

**Universität
Rostock**



Traditio et Innovatio

Vibrational Spectroscopy via Generalized Langevin Dynamics: Applicability and Beyond

Master-Arbeit

**angefertigt am Institut für Physik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Rostock**

von Fabian Gottwald, geb. am 27. Januar 1990 in Vechta

1. Gutachter: Prof. Dr. Oliver Kühn, Universität Rostock

2. Gutachter: Dr. Sergei Ivanov, Universität Rostock

Rostock, den 29. August 2014

Abstract

In this thesis the ability of the Generalized Langevin Equation (GLE) to mimic the dynamics of solute modes in liquid solvents is investigated. The GLE formalism constitutes a reduced description of a system in a non-Markovian environment whose influence is reduced to dissipation and noise. The theoretical background is presented in detail shedding light onto the approximations and models involved as well as the numerical treatment. The dynamics is probed by means of linear absorption spectroscopy. First, the spectra are analyzed with respect to simple test systems illustrating the physical mechanisms of non-Markovian system-bath interactions. Afterwards, the central question about the applicability of the formalism to dynamics of solute modes in liquid solvents is probed by the ability of the GLE to reproduce spectra obtained from explicit Molecular Dynamics simulations. For this purpose a protocol for extracting the herein needed spectral density of the solvent from explicit Molecular Dynamics data is established. The results show that only a linear form of the GLE yields correct spectra, although at a price of projecting any anharmonicity from the system to the environment. Conversely, any non-linear GLE preserving the system anharmonicity suffers from a fundamental mathematical problem and, hence, yields incorrect results. A possible solution of this problem is sketched giving an outlook on how to explicitly include anharmonicity of the system potential into the formalism.

Zusammenfassung

In der vorliegenden Arbeit wird die Anwendbarkeit der verallgemeinerten Langevin Gleichung (VLG) zur Beschreibung von Schwingungsdynamik in Lösungen behandelt. Bei dem VLG-Formalismus handelt es sich um eine reduzierte Beschreibung von Systemen in einer nicht-Markovschen Umgebung, dessen wesentlicher Einfluss durch Dissipation und Rauschen modelliert wird. Die theoretischen Grundlagen werden detailliert dargestellt, die gemachten Näherungen und Modelle erläutert sowie eine Methode zur numerischen Behandlung vorgestellt. Zur Untersuchung der dynamischen Eigenschaften werden Methoden der linearen Absorptionsspektroskopie verwendet. Zunächst werden die Spektren anhand von einfachen Modellsystemen untersucht und die wesentlichen physikalischen Mechanismen einer nicht-Markovschen System-Bad Wechselwirkung analysiert. Anschließend wird die zentrale Frage nach der Anwendbarkeit des Formalismus auf die Dynamik von Schwingungsmoden in Lösungen beleuchtet. Hierbei wird untersucht inwieweit die VLG in der Lage ist, lineare Spektren aus expliziten Molekulardynamik Simulationen zu reproduzieren. Eine Methode zur Bestimmung der dazu benötigten Spektraldichte des Lösungsmittels wird vorgestellt. Insgesamt zeigt sich, dass lediglich eine lineare Form der VLG, wo jedoch sämtliche Anharmonizität im Systempotential formal in die Umgebung projiziert wird, erfolgreiche Ergebnisse liefert. Nichtlineare Formen der VLG, in denen die Anharmonizität explizit gewahrt bleibt, liefern unbrauchbare Ergebnisse, welches durch das Auftreten eines mathematischen Problems begründet werden kann. Ein möglicher Ansatz zur expliziten Einbeziehung von Anharmonizität wird skizziert um den Formalismus für zukünftige Untersuchungen zu erweitern.

Contents

Introduction	3
1. Theoretical Background	7
1.1. The Generalized Langevin Equation	7
1.1.1. An Overview	7
1.1.2. The linear Form of the GLE	11
1.1.3. Non-linear Forms of the GLE	16
1.1.4. Numerical Propagation of the GLE via Colored Noise Thermostats	23
1.2. Linear Absorption Spectroscopy	27
1.3. Concluding Remarks	29
2. Spectral Features of non-Markovian Dynamics	31
2.1. Setting up a Colored Noise Simulation Protocol	32
2.1.1. The Block Averaging scheme	32
2.1.2. Statistical Convergence Analysis	34
2.2. Non-Markovian Spectra and Time-Correlation Functions	36
2.2.1. Spectroscopic Criterion for non-Markovian Dynamics	37
2.2.2. Influence of a non-Markovian Spectral Density	39
2.2.3. Anharmonic spectral Regime	42
2.3. Concluding Remarks	44
3. Applications to vibrational Spectroscopy of Solute Dynamics	47
3.1. Spectral Densities from explicit MD Simulations	47
3.1.1. Method	48
3.1.2. Error Analysis	52
3.1.3. Self-consistency Tests	57
3.2. Vibrational Spectra of Solutes in liquid Solvents	59
3.2.1. Systems	59
3.2.2. Simulation Details	61

3.2.3. Results	62
Conclusions and Outlook	71
Acknowledgements	75
A. Memory kernel, Noise and explicit Force in the LP-GLE	77
B. Memory kernel, Noise and explicit Force in the linearized NLP-GLE	81
C. Integro-differential Equations for Time-Correlation Functions	85
D. Analytic Spectrum for a harmonic Oscillator	87
E. Gaussian Filtering	89
F. Derivation of a second order NLP-GLE	93
G. Short Documentation of the current Colored Noise Implementation	97
G.1. Headers and Object Files	97
G.2. Data Structure and Variables	98
G.3. Routines	99
G.4. Input Files	100
G.5. Example Code	101
Bibliography	107

Abbreviations

Chap.	Chapter
DAF	Dipole autocorrelation function
DOF	Degree of freedom
e.g.	exempli gratia – for example
EOM	Equation of motion
Eq.	Equation
<i>et al.</i>	et alii – and others
etc.	et cetera
FDT	Fluctuation-dissipation theorem
GLE	Generalized Langevin Equation
i.e.	id est – that is
LD	Langevin dynamics
LP-GLE	Generalized Langevin Equation from linear projection operator techniques
MAF	Momentum autocorrelation function
MBO-GLE	Generalized Langevin Equation derived from the Multimode Brownian Oscillator model
MBO	Multimode Brownian Oscillator
MD	Molecular dynamics
MFC	Momentum-force cross-correlation
NLP-GLE	Generalized Langevin Equation from non-linear projection operator techniques
NVE	microcanonical
NVT	canonical
Sec.	Section
TCF	Time-correlation function
TDSE	Time-dependent Schrödinger equation
TISE	Time-independent Schrödinger equation

Introduction

“God does not care about our mathematical difficulties. He integrates empirically.”

— Albert Einstein

Studying complex dynamics of many-particle systems has become one of the main goals in modern molecular physics. The fundamental understanding of the underlying processes requires the interplay of elaborate experimental techniques and sophisticated theoretical approaches. Experimentally, (non-)linear spectroscopy revealed itself as a powerful tool for probing the dynamics and for determining the characteristic timescales for the processes of interest. With the rapid development of the laser technology in the past years it has become more and more feasible to resolve these timescales down to the sub-picosecond range. For interpreting the experimental spectra theoretical models are needed which give insight into the atomistic dynamics on a microscopic level. Specifically, computer simulations provide the bridge between experimental observations and a theoretical understanding of the processes under study. The fast progress in computer facilities allows theoreticians to thrust into the field of many-body systems with increasing size and complexity.

At the beginning of the twentieth century physicists realized that new fundamental concepts like the Heisenberg Uncertainty Principle, a Wave-Particle Duality as well as a Discretization of Energy have to be built into the theoretical description of systems on the atomistic scale. These quantum effects are especially important for light particles at low temperatures and high densities. Their proper inclusion generally requires to solve the full time-dependent Schrödinger Equation (TDSE) accounting for both nuclei and electrons as well as their mutual interactions. Methods like MCTDH [1] can provide an efficient quantum description on the basis of a pure state wavepacket propagation being valid at 0 K. However, simulations of large systems with strong coupling suffer from the curse of dimensionality, which states an exponential increase

of numerical effort with growing system size also known as the exponential wall. This makes a numerical treatment on the basis of the full TDSE impossible for large or even moderate size systems and, hence, meaningful approximate methods are required to overcome this fundamental problem. As a first big step the Born-Oppenheimer approximation is involved [2]. Its basic idea is that since electronic masses are about three orders of magnitudes smaller than nuclei masses, the electronic motion takes place on a much smaller timescale than the nuclei dynamics. This allows one to assume that electrons follow the nuclear configuration nearly instantaneously. The electronic influence can then be reduced to an effective nuclear pair interaction which is given by the electronic charge distribution around the cores. For the nuclear dynamics the remaining task is to solve an effective TDSE being completely freed from explicit electronic contributions. Further approximations are obtained by a systematic expansion of the TDSE in terms of the Planck constant \hbar leading to various semiclassical methods [3]. Finally, the $\hbar \rightarrow 0$ limit amounts to a fully classical description. Treating the nuclei as point particles allows one to calculate the effective nuclear potential by solving the electronic time-independent Schrödinger Equation (TISE) with the nuclei fixed at their present configuration. Here, state-of-the-art quantum chemistry methods like Coupled Cluster, Møller-Plesset Perturbation Theory as well as Density Functional Theory [4] come into play. A parametrization of the effective potentials to force fields finally removes the necessity to repeat time consuming quantum-chemical calculations and the nuclear motion can be described by means of robust Molecular Dynamics (MD) simulations. In many cases, the classical limit is the only feasible way to simulate large many-particle systems like biological molecules with reliable computational effort.

Having in mind the various approximations and simplifications made in deriving classical MD one might be surprised about its undisputed success [2]. Nevertheless, it has been found out by M. L. Koszykowski *et al.* in 1982 that, in the case of the anharmonic Morse oscillator, good agreement between quantum and classical spectral intensities is obtained when employing classical trajectories with a semiclassically chosen action [5]. Further, it has been shown by R. B. Shirts in 1987 that, within the same model, Fourier amplitudes of classical trajectories yield exact agreement with quantum-mechanical expectation values for a selection of quantum observables whereas the classical Fourier amplitudes for off-diagonal elements are shown to be not exact but remarkably accurate [6]. Although these examples show that a classical description of nuclear dynamics can be successful for simple, isolated oscillators it is desirable to extend such quantum vs. classical comparisons to realistic complex many-body systems at finite temperature, where these oscillators are coupled to en-

vironmental degrees of freedom (DOFs). However, as mentioned above, a precise quantum benchmark is hardly affordable via explicit quantum simulations.

Great potential can be expected from implicit formalisms whose basic idea is to select a small subset of DOFs to be followed explicitly, whereas others are considered as unimportant and are described implicitly by modeling their influence on the system. One of the common approaches is based on the idea of Brownian motion, where the environmental influence on the system is reduced to dissipation and fluctuation, which are added to the unperturbed equation of motion (EOM). The most simple formulation of this idea is provided by the Markovian Langevin Equation, where the dissipation is a standard friction and the fluctuation takes the form of white noise [7–10]. This phenomenological equation has found application, for instance, in the theory of chemical reaction rates [11, 12] or diffusion [7–9]. A generalization of this equation to a non-Markovian regime is provided by the Generalized Langevin Equation (GLE) [13–15]. Here, memory effects are explicitly accounted for via a non-Markovian dissipation and a fluctuating force with a finite correlation time. This generalized equation has been involved, for instance, in the theory of vibrational relaxation for estimating characteristic relaxation times [16–19]. The microscopic origin of the GLE can be rationalized starting from different standpoints. In a physical picture it can be derived from the so-called Multimode Brownian Oscillator (MBO) model, where the environment is assumed to be a collection of independent harmonic oscillators bilinearly coupled to the system [10, 14, 20]. This model has been widely used in analyzing and interpreting (non-)linear spectroscopic experiments on systems in condensed phase [21–26]. One of the benefits of this model is that it interpolates between the limits of homogenous and inhomogenous linebroadening providing a microscopic explanation of these effects [20]. Importantly, deriving an implicit description from this model gives rise to memory effects and noise being exactly of GLE form [10, 14, 27]. Another, more formal ansatz for justifying the GLE is to employ projection operator techniques in order to rearrange the system’s EOM into a GLE form [10, 13, 15, 28]. In this approach noise and dissipation can be mathematically defined as projected quantities. Independently of the standpoint from which the GLE is justified, practical use of this equation can only be made in connection with a stochastic model for the noise term being the main assumption in the formalism. The general advantage of the stochastic GLE is that dissipation and the statistical properties of the noise are entirely described by one single function of time, the so-called memory kernel. Due to its simplicity and intuitive physical background the GLE formalism seems to be a promising candidate for a reduced description of systems in macroscopic environments. Moreover, it sets

the stage for a quantum vs. classical comparison, since a quantum version of the GLE is also available [27, 29, 30]. However, the validity of the GLE can be questioned as in the rigorous derivation from projection operator techniques the explicit system forces lose their original form, whereas the possibility to utilize the MBO model, where these forces are preserved, is not apparent in the general case due to the very special assumptions made in this model. The practical applicability of the GLE is the central question in this thesis. Specifically, the possibility to simulate realistic, linear vibrational spectra of systems in liquid solvent environments is the main object of investigations.

In order to utilize the framework of the GLE a memory kernel has to be constructed such that dissipation and noise correctly mimic the environmental behaviour. Common approaches involve classical MD simulations where the environment is explicitly taken into account. Often, one calculates the memory kernel as the time-correlation function (TCF) of the forces exerted on a frozen system coordinate [16, 17, 19]. A formal justification of such an ansatz can only be achieved when the system approaches the high frequency limit [31]. Another family of methods is based on extracting the memory kernel from a Volterra-type integro-differential equation for the momentum-autocorrelation function (MAF). From this equation the memory kernel can be computed from explicit MD TCFs involving either discretization schemes in time-domain [32–35] or Laplace domain techniques [34, 36, 37]. However, the latter suffer from numerical instabilities especially when transforming from Laplace domain back to time domain. In this thesis a Fourier domain protocol is proposed which allows one to directly parametrize the Fourier transform of the memory kernel, the so-called spectral density. Here, the numerical problems in a back transform to time domain are avoided.

The thesis is structured as follows. In the first chapter the theoretical background of the GLE formalism is presented focusing on the models and approximations employed. In the second chapter the principal mechanisms of non-Markovian system-bath interactions and their spectral signatures are discussed and a reliable protocol to set up implicit GLE simulations is established. In the third chapter, the main question of the thesis, the applicability of the GLE to vibrational spectroscopy of real solute dynamics, is probed via the ability to reproduce the explicit spectra obtained from MD simulations where the system-bath interactions are accounted for explicitly. Finally, the findings are summarized and an outlook describing the emerged perspectives is given.

Chapter 1.

Theoretical Background

1.1. The Generalized Langevin Equation

1.1.1. An Overview

The original idea of Brownian motion is based on a very intuitive picture and finds its physical formulation in the (generalized) Langevin equation. The physical situation to be described in this thesis will be a single molecule that is embedded in a large environment. In particular, the main focus will be on a single molecular coordinate x which might correspond to some vibrational mode and which is referred to as *system* in the following. All the environmental DOFs are denoted as $\{Q\}$ and will be called *bath* in the subsequent discussion. Without loss of generality the total potential of system and bath can be decomposed into

$$V(x, \{Q\}) = V_S(x) + V_B(\{Q\}) + V_{S-B}(x, \{Q\}) , \quad (1.1)$$

where V_S labels the system part, V_B the bath part and V_{S-B} their mutual coupling. In principle, the complete knowledge of all positions and momenta of system and bath at one time is sufficient to predict all dynamical properties in the future. The deterministic time evolution of the system's state can be computed from the Hamilton EOMs, which read

$$\begin{aligned} \dot{x}(t) &= \frac{p(t)}{m} \\ \dot{p}(t) &= -\frac{\partial}{\partial x} V_S(x(t)) - \frac{\partial}{\partial x} V_{S-B}(x(t), \{Q(t)\}) , \end{aligned} \quad (1.2)$$

with m being an effective mass associated with the coordinate x . This equation shows that the dynamics of the bath, $\{Q(t)\}$, enters the system's EOM via the system-bath coupling term V_{S-B} which implies that, in order to get a closed solution for $x(t)$, one has to follow this detailed time evolution as well. This is a task that becomes more and more unfeasible when the number of bath DOFs approaches a macroscopic value (i.e. $\sim 10^{23}$). One way to overcome this difficulty is to invoke the central idea of Brownian motion [7–10]. The principal inability of an observer to follow the deterministic bath dynamics amounts to a lack of information being necessary to predict the exact forces stemming from system-bath interactions. Due to this incomplete knowledge the system-bath coupling and, hence, the whole system's motion appears to be of non-deterministic, stochastic nature. This leads to the idea of mimicking the coupling V_{S-B} via stochastic forces $R(t)$. Further, experience teaches one that in many situations the bath hinders the system's motion via a friction force which can be most easily represented by a term $-\gamma p$, with γ denoting a phenomenological friction coefficient. In this simple picture of Brownian motion the difficult coupling term in Eq. (1.2) is reduced to these two effects and the corresponding EOM reads

$$\begin{aligned}\dot{x} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial}{\partial x}V_S(x) - \gamma p + R(t)\end{aligned}\tag{1.3}$$

making the explicit coordinate dependence on bath variables vanishing. The random force $R(t)$ is set up as Gaussian white noise meaning that it is Gaussian distributed, has zero mean and is delta-correlated

$$\begin{aligned}\langle R(t) \rangle &= 0 \\ \langle R(0)R(t) \rangle &= mkT\gamma\delta(t),\end{aligned}\tag{1.4}$$

where T is the temperature, k the Boltzmann constant and $\langle \dots \rangle$ indicate expectation values. The latter relation is also known as *fluctuation-dissipation theorem* (FDT) which states that in thermal equilibrium, the fluctuations expressed by the stochastic forces R must be related to the dissipation determined by the friction coefficient γ [8, 10, 13]. The delta-correlation of the stochastic forces implies that past events have no influence on the force felt at the present time. This means that the system has no memory about its history which is the characteristic of *Markovian* dynamics. If, conversely, memory effects were important, the dynamics would be called *non-Markovian* and a finite correlation of the random forces would have to be accounted for.

The aforementioned stochastic EOM constitutes the original form of Langevin dynamics (LD) as has been used for describing various quantities like chemical reaction rates [11, 12] or collective transport properties like diffusion constants [7–9] on a statistical-mechanical basis. So far, it is a simple but empirical ansatz for a reduced description of a condensed phase system providing physically sensible results for purposes of MD simulations that can be probed, e.g., via vibrational spectroscopy. For instance, LD can serve as a thermostat giving correct canonical distributions of positions and momenta without affecting dynamical properties if the friction coefficient γ is set small enough [38, 39]. In the previous work it has been applied to ensure the canonical ensemble for an HDO molecule in gas phase [39]. The results have shown that spectroscopic features like homogeneous line broadening, that reveals itself via a Lorentzian spectral shape [20, 39, 40], can be obtained from the dynamics induced by Eq. (1.3). However, spectra simulated with an explicit bulk water environment show significant deviations from a Lorentzian lineshape (see [41, 42] or results in Sec. 3.2). This indicates that the simple system-bath coupling description in Eq. (1.3) lacks important effects and has to be extended properly. Specifically, the Markovian assumption of an immediate memory loss in the system-bath interaction becomes questionable as can be illustrated by simple physical arguments. In condensed phase, the surrounding usually forms a characteristic structure located around the system. If at a certain time the system perturbs the environment, it will begin to rearrange its structure on a finite characteristic timescale. Thus, it becomes obvious that the environmental forces exerted on the system at later times would depend on the present state of the system which gives rise to memory effects and thus requires a non-Markovian version of LD. Indeed, such a Generalized Langevin equation (GLE) has been exploited, for instance, in successful studies of vibrational relaxation in bulk systems [16–19]. The herein used GLE reads

$$\begin{aligned}\dot{x} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial}{\partial x} V_S(x) - \int_0^t \xi(t-\tau)p(\tau)d\tau + R(t),\end{aligned}\tag{1.5}$$

where the dissipative force is formed by an integral over all the momenta lying in the system's past, weighted with a memory kernel $\xi(t)$. The random force $R(t)$ used in this

approach is given the statistical properties

$$\begin{aligned}\langle R(t) \rangle &= 0 \\ \langle R(0)R(t) \rangle &= mkT\xi(t),\end{aligned}\tag{1.6}$$

with the latter formula being the generalized FDT that incorporates a finite correlation time [10, 13–15]. Apparently, the Markovian Langevin Equation, Eq. (1.3), is the limiting case of the GLE for a fast memory loss $\xi(t) \rightarrow \gamma\delta(t)$. In the later discussions it will be more convenient to think in terms of spectral densities, which are defined as the half-sided Fourier-transform of the memory kernel

$$J(\omega) \equiv \int_0^{\infty} e^{-i\omega t} \xi(t) dt.\tag{1.7}$$

In the Markovian case the spectral density becomes a constant, i.e. $J(\omega) = \gamma$, whereas for any non-Markovian situation it is a function of finite width. For this reason one often refers to the Markovian noise as *white noise*, whereas its non-Markovian counterpart is often called *colored noise*. Note that in the literature one often includes an additional prefactor ω in the definition of $J(\omega)$, which will not be adopted in this thesis. Further, in subsequent discussions only the real part of $J(\omega)$ will be considered since the imaginary part is related to the real part via Kramers-Kronig relations and, hence, its knowledge completely determines the former.

The main subject of this thesis is the dynamics of a selected vibrational mode of a molecule embedded in liquid surroundings, whose influence shall be described by the GLE and probed via linear absorption spectroscopy. Before applying the formalism presented above, it is necessary to study its microscopic origin since this will demonstrate possible limitations and approximations of the method, which so far has been motivated by physical arguments only. The following subsections are thus aimed at shedding light on the theoretical background of the GLE by presenting formal derivations of its different prototype forms. In anticipation of the subsequent discussion it will turn out that the only exact justification of Eq. (1.5) is based on a linear form in which the system's explicit potential, V_S , is replaced by an effective harmonic one. As already has been put forward in the introduction, an important criterion for a reduced system-bath formalism is a meaningful connection to quantum applications when a quantum vs. classical comparison shall be performed. Unfortunately, this linear GLE will turn out inappropriate for this purpose and, hence, a non-linear GLE should be

preferred. Therefore, three approximate approaches for an explicit inclusion of nonlinearities into the GLE are discussed. After having presented the theoretical background a numerical scheme is described in order to obtain solutions of the GLE. Here, the machinery of *colored noise thermostats* developed by Cerriotti *et al.* [43–46] is involved.

1.1.2. The linear Form of the GLE

First, the derivation of a linear GLE is given involving the mathematically rigorous concept of linear projection operators in the Hilbert space of dynamical functions. An extension to the non-linear case is given in the next section. The derivation starts with the definition of the linear projection operator onto the system's linear subspace. Then, an operator identity is applied to the Hamilton EOMs in order to separate the forces acting on the system into terms which lie in the linear subspace and those orthogonal to it [10, 13, 15]. The EOM obtained this way has a similar form to that of the GLE in Eq. (1.5). Particular differences are discussed after the derivation.

In general, one can understand a classical dynamical variable A to be a function defined on the system's phase space spanned by all positions and their conjugate momenta, $A \equiv A(\Gamma)$ with $\Gamma = (x_1, x_2, \dots, x_f, p_1, p_2, \dots, p_f)^T$. A Hilbert space can be constructed out of these phase space functions by considering a distribution function $f(\Gamma)$ which can in principal be chosen arbitrarily. The only requirement for $f(\Gamma)$ is that the set of observables under interest lies in the following function space

$$\mathcal{H} = \left\{ A(\Gamma) : \int f(\Gamma) |A(\Gamma)|^2 d\Gamma < \infty \right\}. \quad (1.8)$$

The distribution function $f(\Gamma)$ acts as an integration measure and is often chosen to be the canonical distribution function of the total system

$$f(\Gamma) = \frac{1}{Z} \exp \left[-\frac{1}{kT} H(\Gamma) \right], \quad (1.9)$$

with the partition function Z , or to be some non-equilibrium distribution function when a non-equilibrium process is described [10, 28]. The function space defined in

Eq. (1.8) is a Hilbert space with respect to the following scalar product

$$\langle A|B \rangle = \int f(\Gamma) A(\Gamma) B(\Gamma) d\Gamma, \quad (1.10)$$

which now enables full use of Hilbert space properties. Note that in the literature one often includes the complex conjugated variable, $A^*(\Gamma)$, into the definition of the scalar product. Since classical dynamic variables are usually real-valued, the complex conjugation is not included into the definition here.

Given the aforementioned mathematical framework one can define the linear projection operator for a set of N dynamical variables $\{A_i\}$, $i = 1, \dots, N$ whose dynamics is followed explicitly. Others are be graded as unimportant and are regarded as a part of the bath. For shortening the notation in the later derivation it is convenient to introduce the N -dimensional vectors

$$\mathbf{A} \equiv \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix}, \quad \langle \mathbf{A}|B \rangle \equiv \begin{pmatrix} \langle A_1|B \rangle \\ \langle A_2|B \rangle \\ \vdots \\ \langle A_N|B \rangle \end{pmatrix}. \quad (1.11)$$

The former consists of all dynamical variables $\{A_i\}$ and the latter comprises all scalar products of some dynamical function B and the variables $\{A_i\}$. Finally the scalar product of an N -dimensional vector \mathbf{A} with an M -dimensional vector \mathbf{B} is defined as a $N \times M$ matrix

$$\langle \mathbf{A}|\mathbf{B} \rangle \equiv \begin{pmatrix} \langle A_1|B_1 \rangle & \langle A_1|B_2 \rangle & \dots & \langle A_1|B_M \rangle \\ \langle A_2|B_1 \rangle & \langle A_2|B_2 \rangle & & \vdots \\ \vdots & & \ddots & \\ \langle A_N|B_1 \rangle & \dots & & \langle A_N|B_M \rangle \end{pmatrix}. \quad (1.12)$$

With these conventions, the linear projection operator \hat{P} onto the subspace spanned by the variables \mathbf{A} can be obtained by defining its action on an arbitrary $B \in \mathcal{H}$ as

$$\hat{P}B \equiv \langle \mathbf{A}|B \rangle^T \langle \mathbf{A}|\mathbf{A} \rangle^{-1} \mathbf{A}, \quad (1.13)$$

written explicitly as

$$\hat{P}B = \sum_{i,j=1}^N \langle A_i|B \rangle \cdot (\langle \mathbf{A}|\mathbf{A} \rangle^{-1})_{ij} \cdot A_j. \quad (1.14)$$

Note that the set of observables does not need to be orthonormal which means that the matrix $\langle \mathbf{A}|\mathbf{A} \rangle$ can have off-diagonal elements. Now, the time evolution of the vector \mathbf{A} is considered. According to the Hamilton EOM \mathbf{A} propagates via

$$\begin{aligned} \frac{\partial}{\partial t} \mathbf{A}(t) &= \{\mathbf{A}(t), H\} \\ &\equiv \mathcal{L} \mathbf{A}(t), \end{aligned} \quad (1.15)$$

where the action of the Poisson bracket, $\{\bullet, H\}$ with the Hamilton function H defines the Liouville operator $\mathcal{L}\bullet$. Note that the action of the Liouville operator and the time derivative on an N -dimensional vectors \mathbf{A} is defined componentwise and, thus, the result is an N -dimensional vector as well. The formal solution of this equation can be written as

$$\mathbf{A}(t) = e^{t\mathcal{L}} \mathbf{A}(0), \quad (1.16)$$

where $\mathbf{A}(0)$ denotes the initial function, i.e. $\mathbf{A}(0) \equiv \mathbf{A}[\Gamma(0)]$. Armed with the projection operator defined above the derivation of the GLE can be started. The key formula used herein is the following operator identity [10, 15, 28]

$$e^{t\mathcal{L}} = e^{t(1-\hat{P})\mathcal{L}} + \int_0^t e^{(t-\tau)\mathcal{L}} \hat{P} \mathcal{L} e^{\tau(1-\hat{P})\mathcal{L}} d\tau, \quad (1.17)$$

which can be implicitly proven by taking the Laplace transform of both sides of the equation or explicitly via direct differentiation. Applying this identity to $(1 - \hat{P})\mathcal{L}\mathbf{A}(0)$ yields for the left hand side of Eq. (1.17)

$$\begin{aligned} e^{t\mathcal{L}}(1 - \hat{P})\mathcal{L}\mathbf{A}(0) &= \frac{\partial}{\partial t} \mathbf{A}(t) - e^{t\mathcal{L}} \hat{P} \mathcal{L} \mathbf{A}(0) \\ &= \frac{\partial}{\partial t} \mathbf{A}(t) - \langle \mathbf{A}|\mathcal{L}\mathbf{A} \rangle^T \langle \mathbf{A}|\mathbf{A} \rangle^{-1} \mathbf{A}(t) \\ &= \frac{\partial}{\partial t} \mathbf{A}(t) - \boldsymbol{\Omega} \mathbf{A}(t), \end{aligned} \quad (1.18)$$

with the matrix Ω defined as

$$\Omega \equiv \langle \mathbf{A} | \mathcal{L} \mathbf{A} \rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} . \quad (1.19)$$

Note that the product of the two matrices $\langle \mathbf{A} | \mathcal{L} \mathbf{A} \rangle^T$ and $\langle \mathbf{A} | \mathbf{A} \rangle^{-1}$ is a usual matrix multiplication. For the first term on the right hand side of Eq. (1.17) one gets

$$\mathbf{F}(t) \equiv e^{t(1-\hat{P})\mathcal{L}}(1 - \hat{P})\mathcal{L}\mathbf{A}(0) , \quad (1.20)$$

which will turn out to be the counterpart of the fluctuating force $R(t)$ in Eq. (1.5). In the remaining integral one notices that the term $\mathbf{F}(t)$ appears. Upon carrying out the projection one obtains

$$\begin{aligned} \int_0^t e^{(t-\tau)\mathcal{L}} \hat{P} \mathcal{L} e^{\tau(1-\hat{P})\mathcal{L}} (1 - \hat{P}) \mathcal{L} \mathbf{A}(0) d\tau &= \int_0^t e^{(t-\tau)\mathcal{L}} \hat{P} \mathcal{L} \mathbf{F}(\tau) d\tau \\ &= \int_0^t e^{(t-\tau)\mathcal{L}} \langle \mathbf{A} | \mathcal{L} \mathbf{F}(\tau) \rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} \mathbf{A}(0) d\tau \\ &= - \int_0^t \mathbf{K}(\tau) \mathbf{A}(t - \tau) d\tau , \end{aligned} \quad (1.21)$$

with the memory matrix defined as

$$\mathbf{K}(t) \equiv \langle \mathcal{L} \mathbf{A} | \mathbf{F}(t) \rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} . \quad (1.22)$$

To obtain this result the antihermitian property of the Liouville operator with respect to the scalar product defined in Eq. (1.10) has been employed. Combining the results collected so far, one can write down a formal GLE for the dynamical functions \mathbf{A}

$$\frac{\partial}{\partial t} \mathbf{A}(t) = \Omega \mathbf{A}(t) - \int_0^t \mathbf{K}(\tau) \mathbf{A}(t - \tau) d\tau + \mathbf{F}(t) . \quad (1.23)$$

Further, the generalized FDT, that can be directly read off Eq. (1.22), is validated

$$\begin{aligned} \mathbf{K}(t) &= \langle \mathcal{L} \mathbf{A} | \mathbf{F}(t) \rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} \\ &= \left\langle (1 - \hat{P}) \mathcal{L} \mathbf{A} | \mathbf{F}(t) \right\rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} \\ &= \langle \mathbf{F}(0) \mathbf{F}(t) \rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} . \end{aligned} \quad (1.24)$$

Note that the second line has been obtained by inserting a redundant projector $1 - \hat{P}$. The hermitian property of projectors allows one to let $1 - \hat{P}$ act on $\mathbf{F}(t)$, which itself contains the same projector, see Eq. (1.20). The redundancy of this projector then follows from its idempotency property. In the third line the definition of the scalar product, Eq. (1.10) has been used. At this point the derivation of the GLE is finished. To be more concrete one can apply the formalism to a special pair of variables namely the system's coordinate x and momentum p . Here, the vector \mathbf{A} becomes $\mathbf{A} = (x, p)^T$ and, as shown in appendix A, the noise $\mathbf{F}(t)$ as well as Ω and the memory matrix $\mathbf{K}(t)$ take the forms

$$\mathbf{F}(t) = \begin{pmatrix} 0 \\ R(t) \end{pmatrix}, \quad \Omega = \begin{pmatrix} 0 & 1/m \\ -kT/\langle x^2 \rangle & 0 \end{pmatrix}, \quad \mathbf{K}(t) = \begin{pmatrix} 0 & 0 \\ 0 & \xi(t) \end{pmatrix}, \quad (1.25)$$

if the canonical distribution function, Eq. (1.9), is used in the scalar product. These equations define the random force $R(t)$ and the memory kernel $\xi(t)$. The corresponding GLE then reads

$$\begin{aligned} \dot{x}(t) &= \frac{p(t)}{m} \\ \dot{p}(t) &= -\frac{kT}{\langle x^2 \rangle} x(t) - \int_0^t \xi(t - \tau) p(\tau) d\tau + R(t) \\ \langle R(0)R(t) \rangle &= mkT\xi(t). \end{aligned} \quad (1.26)$$

This formally exact GLE derived from linear projection operator techniques looks very similar to the common GLE in Eq. (1.5) motivated by purely physical arguments. Throughout this thesis Eq. (1.26) is given the abbreviation LP-GLE. Its derivation mathematically shows why and how memory effects and noise arise in a reduced description.

However, the LP-GLE has certain differences to the common GLE which are worth discussing further. At first, it has to be stressed that the noise term is given explicitly via Eq. (1.20) and is formally not random at all. Nevertheless, since it propagates according to a physically counter-intuitive propagator $e^{t(1-\hat{P})\mathcal{L}}$ its deterministic time evolution is hard to follow explicitly especially for an increasing size of the environment. In the spirit of the aforementioned Brownian motion idea one can motivate a stochastic model for this term keeping the FDT as its main statistical property. A second difference lies in the explicit force, which in Eq. (1.5) is given by the full system

force $-\partial_x V_S$. In contrast, the force in Eq. (1.26) is always linear, or in other words always appears as an effective harmonic force with the frequency $\tilde{\omega}^2 = kT/m \langle x^2 \rangle$, even if the real system is arbitrarily anharmonic. This is the direct consequence of invoking linear projections, which project out all non-linearities to the bath. Here, several conceptual problems arise. First, it should be stressed that the later applications of the LP-GLE, see Chap. 3, show that the effective harmonic frequency can be hardly connected to the real parameters of the intramolecular potential V_S employed. In contrast, they artificially contain frequency shifts usually being a consequence of system-bath interactions. This mixing of system and bath contributions completely disguises the atomistic picture of the dynamical processes under study. The second problem concerns the intention of a quantum vs. classical comparison. Conceptually, such a comparison can be done by adopting the classical memory kernel for a quantum dynamical simulation. Quantum effects in vibrational spectra are mostly expected for anharmonic systems since, spectroscopically, a classical harmonic oscillator is identical to a quantum harmonic oscillator. If the LP-GLE memory kernel, where all the anharmonicity is included, was used for quantum simulations, one would effectively keep the anharmonicity classical and, hence, no differences in quantum and classical spectra would arise. Finally, in future applications it is desirable to probe the LD via non-linear spectroscopy. Due to the linear structure of the LP-GLE it is expected that only linear response functions are described correctly, whereas the proper inclusion of non-linear effects arising from anharmonicity in the intermolecular potential V_S becomes questionable. For these reasons, the LP-GLE derived above is not of much use and one would prefer a form of the GLE which keeps the explicit system force non-linear. The different strategies on how to explicitly include non-linearities into the GLE are presented in the following section.

1.1.3. Non-linear Forms of the GLE

As it has been discussed in the previous section the linear GLE is not the proper concept in various physical situations. In this section, three ways to explicitly preserve non-linearities of the system potential in the corresponding GLE will be discussed in detail.

Ad hoc anharmonic GLE

The most simple approach to keep the anharmonicity would be to add it to the LP-GLE directly. In particular, the resulting GLE reads

$$\begin{aligned}\dot{x} &= \frac{p}{m} \\ \dot{p} &= -m\tilde{\omega}^2 x + m\chi x^2 - \int_0^t \xi(t-\tau)p(\tau)d\tau + R(t),\end{aligned}\quad (1.27)$$

where the anharmonicity of the potential has been restricted to a cubic term only. Technically, the anharmonicity constant χ can be obtained from a third order Taylor expansion of the real system potential V_S occurring in Eq. (1.1). The memory kernel $\xi(t)$, the harmonic frequency $\tilde{\omega}$ as well as the FDT for the noise $R(t)$ are simply adopted from the LP-GLE introduced in the previous section. A possible argumentation for justifying this approach is that memory kernel and noise should entirely be determined by the system-bath interactions which, in turn, should be independent on the intramolecular system potential V_S according to the system-bath partitioning in Eq. (1.1). However, as is explained in the previous section, using the LP-GLE leads to a rather counter-intuitive partitioning where system properties, namely the anharmonicity, constitute a part of the bath. In contrast, it is rather obvious that the corresponding memory kernel and FDT depend on system properties as well, which can be for instance seen in Eq. (1.22). From this perspective the non-linear GLE presented above lacks a formal justification. Still, in order to elucidate the practical consequences of these shortcomings it will be included into the scope of this thesis and will be referred to as the *ad hoc* anharmonic GLE.

GLE from the Multimode Brownian Oscillator Model

The Multimode Brownian Oscillator (MBO) model introduced in this section has enjoyed a broad application in the analysis of non-linear spectroscopic signals arising in experiments in condensed phase [21–26]. Here, it serves as a model for the system-bath coupling from which a non-linear GLE can be derived following R. Zwanzig in [10, 14].

Within the MBO model the bath is assumed to be a set of harmonic oscillators with masses M_j , positions Q_j , momenta P_j and a bilinear coupling to the system quantified

by the coupling strengths g_j [10, 14, 20]. The total potential reads

$$V = V_S + \sum_j \frac{1}{2} M_j \omega_j^2 \left(Q_j - \frac{g_j}{M_j \omega_j^2} x \right)^2. \quad (1.28)$$

The Hamilton EOMs derived from this potential read for the particle

$$\begin{aligned} \dot{x} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial V_S}{\partial x} + \sum_j g_j \left(Q_j - \frac{g_j}{M_j \omega_j^2} x \right) \end{aligned} \quad (1.29)$$

and for the bath oscillators

$$\begin{aligned} \dot{Q}_j &= \frac{P_j}{M_j} \\ \dot{P}_j &= -M_j \omega_j^2 Q_j + g_j x. \end{aligned} \quad (1.30)$$

From Eq. (1.30) one sees that the bath undergoes independent harmonic oscillations which are driven by the system's trajectory $x(t)$. Thus, the homogenous solutions of the set of equations for the bath oscillators can be written down easily

$$Q_j^H(t) = Q_j(0) \cos(\omega_j t) + \frac{P_j(0)}{M_j \omega_j} \sin(\omega_j t), \quad (1.31)$$

whereas the inhomogenous solution can be expressed in terms of the system's trajectory $x(t)$ as

$$Q_j^I(t) = g_j \int_0^t x(\tau) \frac{\sin[\omega_j(t - \tau)]}{M_j \omega_j} d\tau, \quad (1.32)$$

which can be proven directly by differentiation. Combining the homogenous and inhomogenous solutions into the general solution and integrating the inhomogenous term by parts one can obtain the following relation

$$\begin{aligned} Q_j(t) - \frac{g_j}{M_j \omega_j^2} x(t) &= \left[Q_j(0) - \frac{g_j}{M_j \omega_j^2} x(0) \right] \cos(\omega_j t) + \frac{P_j(0)}{M_j \omega_j} \sin(\omega_j t) \\ &\quad - \frac{g_j}{M_j \omega_j^2} \int_0^t \frac{p(\tau)}{m} \cos[\omega_j(t - \tau)] d\tau. \end{aligned} \quad (1.33)$$

Direct substitution into Eq. (1.29) yields a GLE of the form of Eq. (1.5)

$$\begin{aligned}\dot{x} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial V_S}{\partial x} - \int_0^t \xi(t-\tau)p(\tau)d\tau + R(t).\end{aligned}\quad (1.34)$$

The explicit forms of the memory kernel $\xi(t)$ and the noise $R(t)$ read

$$\xi(t) = \sum_j \frac{g_j^2}{mM_j\omega_j^2} \cos(\omega_j t) \quad (1.35)$$

$$R(t) = \sum_j g_j \frac{P_j(0)}{M_j\omega_j} \sin(\omega_j t) + \sum_j g_j \left(Q_j(0) - \frac{g_j}{M_j\omega_j^2} x(0) \right) \cos(\omega_j t). \quad (1.36)$$

Further, the FDT

$$\langle R(0)R(t) \rangle = mkT\xi(t) \quad (1.37)$$

can be directly verified by carrying out the canonical ensemble average.

The GLE obtained from the MBO model coincides with Eq. (1.5) in all aspects. Especially an inclusion of anharmonic effects is possible, which was the major drawback of the LP-GLE. Another advantage of the MBO model is that it can be treated quantum mechanically in a similar fashion with the only difference of a quantum-corrected FDT [27]. Thus, it provides a proper stage for a quantum vs. classical comparison with the only disadvantage that it is based on a special model whose microscopic justification for an arbitrary system is questionable. Such a justification can be achieved in a normal mode description of the bath as pointed out by Goodyear and Stratt [36]. In this representation the form of the MBO model, Eq. (1.28), can be derived with the Q_i being the normal mode coordinates of the bath. Of course, the validity of this description then depends on the harmonicity of the bath under study, which is not given in the general case. Later, Eq. (1.34) derived from the MBO model is referred to as the MBO-GLE.

GLE from the non-linear Projection Operator Technique

The MBO model introduced in the previous section provides a direct way to preserve a system anharmonicity explicitly in the GLE. However, as has been pointed out before, the MBO model assumes a very special form of bath and system-bath coupling. Hence, it is not clear whether a realistic system can be described by it or not. In this section, a mathematically rigorous derivation of the GLE that includes non-linear effects in the explicit forces is presented. The procedure is very similar to that used in Sec. 1.1.2, with the only difference that non-linear projections are employed [15, 28].

The starting point is the same Hilbert space of dynamic functions as in Sec. 1.1.2 with the very same scalar product defined in Eq. (1.10). Given a vector of dynamical variables $\mathbf{A} \equiv \mathbf{A}(\Gamma) = (A_1(\Gamma), A_2(\Gamma), \dots, A_N(\Gamma))^T$ the subspace of all functions of \mathbf{A}

$$\mathcal{F} = \left\{ F(\mathbf{A}(\Gamma)) : \int f(\Gamma) |F(\mathbf{A}(\Gamma))|^2 d\Gamma < \infty \right\} \quad (1.38)$$

is now considered. Note that in Sec. 1.1.2 the subspace was linearly spanned by the variables, whereas here all non-linear functions are included in \mathcal{F} . In order to define the projector \hat{P} onto \mathcal{F} consider a complete basis $\{\Phi_j(\mathbf{A})\}$, $j = 1, 2, \dots, \infty$ on \mathcal{F} which is assumed to be orthonormal. Then the projector \hat{P} is defined in a similar way as in Sec. 1.1.2

$$\hat{P}B \equiv \sum_{i=1}^{\infty} \langle \Phi_i | B \rangle \Phi_i. \quad (1.39)$$

Now, the further procedure is identical to that in Sec. 1.1.2: the operator identity in Eq. (1.17) is applied to the quantity $(1 - \hat{P})\mathcal{L}\mathbf{A}$ in order to separate the Liouville propagator $e^{t\mathcal{L}}$ into terms lying in \mathcal{F} and those orthogonal to it. The only difference is the non-linear projection, Eq. (1.39), instead of the linear one. According to Kawai *et al.* [28] the respective terms are

$$\begin{aligned} e^{t\mathcal{L}} \hat{P} \mathcal{L} \mathbf{A} &= \langle \partial_t \mathbf{A}; \mathbf{A}(t) \rangle \\ e^{t(1-\hat{P})\mathcal{L}} (1 - \hat{P}) \mathcal{L} \mathbf{A} &\equiv \mathbf{F}(t) \\ e^{(t-\tau)\mathcal{L}} \hat{P} \mathcal{L} \mathbf{F}(\tau) &= - \sum_{i=1}^{\infty} \Phi_i[\mathbf{A}(t-\tau)] \langle \{\nabla_{\mathbf{A}} \Phi_i \cdot \mathbf{F}(0)\} \mathbf{F}(\tau) \rangle, \end{aligned} \quad (1.40)$$

where the first term forms the explicit part, the second term the noise and the last term the memory kernel. In the latter, $\nabla_{\mathbf{A}}$ denotes the gradient with respect to the vector of

variables \mathbf{A} , which being applied to a function Φ gives a vector of same dimensionality as \mathbf{A} . Note that the product in curly braces is understood as a usual vector scalar product. Hence, the whole term in $\langle \dots \rangle$ becomes a vector of the dimensionality of \mathbf{F} . The first line in Eq. (1.40) is formed by the average of the time derivative $\partial_t \mathbf{A}$ under the condition that \mathbf{A} carries the trajectory value $\mathbf{A}(t)$. This conditional average can generally be defined as

$$\langle B; \mathbf{a} \rangle \equiv P(\mathbf{a})^{-1} \int f(\Gamma) \delta(\mathbf{A}(\Gamma) - \mathbf{a}) B(\Gamma) d\Gamma, \quad (1.41)$$

with $\mathbf{a} = \mathbf{A}(t)$ being the condition and $P(\mathbf{a})$ the reduced distribution function for \mathbf{a}

$$P(\mathbf{a}) = \int f(\Gamma) \delta(\mathbf{A}(\Gamma) - \mathbf{a}) d\Gamma. \quad (1.42)$$

For the derivation details of the Eqs. (1.40) consult the paper of Kawai *et al.* [28]. Combining the terms in Eqs. (1.40) yields the non-linear version of the GLE

$$\frac{\partial}{\partial t} \mathbf{A}(t) = \langle \partial_t \mathbf{A}; \mathbf{A}(t) \rangle - \int_0^t \sum_{i=1}^{\infty} \Phi_i [\mathbf{A}(t-s)] \langle \{ \nabla_{\mathbf{A}} \Phi_i \cdot \mathbf{F}(0) \} \mathbf{F}(s) \rangle ds + \mathbf{F}(t), \quad (1.43)$$

which being applied to the special choice of $\mathbf{A} = (x, p)^T$ using the canonical distribution function in the scalar product reduces to

$$\begin{aligned} \dot{x} &= \frac{p}{m} \\ \dot{p} &= F_m(x) - \int_0^t \xi[\tau; x(t-\tau), p(t-\tau)] d\tau + R(t), \end{aligned} \quad (1.44)$$

with the memory kernel

$$\xi[\tau; x(t-\tau), p(t-\tau)] \equiv \sum_{i=1}^{\infty} \Phi_i [p(t-\tau), x(t-\tau)] \left\langle \frac{\partial \Phi_i(x, p)}{\partial p} R(0) R(\tau) \right\rangle \quad (1.45)$$

and the mean-field force

$$F_m(x) \equiv -\frac{\partial V_m}{\partial x} \quad (1.46)$$

$$\equiv \left\langle -\frac{\partial}{\partial x'} \{ V_S(x') + V_{S-B}(x', \{Q\}) \}; x \right\rangle. \quad (1.47)$$

Note that the noise $R(t)$ is a scalar and has no vector character anymore. This is because a particular noise term associated with the x component is vanishing. The details of the derivation of Eq. (1.44) from Eq. (1.43) are given in appendix B. In the canonical ensemble, the mean-field potential $V_m(x)$ can be calculated from the reduced distribution function $P(x)$ of the tagged coordinate x

$$V_m(x) = -kT \cdot \ln P(x), \quad (1.48)$$

as has been shown by Lange *et al.* [34] and as is proven explicitly in appendix B. The non-linear GLE as it is written down in Eq. (1.44) is valid for arbitrary systems and provides the general way to include non-linearities via the mean-field in Eq. (1.46). However, it is much more complex in comparison to the MBO-GLE and of less use in the present form. The latter is due to the functional dependence of the memory kernel on the coordinates of the system, see Eq. (1.45). In order to use this non-linear GLE one should make further approximations. The formal structure of Eq. (1.45) already suggests in which direction to proceed: truncating the (infinite) sum after a certain power in x and p . One can do this systematically by orthonormalizing the functions $x, p, x^2, p^2, xp, \dots$ up to the desired power to construct the orthonormal set $\Phi_j(x, p)$. This will result in a sequence of integral terms

$$- \sum_{i,j=1}^{\infty} \int_0^t \xi_{ij}(\tau) x^i(t-\tau) p^j(t-\tau) d\tau \quad (1.49)$$

and will for each combination i, j introduce an additional memory kernel $\xi_{ij}(t)$, which is itself a pure function of time. As an example, truncating the expansion after the first order (see appendix B) results in

$$\begin{aligned} \dot{x} &= \frac{p}{m} \\ \dot{p} &= F_m(x) - \int_0^t \xi(\tau) p(t-\tau) d\tau + R(t) \\ \langle R(0)R(t) \rangle &= mkT\xi(t), \end{aligned} \quad (1.50)$$

which is exactly of the MBO-GLE form with the only difference that the mean-force replaces the intramolecular system force [28]. In other words, the form of the MBO model is preserved in this first order approximation. Further, it has to be stressed that if one applied the non-linear projection to a system whose Hamiltonian is of

MBO form, i.e. Eq. (1.28), one would end up with the very same GLE as Eq. (1.34) since all higher order contributions in the integral kernel vanish [10]. Additionally, all the mean-field corrections become zero because of the special bilinear form of the system-bath coupling and the fact that bath modes are Gaussian distributed. In later discussions it is referred to the here-presented non-linear projection GLE as NLP-GLE. Additionally, the linearization of the NLP-GLE will be called linearized NLP-GLE.

1.1.4. Numerical Propagation of the GLE via Colored Noise Thermostats

For the numerical treatment of the GLE it is more useful to express it in a different form than that of Eq. (1.5) since a direct propagation requires time consuming operations. These are, on one hand, the calculation of the convolution integral, i.e. the dissipative force, at every time step. On the other hand generating the random forces directly with correct correlation as required by the FDT, Eq. (1.6), implies dealing with correlation matrices whose size grows from timestep to timestep. An elegant approach to overcome these performance problems is to express the GLE as an effective Markovian process by introducing auxiliary DOFs which are coupled to the system's coordinates. This formalism, which is referred to as *Colored Noise*, has been introduced by Ceriotti *et al.* [43–46] and is presented in the remainder of this section.

Consider the following stochastic process

$$\begin{aligned} \begin{pmatrix} \dot{p} \\ \dot{\mathbf{y}} \end{pmatrix} &= \begin{pmatrix} -\frac{\partial V}{\partial x} \\ \mathbf{0} \end{pmatrix} - \mathbf{A} \begin{pmatrix} p \\ \mathbf{y} \end{pmatrix} + \mathbf{B}\boldsymbol{\xi}(t) \\ &= \begin{pmatrix} -\frac{\partial V}{\partial x} \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_{py}^T \\ \mathbf{a}_{yp} & \mathbf{A}_{yy} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{y} \end{pmatrix} + \begin{pmatrix} \mathbf{b}_{p\xi}^T \\ \mathbf{B}_{y\xi} \end{pmatrix} \boldsymbol{\xi}(t), \end{aligned} \quad (1.51)$$

where x and p are the system coordinate and its conjugate momentum and the vector \mathbf{y} denotes auxiliary momenta. All momenta are linearly coupled via the so-called drift matrix \mathbf{A} , which has a block form consisting of a number a_{pp} , two vectors \mathbf{a}_{yp} , \mathbf{a}_{py}^T as well as a submatrix \mathbf{A}_{yy} . The number a_{pp} represents the Markovian friction coefficient. The vectors \mathbf{a}_{yp} and \mathbf{a}_{py}^T have the dimensionality N if N auxiliary momenta are included. These vectors describe the mutual coupling between the physical momentum and the auxiliary ones. Additionally, the auxiliary momenta \mathbf{y} are mutually coupled

via the submatrix \mathbf{A}_{yy} being of dimension $N \times N$. The vector $\boldsymbol{\xi}$ represents Markovian fluctuations, meaning that its components comprise Gaussian white noise with the properties given in Eq. (1.4). These noise components are linearly coupled via the diffusion matrix \mathbf{B} , which consists of an $(N + 1)$ -dimensional vector $\mathbf{b}_{p\xi}^T$ determining the coupling to the physical momentum and an $N \times (N + 1)$ -dimensional submatrix $\mathbf{B}_{y\xi}$ for the coupling to all components of \mathbf{y} . Treating the white noise as well as the trajectory $p(t)$ as an inhomogeneity for \mathbf{y} one can formally write down the solution $\mathbf{y}(t)$ under the assumption that $\mathbf{y}(0) = 0$

$$\mathbf{y}(t) = \int_0^t ds e^{-(t-s)\mathbf{A}_{yy}} [-\mathbf{a}_{yp}p(s) + \mathbf{B}_{y\xi}\boldsymbol{\xi}(t)] . \quad (1.52)$$

Substituting $\mathbf{y}(t)$ back into Eq. (1.51) yields for the remaining EOM for p

$$\dot{p} = -\frac{\partial V}{\partial x} - \int_0^t ds K(t-s)p(s) + R(t) , \quad (1.53)$$

where the memory kernel $K(t)$ and the force $R(t)$ read

$$\begin{aligned} K(t) &= a_{pp}\delta(t) - \mathbf{a}_{py}^T e^{-t\mathbf{A}_{yy}} \mathbf{a}_{yp} \\ R(t) &= \mathbf{b}_{p\xi}^T \boldsymbol{\xi}(t) - \int_0^t ds \mathbf{a}_{py}^T e^{-(t-s)\mathbf{A}_{yy}} \mathbf{B}_{y\xi} \boldsymbol{\xi}(s) . \end{aligned} \quad (1.54)$$

The force $R(t)$ constitutes the new stochastic force taking all auxiliary DOFs into account. Further, one can show that $R(t)$ satisfies the FDT when the matrices \mathbf{A} and \mathbf{B} have the relationship

$$mkT (\mathbf{A} + \mathbf{A}^T) = \mathbf{B}\mathbf{B}^T . \quad (1.55)$$

One might be surprised that, in order to propagate the GLE, one goes back to an extended variable space although the reduction of the dimensionality in the GLE originally was the key ansatz for making the propagation feasible. As will be seen later, just a few auxiliary DOFs are sufficient to cover a wide class of memory kernels.

Based on Eq. (1.51) one can derive an integrator for the time propagation using a Trotter factorization of the Liouville propagator [45, 46]

$$e^{\mathcal{L}\Delta t} = e^{\mathcal{L}_{py}\Delta t/2} e^{\mathcal{L}_p\Delta t/2} e^{\mathcal{L}_x\Delta t} e^{\mathcal{L}_p\Delta t/2} e^{\mathcal{L}_{py}\Delta t/2}. \quad (1.56)$$

The inner steps of the propagation exactly correspond to the usual velocity Verlet propagator

$$\begin{aligned} p &\leftarrow p - \frac{\partial}{\partial x} V(x) \cdot \frac{\Delta t}{2} \\ x &\leftarrow x + \frac{p}{m} \Delta t \\ p &\leftarrow p - \frac{\partial}{\partial x} V(x) \cdot \frac{\Delta t}{2}, \end{aligned} \quad (1.57)$$

whereas the outermost operations, $e^{\mathcal{L}_{py}\Delta t/2}$, can be expressed as

$$\begin{pmatrix} p \\ y \end{pmatrix} \leftarrow \mathbf{C}_1 \begin{pmatrix} p \\ y \end{pmatrix} + \sqrt{mkT} \mathbf{C}_2 \xi(t). \quad (1.58)$$

The propagation matrices can be shown to obey

$$\begin{aligned} \mathbf{C}_1 &= e^{-\Delta t/2 \cdot \mathbf{A}} \\ \mathbf{C}_2^T \mathbf{C}_2 &= \mathbf{I} - \mathbf{C}_1^T \mathbf{C}_1, \end{aligned} \quad (1.59)$$

where the matrix \mathbf{C}_2 is a lower triangular matrix calculated from a Cholesky decomposition given in the second line of Eq. (1.59).

Writing the GLE as an effective Markovian process leads to the advantage that a Verlet-like integrator can be used for its propagation being symmetric and local in time instead of a global and, hence, more time consuming integration scheme as would be required by the original form of the GLE, Eq. (1.5). Also the calculation of the completely uncorrelated random forces, $\xi(t)$, does not involve large correlation matrices to deal with. Nevertheless, one recognizes from Eq. (1.54) that the mapping of the GLE onto the multi-dimensional process, Eq. (1.51), only works for a special class of memory kernels $K(t)$, namely for those which can be written as a superposition of exponential forms

$$K(t) = \sum_i a_i \exp(-b_i \cdot t). \quad (1.60)$$

The fact that the matrix \mathbf{A}_{yy} can in principle have complex eigenvalues also allows for oscillatory damped functions via complex-valued coefficients b_i . As an example, one can show that the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & a \\ -a & b \end{pmatrix} \quad (1.61)$$

corresponds to the memory function

$$K(t) = a^2 \exp(-bt), \quad (1.62)$$

where only one additional degree of freedom must be included [46]. The matrix

$$\mathbf{A} = \begin{pmatrix} 0 & a & a \\ -a & b & \omega \\ -a & -\omega & b \end{pmatrix} \quad (1.63)$$

yields an oscillatory damped memory function of the form

$$K(t) = 2a^2 \exp(-bt) \cos(\omega t) \quad (1.64)$$

and requires two auxiliary momenta [46]. Further, one can show that if the matrices \mathbf{A} and $\tilde{\mathbf{A}}$ provide the memory functions $K(t)$ and $\tilde{K}(t)$ respectively, the matrix

$$\mathcal{A} = \begin{pmatrix} a_{pp} + \tilde{a}_{pp} & \mathbf{a}_{py}^T & \tilde{\mathbf{a}}_{py}^T \\ \mathbf{a}_{yp} & \mathbf{A}_{yy} & 0 \\ \tilde{\mathbf{a}}_{yp} & 0 & \tilde{\mathbf{A}}_{yy} \end{pmatrix} \quad (1.65)$$

corresponds to the sum of both memory kernels [46]. Hence, the matrices in Eqs. (1.61) and (1.63) can be used as building blocks for constructing a general memory kernel of the form given in Eq. (1.60).

1.2. Linear Absorption Spectroscopy

After having discussed the EOM that accounts for the system-bath coupling in a condensed phase environment, a brief summary will be given on how to probe the corresponding dynamics via linear absorption spectroscopy. The details of the underlying theoretical description can be found in the literature at many places, see e.g. [20, 40, 47].

The key quantity in linear absorption spectroscopy is the frequency-dependent absorption coefficient $\alpha(\omega)$, which according to the Lambert-Beer absorption law describes the intensity loss of monochromatic radiation with frequency ω and initial intensity I_0 after propagating through a medium of thickness l

$$I(l, \omega) = I_0 e^{-\alpha(\omega)l}. \quad (1.66)$$

The microscopic mechanism for the intensity loss is the coupling of the electromagnetic field to the particular motion of the molecules in the medium, which can be resonantly excited if the radiation frequency fits the frequency of a particular mode. Hence, the intensity loss probed by varying the radiation frequency unravels the frequency distribution of the atomistic motion within the medium. This is the central idea of linear absorption spectroscopy.

For interpreting the results of such absorption experiments a bridge must be built to a theoretical description of the underlying microscopic processes, which can be achieved by finding theoretical expressions for the absorption coefficient $\alpha(\omega)$. This is usually done starting from a full quantum description of the medium's polarization as a functional of the incoming electromagnetic field treating the light-matter interaction in dipole approximation [20, 40]. Then perturbation theory is performed in order to sort the particular contributions according to their power in the electromagnetic field [20, 40]. By restricting the perturbation series to first order one finds an expression for the linear absorption coefficient [47]

$$\alpha(\omega) \sim \frac{\omega}{\hbar} \left(1 - \exp \left[-\frac{\hbar\omega}{kT} \right] \right) I(\omega), \quad (1.67)$$

with the quantum lineshape function

$$I(\omega) = \frac{1}{2\pi} \text{Re} \int_{-\infty}^{\infty} \exp[-i\omega t] \text{Tr} \{ \hat{\rho}_{\text{eq}} \hat{\mu}(0) \hat{\mu}(t) \} dt. \quad (1.68)$$

Here, $\hat{\mu}$ stands for the dipole moment operator and $\text{Tr}\{\hat{\rho}_{\text{eq}} \bullet\}$ denotes the quantum-mechanical ensemble average with respect to the equilibrium density operator $\hat{\rho}_{\text{eq}}$. The overall expression under the Fourier integral represents a quantum dipole auto-correlation function (DAF). In the classical description of the medium based on a GLE, one needs to take the classical limit, i.e. $\hbar \rightarrow 0$, of Eq. (1.67). Then the quantum DAF becomes a classical one

$$\text{Tr} \{ \hat{\rho}_{\text{eq}} \hat{\mu}(0) \hat{\mu}(t) \} \longrightarrow \langle \mu(0) \mu(t) \rangle_{\text{cl}} = \int f_{\text{eq}}(x_0, p_0) \mu(0; x_0, p_0) \mu(t; x_0, p_0) dx_0 dp_0, \quad (1.69)$$

where it is integrated over the initial conditions x_0, p_0 weighted by the canonical equilibrium distribution function $f_{\text{eq}}(x_0, p_0)$. The classical expression for the absorption coefficient becomes

$$\alpha(\omega) \sim \frac{\omega^2}{kT} I_{\text{cl}}(\omega), \quad (1.70)$$

where $I_{\text{cl}}(\omega)$ is the classical lineshape function defined as the real part of the classical DAF's Fourier transform. In the one-dimensional case the dipole moment of a particular mode is given by

$$\mu(t) = q \cdot x(t), \quad (1.71)$$

with x being the mode coordinate (usually a bondlength). For this case, the DAF can be written as

$$\langle \mu(0) \mu(t) \rangle_{\text{cl}} = q^2 \langle x(0) x(t) \rangle_{\text{cl}}. \quad (1.72)$$

Using the relation

$$-\frac{d^2}{dt^2} \langle x(0) x(t) \rangle_{\text{cl}} = \frac{1}{m^2} \langle p(0) p(t) \rangle_{\text{cl}} \quad (1.73)$$

and the fact that a second time-derivative results in frequency factor ω^2 in Fourier domain, one can also calculate the absorption coefficient from the Fourier transform

of the momentum autocorrelation function (MAF) denoted as $S_{pp}(\omega)$

$$\alpha(\omega) \sim \frac{1}{kT} \text{Re}\{S_{pp}(\omega)\}. \quad (1.74)$$

In this thesis this is the basic formula to calculate the linear absorption spectra from the trajectories provided by explicit or implicit simulations.

1.3. Concluding Remarks

In this chapter the theory of the GLE has been introduced putting emphasis on its statistical-mechanical origin. It has been shown that the only way to justify the common form of the GLE, i.e. Eq. (1.5), is to involve the MBO model, which assumes a harmonic bath and bilinear system-bath coupling. However, in the general case one can derive prototype forms of the GLE by means of linear or non-linear projection operator techniques, which differ from the common GLE form in Eq. (1.5) such that the explicit system force is altered. In the LP-GLE the explicit force is exclusively linear meaning that all the anharmonicity in the real potential is projected into the bath. In the NLP-GLE the explicit force consists of a mean-force obtained from a conditional average of the total force over bath DOFs. Further, a functional dependence of the friction term on system coordinates arises, which can be brought into the common form of the non-Markovian friction, Eq. (1.5), in a first order expansion only.

It has been discussed that in certain situations like in non-linear spectroscopy or in a quantum vs. classical comparison the explicit force should be kept anharmonic and, hence, a non-linear GLE is preferred. In this case, the form of GLE with an anharmonic explicit force **cannot** be justified in full generality since either the friction integral in the NLP-GLE must be expanded up to first order, the MBO model must be invoked or anharmonicity has to be added in an *ad hoc* way after parametrizing the memory kernel for the LP-GLE. The applicability of these three approaches is the main subject of this thesis, and is probed via the ability of the (non-linear) GLEs to reproduce vibrational spectra calculated from explicit MD simulations.

Chapter 2.

Spectral Features of non-Markovian Dynamics

Armed with the theoretical understanding of the GLE as well as with the implementation of the numerical algorithm described in the previous chapter, the formalism is principally ready to be applied to real physical problems. However, it is convenient to first study some general physical properties of non-Markovian dynamics and to get hands-on experience with the physical interpretation of the observed effects. Further, it is sensible to investigate the error behavior of the physical quantities of interest: time-correlation functions (TCFs) and vibrational spectra. These two issues will be addressed in this chapter. Specifically, a statistical convergence analysis of TCFs computed from numerical GLE trajectories is employed in order to establish a reliable simulation protocol. Afterwards, the non-Markovian dynamics as such is investigated putting emphasis on the interplay of the spectral density and the system's dynamical properties in terms of TCFs and vibrational spectra. Throughout this chapter the validity of the GLE in its common form, i.e. Eq. (1.5) is assumed without questioning its principal applicability. The applicability of the GLE, especially in its non-linear forms, will be the subject of the next chapter.

2.1. Setting up a Colored Noise Simulation Protocol

2.1.1. The Block Averaging scheme

In this thesis the dynamical properties of interest are vibrational spectra computed from MAFs according to Eq. (1.74). In the standard protocol the MAFs are calculated employing both time averages along a single trajectory and a subsequent average over an ensemble of independent trajectories

$$C(k \cdot \Delta t) = \frac{1}{N_t} \sum_{i=1}^{N_t} C^{(i)}(k \cdot \Delta t). \quad (2.1)$$

Here, N_t labels the number of trajectories and $C^{(i)}(k \cdot \Delta t)$ the MAF calculated as a time-average along the i th trajectory comprising N timesteps of size Δt each

$$C^{(i)}(k \cdot \Delta t) = \frac{1}{N-k} \sum_{j=1}^{N-k} C_j^{(i)}(k \cdot \Delta t) \quad (2.2)$$

$$\equiv \frac{1}{N-k} \sum_{j=1}^{N-k} p^{(i)}(j \cdot \Delta t) \cdot p^{(i)}([j+k] \cdot \Delta t). \quad (2.3)$$

Principally, one can think of two different types of errors being present in the MAFs calculated in this manner. On one hand there are errors stemming from the numerical integration of the EOMs. On the other hand the GLE data are subject to statistical fluctuations requiring a meaningful number of samples (i.e. timesteps and trajectories) to be sufficiently averaged out. In practice, the numerical integration errors can be kept sufficiently small by properly choosing the timestep Δt of the simulation. In comparison to that, the statistical fluctuations of the individual samples are usually larger and thus the following error analysis will be restricted to this source of errors only. In this thesis, the statistical error of the MAF at the timestep k

$$\epsilon(k) = s \cdot \sigma(k) \quad (2.4)$$

is understood as a multiple of the standard deviation $\sigma(k)$ of the sample average. Note, that here and further the explicit notation of the timestep Δt will be omitted. The factor s determines the tolerance level and is here set to $s = 2$ leading to a tolerance of about 95%. According to Eqs. (2.1) and (2.2) the individual samples

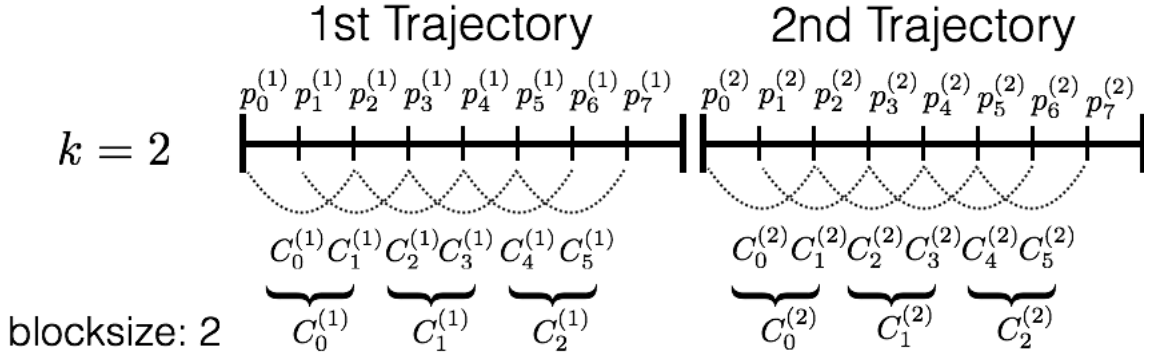


Figure 2.1.: The Block Averaging scheme is shown exemplarily for $k = 2$ and a block size of $b = 2$. The number of samples from which the standard deviation is calculated according to Eq. (2.5) is then reduced by half.

$C_j^{(i)}(k)$ being averaged to give the MAF at the step k are formed by all the products of momenta displaced by k timesteps. The total number of samples is thus given by $N_S(k) = N_t \cdot (N - k)$. A naive way of calculating the standard deviation would be to employ the estimator

$$\sigma(k) = \sqrt{\frac{1}{N_S(k)[N_S(k) - 1]} \sum_{i,j} \left(C_j^{(i)}(k) - C(k) \right)^2}. \quad (2.5)$$

However, this estimator gives reliable standard deviations only if the individual samples are uncorrelated – a condition that is clearly violated in the non-Markovian dynamics studied here. If the standard deviation was calculated from Eq. (2.5), the errors would be underestimated dramatically.

The correct way to estimate the standard deviation in the presence of correlation is to group the samples within a trajectory into n blocks of size b and to calculate the average $C_l^{(i)}(k)$, $l = 1 \dots n$ of all samples within the blocks. Note, that the index l labels blocks within the i th trajectory. The resulting block averages are then treated as uncorrelated numbers and the standard deviation of the sample average is calculated via Eq. (2.5). This procedure known as Block Averaging [48] is illustrated in Fig 2.1. Since the standard deviations depend on the blocksize b one should repeat the Block Averaging for systematically increased values of b . The standard deviation will eventually arrive at a plateau, which can be regarded as the meaningful estimate for $\sigma(k)$ in the presence of correlation between the samples, see Fig. 2.2 in the next section.

2.1.2. Statistical Convergence Analysis

The statistical convergence analysis is performed for a simple guinea pig model being a harmonic oscillator with the frequency $\omega_0 = 0.4$ coupled to a bath described by the memory kernel $\xi(t) = 2a^2 \exp[-bt] \cos(\omega t)$ with the parameters $a = 0.03$, $b = 0.03$ and $c = 0.4$. On one hand this system constitutes a simple model for a solute mode being the main objects of investigations in the next chapter. On the other hand this system provides an exact reference MAF, which can be compared against the numerical GLE results. As shown in appendix D the MAF's Fourier transform for the harmonic oscillator is available analytically and reads

$$S_{pp}(\omega) = \frac{\omega}{\omega J(\omega) + i(\omega^2 - \omega_0^2)}, \quad (2.6)$$

with the spectral density $J(\omega)$ defined in Eq. (1.7). Exact references in time domain can be computed from a numerically exact inversion of the Fourier transform.

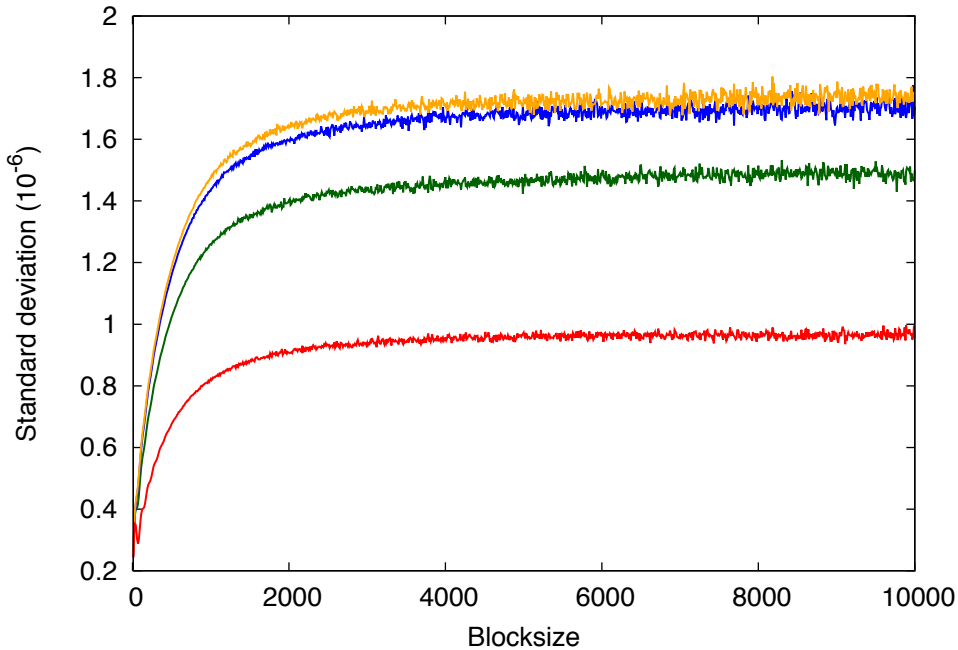


Figure 2.2.: Estimates for the standard deviations at the minima of the MAF as a function of the blocksize, are shown for $k = 350$ (red), $k = 658$ (green), $k = 2612$ (blue) and $k = 3914$ (orange), with k defined in Eq. (2.2).

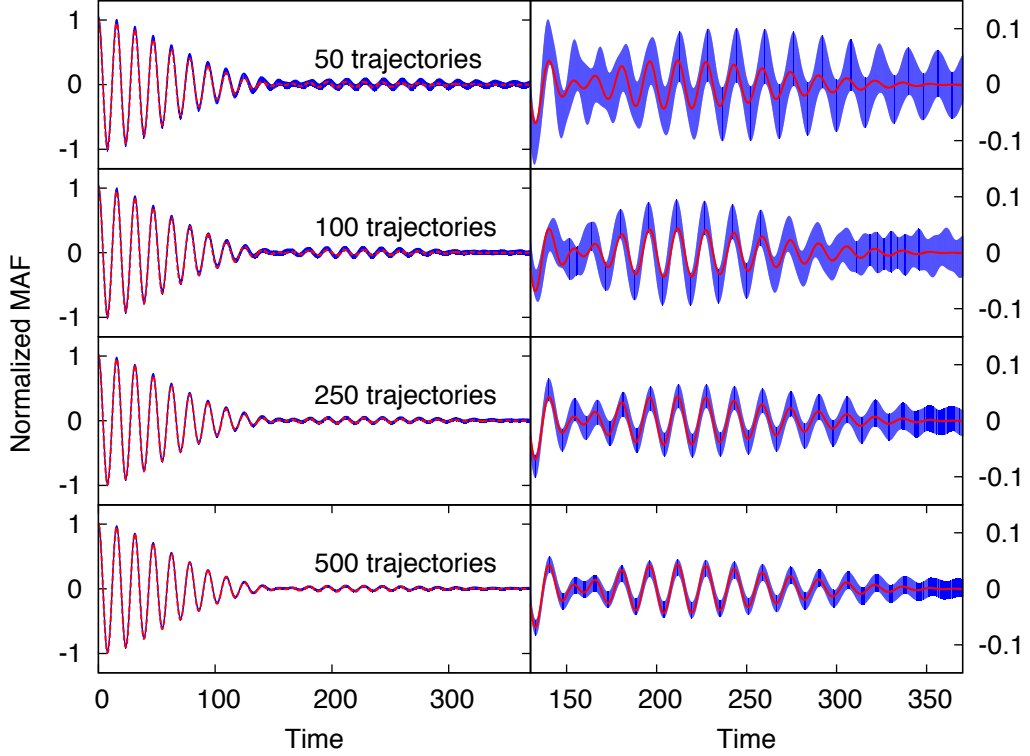


Figure 2.3.: The numerical MAFs (blue) are shown for different numbers of trajectories together with the exact reference (red). The thickness of the blue curves stands for the trusted interval of 2σ around the calculated points and hence measures the statistical error according to Eq. (2.4). The right panels zoom in on the range 150 – 350 where a small revival occurs.

For the analysis of the convergence two numerical parameters have to be taken into account: the number of trajectories and the number of timesteps per trajectory. According to the Block Averaging scheme it is sensible to set the number of timesteps at least to the minimal blocksize needed for a proper estimate of the standard deviation. This minimal blocksize corresponds to the characteristic correlation length in the system and can be determined by blocking a sufficiently long trajectory. In Fig. 2.2 the blocking of one long GLE trajectory of 10,000,000 timesteps has been performed. The standard deviations of the MAF are probed in dependence of the blocksize for a selection of timesteps k . It becomes visible that in all cases the standard deviation reaches a plateau at a blocksize of about $N_{\text{corr}} = 4000$ timesteps, which should be chosen as the minimal length of the trajectories. In order to increase the statistics at the tail of the MAF the number of timesteps is set to $N = 5N_{\text{corr}} = 20000$. For the following convergence analysis with respect to the number of trajectories the statistical errors are

calculated involving a blocksize of $b = N = 20000$, which implies that no additional blocking within one trajectory is required.

In Fig. 2.3 the numerical MAFs for different numbers of trajectories (blue curves) are plotted within their statistical errors together with the exact reference (red curves). In all cases the statistical errors are extremely small at the beginning of the MAF, which shows that for this part already 50 trajectories are more than enough. However, at larger times, where a small revival occurs, the error becomes larger and hence more than 50 trajectories should be involved in order to resolve this characteristic feature with sufficient accuracy. Here, 500 GLE trajectories seem to be the optimal choice. The Colored Noise propagation scheme, hence, reveals itself as an extremely efficient algorithm providing MAFs that converge fast with respect to the number of trajectories to be employed.

2.2. Non-Markovian Spectra and Time-Correlation Functions

Exploring the wide range of non-Markovian effects in dissipative dynamics can be a very far reaching undertaking. This is because the spectral density, the fundamental quantity characterizing the system-bath interaction, can possess very different forms depending on the physical situation under study for which there exist different theoretical models. The Ohmic spectral density, for instance, possesses an exponential decaying form $\text{Re}\{J(\omega)\} \sim e^{-\omega/\omega_c}$ [40] with ω_c representing a characteristic cut-off frequency. A common model for polar solvents, the Debye spectral density, consists of a Lorentzian form $\text{Re}\{J(\omega)\} \sim 1/(\omega^2 + \omega_c^2)$ centered at zero frequency [20, 40]. The two aforementioned models describe a monotonous decay with the main contributions being located at low frequencies and are often used to describe electron or exciton transfer dynamics in dissipative environments [40]. However, for vibrational spectroscopy of solute dynamics in solvents they do not reflect the correct characteristic behavior. In anticipation of the results in Chap. 3 the spectral densities for the solvents discussed in this thesis rather comprise peaked contributions localized around distinct frequencies. Their typical functional form can be better described by

Lorentzian functions

$$\text{Re}\{J(\omega)\} = a^2 \frac{b}{b^2 + (\omega - c)^2} + a^2 \frac{b}{b^2 + (\omega + c)^2}, \quad (2.7)$$

where the parameter c determines the frequency position, b the width and a the magnitude. For this reason, the subsequent discussion of the non-Markovian effects will be restricted to this functional form only. Note, that the memory kernel $\xi(t)$ corresponding to $J(\omega)$ in Eq. (2.7) reads $\xi(t) = 2a^2 e^{-bt} \cos(ct)$ and therefore perfectly matches the function type required for employing the Colored Noise propagation scheme (see Sec. 1.1.4).

For setting the stage of the subsequent investigations a general spectroscopic criterion for the presence of non-Markovianity is defined first. Afterwards, the main question of this section is addressed: how does the Lorentzian spectral density in Eq. (2.7) impact spectra and TCFs in terms of its magnitude, width and position relative to the system frequency? The observed non-Markovian features will be first illustrated on the basis of the same harmonic oscillator as in the previous section. Afterwards, the investigations are extended to an anharmonic regime. Note that, although no units will be specified, the presented parameters and quantities are given values in realistic units for typical solutes. These units are fs for time, fs⁻¹ for frequency and K for temperature. If desired, the reader can think of all parameters and quantities to be given in these special units.

2.2.1. Spectroscopic Criterion for non-Markovian Dynamics

Consider the Markovian dynamics of a harmonic oscillator according to the original Langevin equation, Eq. (1.3), where the spectral density has the special form $J(\omega) = \gamma = \text{const}$. The corresponding spectrum follows from Eq. (2.6) as

$$\begin{aligned} \alpha(\omega) &\sim \text{Re}\{S_{pp}(\omega)\} \\ &= \frac{\omega^2 \gamma}{\omega^2 \gamma^2 + (\omega^2 - \omega_0^2)^2}. \end{aligned} \quad (2.8)$$

It becomes apparent that Markovian dynamics provides a very specific lineshape which is shown in Fig. 2.4 via the black curves. In contrast, the spectral density for non-Markovian dynamics can, in principle, have an arbitrary form and thus provides a variety of qualitatively different lineshapes. Based on this observation one can for-

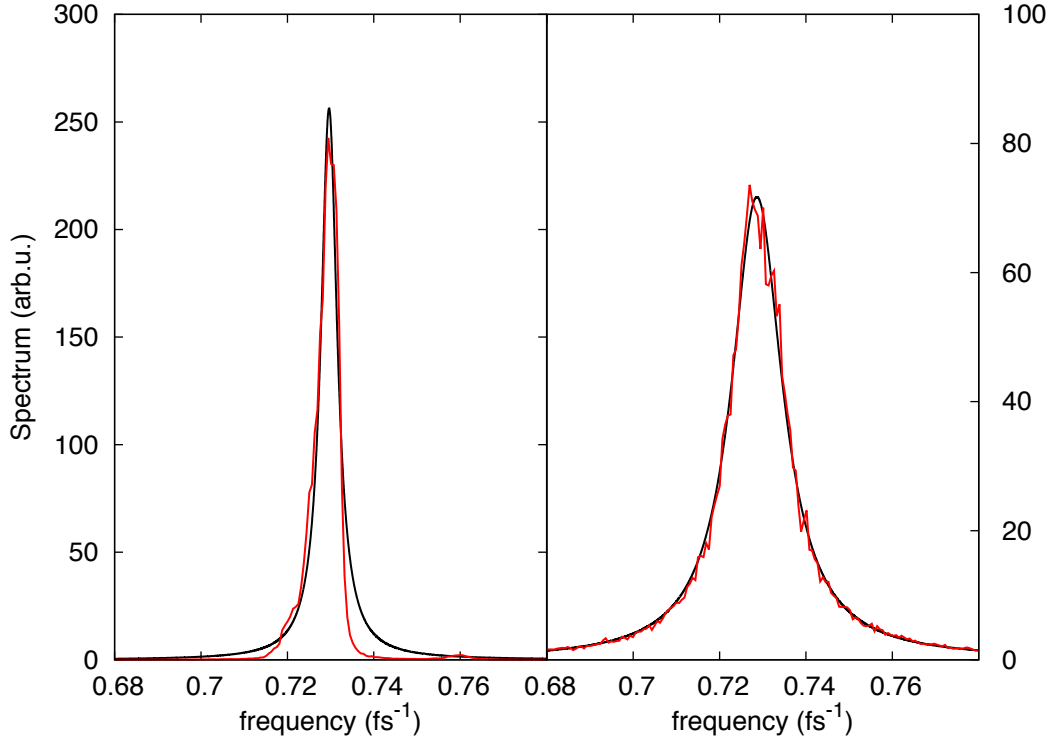


Figure 2.4.: The vibrational spectra of an anharmonic OH-stretch in a Markovian surrounding with $\gamma = 0.001$ (left panel) and $\gamma = 0.025$ (right panel) are plotted in red. The data have been obtained via a LD simulation according to Eq. (1.3) and are adopted from the previous work [39]. The black curves constitute fits to the functional form of the Markovian spectrum for a harmonic oscillator given in Eq. (2.8).

ulate a criterion for the presence of non-Markovianity in the dynamics under study: *whenever the linear absorption spectrum cannot be fitted to the special functional form in Eq. (2.8) the dynamics is non-Markovian*. The opposite of this statement is equally true: *whenever the linear absorption spectrum can be fitted to the special functional form in Eq. (2.8) the dynamics is Markovian*.

Since the derivation of the lineshape in Eq. (2.8) is based on the assumption that the system is harmonic one might argue about the general validity of the criteria formulated above. However, as has been discussed in Sec. 1.1.2 any system can be mapped onto an effective harmonic oscillator via linear projections resulting in a LP-GLE. In this sense, the aforementioned criterion for non-Markovianity can be seen as generally valid for anharmonic systems as well. Nevertheless, special care should be taken in the physical interpretation of the lineshape since the counter-intuitive system-

bath partitioning in the LP-GLE can be misleading in certain situations. To illustrate this further the linear absorption spectrum of an anharmonic OH-stretch vibration in a truly Markovian bath is shown in Fig 2.4 for small (left panel) and high friction (right panel). The spectral lines have been fitted to the Markovian functional form given in Eq. (2.8). In the high friction regime the fit perfectly coincides with the spectral line and thus, the criterion correctly states Markovianity of the underlying dynamics. In contrast, the spectrum in the low friction regime has an asymmetric shape being slightly skewed to the red. As can be seen in Fig. 2.4 it cannot be fitted to the functional form in Eq. (2.8) and thus, according to the criterion given above, the system-bath interactions would be declared non-Markovian. From the physical perspective, however, it is clear that the asymmetric lineshape is a consequence of the anharmonicity in the system potential alone and has nothing to do with a non-Markovian behavior of the bath [39]. From the rather counter-intuitive viewpoint of the effective harmonic LP-GLE, anharmonicity is seen as a part of the bath and has to be accounted for via a non-Markovian GLE.

2.2.2. Influence of a non-Markovian Spectral Density

Having established the criterion for non-Markovianity in harmonic systems, we are in position to discuss the impact of a Lorentzian spectral density on vibrational spectra and MAFs. Fig. 2.5 contains the vibrational spectrum and the corresponding MAF of a harmonic oscillator with frequency $\omega_0 = 0.4$ for the different spectral density parameters a , b and c according to Eq. (2.7). In the panels a) – f) the spectral density is located around the system frequency, describing a resonant coupling of the system to the bath. In this regime the spectrum generally has a double peak structure with the two maxima lying above and below the harmonic frequency. The corresponding MAFs possess several revivals and an overall dephasing behavior. Increasing the spectral density's magnitude, see panels a) – c), amounts to a stronger separation between the two maxima in the spectra and an increased number of revivals in the MAF. The overall dephasing time appears to be independent on the magnitude. Small spectral density widths, see panels d) – f), lead to sharply peaked spectra with clearly distinguishable maxima. When increasing the width the two peaks become more and more indistinguishable and the spectral line gets broader. In the corresponding MAF one then observes less revivals and an overall faster dephasing. In the panels g) – i) the spectral density has been shifted with respect to the system frequency corresponding

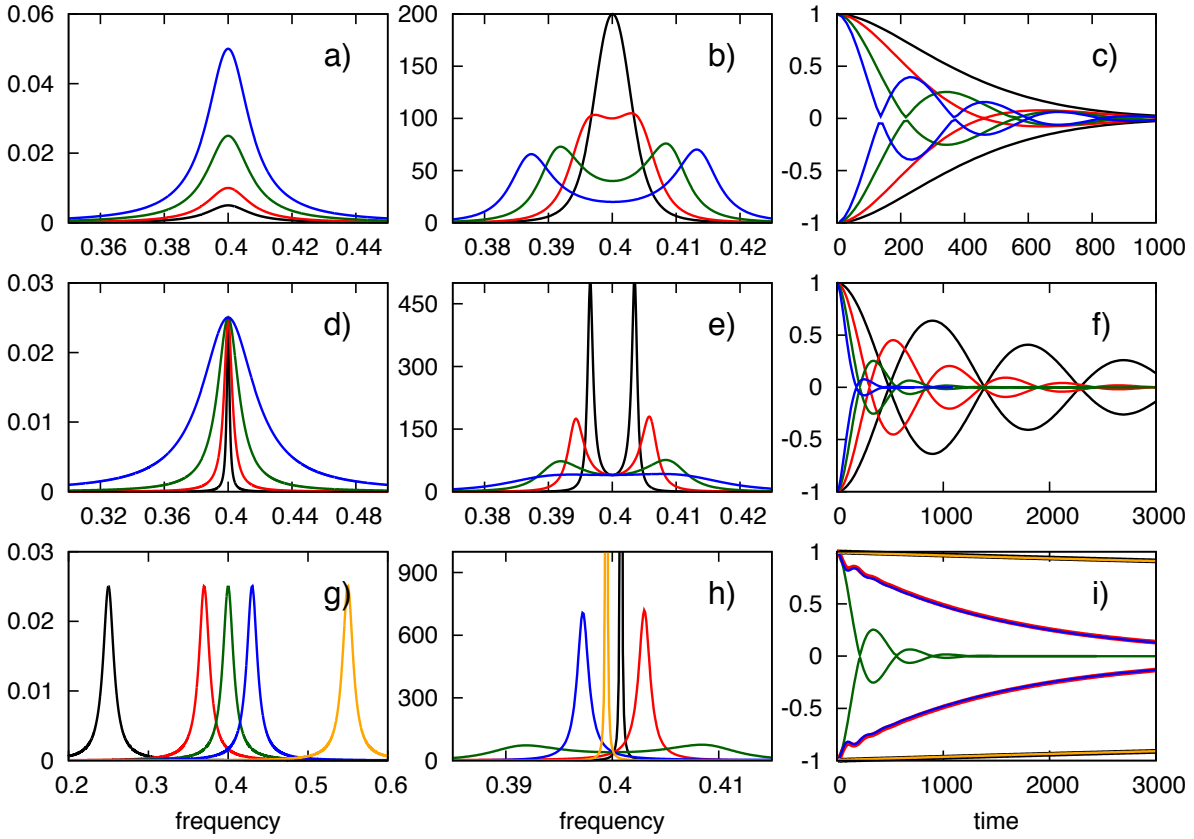


Figure 2.5.: Spectral density (panels a, d, g), vibrational spectrum (panels b, e, h) and MAFs (panels c, f, i) are shown for different spectral density's magnitudes (panels a-c), widths (panels d-f) and peak position (panels g-i). The color code within one line is such that all spectral densities, spectra and MAF carrying the same color belong together. The spectra are obtained from the analytic formula in Eq. (2.6) and the corresponding MAFs by a numerical Fourier back-transform. For better visibility only the MAF's envelopes are plotted. The frequency of the harmonic oscillator is $\omega_0 = 0.4$.

to an off-resonant coupling to the bath. Here, it is visible that the spectrum loses its symmetric double-peak structure with increasing displacement. Further, the spectrum becomes narrow and is more closely located to the oscillator frequency. In the MAF the beat structure disappears and the dephasing gets significantly slower.

The aforementioned observations can be interpreted with the help of a simple physical picture, which is mainly motivated by the imagination of harmonic bath oscillators in the spirit of the MBO model in Sec. 1.1.3. If the bath couples resonantly to the sys-

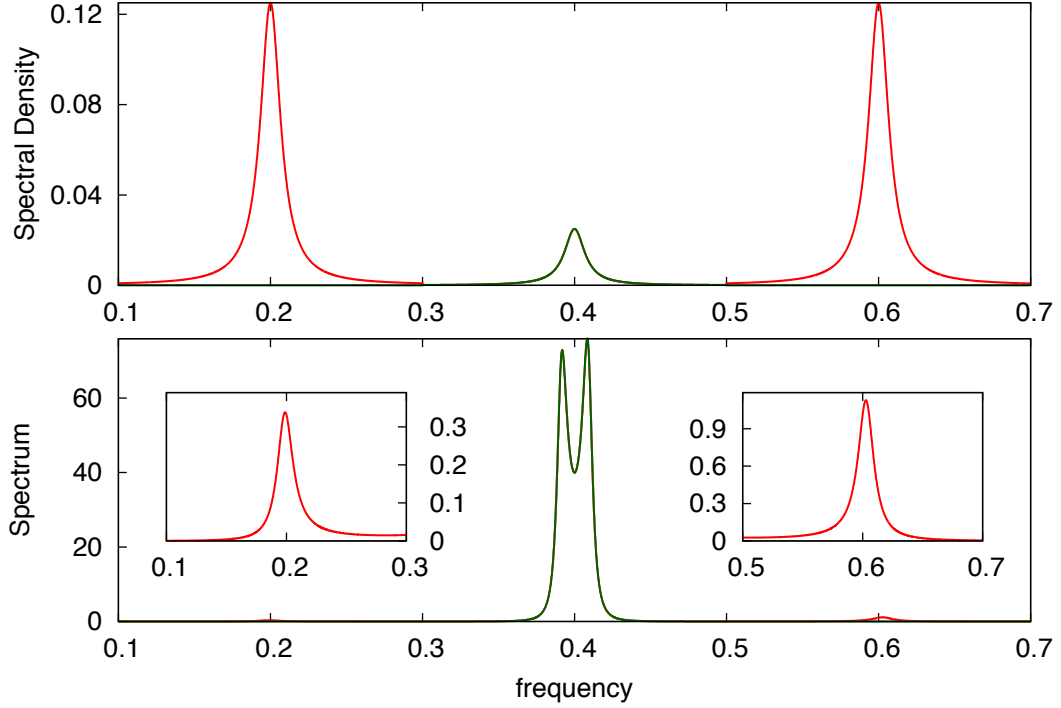


Figure 2.6.: A spectral density having one resonant (green) and two off-resonant Lorentzian contributions (red) is shown in the upper panel. The lower panel shows the corresponding spectrum. The green curve therein entirely stems from the resonant coupling to the bath and remains unchanged when the off-resonant peaks in the spectral density are removed. Adding the off-resonant contributions to the spectral density only amounts to small side peaks in the spectrum (red curves) shown in an enlarged format in the two insets.

tem, panels a) – f), one obtains a regime of energy transfer similar to that of two coupled pendulums: the vibrational energy is exchanged periodically between system (first pendulum) and bath (second pendulum). The particular reflux of energy from the bath into the system can be observed in the corresponding MAFs as periodic revivals. Mathematically, such revivals are obtained by interfering two oscillations of slightly different frequencies. Consequently, the corresponding spectrum has the double peak structure as observed above. The energy exchange is favored by a stronger coupling to the environment, which amounts to a higher frequency of revivals in the MAF and thus to a stronger peak separation in spectra, see panels a) – c) of Fig. 2.5. The impact of the spectral density's width, panels d) – f), can be connected to the MAF's decay: if the frequency distribution of the bath becomes broader, the energy is distributed over neighboring, off-resonant modes as well. The respective off-resonant

back coupling destroys the phase relation between system and resonant bath mode and thus leads to faster dephasing, less revivals and broader spectra. When increasing the width further the revivals in the MAF and the double peaks in the spectra will disappear completely and one approaches the Markovian regime, see blue curve in panel e). In contrast, if the bath's frequency distribution becomes narrow, the energy is more and more concentrated in the resonant bath mode. Then, a fixed phase relation between system and bath is maintained for a longer time leading to a slowly decaying MAF and narrow peaks in the spectrum. Approaching the limit of a vanishing width the dephasing is expected to disappear completely. In the regime of off-resonant coupling, panels g) – i), the oscillators cannot exchange energy efficiently. Hence, their particular motion does not influence each other and the oscillators evolve undisturbed. This results in slowly decaying MAFs and very narrow spectra comparable to those of a gas phase regime.

The distinction whether the bath couples resonantly or off-resonantly to the system is of fundamental importance in what follows. Since off-resonant coupling weakly perturbs the system one can draw the conclusion that off-resonant peaks in the spectral density can in general be neglected. This statement is underlined by Fig. 2.6, where the spectral contributions stemming from off-resonant coupling are compared to those from resonant coupling. Here, it becomes apparent that although the magnitude of the off-resonant parts in the spectral density are five times larger than the resonant ones, the corresponding off-resonant peaks in the spectrum are two orders of magnitude smaller than the resonant ones. This provides a great simplification when spectral densities of realistic systems, which usually comprise a lot of distinct peaks (see Sec. 3.2), shall be fitted. Restricting the fit to the resonant frequency region reduces the number of fit functions as well as the number of auxiliary momenta to be introduced in the Colored Noise propagation scheme significantly.

2.2.3. Anharmonic spectral Regime

So far the interplay of spectral densities and corresponding spectra has been discussed for a harmonic regime only. In this section the investigations are extended to an anharmonic system which is represented by a particle of unit mass in a Morse potential

$$V(x) = D [1 - \exp(-\alpha(x - x_0))]^2, \quad (2.9)$$

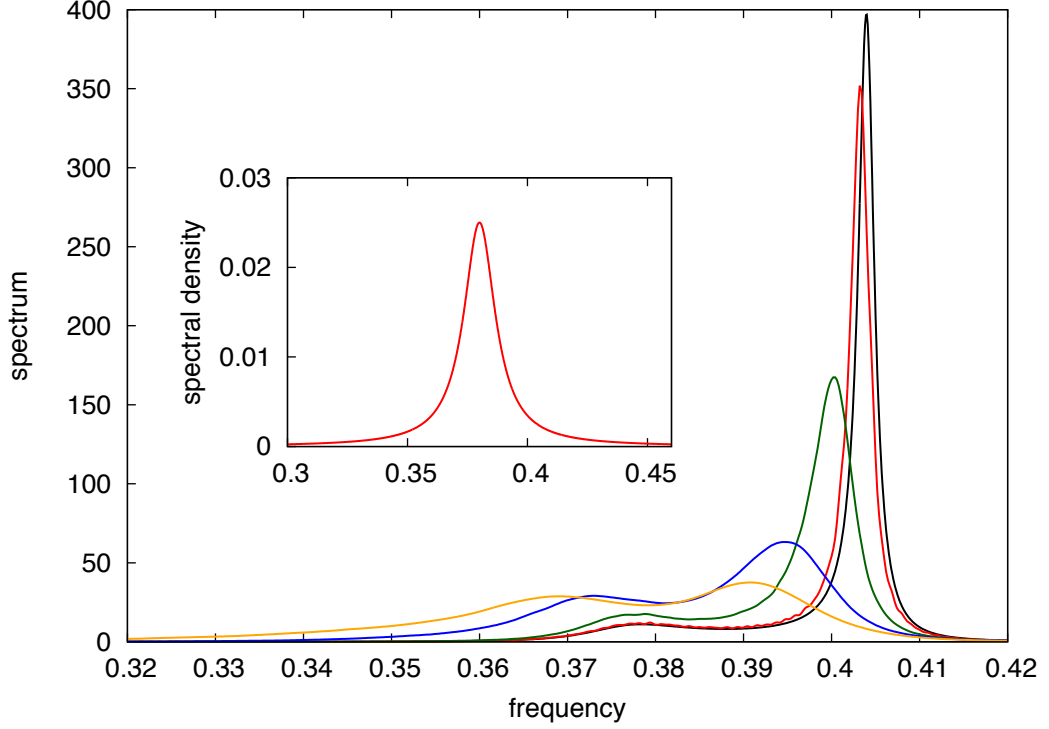


Figure 2.7.: The anharmonic spectra are compared for different temperatures. The spectral density, displayed in the inset, is located at $c = 0.38$ and is thus slightly red-shifted from the harmonic frequency of $\omega_0 = 0.4$. The colors for anharmonic spectra correspond to red: $T = 50$, green: $T = 300$, blue: $T = 900$ and orange: $T = 1500$. The reference harmonic spectrum is shown in black.

with the parameters $D = 0.02$, $\alpha = 2.0$ and $x_0 = 0$. The resulting harmonic frequency $\omega_0 = \alpha\sqrt{2D/m} = 0.4$ coincides with the one used in the previous section for the harmonic oscillator. The main anharmonic feature is that the frequency of the Morse-oscillator depends on its vibrational energy E [39]

$$\omega(E) = \alpha\sqrt{\frac{2(D-E)}{m}}. \quad (2.10)$$

This energy dependence provides the key difference between the harmonic and anharmonic regime: an explicit temperature dependence of the spectral lineshapes. Note that although the temperature enters the GLE formalism through the FDT, Eq. (1.6), the resulting lineshape of a harmonic oscillator in Eq. (2.6) does not depend on it.

In Fig. 2.7 the vibrational spectrum of the Morse oscillator is shown for different temperatures together with the one of its harmonic counterpart. The herein used

spectral density is red-shifted ($c = 0.38$) from the harmonic frequency, see inset in Fig. 2.7. The harmonic spectrum therefore possesses a sharp gas phase peak and a small wing originating from the off-resonant coupling to the bath. The anharmonic oscillator shows almost the same behavior at $T = 50$, see red curve in Fig. 2.7. Upon increasing the temperature the spectrum shifts to lower frequencies, becomes significantly broader and the magnitude of the off-resonant peak grows. This effect is a consequence of the explicit energy dependence of the Morse oscillator frequency. Increasing the temperature implies that, according to Eq. (2.10), lower frequencies of the Morse oscillator are probed. Within the picture that emerged in the previous section this means that the spectral density, which interacted off-resonantly at low temperatures, becomes resonant with increasing temperature. Hence, the spectral line adopts the double-peak structure observed in the previous section for the resonant coupling regime, see blue and orange curves in Fig. 2.7.

2.3. Concluding Remarks

In this chapter two aspects have been discussed. First, a statistical error analysis via the Block Averaging technique was performed in order to check the convergence of TCFs computed from the Colored Noise propagating scheme and to establish a reliable simulation protocol. The number of timesteps per trajectory should be at least equal to the characteristic correlation length. If the number of timesteps was set to five times the correlation length, then 500 GLE trajectories were sufficient in order to achieve TCFs with appropriate accuracy. Second, the properties of a non-Markovian dynamics have been investigated. Here, a spectroscopic criterion for the presence of non-Markovianity in the dynamics under study has been formulated for harmonic systems. According to this criterion, non-Markovianity is present whenever the spectral line cannot be fitted to the functional form given in Eq. (2.8). Further, the interplay between the spectral density and the resulting vibrational spectra and MAFs has been analyzed. The discussions were limited to Lorentzian-type spectral densities, which are appropriate for describing solute dynamics in liquid solvents as will be studied in the next chapter. The non-Markovian features have been investigated on the basis of a purely harmonic system first. It has been found that for a resonant system-bath coupling the energy transfer between system and bath can be interpreted in analogy to that of two coupled pendulums providing an intuitive picture for interpreting the physical observations. Further, it has been shown that the resonant coupling regime

yields much stronger contributions than those from off-resonant coupling, which can be neglected accordingly. This is an advantage since in the particular applications considered later the spectral densities could be restricted to the resonant frequency region only. Finally, a comparison to an anharmonic system has been made. The qualitative understanding of the anharmonic oscillator can be obtained with the same physical picture of resonant and off-resonant coupling between system and bath. The only difference to the harmonic regime is that one has to account for a temperature-dependent distribution of frequencies and hence a temperature dependence of vibrational spectra emerges.

Chapter 3.

Applications to vibrational Spectroscopy of Solute Dynamics

After having investigated general properties of non-Markovian dynamics provided by the GLE it is now time to apply the formalism to realistic dynamics. Specifically, the central question of this thesis, the ability of the three non-linear GLEs to describe vibrational spectra of realistic solutes in solvent environments, is addressed in this chapter. Conceptually, the applicability is determined as follows. First, force field-based MD simulations of solute and solvent are employed taking the full system-bath interactions into account. Then, the system is simulated via the different non-linear GLEs proposed in Chap. 1. Their applicability is judged upon the ability to reproduce explicit MD spectra. Although the LP-GLE does not provide a proper stage for a quantum vs. classical comparison and, hence, is not of main interest in this thesis, it is still employed for cross-checking purposes. In order to employ GLE simulations one needs a parametrization of the spectral density which correctly mimics the solvent's influence on the system. Thus, a method for extracting the proper spectral density from explicit MD data is introduced first. Finding a reliable protocol for this purpose is neither trivial nor straightforward due to a strong accumulation of numerical errors. Therefore the developed method is supported by a detailed error analysis.

3.1. Spectral Densities from explicit MD Simulations

The problem of extracting spectral densities from explicit MD simulations emerges in various branches of physics and chemistry. In particular the spectral density enters

the expressions for chemical reaction rates [12], whereas in studies of vibrational relaxation it is needed for estimating characteristic relaxation times [16–19, 26, 31] to mention but two. The common approach is to calculate the memory kernel as the correlation function of the forces exerted by the surrounding on the frozen vibrational mode [16, 17, 19]. In the framework of the MBO model this procedure is justified by the FDT, Eq. (1.6), as the memory kernels are entirely given by the system-bath coupling strengthes, which are independent on the system. In contrast, for the GLEs derived from projection operator techniques the memory kernel does depend on the system’s behavior and, hence, deviations from the common approach are expected. However, Tuckerman and Berne have shown that the common approach is justified if the system’s frequency is much larger than those of the bath modes [31]. Since the system frequencies studied in this thesis are comparable to that of the bath the validity of this approach becomes questionable. Therefore, an aim of having a more general protocol becomes apparent. A procedure without restrictions to such frequency separations has been developed in the framework of this thesis.

3.1.1. Method

The method which shall be developed here is based on the integro-differential equation for the momentum autocorrelation function (MAF) $C_{pp}(t) = \langle p(0)p(t) \rangle$

$$\dot{C}_{pp}(t) = C_{pF}(t) - \int_0^t \xi(t - \tau) C_{pp}(\tau) d\tau. \quad (3.1)$$

As it is shown in appendix C such an equation can be derived from any GLE being of the form in Eq. (1.5), which, in this thesis, are the LP-GLE, the MBO-GLE and the NLP-GLE with linearized memory functional. In Eq. (3.1) the function $C_{pF}(t) = \langle p(0)F(t) \rangle$ denotes the momentum-force cross-correlation (MFC), with $F(t)$ being the explicit part of the force. This force consists of the effective harmonic force in the LP-GLE, the intramolecular force in the MBO-GLE and the mean-force in the linearized NLP-GLE. Now, the basic idea is to use Eq. (3.1) for calculating the proper memory kernel from the MAF and MFC computed via MD simulations where the system-bath coupling is accounted for explicitly. Note that for the *ad hoc* anharmonic GLE, which is also considered here, the spectral density is simply adopted from the LP-GLE.

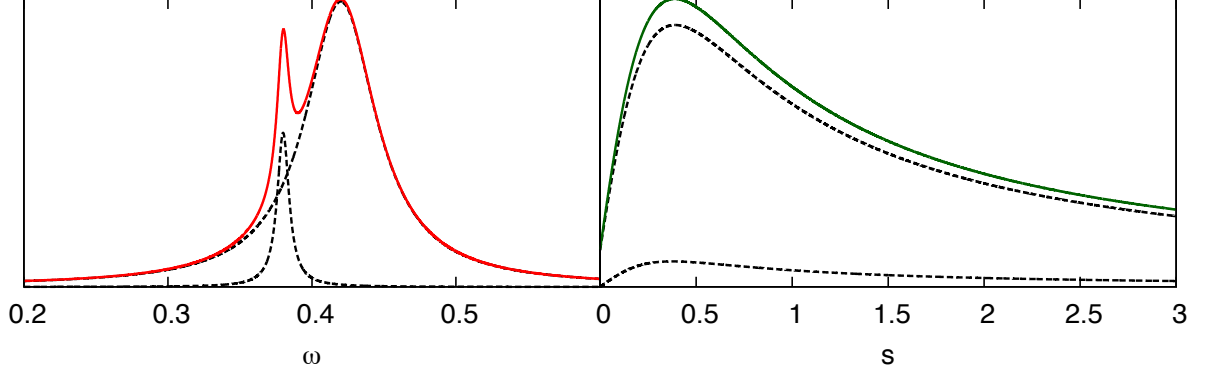


Figure 3.1.: Fourier transform (left panel) and Laplace transform (right panel) of a memory kernel consisting of a superposition of two exponentially damped cosine functions. Their individual contributions are plotted in black dotted lines.

There exist different numerical schemes to solve Eq. (3.1) for the memory kernel $\xi(t)$. On one hand an iterative scheme can be derived from a discretization of Eq. (3.1) in time domain [32–35]. Other methods are based on transforming Eq. (3.1) into Laplace domain, where the convolution becomes a product and, hence, the integro-differential equation turns into a simple algebraic equation, which can be easily solved for the memory kernel [34, 36, 37]. However, the back transform of the memory kernel into the time domain is a rather delicate task, since it involves a contour integration in the complex plane [49]. In practice, one therefore postulates certain functional forms to which the MAF and MFC are fitted and from which the inverse Laplace transform of the memory kernels is available analytically [34].

In this thesis a slight reformulation of the Laplace domain technique is used. Since the Colored Noise scheme for solving the GLE (see Sec. 1.1.4) requires a fit of the kernel to exponentially damped forms, whose Laplace transforms are well-known, one can omit the complicated transform back into time domain and perform the fits directly in Laplace domain. Equation (3.1) transformed into Laplace domain reads

$$sS_{pp}(s) - C_{pp}(0) = S_{pF}(s) - J(s) \cdot S_{pp}(s), \quad (3.2)$$

where S denotes the Laplace-transformed correlation function C

$$S(s) = \int_0^\infty \exp[-st]C(t)dt, \quad (3.3)$$

$J(s)$ stands for the transformed memory kernel and s is the Laplace variable. Note that a time derivative turns into a multiplication by s in Laplace domain. For the sub-

sequent fitting procedure and for a more intuitive interpretation of the memory kernel it is helpful to represent it in Fourier domain, where it is referred to as the spectral density $J(\omega)$. The advantage of the Fourier domain compared to the Laplace domain is illustrated in Fig. 3.1. Here, the Laplace and Fourier transforms of two oscillatory damped memory kernels (Eq. (1.64)) and their superpositions are plotted. It can be seen that the differences of the two contributions are significantly more pronounced in the Fourier domain. Thus, the Laplace transforms of the memory kernels are expected to be numerically less sensitive to such differences. The transition into Fourier domain can be achieved by setting the Laplace variable imaginary, i.e. $s \equiv i\omega$. Then, the Laplace transform becomes equal to the half-sided Fourier transform. Solving Eq. (3.2) for the spectral density yields

$$J(\omega) = \frac{1 + S_{pF}(\omega)}{S_{pp}(\omega)} - i\omega, \quad (3.4)$$

which forms the basic formula for calculating the spectral density from explicit MD simulations. Note that in Eq. (3.4) both $C_{pp}(t)$ and $C_{pF}(t)$ have been normalized to the value $C_{pp}(0)$.

The formula derived above is valid for the GLE with arbitrary (non-linear) forces. A simplification can be achieved when the force is harmonic. In this case the MFC, $C_{pF}(t) = -m\omega_0^2 C_{px}(t)$, can be expressed as

$$C_{pF}(t) = -\omega_0^2 \int_0^t C_{pp}(\tau) d\tau, \quad (3.5)$$

where ω_0 is the harmonic frequency and m is the system's mass. Substitution into Eq. (3.1) yields

$$\dot{C}_{pp}(t) = - \int_0^t K(t - \tau) C_{pp}(\tau) d\tau, \quad (3.6)$$

where the memory kernel $K(t) = \xi(t) + \omega_0^2$ is defined. The corresponding spectral density, labelled $J_K(\omega)$, is then computed via

$$J_K(\omega) = \frac{1}{S_{pp}(\omega)} - i\omega. \quad (3.7)$$

This form provides an advantage when the spectral density for the LP-GLE shall be computed since a calculation of the MFC can be omitted. Further, the effective harmonic frequency $\tilde{\omega}$, see Eq. (1.26), can be calculated from $J_K(\omega)$ without knowledge of $\langle x^2 \rangle$ as from the definition of $K(t)$ it follows that

$$J_K(\omega) = J(\omega) + \pi \tilde{\omega}^2 \delta(\omega) - i \frac{\tilde{\omega}^2}{\omega}. \quad (3.8)$$

The hyperbola in the imaginary part is a consequence of taking the half-sided Fourier transform only. This suggests that the effective harmonic frequency is accessible via fitting the hyperbola in the imaginary part of $J_K(\omega)$.

The practical use of Eqs. (3.4) and (3.7) encounters different sources of numerical errors which need to be reduced. On one hand MAFs and MFCs usually possess large statistical errors in their unconverged tails, which mainly cause noisy results. On the other hand the numerical transform into frequency domain involves integration errors. In the following, two ways are proposed to overcome these numerical problems. As a first approach Gaussian filtering is applied to the MD data in order to reduce the noise level. Here, the TCFs are multiplied with a Gaussian window

$$G(t) = \exp \left[-\frac{t^2}{2T^2} \right], \quad (3.9)$$

which corresponds to a convolution with a Gaussian function of the width $\Delta\omega = 1/T$ in Fourier domain. This width has to be adjusted to build a compromise between optimal noise reduction and minimal smoothing errors. Further details about Gaussian filtering can be found in appendix E.

As a second approach the TCFs are least-squares fitted to superpositions of

$$f(t) = a_1 \exp[-b_1 t] \cos(c_1 t) + a_2 \exp[-b_2 t] \sin(c_2 t), \quad (3.10)$$

whose transforms into frequency domain are known analytically [49]. The charm of this procedure is that one introduces no additional errors on the way to the spectral density once a successful fit has been established. A comparison of these two approaches is made in the next section based on a detailed error analysis. Afterwards, the successful method is tested for self-consistency on a set of various test systems.

3.1.2. Error Analysis

The discussion of the error analysis starts with the first approach based on smoothing the TCFs via Gaussian filtering. Afterwards, the comparison with the second approach, the fit procedure of TCFs in time domain, is made. The error analysis is done with the same test system as already has been used in Sec. 2.1.2: a harmonic oscillator of frequency $\omega_0 = 0.4$ in a Lorentzian bath described by the memory kernel $\xi(t) = 2a^2 \exp[-bt] \cos(ct)$ with parameters $a = 0.03$, $b = 0.03$ and $c = 0.4$. The necessary numerical data are computed from the Colored Noise scheme as presented in Sec. 1.1.4. The successful procedure should exactly reproduce the kernel $\xi(t)$ explicitly set up above. This is the main criterion for the accuracy of the method.

Smoothing Procedure

In the smoothing procedure one encounters two types of errors. The first error originates from the numerical integration in calculating the half-sided Fourier transforms. These are carried out both with the help of the C-library FFTW3 [50] providing a fast $\mathcal{O}(N \log N)$ algorithm and a self-implemented $\mathcal{O}(N^2)$ scheme based on the Simpson integration rule [49]. The second error is the smoothing error stemming from the use of a Gaussian window. Both errors are illustrated in Fig. 3.2.

The left panel of Fig. 3.2 shows that the FFTW3 algorithm yields surprisingly large errors for the timestep of 0.3, which is a reasonable choice for a molecular system with hydrogen atoms in, if the typical units (fs) are employed. Decreasing the timesteps to 0.2 and 0.1 still does not lead to a reasonable result. Contrary to this, the Simpson-based scheme yields a spectral density, that perfectly matches the exact one already for the largest step size. The reason is that the FFTW3 algorithm is based on the simplest single-sided integration scheme, whose accuracy is only of first order in the integration step [49, 50]. In contrast, the Simpson's rule is of fourth order and is, hence, much more accurate. The error behavior of the FFTW3 scheme shows that this algorithm is not applicable for the present purpose and the Simpson-based scheme should be preferred. Since the Fourier transforms have to be performed only once for each TCF the additional computational cost of this slower $\mathcal{O}(N^2)$ algorithm hardly makes a difference compared to the cost of performing explicit MD simulations.

The right panel in Fig. 3.2 contains a comparison of the computed spectral densities for different time widths T of the Gaussian window, see Eq. (3.9). It can be observed

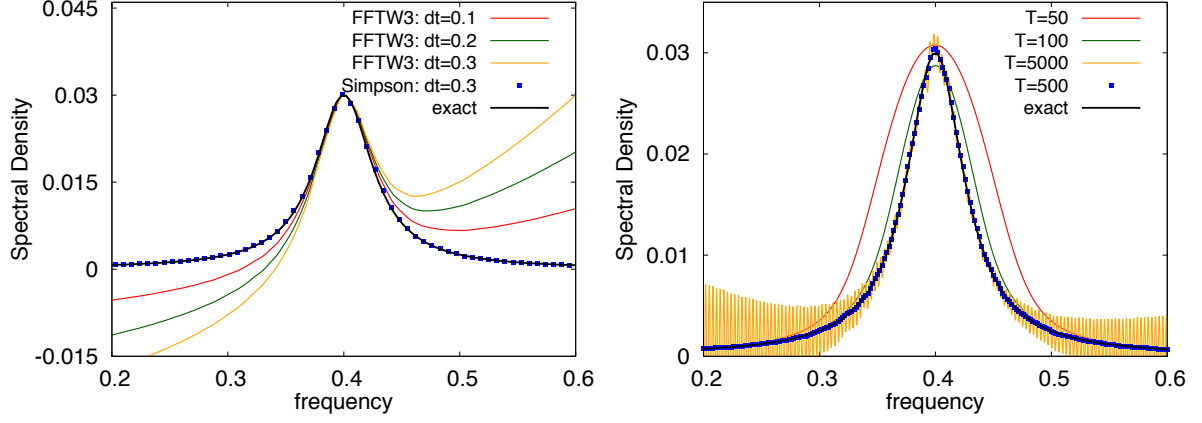


Figure 3.2.: The spectral densities illustrating the two sources of errors emerging in the smoothing approach are shown. In the left panel, FFTW3-based spectral densities are plotted for various time steps compared to the Simpson-based and the exact one. The width of the Gaussian window is $T = 500$, which corresponds to a frequency width of $\Delta\omega = 0.002$. In the right panel, the spectral densities are shown for various widths T of the Gaussian window compared to the exact one. The Fourier transform has been carried out using the Simpson's rule. The numerical parameters of corresponding Colored Noise simulations are the same as in Sec. 2.1.2.

that the curves become smoother the smaller T becomes and more noisy for large T . Too small choices of T yield too broad spectral densities. These findings can be easily understood, since a multiplication with a Gaussian function having the width T in time domain is equivalent to convoluting with a Gaussian function with the width $\Delta\omega = 1/T$ in frequency domain. There exists an optimal value of $T = 500$ (blue dots in right panel) which establishes a compromise between the noise level and the smoothing error. This value is comparable to the characteristic correlation time of the system, see Fig. 2.3 in Sec. 2.1.2, which can be used as a criterion for setting T .

Fit Procedure

Since the fit is performed to functions whose analytical Fourier transform is known, the only errors induced by the entire procedure are the fit errors themselves. In Fig. 3.3 (panel a) the resulting spectral density is compared against the analytic one. The result of the successful smoothing procedure from the previous section is displayed in panel b). Surprisingly, the fit procedure causes errors (panel d), which are unacceptably large. Fig. 3.4 shows the Fourier transforms of the TCF fits (panel a), of the smoothed

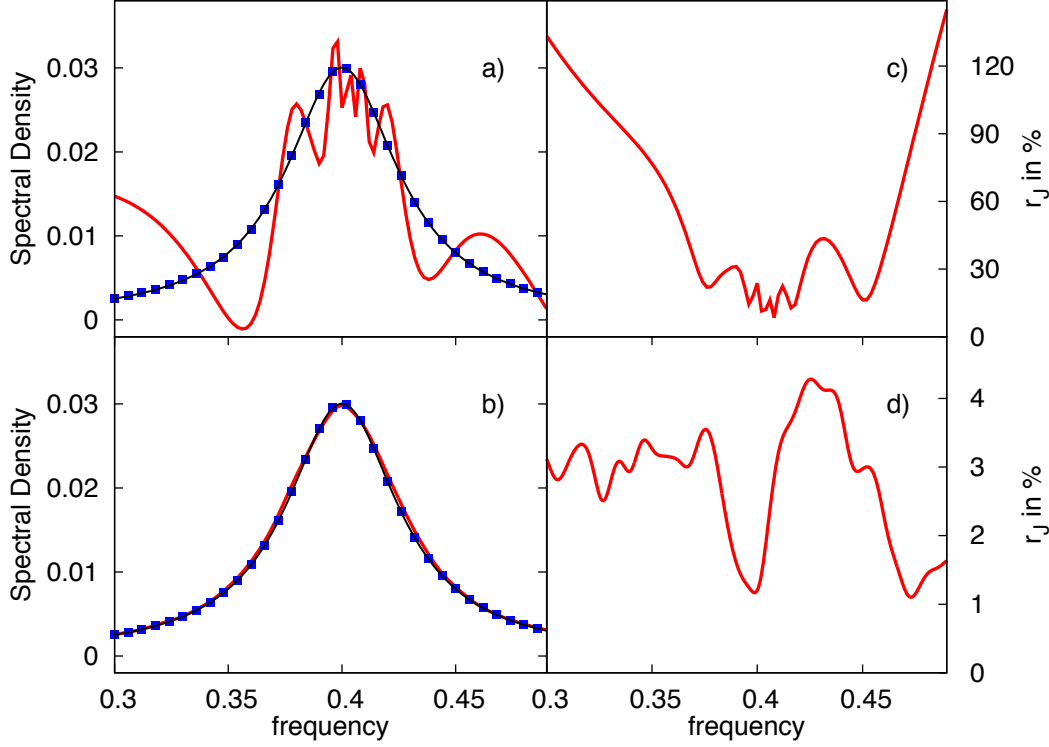


Figure 3.3.: The numerical (red lines), first-order corrected (blue dots) and exact (black lines) spectral densities are shown for a) the fit procedure and b) the smoothing procedure. For the fit procedure 3 fit functions of the form given in Eq. (3.10) have been used. In the smoothing procedure the width of the Gaussian window has been set to $T = 500$. The relative errors r_J of the absolute values are given in panels c) and d). The numerical parameters for corresponding colored noise simulations are the same as in Sec. 2.1.2.

TCFs (panel d) and the particular relative error of the absolute values (panels b and e). Here, fit and smoothing errors have the same order of magnitude although the errors in the spectral densities are dramatically different. This apparent paradox suggests to inspect the error accumulation in more detail.

In the following the numerical TCF-transforms $S^{(\text{num})}(\omega)$ are written in the form

$$S^{(\text{num})}(\omega) = S(\omega) + \epsilon(\omega), \quad (3.11)$$

where $S(\omega)$ denotes the exact function and $\epsilon(\omega)$ the (complex-valued) error. Substituting this into Eq. (3.4) and performing a first order Taylor expansion around $\epsilon(\omega) = 0$

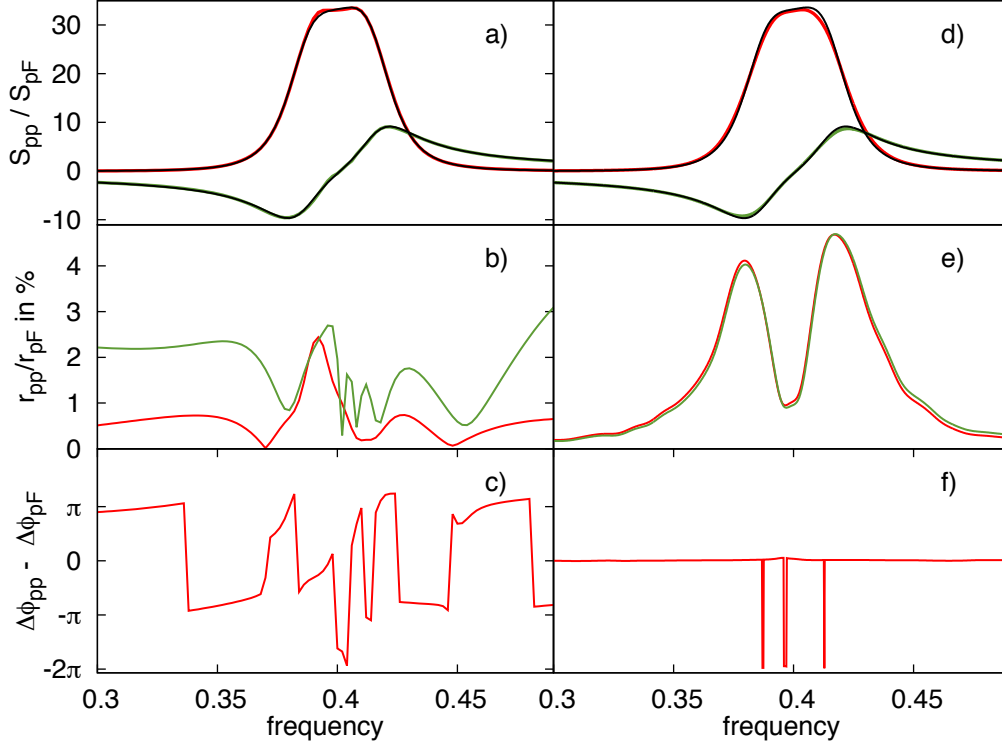


Figure 3.4.: The real parts of S_{pp} and S_{pF} (panels a and d), the relative errors of their absolute values (panels b and e), as well as the phase differences (panels c and f) are shown for the fit procedure (panels a – c) and the smoothing procedure (panels d – f). Red and green curves indicate a correspondence to S_{pp} and S_{pF} , respectively, and the black curve denotes exact results. The numerical parameters are the same as in Fig. 3.3.

one finds for the numerical spectral density $J^{(\text{num})}(\omega)$

$$J^{(\text{num})}(\omega) = J(\omega) - [J(\omega) + i\omega] \frac{\epsilon_{pp}(\omega)}{S_{pp}(\omega)} + \frac{\epsilon_{pF}(\omega)}{S_{pp}(\omega)} + \mathcal{O}(\epsilon^2), \quad (3.12)$$

from which the first order expression for the error $\epsilon_J(\omega) = J^{(\text{num})} - J(\omega)$ can be read off

$$\epsilon_J^{(1)}(\omega) = -[J(\omega) + i\omega] \frac{\epsilon_{pp}(\omega)}{S_{pp}(\omega)} + \frac{\epsilon_{pF}(\omega)}{S_{pp}(\omega)}. \quad (3.13)$$

Upon utilizing the symbol $r(\omega) = \epsilon(\omega)/S(\omega)$ for relative errors and employing the relation $S_{pF}(\omega) = i\omega_0^2/\omega \cdot S_{pp}(\omega)$ valid in the harmonic regime, see Eq. (D.4) in appendix D,

one obtains

$$\epsilon_J^{(1)}(\omega) = -[J(\omega) + i\omega] r_{pp}(\omega) + i \frac{\omega_0^2}{\omega} r_{pF}(\omega), \quad (3.14)$$

where ω_0 represents the oscillator frequency. In Fig. 3.3 the first order-corrected spectral densities are plotted (blue dots) giving very good agreement with the exact functions in both cases. This allows one to estimate the error up to the first order and Eq. (3.14) can be used to draw further conclusions. Note that such a correction can only be made because the exact spectral density $J(\omega)$ in Eq. (3.14) is known in the error analysis performed here. It does not provide a useful correction for practical applications where the exact spectral densities are unknown.

In order to gain more insight into the mechanisms of the error accumulation implied by Eq. (3.14) one can write the relative errors in Euler form $r = |r| \cdot e^{i\Delta\phi}$, where the explicit frequency argument shall be dropped from now on. This yields for the first order correction

$$\epsilon_J^{(1)} = -J|r_{pp}| \cdot e^{i\Delta\phi_{pp}} + i \cdot e^{i\Delta\phi_{pF}} \left[|r_{pF}| \frac{\omega_0^2}{\omega} - |r_{pp}| \omega \cdot e^{i(\Delta\phi_{pp} - \Delta\phi_{pF})} \right], \quad (3.15)$$

where it becomes apparent that the error accumulation is sensitive to the phase difference $\Delta\phi_{pp} - \Delta\phi_{pF}$. Error cancellation is supported in the vicinity of ω_0 , for instance, if the magnitudes of the relative errors are comparable, $|r_{pF}| \approx |r_{pp}| = |r|$ and the phase difference is a multiple of 2π . In such a case the frequency-dependent expression in Eq. (3.15) vanishes and the error reduces to

$$|\epsilon_J^{(1)}| = |J| \cdot |r| \quad (3.16)$$

meaning that the relative error for the absolute value of $J(\omega)$ is the same as for the TCFs in Fourier space. If, in the opposite, the phase differences are close to odd multiples of π , the error strongly accumulates since the term in braces introduces a frequency-dependent error

$$|\epsilon_J^{(1)}| = |r| \cdot \sqrt{|J|^2 + 4\omega^2}. \quad (3.17)$$

In Fig. 3.4 the errors $|r_{pp}|$ and $|r_{pF}|$ (panels b and e) as well as the phase differences $\Delta\phi_{pp} - \Delta\phi_{pF}$ (panels c and f) are shown for the fitted and smoothed TCFs correspondingly. It is evident that the errors and phase differences for the smoothed curves obey

quite strictly the aforementioned conditions, which support error cancellation. Contrary to this, the errors and phase differences for the fit procedure are less regular and hardly predictable. In the worst case the error accumulates strongly as can be seen in Fig. 3.3.

The detailed error analysis presented above suggests that smoothing the raw data via Gaussian filtering is the method of choice since the phases of the errors as well as their magnitudes are changed in a way which is convenient for error cancellation. The fit procedure suffers from large error accumulation and yields useless results. In the next section the smoothing procedure is verified further via self-consistency tests, showing the success of the method for other test systems in addition to the one used throughout this section.

3.1.3. Self-consistency Tests

Since it has been shown in the previous section, that the proposed smoothing procedure favors strong error cancellation, the question may arise, whether this happens in the general case as well. In the remainder of this section the success of the method will be demonstrated for three additional, qualitatively different systems. These are a free particle in a bath described by an exponentially damped kernel $\xi(t) = a^2 e^{-bt}$ with the parameters $a = 1.0$ and $b = 1.0$, an anharmonic oscillator in a bath corresponding to the memory kernel $\xi(t) = 2a^2 e^{-bt} \cos(ct)$ with $a = 0.015$, $b = 0.005$ and $c = 0.37$ as well as an anharmonic oscillator in a bath described by $\xi(t) = 2a_1^2 e^{-b_1 t} \cos(c_1 t) + 2a_2^2 e^{-b_2 t} \cos(c_2 t)$ with $a_1 = 0.02$, $b_1 = 0.03$, $c_1 = 0.42$, $a_2 = 0.015$, $b_2 = 0.007$ and $c_2 = 0.38$. The anharmonic oscillators are modelled by a Morse potential, Eq. (2.9), with $D = 0.05$ and $\alpha = 1.265$ resulting in the harmonic frequency $\omega_0 = 0.4$. The memory kernel for the anharmonic oscillator in the first example is constructed such that the corresponding spectral density is narrow and slightly off-resonant from the harmonic frequency. As has been discussed in Chap. 2 this causes long dephasing times and hence very narrow peaks in the Fourier transforms of the TCF. The reason for the choice of this regime is that the TCFs are expected to be less sensitive to the memory kernels. In the second anharmonic oscillator example the spectral density possesses a sharp peak below the oscillator frequency and a broad peak above.

In Fig. 3.5 the results for the three examples together with the one for the harmonic system studied in the previous section are shown. The coincidence of the exact spectral densities and the numerical ones is reasonable for all systems considered. The largest deviations are obtained at the maxima of sharply peaked contributions since there the curvature is very large, see panels c) and d). Negligible deviations are obtained for broader curves of low curvature, see panels a), b) and d). This illustrates where the principal limitation of the method lies, although it should be stressed that sharply peaked spectral densities are rather untypical for liquid systems as can be seen in the next sections. Thus, the developed method for extracting the spectral densities from explicit simulations, based on smoothing the raw data for noise reduction, can be graded as successful and is ready for application to real solute dynamics.

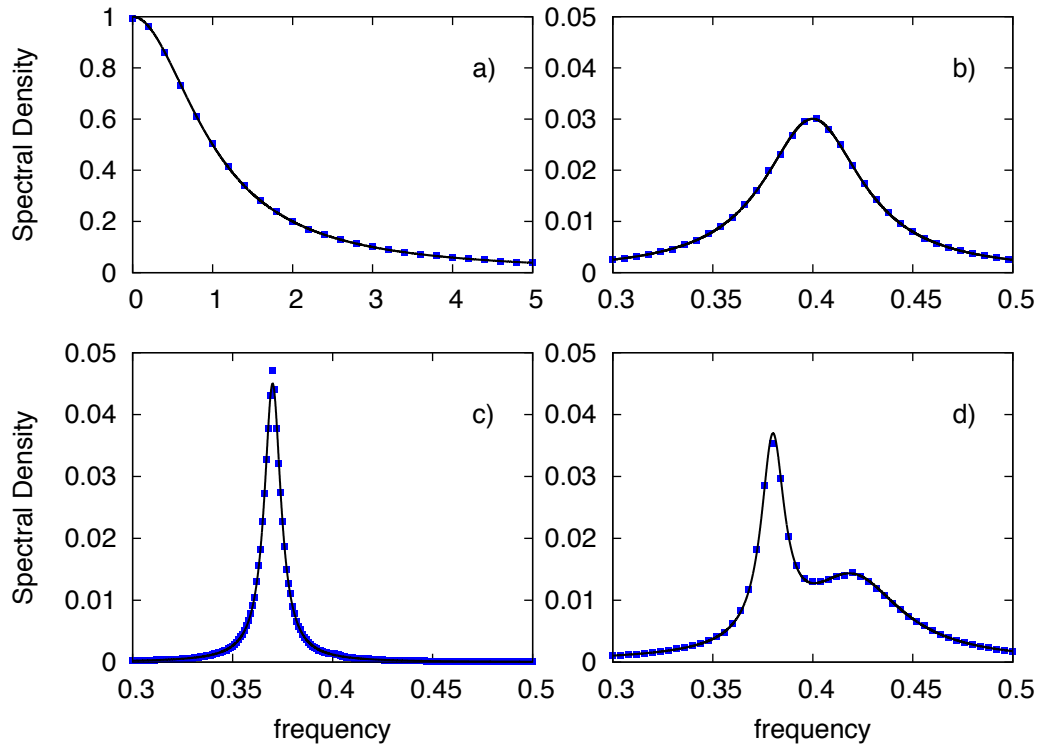


Figure 3.5.: Four examples of qualitatively different spectral densities are shown. Blue dots indicate the numerical results from the smoothing procedure and black lines stand for the exact spectral density. The systems are a free particle (panel a) with $\xi(t) = a^2 e^{-bt}$, the harmonic oscillator from the previous section (panel b), anharmonic oscillators with $\xi(t) = 2a^2 e^{-bt} \cos(ct)$ (panel c) and with $\xi(t) = 2a_1^2 e^{-b_1 t} \cos(c_1 t) + 2a_2^2 e^{-b_2 t} \cos(c_2 t)$ (panel d), see text for the values of the constants. Numerical parameters of Colored Noise simulations and the width T of the Gaussian window are chosen such that numerical convergence is obtained.

3.2. Vibrational Spectra of Solutes in liquid Solvents

3.2.1. Systems

In order to probe the applicability of the non-linear GLEs, representative test-cases have been chosen according to the following guidelines. First, it is aimed at describing realistic solvent environments which, additionally, should have a significant influence on the solute's vibrational mode under investigation. Second, the test-cases should be chosen to represent qualitatively different physical situations in order to draw more or less general conclusions.

In order to meet these requirements one of the most important solvents in the everyday life has been chosen as a first guinea pig: water. As a solute mode, the OH-stretch of an HOD molecule is taken, whose gas phase spectra have already been investigated in a previous work [39]. In bulk water, this molecule participates in a network of hydrogen bridges, which strongly influence the solute stretching motion. Especially, one expects a strong resonant energy flow between the OH-stretches in HOD and water. Additionally to the OH-stretch, HOD possesses internal vibrational modes formed by the OD-stretch and HOD bends. Due to the system-bath partitioning, chosen in this thesis, these motions are considered as a part of the bath resulting in a rather counter-intuitive picture. Hence, it is desirable to consider a solute with no additional internal motion, i.e. whose state is entirely described by one single coordinate. Therefore, as a second system, a diatomic OH molecule embedded in the same bulk water surrounding is studied. The third system has been chosen to represent a qualitatively different situation: the ionic liquid $[C_2mim][NTf_2]$. Here, solute and solvent have a much more complex mode structure than that in the water examples and their Coulomb interaction is remarkably strong. The vibrational mode of interest is the C(2) – H stretch on the imidazolium ring (see Scheme 1 of [51] for a sketch). The three systems are all simulated at an ambient temperature of 300 K.

For describing the interactions in the two aqueous systems the q-SPC/Fw water model has been employed [41], where the harmonic stretching potential has been substituted with an anharmonic one according to the q-TIP4P/F force field [52]. This combination of the two water force fields has been adopted from Paesani and Voth [53]. Particularly for HOD, the OH- and OD interactions are formed by Morse potentials, Eq. (2.9), whereas the potential for HOD bending is harmonic with respect to the

HOD-angle θ

$$V(\theta) = \frac{k_\theta}{2} (\theta - \theta_{\text{eq}})^2 . \quad (3.18)$$

All intermolecular interactions are of Lennard-Jones and Coulomb type

$$V^{(\text{inter})}(r) = \sum_{i,j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^6 \right] + \frac{q_i q_j}{R_{ij}} , \quad (3.19)$$

with a sum that runs over all particle pairs being not within the same molecule. The diatomic OH molecule is simulated with exactly the same setup as for HOD with the only difference that the deuteron has been removed. The ionic liquid [C₂mim] [NTf₂] is parametrized according to the force field given by Ludwig *et al.* [51]. Again, the original harmonic stretching potentials are replaced by Morse potentials whose parameters are chosen according to an *ab initio* (DFT-B3LYP) calculation performed by T. Zentel in [54].

In all cases the solute systems are considered as one-dimensional with the OH/CH-bondlength being the coordinate associated with x in the GLE. According to Sec. 3.1.1 one needs to calculate the MAF and MFC corresponding to this coordinate in order to extract the spectral density from explicit MD simulations. The conjugate momentum p , needed for the MAF, is calculated as

$$p(t) = \mu \vec{n}(t) \cdot (\vec{v}_1(t) - \vec{v}_2(t)) , \quad (3.20)$$

with the normalized bondvector \vec{n} , the individual velocities $\vec{v}_{1,2}$ of the two particles participating in the bond and their reduced mass μ . The forces, needed for the MFC, are calculated depending on the kind of GLE involved. In general, it has the form

$$F(t) = - \left. \frac{\partial V}{\partial x} \right|_{x=x(t)} \quad (3.21)$$

with the potential V being the full Morse potential, Eq. (2.9), in the case of the MBO-GLE and the mean-field potential for the NLP-GLE. For parametrizing the LP-GLE the MFC calculation can be omitted since the spectral density is exclusively determined by the MAF, see Sec. 3.1.1.

3.2.2. Simulation Details

First, explicit MD simulations of the three systems have been carried out in order to obtain the MAF and MFC needed for both calculating explicit vibrational spectra through Eq. (1.74) and extracting the memory kernels according to the procedure developed in the previous section. For the explicit MD simulations the GROMACS simulation package v. 4.6.5 has been used [55]. For HOD and OH in bulk water the solute molecule has been included in a cubic and periodic simulation box with the edge of 2.4 nm together with 465 water molecules forming the bulk surrounding. For the ionic liquid a periodic box of 4.5 nm comprising 216 ion pairs has been used. In all systems the calculation of the Lennard-Jones forces has involved a cut-off radius of 0.9 nm and a switch radius of 0.8 nm. For Coulombic forces the particle-mesh Ewald method with a cut-off radius of 0.9 nm and a switch radius of 0.89 nm has been exploited. The cut-off radius for nearest-neighbor list generation has been set to 1.1 nm, which allows one to update the pair lists every 1000 timesteps, preserving sufficiently good energy conservation. TCFs have been calculated both as time-averages along a single microcanonical (NVE) trajectory and a subsequent average over a swarm of NVE-trajectories whose initial conditions were sampled from the canonical (NVT) ensemble, see Sec. 2.1.1. Specifically, a swarm of 1000 independent trajectories per stretch each of 6 ps length and a timestep of 0.1 fs has been involved. The initial conditions have been sampled in equidistant steps of 0.2 ps from a NVT-trajectory thermostatted via a Langevin thermostat with a coupling strength of $\tau = 0.04$ ps. Data production has started after an equilibration time of 20 ps.

The spectral densities have been extracted from the explicit MD data according to the procedure discussed in the previous section. The Gaussian width T has been chosen to minimize the smoothing errors in the resulting spectral densities (see Sec. 3.1.2). A good choice for T has been the correlation time within the system under study. After extracting the spectral density its real parts have been least-squares fitted to superpositions of

$$\text{Re}\{J(\omega)\} = a^2 b \cdot \left[\frac{1}{b^2 + (c - \omega)^2} + \frac{1}{b^2 + (c + \omega)^2} \right], \quad (3.22)$$

which correspond to superpositions of the memory kernels $\xi(t) = 2a^2 e^{-bt} \cos(ct)$ [49]. The fit coefficients determine the elements of the drift matrices for the Colored Noise simulations according to Sec. 1.1.4. The explicit forces used in the GLE have been set to the explicit Morse force for the MBO-GLE and the mean-force for the linearized

NLP-GLE. A bondlength distribution function $f(x)$ has been calculated from the explicit MD trajectories in order to give access to the mean-field potential via Eq. (1.48). Employing the LP-GLE requires to compute the effective harmonic frequencies. These have been fitted from the hyperbola in the imaginary part of the spectral density as explained in Sec. 3.1.1. Finally, the *ad hoc* anharmonic GLE has been utilized. Therefore, an anharmonic term proportional to x^2 calculated from a third order Taylor expansion of the Morse potential employed in MD simulations has been added to the LP-GLE. For data production Colored Noise simulations of 1000 independent trajectories with 6 ps length and a timestep of 0.1 fs have been employed. Vibrational spectra have been calculated from the MAF in the same manner as in explicit simulations described above.

3.2.3. Results

The presentation of the results starts with showing the spectral densities of the three systems under investigation. Afterwards, the spectra obtained from the different GLE simulations are compared against the explicit MD result. From this comparison, conclusions about the GLE's applicability are drawn.

Spectral Densities

Figure 3.6 contains the spectral densities for the three investigated systems corresponding to the LP-GLE (red), the MBO-GLE (green) and the linearized NLP-GLE (orange). Note that for the *ad hoc* anharmonic GLE the spectral density coincides with that of the LP-GLE by construction. In all cases the spectral densities have a peaked structure instead of being broad and continuously distributed along the frequency axis. The particular contributions can be assigned to vibrational modes of the bath. For instance, the peak at 2600 cm^{-1} in the HOD spectral density (panel a) is connected to the OD vibration and is missing in the OH spectral density (panel b), where this mode has been removed by construction. Further contributions at 1500 cm^{-1} can be assigned to the bending modes and the continuum at lower frequencies corresponds to libration bands of the molecules within the hydrogen bridges network. In the case of the ionic liquid $[\text{C}_2\text{mim}][\text{NTf}_2]$ the low frequency part of the spectral density is much more complex since noticeably more vibrational modes are available.

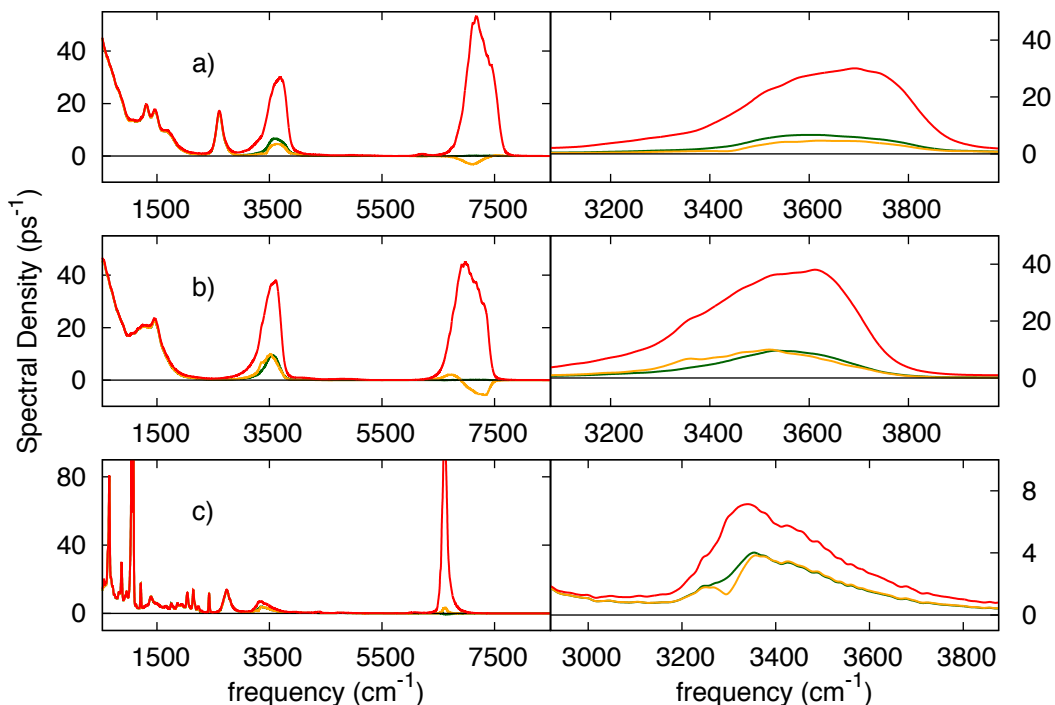


Figure 3.6.: The real parts of the spectral densities are shown for HOD in bulk water, OH in bulk water and the ionic liquid $[\text{C}_2\text{mim}][\text{NTf}_2]$ in the panels a), b) and c), respectively. The right plots within one panel show the spectral density zoomed into the resonant region. The colors correspond to the LP-GLE (red), MBO-GLE (green) and the linearized NLP-GLE (orange).

Comparing the spectral densities derived from the LP-GLE, the MBO-GLE and linearized NLP-GLE for each system, reveals the significant differences only in the region of the spectral density where the system-bath coupling is resonant with the system mode (about 3600 cm^{-1} for HOD and OH and about 3300 cm^{-1} for the ionic liquid, see Fig. 3.7). These regions are shown in the right column of Fig. 3.6. Here, the spectral density turns out to be very sensitive to the presence of anharmonicity in the system. In particular, the spectral densities obtained from LP-GLE and MBO-GLE parametrizations are similar in shape but different in intensity, the former being four and two times larger than the latter for the water systems and the ionic liquid, respectively. As it was discussed in Sec. 2.2, resonant coupling is most important since off-resonant coupling causes spectral amplitudes which are about two orders of magnitudes smaller than those arising from resonant coupling. Fortunately, this allows one to restrict the

least-squares fit of the spectral densities to the resonant region providing a great simplification especially for the ionic liquid.

An interesting feature which is worth discussing at this point, is the strong peak lying at the overtone frequency of the system, being only visible in the LP-GLE spectral densities. The only difference between the LP-GLE and non-linear GLE spectral densities must be in the system anharmonicity, which, has been put entirely into the bath in the former case (see discussions in Sec. 1.1.2). Therefore, the additional overtone in the LP-GLE spectral density must be a consequence of the system anharmonicity. On the contrary, in the MBO spectral density, where the full anharmonic Morse potential is kept explicitly, this contribution is absent. This nicely underlines the fundamental differences of the system-bath partitioning performed in the different kinds of GLEs. The spectral density from the linearized NLP-GLE shows an unphysical behaviour in this overtone region since it becomes negative in the case of HOD and OH. However, this overtone is off-resonant with the system frequency and hence anyway neglected in the GLE simulations as was pointed out before.

According to [17–19] the LP-GLE spectral densities contain further dynamical information: the vibrational relaxation times T_1 . If the solute mode is excited to a non-equilibrium state, this time determines the characteristic energy decay back to thermal equilibrium. The so-called Landau-Teller formula [17–19]

$$T_1^{-1} = \text{Re}\{J(\tilde{\omega})\} \quad (3.23)$$

provides a connection between this relaxation time and the LP-GLE spectral density. The effective harmonic frequencies $\tilde{\omega}$ for the three systems are 3628.79, 3524.32 and 3316.48 cm^{-1} leading vibrational relaxation times of 35 fs, 28 fs and 143 fs for HOD, OH and the ionic liquid, respectively. Note, that the effective frequencies $\tilde{\omega}$ occurring in the LP-GLE must be used instead of the harmonic frequencies ω_0 stemming from a second order Taylor expansion of the employed Morse potentials. These harmonic frequencies are 3886.39 cm^{-1} for HOD/OH and 3326.02 cm^{-1} for the ionic liquid, which, for the aqueous systems, are significantly blue-shifted from the effective ones. Hence, the vibrational relaxation times would be dramatically overestimated if these frequencies were used.

The Applicability of the GLEs to vibrational Spectra

The vibrational spectra of the three systems obtained from explicit MD simulations as well as from the different types of GLEs are displayed in Fig. 3.7. In all systems the spectra obtained from the LP-GLE (red stars) perfectly coincide with the explicit MD results (black curves). This is expected since the LP-GLE is the only one having a mathematically rigorous foundation. This also illustrates that in terms of linear spectroscopy any anharmonic system can be entirely mapped onto an effective harmonic one. However, the effective frequencies, see previous subsection, cannot be connected to the real Morse potentials employed in the explicit simulations. In contrast, they artificially include the redshifts usually being a consequence of the system-bath coupling. This again underlines that a rather counter-intuitive system-bath partitioning is performed when the linear projection operator formalism is used.

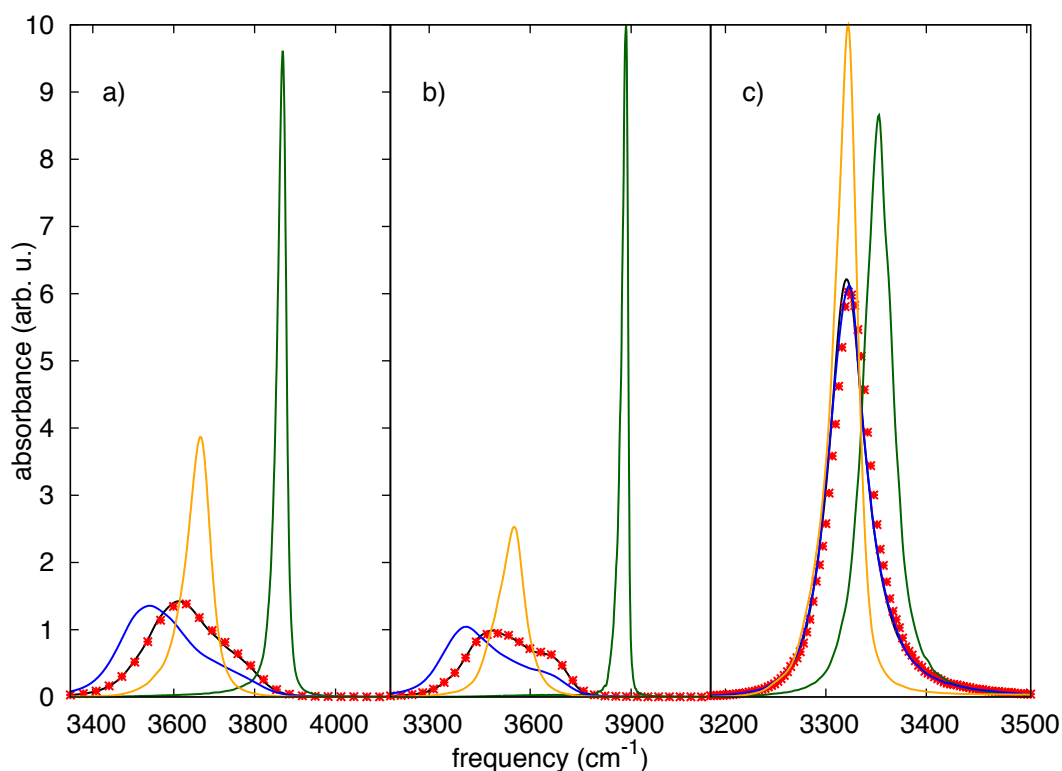


Figure 3.7.: Vibrational spectra shown for a) HOD in bulk water, b) OH in bulk water and c) the ionic liquid $[C_2mim][NTf_2]$. The explicit MD results (black curve) are compared against the results of the LP-GLE (red stars), MBO-GLE (green curve), the linearized NLP-GLE (orange curve) and the the GLE with *ad hoc* anharmonicity (blue curve).

The spectra provided by the MBO-GLE (green curves) dramatically deviate from the MD ones as they are blue-shifted by 250, 330, 30 cm^{-1} in panels a), b), c), respectively, are too narrow and of wrong shape in all cases. This is because the harmonic frequencies of the employed Morse potentials are off-resonant to the spectral density and hence the spectral peak hardly differs from that in gas phase, see discussion in Chap. 2. The results of the linearized NLP-GLE (orange curves) improve in terms of peak positions but still yield incorrect shapes. The spectra from the *ad hoc* anharmonic GLE are very close to the explicit ones although deviations are still visible for HOD and OH. However, given the excellent quality of the LP-GLE results none of the three non-linear GLEs reveals itself as successful. This shows that the attempts to include anharmonicity explicitly into the GLE are not applicable for purposes of linear vibrational spectroscopy. In contrast, there seems to be no room for anharmonicity as the LP-GLE already yields exact spectra on the basis of an effective harmonic description.

The Invertibility Problem

Unfortunately, the only successful GLE, the LP-GLE, is of limited use for the purposes of a classical vs. quantum comparison and non-linear spectroscopy. Therefore, it is important to discover the exact reasons for the breakdown of the employed non-linear GLEs and to find possible improvements. In the following the problem shall be formulated in a formal mathematical way and the deep reason for the breakdown of any of the employed non-linear GLEs will be unraveled from this perspective.

According to the procedure developed in Sec. 3.1.1 one extracts the memory kernel $\xi(t)$ from the MAF and MFC calculated via explicit MD simulations. Mathematically, this defines a mapping

$$A : \{C_{pp}(t); C_{pF}(t)\} \mapsto \xi [C_{pp}(t); C_{pF}(t)] , \quad (3.24)$$

where two functions, $C_{pp}(t)$ and $C_{pF}(t)$ are mapped onto one memory kernel $\xi(t)$. Specifically, the spectral density is calculated from the TCFs Fourier transforms according to Eq. (3.4)

$$J(\omega) = \frac{1 + S_{pF}(\omega)}{S_{pp}(\omega)} - i\omega , \quad (3.25)$$

which uniquely defines the memory kernel in time domain via a back Fourier transform. Conversely, a mapping B can be defined which maps one kernel $\xi(t)$ onto two functions $C'_{pp}(t)$ and $C'_{pF}(t)$

$$B : \xi(t) \mapsto \{C'_{pp}[\xi(t)]; C'_{pF}[\xi(t)]\} . \quad (3.26)$$

This mapping is defined via the (non-linear) GLE itself, since once the memory kernel has been fixed for a given temperature the corresponding TCFs are determined uniquely through the corresponding LD. The requirement for the GLE to be successful is that it should reproduce exactly that pair of $C_{pp}(t)$ and $C_{pF}(t)$ from which the memory kernel has once been extracted. Mathematically speaking, the mapping B should be exactly the inverse mapping of A

$$B \stackrel{!}{=} A^{-1} . \quad (3.27)$$

Having established a mathematical formulation of the problem one easily recognizes from Eq. (3.25) that the mapping A is not injective, since one can think of infinitely many pairs of functions $C_{pp}(t)$ and $C_{pF}(t)$ which yield the same memory kernel $\xi(t)$. In principle, this set of infinite pairs can be artificially constructed by setting one function, say $C_{pF}(t)$, to a specific form and calculating the other, $C_{pp}(t)$, from Eq. (3.25) keeping the spectral density fixed. The missing injectivity implies that the mapping A is not invertible and, hence, the requirement for B to be the inverse mapping of A cannot be fulfilled. This serious problem, referred to as the *invertibility problem* in the following, clearly puts forward the reason why any of the proposed non-linear GLEs cannot meet the requirement for their general applicability as formulated in Eq. (3.27).

In order to exclude the possibility that some yet unknown dependence between $C_{pp}(t)$ and $C_{pF}(t)$ restricts the choice to a single pair, the numerical evidence for the invertibility problem has been obtained. This can be impressively seen with the help of Fig. 3.8 that contains two pairs of $C_{pp}(t)$ and $C_{pF}(t)$, one from explicit MD (plugged into mapping A) and the other coming from the corresponding non-linear GLEs (results of mapping B with the same kernel). Although the two pairs of TCFs are dramatically different they yield exactly the same memory kernels when put into mapping A . One might wonder why the invertibility problem does not occur in the LP-GLE formalism. The reason for this is, that since the effective force is purely harmonic the corresponding MFC can be expressed entirely in terms of the MAF as discussed in Sec. 3.1.1. In mapping A only one function is then needed for calculating $\xi(t)$

according to Eq. (3.7). This fact makes A injective and hence the desired condition formulated in Eq. (3.27) realizable.

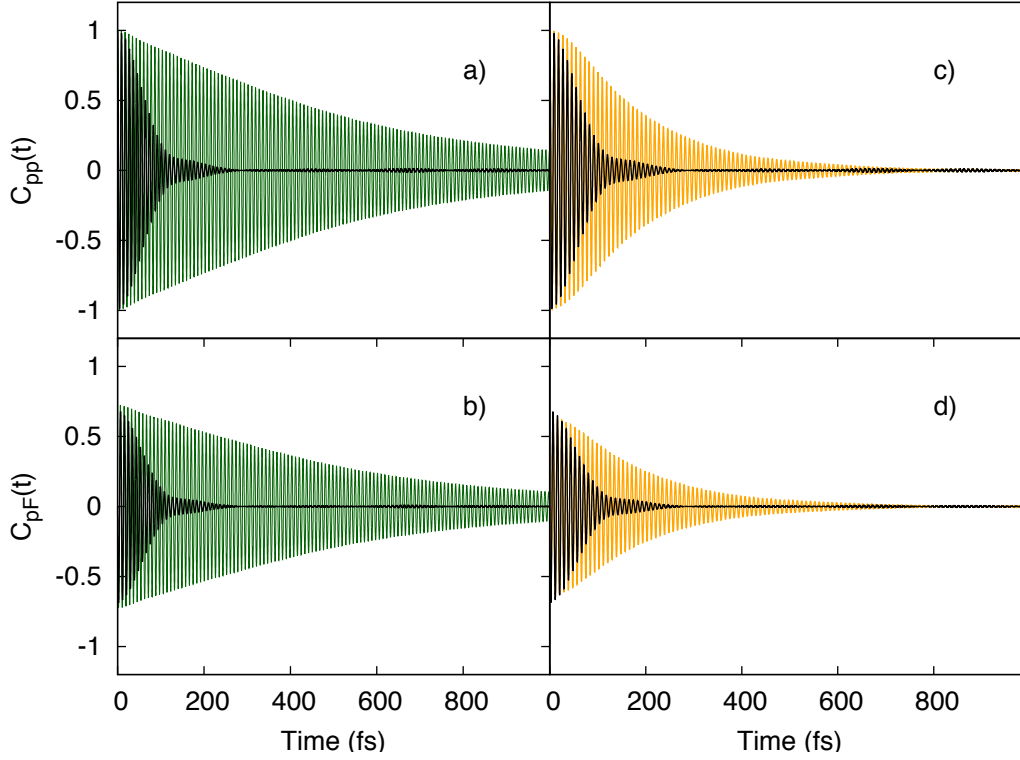


Figure 3.8.: MAF (panels a and c) and MFC (panels b and d) shown for HOD in bulk water. The panels a) and b) correspond to explicit MD (black) vs. MBO-GLE (green). In panels c) and d) explicit MD results (black) are compared against the linearized NLP-GLE results (orange).

Ansatz to solve the Invertibility Problem

The aforementioned invertibility problem prohibits the use of the MBO-GLE in the general case. It can still be successful by chance if for a particular system mapping A happens to be invertible on a special subset of functions $C_{pF}(t)$ and $C_{pp}(t)$. However, for the NLP-GLE there might be a systematic way to avoid the invertibility problem by taking into account higher orders in the expansion of the memory functional in Eq. (1.45). The basic idea is to include so many terms in the expansion that, similar to the LP-GLE, the number of memory kernels introduced by each term coincides with the number of TCFs needed for their extraction. To illustrate this idea further a second

order expansion of the NLP-GLE is performed in appendix F. The resulting second order NLP-GLE reads

$$\begin{aligned}\dot{x} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial V_m}{\partial x} - \int_0^t \xi_1(t-\tau)p(\tau)d\tau - \int_0^t \xi_2(t-\tau)p(\tau)x(\tau)d\tau + R(t),\end{aligned}\quad (3.28)$$

with two memory kernels $\xi_1(t)$, $\xi_2(t)$ and a noise term obeying the two FDT relations

$$\begin{aligned}\langle R(0)R(t) \rangle &= mkT\xi_1(t) \\ \langle x(0)R(0)R(t) \rangle &= mkT\langle x^2 \rangle \xi_2(t).\end{aligned}\quad (3.29)$$

As before, $V_m(x)$ stands for the mean-field potential. The corresponding integro-differential equation for the MAF reads

$$\dot{C}_{pp}(t) = C_{pF}(t) - \int_0^t \xi_1(t-\tau)C_{pp}(\tau)d\tau - \int_0^t \xi_2(t-\tau)C_{pxp}(\tau)d\tau, \quad (3.30)$$

with the additional correlation function $C_{pxp}(t) = \langle p(0)x(t)p(t) \rangle$. So far, this equation does not seem to be helpful, since there will be always one TCF, the MFC $C_{pF}(t)$, more than memory kernels. However, if one can expand the mean-force up to second order

$$F_m(x) = -m\omega_0^2 x + m\chi x^2 \quad (3.31)$$

the MFC can be expressed in terms of $C_{pp}(t)$ and $C_{pxp}(t)$ in the same spirit as the effective harmonic force in the LP-GLE could be entirely expressed in terms of $C_{pp}(t)$. To see this one can write $C_{pF}(t)$ as

$$C_{pF}(t) = -m\omega_0^2 C_{px}(t) + m\chi C_{px^2}(t). \quad (3.32)$$

Using $C_{px}(t) = 1/m \int_0^t C_{pp}(\tau)d\tau$ as well as $C_{px^2}(t) = 2/m \int_0^t C_{pxp}(\tau)d\tau$ one obtains

$$C_{pF}(t) = -\omega_0^2 \int_0^t C_{pp}(\tau)d\tau + 2\chi \int_0^t C_{pxp}(\tau)d\tau \quad (3.33)$$

and, hence, Eq. (3.30) reduces to

$$\dot{C}_{pp}(t) = - \int_0^t K_1(t-\tau)C_{pp}(\tau)d\tau - \int_0^t K_2(t-\tau)C_{pxp}(\tau)d\tau, \quad (3.34)$$

with the new kernels

$$\begin{aligned} K_1(t) &= \xi_1(t) + \omega_0^2 \\ K_2(t) &= \xi_2(t) - 2\chi. \end{aligned} \quad (3.35)$$

The procedure shown above suggests that a successful inclusion of anharmonicity could become possible on the basis of a second order NLP-GLE if the mean-force can be expanded up to the second order as well. Further it is expected, that if the mean-force truncates at n th order, then an expansion of the memory functional must be likewise performed to n th order. This provides a clear mathematical statement on where to truncate the expansion if the fundamental invertibility problem shall be avoided. However, in order to formulate a practical scheme for extracting the two memory kernels in Eq. (3.35) one needs a second equation additionally to Eq. (3.34). Unfortunately, such an equation has not been found up to the present time. Furthermore, the second order NLP-GLE, Eq. (3.28), cannot be translated into the Colored Noise propagation scheme and thus a new algorithm has to be developed for its practical implementation. Nevertheless, the formalism outlined above seems to be promising and suggests the route to explicitly include anharmonicity into the GLE.

Conclusions and Outlook

The main goal of this thesis has been to investigate the applicability of the generalized Langevin equation (GLE) formalism to vibrational spectroscopy of realistic solute dynamics in liquid solvents. The charm of the GLE approach is that only a few (here one) selected degrees of freedom (DOFs) evolve under the explicit influence of the system forces, whereas the influence of the other DOFs, called bath, is described implicitly via non-Markovian dissipation and stochastic fluctuations. Their properties are fully determined by a single function being the so-called memory kernel in time domain, or alternatively its Fourier transform referred to as the spectral density.

In the first chapter the detailed theoretical background of the GLE has been presented in order to explore the models and approximations employed. It has been shown that the common form of the GLE can be obtained in a direct way by postulating the so-called Multimode Brownian Oscillator (MBO) model, which assumes a harmonic bath and a bilinear system-bath coupling. A truly rigorous approach to derive the (non-)linear prototype forms of the GLE from (non-)linear projection operator techniques has been presented. It has been put forward that the linear version deviates from the common form by the explicit system force, which is mapped onto an effective harmonic force. As a consequence of the linear projections involved, any system anharmonicity is projected into the bath leading to a rather counter-intuitive system-bath partitioning. In turn, in the non-linear GLE, the system force is modified by mean-field corrections given as a conditional average of the system-bath coupling over the environmental DOFs. Further, the dissipative term involves a functional dependence on system coordinates, which can be only brought into the common GLE form by a first order expansion. It has been concluded that a microscopic justification of the common form of the GLE, that is with the system anharmonicity preserved, cannot be obtained in a rigorous way since either an MBO model has to be postulated, the memory kernel in the GLE from non-linear projections has to be approximated up to the first order or anharmonicity has to be added in an *ad hoc* way to the rigorously derived GLE from linear projections.

After theoretical studies, a numerical propagator has been implemented based on the method of Colored Noise thermostats developed by Ceriotti *et al.* [43–46]. In the second Chapter, this propagator has been investigated in order to set up a reliable protocol for calculating time-correlation functions (TCFs) and vibrational spectra. It turns out, that 500 trajectories are needed for sufficient convergence, which can be calculated with small numerical effort for the systems studied. In the remainder of the Chapter, the mechanisms of a non-Markovian system-bath coupling have been investigated and a spectroscopic criterion for non-Markovian dynamics has been proposed for harmonic systems.

In the last chapter of this thesis the applicability of the GLE formalism to solute dynamics in solvents has been discussed. A method for extracting the spectral density from explicit MD data has been proposed. In this technique the time-correlation functions (TCFs) are calculated from explicit MD simulations and the resulting integro-differential equation is solved for the spectral density via transforming it to Fourier domain, where it becomes a simple algebraic equation. A big challenge in the practical application of this method is to obtain an appropriate noise reduction in the Fourier domain. Therefore, two protocols have been suggested for this purpose. The first is based on fitting the MD TCFs to functional forms whose Fourier transforms are known analytically. The second involves Gaussian filtering techniques for smoothing MD data in time domain. It has been shown that the fit procedure suffers from a large error accumulation of the resulting fit errors and, hence, should not be used. In contrast, the Gaussian filtering technique has revealed itself as a successful approach for extracting spectral densities corresponding to different physical scenarios. After extracting the spectral densities from explicit MD data the GLE formalism has been applied to stretching modes of three systems which are the OH-stretch of an HOD molecule in bulk water, a diatomic OH molecule in bulk water and a CH-stretch of the ionic liquid [C₂mim][NTf₂]. The quality of the GLE results has been measured upon the ability to reproduce the linear vibrational spectra from explicit MD simulations. Main emphasis has been put onto the performance of non-linear GLEs, namely those derived from the MBO model, from the non-linear projection with linearized memory kernel and from adding anharmonicity in an *ad hoc* way. The quality of all non-linear GLE spectra were found to be bad in all systems studied, with the GLE from the MBO model yielding the largest deviations. In contrast, the linear GLE excellently agrees with explicit spectra as it is the only GLE being derived in a mathematically rigorous way without any approximation or postulation of a particular model. This underlines that in linear vibrational spectroscopy any system can be mapped onto an effective harmonic

system and, hence, there is no need to introduce anharmonicity into the physical picture. However, using an effective harmonic system is possible at a prize of loosing the correct microscopic picture of the dynamical processes under study, since the effective harmonic frequencies can hardly be connected to the real system potentials employed.

The breakdown of the non-linear GLEs has been discussed on a formal, mathematical basis. It has turned out that they suffer from an invertibility problem, which states that the mappings employed to extract the spectral densities are not invertible. This means that a spectral density does not uniquely define the TCFs needed for its extraction. Indeed, it has become numerically evident that significantly different TCFs can correspond to the very same spectral density and, hence, non-linear GLEs cannot reproduce TCFs calculated from MD simulations. This serious problem has been identified to be the deep reason behind the breakdown of the non-linear formalisms and has been shown to be absent when the linear GLE is invoked giving a purely mathematical reason why a mapping onto an effective harmonic system is successful. It can be therefore concluded, that in future applications the invertibility problem needs to be avoided. For the GLE derived from the MBO model a solution of this problem is not foreseen. For the GLE derived from non-linear projection operator techniques a possible solution has been sketched, which amounts to performing a higher order expansion of the memory functional, which would in turn produce higher order memory kernels. It has been shown that a step towards solving the invertibility problem can be done. However, the set of equations needed for extracting the memory kernels could not be closed yet. Furthermore, the practical applicability of higher order non-linear GLEs needs to be investigated further since a completely new algorithm for their numerical treatment has to be developed.

Finally, it can be concluded that the GLE formalism does not provide a promising method for describing realistic solute dynamics in liquid solvents. In contrast, a more sophisticated theoretical framework needs to be employed. The initial charm of this simple reduced formalism is overshadowed by a principal inapplicability stemming from either too crude models or too rough approximations. Thus, the description of macroscopic condensed phase systems remains a challenging task in modern physical research. The Holy Grail of many-particle physics is still to be found.

Acknowledgements

The time of my Master studies has finished and it remains for me to express my thanks to several people without whose support this thesis would not have become possible. At first, I would like to thank Prof. Dr. Oliver Kühn, who gave me the opportunity to work on this thesis in the Molecular Quantum Dynamics group in Rostock and who constructively contributed in discussions about problems and results. I would like to thank Dr. Sergei Ivanov for being a very committed supervisor and teacher, who did not shrink back from supporting me even at an ungodly hour late in the night. Further, I would like to express my gratitude to Sven Karten and Gilbert Grell for providing helpful comments on my work. Zu guter Letzt danke ich auch meinen anderen Freunden und vor allem meinen Eltern für die gute Unterstützung während der Zeit meines Studiums.

Appendix A.

Memory kernel, Noise and explicit Force in the LP-GLE

In this appendix the explicit calculation of the matrix Ω defined in Eq. (1.19) as well as the form of the memory matrix $K(t)$, Eq. (1.22), and the noise $F(t)$, Eq. (1.20), is presented for the specific pair of variables $\mathbf{A} = (x, p)^T$ being the system's coordinate and conjugate momentum. As a first step the matrix $\langle \mathbf{A} | \mathbf{A} \rangle$ needs to be calculated and inverted in order to carry out the linear projections. For x and p this matrix reads

$$\langle \mathbf{A} | \mathbf{A} \rangle = \begin{pmatrix} \langle x^2 \rangle & \langle xp \rangle \\ \langle px \rangle & \langle p^2 \rangle \end{pmatrix} \quad (\text{A.1})$$

where the special form of the scalar product, Eq. (1.10), results in ensemble averages denoted as $\langle \dots \rangle$. If $f(\Gamma)$ in Eq. (1.10) is chosen to be the canonical distribution function, the off-diagonal matrix elements vanish since $\langle xp \rangle = \langle x \rangle \langle p \rangle$ and $\langle p \rangle = 0$. Further, the second moment of the momentum p becomes $\langle p^2 \rangle = mkT$ which results in the matrix

$$\langle \mathbf{A} | \mathbf{A} \rangle = \begin{pmatrix} \langle x^2 \rangle & 0 \\ 0 & mkT \end{pmatrix} \quad (\text{A.2})$$

that can easily be inverted

$$\langle \mathbf{A} | \mathbf{A} \rangle^{-1} = \begin{pmatrix} 1/\langle x^2 \rangle & 0 \\ 0 & 1/mkT \end{pmatrix}. \quad (\text{A.3})$$

Further one needs the vector

$$\begin{aligned}\mathcal{L}\mathbf{A} &= (\mathcal{L}x, \mathcal{L}p)^T \\ &= (p/m, -\partial_x H)^T\end{aligned}\tag{A.4}$$

where the Hamilton EOMs have been involved and H denotes the total Hamilton function of system and bath. With these expressions one finds for Ω

$$\begin{aligned}\Omega &= \langle \mathbf{A} | \mathcal{L}\mathbf{A} \rangle^T \langle \mathbf{A} | \mathbf{A} \rangle^{-1} \\ &= \begin{pmatrix} \langle xp \rangle / m & -\langle x \cdot \partial_x H \rangle \\ \langle p^2 \rangle / m & -\langle p \cdot \partial_x H \rangle \end{pmatrix}^T \begin{pmatrix} 1/\langle x^2 \rangle & 0 \\ 0 & 1/mkT \end{pmatrix} \\ &= \begin{pmatrix} 0 & kT \\ -kT & 0 \end{pmatrix} \begin{pmatrix} 1/\langle x^2 \rangle & 0 \\ 0 & 1/mkT \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1/m \\ -kT/\langle x^2 \rangle & 0 \end{pmatrix},\end{aligned}\tag{A.5}$$

where the equipartition theorem has been used in order to equate $\langle x \partial_x H \rangle = kT$.

For the noise term $\mathbf{F}(t) = e^{t(1-\hat{P})\mathcal{L}}(1-\hat{P})\mathcal{L}\mathbf{A} = e^{t(1-\hat{P})\mathcal{L}}(1-\hat{P})(p/m, -\partial_x H)^T$ one realizes that the projection $(1-\hat{P})$ applied to p vanishes because $\hat{P}p = p$ by construction, see Eq. (1.14). Hence, the first component of the noise acting on positions x vanishes and it remains

$$\mathbf{F}(t) = \begin{pmatrix} 0 \\ R(t) \end{pmatrix}\tag{A.6}$$

with

$$R(t) = -e^{t(1-\hat{P})\mathcal{L}}(1-\hat{P})\partial_x H.\tag{A.7}$$

For the memory matrix $\mathbf{K}(t)$ one needs the matrix $\langle \mathcal{L}\mathbf{A}|\mathbf{F}(t) \rangle$ which reduces to

$$\begin{aligned} \langle \mathcal{L}\mathbf{A}|\mathbf{F}(t) \rangle &= \begin{pmatrix} 0 & \langle p \cdot R(t) \rangle / m \\ 0 & -\langle \partial_x H \cdot R(t) \rangle \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & -\langle \partial_x H \cdot R(t) \rangle \end{pmatrix} \end{aligned} \quad (\text{A.8})$$

where it has been used that, according to the definition given above, $R(t)$ cannot carry an explicit dependence on p . Further, in the canonical ensemble individual momenta are uncorrelated to the other phase space coordinates and hence the average $\langle p \cdot R(t) \rangle = \langle p \rangle \langle R(t) \rangle = 0$ vanishes. Finally, the memory matrix reads

$$\mathbf{K}(t) = \langle \mathcal{L}\mathbf{A}|\mathbf{F}(t) \rangle^T \langle \mathbf{A}|\mathbf{A} \rangle^{-1} \quad (\text{A.9})$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & -\langle \partial_x H \cdot R(t) \rangle / mkT \end{pmatrix} \quad (\text{A.10})$$

which defines the memory kernel

$$\begin{aligned} \xi(t) &= -\langle \partial_x H \cdot R(t) \rangle / mkT \\ &= \left\langle (1 - \hat{P}) \partial_x H \cdot R(t) \right\rangle / mkT \\ &= \langle R(0)R(t) \rangle / mkT \end{aligned} \quad (\text{A.11})$$

where in the latter equation the FDT can be obtained. The possibility to insert a redundant projector $1 - \hat{P}$ is justified through the hermitian and idempotency property of projectors.

Appendix B.

Memory kernel, Noise and explicit Force in the linearized NLP-GLE

In this appendix the linearization of the NLP-GLE will be performed explicitly. Further, the formula for the potential of the mean-force, Eq. (1.48) is proven. As in the linear case, the formalism is applied to the special pair of variables $\mathbf{A} = (x, p)^T$ being the coordinate x and conjugate momentum p . For this special pair of variables the noise term reads

$$\begin{aligned}\mathbf{F}(t) &= e^{t(1-\hat{P})\mathcal{L}}(1 - \hat{P})\mathcal{L}\mathbf{A} \\ &= e^{t(1-\hat{P})\mathcal{L}}(1 - \hat{P})(p/m, -\partial_x H)^T \\ &= e^{t(1-\hat{P})\mathcal{L}}(1 - \hat{P})(0, -\partial_x H)^T,\end{aligned}\tag{B.1}$$

where the Hamilton EOMs for x and p have been used involving the Hamilton functions H of the total system. In the last line it has been used $\hat{P}p = p$ and hence $(1 - \hat{P})p = 0$. Similar to the LP-GLE the noise term has only a component acting on momenta p

$$R(t) = -e^{t(1-\hat{P})\mathcal{L}}(1 - \hat{P})\partial_x H.\tag{B.2}$$

In the expansion of the memory functional (integral term in Eq. (1.43)) one finds

$$\begin{aligned}
-\sum_{i=1}^{\infty} \Phi_i [\mathbf{A}(t-\tau)] \langle [\nabla_{\mathbf{A}} \cdot \mathbf{F}(0)] \mathbf{F}(\tau) \rangle &= \sum_{i=1}^{\infty} -\Phi_i [x(t-\tau), p(t-\tau)] \\
&\quad \times \left\langle \left[\frac{\partial \Phi_i}{\partial x} \cdot 0 + \frac{\partial \Phi_i}{\partial p} R(0) \right] \cdot (0, R(\tau))^T \right\rangle \\
&= \sum_{i=1}^{\infty} -\Phi_i [x(t-\tau), p(t-\tau)] \\
&\quad \times \left(0, \left\langle \frac{\partial \Phi_i}{\partial p} R(0) \cdot R(\tau) \right\rangle \right)^T \quad (\text{B.3})
\end{aligned}$$

which, similar to the noise, only contributes in the EOM for p . Now the general equation for the memory functional is expanded up to first order in x and p . In order to calculate memory kernel and noise one needs to construct an orthogonal set $\{\Phi_j(x, p)\}$ of functions. Since all expansions shall be truncated after first order in x and p , these functions are easily found

$$\begin{aligned}
\Phi_1(x, p) &= \frac{x}{\sqrt{\langle x^2 \rangle}} \\
\Phi_2(x, p) &= \frac{p}{\sqrt{mkT}}, \quad (\text{B.4})
\end{aligned}$$

where x and p have been assumed to be canonically distributed. Note, that x and p are already orthogonal and just need to be normalized. Inserting the two functions Φ_1, Φ_2 one finds for the averages

$$-\Phi_1[x(t-\tau), p(t-\tau)] \left\langle \frac{\partial \Phi_1}{\partial p} R(0) \cdot R(\tau) \right\rangle = 0 \quad (\text{B.5})$$

and

$$-\Phi_2[x(t-\tau), p(t-\tau)] \left\langle \frac{\partial \Phi_2}{\partial p} R(0) \cdot R(\tau) \right\rangle = -\frac{p(t-\tau)}{mkT} \langle R(0)R(\tau) \rangle \quad (\text{B.6})$$

which directly defines the memory kernel $\xi(t)$ via a FDT relation

$$\xi(t) = \frac{1}{mkT} \langle R(0)R(t) \rangle. \quad (\text{B.7})$$

The overall integral term in the linearized NLP-GLE reduces to the familiar form

$$\int_0^t \xi[\tau; x(t-\tau), p(t-\tau)] d\tau = \int_0^t p(t-\tau) \xi(\tau) d\tau. \quad (\text{B.8})$$

The explicit part of the NLP-GLE is given by the conditional average

$$\begin{aligned} \langle \partial_t \mathbf{A}; \mathbf{A}(t) \rangle &= \begin{pmatrix} \langle p; p(t) \rangle / m \\ - \langle \partial_x H; x(t) \rangle \end{pmatrix} \\ &= \begin{pmatrix} p(t)/m \\ F_m[x(t)] \end{pmatrix}, \end{aligned}$$

with the mean-force

$$F_m(x) = \frac{\int f(x', \mathbf{Q}) \delta(x' - x) F(x', \mathbf{Q}) dx' d\mathbf{Q}}{\int f(x', \mathbf{Q}) \delta(x' - x) dx' d\mathbf{Q}} \quad (\text{B.9})$$

$$= [P(x)]^{-1} \int f(x', \mathbf{Q}) \delta(x' - x) F(x', \mathbf{Q}) dx' d\mathbf{Q}. \quad (\text{B.10})$$

Here, the definition in Eq. (1.41) has been involved where the configurational part of the phase space variables Γ has been written out explicitly in terms of bath coordinates \mathbf{Q} and system coordinate x . Further, $F(x, \mathbf{Q})$ denotes the total force exerted on the system, i.e. including system-bath interactions. In the canonical ensemble the potential $V_m(x)$ of the mean-force can be calculated as

$$V_m(x) = -kT \cdot \ln[P(x)] \quad (\text{B.11})$$

with $P(x)$ being the reduced distribution function for the system's coordinate

$$\begin{aligned} P(x) &= \int f(x', \mathbf{Q}) \delta(x - x') dx' d\mathbf{Q} \\ &= \int f(x, \mathbf{Q}) d\mathbf{Q}. \end{aligned} \quad (\text{B.12})$$

This statement can proven via direct differentiation

$$-\frac{\partial}{\partial x} V_m(x) = kT \cdot \frac{\partial P}{\partial x} \cdot [P(x)]^{-1}. \quad (\text{B.13})$$

For the calculating the derivative of $P(x)$ one uses the special form of the canonical distribution function $f(x, \mathbf{Q}) = 1/Z \cdot \exp[-V(x, \mathbf{Q})/kT]$ with Z being the configurational partition function and $V(x, \mathbf{Q})$ the total potential of system and bath. One finds

$$\begin{aligned}
 \frac{\partial P}{\partial x} &= \frac{1}{Z} \int \frac{\partial}{\partial x} \exp[-V(x, \mathbf{Q})/kT] d\mathbf{Q} \\
 &= -\frac{1}{ZkT} \int \exp[-V(x, \mathbf{Q})/kT] \frac{\partial V}{\partial x} d\mathbf{Q} \\
 &= \frac{1}{kT} \int f(x, \mathbf{Q}) F(x, \mathbf{Q}) d\mathbf{Q} .
 \end{aligned} \tag{B.14}$$

Combining this result with Eq. (B.13) one ends up with the definition of the mean-force $F_m(x)$. For further reading it is referred to [10, 15, 28].

Appendix C.

Integro-differential Equations for Time-Correlation Functions

The MAFs obey characteristic integro-differential equations which are connected to the underlying GLE of the common form given in Eq. (1.5) in a simple matter. The resulting equations are of the same GLE structure apart from a missing noise term. The usefulness of these equations becomes apparent in Sec. 3.1.1 where it is discussed how the memory kernel can be extracted from explicit MD MAFs. In this appendix the systematic way to derive these equations is presented. The general idea is to apply the ensemble average $\langle p(0) \cdot \dots \rangle$ to the GLE

$$\begin{aligned}\langle p(0)\dot{p}(t) \rangle &= \langle p(0)F[x(t)] \rangle - \int_0^t \xi(t-\tau) \langle p(0)p(\tau) \rangle d\tau + \langle p(0)R(t) \rangle \\ \dot{C}_{pp}(t) &= C_{pF}(t) - \int_0^t \xi(t-\tau)C_{pp}(\tau)d\tau + \langle p(0)R(t) \rangle .\end{aligned}\tag{C.1}$$

Here, the symbol $C_{AB}(t) = \langle A(0)B(t) \rangle$, which denotes the correlation function of variable A and B has been introduced. Further, it has been used that time integration and time differentiation commute with the ensemble average $\langle \dots \rangle$.

The treatment of the correlation function with the noise term $R(t)$ depends on the type of GLE involved. In the GLEs based on projection operators one needs to calculate

$$\langle p(0)R(t) \rangle = \left\langle p(0)e^{t(1-\hat{P})\mathcal{L}}(1-\hat{P})\mathcal{L}p(0) \right\rangle\tag{C.2}$$

where one should remember that the projector \hat{P} projects onto the closed subspace which includes the variable p . Then, from the Hilbert projection theorem, which states that for projectors \hat{P} onto closed subspaces the projectors $1 - \hat{P}$ project onto the orthogonal complement of \hat{P} , it follows that

$$\langle p(0)(1 - \hat{P})A(0) \rangle = 0 \quad (\text{C.3})$$

for any dynamical variable A . This has the consequence that $p(0)$ is uncorrelated with the noise $R(t)$

$$\langle p(0)R(t) \rangle = 0. \quad (\text{C.4})$$

Note, that in the non-linear projection operator technique, where one projects onto the subspace of all (non-linear) functions $F(x, p)$, this can be generalized to

$$\langle F[x(0), p(0)]R(t) \rangle = 0. \quad (\text{C.5})$$

For the linear projection, in contrast, this is only valid for correlations with linear functions $F(x, p) = c_1x + c_2p$. For the MBO-GLE one can use the explicit expression of the noise term

$$\begin{aligned} \langle p(0)R(t) \rangle &= \sum_j g_j \frac{\langle p(0)P_j(0) \rangle}{M_j\omega_j} \sin(\omega_j t) \\ &+ \sum_j g_j \left(\langle p(0)Q_j(0) \rangle - \frac{g_j}{M_j\omega_j^2} \langle p(0)x(0) \rangle \right) \cos(\omega_j t) \end{aligned} \quad (\text{C.6})$$

where all occurring averages vanish in the canonical ensemble.

In all GLEs involved it can be explicitly shown that the correlation of momentum and noise becomes zero and, hence, the integro-differential equation for the MAF reduces to

$$\dot{C}_{pp}(t) = C_{pF}(t) - \int_0^t \xi(t - \tau) C_{pp}(\tau) d\tau. \quad (\text{C.7})$$

Appendix D.

Analytic Spectrum for a harmonic Oscillator

For the special case of a harmonic oscillator the vibrational spectrum can be analytically expressed in terms of the spectral density. This short derivation is presented in this appendix. According to Eq. (1.74) the absorption coefficient is given by the real part of the MAF's Fourier transform which can be easily obtained by transforming the MAF's integro-differential equation into Fourier domain

$$i\omega S_{pp}(\omega) - C_{pp}(0) = S_{pF}(\omega) - J(\omega)S_{pp}(\omega). \quad (\text{D.1})$$

Note, that the term $C_{pp}(0)$ appears because only the half-sided Fourier transforms are taken. For the harmonic oscillator the force F entering the MFC has the form $F = -m\omega_0^2 x$ and hence $C_{pF}(t) = -m\omega_0^2 C_{px}(t)$. The correlation function $C_{px}(t)$ is connected to $C_{pp}(t)$ via

$$\dot{C}_{px}(t) = \frac{C_{pp}(t)}{m} \quad (\text{D.2})$$

which, transformed into the Fourier space, reads

$$i\omega S_{px}(\omega) - C_{px}(0) = \frac{S_{pp}(\omega)}{m}. \quad (\text{D.3})$$

Taking into account that $C_{px}(0) = \langle p(0)x(0) \rangle = 0$ for the canonical ensemble the Fourier transform $S_{pF}(\omega)$ follows from the above relation

$$S_{pF}(\omega) = i \frac{\omega_0^2}{\omega} S_{pp}(\omega). \quad (\text{D.4})$$

Combining these relations one can find for $S_{pp}(\omega)$

$$S_{pp}(\omega) = \frac{\omega C_{pp}(0)}{\omega J(\omega) + i(\omega^2 - \omega_0^2)}, \quad (\text{D.5})$$

which is entirely given by the spectral density $J(\omega)$. Note, that the MAF is often normalized to its initial value which implies $C_{pp}(0) = 1$.

Appendix E.

Gaussian Filtering

In this appendix the technique of Gaussian filtering, as has been used in this thesis for reducing the noise level in spectra and spectral densities, is presented in greater detail. Consider a TCF $C(t)$ which has been computed numerically and which can be decomposed into

$$C(t) = C_0(t) + \epsilon(t), \quad (\text{E.1})$$

where $C_0(t)$ is the exact TCF and $\epsilon(t)$ represents the statistical error. For calculating the spectrum of $C(t)$ a numerical time integration must be performed

$$\begin{aligned} S(\omega) &= \int_0^{\infty} C(t) e^{-i\omega t} dt \\ &\approx \int_0^{T_{\max}} C(t) e^{-i\omega t} dt \\ &= \int_0^{T_{\max}} C_0(t) e^{-i\omega t} dt + \int_0^{T_{\max}} \epsilon(t) e^{-i\omega t} dt, \end{aligned} \quad (\text{E.2})$$

which cannot be employed to infinity and, hence, must be truncated at a sufficiently large time T_{\max} . This time has to be chosen such that the function $C_0(t)$ has come sufficiently close to zero. However, the statistical error $\epsilon(t)$ usually gets larger in the unconverged tails since the number of samples from which the TCF is calculated linearly decreases with time (see Sec. 2.1.1). Thus a cut of $\epsilon(t)$ at large times T_{\max} causes errors in the spectra which have a highly oscillating behavior. Further, converging the

tail of the TCF is no option especially for systems with large correlation times since this would require a large amount of trajectories and timesteps to be involved.

The central idea to overcome this problem without a significant increase of the sampling length is to average out these highly oscillating errors. In Gaussian filtering the spectrum at frequency ω is averaged with respect to a Gaussian weight

$$\bar{S}(\omega) = \int_{-\infty}^{+\infty} S(\omega') \frac{1}{\sqrt{2\pi\Delta\omega^2}} \exp\left[-\frac{(\omega' - \omega)^2}{2\Delta\omega^2}\right] d\omega' \quad (\text{E.3})$$

with a parameter $\Delta\omega$ that determines the frequency width to which the average is restricted. On one hand this parameter has to be chosen such that the error is averaged over sufficiently many oscillations to give

$$\int_{-\infty}^{+\infty} \epsilon(\omega') \frac{1}{\sqrt{2\pi\Delta\omega^2}} \exp\left[-\frac{(\omega' - \omega)^2}{2\Delta\omega^2}\right] d\omega' \approx 0. \quad (\text{E.4})$$

On the other hand it should have a minimal affect on the exact part of the spectrum

$$\int_{-\infty}^{+\infty} S_0(\omega') \frac{1}{\sqrt{2\pi\Delta\omega^2}} \exp\left[-\frac{(\omega' - \omega)^2}{2\Delta\omega^2}\right] d\omega' = S_0(\omega) + \epsilon_S(\omega) \quad (\text{E.5})$$

with a small smoothing error $\epsilon_S(\omega) \approx 0$. The smoothing error can be calculated directly starting from a Taylor expansion of $S_0(\omega')$ around ω

$$S_0(\omega') = S_0(\omega) + \sum_{n=1}^{\infty} \frac{S_0^{(n)}(\omega)}{n!} (\omega' - \omega)^n \quad (\text{E.6})$$

and subsequent convolution with the Gaussian weight requiring to calculate the central moments of a Gaussian function. These vanish for odd powers of $\Delta\omega$ whereas for even powers one gets [49]

$$\int_{-\infty}^{+\infty} (\omega' - \omega)^{2k} \frac{1}{\sqrt{2\pi\Delta\omega^2}} \exp\left[-\frac{(\omega' - \omega)^2}{2\Delta\omega^2}\right] d\omega' = \frac{(2k)!}{2^k k!} \Delta\omega^{2k}. \quad (\text{E.7})$$

Hence, for the smoothing error one finds

$$\begin{aligned}\epsilon_S(\omega) &= \sum_{k=1}^{\infty} \frac{S_0^{(2k)}(\omega)}{2^k k!} \Delta\omega^{2k} \\ &= \frac{S_0''(\omega)}{2} \Delta\omega^2 + \mathcal{O}(\Delta\omega^4)\end{aligned}\tag{E.8}$$

being at minimum of second order in the frequency width. Apparently the smoothing error is large at high curvatures $S_0''(\omega)$ typically being the situation at the cusps of narrow spectral lines. Here, one is limited to small values of $\Delta\omega$ at a price of a higher noise level. Practically, the two requirements for the choice of $\Delta\omega$ cannot be fulfilled exactly at the same time since averaging out the noise generally requires large $\Delta\omega$ whereas rather small widths are needed in order to minimally affect the exact part $S_0(\omega)$. This means that one has to find the optimal compromise between noise reduction and small smoothing errors depending on the particular system under study.

Technically, the easiest way to perform the average in frequency domain is to multiply the time-domain signal $C(t)$ by a Gaussian window $G(t) = \exp[t^2/2T^2]$ before transforming into Fourier domain. The average in Eq. (E.3) is then performed automatically since due to the convolution theorem a product in time domain becomes a convolution in Fourier domain

$$\int_0^{\infty} C(t) \cdot G(t) e^{-i\omega t} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega') G(\omega' - \omega) d\omega' \dots\tag{E.9}$$

The Fourier transform of the Gaussian window is of Gaussian form as well and reads

$$G(\omega) = \sqrt{2\pi T^2} e^{-\frac{1}{2} T^2 \omega^2},\tag{E.10}$$

where the relation between $\Delta\omega$ and the window width T becomes apparent

$$\Delta\omega = \frac{1}{T}.\tag{E.11}$$

A multiplication with a Gaussian window in time domain acts as a low-pass filter rejecting the contributions located at times $t > T$. This is the reason why this smoothing procedure involving a Gaussian window is referred to as *Gaussian filtering* throughout this thesis.

Appendix F.

Derivation of a second order NLP-GLE

In this appendix the second order NLP-GLE is derived. This means, that the expansion of the memory kernel

$$\xi[\tau; x(t-\tau), p(t-\tau)] = \sum_{i=1}^{\infty} -\Phi_i[x(t-\tau), p(t-\tau)] \cdot \left\langle \frac{\partial \Phi_i}{\partial p} R(0) R(\tau) \right\rangle \quad (\text{F.1})$$

is truncated after quadratic terms in x and p . Technically, the functions x, p, xp, x^2, p^2 are orthonormalized via the Gram-Schmidt orthonormalization scheme [49] in order to give the second order set $\{\Phi_i(x, p)\}$. The first two functions of the orthonormal set can be adopted from appendix B

$$\begin{aligned} \Phi_1(x) &= \frac{x}{\sqrt{\langle x^2 \rangle}} \\ \Phi_2(p) &= \frac{p}{\sqrt{mkT}}. \end{aligned} \quad (\text{F.2})$$

The third function $\Phi_3(x, p)$ is constructed from the product term xp upon invoking the Gram-Schmidt orthonormalization

$$\begin{aligned} \Phi_3(x, p) &= \mathcal{N}_3 (xp - \langle \Phi_1 | xp \rangle \Phi_1 - \langle \Phi_2 | xp \rangle \Phi_2) \\ &= \mathcal{N}_3 \left(xp - \frac{\langle x^2 p \rangle}{\langle x^2 \rangle} x - \frac{\langle xp^2 \rangle}{mkT} p \right) \\ &= \mathcal{N}_3 (xp - \langle x \rangle p), \end{aligned} \quad (\text{F.3})$$

where in the second line the definition of the scalar product in terms of ensemble averages, Eq. (1.10), has been used. The discussions in this thesis are always restricted

to the canonical ensemble. Therefore, the special property of the canonical ensemble that all moments of positions are uncorrelated to moments of the momenta, i.e.

$$\langle x^n p^m \rangle = \langle x^n \rangle \langle p^m \rangle, \quad (\text{F.4})$$

can be used in the last line. Further, the relations $\langle p \rangle = 0$ and $\langle p^2 \rangle = mkT$ were invoked. The normalization factor \mathcal{N}_3 of $\Phi_3(x, p)$ follows after some algebra as

$$\mathcal{N}_3 = \frac{1}{\sqrt{mkT (\langle x^2 \rangle - \langle x \rangle^2)}} \quad (\text{F.5})$$

$$= \frac{1}{\sqrt{mkT \sigma_x^2}}, \quad (\text{F.6})$$

giving the final expression for $\Phi_3(x, p)$

$$\Phi_3(x, p) = \frac{xp - \langle x \rangle p}{\sqrt{mkT \sigma_x^2}}. \quad (\text{F.7})$$

Note, that the symbol $\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$ has been used to denote the variance of x . For constructing the next two basis functions $\Phi_4(x, p)$ and $\Phi_5(x, p)$ out of x^2 and p^2 the Gram-Schmidt orthonormalization scheme must, in principle, be continued accordingly. However, since the terms arising from these two functions will turn out to vanish it will be restricted to give their functional form only. Upon carrying out the orthonormalization and using Eq. (F.4) one finds

$$\begin{aligned} \Phi_4(x) &= C_1 x^2 + C_2 x \\ \Phi_5(x, p) &= D_1 p^2 + D_2 x + D_3 x^2, \end{aligned} \quad (\text{F.8})$$

with coefficients $C_{1/2}$ and $D_{1/2}$ which are not specified further. After having constructed the orthonormal set all expectation values $\left\langle \frac{\partial \Phi_i}{\partial p} R(0) R(\tau) \right\rangle$ need to be calculated. This is a very straightforward task

$$\begin{aligned} \left\langle \frac{\partial \Phi_1}{\partial p} R(0) R(\tau) \right\rangle &= 0 \\ \left\langle \frac{\partial \Phi_2}{\partial p} R(0) R(\tau) \right\rangle &= \frac{1}{\sqrt{mkT}} \langle R(0) R(\tau) \rangle \end{aligned}$$

$$\begin{aligned}
\left\langle \frac{\partial \Phi_3}{\partial p} R(0) R(\tau) \right\rangle &= \frac{1}{\sqrt{mkT\sigma_x^2}} \langle (x - \langle x \rangle) R(0) R(\tau) \rangle \\
\left\langle \frac{\partial \Phi_4}{\partial p} R(0) R(\tau) \right\rangle &= 0 \\
\left\langle \frac{\partial \Phi_5}{\partial p} R(0) R(\tau) \right\rangle &= 2D_1 \langle p R(0) R(\tau) \rangle \\
&= 2D_1 \langle p \rangle \langle R(0) R(\tau) \rangle \\
&= 0.
\end{aligned}$$

Note, that in the last average the random force does not depend on the variable p being completely uncorrelated to all other phase space variables. Therefore, its average $\langle p \rangle = 0$ can be moved in front of $\langle R(0) R(\tau) \rangle$ making the whole expression vanishing. From the relations given above it becomes apparent that the only difference to a first order expansion of the NLP-GLE is one additional term stemming from the function Φ_3 . Combining all the results one can write down the second order NLP-GLE

$$\begin{aligned}
\dot{x} &= \frac{p}{m} \\
\dot{p} &= F_m(x) - \int_0^t p(t-\tau) \xi_1(\tau) d\tau - \int_0^t p(t-\tau) \{x(t-\tau) - \langle x \rangle\} \xi_2(\tau) d\tau + R(t),
\end{aligned}$$

with the two memory kernels that are given by

$$\begin{aligned}
mkT\xi_1(t) &= \langle R(0) R(t) \rangle \\
mkT\sigma_x^2\xi_2(t) &= \langle (x - \langle x \rangle) R(0) R(t) \rangle.
\end{aligned} \tag{F.9}$$

The first memory kernel, $\xi_1(t)$, coincides with the one from the linearized NLP-GLE whereas the latter introduces a second FDT relation. Note, that in the main text of the thesis it has been assumed $\langle x \rangle = 0$ which can be always realized by a proper variable substitution. Note also, that the mean-force $F_m(x)$ does not depend on the expansion of the memory functional and is hence exactly the same as in the linearized NLP-GLE.

Appendix G.

Short Documentation of the current Colored Noise Implementation

In the framework of this thesis an implementation of the Colored Noise thermostat method according to Ceriotti *et al.* [43–46] has been established. A short documentation of the current implementation shall be given here. I have to stress at this point that there is room for optimization especially in the useability of the written packages. However, the correctness of the procedures has been tested comprehensively for the systems investigated in this thesis. In case of questions feel free to contact me. Further, I would be grateful if I receive your messages concerning bugs and errors.

G.1. Headers and Object Files

The Colored Noise package has been written in the programming language C and is constructed in a semi object-oriented way. The written procedures are contained in header and object files which have to be included into the main C-programme. The header files to be included are

1. `mathop.h`: containing mathematical operations like matrix multiplications, a cholesky decomposition scheme, a Gaussian random number generator and much more
2. `dynamics.h`: containing necessary routines and variables needed for the time propagation. Further, it provides a routine for calculating TCFs and for creating distribution functions of dynamical quantities

3. GLE.h: containing the routines needed for performing GLE simulations based on the method of Colored Noise thermostats

The implementation of the routines are contained in the object files `name.c` corresponding to the header files `name.h`. The files are provided by me, so please contact me in the case you want to use them. If you use the GNU-compiler the object files need to be compiled with the `-c` option: `gcc -c name.c -o name.o -lm`.

G.2. Data Structure and Variables

At the current state of the implementation the routines for propagating the system involve a specific set of variables which the user has to adopt exactly. These variables are listed in the following table.

variable name	data type	explanation
Dim	integer	spacial dimensionality
Npart	integer	number of particles
length	double	edge of the cubic simulation box
PBC	integer	periodic boundary conditions (on: 1, off: 0)
dt	double	timestep
NStep	integer	number of timesteps per trajectory
Ntraj	integer	number of trajectories
NEqui	integer	number of equilibration steps
Nnext	integer	sampling distance of trajectories
Ncorr	integer	number of timesteps in TCF
NBins	integer	number of bins for distribution function
xmin	double	lower boundary of bin interval
xmax	double	upper boundary of bin interval

Table G.1.: Variables used in the Colored Noise implementation

The Trajectories of `Npart` particles are stored in arrays of data type `double`. Positions and forces are stored in arrays of size `Npart*Dim` if `Dim` is the spacial dimension. Momenta are stored arrays of size `Npart*Dim*(Ns+1)` with `Ns` auxiliary momenta per degree of freedom. The data structure is organized as follows: if one wants

to access the j th spacial component of the i th particle's position or force one has to dereference the arrays as $x[i*Dim+j]$ or $F[i*Dim+j]$, respectively. The j th component of the momentum corresponding to the i th particle is to be dereferenced as $p[i*Dim*(Ns+1)+j*(Ns+1)]$. The user just needs to declare the position, momentum and force arrays in the main programme. Functions for allocating these arrays in the correct way are provided. The initial positions have to be given by the user taking care of not producing too strong initial forces. Allocating and initializing the drift matrices for the Colored Noise scheme can be done by calling a function as well. Inverse masses (which are used here instead of masses) and charges, being arrays of size N_{part} , are allocated in a function but need to be initialized in the main code.

G.3. Routines

Here, a short explanation of the routines needed in the main programme is given.

1. After declaration of necessary variables the function `SetupDyn(FILE *IN)`, to which a file pointer has to be passed, needs to be called. This function reads the input file 'dynamics_in' from which all variables in table [G.1](#) are extracted.
2. The function `SetupGLE(FILE *IN, double **imass, double **charge, double **F)` needs to be called next. Here, the input file 'GLE_in' specifying the thermostat properties is read and the drift matrices are set up. Further, inverse masses, charges and forces are allocated. Note, that arguments have to be declared as pointers but have to be passed via the `&`-operator
3. Positions and momenta are allocated via `GLEallocate(double **x, double **p)`. Further, this function initializes all momenta to zero. Note, that positions must be initialized in the main programme. Note also, that the arguments have to be passed with the `&`-operator.
4. As has been mentioned above, the function that calculates forces has to be written by the user himself. Here, it is mandatory to explicitly call this function `forces(double *x, double *F)` with x being the position array and F the force array.

5. The function `GLEpropagate(double *x, double *p, double *F, int tag)` performs one GLE-propagation step if `tag` is set to 1. If `tag` is set to 0 the function performs one NVE-propagation step.
6. The function `binning(FILE *Bin, double *binvariable)` creates a distribution function of the data stored in `binvariable`. This array constitutes a list of the values to be binned over timesteps. The distribution function is normalized to unity, i.e. the integral over the whole bin interval is 1. The distribution function is always written to a file called 'Bin.dat'.
7. The function `Correlation(double *obs1, double *obs2, double *corr, int D)` calculates the TCF of the data in `obs1` and `obs2` and stores the result in `corr`. The arrays can have a dimension `D` which is not necessarily equal to `Dim`. The data structure of the arrays must be such that `obs[i*D+j]` dereferences the j th component of `obs` at the i th timestep. Note, that this function calculates the correlation function as a time-average and accumulates this average over the trajectories. One therefore needs to call it in the loop over trajectories.

G.4. Input Files

There are two input files to be prepared. The input file 'dynamics_in' contains a list of the parameters given in the table [G.1](#). The parameters have to be listed in the same order as in they appear in this table. The second input file 'GLE_in' contains the specifications of the Colored Noise thermostat. The first line therein is the Boltzmann parameter $\beta = 1/kT$. The second line specifies the Markovian friction called a_{pp} in Eq. (1.51). In the next three lines one needs to specify the number of exponential functions, damped cosine functions and damped sine functions (in this order!) used to build the memory kernel. Note, that a sine function is included in the implementation but should not be used at the present state because its implications are not comprehensively tested. After specifying the number of functions their corresponding parameters are listed. First one has to list the blocks of coefficients a, b corresponding to an exponentially damped $a^2 \exp[-bt]$. Afterwards, all the blocks of coefficients a, b, c according to the damped cosine functions $2a^2 \exp[-bt] \cos[ct]$ are listed. To give an example, the file 'GLE_in' corresponding to $\beta = 1$, $a_{pp} = 0$, $a_{pp} = 0.0$, two exponential function with $a = 1$, $b = 2$ and $a = 3$, $b = 4$, one damped cosine function with $a = 5$,

$b = 6$, $c = 7$ and zero sine functions reads 1.0 0.0 2 1 0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 with a line break after each number.

G.5. Example Code

In the following example code the MAF of a harmonic oscillator is calculated and its momentum distribution function is printed. When compiling this code, called `colored.c`, via the GNU compiler one has to correctly link it with the object files via `gcc colored.c -o colored.x mathop.o dynamics.o GLE.o -lm`, where it is assumed that the object files are in the same directory as the code. Note, that linking with the math library is necessary.

```

#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>
#include "mathop.h"
#include "dynamics.h"
#include "GLE.h"

double *charge;

/*****
/*                      FORCE FIELD                      */
*****/

// Calculate forces.
// An implementation of a force field library is in progress.
// So far, you have to write forces yourself
// In this example a harmonic oscillator is set up
void force(double *x, double *F)
{
    int i;
    double k=0.16; // force constant

    for(i=0;i<Dim*Npart;i++)
        F[i]=-k*x[i];
}

/*****
/*                      MAIN FUNCTION                      */
*****/

int main()
{
    int i,j,k;
    double *x,*p,*F;           // positions, momenta, force
    double *imass;             // inverse masses
    double *xequi,*pequi;      // position and momenta for equilibration
    double *D_pp;              // Samples for which correlation function
                                // is calculated
    double *binvariable;       // Samples to be binned
    FILE *Bin,*IN;             // FILE pointers
    double *C_pp;              // array that contains correlation
                                // function of D_pp data

    /***** ALLOCATIONS *****/

    // Reads input file "dynamics_in" and sets up variables needed
    // for propagation
    SetupDyn(IN);

    // Reads input file "GLE_in" and sets up propagation matrices
    // for Colored Noise and allocates masses, charges and
    // a force array (note: you need only one force array)
    SetupGLE(IN,&imass, &charge, &F);

    // Allocates two pairs of positions and momenta
    GLEallocate(&x,&p);

```

```

GLEallocate(&xequi,&pequi);

// Set the seed of random numbers (time.h needs to be included!)
srand(time(NULL));

// Allocation of correlation function
C_pp=malloc(NStep* sizeof(double));

// Allocates help arrays to store data sequence from which
// correlation functions shall be calculated
D_pp=malloc(NStep* sizeof(double));

// Data sequence from which a histogram shall be established
binvariable=malloc( (Ntraj*NStep)* sizeof(double));

/***** INITIALISATIONS *****/

// initialize masses of particles (here set to 1)
for(i=0;i<Npart;i++)
    imass[i]=1.0;

// initial positions
// Note: momenta are initialized in GLEallocate
// CAUTION: Choose initial configuration carefully
// not to produce large initial forces
for(i=0;i<Npart*Dim;i++)
    xequi[i]=0.0;

// initialise forces!
// CAUTION: Do not forget this
force(xequi,F);

// initialize correlation function to zero!
for(i=0;i<NStep;i++)
    C_pp[i]=0.0;

/***** TIME PROPAGATION *****/

// 1.) equilibration
for(i=0;i<NEqui;i++)
    GLEpropagate(xequi,pequi,F,imass,1);

for(i=0;i<Ntraj;i++)
{
    // copy equilibration arrays to production arrays
    // function is provided in mathop.h
    Copy(x,xequi,Dim*Npart);
    Copy(p,pequi,(Ns+1)*Dim*Npart);

    // 2.) propagate ensemble member
    for(j=0; j<NStep;j++)
    {

        // stores data for binning
        *(binvariable+j+i*NStep)=*p;

        // stores data for correlation function

```

```
        // here: momentum of first particle
        D_pp[j]=*p;

        // performs one GLE step
        GLEpropagate(x,p,F,imass,1);
    }

    // 3.) progress correlation function
    Correlation(D_pp,D_pp,C_pp,1);

    // 4.) generate next ensemble member
    for(j=0;j<Nnext;j++)
        GLEpropagate(xequi,pequi,F,imass,1);
}

/***** OUTPUT OF DATA *****/

// Output time-correlation function
// prints on stdout. Needs to be redirected into file
for(i=0;i<Ncorr;i++)
    printf("%lf \t %lf\n", i*dt,C_pp[i]);

// Output histogram (stored in File "Bin.dat")
binning(Bin,binvariable);

return 0;
}
```


Bibliography

- 1 M. Beck, A. Jäckle, and G. Worth. The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets. *Phys. Rep.*, 324:1, 2000.
- 2 D. Marx and J. Hutter. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press, 2009.
- 3 M. A. Sepúlveda and F. Grossman. Time-dependent semiclassical mechanics. In I. Prigogine and S. Rice, editors, *Advances in chemical physics*, volume XCVI. John Wiley & Sons, 1996.
- 4 A. Szabo. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Mineola: Courier Dover Publications, 1996.
- 5 M. L. Koszykowski, D. W. Noid, and R. A. Marcus. Semiclassical Theory of Intensities of Vibrational Fundamentals, Overtones, and Combination Bands. *J. Chem. Phys.*, 86:2113, 1982.
- 6 R. B. Shirts. Use of classical Fourier Amplitudes as Quantum Matrix Elements: A Comparison of Morse Oscillator Fourier Coefficients with Quantum Matrix Elements. *J. Phys. Chem.*, 91:2258, 1986.
- 7 D. S. Lemons. Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien", C. R. Acad. Sci. (Paris) 146, 530-533 (1908)]. *Am. J. Phys.*, 65:1079, 1997.
- 8 L. S. Ornstein and G. E. Uhlenbeck. On the Theory of the Brownian Motion. *Phys. Rev.*, 36:823, 1930.
- 9 S. Chandrasekhar. Stochastic Problems in Physics and Astronomy. *Rev. Mod. Phys.*, 15:1, 1943.
- 10 R. Zwanzig. *Nonequilibrium Statistical Mechanics*. Oxford University Press, 2001.

- 11 H. A. Kramers. Brownian Motion in a Field of Force and the Diffusion Model. *Physica*, 7:284, 1940.
- 12 P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory : fifty years after Kramers. *Rev. Mod. Phys.*, 62:251, 1990.
- 13 H. Mori. Transport, Collective Motion, and Brownian Motion. *Prog. Theor. Phys.*, 33:423, 1965.
- 14 R. Zwanzig. Nonlinear Generalized Langevin Equations. *J. Stat. Phys.*, 9:215, 1973.
- 15 K. Kawasaki. Simple derivations of generalized linear and nonlinear Langevin equations. *J. Phys. A Math. Nucl. Gen.*, 6:1289, 1973.
- 16 R. M. Whitnell, K. R. Wilson, and J. T. Hynes. Fast Vibrational Relaxation for a Dipolar Molecule in a Polar Solvent. *J. Phys. Chem.*, 94:8625, 1990.
- 17 I. Benjamin and R. M. Whitnell. Vibrational relaxation of I_2^- in water and ethanol: molecular dynamics simulation. *Chem. Phys. Lett.*, 204:45, 1993.
- 18 M. Tuckerman and B. J. Berne. Vibrational relaxation in simple fluids: Comparison of theory and simulation. *J. Chem. Phys.*, 98:7301, 1993.
- 19 S. Gnanakaran and R. M. Hochstrasser. Vibrational relaxation of HgI in ethanol: Equilibrium molecular dynamics simulations. *J. Chem. Phys.*, 105:3486, 1996.
- 20 S. Mukamel. *Principles of Nonlinear Optical Spectroscopy*. Oxford University Press, 1995.
- 21 S. Palese and S. Mukamel. Interrogation of Vibrational Structure and Line Broadening of Liquid Water by Raman-Induced Kerr Effect Measurements within the Multimode Brownian Oscillator Model. *J. Phys. Chem.*, 3654:10380, 1996.
- 22 K. Okumura and Y. Tanimura. Femtosecond two-dimensional spectroscopy from anharmonic vibrational modes of molecules in the condensed phase. *J. Chem. Phys.*, 107:2267, 1997.
- 23 S. Woutersen and H. J. Bakker. Hydrogen Bond in Liquid Water as a Brownian Oscillator. *Phys. Rev. Lett.*, 83:2077, 1999.
- 24 M. M. Toutounji and G. J. Small. The underdamped Brownian oscillator model with Ohmic dissipation: Applicability to low-temperature optical spectra. *J. Chem.*

Phys., 117:3848, 2002.

- 25 Y. Tanimura and A. Ishizaki. Modeling, calculating, and analyzing multidimensional vibrational spectroscopies. *Acc. Chem. Res.*, 42:1270, 2009.
- 26 T. Joutsuka and K. Ando. Vibrational spectroscopy and relaxation of an anharmonic oscillator coupled to harmonic bath. *J. Chem. Phys.*, 134:204511, 2011.
- 27 E. Cortes, B. J. West, and K. Lindenberg. On the generalized Langevin equation: Classical and quantum mechanical^{a)}. *J. Chem. Phys.*, 82:2708, 1985.
- 28 S. Kawai and T. Komatsuzaki. Derivation of the generalized Langevin equation in nonstationary environments. *J. Chem. Phys.*, 134:114523, 2011.
- 29 H. K. McDowell. Quantum generalized Langevin equation: Explicit inclusion of nonlinear system dynamics. *J. Chem. Phys.*, 112:6971, 2000.
- 30 D. Banerjee, B. C. Bag, S. K. Banik, and D. S. Ray. Solution of quantum Langevin equation: approximations, theoretical and numerical aspects. *J. Chem. Phys.*, 120:8960, 2004.
- 31 B. J. Berne, M. E. Tuckerman, John E. Straub, and A. L. R. Bug. Dynamic friction on rigid and flexible bonds. *J. Chem. Phys.*, 93:5084, 1990.
- 32 B. J. Berne and G. D. Harp. On the calculation of time correlation functions. *Adv. Chem. Phys.*, 17:63, 1970.
- 33 M. Berkowitz, J. D. Morgan, D. J. Kouri, and J. A. McCammon. Memory kernels from molecular dynamics. *J. Chem. Phys.*, 75:2462, 1981.
- 34 O. F. Lange and H. Grubmüller. Collective Langevin dynamics of conformational motions in proteins. *J. Chem. Phys.*, 124:214903, 2006.
- 35 H. K. Shin, C. Kim, P. Talkner, and E. K. Lee. Brownian motion from molecular dynamics. *Chem. Phys.*, 375:316, 2010.
- 36 G. Goodyear and R. M. Stratt. The short-time intramolecular dynamics of solutes in liquids. I. An instantaneous-normal-mode theory for friction. *J. Chem. Phys.*, 105:10050, 1996.
- 37 A. D. Baczewski and S. D. Bond. Numerical integration of the extended variable generalized Langevin equation with a positive Prony representable memory kernel. *J. Chem. Phys.*, 139:044107, 2013.

- 38 J. E. Basconi and M. R. Shirts. Effects of temperature control algorithms on transport properties and kinetics in molecular dynamics simulations. *J. Chem. Theory Comput.*, 9:2887, 2013.
- 39 F. Gottwald. Why does classical molecular dynamics yield quantum-mechanical infrared spectra? Bachelor's thesis, Rostock University, Rostock, Germany, 2012.
- 40 V. May and O. Kühn. *Charge and Energy Transfer Dynamics in Molecular Systems*. Wiley-VCH, 2011.
- 41 F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham, and G. A. Voth. An accurate and simple quantum model for liquid water. *J. Chem. Phys.*, 125:184507, 2006.
- 42 S. D. Ivanov, A. Witt, M. Shiga, and D. Marx. Communications: On artificial frequency shifts in infrared spectra obtained from centroid molecular dynamics: Quantum liquid water. *J. Chem. Phys.*, 132:031101, 2010.
- 43 M. Ceriotti, G. Bussi, and M. Parrinello. Langevin Equation with Colored Noise for Constant-Temperature Molecular Dynamics Simulations. *Phys. Rev. Lett.*, 102:020601, 2009.
- 44 M. Ceriotti, G. Bussi, and M. Parrinello. Nuclear Quantum Effects in Solids Using a Colored-Noise Thermostat. *Phys. Rev. Lett.*, 103:030603, 2009.
- 45 M. Ceriotti, G. Bussi, and M. Parrinello. Colored-Noise Thermostats à la Carte. *J. Chem. Theory Comput.*, 6:1170, 2010.
- 46 J. A. Morrone, T. E. Markland, M. Ceriotti, and B. J. Berne. Efficient multiple time scale molecular dynamics: Using colored noise thermostats to stabilize resonances. *J. Chem. Phys.*, 134:014103, 2011.
- 47 R. Ramírez, T. López-Ciudad, P. Kumar P, and D. Marx. Quantum corrections to classical time-correlation functions: hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.*, 121:3973, 2004.
- 48 D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2 edition, 2004.
- 49 I. N. Bronstein, K. A. Semendjajew, G. Musiol, and H. Mühlig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 7 edition, 2008.

-
- 50 M. Frigo and S. G. Johnson. FFTW for version 3.3.3, 2012. URL: <http://www.fftw.org/fftw3.pdf>.
- 51 T. Köddermann, D. Paschek, and R. Ludwig. Molecular dynamic simulations of ionic liquids: a reliable description of structure, thermodynamics and dynamics. *Chemphyschem*, 8:2464, 2007.
- 52 S. Habershon, T. E. Markland, and D. E. Manolopoulos. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.*, 131:024501, 2009.
- 53 F. Paesani and G. A. Voth. A quantitative assessment of the accuracy of centroid molecular dynamics for the calculation of the infrared spectrum of liquid water. *J. Chem. Phys.*, 132:014105, 2010.
- 54 T. Zentel. (Non-)linear Spectroscopy Based on Classical Trajectories. Master's thesis, Rostock University, Rostock, Germany, 2012. URL: http://rosdok.uni-rostock.de/resolve/id/rosdok_thesis_0000000013.
- 55 B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, 4:435, 2008.

Selbständigkeitserklärung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig angefertigt und ohne fremde Hilfe verfasst habe, keine außer den von mir angegebenen Hilfsmitteln und Quellen dazu verwendet habe und die den benutzten Werken inhaltlich und wörtlich entnommenen Stellen als solche kenntlich gemacht habe.

Rostock, 29.08.2014