psi^0$ is a pure electron eigenstate $e^{i\gamma_0}\tilde{\psi}_j$ with Gibbs-Boltzmann probability

$$q_j := e^{-\tilde{\lambda}_j/T} \left( \sum \lambda e^{-\lambda/T} \right)^{-1}$$

and independent random phase shifts $\alpha_j$ uniformly distributed on $[0, 2\pi]$ for $j = 0, \ldots, J$, and write

$$\Psi^0 = \sum_{j \geq 0} \tilde{\gamma}_j \tilde{\psi}_j$$

which has the covariance $\mathbb{E}[(\tilde{\phi}_j)^* \tilde{\gamma}_k] = q_j \delta_{jk}$. Let $\Psi_0(X) = \gamma_0 \tilde{\psi}_0(X)$. Then the fluctuations are very different from the case in Theorem 1.1, since $E\geq0$ for gap condition (1.14), and with $\partial X\lambda$ dynamics by (3.1).

The Ehrenfest dynamics can be written in the slow time scale

$$\dot{X}^\tau = p^\tau$$

$$\dot{p}^\tau = -\partial X\lambda_0 - \langle \tilde{\phi}^\tau, \partial X\tilde{H}(X^\tau)\tilde{\phi}^\tau \rangle$$

$$\dot{\tilde{\phi}}^\tau = \tilde{H}(X^\tau)\phi^\tau$$

with the Hamiltonian

$$\frac{|p|^2}{2} + \langle \phi^\tau, \tilde{H}\phi^\tau \rangle + \lambda_0 = E,$$

using the coordinates $(X, \psi^\tau; p, \psi)$ and $\psi := \psi^\tau + i\psi^\lambda = (2/M)^{1/2}\phi$. We shall approximate the Ehrenfest dynamics by $(X_L, p_L)$ defined from the Ito-Langevin dynamics

$$\dot{X}_L = p_L$$

$$\dot{p}_L = -\partial X\lambda_0(X_L) - M^{-1/2}K(X_L)p_L + (2TM^{-1/2})^{1/2}M^{-1/2}K^{1/2}(X_L)\dot{W},$$

for gap condition (1.14), and with $\partial X(\lambda_0 + \text{trace}^{-1}\tilde{H})$ replacing the drift $\partial X\lambda_0$ for gap condition (1.15).

To simplify the analysis of the coupling between $(X, p)$ and the thermal fluctuations, introduce the electron wave functions $\{\tilde{\psi}_n\}$ resembling individual states

$$\frac{i}{M^{1/2}} \tilde{\psi}_n^\tau = \tilde{H}(X^\tau)\tilde{\psi}_n^\tau, \quad \tilde{\psi}_n^0 = \tilde{\psi}_n(X^0),$$

which implies

$$\phi^\tau = \sum_{n=0}^J \gamma_n \tilde{\psi}_n,$$

where $\gamma_n$ are independent normal distributed with mean zero and variance $T/\tilde{\lambda}_n(X^0)$. The Schrödinger dynamics (3.2) shows that $\{\tilde{\psi}_n^\tau | n = 0, 1, \ldots\}$ forms an orthogonal set

$$\frac{d}{d\tau} \langle \tilde{\psi}_n^\tau, \tilde{\psi}_m^\tau \rangle = \langle -iM^{1/2}\tilde{H}\tilde{\psi}_n^\tau, \tilde{\psi}_m^\tau \rangle + \langle \tilde{\psi}_n^\tau, -iM^{1/2}\tilde{H}\tilde{\psi}_m^\tau \rangle$$

$$= \langle \tilde{\psi}_n^\tau, iM^{1/2}\tilde{H}\tilde{\psi}_m^\tau \rangle + \langle \tilde{\psi}_n^\tau, -iM^{1/2}\tilde{H}\tilde{\psi}_m^\tau \rangle = 0,$$

since the initial data $\{\tilde{\psi}_n^0 | n = 0, 1, \ldots\}$ is orthogonal.
3.2. An Error Representation. Define for the given observable \( g : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R} \) and the Langevin dynamics \((X_L^\tau, p_L^\tau), \) in (3.1), the value function
\[
u(y, \tau) := E[g(X_L^\tau, p_L^\tau) | (X_L^\tau, p_L^\tau) = y],
\]
which solves the Kolmogorov backward equation
\[
\begin{aligned}
\partial_\tau u + p \circ \partial X u - (\partial_X \lambda_0(X) + M^{-1/2}K(X)p) \circ \partial_p u \\
+ \sum_j M^{-1/2}K(X)\partial_{p_j}u = 0 \quad \tau < Y
\end{aligned}
\]
(3.4)
where \( v \circ w \) is the usual Euclidean scalar product in \( \mathbb{R}^{3N} \). The goal is to analyze the error \( E[g(X_L^\tau, p_L^\tau) - g(X_L^\tau, p_L^\tau)] \) of the heavy nuclei particles in the Ehrenfest dynamics approximated by the Langevin dynamics. This error can be written as the residual of the Langevin Kolmogorov solution (3.4) along the Ehrenfest dynamics \((X^\tau, p^\tau)\)
\[
E[g(X_L^\tau, p_L^\tau) - g(X_L^\tau, p_L^\tau)] = E[u(X_L^\tau, p_L^\tau, Y) - u(X_L^\tau, p_L^\tau, 0)]
\]
\[
= E[u(X_L^\tau, p_L^\tau, Y) - u(X_0^\tau, p_0^\tau, 0)] - E[u(X_0^\tau, p_0^\tau, 0) - u(X_0^\tau, p_0^\tau, 0)].
\]
We assume that initial data for the Langevin dynamics is the same as for the Ehrenfest dynamics. Therefore the last term is zero. The first term in the right hand side will be written as an integral over time and the assumption (1.17) makes the integral bounded.

We will see that the fluctuations in Ehrenfest dynamics behaves approximately as the Wiener process fluctuations in the Langevin dynamics only on sufficiently large time intervals \( \Delta \tau \). Therefore we perform the analysis on such large time steps and introduce for time steps \( \Delta \tau \) the notation \( X(j\Delta \tau) := X_j, dX^\tau/d\tau|_{\tau=j\Delta \tau} = p(j\Delta \tau) =: p^j \) and \( \tau_j := j\Delta \tau \). Telescoping cancelation implies
\[
E[u(X_0^\tau, p_0^\tau, Y) - u(X_0^\tau, p_0^\tau, 0)] = \sum_j E\left[u(X_j+1, p^{j+1}, \tau_{j+1}) - u(X_j, p^j, \tau_j)\right]
\]
and Taylor expansion, with the position and velocity increments
\[
\begin{aligned}
\Delta X_j &:= X^{j+1} - X_j = \int_{\tau_j}^{\tau_{j+1}} p^\sigma \, d\sigma, \\
\Delta p^j &:= p^{j+1} - p^j = -\int_{\tau_j}^{\tau_{j+1}} \langle \phi^\sigma, \partial_X H(X^\sigma)\phi^\sigma \rangle \, d\sigma.
\end{aligned}
\]
shows that
\[
E[u(X_0^\tau, p_0^\tau, Y) - u(X_0^\tau, p_0^\tau, 0)]
= \sum_j E\left[\partial_\tau u(X_j, p^j, \tau_{j+1}) \Delta \tau + \partial_x u(X_j, p^j, \tau_{j+1}) \Delta X_j + \partial_p u(X_j, p^j, \tau_{j+1}) \Delta p + \frac{1}{2} \partial_{pp} u(X_j, p^j, \tau_{j+1})(\Delta p)^2\right] + \text{error terms}_1,
\]
for some point \((\tilde{X}^j, \tilde{p}^j)\) on the line between \((X_j, p^j)\) and \((X^{j+1}, p^{j+1})\); here the error terms are
\[
\text{error terms}_1 := \sum_j E\left[\partial_\tau u(X_j, p^j, \tau_{j+1}) \Delta \tau^2/2 + \partial_x u(X_j, p^j, \tau_{j+1}) \Delta X_j^2/2 + \partial_p u(X_j, p^j, \tau_{j+1}) \Delta X \Delta p\right]
\]
for some \( \tilde{\tau}_j \) between \( \tau_j \) and \( \tau_{j+1} \), using the notation
\[
\begin{aligned}
w_{\tau_j} u &= \sum_j w_j \partial_{\tau_j} u, \\
\partial_{xy} u \Delta X \Delta Y &= \sum_{jk} \partial_{xy} u \Delta X_j \Delta Y_k, \\
\partial_{yy} u \mathbb{I} &= \sum_j \partial_{y_j y_j} u.
\end{aligned}
\]
Apply the Kolmogorov equation to eliminate the $\partial_x u$ term

$$
\partial_x u(X^j, p^j, \tau_{j+1}) = -\partial_x u(X^j, p^j, \tau_{j+1}) p^j \\
-\partial_p u(X^j, p^j, \tau_{j+1}) \left( -\partial_x \lambda_0(X^j) - M^{-1/2} K(X^j) p^j \right) \\
-\frac{T}{M^{1/2}} \partial_{pp} u(X^j, p^j, \tau_{j+1}) K(X^j)
$$

in the error representation to obtain

$$
\begin{align*}
\mathbb{E}[u(X^j, p^j, \tau) - u(X^0, p^0, 0)] &= \sum_j \mathbb{E}\left[ \partial_x u(X^j, p^j, \tau_{j+1})(\Delta X - p^j \Delta \tau) \right] \\
&+ \partial_p u(X^j, p^j, \tau_{j+1}) \left( \Delta p - \left( -\partial_x \lambda_0(X^j) - M^{-1/2} K(X^j) p^j \right) \Delta \tau \right) \\
&+ \frac{1}{2} \left( \partial_{pp} u(X^j, p^j, \tau_{j+1})(\Delta p)^2 - \partial_{pp} u(X^j, p^j, \tau_{j+1}) T M^{1/2} K(X^j) \Delta \tau \right) \\
&+ \text{error terms}_i.
\end{align*}
$$

These expected values are determined in Sections 4.1–4.3. Section 4.4 estimates the three error terms

$$
\mathbb{E}[\partial_{\tau_1} u(\Delta \tau)^2], \quad \mathbb{E}[\partial_{x x} u(\Delta X)^2], \quad \mathbb{E}[\partial_{x v} u \Delta X \Delta p].
$$

We have

$$
\Delta p_L := - \int_\sigma^\tau \left( \partial_x \lambda_0(X^s_L) + M^{-1/2} K(X^s_L) p^s_L \right) ds + \int_\sigma^\tau \left( M^{-1/2} T K(X^s_L) \right)^{1/2} dW^s
$$

and

$$
\Delta p = - \int_\sigma^\tau \left( \partial_x \lambda_0(X^s) + \sum_{n,m} \gamma_n \gamma_m^* \langle \tilde{\psi}^n_s, \partial_x \hat{H} \tilde{\psi}_m^s \rangle \right) ds \\
= - \int_\sigma^\tau \left( \partial_x \lambda_0(X^s) + |\gamma_0|^2 \langle \tilde{\psi}_0, \partial_x \hat{H} \tilde{\psi}_0 \rangle + 2 \sum_{n \neq 0} R(\gamma_n \tilde{\psi}_n, \partial_x \hat{H} \tilde{\psi}_0) \right) ds \\
- \int_\sigma^\tau \left( \sum_{n \neq 0} |\gamma_n|^2 \langle \tilde{\psi}_n, \partial_x \hat{H} \tilde{\psi}_n \rangle + \sum_{n \neq 0, m \neq 0} \gamma_n \gamma_m^* \langle \tilde{\psi}_n, \partial_x \hat{H} \tilde{\psi}_m \rangle \right) ds.
$$

In the next section we show that the last integral with the quadratic terms is negligible small, due to our assumption $\sum_{n > 0} \tilde{\lambda}_n^{-1} = o(M^{-1/2})$ in condition (1.14), and yields the correction of the drift for condition (1.15).

Introduce the orthogonal decomposition

$$
\tilde{\psi}_n = \tilde{\psi}_n \oplus \psi_n^l,
$$

based on the electron eigenvectors $\tilde{\psi}_n$ satisfying $\hat{H} \tilde{\psi}_n = \tilde{\lambda}_n \tilde{\psi}_n$. The Born-Oppenheimer approximation (4.14) shows that

$$
\psi_n^l(t) = O(M^{-1/2})
$$

and since $\hat{H} \tilde{\psi}_0 = 0$ we have by (4.14) the following representation in the slow time scale

$$
\psi_n^l(\tau) = \tilde{S}_{\tau,0} \psi_n^l(0) - i \int_0^\tau \tilde{S}_{\tau,\sigma} (\tilde{\psi}_0^l)^2 d\sigma,
$$

where $\tilde{\psi}_0 = \langle \tilde{\psi}_0, \psi_0 \rangle \psi_0 \oplus \tilde{\psi}_0^l$ is the orthogonal decomposition and $\tilde{S}$ is the solution operator $\tilde{S}_{\tau,\sigma} \tilde{\psi}_0(\sigma) = \tilde{\psi}_0(\tau)$ in the slow time scale. The ground state relation $\hat{H} \tilde{\psi}_0 = 0$ implies $\partial_x \hat{H} \tilde{\psi}_0 + \hat{H} \partial_x \tilde{\psi}_0 = 0$, which combined
with \( \psi_0^\perp = O(M^{-1/2}) \) yield the main friction term
\[
(3.10) \quad \langle \tilde{\psi}_0, \partial_X \tilde{H} \tilde{\psi}_0 \rangle = \langle \tilde{\psi}_0, \partial_X \tilde{H} \psi_0 \rangle \tag{3.10}
\]
\[
= 2\Re(\psi_0^\perp, \partial_X \tilde{H} \psi_0) + o(M^{-1/2})
\]
\[
= -2\Re(\psi_0^\perp, \tilde{H} \partial_X \psi_0) + o(M^{-1/2})
\]
\[
= -2\Re(-i \int_0^1 \tilde{S}_{l,s} \langle \psi_0, \nabla_x \tilde{H} \psi_0 \rangle - 2\Re(\tilde{S}_{l,0} \psi_0^\perp(0), \tilde{H} \partial_X \psi_0^\perp) = o(M^{-1/2}).
\]

4. Estimates of the Friction, Fluctuation and Coupling

In the error estimate there are two main terms:

\[
\text{(4.1)} \quad \text{a time discretization error } O(\Delta \tau), \text{ and}
\]
\[
\text{a coupling error } \lambda^{-1}_1 O((M \Delta \tau)^{-1}) + o(M^{-1/2}).
\]

These two errors combine to the bound \( O(\lambda^{-1/2}_1 M^{-1/2}) + o(M^{-1/2}) \), for the optimal choice \( \Delta \tau = \tilde{\lambda}^{-1/2}_1 M^{-1/2} \), to obtain the error estimate \( o(M^{-1/2}) \) in (1.10).

The error representation (3.5) consists of a \( X \)-drift term with the factor \( \partial_X u \), a \( p \)-drift term with the factor \( \partial_p u \), a diffusion term with the factor \( \partial_{pp} u \) and the three error terms (3.6) analyzed in the following sections.

4.1. The Friction Term. This section evaluates the friction term (3.10) using the solution operator (1.10), its time derivative and integration by parts (in the slow time scale) to remove the factor \( \tilde{H} \): study the first term in the right hand side of (3.10)

\[
\begin{align*}
-2\Re&\int_0^t \tilde{S}_{l,s} \psi_0^\perp ds, \tilde{H}^t \partial_X \psi_0^\perp \rangle \\
&= -2M^{-1/2} \Re(\int_0^t \int_0^s \tilde{S}_{l,s} \langle \psi_0, \partial_X \psi_0 \rangle ds, \tilde{H}^t \partial_X \psi_0^\perp) \\
&+ 2M^{-1/2} \Re(\int_0^t \int_0^s \tilde{S}_{l,s} \partial_X \psi_0 \psi_0^\perp ds, \tilde{H}^t \partial_X \psi_0^\perp) \\
&= -2M^{-1/2} \Re(\tilde{S}_{l,0} \langle \psi_0, \partial_X \psi_0 \rangle ds, \tilde{H}^t \partial_X \psi_0^\perp) \\
&+ 2M^{-1/2} \Re(\tilde{S}_{l,0} \partial_X \psi_0 \psi_0^\perp ds, \tilde{H}^t \partial_X \psi_0^\perp) \\
&= -2M^{-1/2} \Re(\tilde{S}_{l,0} \langle \psi_0, \partial_X \psi_0 \rangle ds, \tilde{H}^t \partial_X \psi_0^\perp) \\
&+ 2M^{-1/2} \Re(\tilde{S}_{l,0} \partial_X \psi_0 \psi_0^\perp ds, \tilde{H}^t \partial_X \psi_0^\perp).
\end{align*}
\]

Here we use that the null space of \( \tilde{H}(X) \) is spanned by \( \psi_0(X) \) which is orthogonal to \( \tilde{\psi}_0(X) = \partial_X \psi_0(X) \), so that \( \tilde{H}^{-1}(X) \psi_0(X) = \sum_{n>0} \tilde{\psi}_0(\psi_n, \psi_0) \) bounded by the assumption of a spectral gap (1.14) or (1.18). The first term in the right hand side yields the main friction term \( 2M^{-1/2} \int_{t=1}^t \langle \partial_X \psi_0, \partial_X \psi_0 \rangle ds \).

The integral over one time step of the second term above can also be integrated by parts as above to gain a power of \( M^{-1/2} \), loosing one power of \( (\Delta \tau)^{-1} \) and gaining one power of \( \tilde{\lambda}^{-1/2}_1 = o(1) \) from \( \tilde{H}^{-1} \), contributing in total to the coupling error \( O(\tilde{\lambda}^{-1}_1 M^{-1} \Delta \tau^{-1}) \).

The last term in the right hand side above can be integrated by parts again to give a negligible error contribution of order \( O(M^{-1}) \).

4.2. The Fluctuation-Dissipation Property. This section derives the fluctuation dissipation property based on the representation \( \phi = \sum_n \gamma_n \tilde{\psi}_n \), with independent normal distributed random variables \( \gamma_n \), and the orthonormal set \( \{ \tilde{\psi}_n \mid n = 1, \ldots \} \), orthogonal to \( \tilde{\psi}_0 \). We use the notation \( a \asymp b \), meaning \( a = b + o(1) \), at the following five steps in the calculation below:
1. in the first step the error term comes from $(X, p, φ)$ being slightly dependent on $γ_n$ – this coupling yields a small error term estimated in Section 4.3.

2. the second step uses the spectral gap condition (1.14) to relate $\tilde{λ}_n^0$ and $\tilde{λ}_n$, explained more below,

3. the third step uses that $\dot{ψ}_n = \tilde{ψ}_n + O(M^{-1/2})$ in (4.38) and (4.39) to replace a factor of $\tilde{λ}_n^{-1}$ with $\tilde{H}^{-1}$,

4. the fourth step applies assumption (1.13), that $\tilde{λ}_n \rightarrow \infty$, and the orthogonal set $\{\tilde{ψ}_n\}_{n=1}^\infty$ in (3.3) to have a sum of eigenstates forming a basis in the $L^2(T^3)$ orthogonal complement of $ψ_0$, and

5. the fifth step uses that $ψ_0$ is orthogonal to $\partial_X Ψ_0$ and $ψ_0 \simeq ψ_0$;

\[ E[2R(\phi^*, \tilde{H}\partial_X Ψ_0^*)] \leq E[γ_n^* γ_m^*] = 2T(\tilde{λ}_0^{-1} δ_{nm}) \]

\[ = 4R \sum_{n=1}^∞ \langle \tilde{S}_{τ,σ} \tilde{ψ}_n^*, \tilde{H}\partial_X Ψ_0^* \rangle \langle \tilde{H}\partial_X Ψ_0^* \tilde{ψ}_m^* \rangle \]  

\[ = 4R \sum_{n=1}^∞ \langle \tilde{S}_{τ,σ} \tilde{ψ}_n^*, \tilde{H}\partial_X Ψ_0^* \rangle \langle \tilde{H}\partial_X Ψ_0^* \tilde{ψ}_m^* \rangle \]  

\[ = 4R \sum_{n=1}^∞ \langle \tilde{S}_{τ,σ} \tilde{ψ}_n^*, \tilde{H}\partial_X Ψ_0^* \rangle \langle \tilde{H}\partial_X Ψ_0^* \tilde{ψ}_m^* \rangle \]  

Integration $\int_{τ_n-Δτ}^{τ_n} \ldots dσ$ gives $2T$ times the friction term we have in the left hand side of (1.2)

\[
\int_{τ_n-Δτ}^{τ_n} 4R \langle \tilde{S}_{τ,σ} \partial_X Ψ_0^* \rangle \tilde{H}^\dagger \partial_X Ψ_0^* dσ
\]

without the factor $p = \tilde{X}$, with the difference that the integration is over a time interval of length $Δτ$ instead of $t$. Since this is also integrated over a time interval with respect to $τ$, we conclude that the main contribution and the error terms are the same as for the friction case.

We see that this fluctuation-dissipation result requires precisely that the variance of $γ_n$ is $T/\tilde{λ}_n$ (to leading order), that $E[γ_n γ_m]$ is zero for $n \neq m$ and that $\{\tilde{ψ}_n\}_{n≥0}$ forms an orthonormal basis. In the second step we used that

\[ \frac{1}{\tilde{λ}_n(X^0)} = \frac{1}{\tilde{λ}_n(X^σ)} (1 - \frac{(X^0 - X^σ)}{\tilde{λ}_n}) \]

where the error term

\[ \sum_{n>0} \frac{(X^0 - X^σ)}{λ_n} \simeq o(1) \]

is negligible by the spectral gap assumption (1.14) or (1.18).

4.3. Couplings. The $p$–drift has a fluctuation part

\[ \Delta \dot{p} := -2 \int_{τ_n-Δτ}^{τ_n} R(\sum_{n \neq 0} γ_n \tilde{ψ}_n \tilde{H}(X^τ) \partial_X Ψ_0^* ) dτ. \]

This integral would have vanishing expected value if $X^τ$ and $\tilde{ψ}_n$ would be independent of $γ_n$, since $E[γ_n] = 0$.

In this section we show that the dependence between $(X, p, φ)$ and $γ_n$ is asymptotically negligible, so that
the expectation of this fluctuation part of the drift becomes small

\begin{equation}
\mathbb{E}[\partial_p u(X^{\tau_n}, p^{\tau_n}, \tau_{n+1}) \Delta \hat{p}] = \Delta \tau o(M^{-1/2}). \tag{4.7}
\end{equation}

The estimate is subtle in the sense that we expect that \( \Delta \hat{p} \) behaves as

\begin{equation}
(2TM^{-1/2}K^{1/2})^{1/2}(W^{\tau_n+\Delta \tau} - W^{\tau_n})
\end{equation}

and we will need to use that \( \partial_p u \) is evaluated in the point \( p^{\tau_n} \) and not \( p^{\tau_n+\Delta \tau} \), since e.g. a factor \( p \) in \( \partial_p u(X, p, \cdot) \) would give the coupling error \( \mathbb{E}[\Delta \hat{p} \otimes \Delta \hat{p}] = 2TM^{-1/2}K^{1/2} \Delta \tau \) by \((4.3)\), which is too large compared to \((4.7)\). To establish the improved estimate \((4.7)\) we use the oscillatory behavior of \( \bar{\psi}_n^\tau \) in integration by parts, as in the two error terms of \((4.2)\), which distinguishes between the arguments \( p^{\tau_n+\Delta \tau} \) and \( \bar{p}^{\tau_n} \) (since \( \tau \) will always be after \( \tau_n \)). In this sense the product of the fluctuation term \((4.6)\) and \( \partial_p u \) in \((4.8)\) generates one of the most important terms – the one corresponding to Itô martingales.

Define the first variation \( \partial_{\tau_n} (X, p, \phi) =: (X'_k, p'_k, \phi'_k) \), which satisfies the linearized Ehrenfest system

\begin{equation}
\begin{align*}
X' &= p', \\
p' &= -X' \circ \partial_{XX} \lambda_0(X) - 2\Re(\bar{\psi}_n \partial_X \tilde{H}(X) \phi) - \langle \phi, X' \circ \partial_{XX} \tilde{H} \phi \rangle, \\
\bar{\psi}_n' &= -2\Re\left(\sum_n \gamma_n \bar{\psi}_n^\tau, \partial_X \tilde{H} \phi\right), \\
\langle X, p, \phi' \rangle(0) &= 0,
\end{align*}
\end{equation}

where \( \phi = \sum_n \gamma_n \bar{\psi}_n \). Duhamel’s principle shows that

\begin{equation}
\langle X, p, \phi \rangle_n(\tau) = -2 \int_0^\tau f(\tau, \sigma) \Re(\bar{\psi}_n \partial_X \tilde{H}(X) \phi)(\sigma) d\sigma 
\end{equation}

and \( f \) is the linear solution operator for \((4.8)\), which is bounded \( O(1) \). Use this and the first variation to write the coupling as

\begin{equation}
\mathbb{E}[\partial_p u(X^{\tau_n}, p^{\tau_n}, \tau_{n+1}) \Delta \hat{p}] = \int_{\tau_n}^{\tau_{n+\Delta \tau}} \sum_{n \geq 0} \gamma_n^* \int_{[0, \gamma_n]} g d\gamma_n d\tau,
\end{equation}

where the function \( g \) is the first variation of \( \partial_p u(\bar{\psi}_n, \tilde{H} \partial_X \Psi_0) \) and by the chain rule it is a function of \((X, p, \phi)\)-derivatives of \( \partial_p u(\bar{\psi}_n, \tilde{H} \partial_X \Psi_0) \) and of \( f \). The coupling is then similar to the fluctuation term determined in Section \((4.2)\) with a product \( \gamma_n \bar{\psi}_n^\tau, \tilde{H} \partial_X \Psi_0^\tau \gamma_n \bar{\psi}_n^\tau, \tilde{H} \partial_X \Psi_0^\tau \) for the coupling with \( \partial_p u \): the difference that makes this coupling smaller is that \( \tau \) and \( \sigma \) are not in the same time interval, so the main term in \((4.2)\) will not appear, instead there will be two terms like the second term in the right hand side of \((4.2)\) containing the solution operator, generating an extra factor \( o(1) \) of decay in the case \( \Delta \tau \gg M^{-1/2} \). The coupling \( \gamma_n \bar{\psi}_n^\tau, \partial_X (\tilde{H} \partial_X \Psi_0^\tau) \gamma_n \bar{\psi}_n^\tau, \tilde{H} \partial_X \Psi_0^\tau \) introduced from \( \bar{p} \) itself can be handled by noting that

\begin{equation}
X'(\tau) = X'(\tau_n) + \int_{\tau_n}^\tau p'(s) ds
\end{equation}

where the \( X'(\tau_n) \) term behaves as above; the second term now has coupling in the same time interval, but here there is an additional factor of \( \Delta \tau \) coming from the time integral yielding the total coupling bound \( O(M^{-1/2}(\Delta \tau)^2) \) satisfying \((4.7)\).

We can similarly determine the coupling of quadratic terms. The diffusion term

\begin{equation}
\mathbb{E}[\partial_{pp} u(X^n, p^n, \tau_{n+1}) \Delta \hat{p} \otimes \Delta \hat{p}]
\end{equation}

has the main contribution from

\begin{equation}
\mathbb{E}[\partial_{pp} u(X^n, p^n, \tau_{n+1}) \Delta \hat{p} \otimes \Delta \hat{p}],
\end{equation}
To analyze the diffusion term similarly as the drift term, we differentiate (4.9) with respect to $\gamma_m$ to obtain a linear system for the second variation $\partial_{\gamma_n\gamma_m}(X,p,\phi)$

$$(X,p,\phi)^{\prime\prime}_{nm}(\tau) = -2\int_0^\tau \tilde{f}(\tau,\sigma)R(\tilde{\psi}_n,\partial_X\tilde{H}(X)\tilde{\psi}_m)(\sigma) d\sigma$$

and we obtain as in (4.10)

$$E[\partial_{pp}u(\tilde{X}^n,\tilde{p}^n,\tau_{n+1})\Delta\tilde{p} \otimes \Delta\tilde{p}] = \int_{\tau_n}^{\tau_n+\Delta\tau} \sum_{n\neq m>0} \int_{[0,\gamma_n]} \int_{[0,\gamma_m]} \gamma_n^*\gamma_m^* \tilde{d}_n\gamma_md\tau.$$

The second variation shows also that the remaining coupling terms

$$E[\partial_{pp}u(X^n,p^n,\tau_{n+1})\Delta p \otimes \Delta p]$$

$$E[\partial_{pp}u(X^n,p^n,\tau_{n+1})\Delta \tilde{p} \otimes \Delta \tilde{p}]$$

$$E[\partial_{pp}u(X^n,p^n,\tau_{n+1})(\Delta \tilde{p} \otimes \Delta \tilde{p} - 2TK^{1/2}M^{-1/2}\Delta \tau)]$$

with

$$\Delta\tilde{p} := \int_{t-\Delta t}^t \left( \sum_{n\neq 0} |\gamma_m|^2\langle \tilde{\psi}_n,\partial_X\tilde{H}\tilde{\psi}_n \rangle + \sum_{n\neq 0,m\neq 0,n\neq m} \gamma_n^*\gamma_m^* \langle \tilde{\psi}_n,\partial_X\tilde{H}\tilde{\psi}_m \rangle \right) ds,$$

are bounded by $\Delta\tau o(M^{-1/2})$.

The second term in $\tilde{p}$ uses $E[\gamma_n\gamma_m] = 0$ and coupling as follows. The solutions $\tilde{\psi}_n$ are the eigenfunctions $q_n$ of the essentially self-adjoint operator $Q$, defined by the backward solution operator $	ilde{S}_{0,t} = e^{-iM^{1/2}Q_{0,t}}$ (i.e. $\tilde{\psi}_n(0) = \tilde{S}_{0,t}\tilde{\psi}_n(t)$) with real eigenvalues $\lambda_n^*$, explained as follows. Let $f(0) = e^{iM^{1/2}Q_{0,t}}q_n(t)$. The continuity at zero gives $f(0) = q_n(0)$ and the limit $t \to 0+$ in

$$t^{-1}Q_{0,t}q_n(t) = \tilde{\lambda}_n(t)q_n(t)$$

combined with $t^{-1}Q_{0,t} \to \tilde{H}^0$ shows that $q_n(0)$ are the eigenfunctions of $\tilde{H}^0$. We have obtained

$$q_n(0) = \tilde{\psi}_n(0) = e^{iM^{1/2}Q_{0,t}}q_n(t)$$

which shows that $q_n(t)$ is the solution of the Ehrenfest dynamics along $X$ with initial data $\tilde{\psi}_n(0)$ and consequently

$$\tilde{\psi}_n(t) = q_n(t) = e^{-iM^{1/2}\tilde{\lambda}_n^*}\tilde{\psi}_n(0).$$

Differentiation of the relation $Q_{0,t}q_n(t) = \tilde{\lambda}_n(t)q_n(t)$ and $\dot{Q} = \tilde{H}$ shows that $\dot{\tilde{\lambda}}_n(0) = \tilde{\lambda}_n(0)$. Therefore

$$\langle \tilde{\psi}_n^s,\partial_X\tilde{H}\tilde{\psi}_m^s \rangle = e^{iM^{1/2}(\tilde{\lambda}_n^* - \tilde{\lambda}_m^*)} \langle \tilde{\psi}_n^0,\partial_X\tilde{H}\tilde{\psi}_m^0 \rangle$$

is oscillatory as a function of $s$ and integration by parts in the coupling term

$$I_{nm} := \int_0^\tau g^s |\gamma_n|^2|\gamma_m|^2 \langle \tilde{\psi}_n,\partial_X\tilde{H}\tilde{\psi}_m \rangle^\tau \langle \tilde{\psi}_n,\partial_X\tilde{H}\tilde{\psi}_m \rangle^s d\sigma$$

yields by the gap condition (1.14) or (1.18)

$$E[\sum_{nm} I_{nm}] = E[\sum_n |\gamma_n|^2 \sum_m |\gamma_m|^2] O(M^{-1/2}) = o(M^{-1/2}),$$

provided the phases $\tilde{\lambda}_n - \tilde{\lambda}_m$ have no critical point, which holds if the eigenvalues $\tilde{\lambda}_n$ do not cross as assumed.

The first term in $\tilde{p}$ for case (1.14) can be handled without coupling

$$E[\sum_{n>0} |\gamma_n|^2|\tilde{\psi}_n,\partial_X\tilde{H}\tilde{\psi}_n|] = E[\sum_{n>0} |\gamma_n|^2(\langle \tilde{\psi}_n,\partial_X\tilde{H}\tilde{\psi}_n \rangle + O(M^{-1/2})]$$

$$\leq C T \sum_{n>0} \frac{|\gamma_n|^2|\partial_X\tilde{\lambda}_n + o(M^{-1/2})|}{\tilde{\lambda}_n} = o(M^{-1/2}).$$
The first term in $\tilde{p}$ for case (1.18). The first term gives the leading order contribution

$$E\left[\sum_{n>0} |\gamma_n|^2 \langle \tilde{\psi}_n, \partial_X \tilde{H} \tilde{\psi}_n \rangle \right] = E\left[\sum_{n>0} \frac{T}{X_n} \langle \tilde{\psi}_n, \partial_X \tilde{H} \tilde{\psi}_n \rangle \right] = E\left[\sum_{n>0} \frac{T}{X_n} \left( \langle \tilde{\psi}_n, \partial_X \tilde{H} \tilde{\psi}_n \rangle + O(M^{-1/2}) \right) \right] = E\left[T \sum_{n>0} \frac{\lambda_n^t}{\lambda_n^t} \right] + o(M^{-1/2}).$$

The gap condition (1.18) and (4.4) show that

$$\frac{\partial_X \tilde{\lambda}_n^t}{\lambda_n^t} = \frac{\partial_X \tilde{\lambda}_n^t}{\lambda_n^t} + o(M^{-1/2})$$

which establishes the leading order correction to the drift of the Langevin equation

$$\frac{1}{2} T \sum_{n>0} \frac{\partial_X \tilde{\lambda}_n(X^t)}{\tilde{\lambda}_n(X^t)} = \frac{1}{2} T \partial_X \sum_{n>0} \log \tilde{\lambda}_n(X^t) = \frac{1}{2} T \partial_X \text{trace}^\perp \log \tilde{H}(X^t)$$

where trace$^\perp$ is the trace in the orthogonal complement of $\Psi_0$ and the factor $1/2$ comes from $\partial_{pp} u(\Delta p)^2/2$ in (3.5).

The coupling for this trace drift has, as in (4.11), the factor

$$\sum_{n>0} |\gamma_n|^2 \langle \tilde{\psi}_n, \partial_X \tilde{H} \tilde{\psi}_n \rangle^* \langle \tilde{\psi}_n, \partial_X \tilde{H} \tilde{\psi}_n \rangle$$

with its expected value bounded by $\sum_{n>0} (\partial_X \tilde{\lambda}_n \tilde{\lambda}_n^{-1})^2 = o(M^{-1/2})$, which makes the coupling negligible.

### 4.4. The Three Error Terms

Our main assumption, that the nuclear force is bounded $\langle \Psi, \partial_X H(X) \Psi \rangle = O(1)$, implies directly that the three error terms in the error representation (3.5) have the bound

$$|E[\partial_{tr} u(\Delta \tau)^2]| + |E[\partial_{tr} X u(\Delta X)^2]| + |E[\partial_{tr} u \Delta X \Delta p]| = O(\Delta \tau^2).$$

When the mean-field drift for the Langevin dynamics is small, the energy $e^* := |p_L|^2/2 + \lambda_0(X_L^t)$ is low; for instance at the minimal energy the mean-field drift is zero, which defines the equilibrium points. This energy evolves by Ito’s formula as

$$de^* = p_L^t \circ dp_L^t + \partial_X \lambda_0(X_L^t) \circ dX_L^t + 3TM^{-1/2}K^* N d\tau = M^{-1/2}K^* (3TN - |p_L|^2/2) d\tau + (2TM^{-1/2}K^*)^1/2 p_L^t \circ dW^t$$

so that its expected value satisfies

$$dE[e] = E\left[M^{-1/2}K^* (3TN - |p_L|^2/2) \right] d\tau.$$

To have initial data near equilibrium means that the energy is low (i.e. the mean-field drift is small) and (4.12) shows that the mean energy grows until the kinetic energy $|p_L|^2/2$ reaches the value $3TN$. This growth takes long time of the order $M^{1/2}$ in the slow time scale. Therefore the Langevin dynamics stays long time near equilibrium, where the mean-field drift is small, it starts from an equilibrium state where the energy is low.

### 4.5. The Born-Oppenheimer approximation

The purpose of this section is to study the evolution (3.2) of $\tilde{\psi}_n$ and explain $\tilde{\psi}_n = O(M^{-1/2})$ in (3.9) for the decomposition $\tilde{\psi}_n = \tilde{\psi}_n^t \oplus \tilde{\psi}_n^\perp$, where $\tilde{\psi}_n^t$ is an eigenvector of $\tilde{H}(X^t)$, with norm of order one, satisfying $\tilde{H}^* \tilde{\psi}_n^t = \tilde{\lambda}_n^t \tilde{\psi}_n^t$ for an eigenvalue $\tilde{\lambda}_n^t \in \mathbb{R}$. A more careful study of the Born-Oppenheimer approximation is in [24]. Let $\psi_n := e^{i(M^{1/2} \int_0^t \tilde{\lambda}_n^s ds)} \psi_n^t$. This Ansatz is motivated by the zero residual

$$R\psi_n := i\psi_n - M^{1/2}(\tilde{H} - \tilde{\lambda}_n)\psi_n = 0$$
and the small residual for the eigenvector
\[ \langle \tilde{\psi}^*_n, \tilde{\psi}_n \rangle = 0 \]
\[ M^{1/2}(\tilde{H} - \lambda_n)\tilde{\psi}_n = 0, \]
where \( w = |\tilde{\psi}_n|^{-2}(\tilde{\psi}_n, w)\tilde{\psi}_n + w^2 \) denotes the orthogonal decomposition in the eigenfunction direction and its orthogonal complement, as in Section 4.3. We have
\[ i\tilde{\psi}^+_n = M^{1/2}(\tilde{H} - \tilde{\lambda}_n)\tilde{\psi}^+_n - (R\tilde{\psi}_n)^2 \]
and by (4.2)
\[ \psi^+_n(t) = \tilde{S}_{t,0} \psi^+_n(0) - \int_0^t \tilde{S}_{t,s}(R\tilde{\psi}_n(s))^2 ds, \]
where \( \tilde{S} \) is the solution operator \( \tilde{S}_{t,0}\psi^+_n = \tilde{\psi}^+_n \). Therefore the integration by parts (4.2) can be applied and we obtain
\[ \psi^+_n(t) = \mathcal{O}(M^{-1/2}), \]
provided \( \tilde{\lambda}_m - \tilde{\lambda}_n \) does not vanish, which we assume.

5. The Constrained Stochastic Initial Data

As in models of heat baths we assume that the initial data of the light particles (here the electrons) are stochastic, sampled from an equilibrium distribution, see [12, 11, 28]. We use the Gibbs-Boltzmann distribution proportional to \( \exp(-\tilde{E}/T) \), which is an equilibrium solution of the Liouville equation of the Ehrenfest dynamics, as explained in Section 1.2. Let us now determine precise properties of this distribution generated by the Hamiltonian \( \tilde{H}_E \). Diagonalize the electron operator \( \tilde{H}(X^i) \) by the normalized eigenvectors and eigenvalues \( \{\tilde{\psi}_j, \tilde{\lambda}_j\} \)
\[ \langle \Psi, \tilde{H}(X^i)\Psi \rangle = \sum_{j>0} \tilde{\lambda}_j |\tilde{\gamma}_j|^2 \]
where
\[ \psi = \sum_{j \geq 0} \tilde{\gamma}_j \tilde{\psi}_j; \]
\[ \tilde{\lambda}_0 = 0, \]
with real and imaginary parts \( \tilde{\gamma}_j = \gamma_j^r + i\gamma_j^i \). We show in Section 5.1 that the equilibrium density can be approximated by the decoupled density
\[ D := \frac{\prod_{j>0} e^{-\tilde{\lambda}_j |\tilde{\gamma}_j|^2/T} d\gamma_j^r d\gamma_j^i dX dp}{\left(\int_{\mathbb{R}^N} \prod_{j>0} e^{-\tilde{\lambda}_j |\tilde{\gamma}_j|^2/T} d\gamma_j^r d\gamma_j^i dX dp\right)^{1/2}}, \]
which implies that \( \{\gamma_j^r, \gamma_j^i; j > 0\} \) are independent and each \( \gamma_j^r \) and \( \gamma_j^i \) is normally distributed \( \gamma_j^m \sim N(0, T/\tilde{\lambda}_j) \) for \( j > 0, m = r, i \). Note that \( \mathbb{E}[\gamma_j] = 0 \) for \( j > 0 \) and let \( \Psi_0 \) be the mean field approximation
\[ \Psi_0(X) := \sum_{j \geq 0} \mathbb{E}[\gamma_j] \tilde{\psi}_j(X) = \mathbb{E}[\gamma_0] \tilde{\psi}_0(X) \simeq \mathbb{E}[1 - \sum_{j>0} |\gamma_j|^2] \tilde{\psi}_0(X), \]
which requires the upper bound on the temperature, \( T \sum_{j>0} \tilde{\lambda}_j^{-1} < 1 \).

Note that our initial state is a pure eigenstate of the full Schrödinger operator with energy \( E \) and since there are many such states we use the model where the initial data is in any state \( \Psi \), for \( ||\Psi|| = 1 \), with probability weight
\[ e^{-\langle \Psi, \tilde{H}\Psi \rangle/T} d\Psi^* d\Psi. \]
The orthogonal transformation $\Psi = \sum_j \gamma_j \bar{\psi}_j$ shows that this probability is approximated by $D$ in (5.2), using that the determinant of the matrix of eigenvectors is one, where the approximation error comes from neglecting the constraint $\|\Psi\| = 1$. The advantage with the initial distribution derived by $D$ is that it generates a Langevin equation (which also satisfies the classical fluctuation-dissipation result), while initial data given by (2.5), or a mixed-state variant of that, lead to different fluctuations, see Remark 2.1 and Remark 5.2. Note that if we consider densities that are functions of the Hamiltonian $\tilde{H}$, it is only the exponential function (i.e. the Gibbs distribution) that yields independent electron states in this sense, since independence means that the joint distribution must be a product of individual distributions.

5.1. The Constraint $\|\Psi\| = 1$. The distribution (5.2) is determined by neglecting the constraint $\|\Psi\| = 1$. Let us now motivate it also with the constraint, asymptotically as $M$ tends to infinity. The constraint can, by homogeneity of the norm, be explicitly removed if one instead considers the density generated by

$$e^{-\sum_{j>0} \lambda_j |\bar{\gamma}_j|^2 / (T \sum_{j>0} |\bar{\gamma}_j|^2)},$$

for all values of $\bar{\gamma}_j$. This gives a uniform distribution of $\bar{\gamma}_0$ (conditioned on $\bar{\gamma}_j$, $j > 0$) on the domain of $\bar{\gamma}_0$, which is arbitrary large. Therefore the coupling sum $\sum_{j>0} |\bar{\gamma}_j|^2$ in the denominator is dominated by the term $|\bar{\gamma}_0|^2$. Normalizing $\bar{\gamma}_j$, $j > 0$ by $\bar{\gamma}_j =: \gamma_j \bar{\gamma}_0$ reduces (5.3) to

$$e^{-\sum_{j>0} \lambda_j |\gamma_j|^2 / (T + T \sum_{j>0} |\gamma_j|^2)}.$$

If we let $\gamma_j$ be independent normal distributed with zero mean and variance $\alpha T/\lambda_j$, for a certain constant $\alpha$, to be determined, the coupling term $1 + \sum_{j>0} |\gamma_j|^2$ is $o(1)$ close to the expected value $1 + \alpha T \sum_{j>0} \lambda_j^{-1}$, since its variance $\sum_{j>0} 2T^2 \lambda_j^{-2}$ is $o(1)$, as $M$ tends to infinity. Therefore the condition $1 + \alpha T \sum_{j>0} \lambda_j^{-1} = \alpha$, with the positive solution $\alpha = 1/(1 - T \sum_{j>0} \lambda_j^{-1})$, requiring the condition

$$T \sum_{j>0} \lambda_j^{-1} < 1,$$

implies

$$e^{-\sum_{j>0} \lambda_j |\gamma_j|^2 / (T + T \sum_{j>0} |\gamma_j|^2)} = e^{-\sum_{j>0} \lambda_j |\gamma_j|^2 / (\alpha T) (1 + o(1))}.$$

We have $\sum_{j>0} |\bar{\gamma}_j|^2 = |\gamma_0|^2 \alpha$ and the normalization leads to

$$\sum_{j>0} |\gamma_j|^2 / (|\gamma_0|^2) = T \sum_{j>0} \lambda_j^{-1},$$

$$\sum_{j>0} |\gamma_j|^2 / (\alpha |\gamma_0|^2) = 1 - T \sum_{j>0} \lambda_j^{-1},$$

which is the same as choosing $\gamma_j$ from (5.2) for $j>0$ and let $|\gamma_0|^2 = 1 - \sum_{j>0} |\gamma_j|^2$.

The normalization $\sum_{j>0} |\gamma_j|^2 = 1$ couples the variable $\gamma_0$ with the others

$$\gamma_0 = \sqrt{1 - \sum_{j>0} |\gamma_j|^2}.$$

At zero temperature $\gamma_j = 0$ for $j>0$ and $\gamma_0 = 1$. This normalization only makes sense for $1 - \sum_{j>0} |\gamma_j|^2 \geq 0$, which will be violated almost surely with normally distributed variables $\gamma_j$. Therefore we make a tiny cut-off in the probability density to have $1 - \sum_{j>0} |\gamma_j|^2 > 0$ with probability one. In fact the analysis uses only the mean, the variance and independence of $\gamma_j$, $j>0$. 

5.2. **Entropy and the Standard Canonical Density Distribution.** Let \( q_j \) denote the density of state \( j \) in the initial data \( \Psi^0 \). In the usual setting of a canonical Gibbs-Boltzmann distribution \( q_j = e^{-\lambda_j / T} / \sum_j e^{-\lambda_j / T} \), as in [23], which follows from maximizing the von Neumann entropy defined by \(-\sum_j q_j \log q_j\), with the probability and energy constraints \( \sum_j q_j = 1 \) and \( \sum_j \lambda_j q_j = \text{constant} \), see [13]. The stochastic model for the variable \( |\gamma_j|^2\), measuring in (5.2) and (5.1) the probability to be in electron state \( j \), is different: the chi-square distribution of \( \lambda_j |\gamma_j|^2 / T \) contains both the weight to be in electron state \( j \) and the spacial density of this state, see [24].

**References**

[1] G. Ariel and E. Vanden-Eijnden, A strong limit theorem in Kac-Zwanzig model, Nonlinearity 22 (2009) 145–162.
[2] F.A. Bornemann, P. Nettesheim and C. Schütte, Quantum-classical molecular dynamics as an approximation to full quantum dynamics, J. Chem. Phys. 105 (1996) 1074–1083.
[3] B. Cano, A.M. Stuart, Under-resolved simulations of heat-baths, J. Comp. Phys. 169 (2001) 193–214.
[4] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. LeBris and Y. Maday, Computational Chemistry: a primer, in Handbook of Numerical Analysis, X (North-Holland 2003).
[5] E. Cancès, F. Legoll and G. Stolz, Theoretical and numerical comparison of some sampling methods for molecular dynamics, Math. Model. Num. Anal. 41 (2007) 351–389.
[6] A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, Ann. Phys. 17 (1905) 549–560.
[7] P.A.M. Dirac, Note on exchange phenomena in the Thomas atom, Proc. Cambridge Phil. Soc. 26 (1930) 376–385.
[8] D. Dürr, S. Goldstein and J.L. Lebowitz, A mechanical model for the Brownian motion of a convex body, Z. Wahrscheinlichkeitstheori verw. Gebiete 62 (1983) 427–448.
[9] D. Dürr, S. Goldstein and J.L. Lebowitz, A mechanical model of Brownian motion, Commun. Math. Phys. 78 (1981) 597–530.
[10] R.F. Feynman, Statistical Mechanics: A Set of Lectures (Westview Press 1998).
[11] G.W. Ford and M. Kac, On the quantum Langevin equation, J. Statist. Phys. 46 (1987), 803–810.
[12] G.W. Ford, M. Kac and P. Mazur, Statistical mechanics of assemblies of coupled oscillators, J. Mathematical Phys. 6 (1965) 504–515.
[13] C.W. Gardiner, Quantum Noise (Springer-Verlag 1991).
[14] G.A. Hagedorn, A time-dependent Born-Oppenheimer approximation, Commun. Math. Phys. 77 (1980) 1–19.
[15] O.H. Hald and R. Kupferman, Asymptotic and numerical analysis for mechanical models of heat baths, J. Statist. Phys. 106 (2002) 1121–1184.
[16] P. Häussler, Interrelations between atomic and electronic structures – Liquid and amorphous metals as model systems, Physics Reports 222 (1992) 65–143.
[17] L. Kadanoff, Statistical Physics: Statistics, Dynamics and Renormalization (World Scientific 2000).
[18] R. Kupferman and A.M. Stuart, Fitting SDE models to nonlinear Kac-Zwanzig heat bath models, Physica D 199 (2004) 279–316.
[19] R. Kupferman, A.M. Stuart A.M., J.R. Terry and P.F. Tupper, Long-term behaviour of large mechanical systems with random initial data, J. Stoch. Dyn. 2 (2002) 533–562.
[20] P. Langevin, On the theory of Brownian movement, C.R. Acad. Sci. 146 530 (1908), (translation Am. J. Phys. 65 1079, 1997).
[21] E. Lieb, Thomas-Fermi and related theories for atoms and molecules, Rev. Modern Phys. 53 (1981) 603–641.
[22] E. Madelung, Z. Phys. 40 (1926) 549–560.
[23] D. Marx and J. Hutter, Ab initio molecular dynamics: Theory and implementation, Modern Methods and Algorithms of Quantum Chemistry, J.Grotendorst(Ed.), John von Neumann Institute for Computing, Jülich, NIC Series, Vol. 1, ISBN 3-00-005618-1 (2000) 301–449.
[24] A. Szepessy, Stochastic and deterministic molecular dynamics derived from the time-independent Schrödinger equation, arXiv:0812.4338.
[25] A.M. Stuart and J.O. Warren, Analysis and experiments for a computational model of a heat bath, J. Stat. Phys. 97 (1999) 687–723.
[26] J.C. Tully, Mixed quantum-classical dynamics, Faraday Discuss. 110 (1998) 407–419.
[27] C.M. Isborn, X. Li, J.C. Tully, Time-dependent density functional theory Ehrenfest dynamics: Collisions between atomic oxygen and graphite clusters, J. Chem. Phys. 126 (2007) 134307.
[28] R. Zwanzig, Nonlinear generalized Langevin equations, J. Stat. Phys. 9 (1973) 215–220.
Department of Mathematics, Kungl. Tekniska Högskolan, 100 44 Stockholm, Sweden
E-mail address: szepessy@kth.se