LANGEVIN MOLECULAR DYNAMICS DERIVED FROM EHRENFEST DYNAMICS

ANDERS SZEPESSY

ABSTRACT. Stochastic Langevin molecular dynamics for nuclei is derived from the Ehrenfest Hamiltonian system (also called quantum classical molecular dynamics) in a Kac-Zwanzig setting, with the initial data for the electrons stochastically perturbed from the ground state and the ratio, M, of nuclei and electron mass tending to infinity. The Ehrenfest nuclei dynamics is approximated by the Langevin dynamics with accuracy $o(M^{-1/2})$ on bounded time intervals and by $o(1)$ on unbounded time intervals, which makes the small $\mathcal{O}(M^{-1/2})$ friction and $o(M^{-1/2})$ diffusion terms visible. The initial electron probability distribution is a Gibbs density at low temperture, derived by a stability and consistency argument: starting with any equilibrium measure of the Ehrenfest Hamiltonian system, the initial electron distribution is sampled from the equilibrium measure conditioned on the nuclei positions, which after long time leads to the nuclei positions in a Gibbs distribution (i.e. asymptotic stability); by consistency the original equilibrium measure is then a Gibbs measure. The diffusion and friction coefficients in the Langevin equation satisfy the Einstein's fluctuation-dissipation relation.

CONTENTS

1. Introduction to Ab Initio Molecular Dynamics 1
2. Zwanzig's Model and Derivation of Langevin Dynamics 5
2.1. The Gibbs Distribution Derived from Dynamic Stability 10
3. Data for the Ehrenfest Model and the Main Results 12
3.1. The Force on Nuclei 13
3.2. Stochastic Electron Initial Data 14
3.3. The Main Result 18
4. Error Estimates 20
4.1. The Dynamics 20
4.2. Proof of the Theorem 21
4.3. The Born-Oppenheimer Approximation 29
4.4. Proof of Theorem 3.3 30
4.5. A Motivation for $\lim_{\tau\to\infty}\int_0^\tau \partial_p u(X^\sigma, p^\sigma, \sigma; \tau)d\sigma = \mathcal{O}(M^{1/2})$ 31
4.6. A Motivation for Non Zero Friction 33
4.7. Other Initial Electron Distributions 33
References 34

1. INTRODUCTION TO AB INITIO MOLECULAR DYNAMICS

One method to simulate molecular motion is to use quantum classical molecular dynamics (QCMD), also called *Ehrenfest dynamics*, where the nuclear positions $X_n : [0, \infty) \to \mathbb{R}^3,\ n =$

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

$1,\dots,N$ and the electron wave function $\bar\psi:[0,\infty)\times\mathbb{R}^{3J}\to\mathbb{C}$ solve the Hamiltonian system

$$M\ddot X_n^t = -\langle\bar\psi^t,\partial_{X_n}H(X^t)\bar\psi^t\rangle, \tag{1.1}$$

$$i\frac{d}{dt}\bar\psi^t = H(X^t)\bar\psi^t, \tag{1.2}$$

see [26], [17]. The wave function belongs to a Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and the operator $H(X)$ is self-adjoint in that Hilbert space. In computational chemistry the operator H, the electron Hamiltonian, is precisely determined by the sum of kinetic energy of the electrons and the Coulomb interaction between nuclei and electrons; the Hilbert space is then a subset of $L^2(\mathbb{R}^{3J})$ with symmetry conditions based on the Pauli exclusion principle for electrons, see [25], [4]. Therefore the complex valued L^2 inner product

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

is used here. 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 (1.1-1.2) is a Hamiltonian system with the Hamiltonian

$$M^{-1}\frac{|p|^2}{2}+\frac12\langle\phi,H\phi\rangle=:H_E$$

in the real variable $\big((X,\phi^r),(p,\phi^i)\big)=:(r_E,p_E)$, where $p=M\dot X$,

(1.3) $\phi=:\phi^r+i\phi^i$ is the decomposition into real ϕ^r and imaginary ϕ^i parts,

$\langle\phi,\phi\rangle=2$ and $\bar\psi:=\phi/\langle\phi,\phi\rangle^{1/2}$; the normalization $\phi/\langle\phi,\phi\rangle^{1/2}$ is possible since the Schrödinger equation (1.2) is linear and the norm $\langle\phi^t,\phi^t\rangle$ remains constant in time. Conequently, the Ehrenfest dynamics conserves the energy H_E. 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 for the full nuclei-electron system, see [27], and from the time-dependent Schrödinger equation, cf. [3, 26]. The work [29] describes with examples some of the weaknesses and strengths of the Ehrenfest approximation.

The Ehrenfest dynamics can be coarse-grained by eliminating the electron dynamics and assuming the electron wave function is in its ground state Ψ_0, satisfying the eigenvalue problem $H(X)\Psi_0(X)=\lambda_0(X)\Psi_0(X)$ for the lowest eigenvalue $\lambda_0(X)$, which reduces the Ehrenfest dynamics (1.1)-(1.2) to so called *Born-Oppenheimer dynamics* [26]

$$M\ddot X_{BO}^t=-\partial_X\lambda_0(X_{BO}^t).$$

The large mass $M\gg1$ together with the bounded forces $\partial_X\lambda_0(X_{BO}^t)\sim1$ make the dynamics slow and position increments are better studied in the slower time scale, where $\tau:=M^{-1/2}t$ and $\hat\tau:=M^{1/2}\tau=t$,

$$\frac{d^2}{d\tau^2}X_{BO}^{\hat\tau}=-\partial_X\lambda_0(X_{BO}^{\hat\tau}),$$

since this dynamics does not depend on M; we will use greek letters $\tau,\sigma,\dots$ to denote time in the slow scale and latin letters $t,s,\dots$ for time in the fast scale. The Born-Oppenheimer approximation leads to accurate approximation of observables for the *time-independent Schrödinger equation*,

$$\big(-(2M)^{-1}\Delta_X+H(X)\big)\Phi(X,\cdot)=E\,\Phi(X,\cdot),$$

in the case electron eigenvalues do not cross:

$$|\int_{\mathbb{R}^{3N}} g(X)\langle\Phi(X,\cdot),\Phi(X,\cdot)\rangle dX - \int_{\mathbb{R}^{3N}} g(X)\rho_{BO}(dX)| = \mathcal{O}(M^{-1}) \tag{1.4}$$

for smooth functions g, see [27]. We use the big $\mathcal{O}$ and little o notation for

$$h(M) = \mathcal{O}\big(f(M)\big) \iff \exists\, C, M_0 \in \mathbb{R} \text{ such that } |h(M)| \le Cf(M) \text{ for } M > M_0,$$

$$h(M) = o\big(f(M)\big) \iff \lim_{M\to\infty} \frac{h(M)}{f(M)} = 0.$$

The first integral in (1.4) is the quantum mechanical measure of the nuclei-position *observable* $g(X)$ and $\int_{\mathbb{R}^{3N}} g(X)\rho_{BO}(dX)$ is the the micro canonical ensemble average of $g(X_{BO})$, which would be equal to $\lim_{\mathcal{T}\to\infty} \mathcal{T}^{-1}\int_0^{\mathcal{T}} g(X^t_{BO})dt$ if X_{BO} would be ergodic. The two densities are normalized so that

$$1 = \int_{\mathbb{R}^{3N}} \langle\Phi(X,\cdot),\Phi(X,\cdot)\rangle dX = \int_{\mathbb{R}^{3N}} \rho_{BO}(dX).$$

The same result holds for the Ehrenfest dynamics replacing the Born-Oppenheimer dynamics; however, this accuracy requires to know the initial electron wave function very precisely. It seems reasonable to study randomness in the initial data of the electron wave function – the modeling and accuracy with stochastic electron data $\bar\psi$ in (1.2) is the purpose of this work.

The main inspiration for the stochastic model here is [12]-[11], by Kac, Ford and Mazur, and in particular [30], where Zwanzig derives a Langevin equation for a heavy particle, from a Hamiltonian system with the heavy particle coupled through a harmonic interaction potential to a heat bath particle system. The stochastics enters by having the heat bath degrees of freedom initially Gibbs distributed at a certain temperature, which after elimination of the heat bath degrees of freedom yields a generalized Langevin equation for the heavy particle, including an integral operator for the friction term – a so called memory term. An assumption of a limiting continuous Debye heat bath frequency distribution and certain linear weak coupling behavior then reduces the dynamics to a proper Langevin equation, where the integral kernel becomes a point mass, i.e. an equation without memory terms.

This work shows, in Section 2, that Zwanzig's model is closely related to the Ehrenfest Hamiltonian system and Sections 3-4 extends the ideas in [30] to the *ab initio* Ehrenfest dynamics (1.1) for nuclei and the Schrödinger equation (1.2) for electrons (or other light particles). This means that the electron wave function plays the role of the "heat bath system", the stochastics enters as Gibbs distributed initial data for the electron wave function $\bar\psi$, and the Gibbs measure is parametrized by temperature. The approximating *Langevin dynamics*, for the nuclei positions X_L^τ in the slow time scale, takes the form

$$\begin{aligned}\dot X_L^\tau &= p_L^\tau\\ \dot p_L^\tau &= -\partial_X\lambda_0(X_L^\tau) - \hat K p_L^\tau + (2T)^{1/2}\hat K^{1/2}\dot W^\tau,\end{aligned}$$

at low temperature $T \ll 1$, where W is the standard Wiener process in $\mathbb{R}^{3N}$ and the positive $3N\times 3N$ friction/diffusion matrix $\hat K = \hat K(X_L^\tau)$ is a certain small function (of the ground state Ψ_0) of order $M^{-1/2}$.

The purpose of this work is to study some observables including time-correlation to precisely determine the friction/diffusion matrix $\hat K$ for the Langevin dynamics. The main idea is to use the Hamiltonian structure of the Ehrenfest dynamics to formulate and determine stochastic molecular dynamics in the canonical ensemble of constant number of particles constant volume and constant temperature. Theorem 3.1 determines a specific friction matrix so that Langevin dynamics approximates Ehrenfest dynamics for some 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 $\mathcal{O}(M^{-1/2})$ small friction term visible. Theorem 3.3 shows that the accuracy is $o(1)$ on unbounded time intervals, assuming the time correlation length, of the first variation of the observable with respect to momentum, is at most of order $M^{1/2}$; this result, valid for long time approximation, detects also the $o(M^{-1/2})$ small diffusion term. The main assumption for the results is related to the reduction to a point mass for the integral kernel in the friction term. In contrast to [30] no explicit expression of the frequency distribution of the coupling is used. Instead it is more generally assumed that the electron eigenvalues form a continuum in the limit as M, J tend to infinity, so that the zero frequency limit of the spectral representation of $(H-\lambda_0)\partial_X\Psi_0(X)$ exists, in the sense described in (3.13) based on random matrices. A second assumption is that the ℓ^1 norm of the first variation of the observable is bounded, which restricts the study to observables that are stable in the sense discussed in Remark 3.2. The results also require the temperature to be low $T \ll 1$, all electron eigenvalues to be separated and $\tilde H$ to depend smoothly on X.

The particles with coordinates x, in (1.1)-(1.2), 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 (3.18) also describes approximately the dynamics of heavy so called Brownian particles. Theorems 3.1 and 3.3 therefore contribute to 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, cf. [18]. Such heat bath studies initiates from the early pioneering work [7],[22] and continues with more precise heat bath models, based on harmonic interactions, in [12, 11] [30]. More recently these models of a heavy particle coupled to a heat bath are also used for numerical analysis studies related to coarse-graining in molecular dynamics and weak convergence analysis [28, 20] [16], for strong convergence analysis [2], and for computational studies on nonlinear heat bath models [6, 19]. 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].

Five new issues here are:

- why shall the electron initial data be Gibbs distributed? If the electron data could have other distributions it would lead to different results and one would have a problem to predict the dynamics. Here the initial Gibbs distribution for the electrons is derived from a stability and consistency argument in Section 2.1 providing the following closure result: starting with any equilibrium measure of the Ehrenfest Hamiltonian system, the initial electron distribution is sampled from the equilibrium measure conditioned on the nuclei positions, which turns out to be a Gibbs density when the ground state energy dominates and after long time it leads to the nuclei positions in a Gibbs distribution; by consistency the original equilibrium measure is then a Gibbs measure.
- The slow nuclei dynamics compared to the fast electron dynamics is exploited, to find a proper Langevin equation without using explicit heat bath frequencies.
- The error analysis uses the residual in the Kolmogorov equation, of the Langevin dynamics, evaluated along the Ehrenfest dynamics, instead of the explicit solution available for harmonic oscillators.
- A long time result uses exponential decay in time of the first variation of the observable with respect to perturbations in the momentum.

- The wave function $\bar{\psi}$ needs to be normalized, which yields a global coupling of the random initial data that is studied as a perturbation by an assumption of low temperature.

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 Gibbs fluctuations in initial data of the light particles – were already the basis in [12, 11] and [30].

The outline of the paper is the following. Section 2 connects the Ehrenfest system to the Zwanzig model and presents a uniqueness argument for the stochastic initial data. Section 3 formulates the stochastic data for the Ehrenfest model, including the normalization constraint, and states the two theorems proved in Section 4.

2. Zwanzig's Model and Derivation of Langevin Dynamics

This section reviews the heat bath model of Zwanzig [30], with his derivation of stochastic Langevin dynamics from a Hamiltonian system, related to the earlier work [12]. The model consists of heavy particles interacting with many light particles which are initially in a Gibbs distribution; in this sense, it models heavy particles in a heat bath of light particles. A modification of the general formulation in [30] is here to choose a special case closely resembling the Ehrenfest dynamics. The model is as simple as possible to have the desired qualitative properties of a system interacting with a heat bath.

The aim is to show that the Ehrenfest model is closely related to the Zwanzig model and to give some understanding of simulating, at constant temperature, the coarse-grained molecular dynamics of the heavy particles without resolving the lighter particles, using Langevin dynamics. It is an example how stochastics enter into a coarse-grained model through elimination of some degrees of freedom in a determinstic model, described by a Hamiltonian system. The original model is time reversible while the coarse-grained model is not.

We consider Zwanzig's model in [30] with N heavy particles and the particle positions $X \in \mathbb{R}^{3N}$ and momentum $p \in \mathbb{R}^{3N}$ in a heat bath with J light particles modes, with "position" coordinate $x \in \mathbb{R}^J$ and dual "momentum" coordinate $q \in \mathbb{R}^J$. Define, as a special case of *Zwanzig's general model*, the Hamiltonian

$$H_Z(X,p,x,q) := \frac{1}{2}|p|^2 + \lambda(X) + \frac{m}{2}\langle x - \hat{\Psi}(X), \hat{H}(x - \hat{\Psi}(X))\rangle + \frac{1}{2m}\langle q, \hat{H}q\rangle \tag{2.1}$$

where the operator $\hat{H}$ is a constant positive definite Hermitian $J \times J$ matrix, the coupling is represented by the function $\hat{\Psi} : \mathbb{R}^{3N} \to \mathbb{R}^J$, the bilinear form $\langle \cdot, \cdot \rangle$ is now the Euclidian scalar product in $\mathbb{R}^J$, the light particle mass is m, the heavy particle mass is one and $\lambda : \mathbb{R}^{3N} \to \mathbb{R}^{3N}$ is a given potential. This Hamiltonian yields the dynamics (in the slow time scale of the heavy particles)

$$\ddot{X}^\tau = -\lambda'(X^\tau) + \langle m\hat{H}(x - \hat{\Psi}(X^\tau)), \hat{\Psi}'(X^\tau)\rangle, \tag{2.2}$$

$$\dot{x}^\tau = m^{-1}\hat{H}q^\tau \tag{2.3}$$

$$\dot{q}^\tau = -m\hat{H}(x^\tau - \hat{\Psi}(X^\tau)), \tag{2.4}$$

where we use the notation $\Psi_0'(X) := \partial_X \Psi_0(X)$. To resemble the Ehrenfest dynamics, we define $y := m^{-1}q$ and obtain

$$\dot{x}^\tau = \hat{H}y^\tau$$
$$\dot{y}^\tau = -\hat{H}(x^\tau - \hat{\Psi}(X^\tau))$$

which using $\psi := x - \hat\Psi(X) + iy$ shows the Zwanzig model written as a Schrödinger equation coupled to heavy particle dynamics

$$\begin{aligned}\ddot X^\tau &= -\lambda'(X^\tau) + \Re\langle m\hat H\Psi^\tau, \hat\Psi'(X^\tau)\rangle,\\ i\dot\psi^\tau &= \hat H\psi^\tau - i\hat\Psi'(X^\tau)\dot X^\tau,\end{aligned} \tag{2.5}$$

where $\Re w$ denotes the real part of $w \in \mathbb{C}$.

Let us compare the Zwanzig model to the Ehrenfest dynamics:

Lemma 2.1. *The Ehrenfest dynamics* (1.1)-(1.2) *can be written in the slow time scale as*

$$\begin{aligned}\ddot X^\tau &= -\lambda_0'(X^\tau) + 2\Re\langle \tilde H\tilde\psi^\tau, \Psi_0'(X^\tau)\rangle - \langle\tilde\psi^\tau, \tilde H'(X^\tau)\tilde\psi^\tau\rangle,\\ i\dot{\tilde\psi}^\tau &= M^{1/2}\tilde H(X^\tau)\tilde\psi^\tau - i\Psi_0'(X^\tau)\cdot\dot X^\tau,\end{aligned} \tag{2.6}$$

using the definitions

$H\Psi_0 = \lambda_0\Psi_0,$	*eigenvalue*
$H_E(X,\phi_r;p,\phi_i) := \frac{\|p\|^2}{2} + \lambda_0(X) + \frac{M^{1/2}}{2}\langle\phi,\tilde H\phi\rangle,$	*Hamiltonian*
$\tilde H := H - \lambda_0,$	*operator splitting*
$\psi := \frac{\phi}{\sqrt{\phi\cdot\phi}}, \quad M^{1/2}/2 = 1/\langle\phi,\phi\rangle,$	*normalization*
$\tilde\psi := \psi - \Psi_0,$	*wave function splitting,*
$\tau = M^{-1/2}t,$	*slow time scale.*

The lemma is proved in Section 3, see references (3.4) and (3.8), by direct application of the definitions in the lemma. We note that the Zwanzig dynamics (2.5) and the Ehrenfest dynamics (2.6) are similar and by choosing

$$\begin{aligned}\hat\Psi(X) &= \Psi_0(X),\\ \hat H &= M^{1/2}\tilde H,\\ m &= 2M^{-1/2},\\ \lambda &= \lambda_0,\end{aligned} \tag{2.7}$$

in the case that $\tilde H$ is constant (in the subspace orthogonal to Ψ_0), the two models are identical, since the quadratic interaction $-\langle\psi,\tilde H'\psi\rangle$ vanishes. We also note that extending the Zwanzig model (2.1) to a case when the matrix $\hat H : \mathbb{R}^{3N} \to \mathbb{R}^J\times\mathbb{R}^J$ is a function of X yields exactly the Ehrenfest model (2.6), including the quadratic interaction. We will see that the quadratic term is negligible for sufficiently low temperature. Consequently, the results for the general Zwanzig model are relevant for the Ehrenfest model; this section reviews the results from [30] and the other sections extends them to the Ehrenfest model (with non constant $\tilde H$).

It seems reasonable to assume that the many initial positions and velocities of the light particles are impossible to measure and determine precisely. Clearly, to predict the dynamics of the heavy particle some information of the light particle initial data is necessary: we shall use an equilibrium probability distribution for the light particles depending only on one parameter – the temperature T. Section 2.1 presents a motivation for stochastic sampling of $\psi(0)$ from the *Gibbs probability measure*

$$\begin{aligned}&Z^{-1}\exp\big(-H_Z(X^0,p^0,x,q)/T\big)dx_1\ldots dx_J dq_1\ldots dq_J,\\ &Z := \int_{\mathbb{R}^{2J}}\exp\big(-H_Z(X^0,p^0,x,q)/T\big)dx_1\ldots dx_J dq_1\ldots dq_J.\end{aligned} \tag{2.8}$$

The distribution of the initial data $\psi(0) = x^0 - \hat\Psi(X^0) + iq^0/m$, using for simplicity $\hat\Psi(X^0) = 0$, becomes clear in an orthonormal basis $\{\Psi_j\}_{j=1}^J$ of eigenvectors to $\hat H$, with the corresponding eigenvalues $\{\lambda_j\}_{j=1}^J$, since writing $\psi(0) = \sum_{j=1}^J \gamma_j\Psi_j$ implies

$$\frac{m}{2}\langle\psi^0, \hat H\psi^0\rangle = \sum_{j=1}^J \frac{m\lambda_j}{2}|\gamma_j|^2 = \sum_{j=1}^J \frac{m\lambda_j}{2}(|\gamma_j^r|^2 + |\gamma_j^i|^2),$$

so that

$$e^{-\frac{m}{2T}\langle\psi^0,\hat H\psi^0\rangle} dx_1\dots dx_J dq_1\dots dq_J = e^{-\sum_{j=1}^J \frac{m\lambda_j}{2T}(|\gamma_r^j|^2+|\gamma_i^j|^2)} d\gamma_1^r\dots d\gamma_J^r d\gamma_1^i\dots d\gamma_J^i, \tag{2.9}$$

where it is used that the orthogonal transformation to the coordinates $x = \sum_j \gamma_j^r\Psi_j$ and $q = \sum_j \gamma_j^i\Psi_j$ has the Jacobian determinant equal to one and $\gamma_j = \gamma_j^r + \gamma_j^i$ is the decomposition into real and imaginary parts as in (1.3). We conclude that the complex valued γ_j are all independent with independent normal $N(0, 2T/(m\lambda_j))$ distributed real and imaginary parts, γ_r^j and γ_i^j.

Given the path X, the linear equation (2.4) can by Duhamel's representation be solved explicitly, with the solution

$$\psi(\tau) = -\int_0^\tau e^{-i(\tau-\sigma)\hat H}\hat\Psi'(X^\sigma)\dot X^\sigma d\sigma + \underbrace{e^{-i\tau\hat H}\psi(0)}_{z^\tau}. \tag{2.10}$$

Insert this representation into equation (2.2), for the heavy particle, to obtain

$$\ddot X^\tau = -\lambda'(X^\tau) - \int_0^\tau \langle m\hat H\cos\big((\tau-\sigma)\hat H\big)\hat\Psi'(X^\sigma)\dot X^\sigma, \hat\Psi'(X^\tau)\rangle d\sigma + \underbrace{\Re\langle m\hat H z^\tau, \hat\Psi'(X^\tau)\rangle}_{\zeta^\tau}. \tag{2.11}$$

To determine the covariance $\mathbb{E}_\gamma[\zeta^\sigma\otimes\zeta^{\tau *}]$, where $\mathbb{E}_\gamma$ denotes the expected value with respect to the initial data γ, note that the product of the fluctuation terms

$$\big(\langle\dots{}^\tau\rangle + \langle\dots{}^\tau\rangle^*\big)\big(\langle\dots{}^\sigma\rangle + \langle\dots{}^\sigma\rangle^*\big),$$

where $\langle\dots{}^\tau\rangle := \langle m\hat H z^\tau/2, \Psi'(X^\tau)\rangle$, yields four terms. This sum of four terms can be written as follows; use first that the initial data is $\psi^0 = \sum_j \gamma_j\Psi_j$, where $\{\gamma_j\}_{j=1}J$ forms a set of independent normal distributed random variables, and then that $\{\Psi_j\}_{j=1}^J$ is an orthonormal basis to obtain

$$\begin{aligned}
\mathbb{E}_\gamma[\zeta^\sigma\otimes\zeta^{\tau *}] &= \mathbb{E}_\gamma\big[\frac12\Re\langle m\hat H e^{-i\sigma\hat H}\psi^0, \hat\Psi'(X^\sigma)\rangle\langle\hat\Psi'(X^\tau), m\hat H e^{-i\tau\hat H}\psi^0\rangle\big]\\
&\quad + \mathbb{E}_\gamma\big[\frac12\Re\langle m\hat H e^{-i\sigma\hat H}\psi^0, \hat\Psi'(X^\sigma)\rangle\langle m\hat H e^{-i\tau\hat H}\psi^0, \hat\Psi'(X^\tau)\rangle\big]\\
&\simeq \frac12\Re\sum_{j,k}\langle m\lambda_j\Psi_j, e^{i\sigma\hat H}\hat\Psi'(X^\sigma)\rangle\langle m\hat H e^{i\tau\hat H}\hat\Psi'(X^\tau), \Psi_k\rangle \underbrace{\mathbb{E}_\gamma[\gamma_j^*\gamma_k]}_{=4T\delta_{jk}/(m\lambda_j)}\\
&\quad + \frac12\Re\sum_{j,k}\langle m\lambda_j\Psi_j, e^{i\sigma\hat H}\hat\Psi'(X^\sigma)\rangle\langle\Psi_k, m\hat H e^{i\tau\hat H}\hat\Psi'(X^\tau),\rangle \underbrace{\mathbb{E}_\gamma[\gamma_j^*\gamma_k^*]}_{=0}\\
&= 2T\Re\sum_j\langle\Psi_j, e^{i\sigma\hat H}\hat\Psi'(X^\sigma)\rangle\langle m\hat H e^{i\tau\hat H}\hat\Psi'(X^\tau), \Psi_j\rangle\\
&= 2T\langle m\hat H\cos\big((\tau-\sigma)\hat H\big)\Psi'(X^\sigma), \Psi'(X^\tau)\rangle;
\end{aligned} \tag{2.12}$$

we have in the second step, where $a\simeq b \iff a = b(1+o(1))$, used that X depends on γ_n in such a way that the coupling between X and γ_n leads to lower order terms $\mathcal{O}(M^{-1/2})$, which is verified in Lemma 4.1. We see that the covariance of the Gaussian process, $\zeta : [0,\infty)\times\{\text{probability outcomes}\}\to\mathbb{R}^{3N}$,

$$\mathbb{E}_\gamma[\zeta^\sigma\otimes\zeta^{\tau *}] \simeq 2T\langle m\hat H\cos\big((\tau-\sigma)\hat H\big)\hat\Psi'(X^\sigma),\hat\Psi'(X^\tau)\rangle =: 2Tf(\tau-\sigma), \tag{2.13}$$

is the multiple $2T$ of the integral kernel for the friction term in the generalized Langevin equation (2.11), forming a version of Einstein's fluctuation-dissipation result.

The next step in Zwanzig's modeling is to derive a pure Langevin equation for a special choice of oscillation frequencies $\hat H$ and coupling $\hat\Psi$. Write $\hat\Psi'(X) =: \sum_{j=1}^J \hat\Psi_j'(X)\Psi_j$ to obtain

$$\langle\hat H\cos\big((\tau-\sigma)\hat H\big)\hat\Psi'(X^\sigma),\hat\Psi'(X^\tau)\rangle = \sum_{j=1}\lambda_j\cos\big((\tau-\sigma)\lambda_j\big)\hat\Psi_j'(X^\sigma)\hat\Psi_j'(X^\tau).$$

Assume now that the frequencies λ_j are distributed so that the sum over particles is in fact an integral over frequencies with a *Debye distribution*, i.e. for any continuous function h

$$J^{-1}\sum_{j=1}^J h(\lambda_j)\to\int_0^{\lambda_d} h(\lambda)\frac{3\lambda^2}{\lambda_d^3}d\lambda, \tag{2.14}$$

and let the coupling between the heavy particle and the heat bath be linear satisfying

$$3\lambda_j^3\partial_{X_i}\hat\Psi_j(X^\sigma)\partial_{X_k}\hat\Psi_j(X^\tau) = \kappa_{ik}J^{-1} \tag{2.15}$$

to obtain

$$f_{ik}(\tau) = \frac{m\kappa_{ik}}{\lambda_d^3}\frac{\sin(\lambda_d\tau)}{\tau},\qquad \text{as } J\to\infty.$$

Suppose that $\frac{\pi m}{2\lambda_d^3}\kappa\to\hat K$ as $\lambda_d\to\infty$, then the integral kernel $f(\tau)$ converges weakly to a point mass, since the change of variables $\omega=\lambda_d\tau$ yields

$$\int_0^\infty\frac{\sin(\lambda_d\tau)}{\tau}h(\tau)\,d\tau = \int_0^\infty\frac{\sin\omega}{\omega}h(\frac{\omega}{\lambda_d})d\omega\to\underbrace{\int_0^\infty\frac{\sin\omega}{\omega}\,d\omega}_{\pi/2}\ h(0)$$

which formally leads to the Langevin equation

$$\begin{aligned} dX^\tau &= p^\tau\,d\tau,\\ dp^\tau &= \Big(-\lambda'(X^\tau)-\hat K(X^\tau)p^\tau\Big)d\tau+\sqrt{2T\hat K(X^\tau)}\,dW^\tau,\end{aligned} \tag{2.16}$$

as $\lambda_d\to\infty$ where W is the standard Wiener process with independent components in $\mathbb{R}^{3N}$ and $\hat K$ is a constant $3N\times 3N$ matrix; Theorems 3.1 and 3.3 verify a related limit carefully for the Ehrenfest model by giving precise conditions, without using the Debye distribution and the linear coupling. The relation (2.15) defines κ as a rank one matrix so that $\hat K$ becomes positive semi-definite.

This Langevin equation satisfies the ergodic limit

$$\lim_{\mathcal{T}\to\infty}\mathcal{T}^{-1}\int_0^{\mathcal{T}} g(X^t,p^t)\,dt = \int_{\mathbb{R}^{6N}} g(X,p)\mu(dX,dp)$$

for instance if the invariant measure

$$\mu(dX,dp) = \frac{e^{-(|p|^2/2+\lambda(X))/T}\,dX\,dp}{\int_{\mathbb{R}^{6N}} e^{-(|p|^2/2+\lambda(X))/T}\,dX\,dp},$$

exists, λ and $\hat K$ are smooth and $\hat K$ has full rank cf. [23]. The assumption of a full rank positive definite matrix $\hat K$ is hence not satisfied in (2.16), but by considering a slightly perturbed friction matrix, adding ϵI, the friction matrix becomes positive definite; we can also add ϵI, with $0<$

$\epsilon = o(M^{-1/2})$, to the friction matrix K in (3.13) and the approximation results in Theorems 3.1 and 3.3 remain valid, i.e. such a small perturbation of K does not affect the convergence result. The probability measure μ is then also the unique solution of the corresponding Kolmogorov forward equation and it is the heavy particle marginal distribution of the Gibbs distribution

$$\frac{e^{-H_Z(X,p,x,q)/T}\, dX\, dp\, dx\, dq}{\int_{\mathbb{R}^{6N+2J}} e^{-H_Z(X,p,x,q)/T}\, dX\, dp\, dx\, dq}$$

in (2.8). We conclude that sampling the light particles from the Gibbs distribution, conditioned on the heavy particle coordinates, leads time asymptotically to having the heavy particle in the heavy particle marginal of the Gibbs distribution: this fundamental stability and consistency property is unique to the Gibbs distribution, as explained in the next subsection.

The purpose of this work is to apply Zwanzig's derivation to the Ehrenfest model of nuclei and electrons, acting as heavy and light particles, respectively. As we have seen the two models are similar and the difference is that the operator $\tilde H$ is a function of X in the Ehrenfest model. To still be in a setting close to the Zwanzig model, we assume therefore that the temperature is low, so that the ground state dominates over small fluctuations, see (3.16). Three new issues in the Ehrenfest model are the lack of an explicit solution for the electron wave function, the non linear coupling of the wave function to the nuclei and the additional constraint to have a normalized wave function $\langle\psi^0,\psi^0\rangle = 1$, changing the initial distribution. The extension to a non constant operator $\tilde H$ requires different mathematical tools as compared to the case with an explicit solution. Here we will use estimates of the residual in the Kolmogorov backward equation for the Langevin equation evaluated along the Ehrenfest dynamics. A main result is to determine the diffusion/friction matrix without assuming explicit properties of the spectrum of $\tilde H$, but instead use the separation of time scales in the nuclei and electron dynamics and a continuous (possibly random) spectrum assumption of the limit $\tilde H_\infty$ of $\tilde H$ as $M, J \to \infty$.

To explain how this works, the idea is first sketched in the simpler setting of a constant $\tilde H$ and no random matrix, i.e. the Zwanzig model using the identification (2.7). The friction term, from the generalized Langevin equation (2.11), is assumed to satisfy the equalities

$$\begin{aligned}
&\lim_{M\to\infty} 2M^{1/2}\int_0^\tau \Re\langle\cos(\sigma M^{1/2}\tilde H)\frac{d}{d\tau}\Psi_0(X^{\tau-\sigma}), \tilde H\partial_X\Psi_0(X^\tau)\rangle d\sigma\\
&= \lim_{M\to\infty} 2\int_0^{M^{1/2}\tau} \Re\langle\cos(\hat\sigma\tilde H)\frac{d}{d\tau}\Psi_0(X^{\tau-M^{-1/2}\hat\sigma}), \tilde H\partial_X\Psi_0(X^\tau)\rangle d\hat\sigma\\
&= 2\int_0^\infty \lim_{M\to\infty} \Re\langle\underbrace{1_{\hat\sigma\le M^{1/2}\tau}\tilde H\cos(\hat\sigma\tilde H)\frac{d}{d\tau}\Psi_0(X^{\tau-M^{-1/2}\hat\sigma})}_{=:\Gamma_M(\hat\sigma)}, \partial_X\Psi_0(X^\tau)\rangle\, d\hat\sigma\\
&= 2\int_0^\infty \langle\tilde H_\infty\cos(\hat\sigma\tilde H_\infty)\partial_X\Psi_0(X^\tau), \partial_X\Psi_0(X^\tau)\rangle d\hat\sigma\, \dot X^\tau\\
&=: K(X^\tau)\dot X^\tau
\end{aligned} \tag{2.17}$$

where the first equality applies the change of variables $\hat\sigma = M^{1/2}\sigma$, the second follows by dominated convergence, if the sequence Γ_M converges pointwise and is bounded by an $L^1(\mathbb{R})$-function $|\Gamma_M(\hat\sigma)| \le \Gamma(\hat\sigma)$, and the third equality uses the pointwise convergence of

$$\tilde H\cos(\hat\sigma\tilde H)\frac{d}{d\tau}\Psi_0(X^{\tau-M^{-1/2}\hat\sigma}),$$

based on $\sigma\mapsto \frac{d}{d\tau}\Psi_0(X^{\tau-\sigma})$ being continuous, $X^{\tau-M^{-1/2}\hat\sigma}\to X^\tau$, and an assumption on the spectrum of $\tilde H\to\tilde H_\infty$ as $J\to\infty$, for example the Debye distribution (2.14). In (3.14) the last integral is described by the spectrum. Note that K is a $3N \times 3N$ matrix and that the friction term in the Langevin equation (2.16), approximating Ehrenfest dynamics, becomes small $\hat{K} = M^{-1/2}K = \mathcal{O}(M^{-1/2})$.

2.1. **The Gibbs Distribution Derived from Dynamic Stability.** At the heart of Statistical Mechanics is the Gibbs distribution

$$\frac{e^{-H(Y,Q)/T}dYdQ}{\int_{\mathbb{R}^{6N}} e^{-H(Y,Q)/T}dYdQ}$$

for an equilibrium probability distribution of a Hamiltonian dynamical system

$$\begin{aligned} \dot{Y}^\tau &= \partial_Q H(Y^\tau, Q^\tau) \\ \dot{Q}^\tau &= -\partial_Y H(Y^\tau, Q^\tau) \end{aligned} \tag{2.18}$$

in the canonical ensemble of constant number of particles N, volume and temperature T. Every book on Statistical Mechanics gives a motivation of the Gibbs distribution, often based on entropy considerations, cf. [10], [21]. Here we motivate the Gibbs distribution instead from dynamic stability reasons. Consider a Hamiltonian system with light and heavy particles, with position $Y = (X, x)$, momentum $Q = (p, q)$ and the Hamiltonian $H = H_h(X, p) + H_l(X, x, q)$, as in (2.1) with

$$H_h = |p|^2/2 + \lambda(X) \text{ and } H_l = m\big\langle x - \hat{\Psi}(X), \hat{H}\big(x - \hat{\Psi}(X)\big)\big\rangle/2 + \langle q, \hat{H}q\rangle/(2m). \tag{2.19}$$

Assume that it is impractical or impossible to measure and determine the initial data (x^0, q^0) for the light particles. Clearly it is necessary to give some information on the data to determine the solution at a later time. In the case of molecular dynamics it is often sufficient to know the distribution of the particles to determine thermodynamic relevant properties, as e.g. the pressure-law. We saw in Section 2 that if the light particles have an initial probability distribution corresponding the Gibbs distribution conditioned on the heavy particle, then the invariant distribution for the heavy particle is unique (in the limit of the Langevin equation) and given by the Gibbs distribution for the heavy particle

$$\frac{e^{-H_h(X,p)/T}dXdp}{\int_{\mathbb{R}^{6N}} e^{-H_h(X,p)/T}dXdp},$$

which, asymptotically for $H_\ell \ll H_h$, approaches the heavy particle *marginal distribution*

$$\frac{\int_{\mathbb{R}^{2J}} e^{-H/T}dXdpdxdq}{\int_{\mathbb{R}^{6N+2J}} e^{-H/T}dXdpdxdq}$$

of the Gibbs measure for the full Hamiltonian integrated over the light particle phase-space. This stability that an equilibrium distribution of light particles leads to the marginal distribution of the heavy particles holds only for the Gibbs distribution in the sense we shall verify below. Define the set $\mathcal{S}$ of equilibrium measures that have this desired stability and consistency property more precisely as follows:

(S) an equilibrium measure $\nu(X, p, x, q)dXdpdxdq$ belongs to $\mathcal{S}$ if the dynamics of the heavy particles, with the light particles initially distributed according to $\nu(X^0, p^0, \cdot, \cdot)dxdq$ (the equilibrium distribution conditioned on the heavy particle initial data), asymptotically after long time ends up with the heavy particles distributed according to $\int_{\mathbb{R}^2 J} \nu(X, p, x, q)dxdq\, dXdp$ (the heavy particle marginal of the original equilibrium measure).

Consequently the heavy particle behavior over long time, using $\nu \in \mathcal{S}$, is consistent with the assumption to start the light particles with the equilibrium distribution ν. Below we give conditions so that $\mathcal{S}$ only contains the Gibbs measure. It is in fact the uniqueness of the Gibbs initial probability distribution that makes a stochastic model of the dynamics useful: if we would have to seek the initial distribution among a family of many distributions we could not predict the dynamics in a reasonable way.

The consistency property (S) is in an abstract setting related to the study on Gibbs measures for lattice systems in [21], where dynamics and conditioning with respect to initial heavy particles is replaced by invariance for conditioning with respect to complement sets on the lattice, as the lattice growths to infinity.

To derive this uniqueness of the Gibbs density, we consider first all equilibrium densities of the the Hamiltonian dynamics and then use the consistency check (S) to rule out all except the Gibbs density. There are many equilibrium distributions for a Hamiltonian system: the *Liouville equation* (i.e. the Fokker-Planck equation for zero diffusion)

$$\underbrace{\partial_t f(H)}_{=0} + \partial_Y(\partial_Q H f(H)) - \partial_Q(\partial_Y H f(H)) = 0$$

shows that any positive (normalizable) function f, depending only on the Hamiltonian H and not on time, is an invariant probability distribution

$$\frac{f\big(H(Y,Q)\big)dY\,dQ}{\int_{\mathbb{R}^{6N}} f\big(H(Y,Q)\big)dY\,dQ}$$

for the Hamiltonian system (2.18). There may be other invariant solutions which are not functions of the Hamiltonian but these are not considered here. Our basic question is now – which of these functions f have the fundamental property that their light particle distribution generates a unique invariant measure given by the heavy particle marginal distribution? We have seen that the Gibbs distribution is such a solution. Are there other?

Write $H = H_h + H_l$ and assume that the heavy particle Hamiltonian H_h dominates the light particle Hamiltonian H_l, so that

$$\frac{H_l}{H_h} \ll 1. \tag{2.20}$$

In a case based on N heavy particles and J light particles we typically have

$$\frac{H_l}{H_h} = \mathcal{O}(\frac{J}{N})$$

which leads to the condition $J \ll N$; in the next section we study the Ehrenfest model and show that H_l/H_h is small at low temperature. Let

$$-\log f(H) =: g(H)$$

and consider perturbations of the Gibbs distribution in the sense that the function g satisfies for a constant C

$$\begin{aligned} \overline{\lim}_{H\to\infty}\Big|\frac{g''(H)H}{g'(H)}\Big| \le C \\ \overline{\lim}_{H\to\infty}\Big|\frac{g'(H)H}{g(H)}\Big| \le C \end{aligned} \tag{2.21}$$

for instance, any monomial g satisfies (2.21). Taylor expansion yields for some $\alpha \in (0,1)$

$$\begin{aligned} -\log f(H) &= g(H_h + H_l) \\ &= g(H_h) + H_l\big(g'(H_h) + 2^{-1}g''(H_h + \alpha H_l)H_l\big) \end{aligned}$$

and (2.20) and (2.21) imply the leading order term

$$-\log f(H) \simeq g(H_h) + H_l g'(H_h); \tag{2.22}$$

in words, this means that the heavy particle energy dominates and acts as a heat bath to find the distribution for the light particles. Define the constant

$$T = 1/g'\big(H_h(X_0, p_0)\big). \tag{2.23}$$

The light particle distribution is then initially approximately given by

$$\frac{e^{-H_l/T}dxdq}{\int e^{-H_l/T}dxdq}.$$

This initial distribution corresponds to a Gibbs distribution with the temperature

$$T = 1/g'\big(H_h(X_0, p_0)\big)$$

and the derivation of (2.16) (alternatively Theorem (3.3)) leads to the heavy particle equilibrium distribution

$$\frac{e^{-H_h/T}dXdp}{\int e^{-H_h/T}dXdp}. \tag{2.24}$$

The equilibrium density f has by (2.20) and (2.21) the leading order expansion

$$\begin{aligned} -\log f(H) &= g(H_h + H_l) \\ &= g(H_h) + g'(H_h + \alpha H_l)H_l \\ &\simeq g(H_h), \end{aligned}$$

which leads to the heavy particle marginal distribution

$$\frac{e^{-g(H_h)}dXdp}{\int e^{-g(H_h)}dXdp}. \tag{2.25}$$

The consistency requirement to have the heavy particle distribution (2.24) equal to the heavy particle marginal distribution (2.25) implies that

$$g(H_h) = H_h/T.$$

We conclude that the quotient $-H/\log f(H)$ is constant, where $-H/\log f(H) = T$ is called the temperature, and we have derived the Gibbs density $f(H) = e^{-H/T}$. The next section applies this derivation to the Ehrenfest model.

3. Data for the Ehrenfest Model and the Main Results

At low temperature one expects 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* Ψ_0, which solves the electron eigenvalue problem

$$H\Psi_0 = \lambda_0\Psi_0, \tag{3.1}$$

and is normalized $\langle\Psi_0, \Psi_0\rangle = 1$; here $\lambda_0 = \lambda_0(X^t)$ is the smallest eigenvalue of $H = H(X^t)$ in the considered Hilbert space (a subset of $L^2(dx)$). The function

$$\hat\Psi^t := \exp\big(-i\int_0^t \lambda_0(X^s)ds\big)\Psi_0(x; X^t)$$

satisfies

$$i\,\dot{\hat\Psi}^t - H\hat\Psi^t = i\,e^{-i\int_0^t \lambda_0(X^s)ds}\frac{d}{dt}\Psi_0(x; X^t),$$

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 $\|\dot\Psi_0\|$ is small, the function $\hat\Psi$ is an approximate solution to the Ehrenfest dynamics. The approximation $(X^t,\hat\Psi)$, with $\hat\Psi$ replacing $\bar\psi$ in (1.1), is called Born-Oppenheimer molecular dynamics [26, 15] and it approximates observables of the time-independent Schrödinger equation for the electron-nuclei system with accuracy $\mathcal{O}(M^{-1})$, when there is a spectral gap and no caustics, see [27]. The aim of this work is to study molecular dynamics when the electron states $\bar\psi$ are randomly perturbed from the ground state, with a Gibbs distribution at positive temperature. To study the perturbation of the ground state, make the transformation

$$\bar\psi(t,x) = e^{-i\int_0^t \lambda_0(X^s)ds}\psi(t,x) \tag{3.2}$$

which implies that ψ solves the Schrödinger equation

$$i\frac{d}{dt}\psi^t = \underbrace{\big(H(X^t)-\lambda_0(X^t)\big)}_{=:\tilde H(X^t)}\psi^t, \tag{3.3}$$

with the translated Hamiltonian $\tilde H$, and then use the Ansatz

$$\psi = \tilde\gamma_0\Psi_0 + \tilde\psi$$

for some constant $\tilde\gamma_0 \simeq 1$ and expect $\tilde\psi : [0,\infty)\times\mathbb{R}^{3J}\to\mathbb{C}$ to be small. The Ansatz implies that $\tilde\psi$ solves

$$i\frac{d}{dt}\tilde\psi = \tilde H(X^t)\tilde\psi - i\tilde\gamma_0\frac{d}{dt}\Psi_0, \tag{3.4}$$

which yields the solution representation (of the same qualitative form as (2.10) in Zwanzig's model)

$$\tilde\psi^t = \tilde S_{t,0}\tilde\psi^0 - \tilde\gamma_0\int_0^t \tilde S_{t,s}\underbrace{\dot\Psi_0^s}_{\partial_X\Psi_0\dot X}\,ds,$$

with the solution operator $\tilde S$ defined by

$$\tilde S_{t,s}\varphi^s := \varphi^t \tag{3.5}$$

for the solution in the fast time scale

$$i\frac{d}{dt}\varphi^t = \tilde H(X^t)\varphi^t \quad t>s.$$

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 (1.1) for the nuclei, eliminates formally the electrons and generates fluctuations, from stochastic initial data $\tilde\Psi^0$, and friction, through the coupling to $\dot\Psi_0 = \partial_X\Psi_0\dot X$, as explained in Section 2.

3.1. **The Force on Nuclei.** It is convenient to split the force on nuclei in (1.1) into two parts, using the definition of ψ in (3.2),

$$\begin{aligned}\langle\bar\psi,\partial_X H(X)\bar\psi\rangle &= \langle\psi,\partial_X H(X)\psi\rangle\\ &= \langle\psi,\partial_X\lambda_0(X)\psi\rangle + \langle\psi,\partial_X\big(H(X)-\lambda_0(X)\big)\psi\rangle\\ &= \partial_X\lambda_0(X) + \langle\psi,\partial_X\tilde H(X)\psi\rangle.\end{aligned}$$

With the Ansatz $\psi = \tilde\gamma_0\Psi_0 + \tilde\psi$, for a constant $\tilde\gamma_0 \simeq 1$, the second term in the nuclear force becomes

$$\begin{aligned}\langle\psi, \partial_X\tilde H(X)\psi\rangle &= |\tilde\gamma_0|^2\langle\Psi_0, \partial_X\tilde H(X)\Psi_0\rangle\\ &\quad+\langle\tilde\psi, \partial_X\tilde H(X)\tilde\gamma_0\Psi_0\rangle + \langle\tilde\gamma_0\Psi_0, \partial_X\tilde H(X)\tilde\psi\rangle\\ &\quad+\langle\tilde\psi, \partial_X\tilde H(X)\tilde\psi\rangle.\end{aligned}$$

Use $\tilde H\Psi_0 = 0$ and consequently

$$\partial_X\tilde H\Psi_0 + \tilde H\partial_X\Psi_0 = 0 \tag{3.6}$$

to obtain for the first term

$$\begin{aligned}\langle\Psi_0, \partial_X\tilde H(X)\Psi_0\rangle &= -\langle\Psi_0, \tilde H(X)\partial_X\Psi_0\rangle\\ &= -\langle\underbrace{\tilde H\Psi_0}_{=0}, \partial_X\Psi_0\rangle = 0.\end{aligned}$$

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

$$\begin{aligned}\langle\tilde\psi, \partial_X\tilde H(X)\tilde\gamma_0\Psi_0\rangle + \langle\tilde\gamma_0\Psi_0, \partial_X\tilde H(X)\tilde\psi\rangle &= 2\Re\langle\tilde\psi, \partial_X\tilde H(X)\tilde\gamma_0\Psi_0\rangle\\ &= -2\Re\langle\tilde\psi, \tilde H(X)\tilde\gamma_0\partial_X\Psi_0\rangle,\end{aligned} \tag{3.7}$$

and we have obtained the forcing

$$\langle\bar\psi, \partial_X H(X)\bar\psi\rangle = \partial_X\lambda_0(X) - 2\Re\langle\tilde\psi, \tilde H(X)\tilde\gamma_0\partial_X\Psi_0(X)\rangle + \langle\tilde\psi, \partial_X\tilde H(X)\tilde\psi\rangle. \tag{3.8}$$

Lemma 4.1 shows that the third term, $\langle\tilde\psi, \partial_X\tilde H\tilde\psi\rangle$, is negligible small at low temperature.

3.2. **Stochastic Electron Initial Data.** The next step, to determine the stochastic initial data for $\tilde\psi$, applies the same reasoning as for the Zwanzig model in Sections 2.1. Similar to the partition of the Zwanzig-Hamiltonian in (2.19), split the Ehrenfest-Hamiltonian into

$$H_E = \underbrace{\frac{|p|^2}{2} + \lambda_0(X)}_{H_h} + \underbrace{\langle\tilde\psi, \tilde H\tilde\psi\rangle/\langle\psi,\psi\rangle}_{H_l}$$

and assume that the "light" particle Hamiltonian H_l, associated to the perturbation of the wave function, is much smaller than the heavy nuclei Hamiltonian H_h so that

$$\frac{H_l}{H_h} \ll 1,$$

as implied by (3.16). Note that it is the Hamiltonian $H_E = |p|^2/2 + \lambda_0(X) + C\langle\tilde\psi, \tilde H(X)\tilde\psi\rangle$ that generates a Hamiltonian system, for any constant C, and not the normalized form $|p|^2/2 + \lambda_0(X) + \langle\tilde\psi, \tilde H(X)\tilde\psi\rangle/\langle\psi,\psi\rangle$; since the Hamiltonian system conserves the L^2-norm $\langle\psi,\psi\rangle$ along solution paths, the normalization can be put in the equation for p by taking $C = 1/\langle\psi,\psi\rangle$, to get the correct forces. In this section we therefore first seek the distribution of the non normalized $\tilde\psi$ and then when this distribution is found we normalize. Diagonalize and make a cut-off allowing only the first J modes

$$\begin{aligned}\tilde\psi^0 &= \sum_{j=1}^J \gamma_j\Psi_j(X^0),\\ \psi^0 &= \gamma_0\Psi_0(X^0) + \tilde\psi^0,\\ C\langle\psi, \tilde H(X^0)\psi\rangle &= C\langle\tilde\psi, \tilde H(X^0)\tilde\psi\rangle = C\sum_{j=1}^J \tilde\lambda_j(X^0)|\gamma_j|^2,\end{aligned} \tag{3.9}$$

with the orthonormal eigenfunctions $\Psi_j(X^0)$ and eigenvalues $\tilde\lambda_j(X^0)$ to $\tilde H(X^0)$. We see that the initial light particle probability density satisfies

$$e^{-C\langle\tilde\psi,\tilde H\tilde\psi\rangle/T}d\tilde\psi^r d\tilde\psi^i = e^{-\sum_{j=1}^J \frac{C\tilde\lambda_j}{T}(|\gamma_r^j|^2+|\gamma_i^j|^2)}d\gamma_1^r\dots d\gamma_J^r d\gamma_1^i\dots d\gamma_J^i,$$

since the orthogonal transformation from $(\tilde\psi^r,\tilde\psi^i)$ to $(\gamma_1^r,\dots,\gamma_J^r,\gamma_1^i,\dots,\gamma_J^i)$ has the Jacobian determinant equal to one, as in (2.9). Consequently,

$$\begin{aligned}&\{\gamma_j \mid j=1,\dots J\} \text{ are independent, with independent real and imaginary parts}\\ &\text{normal } \mathcal{N}\big(0,T/(\tilde\lambda_j(X^0)C)\big) \text{ distributed.}\end{aligned} \tag{3.10}$$

Assume now that the ground state energy dominates, i.e. T is chosen low enough so that

$$\frac{\sum_{j=1}^J|\gamma_j|^2}{|\gamma_0|^2}=o(1), \tag{3.11}$$

and take

$$C^{-1}=\sum_{j=0}^J|\gamma_j|^2\simeq|\gamma_0|^2.$$

This means that the forces generated by $\psi=\tilde\psi+\gamma_0\Psi_0=\sum_{j=0}^J\gamma_j\Psi_j(X^0)$ should be normalized by dividing by $(\sum_{j=0}^J|\gamma_j|^2)=\langle\psi,\psi\rangle$. Since there are very many eigenvalues this is almost the same as dividing by the expected value

$$C^{-1}=\sum_{j=0}^J|\gamma_j|^2\simeq\mathbb{E}_\gamma[\langle\psi,\psi\rangle]=|\gamma_0|^2+\sum_{j=1}^J\frac{T}{C\tilde\lambda_j(X^0)},$$

which is solvable for the requirement

$$\sum_{j=1}^J T/\tilde\lambda_j\ll 1, \tag{3.12}$$

included in assumption (3.11). The normalized wave initial function becomes

$$\sum_{j=0}^J\tilde\gamma_j\Psi_j(X^0):=\sum_{j=0}^J\frac{\gamma_j}{(\sum_{k=0}^J|\gamma_k|^2)^{1/2}}\Psi_j(X^0)\simeq\sum_{j=0}^J\frac{\gamma_j}{|\gamma_0|}\Psi_j(X^0)$$

where in particular $\tilde\gamma_0\simeq 1$.

The size of J is chosen so that the condition $H_l/H_h\ll 1$ holds, implying by (2.22) the light particles to be approximately Gibbs distributed: since the normalized light and heavy particle Hamiltonian satisfies

$$\mathbb{E}_\gamma[H_l]\simeq\mathbb{E}_\gamma[\sum_{j=1}^J\tilde\lambda_j|\gamma_j|^2/|\gamma_0|^2]\simeq TJ$$

and $H_h=|p|^2/2+\lambda_0\sim N$, the condition $H_l/H_h\ll 1$ holds provided $N\gg TJ$.

To derive the Langevin equation the spectrum of $\tilde H$ will be used and the main assumption is that the distribution of eigenvalues $\tilde\lambda_j$ approaches a continuum limit as $J\to\infty$. Such a continuum can be achieved in at least two different ways: either each event of $\tilde H(X)$ generates a continuum density in a deterministic way, as in (2.17), or the spectrum of $\tilde H(X)$ becomes sensitive towards small perturbations of X and generates a random matrix $\tilde H_\infty$ which has a probability distribution with a continuous distribution of eigenvalues. The stochastic continuum limit based on random matrices can be formulated as follows. Assume that $\tilde H$ satisfies the following limit, cf. (2.11) and (2.17),

$$
\begin{aligned}
&K_{mn}(X^\tau)\\
&:= \mathbb{E}_\gamma\big[2\lim_{M,J\to\infty}\int_0^{\tau M^{1/2}} \Re\langle \tilde S_{\hat\tau,\hat\tau-\hat\sigma}\partial_{X_n}\Psi_0(X^{\tau-\hat\sigma M^{-1/2}}), \tilde H(X^\tau)\partial_{X_m}\Psi_0(X^\tau)\rangle d\hat\sigma | X^\tau\big],
\end{aligned}
\tag{3.13}
$$

so that the limit K forms $3N\times 3N$ matrix. In the fast time scale $\hat\sigma$ the nuclei positions $\hat\sigma \mapsto X^{\tau-M^{-1/2}\hat\sigma}$ change slowly so that $X^{\tau-M^{-1/2}\hat\sigma}\to X^\tau$, as $M\to\infty$, and the solution operator would satisfy $\tilde S_{\hat\tau,\hat\tau-\hat\sigma}\to e^{-i\hat\sigma\tilde H(X^\tau)}$, for $\hat\sigma\sim 1$, if $\tilde H$ would be independent of J and M, so that by dominated convergence as in (2.17) the integral would satisfy

$$
\begin{aligned}
&\int_0^{\tau M^{1/2}} \Re\langle \tilde S_{\hat\tau,\hat\tau-\hat\sigma}\Psi_0'(X^{\tau-\hat\sigma M^{-1/2}}), \tilde H\Psi_0'(X^\tau)\rangle d\hat\sigma\\
&\qquad \to \int_0^\infty \langle\cos(\hat\sigma\tilde H)\Psi_0'(X^\tau), \tilde H\Psi_0'(X^\tau)\rangle d\hat\sigma.
\end{aligned}
$$

We may instead assume that $\tilde H(X^{\tau-M^{-1/2}\hat\sigma})$ does not converge but its average along the solution path becomes random in the sense that the solution operator $\tilde S_{\hat\tau,\hat\tau-\hat\sigma}$ tends to $e^{-i\hat\sigma\tilde H_\infty}$, where $\tilde H_\infty$ is a random matrix, as $M\to\infty$. That is, as M and J tend to infinity, the solution operator's average of the matrix $\tilde H(X^{\tau-M^{-1/2}\hat\sigma})$ becomes a random matrix $\tilde H_\infty(X^\tau)$ and its probability distribution has a continuum spectral decomposition with eigenvalues $\tilde\lambda_j^\infty$ arbitrary close. An assumption of such a random matrix seems consistent with our model:

- assume that the eigenvalues $\tilde\lambda_n$ are a distance of order c apart, where $c>0$ may be small and tending to zero as J tends to infinity (giving the continuum of eigenvalues in the deterministic case);
- in the limit (3.13) a small time step of size $M^{-1/2}$ makes an increment $M^{-1/2}p$ in the position X which yields an increment of size $\mathcal{O}(M^{-1/2})$ in each matrix component $\tilde H_{ij}$;
- since the size J of the matrix $\tilde H(X)$ is at the same time growing, the eigenvalue gaps could asymptotically be filled with a continuous density, if J is sufficiently large.

Assuming the existence of the following spectral representation, the limit can by the Fourier transform, $\mathcal{F}h(\hat\sigma) := \int_{\mathbb{R}} e^{i\hat\sigma\lambda}h(\lambda)d\lambda$ and its inverse $h(\lambda) = 2\pi\int_{\mathbb{R}} e^{-i\hat\sigma\lambda}\mathcal{F}h(\hat\sigma)d\hat\sigma$, be written

$$
\begin{aligned}
K(X^\tau) &= \int_0^\infty \underbrace{\int_0^\infty 2\cos(\hat\sigma\lambda)\Gamma(\lambda,X^\tau)d\lambda}_{=\mathcal{F}\,\Gamma(\hat\sigma,X^\tau)}\, d\hat\sigma\\
&= \int_0^\infty \mathcal{F}\,\Gamma(\hat\sigma,X^\tau)d\hat\sigma\\
&= \pi\,\Gamma(0,X^\tau),
\end{aligned}
\tag{3.14}
$$

where $\Gamma \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \cap \mathcal{C}(\mathbb{R})$ is defined by a spectral decomposition, based on the eigenfunctions $\{\Psi_j^\infty\}_{j=1}^\infty$ and eigenvalues $\{\tilde\lambda_j\}_{j=1}^\infty$ of $\tilde H_\infty$,

$$
\begin{aligned}
&\mathbb{E}[\langle \cos(\hat\sigma \tilde H_\infty)\partial_{X_n}\Psi_0(X), \tilde H_\infty(X)\partial_{X_m}\Psi_0(X)\rangle \mid X] \\
&= \sum_{j=1}^\infty \mathbb{E}[\cos(\hat\sigma\tilde\lambda_j)\langle \partial_{X_n}\Psi_0(X), \Psi_j^\infty\rangle \langle \tilde H_\infty \partial_{X_m}\Psi_0(X), \Psi_j^\infty\rangle \mid X] \\
&= \int_{\mathbb{R}} \cos(\hat\sigma\lambda) \underbrace{\sum_{j\in\{n:\,\tilde\lambda_n\in[\lambda,\lambda+d\lambda)\}} \mathbb{E}[\langle \partial_{X_n}\Psi_0, \Psi_j^\infty\rangle\langle \tilde H_\infty\partial_{X_m}\Psi_0, \Psi_j^\infty\rangle \mid X]}_{=:\Gamma_{nm}(\lambda,X)d\lambda},
\end{aligned}
$$

and the expected value is with respect to the random matrix $\tilde H_\infty$. An indication of this size of J is provided by Wigners Semi-Circle Law saying that symmetric matrices of size J, with independent identically distributed entries with mean zero, variance J^{-1} and bounded moments, have the density of eigenvalues semi-circle distributed asymptotically as $J \to \infty$, i.e. the continuum density is $\sqrt{4-\rho^2}1_{|\rho|\le 2}\, d\rho/(2\pi)$, cf. [1]. Since in our case the variance of the matrix elements is $\mathcal{O}(M^{-1})$, we could chose $J \sim M$ to fill gaps of size one, if the matrix elements were independent identically distributed; this would then fill the smaller gaps of size c. It is a challenge to really determine the distributions for the matrix elements.

For assumption (3.13) to make sense the eigenvalue distribution must approach a continuum density, so that the integral kernel decays in time and $\Gamma(0, X)$ is well defined; to have a positive limit it is necessary that the density of states does not vanish around the origin, which excludes a spectral gap in $\tilde H_\infty$ around the origin. The condition $\Gamma(0, X)$ being positive semi-definite can also imply that $\langle \Psi_0', \Psi_0'\rangle$ is infinite – Section 4.6 gives a motivation for such a setting based on the Thomas-Fermi model.

3.2.1. *Slow Nuclear Dynamics.* The Ehrenfest dynamics (1.1), (1.2) can be written as a Hamiltonian system in the slow time scale

$$
\begin{aligned}
\dot X^\tau &= p^\tau \\
\dot p^\tau &= -\partial_X\lambda_0 - \langle \phi^\tau, \partial_X \tilde H(X^\tau)\phi^\tau\rangle/\langle\phi,\phi\rangle \\
\frac{i}{M^{1/2}}\dot\phi^\tau &= \tilde H(X^\tau)\phi^\tau,
\end{aligned}
$$

with the Hamiltonian

$$
\frac{|p|^2}{2} + C\langle \phi, \tilde H(X)\phi\rangle + \lambda_0(X),
$$

using the constant $C = M^{1/2}/2 = 1/\langle\phi,\phi\rangle$ and letting $\psi^{\hat\tau} = \phi^\tau/\langle\phi^\tau,\phi^\tau\rangle^{1/2}$ for the initial data $\psi^0 = \sum_{n=0}^J(\sum_{k=0}^J|\gamma_k|^2)^{-1}\gamma_n\Psi_n(X^0)$; here we change to the simpler notation (X^τ, p^τ) for position and momentum in the slow time scale, although to be consistent with (1.1)-(1.2) it should have been $(X^{\hat\tau}, p^{\hat\tau})$. The change to the variable ψ leads to

$$
\begin{aligned}
\frac{dX^\tau}{d\tau} &= p^\tau, \\
\frac{dp^\tau}{d\tau} &= -\langle \psi^\tau, \partial_X H(X^\tau)\psi^\tau\rangle, \qquad (3.15)\\
\frac{i}{M^{1/2}}\frac{d\psi^\tau}{d\tau} &= \tilde H(X^\tau)\psi^\tau.
\end{aligned}
$$

3.3. **The Main Result.** Let W^τ denote the standard Brownian process (at time τ) in $\mathbb{R}^{3N}$ with independent components. To simplify the notation, 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 $\psi(x,X) = \mathcal{O}(M^{-\alpha})$ also for complex valued functions, meaning that $|\psi(x,X)| = \mathcal{O}(M^{-\alpha})$ holds uniformly in x and X.

We will use the following assumptions for the electron eigenvalues $\tilde\lambda_j$ and electron Hamiltonian $\tilde H$

$$\begin{aligned} T\sup_X \sum_{j=1}^J \frac{|\partial_X\tilde\lambda_j(X)|_{\ell^1(\mathbb{R}^{3N})}}{\tilde\lambda_j(X)} &= o(M^{-1/2}),\\ \sup_X \big\||\tilde H\partial_X\Psi_0(X)|_{\ell^1(\mathbb{R}^{3N})}\big\|_{L^1(dx)} &= \mathcal{O}(1),\\ \sup_X |\partial_{X_k}\tilde H(X)| + \sup_X |\partial_{X_jX_k}\tilde H(X)| &= \mathcal{O}(1),\\ \sigma \mapsto\ & \tilde H(X^\sigma) \text{ is real analytic on } (0,\infty),\\ |\tilde\lambda_n - \tilde\lambda_m| &> c \text{ for } n\neq m \text{ and } c^{-1} = o(M^{1/5}),\\ \frac{\sum_{j=1}^J|\gamma_j|^2}{|\gamma_0|^2} &= o(1). \end{aligned} \tag{3.16}$$

The conditions are motivated and used as follows. The first condition in (3.16) implies that the quadratic forcing term $\langle\tilde\psi, \tilde H'\tilde\psi\rangle$ becomes negligible. The fourth and fifth assumption (that the electron eigenvalues do not cross conditions along solution paths) are used in the Born-Oppenheimer approximation Lemma 4.3. The last condition used for the initial wave function distribution in (3.12) means that T is small enough so that

$$1 \simeq |\gamma_0|^2 \gg \sum_{j=1}^J T\tilde\lambda_j^{-1} \simeq \sum_{j=1}^J |\gamma_j|^2, \tag{3.17}$$

which implies that the condition $H_l/H_h \ll 1$ holds and sampling from the Gibbs distribution becomes the only option, in the sense of Section 2.1. If the density of states would be uniform and $\tilde\lambda_1 = J^{-1}$, we would need $T = o(1/\log J)$ since (3.17) yields

$$1 \gg T\sum_{j=1}^J \tilde\lambda_j^{-1} \simeq T\int_{J^{-1}}^J \frac{d\lambda}{\lambda} = 2T\log J,$$

and in general we expect that (3.17) implies that $T \to 0$ as $J \to \infty$.

As for the Zwanzig model studied in [28, 20], [16], the following result compares expected values of Hamiltonian dynamics, having stochastic initial data γ for the light particles, with Langevin dynamics where the stochasticity enters through the Wiener process W. We use the notation $\mathbb{E}_z[w]$ for the expected value of w with respect to the distribution generated by the random z, the eigenvalue λ_0 denotes the ground state electron energy (3.1) of H with normalized ground state Ψ_0 and $\tilde\lambda_j$ are the translated eigenvalues of $H - \lambda_0 =: \tilde H$.

Theorem 3.1. *Assume that condition* (3.16) *and the friction limit* (3.13) *hold, the temperature* T *is low enough to satisfy* $T\sum_{j=1}^J \tilde\lambda_j^{-1} = o(1)$, *and the Ehrenfest electron initial data is given by* (3.9)-(3.10), *then the Itô Langevin dynamics*

$$\begin{aligned} d\dot X_L^\tau &= -\partial_{X_L}\lambda_0(X_L^\tau)d\tau - M^{-1/2}K(X_L^\tau)\dot X_L^\tau d\tau + \sqrt{2TM^{-1/2}}K^{1/2}(X_L^\tau)dW^\tau,\\ \dot X_L^\tau &:= \frac{dX_L^\tau}{d\tau}, \quad 0<\tau<\mathcal{T}, \end{aligned} \tag{3.18}$$

with deterministic initial data $X_L^0 = X$ and $p_L^0 = p$, approximates Ehrenfest dynamics (3.15) with accuracy

$$\begin{aligned}\Big|\mathbb{E}_\gamma\big[g(X^\mathcal{T}, p^\mathcal{T}; X, p) \mid X^0 = X, \dot X^0 = p\big]\\ - \mathbb{E}_W\big[g(X_L^\mathcal{T}, \dot X_L^\mathcal{T}; X, p) \mid X_L^0 = X, \dot X_L^0 = p\big]\Big| = o(M^{-1/2}),\end{aligned} \tag{3.19}$$

as $M \to \infty$, for any bounded function $g : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R}$, provided the Langevin expected value function

$$u(Y, q, \tau; \mathcal{T}) := \mathbb{E}_W[g(X_L^\mathcal{T}, \dot X_L^\mathcal{T}; X, p) \mid X_L^\tau = Y,\ \dot X_L^\tau = q]$$

has bounded derivatives of order one and two along the Ehrenfest solution

$$\begin{aligned}&\int_0^\mathcal{T} \sum_j |\partial_{p_j} u(X^\sigma, p^\sigma, \sigma; \mathcal{T})|\, d\sigma = \mathcal{O}(1),\\ &\int_0^\mathcal{T} \sum_{j,k} \big(|\partial_{p_j p_k} u(X^\sigma, p^\sigma, \sigma; \mathcal{T})| + |\partial_{X_j p_k} u(X^\sigma, p^\sigma, \sigma; \mathcal{T})|\big)\, d\sigma = \mathcal{O}(1).\end{aligned} \tag{3.20}$$

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 D can be studied

$$D = \mathbb{E}_W[\underbrace{(6N\mathcal{T})^{-1}|X_L^\mathcal{T} - X_L^0|^2}_{=g(X_L^\mathcal{T}, \dot X_L^\mathcal{T}; X_L^0, p_L^0)} \mid X^0].$$

Since (X_L^0, p_L^0) is deterministic and fixed for each path, we simplify the notation and write $g(Y, q)$ instead of $g(Y, q; X_L^0, p_L^0)$.

Remark 3.2. It is well known that not all functions $g(X_L^\mathcal{T}, \dot X_L^\mathcal{T})$ (such as $g(X, p) = \max_j |p_j|$) are accurately computable in a molecular dynamics simulation with many particles; assumption (3.20) restricts the study to observables g that are stable with respect to perturbations in the initial data $(X_L^0, \dot X_L^0)$ in the norm ℓ^1. We consider expected values $E[g(X_L^\mathcal{T}, p_L^\mathcal{T})]$ where typically $g = \sum_{j=1}^N g_j/N$ is a mean over (particle related) real-valued functions g_j with $g_j(X, p) = \mathcal{O}(1)$ and

$$\sum_{n=1}^N |\partial_{X_n} g(X, p)| = \sum_{n=1}^N |\sum_{j=1}^N \frac{\partial_{X_n} g_j(X, p)}{N}| = \mathcal{O}(1).$$

Here, for each derivative ∂_{X_n}, mainly finitely many j contribute to the sum. Similarly we assume that $\sum_{n=1}^N |\partial_{X_n} u(X, p, \sigma)| = \mathcal{O}(1)$. This means that g measures global properties, related to thermodynamic quantities, e.g. the diffusion D with $g = \sum_{j=1}^N |X_j - X_j^0|^2/(6N\mathcal{T})$.

Assumption (3.20) limits $\mathcal{T}$ to bounded times, since the correlation time is expected to grow as $M^{1/2}$, explained as follows. Let K be constant, then $\dot X_L =: p_L$ solves an Ohrnstein-Uhlenbeck equation, which means that

$$\begin{aligned}p_L^\tau = (2TM^{-1/2}K)^{1/2} \int_0^\tau e^{M^{-1/2}(\sigma-\tau)K} dW^\sigma\\ - \int_0^\tau \lambda_0'(X^\sigma) e^{M^{-1/2}(\sigma-\tau)K} d\sigma + p^0 e^{-M^{-1/2}K\tau}\end{aligned}$$

and the dependence of p^τ on p^σ decays exponentially $e^{M^{-1/2}(\sigma-\tau)K}$. If also $\lambda_0' = 0$, the process p^τ is Gaussian with mean zero and covariance

$$2Te^{-M^{-1/2}K|\tau-\sigma|} + e^{-M^{-1/2}K(\tau+\sigma)}(\mathbb{E}[|p^0|^2] - 2T)$$

and the corresponding $\partial_p u(\cdot,\sigma;\tau)$ decays exponentially $e^{-M^{-1/2}|\tau-\sigma|}$, giving the correlation length

$$\int_0^\tau \partial_p u(\cdot,\sigma;\tau)d\sigma = \mathcal{O}(M^{1/2});$$

Section 4.5 presents a motivation for this property also in the case with a general force $\lambda_0'(X)$, using the theory of large deviations for the rare events of escapes from equilibria at low temperature.

Theorem 3.3. *Assume that condition* (3.20) *in Theorem 3.1 is replaced by*

$$\lim_{\tau\to\infty}\int_0^\tau |\mathcal{D}u(X^\sigma,p^\sigma,\sigma;\tau)|_{\ell^1}\, d\sigma = \mathcal{O}(M^{1/2}), \quad \mathcal{D} := \partial_p, \partial_{pp}, \partial_{Xp}, \tag{3.21}$$

i.e. the time-correlation length with respect to sensitivity in p is at most of size $M^{1/2}$, then Langevin dynamics approximates long time observables of Ehrenfest dynamics

$$\lim_{\mathcal{T}\to\infty}\mathcal{T}^{-1}\Big|\int_0^{\mathcal{T}} \mathbb{E}_\gamma\big[g(X^\tau,p^\tau)\big] - \mathbb{E}_W\big[g(X_L^\tau,\dot X_L^\tau)\big]\ d\tau\Big| = o(1) \quad \textit{as } M\to\infty. \tag{3.22}$$

The combination of the two theorems shows that this Langevin dynamics is in a sense an accurate approximation of Ehrenfest dynamics for both short and long time. The small dissipation term $M^{-1/2}Kp_L$ in (3.18) is visible in the convergence rate $o(M^{-1/2})$; the small diffusion parameter $TM^{-1/2}K = o(M^{-1/2})$ can be detected in the long time convergence rate $o(1)$, since the invariant measure is

$$\frac{e^{-(|p|^2/2+\lambda_0(X))/T}dXdp}{\int_{\mathbb{R}^{3N+J}} e^{-(|p|^2/2+\lambda_0(X))/T}dXdp}$$

and this holds if and only if the diffusion coefficient is $(2TM^{-1/2}K)^{1/2}$, provided the dissipation term is $M^{-1/2}Kp_L$. That is, the Langevin dynamics (3.18) satisfies *Einstein's fluctuation-dissipation* result: the square of the diffusion coefficient is the dissipation coefficient times twice the temperature, as in (2.13).

Remark 3.4. The Hamiltonian dynamics generated from the time-independent Schrödinger equation, studied in [27], is closely related to the Ehrenfest dynamics where the Hamiltonian H_E is essentially replaced by the slightly perturbed Hamiltonian

$$H_E + (2M)^{-1}G(X)\sum_n \Delta_{X_n}(\langle\phi,\phi\rangle/G(X))$$

for a certain bounded function $G(X)$, which is different for caustic and non caustic states. Therefore, it seems likely that similar approximation results hold when the Langevin dynamics is compared to this Schrödinger Hamiltonian dynamics.

4. Error Estimates

4.1. **The Dynamics.** We shall approximate the nuclei motion in the Ehrenfest dynamics by (X_L,p_L) defined from the Ito-Langevin dynamics

$$\begin{aligned} \dot X_L &= p_L \\ \dot p_L &= -\partial_X\lambda_0(X_L) - M^{-1/2}K(X_L)p_L + (2TM^{-1/2})^{1/2}K^{1/2}(X_L)\dot W. \end{aligned} \tag{4.1}$$

To simplify the analysis of the coupling between (X,p) and the thermal fluctuations induced by $\tilde\gamma_n$, introduce the electron wave functions $\{\tilde\psi_n \mid n=0,1,\ldots\}$ which initially are eigenfunctions and solve

$$\frac{i}{M^{1/2}}\dot{\tilde\psi}_n^\tau = \tilde H(X^\tau)\tilde\psi_n^\tau, \quad \tilde\psi_n^0 = \Psi_n(X^0), \tag{4.2}$$

to obtain the normalized description

$$\psi^{\hat\tau} = \sum_{n=0}^{J} \frac{\gamma_n}{\left(\sum_{k=0}^{J}|\gamma_k|^2\right)^{1/2}}\tilde\psi_n =: \sum_{n=0}^{J}\tilde\gamma_n\tilde\psi_n \tag{4.3}$$

where γ_n, defined in (3.9), are independent normal distributed with mean zero and variance $T/(C\tilde\lambda_n(X^0)) \simeq (|\gamma_0|T)/\tilde\lambda_n$ for $n = 1,\dots J$, using (3.12) and (3.11). The Schrödinger dynamics (4.2) shows that $\{\tilde\psi_n^\tau \mid n = 0, 1, \dots\}$ forms an orthogonal set

$$\begin{aligned}\tfrac{d}{d\tau}\langle\tilde\psi_n,\tilde\psi_m\rangle &= \langle -iM^{1/2}\tilde H\tilde\psi_n,\tilde\psi_m\rangle + \langle\tilde\psi_n, -iM^{1/2}\tilde H\tilde\psi_m\rangle\\ &= \langle\tilde\psi_n, iM^{1/2}\tilde H\tilde\psi_m\rangle + \langle\tilde\psi_n, -iM^{1/2}\tilde H\tilde\psi_m\rangle = 0,\end{aligned} \tag{4.4}$$

since the initial data $\{\tilde\psi_n^0 \mid n = 0, 1, \dots\}$ is orthogonal.

4.2. **Proof of the Theorem.** Define for the given observable $g : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R}$ and the Langevin dynamics (X_L^τ, p_L^τ), in (4.1), the expected value function

$$u(y,\tau) := \mathbb{E}_W[g(X_L^{\mathcal T}, p_L^{\mathcal T}) \mid (X_L^\tau, p_L^\tau) = y],$$

which solves the Kolmogorov backward equation

$$\begin{gathered}\partial_\tau u + p\cdot\partial_X u - \big(\partial_X\lambda_0(X) + M^{-1/2}K(X)p\big)\cdot\partial_p u\\ +\sum_{n=1}^{3N}\sum_{n=1}^{3N} M^{-1/2}TK_{mn}(X)\partial_{p_m p_n}u = 0 \quad \tau < \mathcal T\\ u(\cdot,\mathcal T) = g.\end{gathered} \tag{4.5}$$

The goal is to analyze the error $\mathbb{E}_\gamma[g(X^{\mathcal T}, p^{\mathcal T})] - \mathbb{E}_W[g(X_L^{\mathcal T}, p_L^{\mathcal T})]$ of the heavy nuclei in the Ehrenfest dynamics approximated by the Langevin dynamics, for given deterministic initial data $X^0 = X_L^0, p^0 = p_L^0$. This error can be written as the residual of the Langevin Kolmogorov solution (4.5) along the Ehrenfest dynamics (X^τ, p^τ)

$$\begin{aligned}&\mathbb{E}_\gamma[g(X^{\mathcal T}, p^{\mathcal T}) \mid X^0 = X, p^0 = p] - \mathbb{E}_W[g(X_L^{\mathcal T}, p_L^{\mathcal T}) \mid X_L^0 = X, p_L^0 = p]\\ &= \mathbb{E}_\gamma[u(X^{\mathcal T}, p^{\mathcal T}, \mathcal T) \mid X^0 = X, p^0 = p] - u(X, p, 0)\\ &= \mathbb{E}_\gamma[u(X^{\mathcal T}, p^{\mathcal T}, \mathcal T) - u(X^0, p^0, 0) \mid X^0 = X, p^0 = p].\end{aligned}$$

The expected value in right hand side will be written as an integral over time where the assumption (3.20) makes the integral bounded. Since all remaining expected values are with respect to the initial data γ we simplify by writing $\mathbb{E} = \mathbb{E}_\gamma$.

Telescoping cancelation, the Kolmogorov equation and the nuclei forces

$$\dot p = -\partial_X\lambda_0(X) + 2\Re\langle\tilde\psi, \tilde H(X)\tilde\gamma_0\partial_X\Psi_0(X)\rangle - \langle\tilde\psi, \partial_X\tilde H(X)\tilde\psi\rangle$$

in (3.8) imply

$$
\begin{aligned}
&\mathbb{E}[u(X^{\mathcal{T}},p^{\mathcal{T}},\mathcal{T})-u(X^0,p^0,0)\mid X^0,p^0]=\int_0^{\mathcal{T}}\mathbb{E}[\,du(X^\tau,p^\tau,\tau)\mid X^0,p^0]\\
&=\int_0^{\mathcal{T}}\mathbb{E}[\partial_\tau u(X^\tau,p^\tau,\tau)+\dot X^\tau\cdot\partial_X u(X^\tau,p^\tau,\tau)+\dot p^\tau\cdot\partial_p u(X^\tau,p^\tau,\tau)\mid X^0,p^0]\,d\tau\\
&=\int_0^{\mathcal{T}}\mathbb{E}\Big[\underbrace{(\dot X^\tau-p^\tau)}_{=0}\cdot\partial_X u(X^\tau,p^\tau,\tau)\\
&\qquad+\big(\dot p^\tau+\lambda_0'(X^\tau)+M^{-1/2}K(X^\tau)p^\tau\big)\cdot\partial_p u(X^\tau,p^\tau,\tau)\\
&\qquad-TM^{-1/2}K(X^\tau)\partial_{pp}u(X^\tau,p^\tau,\tau)\mid X^0,p^0\Big]\,d\tau\\
&=\int_0^{\mathcal{T}}\mathbb{E}\Big[\big(2\Re\langle\tilde\psi,\tilde\gamma_0\tilde H\Psi_0'\rangle-\langle\tilde\psi,\tilde H'\tilde\psi\rangle-M^{1/2}K(X^\tau)p^\tau\big)\cdot\partial_p u(X^\tau,p^\tau,\tau)\\
&\qquad-TM^{1/2}K(X^\tau)\partial_{pp}u(X^\tau,p^\tau,\tau)\mid X^0,p^0\Big]\,d\tau.
\end{aligned}
\tag{4.6}
$$

Consider the solution (X,p,ψ) to the Ehrenfest dynamics (3.15) and define for the Schrödinger equation

$$
\begin{aligned}
i\dot\varphi^\tau&=M^{1/2}\tilde H(X^\tau)\varphi^\tau,\quad \tau\ge\sigma\\
\varphi^\sigma&=w
\end{aligned}
$$

the solution operator in the slow time scale

$$S_{\tau,\sigma}w:=\varphi^\tau.$$

The definition $\psi=\tilde\psi+\tilde\gamma_0\Psi_0$ yields the representation

$$
\begin{aligned}
\tilde\psi^t&=-\tilde\gamma_0\int_0^\tau S_{\tau,\sigma}\dot\Psi_0(X^\sigma)d\sigma+S_{\tau,0}\tilde\psi^0\\
&=-\tilde\gamma_0\int_0^\tau S_{\tau,\sigma}\dot\Psi_0(X^\sigma)d\sigma+\sum_{n>0}\tilde\gamma_n\tilde\psi_n^\tau
\end{aligned}
\tag{4.7}
$$

which implies

$$\Re\langle\tilde\psi,\tilde H\Psi_0'\rangle=-|\tilde\gamma_0|^2\Re\langle\int_0^\tau S_{\tau,\sigma}\dot\Psi_0(X^\sigma)d\sigma,\tilde H\Psi_0'\rangle+\Re\langle\sum_{n=1}^J\tilde\gamma_n\tilde\psi_n^\tau,\tilde H\Psi_0'\rangle.$$

Lemma 4.1. *There holds*

(4.8)

$$
\begin{aligned}
&\lim_{M\to\infty}-M^{1/2}\mathbb{E}\big[2|\tilde\gamma_0|^2\Re\langle\int_0^\tau S_{\tau,\sigma}\dot\Psi_0(X^\sigma)d\sigma,\tilde H\Psi_0'(X^\tau)\rangle\mid X^\tau,p^\tau\big]=K(X^\tau)p^\tau,\\
&\lim_{M\to\infty}M^{1/2}\mathbb{E}\big[2\Re\langle\sum_{n=1}^J\tilde\gamma_n\tilde\psi_n^\tau,\tilde H\Psi_0'(X^\tau)\rangle\cdot\partial_p u(X^\tau,p^\tau,\tau)\,|\,X^\tau,p^\tau\big]=TK(X^\tau)\partial_{pp}u(X^\tau,p^\tau,\tau),\\
&\lim_{M\to\infty}M^{1/2}\mathbb{E}\big[\langle\tilde\psi^\tau,\tilde H'\tilde\psi^\tau\rangle\cdot\partial_p u(X^\tau,p^\tau,\tau)\mid X^\tau,p^\tau\big]=0.
\end{aligned}
$$

The lemma and (4.6) imply

$$\mathbb{E}[u(X^{\mathcal{T}},p^{\mathcal{T}},\mathcal{T})-u(X^0,p^0,0)\mid X^0=X,p^0=p]=o(M^{-1/2})$$

which proves Theorem 3.1.

Proof of the Lemma. *The first limit: the friction term.* We have

$$|\tilde\gamma_0|^2 = \frac{|\gamma_0|^2}{|\gamma_0|^2+\sum_{k=1}^J|\gamma_k|^2} = 1+\mathcal{O}(\frac{\sum_{k=1}^J|\gamma_k|^2}{|\gamma_0|^2}).$$

The last condition in (3.16) implies

$$\frac{\sum_{k=1}^J|\gamma_k|^2}{|\gamma_0|^2} = o(1). \tag{4.9}$$

The first statement follows then directly from the definition of K in (3.13), by the change of variables $M^{1/2}(\tau-\sigma)=\hat\sigma$ and integration of solution operator $\tilde S_{\hat\tau,\hat\sigma}=S_{\tau,\sigma}$ in the fast time scale

$$\begin{aligned}&-M^{1/2}\Re\langle\int_0^\tau S_{\tau,\sigma}\dot\Psi_0(X^\sigma)d\sigma,\tilde H\Psi_0'\rangle\\ &\quad=\int_0^{\tau M^{1/2}}\Re\langle\tilde S_{\hat\tau,\hat\tau-\hat\sigma}\Psi_0'(X^{\tau-\hat\sigma M^{-1/2}}),\tilde H\Psi_0'(X^\tau)\rangle d\hat\sigma.\end{aligned}$$

The second limit: the diffusion term. Define the first variation

$$\partial_{\tilde\gamma_k}(X,p,\psi)=:(X_k',p_k',\psi_k'),\ \ k>0,$$

which satisfies the linearized Ehrenfest system

$$\begin{aligned}\dot X' &= p',\\ \dot p_k' &= -X_k'\cdot\partial_{XX}\lambda_0(X)-2\Re\langle\tilde\psi_k,\partial_X\tilde H(X)\psi\rangle-\langle\psi,X_k'\cdot\partial_{XX}\tilde H\psi\rangle\\ &\qquad -2\Re\langle\textstyle\sum_n\tilde\gamma_n\partial_{\tilde\gamma_k}\tilde\psi_n,\partial_X\tilde H\psi\rangle,\\ i\dot{\tilde\psi}_m' &= M^{1/2}\tilde H\tilde\psi_m'+M^{1/2}X'\cdot\partial_X\tilde H\tilde\psi_m,\\ (X,p,\psi)'(0) &= 0,\end{aligned} \tag{4.10}$$

where $\psi=\sum_{n=0}^J\tilde\gamma_n\tilde\psi_n$. The asymptotic result $\psi\simeq\tilde\gamma_0\tilde\psi_0\simeq\Psi_0$, obtained from (4.9) and the Born-Oppenheimer approximation $\tilde\psi_0\simeq\Psi_0$ in Lemma 4.3, together with Duhamel's representation show that

$$\begin{aligned}(X,p,\psi)_n'(\tau) &= -2\textstyle\int_0^\tau G_{\cdot p}(\tau,\sigma)\Re\langle\tilde\psi_n,\partial_X\tilde H(X)\psi\rangle(\sigma)d\sigma\\ &\simeq -2\textstyle\int_0^\tau G_{\cdot p}(\tau,\sigma)\Re\langle\tilde\psi_n^\sigma,\tilde H\partial_X\Psi_0^\sigma\rangle d\sigma\end{aligned} \tag{4.11}$$

where G is the linear solution operator for (4.10) solving

$$\dot G = AG,\ \tau>\sigma \text{ and } G(\sigma,\sigma)=I;$$

here

$$A=A(X^\tau,\psi^\tau):=\begin{bmatrix}0 & I & 0\\ -\lambda_0''-\langle\psi,\tilde H''\psi\rangle & 0 & -2\Re\langle(\tilde\gamma_1,\ldots,\tilde\gamma_J),\tilde H'\psi\rangle\\ -iM^{1/2}\tilde H'B & 0 & -iM^{1/2}\tilde HI\end{bmatrix}$$

is the matrix in the right hand side of (4.10) and the $J\times J$ matrix B is diagonal with $\tilde\psi_j,\ j=1,\ldots J$ in the diagonal. We need information about the short time behavior of $G(\tau,\sigma)$ as $\sigma\to\tau$. For short time the Green's function takes the form $G(\tau,\sigma)=e^{(\tau-\sigma)A}$ and we have

Lemma 4.2. *There holds*

$$G_{pp}(\tau,\sigma)\to I \quad \text{as } \sigma\to\tau$$
$$G_{Xp}(\tau,\sigma)=\mathcal{O}(\tau-\sigma) \text{ for } 0\le\tau-\sigma=o(M^{-1/4}),$$
$$\partial_{\tilde\gamma_m}\tilde\psi_n^\tau=\sum_{k=1}^{3N}M^{1/2}\int_0^\tau\underbrace{\int_\varsigma^\tau\partial_{X_k}\tilde H(X^\sigma)\int_\varsigma^\sigma G_{pp}(\upsilon,\varsigma)d\upsilon d\sigma}_{\mathcal{O}((\tau-\varsigma)^2)}\langle\tilde\psi_m^\varsigma,\partial_{X_k}\tilde H(X^\varsigma)\psi^\varsigma\rangle d\varsigma\,\tilde\psi_n^\tau. \tag{4.12}$$

The proof is in the end of this section. We will now use the first variation with respect to $\tilde\gamma_m$ in (4.11) to verify the second limit in (4.8). We know that

$$\mathbb{E}[\tilde\gamma_n]=\mathbb{E}[\frac{\gamma_n}{(|\gamma_0|^2+\sum_{j=1}^J|\gamma_j|^2)^{1/2}}]=0,$$

since γ_n is normal distributed with mean zero. Consequently, if (X,p,ψ) would be independent of $\tilde\gamma_n$ the expected value

$$M^{1/2}\mathbb{E}[2\Re\langle\tilde\gamma_n\tilde\psi_n^\tau,\tilde\gamma_0\tilde H\Psi_0'(X^\tau)\rangle\cdot\partial_p u(X^\tau,p^\tau,\tau)]$$

would be zero. We shall use the first variation with respect to $\tilde\gamma_n$ to determine how (X,p,ψ) depend on $\tilde\gamma_n$. The coupling can be split into two terms – one term that considers the coupling between the two factors $\langle\tilde\gamma_n\tilde\psi_n^\tau,\tilde\gamma_0\tilde H\Psi_0'(X^\tau)\rangle$ and $\partial_p u(X^\tau,p^\tau,\tau)$ and one term with the intrinsic coupling in the first factor $\langle\tilde\gamma_n\tilde\psi_n^\tau,\tilde\gamma_0\tilde H\Psi_0'(X^\tau)\rangle$. The coupling between the two factors is

$$\begin{aligned}&\sum_{m=1}^J M^{1/2}\int_0^\tau\int_{\mathbb{R}}\mathbb{E}\Big[2\Re\langle\tilde\gamma_n\tilde\psi_n^\tau,\tilde\gamma_0\tilde H\Psi_0'(X^\tau)\rangle 2\Re\langle\tilde\psi_m^\sigma,\tilde\gamma_0\tilde H'\psi^\sigma\rangle\times\\&\big(G_{pp}(\tau,\sigma)\partial_{pp}u(X^\tau,p^\tau,\tau)+G_{Xp}(t,\sigma)\partial_{Xp}u(X^\tau,p^\tau,\tau)\big)\;\Big|\;\underbrace{\tilde\gamma_m^r+i\tilde\gamma_m^i}_{=\tilde\gamma_m}\Big](d\tilde\gamma_m^r+id\tilde\gamma_m^i)d\sigma.\end{aligned}\tag{4.13}$$

The intrinsic coupling is equal to

$$\begin{aligned}&\sum_{m=1}^J M^{1/2}\int_0^\tau\int_{\mathbb{R}}\mathbb{E}\Big[2\Re\langle\tilde\gamma_n\tilde\psi_n^\tau,\tilde\gamma_0\partial_X(\tilde H\Psi_0')G_{Xp}(\tau,\sigma)\rangle\\&\times 2\Re\langle\tilde\psi_m^\sigma,\tilde\gamma_0\tilde H'\psi^\sigma\rangle\partial_p u(X^\tau,p^\tau,\tau)\;\Big|\;\tilde\gamma_m^r+i\tilde\gamma_m^i\Big](d\tilde\gamma_m^r+id\tilde\gamma_m^i)d\sigma\\&+\sum_{m=1}^J M^{1/2}\int_{\mathbb{R}}\mathbb{E}\big[2\Re\langle\tilde\gamma_n\partial_{\tilde\gamma_m}\tilde\psi_n^\tau,\tilde\gamma_0\tilde H\Psi_0'(X^\tau)\rangle\cdot\partial_p u(X^\tau,p^\tau,\tau)\;\big|\;\tilde\gamma_m^r+i\tilde\gamma_m^i\big](d\tilde\gamma_m^r+id\tilde\gamma_m^i)\end{aligned}\tag{4.14}$$

where the first variation $\partial_{\tilde\gamma_m}\tilde\psi_n$ is expressed in terms of G in (4.12).

Let us first study the coupling (4.13) between the two factors. The forcing has the asymptotics

$$\begin{aligned}\tilde H'\psi&=\sum_{n=0}^J\tilde\gamma_n\tilde H'\tilde\psi_n\\&\simeq\tilde\gamma_0\tilde H'\tilde\psi_0\\&\simeq\tilde\gamma_0\tilde H'\Psi_0\\&=-\tilde H\Psi_0'\end{aligned}$$

since $\sum_{n=1}^J\tilde\gamma_n\tilde\psi_n$ is asymptotically smaller than $\tilde\gamma_0\tilde\psi_0$ in $L^2(dx)$, by the last condition in (3.16), and $\tilde\psi_0\simeq\Psi_0$ by Lemma 4.3. This simplified forcing is used to evaluate to coupling below. The first variation shows that the first term in the perturbation in p due to γ_m becomes

$$\int_0^\tau \int_0^{\tilde\gamma_m} \underbrace{\big(G_{pp}(\tau,\sigma)\partial_{pp}u + G_{Xp}(\tau,\sigma)\partial_p u\big)\langle\tilde\psi_n^\sigma, \tilde H\Psi_0'(X^\sigma)\rangle\langle\tilde H\Psi_0'(X^\tau),\tilde\psi_m^\tau\rangle}_{=:F(\hat\gamma_m)} \times |\tilde\gamma_0|^2\tilde\gamma_n d\hat\gamma_m \, d\sigma. \tag{4.15}$$

The last condition in (3.16) implies that $|\tilde\gamma_0|^2 \simeq 1$, so that

$$|\tilde\gamma_0|^2\hat\gamma_n^* d\hat\gamma_n \simeq \hat\gamma_n^* d\hat\gamma_n$$

and the expected perturbation, for $m = n$, becomes

$$\begin{aligned}
&\int_{\mathbb{R}}\int_0^\tau \int_0^{\tilde\gamma_n} F(\hat\gamma_n)\hat\gamma_n^r d\hat\gamma_n^r \, d\sigma \; \frac{e^{-C\tilde\lambda_n|\gamma_n^r|^2/T}d\gamma_n^r}{\sqrt{2\pi T/(C\tilde\lambda_n)}}\\
&\simeq \int_{\mathbb{R}}\int_0^\tau \int_0^{\tilde\gamma_n^r} F(0)\hat\gamma_n^r d\hat\gamma_n^r \, d\sigma \; \frac{e^{-C\tilde\lambda_n|\gamma_n^r|^2/T}d\gamma_n^r}{\sqrt{2\pi T/(C\tilde\lambda_n)}}\\
&= \int_0^\tau \int_{\mathbb{R}} F(0)\frac{|\tilde\gamma_n^r|^2}{2}\frac{e^{-C\tilde\lambda_n|\gamma_n^r|^2/T}d\gamma_n^r}{\sqrt{2\pi T/(C\tilde\lambda_n)}}\, d\sigma\\
&\simeq \int_0^\tau \int_{\mathbb{R}} F(0)\frac{|\gamma_n^r|^2}{2|\gamma_0|^2}\frac{e^{-\tilde\lambda_n|\gamma_n^r|^2/(T|\gamma_0|^2)}d\gamma_n^r}{\sqrt{2\pi T|\gamma_0|^2/\tilde\lambda_n}}\, d\sigma\\
&= \int_0^\tau F(0)\frac{T}{2\tilde\lambda_n}\, d\sigma.
\end{aligned} \tag{4.16}$$

The imaginary part γ_n^i gives an identical contribution. To see that that the approximation $F(\tilde\gamma_n) \simeq F(0)$ holds, we can similarly use the coupling applied to F to find its dependence on $\tilde\gamma_n$

$$F(\tilde\gamma_n) = F(0) + \int_0^\tau \int_0^{\tilde\gamma_n} G_{pp}(\tau,\sigma)\partial_p F\langle\tilde H\Psi_0'(X^\tau),\tilde\psi_n^\tau\rangle\gamma_0 d\hat\gamma_n \, d\sigma + \dots.$$

We see that the dependence on $\tilde\gamma_n$ is small and proportional to

$$\langle\tilde H\Psi_0', \tilde\psi_n\rangle = \mathcal{O}(\sqrt{\Delta\lambda}) = o(1)$$

where $\Delta\lambda$ is the distance between the eigenvalues. Therefore the perturbation (4.15) in F due to $\tilde\gamma_n$ is small

$$F(\gamma_n) = F(0) + \mathcal{O}((\Delta\lambda)^{1/2}) \simeq F(0).$$

The expected value of the coupling between $\tilde\gamma_n$ and $\tilde\gamma_m$, for $m \neq n$, is to leading order equal to zero, since

$$\int_{\mathbb{R}} F(0)\frac{\gamma_n^*\gamma_m}{2|\gamma_0|^2}\frac{e^{-\tilde\lambda_m|\gamma_m|^2/(T|\gamma_0|^2)}d\gamma_m}{\sqrt{2\pi T|\gamma_0|^2/\tilde\lambda_m}} = 0\,.$$

The first non zero expected value for $n \neq m$ can again be obtained by expanding F with respect to both $\tilde\gamma_m$ and $\tilde\gamma_n$

$$
\begin{aligned}
F(\tilde\gamma_m) &= F(0) + \int_0^\tau \int_0^{\tilde\gamma_m} G_{pp}(\tau,\sigma)\partial_p F\langle \tilde H \Psi_0'(X^\tau), \tilde\psi_m^\tau\rangle \tilde\gamma_0 d\hat\gamma_m \, d\sigma + \dots \\
&= F(0) + \int_0^\tau \int_0^{\tilde\gamma_m} \int_0^\tau \int_0^{\tilde\gamma_n} G_{pp}(\tau,\sigma)\partial_{pp} F\langle \tilde H \Psi_0'(X^\tau), \tilde\psi_m^\tau\rangle \tilde\gamma_0 \\
&\qquad \times G_{pp}(\tau,\varsigma)\langle \tilde H \Psi_0'(X^\tau), \tilde\psi_n^\tau\rangle \tilde\gamma_0 d\hat\gamma_n \, d\varsigma d\hat\gamma_m \, d\sigma
\end{aligned} \tag{4.17}
$$

and it takes a similar form as the coupling with only one factor $\tilde\gamma_n$ but now a product of two such coupling factors appear. We will see below that the each integral gains a factor of $M^{-1/2}$

$$
\int_0^\tau F(0)\frac{T}{\tilde\lambda_n} d\sigma = \mathcal{O}(M^{-1/2}), \tag{4.18}
$$

so that the quadratic term with both factors $\tilde\gamma_n$ and $\tilde\gamma_m$ is negligible small $\mathcal{O}(M^{-1})$.

The function F contains the product

$$
\big(\langle \dots {}^\tau_n\rangle + \langle \dots {}^\tau_n\rangle^*\big)\big(\langle \dots {}^\sigma_m\rangle + \langle \dots {}^\sigma_m\rangle^*\big) \tag{4.19}
$$

and can be analyzed by the four terms $\langle \dots {}^\tau_n\rangle\langle \dots {}^\sigma_m\rangle$ similar as in (2.12), now using the solution operator S_{ts} instead of the explicit solution $e^{-iM^{1/2}(\tau-\sigma)\tilde H}$. The following five steps show that the expected value of the fluctuations takes the same form as the friction term in the first statement of (4.8):

1. in the first step the error term comes from (X, p) being slightly dependent on $\tilde\gamma_n$ – this coupling yields a small error term as estimated in (4.16),
2. the second step uses the first condition in (3.16) to deduce that $1/\tilde\lambda_n^0$ and $1/\tilde\lambda_n^\tau$ are close, as explained in (4.21)-(4.22),
3. the third step uses that $\tilde\psi_n = \Psi_n + o(1)$, as derived in Lemma 4.3, to replace a factor of $\tilde\lambda_n^{-1}$ with $\tilde H^{-1}$,
4. the fourth step applies $J \to \infty$, that $\langle (S_{\tau\sigma})^* \tilde H \Psi_0'(X^\tau), \Psi_0'(X^\sigma)\rangle$ is finite, and that the orthonormal set $\{\tilde\psi_n\}_{n=1}^\infty$ in (4.4) forms a basis in the $L^2(\mathbb{R}^J)$ orthogonal complement of $\tilde\psi_0$, and
5. the fifth step uses that Ψ_0 is orthogonal to $\partial_X \Psi_0$ and $\tilde\psi_0 \simeq \Psi_0$, proved in Lemma 4.3,

we write the four terms formed from (4.19) as the sum of two real parts as follows

$$
\begin{aligned}
&2\Re\,\mathbb{E}[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\gamma_n\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \tilde H\partial_X\Psi_0^\sigma, \tilde\gamma_n\tilde\psi_n^\sigma\rangle \mid X^\tau]\\
&\quad +2\Re\,\mathbb{E}[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\gamma_n\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \tilde\gamma_n\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\sigma\rangle \mid X^\tau]\\
&\underbrace{\simeq}_{1.} 2\Re\mathbb{E}\big[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \tilde H\partial_X\Psi_0^\sigma, \tilde\psi_n^\sigma\rangle \underbrace{\mathbb{E}[\tilde\gamma_n^*\tilde\gamma_n] \mid X^\tau\big]}_{\simeq 2T(\tilde\lambda_n^0)^{-1}}\\
&\quad +2\Re\mathbb{E}\big[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\sigma\rangle \mid X^\tau\big]\underbrace{\mathbb{E}[\tilde\gamma_n^*\tilde\gamma_n^*]}_{=0}\\
&= 4\Re\mathbb{E}\big[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \tilde H\partial_X\Psi_0^\sigma, \tilde\psi_n^\sigma\rangle \frac{T}{\tilde\lambda_n^0} \mid X^\tau\big]\\
&\simeq 4\Re\mathbb{E}\big[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \partial_X\Psi_0^\sigma, \tilde H\tilde\psi_n^\sigma\rangle \frac{T}{\tilde\lambda_n^\sigma} \mid X^\tau\big]\\
&\simeq 4T\Re\mathbb{E}\big[\textstyle\sum_{n=1}^J \langle S_{\tau,\sigma}\tilde\psi_n^\sigma, \tilde H\partial_X\Psi_0^\tau\rangle\langle \partial_X\Psi_0^\sigma, \tilde\psi_n^\sigma\rangle \mid X^\tau\big]\\
&\simeq 4T\Re\mathbb{E}\big[\textstyle\sum_{n=1}^\infty \langle \tilde\psi_n^\sigma, (S_{\tau,\sigma})^*\tilde H\partial_X\Psi_0^\tau\rangle\langle \partial_X\Psi_0^\sigma, \tilde\psi_n^\sigma\rangle \mid X^\tau\big]\\
&\underbrace{\simeq}_{5.} 4T\Re\mathbb{E}\big[\textstyle\sum_{n=0}^\infty \langle \tilde\psi_n^\sigma, (S_{\tau,\sigma})^*\tilde H\partial_X\Psi_0^\tau\rangle\langle \partial_X\Psi_0^\sigma, \tilde\psi_n^\sigma\rangle \mid X^\tau\big]\\
&= 4T\Re\mathbb{E}\big[\langle \partial_X\Psi_0^\sigma, (S_{\tau,\sigma})^*\tilde H^\tau\partial_X\Psi_0^\tau\rangle \mid X^\tau\big]\\
&= 4T\Re\mathbb{E}\big[\langle S_{\tau,\sigma}\partial_X\Psi_0^\sigma, \tilde H^\tau\partial_X\Psi_0^\tau\rangle \mid X^\tau\big].
\end{aligned}
\tag{4.20}
$$

In the second step we used that

$$
\frac{1}{\tilde\lambda_n(X^0)} = \frac{1}{\tilde\lambda_n(X^\sigma)}\Big(1 - \frac{(X^0-X^\sigma)\cdot\partial_X\tilde\lambda_n}{\tilde\lambda_n}\Big).
\tag{4.21}
$$

where the error term

$$
T\sum_{n>0}\frac{(X^0-X^\sigma)\cdot\partial_X\tilde\lambda_n}{\tilde\lambda_n} \le |X^0-X^\sigma|_{\ell^\infty}|T\lambda_n^{-1}\lambda_n'|_{\ell^1} = o(M^{-1/2})
\tag{4.22}
$$

is negligible by the first and last assumption in (3.16). We see that the main fluctuation term takes the same form as the friction term, so that the first limit in (4.8) proves the second limit and we have obtained the leading order contribution in the second statement of (4.8). It remains the verify that the other terms in (4.13) are negligible, but first comes the proof of Lemma 4.2.

Proof. **(Lemma 4.2)** A way to understand that the large $M^{1/2}$ factor in the $\tilde\psi'$ equation of (4.10) does not pollute the estimate of G in short time intervals is to eliminate ψ' through the representation

$$
\tilde\psi_n'(\tau) = M^{1/2}\int_0^\tau X'(\sigma)\cdot\partial_X\tilde H(X^\sigma)d\sigma\ \tilde\psi_n^\tau,
$$

obtained from the last equation in (4.10). The ψ'-term in the $\dot p'$ equation can then be written as

$$
2\Re\langle\sum_n\tilde\gamma_n\tilde\psi_n', \partial_X\tilde H\psi\rangle = 2M^{1/2}\Re\langle\int_0^\tau X'(\varsigma)\cdot\partial_X\tilde H(X^\varsigma)d\varsigma\ \psi^\tau, \partial_X\tilde H\psi^\tau\rangle,
$$

which yields the following additional source term in the right hand side to the equation for the Green's function

$$
\begin{aligned}
R(\tau,\sigma) :=& 2M^{1/2}\Re\langle\int_\sigma^\tau\int_\sigma^\varsigma G_{pp}(\varsigma,\upsilon)\cdot\partial_X\tilde H(X^\varsigma)\ \psi^\tau, \partial_X\tilde H\psi^\tau\rangle\ d\upsilon d\varsigma\\
&= M^{1/2}\int_\sigma^\tau\int_\upsilon^\tau\langle G_{pp}(\varsigma,\upsilon)\cdot\partial_X\tilde H(X^\varsigma)\ \psi^\tau, \partial_X\tilde H\psi^\tau\rangle d\varsigma d\upsilon\\
&= M^{1/2}\mathcal{O}((\tau-\sigma)^2)G_{pp}(\tau,\tau).
\end{aligned}
$$

This remainder leads by Duhamel's representation to a small contribution

$$\int_\sigma^\tau e^{(\tau-\sigma)\tilde A}R(\tau,\sigma)d\sigma = M^{1/2}\mathcal{O}((\tau-\sigma)^3)$$

to $G\simeq I+(\tau-\sigma)\tilde A$, provided $M^{1/2}(\tau-\sigma)^2=o(1)$, where $\tilde A$ is the submatrix in (4.10) with ψ' eliminated

$$\tilde A := \begin{bmatrix} 0 & I\\ -\lambda_0''-\langle\psi,\tilde H''\psi\rangle & 0\end{bmatrix}.$$

Therefore, the bound (4.12) holds for $M^{1/2}(\tau-\sigma)^2=o(1)$ and we will use it for much shorter time intervals, satisfying $M^{1/2}(\tau-\sigma)=\mathcal{O}(1)$.

The other coupling in (4.13) based on $\langle\tilde\gamma_n\tilde\psi_n',\tilde H\Psi_0'\rangle\partial_p u$ with

$$\tilde\psi_n'^\tau = \sum_{k=1}^{3N} M^{1/2}\int_0^\tau X_k'(\sigma)\cdot\partial_{X_k}\tilde H(X^\sigma)d\sigma\ \tilde\psi_n^\tau$$

and

$$X'(\sigma)=\int_0^\sigma \tilde p'(\upsilon)d\upsilon = \int_0^\sigma\int_0^\upsilon G_{pp}(\upsilon,\varsigma)\langle\tilde\psi_n^\varsigma,\tilde H\partial_X\Psi_0^\varsigma\rangle d\varsigma d\upsilon$$

yields by the change of the order of integration

$$\begin{aligned}\tilde\psi_n'^\tau &= M^{1/2}\int_0^\tau\int_0^\sigma\int_0^\upsilon G_{pp}(\upsilon,\varsigma)\partial_X\tilde H^\sigma\langle\tilde\psi_n^\varsigma,\tilde H\partial_X\Psi_0(X^\varsigma)\rangle d\varsigma d\upsilon d\sigma\ \tilde\psi_n^\tau\\ &= \sum_k M^{1/2}\int_0^\tau \underbrace{\int_\varsigma^\tau \partial_{X_k}\tilde H(X^\sigma)\int_\varsigma^\sigma G_{pp}(\upsilon,\varsigma)d\upsilon d\sigma}_{\mathcal{O}((\tau-\varsigma)^2)}\langle\tilde\psi_n^\varsigma,\tilde H\partial_{X_k}\Psi_0(X^\varsigma)\rangle d\varsigma\,\tilde\psi_n^\tau,\end{aligned}\tag{4.23}$$

which finishes the proof of (4.12). □

The quadratic factor $(\tau-\varsigma)^2 d\varsigma$ in (4.23) yields by a change to the fast time scale a factor $M^{-3/2}$, where $(\tau-\varsigma)^2$ gives the extra factor M^{-1} as compared to (4.18), and the component wise bound on $\partial_{X_k}\tilde H_{ij}$ together with the ℓ^1 bound on $\tilde H\Psi_0'$ in (3.16) show that the $\tilde\psi'$-term in the intrinsic coupling (4.14) vanishes asymptotically.

The other terms in (4.13) and (4.14) depending on X' includes the factor $G_{Xp}(\tau,\sigma)$, which introduces the additional factor $\tau-\sigma$ (as compared to the p' term with the factor $G_{pp}\sim 1$) in the slow time scale, which in the integration of the fast scale $\tau-\sigma=M^{-1/2}\hat\sigma$ yields an extra factor of $M^{-1/2}$ as compared to the case (4.20) with $G_{pp}(\tau,\sigma)\simeq 1$. Therefore, the terms depending on X' in (4.13) and (4.14) are asymptotically negligible.

The third limit: the quadratic term. The last statement in (4.8) has by the first condition in (3.16) and (4.22) its main contribution from the diagonal part

$$\begin{aligned}&\sum_{n=1}^J \mathbb{E}[|\tilde\gamma_n|^2\langle\tilde\psi_n^\tau,\partial_X\tilde H(X^\tau)\tilde\psi_n^\tau\rangle\cdot\partial_p u(X^\tau,p^\tau,\tau)]\\ &=\sum_{n=1}^J\mathbb{E}[T\tilde\lambda_n^{-1}(X^0)\langle\tilde\psi_n^\tau,\partial_X\tilde H(X^\tau)\tilde\psi_n^\tau\rangle\cdot\partial_p u\big(1+o(1)\big)]\\ &=\sum_{n=1}^J T\,\mathbb{E}[\langle\tilde\psi_n^\tau,\tilde\lambda_n^{-1}\tilde\lambda_n'(X^\tau)\tilde\psi_n^\tau\rangle\cdot\partial_p u\big(1+o(1)\big)]\\ &=o(M^{-1/2}).\end{aligned}$$

The off diagonal contribution yields as in (4.17) two factors

$$\sum_{n,m} T\int_0^\tau F\tilde\lambda_n^{-1}d\sigma\ T\int_0^\tau F\tilde\lambda_m^{-1}d\sigma$$

which gives a factor M^{-1} so that the off diagonal part is negligible.

4.3. **The Born-Oppenheimer Approximation.** The purpose of this section is to study the evolution (4.2) of $\tilde\psi_n$:

Lemma 4.3. *Assume that $i\dot{\tilde\psi}_n = M^{1/2}\tilde H\tilde\psi_n$ holds with the initial data $\tilde\psi_n^0 = \Psi_n(X^0)$, then the orthogonal decomposition $\tilde\psi_n = \bar\psi_n \oplus \psi_n^\perp$, where $\bar\psi_n = \alpha\Psi_n$ for some $\alpha\in\mathbb{C}$, satisfies*

$$\langle \psi_n(t)^\perp, \psi_n(t)^\perp\rangle^{1/2} = o(1). \tag{4.24}$$

Proof. Let $\psi_n^\tau := e^{iM^{1/2}\int_0^\tau \tilde\lambda_n^\sigma d\sigma}\tilde\psi_n^\tau$ and make the decomposition $\psi_n = \bar\psi_n \oplus \psi_n^\perp$, where $\bar\psi_n^\tau$ is an eigenvector of $\tilde H(X^\tau)$, satisfying $\tilde H^\tau\bar\psi_n^\tau = \tilde\lambda_n^\tau\bar\psi_n^\tau$ for the eigenvalue $\tilde\lambda_n^\tau\in\mathbb{R}$. The similar decomposition of $\tilde\psi_n$ in the lemma is related by the factor $e^{iM^{1/2}\int_0^\tau\tilde\lambda_n^\sigma d\sigma}$. This *Ansatz* is motivated by the zero residual

$$R\psi_n := \dot\psi_n + iM^{1/2}(\tilde H - \tilde\lambda_n)\psi_n = 0 \tag{4.25}$$

and the small residual for the eigenvector

$$\begin{aligned} \langle (\dot{\bar\psi}_n)^\natural, \bar\psi_n\rangle &= 0\\ M^{1/2}(\tilde H - \tilde\lambda_n)\bar\psi_n &= 0, \end{aligned}$$

where

$$w(X) = \langle \Psi_n(X), w(X)\rangle\Psi_n(X) \oplus w(X)^\natural \tag{4.26}$$

denotes the orthogonal decomposition in the eigenfunction direction Ψ_n and its orthogonal complement. The constructions of the linear operator R in (4.25), the orthogonal splitting $\psi_n = \bar\psi_n\oplus\psi_n^\perp$ and the projection $\natural$ in (4.26) imply

$$\begin{aligned} 0 &= \big(R(\bar\psi_n + \psi_n^\perp)\big)^\natural\\ &= R(\bar\psi_n)^\natural + R(\psi_n^\perp)^\natural\\ &= R(\bar\psi_n)^\natural + R(\psi_n^\perp) \end{aligned}$$

so that

$$\dot\psi_n^\perp = -iM^{1/2}(\tilde H - \tilde\lambda_n)\psi_n^\perp - (R\bar\psi_n)^\natural$$

and we have the solution representation, as in (4.7),

$$\psi_n^\perp(\tau) = S_{\tau,0}\underbrace{\psi_n^\perp(0)}_{=0} - \int_0^\tau S_{\tau,\sigma}\underbrace{\big(R\bar\psi_n(\sigma)\big)^\natural}_{=\alpha p\cdot\partial_X\Psi_n} d\sigma, \tag{4.27}$$

where S is the solution operator $S_{\tau,0}\psi_n^0 = \psi_n^\tau$ in the slow time scale. Integration by parts introduces the factor $M^{-1/2}$ we want

$$\begin{aligned}
\int_0^\tau S_{\tau,\sigma}R\bar\psi_n(\sigma)^\natural d\sigma &= \int_0^\tau iM^{-1/2}\frac{d}{d\sigma}(S_{\tau,\sigma})(\tilde H-\tilde\lambda_n)^{-1}R\bar\psi_n(\sigma)^\natural d\sigma\\
&= \int_0^\tau iM^{-1/2}\frac{d}{d\sigma}\big(S_{\tau,\sigma})(\tilde H-\tilde\lambda_n)^{-1}R\bar\psi_n(\sigma)^\natural\big)d\sigma\\
&\qquad -\int_0^\tau iM^{-1/2}S_{\tau,\sigma}\frac{d}{d\sigma}\big((\tilde H-\tilde\lambda_n)^{-1}(X^\sigma)R\bar\psi_n(\sigma)^\natural\big)d\sigma\\
&= iM^{-1/2}(\tilde H-\tilde\lambda_n)^{-1}R\bar\psi_n(\tau)^\natural - iM^{-1/2}S_{\tau,0}(\tilde H-\tilde\lambda_n)^{-1}R\bar\psi_n(0)^\natural\\
&\qquad -\int_0^t iM^{-1/2}S_{\tau,\sigma}\frac{d}{d\sigma}\big((\tilde H-\tilde\lambda_n)^{-1}(X^\sigma)R\bar\psi_n(\sigma)^\natural\big)d\sigma.
\end{aligned}\tag{4.28}$$

The spectral gap assumption in (3.16), i.e. $|\tilde\lambda_m-\tilde\lambda_n|>c$ for $m\neq n$ (which excludes multiple eigenvalues), implies by diagonalization on the orthogonal complement of Ψ_n

$$\begin{aligned}
\|(\tilde H-\tilde\lambda_n)^{-1}R\bar\psi_n(0)^\natural\|^2_{L^2(dx)} &= \sum_{m\neq n}^\infty(\tilde\lambda_m-\tilde\lambda_n)^{-2}|\langle\Psi_m,R\bar\psi_n(0)^\natural\rangle|^2\\
&= \sum_{m\neq n}^\infty(\tilde\lambda_m-\tilde\lambda_n)^{-2}\tilde\lambda_m^{-1}|\langle\Psi_m,\tilde H^{1/2}R\bar\psi_n(0)^\natural\rangle|^2\\
&\le c^{-3}\|\tilde H^{1/2}R\bar\psi_n(0)^\natural\|^2_{L^2(dx)} = \mathcal O(c^{-3})
\end{aligned}$$

and the analogous estimate

$$\|\frac{d}{d\sigma}\big((\tilde H-\tilde\lambda_n)^{-1}(X^s)R\bar\psi_n(s)^\natural\big)\|_{L^2(dx)} = \mathcal O(c^{-5/2}),$$

which inserted in (4.27) proves the Lemma for bounded time intervals.

The evolution on longer times requires another idea: one can integrate by parts recursively in (4.28) and assume that the constructed expansion

$$\begin{aligned}
\int_0^\tau S_{\tau,\sigma}R\bar\psi_n(\sigma)^\natural d\sigma &= \Big[\beta\tilde R-\beta\frac{d}{d\sigma}(\beta\tilde R)+\beta\frac{d}{d\sigma}\big(\beta\frac{d}{d\sigma}(\beta\tilde R)\big)-\ldots\Big]_{\sigma=0}^{\sigma=\tau},\\
\beta &:= iM^{-1/2}S_{\tau,\sigma}(\tilde H-\tilde\lambda_n)^{-1},\\
\tilde R &:= R\bar\psi_n(\sigma)^\natural,
\end{aligned}$$

converges, which we do by requiring $\sigma\mapsto\tilde H(X^\sigma)$ to real analytic.

□

4.4. **Proof of Theorem 3.3.**

Proof. Use the estimate

$$\begin{aligned}
&\mathcal T^{-1}\int_0^{\mathcal T}\mathbb E_\gamma\big[g(X^\tau,p^\tau)\big]-\mathbb E_W\big[g(X_L^\tau,\dot X_L^\tau)\big]\ d\tau\\
&= o(M^{-1/2})\mathcal T^{-1}\int_0^{\mathcal T}\int_0^\tau\mathbb E_\gamma[|\mathcal Du(X^\sigma,p^\sigma,\sigma;\tau)|_{\ell^1}]d\sigma d\tau
\end{aligned}$$

from (4.6) and (4.13) in the proof of Theorem 3.1 to obtain

$$\mathcal T^{-1}\int_0^{\mathcal T}\mathbb E_\gamma\big[g(X^\tau,p^\tau)\big]-\mathbb E_W\big[g(X_L^\tau,\dot X_L^\tau)\big]\ d\tau = o(1),$$

based on the assumption

$$\int_0^\tau |\mathcal{D}u(X^\sigma,p^\sigma,\sigma;\tau)|_{\ell^1}\,d\sigma=\mathcal{O}(M^{1/2}),$$

where $\mathcal{D}u$ is used for the combination of derivatives $\partial_p u, \partial_{pp}u, \partial_{Xp}u$ appearing in the estimates of the proof of Theorem 3.1. □

4.5. **A Motivation for** $\lim_{\tau\to\infty}\int_0^\tau \partial_p u(X^\sigma,p^\sigma,\sigma;\tau)d\sigma=\mathcal{O}(M^{1/2})$**.** This section gives a motivation for the assumption in Theorem 3.3 that $\int_0^\infty \partial_p u(X^\sigma,p^\sigma,\sigma)d\sigma=\mathcal{O}(M^{1/2})$, based on stochastic flows.

Using the expected value $\mathbb{E}=\mathbb{E}_W$ with respect to the Wiener process in this section, we have the representation,

$$\begin{aligned}\partial_p u(X,p,\sigma;\tau)&=\partial_p\mathbb{E}[g(X_L^\tau,p_L^\tau)\mid X_L^\sigma=X,p_L^\sigma=p]\\&=\mathbb{E}[\partial_{p_L^\sigma}g(X_L^\tau,p_L^\tau)\mid X_L^\sigma=X,p_L^\sigma=p]\\&=\mathbb{E}[\partial_X g(X_L^\tau,p_L^\tau)\frac{\partial X_L^\tau}{\partial p_L^\sigma}+\partial_p g(X_L^\tau,p_L^\tau)\frac{\partial p_L^\tau}{\partial p_L^\sigma}\mid X_L^\sigma=X,p_L^\sigma=p]\end{aligned}\tag{4.29}$$

where the stochastic flow $(\frac{\partial X_L^\varsigma}{\partial p_L^\sigma},\frac{\partial p_L^\varsigma}{\partial p_L^\sigma})=:\big(X_L'(\varsigma;\sigma),p_L'(\varsigma;\sigma)\big)$ solves the linear equation

$$\begin{aligned}&\frac{d}{d\varsigma}X_L'(\varsigma;\sigma)=p_L'(\varsigma;\sigma)\\&\frac{d}{d\varsigma}p_L'(\varsigma;\sigma)=-\partial_{XX}\lambda_0(X_L^\varsigma)X_L'(\varsigma;\sigma)-\bar K p_L'(\varsigma;\sigma)\end{aligned}$$

in the special case when diffusion coefficient $\bar K:=M^{-1/2}K=M^{-1/2}k\,I$ is a constant multiple of the identity, with $k>0$. To simplify the writing we use the notation $\partial_{XX}\lambda_0(X)=:\lambda_0''(X)$ for the Hessian of λ_0. Let the matrix in the right hand side be denoted by

$$\hat A(\varsigma):=\begin{bmatrix}0 & I\\ -\lambda_0''(X_L^\varsigma) & -\bar K\end{bmatrix}.$$

The discrete time levels $\sigma=\varsigma_0,\ldots,\varsigma_{\hat N}=\tau$ yields for sufficiently small time step $\Delta t:=\varsigma_{n+1}-\varsigma_n$, the representation

$$\begin{bmatrix}X_L'(\varsigma;\sigma)\\ p_L'(\varsigma;\sigma)\end{bmatrix}\simeq\Big(\prod_{n=1}^{\hat N}e^{\Delta t\hat A(\varsigma_n)}\Big)\begin{bmatrix}0\\ I\end{bmatrix}:=e^{\Delta t\hat A(\varsigma_N)}e^{\Delta t\hat A(\varsigma_{N-1})}\ldots e^{\Delta t\hat A(\varsigma_1)}\begin{bmatrix}0\\ I\end{bmatrix}.\tag{4.30}$$

To estimate this product we will use the Euclidian matrix norm $\|A\|:=\sup_{Y\in\mathbb{R}^{6N}}|AY|/|Y|$, which is bounded by the square root of the largest eigenvalue of A^*A and satisfies the product rule

$$\|\prod_{n=1}^{\hat N}e^{\Delta t\hat A(\varsigma_n)}\|\le\prod_{n=1}^{\hat N}\|e^{\Delta t\hat A(\varsigma_n)}\|.$$

To study these exponentials, we need information about the spectrum and therefore we diagonalize the matrix $\hat A(\varsigma_n)$ and define the matrix

$$Q=\begin{bmatrix}a_+ & a_-\\ I & I\end{bmatrix},$$

where

$$a_\pm:=-\frac12\bar K\pm i\big(\lambda_0''(\varsigma_n)-4^{-1}\bar K^2\big)^{1/2}.$$

The matrix Q transforms $\hat{A}(\varsigma_n)$ into block diagonal form

$$Q^{-1}\hat{A}(\varsigma_n)Q = \begin{bmatrix} a_+ & 0 \\ 0 & a_- \end{bmatrix},$$

since $\bar{K}$ and λ_0'' commute (it is only here we use that $\bar{K}$ is a multiple of the identity). Write the eigenvalues of the Hermitian matrices $a_\pm$ as $-k/2 \pm i(\bar{\lambda}_m - k^2/4)^{1/2}$. When $\bar{\lambda}_m - k^2/4$ is positive, the real part of the eigenvalue is negative equal to $-k/2$ and when $\bar{\lambda}_m - k^2/4$ is negative the real part of the eigenvalue is bounded by $(k^2/4 - \bar{\lambda}_m)^{1/2} - k/2$. Introduce therefore the function

$$-k/2 + (k^2/4 - \bar{\lambda}(\varsigma_n))_+^{1/2} := -k/2 + \sqrt{\max_m(0, k^2/4 - \bar{\lambda}_m(\varsigma_n))},$$

which bounds the real part of the eigenvalue, to obtain the spectral bound

$$\|e^{\Delta t \hat{A}(\varsigma_n)}\| \le \exp\Big(\Delta t\big(-k/2 + (k^2/4 - \bar{\lambda}(\varsigma_n))_+^{1/2}\big)\Big)$$

and consequently

$$\begin{aligned}
\|\prod_{n=1}^{\hat{N}} e^{\Delta t \hat{A}(\varsigma_n)}\| &\le \prod_{n=1}^{\hat{N}} \exp\Big(\Delta t\big(-k/2 + (k^2/4 - \bar{\lambda}(\varsigma_n))_+^{1/2}\big)\Big) \\
&= \exp\Big(\sum_{n=1}^{\hat{N}} \Delta t\big(-k/2 + (k^2/4 - \bar{\lambda}(\varsigma_n))_+^{1/2}\big)\Big) \\
&\simeq \exp\Big(\int_\sigma^\tau \big(-k/2 + (k^2/4 - \bar{\lambda}(\varsigma))_+^{1/2}\big) d\varsigma\Big).
\end{aligned} \tag{4.31}$$

The theory of large deviations tells us that for low temperature $T \ll 1$, Langevin solution paths X_L^ς spend long time around stable equilibria, where $\bar{\lambda}_m > 0$, and at some rare events they make short time τ_e (of order one in the slow scale) excursions between such equilibria, see [13]. The number of such rare events in a time interval $[0, \tau - \sigma]$ can be approximately modelled by a Poisson process $m_{\tau-\sigma}$ with the intensity ξ, proportional to $e^{\Delta\lambda_0/T} \sim e^{-1/T}$ (for a negative potential difference $-\Delta\lambda_0 := \lambda_0(X) - \lambda_0(Y) \sim 1$). Let $\kappa := \max_X \big(k^2/4 - \bar{\lambda}(X)\big)_+^{1/2}$, and $\beta := t_e\kappa$.

The estimates (4.30) and (4.31) together show that

$$\Big|\begin{bmatrix} \partial_{p_k} X_L(\varsigma;\sigma) \\ \partial_{p_k} p_L(\varsigma;\sigma) \end{bmatrix}\Big| \le \exp\Big(\int_\sigma^\tau \big(-k/2 + (k^2/4 - \bar{\lambda}(\varsigma))_+^{1/2}\big) d\varsigma\Big)$$

so that representation (4.29) implies the bound

$$\begin{aligned}
&\lim_{\mathcal{T}\to\infty} \mathcal{T}^{-1} \int_0^{\mathcal{T}} \int_0^\tau |\partial_p u(X^\sigma, p^\sigma, \sigma; \tau)|_{\ell^1} d\sigma d\tau \\
&\quad \le C \lim_{\mathcal{T}\to\infty} \mathcal{T}^{-1} \int_0^{\mathcal{T}} \int_0^\tau e^{\int_\sigma^\tau -k/2 + (k^2/4 - \bar{\lambda}(\varsigma))_+^{1/2} d\varsigma} d\sigma d\tau \\
&\quad = C \lim_{\tau\to\infty} \int_0^\tau \mathbb{E}[e^{\int_\sigma^\tau -k/2 + (k^2/4 - \bar{\lambda}(\varsigma))_+^{1/2} d\varsigma}] d\sigma
\end{aligned}$$

for some constant C bounding Q, where the expected value is taken with respect to the Poisson process. This expected value can then roughly by estimated by

$$\begin{aligned}\mathbb{E}[e^{\int_\sigma^\tau -k/2+(k^2/4-\bar\lambda(\varsigma))_+^{1/2}d\varsigma}] &\le \mathbb{E}[e^{-k(\tau-\sigma)/2+\beta m_{\tau-\sigma}}]\\ &= e^{-(\xi+k/2)(\tau-\sigma)}\sum_{m=0}^\infty e^{\beta m}\frac{\big(\xi(\tau-\sigma)\big)^m}{m!}\\ &= e^{\big((e^\beta-1)\xi-k/2\big)(\tau-\sigma)}.\end{aligned}$$

Since $T \ll \log M$, we have a tiny intensity $\xi \sim e^{-1/T} \ll k \sim M^{-1/2}$, which implies

$$(e^\beta-1)\xi - k/2 < -k/3$$

and we conclude that

$$\lim_{\mathcal{T}\to\infty}\mathcal{T}^{-1}\int_0^{\mathcal{T}}\int_0^\tau |\partial_p u(X^\sigma,p^\sigma,\sigma;\tau)|_{\ell^1}\,d\sigma d\tau = \mathcal{O}(M^{1/2}).$$

4.6. **A Motivation for Non Zero Friction.** Note that to have non zero friction and dissipation requires $\Gamma(0,X)>0$, in (3.13), which together with $\langle \tilde H^{1/2}\Psi_0', \tilde H^{1/2}\Psi_0'\rangle \simeq \int_{\mathbb{R}}\Gamma(\lambda,X)d\lambda$ being finite implies that $\langle\Psi_0',\Psi_0'\rangle \simeq \int_{\mathbb{R}}\lambda^{-1}\Gamma(\lambda,X)d\lambda=\infty$. To have an infinite L^2-norm is possible with a slow decay as $|x_j|\to\infty$, for instance that $\partial_{X_n}\Psi_0$ decays as $|x_j-X_n|^{-\alpha}$ with $1/2<\alpha\le 3/2$ for large $|x_j|$ $j=1,\dots J$, using spherical symmetry. Is such a decay reasonable in reality? The Thomas-Fermi model can be used to motivate that $\partial_X\Psi_0$ could decay as $|x|^{-1}$.

The *Thomas-Fermi model* for ground state energies, using an electron density in $\mathbb{R}^3$, is asymptotically exact for large systems where the number of electrons, J, and the total nuclear charge, Z, tend to infinity, see [24]. Its electron density

$$\rho = \Phi^{3/2}/\gamma, \qquad \gamma := (3\pi^2)^{2/3}, \tag{4.32}$$

satisfies for the neutral case $J=Z$ the Thomas-Fermi equation

$$-\Delta_x\Phi + \frac{4}{3\pi}\Phi^{3/2} = 4\pi\sum_{n=1}^N Z_n\delta(x-X_n),$$

and for the derivative

$$\partial_X\rho = 3\Phi^{1/2}\partial_X\Phi/(2\gamma), \tag{4.33}$$

we obtain

$$-\Delta_x\partial_{X_n}\Phi + \frac{2}{\pi}\Phi^{1/2}\partial_{X_n}\Phi = -4\pi Z_n\delta'(x-X_n).$$

The solution $\partial_X\Phi$ then roughly behaves as the derivative of the Green's function, which decays as $|x|^{-2}$. One can observe the decay $|\Psi_0| = \rho^{1/2}\sim |x|^{-3}$, see [24]. This implies by (4.33) and (4.32) that $\partial_X\Psi_0$ has the desired decay

$$\partial_X\rho/\rho^{1/2} \sim \Phi^{1/2}\partial_X\Phi/\Phi^{3/4} \sim \partial_X\Phi/\Phi^{1/4} \sim |x|\partial_X\Phi \sim |x|^{-1}.$$

4.7. **Other Initial Electron Distributions.** This section compares our model of initial data with two other models of electron initial data, having given probabilities to be in mixed states or in pure eigenstates. It turns out that Einstein's fluctuation-dissipation result, which holds for our Hamiltonian system model using the Gibbs distribution does not hold for the traditional canonical distribution. To sample from the Gibbs equilibrium is standard in classical Hamiltonian statistical mechanics but it seems so far non standard for Ehrenfest quantum dynamics.

4.7.1. *Canonical Mixed States.* Let q_j denote the density of state j in the initial data ψ^0 composed of mixed states. In the usual setting of a canonical Gibbs-Boltzmann distribution $q_j = e^{-\tilde\lambda_j/T}/\sum_j e^{-\tilde\lambda_j/T}$, 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 $\sum_j \tilde\lambda_j q_j =$ constant, see [14]. The stochastic model for the variable $|\gamma_j|^2$, measuring in (3.9) and (3.10) the probability to be in electron state j, is different: the chi-square distribution of $\tilde\lambda_j|\gamma_j|^2/T$ contains both the weight to be in electron state j and the spatial density of this state.

4.7.2. *Canonical Pure States.* Assume, instead of (3.9), that ψ^0 is a *pure electron eigenstate* $e^{i\alpha_j}\Psi_j$ with probability

$$q_j := e^{-\bar\lambda_j/T}(\sum_\ell e^{-\bar\lambda_\ell/T})^{-1} \tag{4.34}$$

(and independent random phase shifts α_j uniformly distributed on $[0,2\pi]$) for $j = 0,\ldots,\tilde J$, and write $\psi^0 =: \sum_{j\ge 0}\check\gamma_j\Psi_j$ which has the covariance $\mathbb{E}[(\check\gamma_j)^*\check\gamma_k] = q_j\delta_{jk}$. Let $\bar\Psi_0(X) = \check\gamma_0\,\Psi_0(X)$. Then the fluctuations are very different from the case in Theorem 3.1, since $\mathbb{E}[(\Re\langle\tilde\psi^0, \tilde H\partial_X\bar\Psi_0\rangle)^2]$ is zero due to $\mathbb{E}[\check\gamma_j^*\check\gamma_k\ \check\gamma_0^*\check\gamma_0] = 0$ for $j,k>0$. We assume in (3.9) that the electron-nuclei system is in a pure eigenstate of the full Schrödinger equation, cf. [27], and that does not mean that the electrons have to be in eigenstates of the electron operator H for fixed nuclei positions.

References

[1] G.W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices* (Cambridge University Press 2009).

[2] G. Ariel and E. Vanden-Eijnden, *A strong limit theorem in Kac-Zwanzig model*, Nonlinearity **22** (2009) 145–162.

[3] 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.

[4] E. Cances, M. Defranceschi, W. Kutzelnigg, C. LeBris and Y. Maday, *Computational Chemistry: a primer, in Handbook of Numerical Analysis*, X (North-Holland 2003).

[5] E. Cances, 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] B. Cano, A.M. Stuart, *Under-resolved simulations of heat-baths*, J. Comp. Phys. **169** (2001) 193–214.

[7] A. Einstein, *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilschen*, Ann. Phys. **17** (1905) 549–560.

[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) 507–530.

[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] M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer-Verlag 1998).

[14] C.W. Gardiner, *Quantum Noise* (Springer-Verlag 1991).

[15] G.A. Hagedorn, *A time-dependent Born-Oppenheimer approximation*, Commun. Math. Phys. **77** (1980) 1–19.

[16] O.H. Hald and R. Kupferman, *Asymptotic and numerical analysis for mechanical models of heat baths*, J. Statist. Phys. **106** (2002) 1121–1184.

[17] 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.

[18] L. Kadanoff, Statistical Physics: Statistics, Dynamics and Renormalization (World Scientific 2000).

[19] R. Kupferman and A.M. Stuart, *Fitting SDE models to nonlinear Kac-Zwanzig heat bath models*, Physica D **199** (2004) 279–316.

[20] R. Kupperman, 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.
[21] O.E. Lanford, *Entropy and equilibrium states in classical statistical mechanics*, Statistical Mechanics and Mathematical Problems, Lecture Notes in Physics, Volume 20 (Springer-Verlag 1973).
[22] P. Langevin, *On the theory of Brownian movement*, C.R.Acad. Sci. **1**46 530 (1908), (translation Am. J. Phys. **65** 1079, 1997).
[23] T. Leliévre, M.Rousset and G.Stoltz, *Free Energy Computation: a Mathematical Perspective* (Imperial College Press 2010).
[24] E. Lieb, *Thomas-Fermi and related theories for atoms and molecules,* Rev. Modern Phys. **53** (1981) 603–641.
[25] E. Lieb and R. Seiringer, *The Stability of Matter in Quantum Mechanics* (Cambridge University Press 2010).
[26] 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.
[27] A. Szepessy, *Stochastic and deterministic molecular dynamics derived from the time-independent Schrödinger equation*, arXiv:0812.4338.
[28] 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.
[29] J.C. Tully, *Mixed quantum-classical dynamics*, Faraday Discuss. **110** (1998) 407–419.
[30] 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