Chapter 2

Theory

In this chapter the theoretical background for the numerical calculations performed in this thesis are presented. To provide an as closed picture as possible the derivation of standard techniques (e.g. the Born–Oppenheimer approximation) is sketched out, while a closer look is taken at the less widely used theories concerning dissipative effects. The presentation follows the steps necessary for the theoretical analysis of a molecule, for which one wants to know the dynamics of its intramolecular bonds. First, in Section 2.1, the quantum chemical methods are described, which are used to calculate the geometries, the molecular potentials and other static properties of the system (e.g. the dipole moment). In Section 2.2 the Cartesian reaction Hamiltonian, which is used to study the dynamics of the system, and how to construct it from the *ab initio* data of the preceding section, is discussed. Section 2.3 describes the methods necessary for the description of quantum systems which are either isolated, or able to exchange energy with a bath. In Section 2.4 possible numerical implementation of this description are introduced. In the final Section 2.5 of this chapter various methods proposed in literature for the control of molecular dynamics are presented.

2.1 Quantum Chemistry

This Section will only give a short sketch of the quantum chemical techniques used in this work. A more detailed treatment of this topic is extensively covered in the literature [69, 70, 71, 72].

To calculate the stationary properties of a molecule, especially its geometry, it is necessary to solve the time independent Schrödinger equation for the molecular Hamiltonian $\hat{\mathbf{H}}_{\mathrm{mol}}$:

$$\hat{\mathbf{H}}_{\mathrm{mol}}\Psi = E\Psi, \tag{2.1}$$

which describes a many body system consisting of nuclei and electrons. An analytical solution for this problem is only possible for very few simple systems, e.g. the harmonic oscillator or the Hydrogen atom. To describe the molecules relevant for this work one

has to apply several approximations before a solution of Eq. (2.1) is possible. The most prominent among them is the Born–Oppenheimer approximation, used to separate the motion of the heavy nuclei from that of the light electrons. Even this separated problem cannot be solved exactly, but has to be evaluated with numerical methods. In this section a short sketch of the Born–Oppenheimer approximation will be presented. After this the standard quantum chemical methods used in this work to calculate molecular properties will be presented. The basis of most of the numerical solutions for the electronic system is the Hartree–Fock self consistent field approach. This can be refined via perturbation theory based on the work of Møller and Plesset. Another approach to solve the electronic problem of the molecule are density functional methods, which are reviewed in the final part of this section.

2.1.1 The Born–Oppenheimer Approximation

The Born–Oppenheimer approximation [73] is used to separate the motion of the heavy nuclei of a molecule from the much faster dynamics of the electrons. The complete molecular Hamiltonian of a system consisting of N_{nuc} atoms with nuclear charges $Z_1, \ldots, Z_{N_{\mathrm{nuc}}}$, Cartesian positions R_n and momenta P_n and N_{el} electrons with positions and momenta written as r_n and p_n, respectively, is generally given in the form

$$\hat{\mathbf{H}}_{\mathrm{mol}} = \hat{\mathbf{T}}_{\mathrm{el}} + \hat{\mathbf{V}}_{\mathrm{el-el}} + \hat{\mathbf{T}}_{\mathrm{nuc}} + \hat{\mathbf{V}}_{\mathrm{nuc-nuc}} + \hat{\mathbf{V}}_{\mathrm{el-nuc}}. \tag{2.2}$$

The single terms are the kinetic energy for the electrons, $\hat{\mathbf{T}}_{\mathrm{el}}$, and for the nuclei, $\hat{\mathbf{T}}_{\mathrm{nuc}}$, as

$$\hat{\mathbf{T}}_{\mathrm{el}} = \sum_{j=1}^{N_{\mathrm{el}}} \frac{p_j^2}{2m_e} \quad \text{and} \quad \hat{\mathbf{T}}_{\mathrm{nuc}} = \sum_{j=1}^{N_{\mathrm{nuc}}} \frac{P_j^2}{2M_j},$$

where m_e is the electron mass and M_j is the mass of the jth nucleus. The interaction between these particles via the Coulomb force results in the potential terms for the electron–electron interaction $\hat{\mathbf{V}}_{\mathrm{el-el}}$, the nucleus–nucleus interaction $\hat{\mathbf{V}}_{\mathrm{nuc-nuc}}$ and for the electron–nucleus interaction $\hat{\mathbf{V}}_{\mathrm{el-nuc}}$, as

$$\hat{\mathbf{V}}_{\mathrm{el-el}} = \frac{1}{2}\sum_{i\neq j} \frac{e^2}{|r_i - r_j|}\ ,\ \hat{\mathbf{V}}_{\mathrm{nuc-nuc}} = \frac{1}{2}\sum_{i\neq j} \frac{Z_i Z_j e^2}{|R_i - R_j|} \text{ and } \hat{\mathbf{V}}_{\mathrm{el-nuc}} = -\sum_{i,j} \frac{Z_j e^2}{|r_i - R_j|}.$$

The solution of the Schrödinger equation for this complex many body problem is then simplified by the application of the Born–Oppenheimer approximation, which assumes that the electrons move in the electro-static field generated by a fixed geometry of the nuclei. This is motivated by the fact, that due to the large mass difference between electrons and nuclei the electrons will be able to respond instantaneously to any change in the nuclear configuration. Therefore it is possible to represent the electronic Hamiltonian in a form, which only depends parametrically on the nuclear coordinates $\mathbf{R}$:

$$\hat{\mathbf{H}}_{\mathrm{el}}(\mathbf{R}) = \hat{\mathbf{T}}_{\mathrm{el}} + \hat{\mathbf{V}}_{\mathrm{el-el}} + \hat{\mathbf{V}}_{\mathrm{el-nuc}} \tag{2.3}$$

The solution of the stationary Schrödinger equation Eq. (2.1) with this electronic Hamiltonian is the objective of quantum chemistry programs. It results in electronic energies $E_n^{\mathrm{el}}(\mathbf{R})$ and wavefunctions $|\phi_n(\mathbf{r},\mathbf{R})\rangle$, which will depend prametrically on the nuclear geometry:

$$\hat{\mathbf{H}}_{\mathrm{el}} |\phi_n(\mathbf{r},\mathbf{R})\rangle = E_n^{\mathrm{el}}(\mathbf{R}) |\phi_n(\mathbf{r},\mathbf{R})\rangle \tag{2.4}$$

The solution of the electronic part of the Schrödinger equation for different nuclear coordinates $\mathbf{R}$ results in the potential hypersurface $V(\mathbf{R})$ when the inter-nuclear repulsion is added. This part of the potential is constant with respect to the electronic coordinates. A further potential term, which results from the non-adiabatic coupling operator, describing the interaction between the different electronic states [61] is neglected here, as this work concentrates on the processes in a single adiabatic potential energy surface, which can be written as:

$$V_n(\mathbf{R}) = E_n^{\mathrm{el}}(\mathbf{R}) + \hat{\mathbf{V}}_{\mathrm{nuc-nuc}} \tag{2.5}$$

If the electronic states of the molecule are well separated from each other this approach is justified. The elimination of the non-adiabatic electronic coupling is the core of the Born–Oppenheimer approximation, which leads to a nuclear Schrödinger equation

$$\hat{\mathbf{H}}_{\mathrm{nuc}} |\psi_n(\mathbf{R})\rangle = \left(\hat{\mathbf{T}}_{\mathrm{nuc}} + V_n(\mathbf{R})\right) |\psi_n(\mathbf{R})\rangle = E_n^{\mathrm{nuc}} |\psi_n(\mathbf{R})\rangle \tag{2.6}$$

This describes the geometry of the nuclei in the average field generated by the fast moving electrons. One has to note, that now there will be a different nuclear potential for each electronic state. The dynamics of the nuclei is described reasonably well with the nuclear potential $V(\mathbf{R})$ within the Born–Oppenheimer approximation, as long as the potential surfaces belonging to different states stay well separated.

2.1.2 The Hartree–Fock (HF) Method

While taking a look at the Born–Oppenheimer approximation it was noted that the potential hypersurface and therefore the nuclear dynamics is mainly determined by the solution of the electronic Schrödinger equation (2.4) for a fixed nuclear configuration. While this problem is easier to solve than the complete molecular Hamiltonian, it is still not possible to calculate the electronic orbitals and energies exactly for anything but the most simple systems. To evaluate more complicated molecules, it is again necessary to resort to some approximations. One of the oldest methods, which is also the basis of other, more refined theories, is the Hartree–Fock self consistent field method (HF-SCF) [74]. This method gives an approximate solution of the electronic Schrödinger equation (2.4) using the Hamiltonian $\hat{\mathbf{H}}_{\mathrm{el}}$ as obtained from the Born–Oppenheimer approximation.

The HF method is a non-relativistic approach, in which the single electrons are described by single particle functions $\chi_n(x)$ consisting of a product of a spatial orbital $\psi(r)$, depending on the position of the electron and a spin orbital $\alpha(\omega)$ or $\beta(\omega)$ depending only

on the spin coordinate:

$$\chi_n(x) = \chi_n(r,\omega) = \psi(r) \cdot \begin{cases} \alpha(\omega) \\ \beta(\omega) \end{cases} \tag{2.7}$$

The total electronic wavefunction of the N_{el} electron molecular system can then be described – at least within the HF approximation – as a single, anti-symmetric Slater determinant:

$$\Psi_{\mathrm{el}}(\mathbf{x}) = \frac{1}{\sqrt{N_{\mathrm{el}}!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_1(x_{N_{\mathrm{el}}}) \\ \chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_1(x_{N_{\mathrm{el}}}) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_{N_{\mathrm{el}}}) & \chi_2(x_{N_{\mathrm{el}}}) & \cdots & \chi_{N_{\mathrm{el}}}(x_{N_{\mathrm{el}}}) \end{vmatrix}. \tag{2.8}$$

The single particle functions $\chi_n(x)$ are then determined by minimizing the functional for the energy expectation value:

$$E_{\mathrm{el}}\left[\Psi_{\mathrm{el}}\right] = \frac{\langle \Psi_{\mathrm{el}} | \, \hat{\mathbf{H}}_{\mathrm{el}} \, | \Psi_{\mathrm{el}} \rangle}{\langle \Psi_{\mathrm{el}} | \, \Psi_{\mathrm{el}} \rangle}. \tag{2.9}$$

To perform this minimization it is convenient to split the electronic Hamiltonian $\hat{\mathbf{H}}_{\mathrm{el}}$ Eq. (2.3) into a sum of single electron operators $\hat{\mathbf{h}}(r_i)$, which affect only the ith electron and the non-separable interaction potential $\hat{\mathbf{V}}_{\mathrm{el-el}}$ between all electrons:

$$\hat{\mathbf{H}}_{\mathrm{el}} = \sum_i \hat{\mathbf{h}}(r_i) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|r_i - r_j|}, \tag{2.10}$$

where

$$\hat{\mathbf{h}}(r_i) = \frac{p_i^2}{2m_e} + \hat{\mathbf{V}}_{\mathrm{el-el}}. \tag{2.11}$$

In addition the orthonormality of the single particle functions $\langle \chi_n | \, \chi_m \rangle = \delta_{mn}$ is assumed.

After performing the variational optimization, one obtains the Hartree–Fock equations for the single particle spin orbitals $\chi_n(x)$:

$$\hat{\mathbf{F}} \chi_n(x) = \epsilon_n \chi_n(x) \tag{2.12}$$

with the Fock operator $\hat{\mathbf{F}}$. It consists of the single particle operator $\hat{\mathbf{h}}$ and the so called Coulomb- and exchange operators $\hat{\mathbf{J}}$ and $\hat{\mathbf{K}}$, respectively:

$$\begin{aligned} \hat{\mathbf{F}}\chi_i(x) &= \hat{\mathbf{h}}\chi_i(x) + \underbrace{e^2 \sum_j^{N_{\mathrm{el}}} \int \frac{\chi_j^*(x')\chi_j(x')}{|r-r'|} \, dr' \chi_i(x)}_{=\hat{\mathbf{J}}\chi_i(x)} - \underbrace{e^2 \sum_j^{N_{\mathrm{el}}} \int \frac{\chi_j^*(x')\chi_j(x)}{|r-r'|} \chi_i(x') \, dr'}_{=\hat{\mathbf{K}}\chi_i(x)} \\ &= \epsilon_i \chi_i(x). \end{aligned} \tag{2.13}$$

This form can be interpreted as a single particle operator consisting of the basic operator $\hat{\mathbf{h}}$ for one electron with an additional effective Hartree–Fock potential term

$$V^{HF}(x) = e^2 \sum_i^{N_{\mathrm{el}}} \left(\hat{\mathbf{J}}_i(x) - \hat{\mathbf{K}}_i(x) \right), \tag{2.14}$$

which can be written in the simple form

$$\left(\hat{\mathbf{h}} + V^{HF}(x)\right) \chi_i(x) = \epsilon_i \chi_i(x). \tag{2.15}$$

The interpretation for the Coulomb part of this potential is straightforward, it just describes the interaction of one electron with the charge of all other electrons located in all the other single particle orbitals. The exchange part on the other hand has no classical counterpart and is a purely quantum mechanical effect, caused by the anti-symmetric *ansatz* of the wavefunction. The energy of the HF molecular orbital is then calculated via

$$E_{HF} = \sum_i h_{ii} + \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}) \tag{2.16}$$

with $h_{ii} = \langle \chi_i | \hat{\mathbf{h}} | \chi_i \rangle$ and the Coulomb and correlation energies as derived from the corresponding operators

$$J_{ij} = \int \frac{\chi_j^*(x')\chi_j(x')\chi_i^*(x)\chi_i(x)}{|r - r'|} \, dr'dr \tag{2.17}$$

and

$$K_{ij} = \int \frac{\chi_j^*(x')\chi_i(x')\chi_i^*(x)\chi_j(x)}{|r - r'|} \, dr'dr. \tag{2.18}$$

To solve these equations an iterative method has to be employed, as the operator generating the single particle functions $\chi_i(x)$ itself depends on these functions via the Hartree–Fock potential $V^{HF}(x)$. Therefore it is necessary to start with an initial guess of the $\chi_i(x)$, use them to calculate an approximate $V^{HF}(x)$ and from this the Fock operator $\hat{\mathbf{F}}$, recalculate the $\chi_i(x)$ with this operator and repeat this procedure until the single particle functions converge to a stable solution.

Basis Sets

To further simplify the problem and reduce the numerical effort the standard quantum chemistry programs represent the spatial part of the single particle functions $\psi(r)$ from which the total molecular wavefunction is built as a linear combination of fixed basis functions:

$$\psi(r) = \sum_k C_k \varphi_k(r). \tag{2.19}$$

This leads to a representation of the problem as a set of linear equations. Within this approximation the electronic Schrödinger equation can now be solved via the determination of the coefficients of these combinations.

Historically this linear combinations were built from atomic orbitals or from a linear combination of Slater–type orbitals (STOs) in the form

$$\varphi_k \propto r^{k-1} e^{-\zeta r} Y_l^m(\theta, \phi), \tag{2.20}$$

which are centered on each atom. This approach is named linear combination of atomic orbitals (LCAO), and leads to an equation for the coefficients of the molecular orbitals (Roothaan equation). The solution of these equations for larger molecules requires an evaluation of three and four center integrals (i.e. integrals over basis functions centered on up to four different atoms), which is very time consuming for STOs. In modern quantum chemistry programs the atomic or Slater–type orbitals normally are not used. Instead a collection of square integrable functions, usually easy to integrate Gaussians, are used. A typical Cartesian Gaussian basis function is defined as

$$g_{ijk} = N x^i y^j z^k e^{-\alpha r^2}, \tag{2.21}$$

where i, j and k are non-negative integers, α is a positive orbital exponent and x, y and z are Cartesian coordinates centered on a nucleus. For $i + j + k = 0$ this results in a s-type Gaussian, for $i+j+k = 1$ in three p-type Gaussians, and so forth. The fact that the radial component of these functions varies with e^{-r^2} and not with the correct exponential factor e^{-r} is compensated by choosing a linear combination of these Gaussian functions for each orbital. Different sets of these linear combinations, the so called contracted Gaussian type functions (CGTF), are used to build up the atomic orbitals of the system. Sets of these CGTF for the atoms are called the basis sets of the calculation. The bases used to evaluate the electronic ground state energy of specific molecules in this work are introduced later in Chapter 3.

Perturbation Theory

The Hartree–Fock approximation has one serious shortcoming – even with an infinite number of single particle functions in the so called HF–limit a single determinant is not able to represent the electron density accurately. The main problem is that the electrons are allowed to approach closer to each other than the true quantum mechanical description of the correlated movement of the electrons would allow. To overcome this problem one naturally has to move away from the HF description toward a model using not one but several Slater determinants to describe the total molecular wavefunction. One possible approach is described in the theory by Møller and Plesset [75]. This approach includes the additional determinants via a perturbation scheme using the original spin orbitals generated by the HF method and is not variational itself.

The so called Møller–Plesset (MP) perturbation theory starts by dividing the Hamiltonian into a main, unperturbed part, and an additional perturbation operator. In this case the unperturbed operator is defined as the Hartree–Fock operator (the sum overall single particle Fock operators):

$$\hat{\mathbf{H}}^{(0)} = \hat{\mathbf{H}}^{HF} = \sum_i \hat{\mathbf{F}}(x_i) = \sum_i \left(\hat{\mathbf{h}}(x_i) + V^{HF}(x_i) \right), \tag{2.22}$$

while the perturbation operator is set to the difference between this operator and the exact

electronic Hamiltonian:

$$\hat{\mathbf{H}}^{(1)} = \hat{\mathbf{H}}_{\mathrm{el}} - \hat{\mathbf{H}}^{(0)} = \hat{\mathbf{V}}_{\mathrm{el-el}} - \sum_i V^{HF}(x_i). \tag{2.23}$$

As the Hartree–Fock potential already gives a quite good approximation of the true electron–electron interaction (typically around 99%), the energy correction generated by this perturbation operator is relatively small. The additional Slater determinants used to describe the molecular orbitals are generated by replacing one or two of the occupied single particle functions from the ground state HF determinant with unoccupied, so called virtual, orbitals. These determinants are called singly or doubly excited, as they represent molecular orbitals in a higher energy state.

An important property of this approximation is, that in first order it just reproduces the result of the Hartree–Fock calculations. To achieve any refinement one has to calculate the corrections in second and higher orders. The method is labeled according to the order of the perturbation treatment used. The MP2 method mostly used in this work therefore is the Møller–Plesset perturbation theory in second order.

2.1.3 Density Functional Theory

The density functional theory (DFT) offers a completely different approach to the calculation of molecular potentials. In contrast to the perturbation methods this theory is not based on the refinement of a result obtained via Hartree–Fock, but takes a different route to calculate the molecular energies [72, 76].

The DFT methods derive these from a charge density $\rho(r)$, which depends only on the coordinates x, y and z. The proof that this much simpler quantity indeed provides enough information to calculate the molecular energies was found by Hohenberg and Kohn and presented in their famous paper from 1964, which started the whole field of DFT methods [77]. In it the first Hohenberg–Kohn theorem is proven, which states that "*[...]the external potential* $V_{\mathrm{ext}}(r)$ (i.e. the complete molecular potential) *is (to within a constant) a unique functional of* $\rho(r)$*; since, in turn* $V_{\mathrm{ext}}(r)$ *fixes* $\hat{\mathbf{H}}$ *we see that the full many particle ground state is a unique functional of* $\rho(r)$*[...]*". Therefore, instead of using a Slater determinant of spin orbitals, by virtue of this theorem it is possible to calculate the total energy via the minimization of the charge density functional $E_{\mathrm{el}}[\rho_{\mathrm{el}}]$ which depends on the electron density

$$\rho_{\mathrm{el}}(r) = \sum_i |\psi_i(r)|^2 . \tag{2.24}$$

This is stated in the second Hohenberg–Kohn theorem, which proves that the energy obtained from a trial density $\tilde{\rho}$ represents an upper bound to the true ground state energy, as obtained from the exact ground state density ρ_0. The energy in the DFT approach is not given as the expectation value of an operator, like in the HF approach (Eq. (2.16)), but as a sum of energy functionals depending on the electron density:

$$E_{DFT}[\rho] = T(\rho) + V(\rho) + U(\rho) + E_{XC}(\rho), \tag{2.25}$$

where $T(\rho)$ describes the kinetic energy of the electrons, $V(\rho)$ the interaction with the nuclei, $U(\rho)$ the Coulomb repulsion between the electrons and $E_{XC}(\rho)$ the effects generated by the electron correlation which have no classical counterpart. This so called exchange-correlation energy is treated as a sort of "*junkyard*", as it is used to collect all parts of the energy which cannot be handled exactly. According to the second Hohenberg–Kohn theorem the total energy given by Eq. (2.25) obeys the relation:

$$E_0 \leq E_{DFT}[\rho] \tag{2.26}$$

where E_0 is the true ground state energy. The equality holds only, if the density inserted into Eq. (2.25) is the exact ground state density.

Similar to the HF equations (2.13) the Kohn–Sham approach leads to a set of one-electron equations, which have to be solved iteratively. The difference to Eq. (2.15) lies in the form of the effective potential, which now is of course no longer given by V^{HF}, but by an effective DFT potential defined by

$$V^{DFT}(r) = \int \frac{\rho(r')}{|r - r'|}\, dr' + V_{XC}(r), \tag{2.27}$$

where the first term is equivalent to the Coulomb–term of the HF equations, while V_{XC} is the potential due to the non-classical exchange-correlation energy E_{XC}. This is simply defined via the functional derivative of E_{XC}:

$$V_{XC} \equiv \frac{\delta E_{XC}}{\delta \rho}. \tag{2.28}$$

If the exact form of the exchange-correlation energy E_{XC} were known, the solution of the Kohn–Sham equation would generate the correct energy eigenvalue of the total Hamiltonian of the Schrödinger equation. So while the HF model started with the approximation that the total wavefunction can be described by a single Slater determinant, and therefore cannot result in an exact solution, the Kohn–Sham approach is in principle exact. Unfortunately the correct form of E_{XC} is not known, so the art of DFT calculations is to find good functional forms for this energy.

A commonly used used pair of functionals is Becke's 1988 exchange functional (B88 or B) [78] and the Lee–Yang–Parr (LYP) [79] correlation functional, or the so called Becke3LYP (B3LYP) hybrid functional, which combines the B88 and LYP functionals via three parameters (indicated by the 3) with three additional functionals. The parameters in these functionals are determined by fitting the results of the calculations for small molecular test systems to well established experimental molecular data.

As a reference the properties of DFT and HF based methods are summarized in Table 2.1. In computer based calculations DFT methods, for which the numerical effort is of the same order of magnitude as that of the bare HF calculation, will normally deliver results much faster than advanced methods based on HF theory, which require the calculation of

	HF	**DFT**
system definition	$\Psi_{\mathrm{el}} = \lvert\chi_1(x_1), \chi_2(x_2) \dots \chi_n(x_n)\rvert$	$\rho(r) = \sum_i \lvert\psi_i(r)\rvert^2$
variational condition	$\delta E[\Psi_{\mathrm{el}}] = \frac{\partial E}{\partial \Psi} = 0$	$\delta E[\rho] = \frac{\partial E}{\partial \rho} = 0$
energy	$E = \sum_i h_{ii} + \frac{1}{2}\sum_{i,j}(J_{ij} - K_{ij})$	$E = T(\rho) + V(\rho) + U(\rho) + E_{XC}(\rho)$
orbital equation	$\left(\hat{\mathbf{h}} + V^{HF}(x)\right)\chi_i(x) = \epsilon_i \chi_i(x)$	$\left(\hat{\mathbf{h}} + V_{XC}(x)\right)\chi_i(x) = \epsilon_i \chi_i(x)$

Table 2.1: Comparison of the main equations used for Hartree–Fock and density functional theory. [80]

additional determinants and their correlation via perturbation or variational methods. The drawbacks are that one cannot be certain that a given functional used for the exchange-correlation energy will produce good results with any given molecule, so that it is normally necessary to cross check the results with other data.

2.1.4 Quantum Chemistry of the H–Bond

The hydrogen bond examined in this work describes an attractive interaction between a proton donor $\mathbf{X} - \mathbf{H}$ and a proton acceptor $\mathbf{Y}$ on the same or a different molecule. (This work will only examine intramolecular hydrogen bonds.) The resulting complex will have the form

$$\mathbf{X} - \mathbf{H} \cdots \mathbf{Y}$$

where $\mathbf{X}$ and $\mathbf{Y}$ usually are electro-negative atoms like O, N or F. Depending on the relative strength of the $\mathbf{H}\cdots\mathbf{Y}$ bond, this will result in a shortening of the distance between $\mathbf{X}$ and $\mathbf{Y}$ and in a red–shift in the fundamental $\mathbf{X}$–$\mathbf{H}$ stretching vibration, caused by the lengthening of the $\mathbf{X}$–$\mathbf{H}$ bond. In some some hydrogen bonds between carbon and benzene complexes one can also observe an improper H–bond, which results in a blue shift of the fundamental $\mathbf{X}$–$\mathbf{H}$ stretching vibration, but this type of bond is less numerous than the red–shifting ones [81].

As a general rule, the treatment of this type of hydrogen bonded systems with quantum chemistry requires the use of large and flexible basis sets and the inclusion of electron correlation is especially important [10, 82]. This is mainly due to the fact, that the $\mathbf{H}\cdots\mathbf{Y}$ bond is relatively weak and describes a long–range interaction, therefore requiring a highly accurate description. Of the quantum chemical methods presented in the last sections, the pure HF method is unsuitable to deliver this accuracy, making it necessary to use at least MP2 to treat the correlation effects in the hydrogen bond. Within the DFT method, the B3LYP functional produces accurate results, in particular for the barrier heights [82]. For both methods a large enough basis set has to be used. The exact requirements of the basis are given in Section 3.2, where the *ab initio* calculations are presented.

2.2 The Cartesian Reaction Surface Hamiltonian

As was demonstrated in the previous section a molecular Hamiltonian consists of a term $\hat{\mathbf{T}}$ describing the kinetic energy of the nuclei, and of a term $\hat{\mathbf{V}}$, which describes the potential they are moving in. With the tools provided by quantum chemistry it is in principle possible to calculate the Hamiltonian responsible for the dynamics of an arbitrary large system. To do this one just has to generate the potential hypersurface given by Eq. (2.5) for all relevant configurations of the molecule. However a system consisting of N nuclei has $3N$ degrees of freedom (DOF), from which $3N-6$ represent internal movements in the molecule ($3N-5$ for linear molecules), while the remaining DOF describe the rotations and translations of the molecule as a whole. If each DOF is represented on a grid of only 10 points, which normally is far from sufficient, it becomes clear, that the complete description of a simple three body system already requires 1000 data-points to describe the three dimensional potential hypersurface. Any additional particle in the molecule adds three more degrees of freedom, increasing this number a thousand-fold. If one in addition bears in mind that a larger system also requires a bigger effort to calculate each single point via the *ab initio* methods it becomes quite clear that a complete *ab initio* quantum mechanical description with the computing power available today is at least unwieldy and in most cases simply not possible.

Due to the problems mentioned above, it is necessary to develop approximate Hamiltonians which require less effort to calculate, but still describe the system reasonably well. After giving some general properties of the potential energy surface, in the next sections two possible simplifications are sketched together with their limitations, before the Cartesian Reaction Surface (CRS) approach, which is used in this work for the larger systems, is described in more detail.

General Properties of the Potential

The potential energy hypersurface for the molecular dynamics is often calculated within the Born–Oppenheimer approximation (Section 2.1.1) and given as an adiabatic potential for each electronic state (Eq. (2.5)). This potential $V(\mathbf{R})$ is a function of all $3N$ nuclear coordinates. The gradient of this potential with respect to the nuclear coordinates gives the force F_{mol} acting on the atoms:

$$F_{\mathrm{mol}} = -\nabla_{\mathbf{R}} V(\mathbf{R}) = \{\partial V(\mathbf{R})/\partial R_1, \ldots, \partial V(\mathbf{R})/\partial R_{3N}\}. \tag{2.29}$$

If there is no force ($F_{\mathrm{mol}} = 0$) the molecular configuration is at a stationary point, i.e. in a local extremum or saddle point. For further information about the potential the calculation of the matrix of the second derivatives (the force constants or Hessian matrix) is necessary:

$$K_{mn} = \frac{\partial^2 V(\mathbf{R})}{\partial R_m \partial R_n} \quad (m, n = 1, \ldots, 3N). \tag{2.30}$$

After mass-weighting this matrix, it can be diagonalized at a stationary point, which results in $3N$ eigenvalues:

- 6 of them are zero, these correspond to the overall translations and rotations of the molecule
- Either all remaining $3N-6$ eigenvalues are positive, which means that the stationary point is a local minimum.
- Otherwise M eigenvalues are negative, the rest positive, which means that the stationary point is a saddle point of order M.

The eigenvectors resulting from the diagonalization describe a coordinate system in which, for small displacements, the movement of the nuclei is uncoupled, i.e. around a stationary point the molecule can be modeled as a collection of $3N-6$ non-interacting harmonic oscillators with frequencies given by the square root of the corresponding eigenvalues.

The harmonic normal modes given by the Hessian are valid for small displacements of all molecular DOF, i.e. with this information one is able to construct a potential surface in full dimensionality. This surface obviously is only a harmonic approximation to the real potential around the stationary point. Additionally, the evaluation of the second derivative matrix takes several orders of magnitude more effort in computer time than the simple calculation of the potential $V(\mathbf{R})$ or the forces F_{mol}. So even though this derivatives give valuable information about the full dimensional potential surface, it is not feasible to calculate this information for too many molecular geometries. Therefore it is necessary to reduce the dimensionality of the molecular potential to a manageable size.

Fixed Molecular Geometries

The easiest way to do this reduction is to keep several of the internal molecular degrees of freedom fixed. The potential hypersurface is then reduced to a number of dimensions equal to the number of molecular coordinates allowed to vary. This results in potentials which are quite easy to calculate if the number of non-fixed DOF is kept low enough. Another possibility is to move away from the internal molecular coordinates and define some new DOF describing, for instance, the movement of the center of mass of two subgroups within a larger molecule, which are otherwise kept fixed. This general approach is widely used to generate low dimensional potential surfaces of larger molecular systems, but it can run into problems quite easily. The main concerns are:

1. The main dynamics of the system has to be located in very few molecular DOF (i.e. the vibration of a single bond within a larger complex). Any large amplitude motion involving several DOF in the molecular coordinate system cannot be accurately represented without again producing too many dimensions for the potential.

2. After moving the variable molecular DOF away from their equilibrium position all the fixed DOF will in general no longer be in any kind of equilibrium as well. As they are fixed they will have no chance to relax to their new minimum energy. This leads to a potential surface energetically higher than the "real" one, so that energy barriers can block configurations, which could otherwise be easily reached.

3. In general the internal coordinates of the system are not Cartesian, even a selection of few DOF often leads to couplings between them located in the kinetic part of the Hamiltonian (see below).

From these points the 2nd one is a feature which sometimes can even be desirable. This happens when the dynamics of the DOF considered to be flexible is much faster than the expected dynamics of the rest of the molecule. In this case it is not expected that any redistribution of the energy to other DOF (so called Intramolecular Vibrational Redistribution (IVR)) takes place, and the artificial barriers created by this approach are realistic in the sense that the system has no time to reconfigure itself to find a path with lower energy. However, as soon as the time scale of the dynamics gets longer than the time expected for IVR this approach will fail to produce realistic results.

This method only requires the calculation of single point energies along the desired flexible coordinates. No forces or second derivatives are needed. This allows one to produce several hundreds of points without too much effort. To get a sufficiently dense distribution of grid points, this allows the inclusion of around three dimensions. The potential generated will contain no information at all about the remaining degrees of freedom.

Relaxed Molecular Surfaces

The relaxed surface approach is another possibility to include only few degrees of freedom of a complex system. For this potential the selected DOF are varied, and for each point the remaining DOF are relaxed to their minimum energy state. For this, the points 1 and 3 given for the fixed molecular geometries remain valid, while point 2 is turned around – now all DOF except for the ones kept flexible are always at their minimum energy configuration. Therefore this method will generate the lowest possible energy surface for the chosen flexible DOF. This approach is valid for slow dynamics in the large amplitude DOF, giving all the other coordinates time to relax. For fast motions the fixed molecule potential of the last section normally is more suited to describe the problem (see the note to point 2 in the last section).

This approach needs a optimization (i.e. a minimum search) of the molecular potential for each point. While this does only need the forces and not the second derivatives, it still takes much longer than only a single point calculation. Therefore the possible grid-size for this method is smaller than for the fixed geometry. On the other hand, the movement of the scaffold atoms gives some information about the DOF which are excluded from the calculation, i.e. one sees if the scaffold remains fixed or if there are still some large

amplitude motions left. This would point to strong couplings, which have to be taken into account for the calculations.

Minimum Energy Path Method

Another approach is the minimum energy path method, which is a standard technique from quantum chemistry [67, 83]. Here two stable configurations of a molecule are connected along a single reaction coordinate, following the valley of minimal energy between those two. This path is generated by first localizing the stationary points on the potential surface. The path then starts from the transition state separating the reactant and the product and follows the path of steepest descent in mass-weighted Cartesian coordinates toward the local energy minima which is separated by this point. This means following a gradient vector with the the components

$$G_i \propto \frac{1}{\sqrt{M_i}} \frac{\partial V(\mathbf{R})}{\partial R_i}, \tag{2.31}$$

which at the transition state will coincide with the normal mode with the highest negative eigenvalue.

In contrast to the approach with fixed molecular geometry here all molecular coordinates are allowed to move and relax to a minimum energy configuration with regard to the position on the reaction path, which is always one dimensional. This path cannot be described by a single coordinate and may involve motions in all the available dimensions. The one dimensional character of the path allows an easy description of the potential, as only the energies along a one dimensional pathway on the multidimensional hypersurface are required. The difficulty, however, is now to find an expression for the kinetic energy operator. The reaction coordinate represents the collective, coupled motion of several atoms in the molecule, and therefore it will introduce kinetic couplings between them, which are quite difficult to treat. They have to be calculated via the transformation from $x \rightarrow q$ of the Laplacian from Cartesian into an arbitrary set of coordinates, as given by

$$\begin{aligned} \Delta &= \frac{1}{\sqrt{g}} \partial_\mu g^{\mu\nu} \sqrt{g} \partial_\nu \\ g &= \det(g_{\mu\nu}) \end{aligned}$$

with the metric tensor defined as

$$\begin{aligned} g_{\mu\nu} &= \sum_i \frac{\partial x_i}{\partial q^\mu} \frac{\partial x_i}{\partial q^\nu} \\ g^{\mu\nu} &= \sum_i \frac{\partial q^\mu}{\partial x_i} \frac{\partial q^\nu}{\partial x_i}. \end{aligned}$$

For a curved reaction path, this transformation will introduce couplings between different coordinates in the kinetic part of the Hamiltonian. As it is a large amplitude motion, it

will also not be possible to use something like the normal mode transformation to uncouple the equations.

The numerical effort for this method is moderate. While it requires an optimization of the molecular geometry at each potential point, the potential is by definition only one dimensional. Therefore one can generate it relatively easily. The main problem of this approach are the kinetic couplings it introduces.

An additional point has to be mentioned. The reaction path approximation requires exactly opposite assumptions about the system than those needed for the fixed coordinates approach (and is in this similar to the relaxed molecular surface): While there it is necessary that the dynamics of the system is fast enough, so that the frozen modes remain unimportant on the relevant time scale, for the reaction path method it is necessary to make certain that the movement along the reaction path is slow enough to allow for relaxation of all other degrees of freedom. Another reason to look for slow motions is, that the kinetic (or adiabatic) couplings between the different coordinates depend on this velocity and become more prominent for higher speeds. Therefore this description is not well suited for the treatment of fast proton transfer reactions. Due to the low weight of the proton, compared to the rest of the molecular scaffold, the reaction is taking place at high velocities, and along a strongly curved reaction path. The proton has not enough momentum to "force" its way along a straight path. As soon, as it hits a potential barrier, it will change direction, rather than move aside the heavier, intervening particles. This behaviour is typical for the heavy–light–heavy reaction the proton transfer represents.

Reaction Surface Method

An extension of the minimum energy path approach described above is the reaction surface method [84]. In addition to the one dimensional reaction path harmonic motions orthogonal to it are now accounted for as well. This leads to a description based on the coordinate s (the length along the minimum energy path) and $3N-7$ coordinates forming a harmonic "valley" around this path. So, if $\mathbf{R}^s$ is a point on the reaction path and $\Delta\mathbf{R} = \mathbf{R} - \mathbf{R}^s$ is a small displacement from it, one can do a local Taylor expansion of the potential around $\mathbf{R}^s$ as

$$\begin{aligned} V(\mathbf{R}) &= V(\mathbf{R}^s) + \nabla V(\mathbf{R})|_{\mathbf{R}=\mathbf{R}^s} \cdot \Delta\mathbf{R} \\ &+ \frac{1}{2}\Delta\mathbf{R} \cdot \mathbf{K}^s \cdot \Delta\mathbf{R}, \end{aligned} \tag{2.32}$$

with $\mathbf{K}^s$ the force constant matrix at $\mathbf{R} = \mathbf{R}^s$. To generate a set of normal modes orthogonal to the reaction path, which will eliminate the linear term from Eq. (2.32), the matrix $\mathbf{K}$ has to be diagonalized at every point of the path. As the normal modes corresponding to the rotations and translations of the molecule will not emerge automatically if the Hessian is diagonalized at an arbitrary point of the potential, as it would happen at a stationary point, it is necessary to project out the overall motions of the molecule together with the

movement along the reaction path. This projected force constant matrix $\mathbf{K}^P$ is given as

$$\mathbf{K}^P = (\mathbf{1} - \mathbf{P}) \cdot \mathbf{K} \cdot (\mathbf{1} - \mathbf{P}) \tag{2.33}$$

with the projection matrix $\mathbf{P}$ consisting of the unit vectors for the infinitesimal overall rotations and translations and the unit vector along the reaction path. This projected matrix now has 7 zero eigenvalues and normal modes oriented orthogonal to the reaction path. So, if q_i describes a displacement along the i^{th} normal mode, with an eigenvalue of ω_i^2, then the potential energy surface can be written as

$$V(s, q_1, \ldots q_{3N-7}) = V(s) + \frac{1}{2} \sum_{\xi=1}^{3N-7} \omega_\xi^2(s) q_\xi^2. \tag{2.34}$$

This means, that in addition to the energies of the reaction path, the force constants along this one dimensional pathway on the multidimensional hypersurface are required. The difficulty of finding an expression for the kinetic energy operator remains. While the harmonic normal modes at a single point of the reaction path are by definition all orthogonal to each other and can be represented by uncoupled kinetic terms, the reaction coordinate s is a collective motion and will also introduce couplings between the harmonic normal modes. This is due to the fact that now the normal mode coordinates will depend on s (the reaction path coordinate), and that the kinetic energy operator for this system will contain couplings between s and $q_\xi(s)$ [67, 84]. An example for a reaction path in a double minimum potential is shown in Fig. 2.2 as the curved, solid arrow.

For this approach it is not necessary to assume a slow motion along the reaction path to allow for a complete relaxation of all other DOF around it, as the system is allowed to deviate along the normal modes surrounding this path. But as the potential surface around the true minimum energy path is only included in a harmonic approximation too strong deviations from it will lead to motions on parts of the potential, which are not approximated very well. To increase the accuracy, it would be necessary to include the reaction valley with anharmonic corrections to the harmonic normal modes. This would increase the numerical effort drastically and is limited to very small molecules

Both this reaction surface method, and the minimum energy path method in the previous section can only model large amplitude motions along a one–dimensional path on the multidimensional PES of a larger molecule. If the system would allow different possible pathways, and the dynamics requires the wavepacket to split, taking different, branching directions, these methods are not suited to model it.

Another approach to the reaction surface can be made by not choosing the minimum energy path as a reference, but just taking a set of internal molecular coordinates, calculating a potential surface for them and then adding the remaining DOF in the harmonic approximation as explained above [68, 84, 85]. In more recent works, the reaction path method has been used to estimate the tunneling splitting using semiclassical methods and an "on the fly" calculation of the potential energy surface for different semicalssical pathways [86]. A recent application of the reaction surface method, which generates a potential

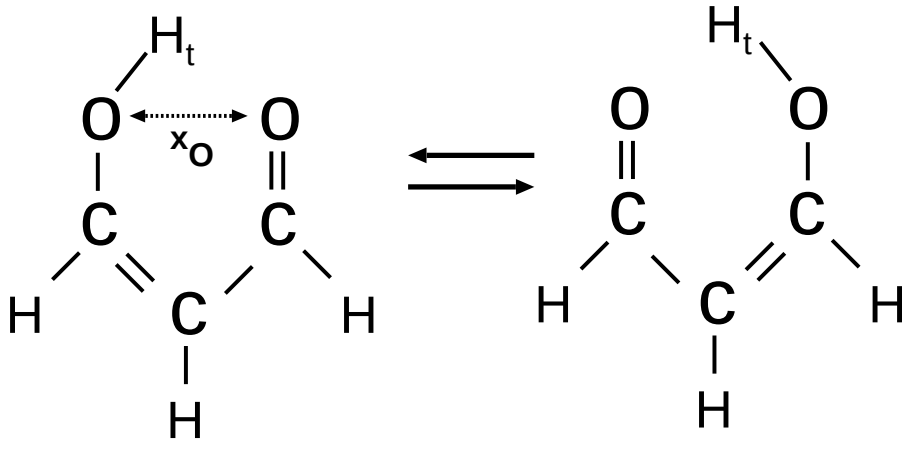

Figure 2.1: In malonaldehyde the hydrogen H_t can be localized at either of the two oxygens. The transfer process is strongly influenced even by slight variations of the distance x_{O} between the oxygen atoms, caused for instance by a low frequency vibration mode of the molecular scaffold.

energy surface in full dimensionality is presented in [87]. A more extensive overview over the different approaches to the reaction path Hamiltonian can be found, e.g., in [88, 89].

The Cartesian Reaction Surface (CRS) Approach

To describe the dynamics of a proton coupled to the heavy atoms in a molecule one needs some aspects of both approaches described above. The movement of the proton will generally be on a much faster time scale than the dynamics of the much heavier molecular scaffold. This allows the assumption that these heavy scaffold atoms will remain close to their equilibrium positions during the transfer process. On the other hand, proton transfer reactions are often strongly influenced by slight movement of the heavy fragments involved in the transfer. For instance, the potential for the hydrogen in the classical example of malonaldehyde shown in Fig. 2.1, changes drastically even for slight motions of the oxygen atoms toward or away from each other. For such systems, a treatment with the Cartesian Reaction Surface (CRS) Hamiltonian approach is well suited [90, 91]. This concept produces a model Hamiltonian which has no kinetic couplings, because all the interactions are included in the potential energy operator. The starting point for this model again is the potential with all $3N$ spatial dimensions. Among these, those Cartesian coordinates are chosen, which will undergo arbitrarily large displacements during the reaction. For a proton transfer reaction these are the three DOF of the hydrogen atom. These N_{rc} DOF are the active or reaction coordinates. The remaining $3N - N_{\mathrm{rc}}$ DOF are the spectator or substrate coordinates. During the reaction the substrate atoms are considered to remain close to their equilibrium positions. This makes it possible to express

the potential for their movement with a Taylor expansion to second order. Hence the total potential energy surface can be approximated as:

$$\begin{aligned} V(\mathbf{R}) &\approx V(\mathbf{x}, \mathbf{Z}^{(0)}(\mathbf{x})) + \left(\frac{\partial V(\mathbf{x}, \mathbf{Z})}{\partial \mathbf{Z}}\right)_{\mathbf{Z}=\mathbf{Z}^{(0)}(\mathbf{x})} \cdot (\mathbf{Z} - \mathbf{Z}^{(0)}(\mathbf{x})) \\ &+ \frac{1}{2}(\mathbf{Z} - \mathbf{Z}^{(0)}(\mathbf{x})) \cdot \left(\frac{\partial^2 V(\mathbf{x}, \mathbf{Z})}{\partial \mathbf{Z} \partial \mathbf{Z}}\right)_{\mathbf{Z}=\mathbf{Z}^{(0)}(\mathbf{x})} \cdot (\mathbf{Z} - \mathbf{Z}^{(0)}(\mathbf{x})), \end{aligned} \tag{2.35}$$

where $\mathbf{R}$ is the $3N$ dimensional coordinate vector, $\mathbf{x}$ is the N_{rc} dimensional vector describing the reaction coordinates, $\mathbf{Z}$ contains the remaining $3N - N_{\mathrm{rc}}$ scaffold DOF and $\mathbf{Z}^{(0)}$ describes the scaffold geometry in equilibrium configuration. If the reactant and product geometries for the scaffold are similar, it is sufficient to keep $\mathbf{Z}^{(0)}(\mathbf{x})$ fixed at $\mathbf{x} = \mathbf{x}_{\mathrm{eq}}$ for the whole potential. The more general approach of flexible reference, which depend on the reaction coordinates, is described in [90]. With this it is possible to treat molecular systems, where the scaffold moves significantly during the proton transfer. Normally, the harmonic approximation for the heavy atom modes of the scaffold would break down in these cases, when their movement is only treated in reference to a single configuration.

In this work, only fixed reference CRS will be treated. This limitation to a single reference is sufficient, if the whole dynamics of the molecule happens in the hydrogen bond, while the scaffold remains nearly stationary. In Fig. 2.2 the straight arrow gives an example for a Cartesian reaction path.

The three terms in Eq. (2.35) are:

1. $V(\mathbf{x}, \mathbf{Z}^{(0)}(\mathbf{x}))$ is the potential energy along the Cartesian N_{rc} dimensional reaction path, calculated with the scaffold modes frozen in some reference geometry.

2. The second term describes the forces acting on the atoms of the molecular scaffold. These are caused by the reaction coordinates being moved away from their equilibrium, $\mathbf{f}(\mathbf{x}) = -(\partial V(\mathbf{x}, \mathbf{Z})/\partial \mathbf{Z})_{\mathbf{Z}=\mathbf{Z}^{(0)}(\mathbf{x})}$

3. The third term describes the changes in the Hessian matrix generated by the motion along the reaction coordinate $\mathbf{x}$, $\mathbf{K}(\mathbf{x}) = -(\partial^2 V(\mathbf{x}, \mathbf{Z})/\partial \mathbf{Z} \partial \mathbf{Z})_{\mathbf{Z}=\mathbf{Z}^{(0)}(\mathbf{x})}$. This will generate changes in the vibrational frequencies of the molecule and induce couplings between the different normal modes.

The Hessian matrix $\mathbf{K}(\mathbf{x})$ for the system has to be diagonalized, to transform the molecule to the normal mode picture. For the CRS method this is done only for the molecular scaffold, so the resulting modes will exclude all DOF which are part of the reaction coordinates $\mathbf{x}$. Before the reduced force constant matrix is diagonalized, the infinitesimal rotations and translations of the molecular scaffold have to be eliminated. This is done for the stationary state of the molecule describing the reactant configuration, i.e. $\mathbf{x} = \mathbf{x}_0$ and $\mathbf{Z} = \mathbf{Z}^{(0)}$. At this point the matrix $\mathbf{K}_0$ is defined as the $3(N-1)$ dimensional

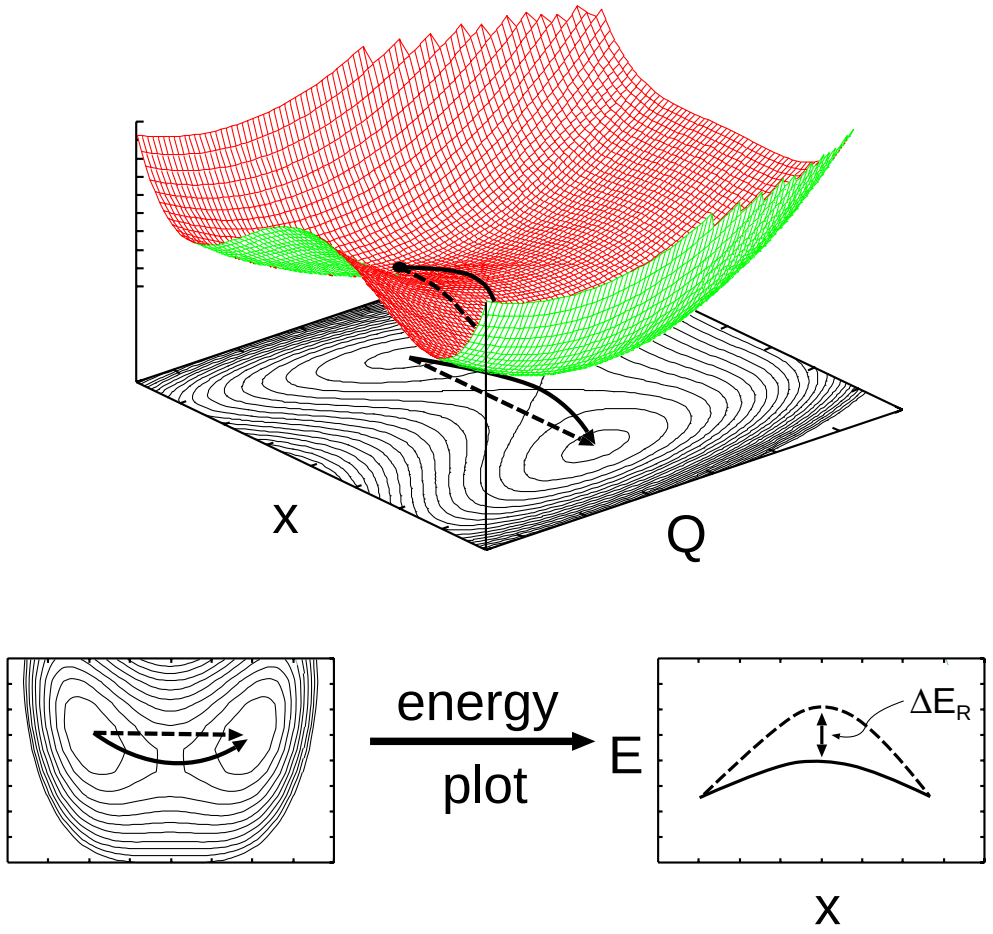

Figure 2.2: The upper panel shows a schematic view of a 2D potential surface describing a 1D movement along x coupled to a normal coordinate Q. A minimum energy reaction path would follow the solid arrow, which entails movement along both x and Q, and therefore leads to a coupled kinetic energy operator. The CRS method would follow the straight path along the dashed arrow and compensates for the higher energy away from the true transition state by including the energy of a shifted normal mode oscillator along Q. The lower two panels show a top view of this potential (left), and a plot of the potential energy along the two different rection pathways. ΔE_R is the reorganization energy, which is present in the system when follwing the CRS path, as compared to the minimum energy path.

force constant matrix for the molecular scaffold. The matrix $\mathbf{U}$ contains the $3(N-1)-6$ normal modes of

$$(1-\mathbf{P})\,\mathbf{m}^{-1/2}\mathbf{K}_0\,\mathbf{m}^{-1/2}(1-\mathbf{P}),$$

i.e. the mass weighted Hessian, from which $\mathbf{P}$ has projected out the rotations and translations. This transformation defines the mass weighted normal modes $\mathbf{q}$.

The Cartesian coordinates of the molecular scaffold are connected to these normal modes via the relation

$$\mathbf{Z}-\mathbf{Z}^{(0)} = \mathbf{m}^{-1/2}\mathbf{U}\,\mathbf{q} \tag{2.36}$$

The transformation matrix $\mathbf{U}$ is then used as the reference transformation for all other configurations with $\mathbf{x} \neq \mathbf{x}_0$ as well. The resulting second derivative matrix will therefore only be diagonal at this stationary point, as it is dependent on $\mathbf{x}$. When the system moves away from this configuration there will be off-diagonal elements generating a coupling between the normal modes of the stationary point. Taking all this into account and writing the projected Hessian in the normal mode basis as $\mathbf{K}^{\text{eff}}$, the complete molecular Cartesian Reaction Hamiltonian can be written as:

$$\begin{aligned}
\hat{\mathbf{H}}_{\text{mol}} &= \sum_{n=1}^{N_{\text{rc}}} \hat{\mathbf{T}}_n + \hat{\mathbf{V}}_{\text{ref}}(\mathbf{x}) \\
&\quad + \sum_{k=1}^{N_{\text{sub}}} \left(\hat{\mathbf{t}}_k + \frac{1}{2}\omega_k^2(\mathbf{x})q_k^2\right) \\
&\quad - \sum_{k=1}^{N_{\text{sub}}} \left[\left(f_k^{\text{eff}}(\mathbf{x}) - \frac{1}{2}\sum_{l\neq k}^{N_{\text{sub}}} K_{kl}^{\text{eff}}(\mathbf{x})q_l\right) q_k\right] \\
&= \sum_{n=1}^{N_{\text{rc}}} \hat{\mathbf{T}}_n + \hat{\mathbf{V}}_{\text{ref}}(\mathbf{x}) + \sum_{k=1}^{N_{\text{sub}}} \hat{\mathbf{t}}_k \\
&\quad + \frac{1}{2}\,\mathbf{q}\,\mathbf{K}^{\text{eff}}(\mathbf{x})\,\mathbf{q} - \mathbf{f}^{\text{eff}}(\mathbf{x})\,\mathbf{q}\,.
\end{aligned} \tag{2.37}$$

Here $\hat{\mathbf{V}}_{\text{ref}}(\mathbf{x}) = V(\mathbf{x}, \mathbf{Z}^{(0)})$ is the potential for the reaction modes, calculated for the fixed reference geometry, $\hat{\mathbf{T}}_n = P_n^2/2M_n$ is the kinetic energy for these modes, $N_{\text{sub}} = 3N - N_{\text{rc}} - 6$ is the number of remaining substrate normal modes, $\hat{\mathbf{t}}_k = P_k^2/2\mu_0$ is the kinetic energy of the substrate, q_k describes the displacements along the substrate normal modes, $\omega_k^2(\mathbf{x}) = \mathbf{K}_{kk}^{\text{eff}}(\mathbf{x})$ describes an approximate normal mode frequency (only approximate due to the coupling terms) and $f_k^{\text{eff}}(\mathbf{x}) = -\nabla_Z V(\mathbf{x}, \mathbf{Z}^{(0)})\mathbf{m}^{-1/2}\mathbf{U}$ is the force acting on the normal mode (generated by transforming the Cartesian forces into the mass weighted normal mode basis). The mass μ_0 describes the effective mass of a normal mode. Due to the mass weighting of the coordinates, all modes will have an equal mass, depending on the weighting factor used. While normally the atomic mass unit (amu) is used for this, in this work the electron mass (m_e) is employed, which is the atomic unit of mass. The first two terms in Eq. (2.37) describe the complete system of reaction modes, resulting

in a Hamiltonian $\hat{\mathbf{H}}_{\mathrm{sys}} = \sum_{n=1}^{N_{\mathrm{rc}}} \hat{\mathbf{T}}_n + \hat{\mathbf{V}}_{\mathrm{ref}}(\mathbf{x})$ for a molecular system with a fixed scaffold geometry (see the previous Section, page 19). The other terms give the additional kinetic energy of the normal modes and an additional CRS potential V_{CRS}.

In a further simplification step it is often assumed, that the frequencies ω_k of the normal modes are independent of the system coordinate. Additionally, if the off diagonal parts of the force constant matrix $\mathbf{K}^{\mathrm{eff}}$ are small it is possible to neglect these couplings. So in general, one can define normal mode displacements as $\mathbf{q}_0 = -\mathbf{f}^{\mathrm{eff}}(\mathbf{x})[\mathbf{K}^{\mathrm{eff}}(\mathbf{x})]^{-1}$. This leads to a representation of the Hamiltonian in the form

$$\begin{aligned}
\hat{\mathbf{H}}_{\mathrm{mol}} &= \sum_{n=1}^{N_{\mathrm{rc}}} \hat{\mathbf{T}}_n + \hat{\mathbf{V}}_{\mathrm{ref}}(\mathbf{x}) + \sum_{k=1}^{N_{\mathrm{sub}}} \hat{\mathbf{t}}_k \\
&\quad + \frac{1}{2}(\mathbf{q}-\mathbf{q}_0)\mathbf{K}^{\mathrm{eff}}(\mathbf{q}-\mathbf{q}_0) - E_{\mathrm{R}}(\mathbf{x}) \\
&= \sum_{n=1}^{N_{\mathrm{rc}}} \hat{\mathbf{T}}_n + \hat{\mathbf{V}}_{\mathrm{ref}}(\mathbf{x}) + \sum_{k=1}^{N_{\mathrm{sub}}} \hat{\mathbf{t}}_k \\
&\quad + \frac{1}{2}\sum_{l=1}^{N_{\mathrm{sub}}} \left[\omega_l^2 \left(q_l - f_l^{\mathrm{eff}}(\mathbf{x})/\omega_l^2\right)^2\right] - E_{\mathrm{R}}(\mathbf{x}),
\end{aligned} \tag{2.38}$$

where the reorganization energy is

$$E_{\mathrm{R}}(\mathbf{x}) = \frac{1}{2}\mathbf{q}_0\mathbf{K}^{\mathrm{eff}}(\mathbf{x})\mathbf{q}_0. \tag{2.39}$$

With a diagonal matrix $\mathbf{K}^{\mathrm{eff}}$ and normal mode frequencies independent of $\mathbf{x}$ this simplifies to

$$E_{\mathrm{R}}(\mathbf{x}) = \sum_{n=1}^{N_{\mathrm{sub}}} \frac{f_n^2(\mathbf{x})}{2\omega_n^2}. \tag{2.40}$$

This represents the energy which is required to compensate for the force exerted by the active degrees of freedom on the reservoir oscillators. The lower panel of Fig. 2.2 shows a graphic representation of this reorganization energy. Equation (2.38) represents a system coupled to a set of shifted harmonic oscillators. The equilibrium positions of these oscillators depend parametrically on the state of the system coordinate, which means that all couplings in this Hamiltonian are located in the potential terms. The kinetic energy can be represented by uncoupled kinetic energy operators in Cartesian coordinates for the oscillations in the normal modes.

2.3 Quantum Dynamics

After generating a suitable Hamiltonian for the molecular system under investigation the goal is to examine the dynamics of quantum mechanical nuclear wavepackets described by this operator. This dynamics can be separated into two main categories: Closed systems, in

which the energy is conserved and open systems, which can loose energy to the surrounding environment [57, 61, 92]. While the first kind can, depending on the initial conditions, be described by normal wavefunctions or a density matrix, the density matrix formalism is required for an open system. If the description with wavefunctions is sufficient, the dynamics can be calculated with the help of the time dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle = \hat{\mathbf{H}}\left|\Psi(t)\right\rangle \tag{2.41}$$

and the equivalent equation for the adjoint state vector

$$-i\hbar\frac{\partial}{\partial t}\left\langle\Psi(t)\right| = \left\langle\Psi(t)\right|\hat{\mathbf{H}} \tag{2.42}$$

where $\hat{\mathbf{H}} = \hat{\mathbf{T}} + \hat{\mathbf{V}}$ is the Hamilton operator describing the total energy of the system. Usually this operator will contain a time dependence in the potential term, $\hat{\mathbf{V}} = V(t)$. In the next section the time dependent Hamilton operator is discussed.

The Time Evolution Operator

To develop the formalism further, one has to take a closer look at the Schrödinger equation (2.41) and its counterpart for the adjoint state (2.42).

For a time independent $\hat{\mathbf{H}}$ one can solve the stationary Schrödinger equation for this operator, resulting in a set of eigenstates $|\varphi_n\rangle$ with energies E_n:

$$\hat{\mathbf{H}}\left|\varphi_n\right\rangle = E_n\left|\varphi_n\right\rangle \tag{2.43}$$

A system starting out in one of these eigenstates at $t = 0$ will then evolve according to Eq. (2.41), which results in the state

$$\left|\varphi_n(t)\right\rangle = e^{-(i/\hbar)E_n t}\left|\varphi_n\right\rangle, \tag{2.44}$$

which is obviously a solution of Eq. (2.41). In a more abstract form, this can be written as

$$e^{-(i/\hbar)E_n t}\left|\varphi_n\right\rangle = e^{-(i/\hbar)\hat{\mathbf{H}}t}\left|\varphi_n\right\rangle, \tag{2.45}$$

where the exponential operator function can be evaluated via the series expansion of the exponential. In addition to that it is known, that all solutions of the Schrödinger equation can be represented as a linear combination of the eigenstates given above:

$$\left|\psi\right\rangle = \sum_n C_n\left|\varphi_n\right\rangle \tag{2.46}$$

Taking this together with Eq. (2.45), it follows that a state vector evolving dynamically under the influence of $\hat{\mathbf{H}}$ can be written as

$$\left|\psi(t)\right\rangle = e^{-(i/\hbar)\hat{\mathbf{H}}t}\left|\psi(0)\right\rangle. \tag{2.47}$$

As can be seen from this, the wavefunction at time t is generated by applying an operator to the initial wavefunction at time $t = 0$. This operator is usually referred to as the time evolution operator $\hat{\mathbf{U}}(t, t_0)$, which is the operator evolving the state at time t_0 into the state at time t. In the case of the time independent $\hat{\mathbf{H}}$ shown above this is simply

$$\hat{\mathbf{U}}(t, t_0) = e^{-(i/\hbar)\hat{\mathbf{H}}(t-t_0)}, \tag{2.48}$$

as obtained from formally integrating the time dependent Schrödinger equation (2.41). So a wavepacket evolving in time can formally be written as

$$|\psi(t)\rangle = \hat{\mathbf{U}}(t, 0)\,|\psi(0)\rangle\,. \tag{2.49}$$

Substituting this into the Schrödinger equation leads to

$$i\hbar\frac{\partial}{\partial t}\hat{\mathbf{U}}(t, 0)\,|\psi(0)\rangle = \hat{\mathbf{H}}(t)\hat{\mathbf{U}}(t, 0)\,|\psi(0)\rangle \tag{2.50}$$

and in turn, as this holds for all states $|\psi(0)\rangle$ and all initial times t_0, to the operator equation

$$i\hbar\frac{\partial}{\partial t}\hat{\mathbf{U}}(t, t_0) = \hat{\mathbf{H}}(t)\hat{\mathbf{U}}(t, t_0) \tag{2.51}$$

with the initial condition

$$\hat{\mathbf{U}}(t_0, t_0) = \mathbf{1} \tag{2.52}$$

(at time zero the operator should transform $|\psi(0)\rangle$ into $|\psi(0)\rangle$). Repeating the same formalism, but starting from the adjoint Schrödinger equation will result in the equivalent equation for the adjoint operator

$$-i\hbar\frac{\partial}{\partial t}\hat{\mathbf{U}}^{\dagger}(t, t_0) = \hat{\mathbf{U}}^{\dagger}(t, t_0)\hat{\mathbf{H}}(t) \tag{2.53}$$

So the time evolution of a quantum mechanical state $|\psi(t)\rangle$ can be calculated either by solving the time dependent Schrödinger equation or by determining the necessary time evolution operator via the formal integration of Eq. (2.51).

A time dependent Hamiltonian will lead to a formal solution for this operator in the form

$$\hat{\mathbf{U}}(t, t_0) = \mathbf{1} - \int_{t_0}^{t} d\tau\,\hat{\mathbf{H}}(\tau)\hat{\mathbf{U}}(\tau, t_0), \tag{2.54}$$

which takes into account the initial condition (2.52). This has to be solved iteratively by inserting it into itself, as the solution for time t requires the result from time $t - dt$. With the limit for small time steps

$$\hat{\mathbf{U}}(t + dt, t) = \mathbf{1} - \frac{i}{\hbar}\hat{\mathbf{H}}(t)dt \tag{2.55}$$

one obtains after iterating:

$$\hat{\mathbf{U}}(t,t_0) = \mathbf{1} + \sum_{n=1}^{\infty} \left(\frac{i}{\hbar}\right)^n \int_{t_0}^{t} d\tau_n \int_{t_0}^{\tau_n} d\tau_{n-1} \ldots \int_{t_0}^{\tau_2} d\tau_1 \, \hat{\mathbf{H}}(\tau_n)\hat{\mathbf{H}}(\tau_{n-1}) \ldots \hat{\mathbf{H}}(\tau_1). \tag{2.56}$$

It has to be noted that due to the fact that $\hat{\mathbf{H}}$ is an operator, which does not commute with itself at different times (i.e. $[\hat{\mathbf{H}}(\tau_n), \hat{\mathbf{H}}(\tau_{n-1})] \neq 0$), the product in the integral cannot be reordered. Therefore the time variables have to be fully ordered with $t \leq \tau_n \leq \ldots \leq \tau_1 \leq t_0$. Without this it would have been possible to write the sum as an exponential, so a common convention is to call this expansion a fully time ordered exponential and write it in the abbreviated notation

$$\hat{\mathbf{U}}(t,t_0) = \exp_+ \left[-\frac{i}{\hbar}\int_{t_0}^{t} d\tau \, \hat{\mathbf{H}}(\tau)\right]. \tag{2.57}$$

This equation is, of course, much more difficult to solve than a simple exponential function, but the system formally will still just be propagated by the application of $\hat{\mathbf{U}}(t_0,t_1)$ (Eq. (2.54)). The main difference is that due to the time dependence of $\hat{\mathbf{H}}(t)$ it becomes now important to specify the exact time interval which is spanned by the operator, as, due to the time dependence of $\hat{\mathbf{H}}$ it is no longer just the time difference which is important.

2.3.1 General Ensemble Theory

If a system cannot be described with a single quantum mechanical wavefunction, e.g. a thermal ensemble consisting of several particles, it is necessary to add a level of classical statistics to the quantum description. This leads to the representation of the system in the density matrix picture [57, 93, 94]. If there are n states present, each prepared independently from each other to occur with a probability of W_n, then the statistical operator for this total state is given by

$$\hat{\rho} = \sum_n W_n \left|\phi_n\right\rangle \left\langle\phi_n\right|, \tag{2.58}$$

which is called the statistical operator. Note that this indeed is an operator, while the object referred to as density matrix is in fact the representation of this operator in an arbitrary basis set. So if the functions $|\varphi_a\rangle$ are elements of an orthonormal basis set then the elements of this matrix are given by

$$\rho_{ab} = \left\langle\varphi_a\right| \hat{\rho} \left|\varphi_b\right\rangle. \tag{2.59}$$

The expectation value of an operator for a mixture of pure states is then calculated by taking the weighted average of the expectation values for each single state,

$$\left\langle \hat{\mathbf{O}} \right\rangle = \sum_n W_n \left\langle\phi_n\right| \hat{\mathbf{O}} \left|\phi_n\right\rangle = \mathrm{tr}\left\{\hat{\mathbf{O}}\hat{\rho}\right\}. \tag{2.60}$$

	Wavefunction	**Density Operator**
pure state	$\|\Psi\rangle = \sum_i c_i \|\phi_i\rangle$	$\hat{\rho} = \|\Psi\rangle\langle\Psi\|$
mixed state	–	$\hat{\rho} = \sum_n W_n \|\Psi_n\rangle\langle\Psi_n\|$
expectation value	$\left\langle \hat{\mathbf{O}} \right\rangle = \langle\Psi\| \hat{\mathbf{O}} \|\Psi\rangle$	$\left\langle \hat{\mathbf{O}} \right\rangle = \mathrm{tr}\{\hat{\mathbf{O}}\hat{\rho}\}$
equation of motion	$i\hbar\frac{\partial}{\partial t}\|\Psi\rangle = \hat{\mathbf{H}}\|\Psi\rangle$	$i\hbar\frac{\partial}{\partial t}\hat{\rho} = \left[\hat{\mathbf{H}}, \hat{\rho}\right]$

Table 2.2: Comparison of the wavefunction picture with the description via a density operator.

Here the "tr" means taking the trace, i.e. summing over all diagonal elements of the matrix representation of the operators in a orthonormal basis. The normalization condition for the density operator is given by $\mathrm{tr}\{\hat{\rho}\} = 1$.

If the system is closed, i.e. its energy is conserved, then there will be no transfer of probability between the states, i.e. the W_n in Eq. (2.58) are time independent. With this information one can derive the Liouville–von Neumann equation, which describes the time evolution of the statistical operator. Starting with Eq. (2.58) for the density operator of a statistical mixture of states which varies according to Eq. (2.49) one gets

$$\begin{aligned} \hat{\rho}(t) &= \sum_n W_n \hat{\mathbf{U}}(t) \|\phi(0)_n\rangle\langle\phi(0)_n\| \hat{\mathbf{U}}^\dagger(t) \\ &= \hat{\mathbf{U}}(t)\hat{\rho}(0)\hat{\mathbf{U}}^\dagger(t) \end{aligned} \tag{2.61}$$

Using Eq. (2.51), Eq. (2.53) and Eq. (2.61) the equation of motion for the density operator is obtained as

$$i\hbar\frac{\partial}{\partial t}\hat{\rho}(t) = \hat{\mathbf{H}}(t)\hat{\mathbf{U}}(t)\hat{\rho}(0)\hat{\mathbf{U}}^\dagger(t) - \hat{\mathbf{U}}(t)\hat{\rho}(0)\hat{\mathbf{U}}^\dagger(t)\hat{\mathbf{H}}(t) = \left[\hat{\mathbf{H}}, \hat{\rho}(t)\right], \tag{2.62}$$

with the commutator $[\hat{\mathbf{H}}, \hat{\rho}(t)] = \hat{\mathbf{H}}\hat{\rho}(t) - \hat{\rho}(t)\hat{\mathbf{H}}$. Eq. (2.62) is the Liouville–von Neumann equation.

2.3.2 The Interaction Picture

After having discussed the basic concepts necessary to describe the dynamics of quantum systems it is time to take a closer look at the approximate determination of the time evolution operator for the time dependent Hamiltonian. Up to now the evolution of the quantum mechanical states was treated in the so called Schrödinger picture. Now a different picture, the interaction picture, is presented [57, 61].

While in general an exact solution of the operator equation (2.51) is not possible, one can obtain a good approximation using the methods of perturbation theory. It is assumed,

that the total Hamiltonian of the system can be split into

$$\hat{\mathbf{H}}(t) = \hat{\mathbf{T}} + \hat{\mathbf{V}}_0 + \hat{\mathbf{V}}(t) \tag{2.63}$$

where $\hat{\mathbf{V}}(t)$ is considered as a small perturbation compared to the time independent part $\hat{\mathbf{H}}_0 = \hat{\mathbf{T}} + \hat{\mathbf{V}}_0$. In this case most of the time dependence of the state vector $|\psi(t)\rangle$ will be due to the action of $\hat{\mathbf{H}}_0$. Then it is possible to define a state vector in the interaction picture as

$$|\psi_I(t)\rangle = e^{(i/\hbar)\hat{\mathbf{H}}_0 t} |\psi(t)\rangle \tag{2.64}$$

or, transformed to $|\psi(t)\rangle$

$$|\psi(t)\rangle = e^{-(i/\hbar)\hat{\mathbf{H}}_0 t} |\psi_I(t)\rangle . \tag{2.65}$$

After substituting this into the Schrödinger equation (2.41) with a Hamiltonian defined by Eq. (2.63) the terms containing $\hat{\mathbf{H}}_0$ will cancel out and the result is an equation of motion for the interaction state vector $|\psi_I(t)\rangle$:

$$i\hbar \frac{\partial}{\partial t} |\psi_I(t)\rangle = \hat{\mathbf{V}}_I(t) |\psi_I(t)\rangle \tag{2.66}$$

with

$$\hat{\mathbf{V}}_I(t) = e^{(i/\hbar)\hat{\mathbf{H}}_0 t} \hat{\mathbf{V}}(t) e^{-(i/\hbar)\hat{\mathbf{H}}_0 t} \tag{2.67}$$

as the perturbation transformed to the interaction picture. This shows that for $\hat{\mathbf{V}}(t) = 0$ the state vector in the interaction picture will remain constant. What in fact has happened is that the rapidly varying factors caused by the action of $\hat{\mathbf{H}}_0$ are removed by applying the operator

$$\hat{\mathbf{U}}_0^\dagger(t, t_0) = e^{+(i/\hbar)\hat{\mathbf{H}}_0 (t - t_0)} \tag{2.68}$$

to all states in the Schrödinger picture. All operators are transformed by applying this operator and its complex conjugate from the left and from the right:

$$\hat{\mathbf{O}}_I(t) = \hat{\mathbf{U}}_0^\dagger(t, t_0) \hat{\mathbf{O}}(t) \hat{\mathbf{U}}_0(t, t_0) \tag{2.69}$$

This transformation shifts most of the time dependence from the state vectors to the operators, which is the key feature of this description.

After noting the fact, that for $t = t_0$ both $\hat{\mathbf{U}}_0(t, t_0) = \mathbf{1}$ and $\hat{\mathbf{U}}(t, t_0) = \mathbf{1}$ and therefore at this time both pictures are the same ($|\psi(t_0)\rangle = |\psi_I(t_0)\rangle$), and defining the time evolution operator in the interaction picture as

$$|\psi_I(t)\rangle = \hat{\mathbf{U}}_I(t, 0) |\psi_I(0)\rangle , \tag{2.70}$$

this can be inserted into Eq. (2.66). The same arguments given for the derivation of Eq. (2.51) and Eq. (2.56) then lead to an expansion for this operator as a time ordered exponential

$$\hat{\mathbf{U}}_I(t, t_0) = \exp_+ \left[-\frac{i}{\hbar} \int_{t_0}^{t} d\tau \, \hat{\mathbf{V}}_I(\tau) \right] . \tag{2.71}$$

Due to the fact that $\hat{\mathbf{V}}(t)$ is considered to be a small perturbation of the full Hamiltonian it is often possible to truncate the time ordered expansion at a much lower order than the one given for the total Hamiltonian in Eq. (2.57).

Applying the formalism described above to the density matrix, as described in Section 2.3.1, one obtains the equations

$$\hat{\rho}_I(t) = \hat{\mathbf{U}}_0^\dagger(t, t_0)\hat{\rho(t)}\hat{\mathbf{U}}_0(t, t_0) \tag{2.72}$$

and

$$\hat{\rho}_I(t) = \hat{\mathbf{U}}_I(t, t_0)\hat{\rho}_I(t_0)\hat{\mathbf{U}}_I^\dagger(t, t_0) \tag{2.73}$$

with

$$\hat{\rho}(t_0) = \hat{\rho}_I(t_0) \tag{2.74}$$

for the density matrix in interaction representation and

$$i\hbar\frac{\partial}{\partial t}\hat{\rho}_I(t) = \left[\hat{\mathbf{V}}_I(t), \hat{\rho}_I(t)\right], \tag{2.75}$$

as the Liouville–von Neumann equation in the interaction picture.

2.3.3 Reduced Density Matrix Theory

If only a small part of the total system is of actual interest, the statistical operator provides a convenient formalism for treating this problem [57, 61, 95]. The starting point for the treatment of such a system is a Hamiltonian which can be separated in the form

$$\hat{\mathbf{H}}(\mathbf{x}, \mathbf{Z}) = \hat{\mathbf{H}}_\mathrm{S}(\mathbf{x}) + \hat{\mathbf{H}}_{\mathrm{S-B}}(\mathbf{x}, \mathbf{Z}) + \hat{\mathbf{H}}_\mathrm{B}(\mathbf{Z}), \tag{2.76}$$

where $\hat{\mathbf{H}}_\mathrm{S}$ describes the relevant system with coordinates $\mathbf{x}$, $\hat{\mathbf{H}}_\mathrm{B}$ the bath with coordinates $\mathbf{Z}$ and $\hat{\mathbf{H}}_{\mathrm{S-B}}$ the interaction between the two parts. As long as the coupling between the two parts of the system is non–zero, it is not possible to factorize the total wavefunction into a system part $\phi_\mathrm{S}(\mathbf{x})$ and another one for the bath $\chi_\mathrm{B}(\mathbf{Z})$.

$$\Psi(\mathbf{x}, \mathbf{Z}) \neq \phi_\mathrm{S}(\mathbf{x})\chi_\mathrm{B}(\mathbf{Z}) \tag{2.77}$$

To attain the goal of representing the dynamics of the relevant system without making explicit use of the dynamics of the bath the total density operator $\hat{\rho}_\mathrm{SB}$ is first written as a density matrix using a complete orthonormal basis set for the system only and another one, which represents the bath, i.e. there is one basis set used to represent

$$\phi_\mathrm{S}(\mathbf{x}) = \sum_n c_n \left|\xi_n\right\rangle \tag{2.78}$$

and one for

$$\chi_\mathrm{B}(\mathbf{Z}) = \sum_n C_n \left|\zeta_n\right\rangle \tag{2.79}$$

from expansion (2.77). Using this representation one then introduces the reduced density matrix for the relevant system $\hat{\rho}_{\mathrm{S}}$ by taking the trace of the full operator $\hat{\rho}_{\mathrm{SB}}$ over the states representing the bath:

$$\hat{\rho}_{\mathrm{S}} = \sum_{\zeta_n} \langle \zeta_n | \, \hat{\rho}_{\mathrm{SB}} \, | \zeta_n \rangle = \mathrm{tr}_{\mathrm{B}} \left\{ \hat{\rho}_{\mathrm{SB}} \right\} . \tag{2.80}$$

With this partitioned Hamiltonian and the definition of the reduced density matrix one can derive an expression for the time evolution in the reduced system:

$$\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}(t) = -\frac{i}{\hbar}\left[\hat{\mathbf{H}}_{\mathrm{S}}, \hat{\rho(t)}_{\mathrm{S}}\right] - \frac{i}{\hbar}\mathrm{tr}_{\mathrm{B}}\left\{\left[\hat{\mathbf{H}}_{\mathrm{S-B}}, \hat{\rho(t)}_{\mathrm{SB}}\right]\right\} . \tag{2.81}$$

In this equation there is still a term containing the statistical operator for the complete system, so it does not provide a closed solution for the reduced density operator $\hat{\rho}_{\mathrm{S}}$. The next task therefore is to (approximately) eliminate the explicit references to the bath and relate the second term of Eq. (2.81) to the relevant system only.

The Mean Field Approximation

To continue without approximations, one can use the path integral formulation described in [96, 97], but in this work the mean field approximation is used. The main assumptions made for this are:

- The coupling between the system and the bath has been switched on at time $t = t_0$. Up to this point the density operator for the total system is represented by a direct product of operators from the system and the bath only:

$$\hat{\rho}_{\mathrm{SB}}(t_0) = \hat{\rho}_{\mathrm{S}}(t_0)\hat{\rho}_{\mathrm{B}}(t_0). \tag{2.82}$$

- The system–bath interaction Hamiltonian can be factorized into a system part $X_n(\mathbf{x})$ and a bath part $Q_n(\mathbf{Z})$,

$$\hat{\mathbf{H}}_{\mathrm{S-B}} = \sum_n X_n(\mathbf{x})Q_n(\mathbf{Z}). \tag{2.83}$$

  As the only restriction on these operators is, that they are exclusively defined in their respective system and bath coordinates, this form is general enough to cover all cases of practical importance (e.g. the CRS Hamiltonian has this form).

- The bath is considered to be much larger than the relevant system, which means that any effects of the interaction with the system will not change it significantly. This leads to the assumption that the bath remains in thermal equilibrium at a constant temperature, even after the interaction is switched on:

$$\hat{\rho}_{\mathrm{B}}(t) = \hat{\rho}_{\mathrm{B}}(t_0) = \hat{\rho}_{\mathrm{B}}^{\mathrm{eq}} = \exp\left\{-\hat{\mathbf{H}}_{\mathrm{B}}/k_B T\right\} / \mathrm{tr}_{\mathrm{B}}\left\{e^{-\hat{\mathbf{H}}_{\mathrm{B}}/k_B T}\right\}, \tag{2.84}$$

  irrespective of the energy diffusing into it from the system. This results in the irreversibility of the process taking place, as anything put into the bath cannot interact back on the relevant system.

Going to the interaction picture with the Hamiltonian $\hat{\mathbf{H}}_{\mathrm{S-B}}$, which describes the coupling between the system and the bath as the perturbation of the total Hamiltonian will give an equation of motion for the reduced density operator in the form

$$\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}^{I}(t) = -\frac{i}{\hbar}\mathrm{tr}_{\mathrm{B}}\left\{\left[\hat{\mathbf{H}}_{\mathrm{S-B}}^{I}(t), \hat{\rho}_{\mathrm{SB}}^{I}(t)\right]\right\}. \tag{2.85}$$

Inserting the factorized form of the interaction Hamiltonian, performing the trace operation and changing back to the Schrödinger picture results in the following equation of motion of the reduced density operator:

$$\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}(t) = -\frac{i}{\hbar}\left[\hat{\mathbf{H}}_{\mathrm{S}} + \sum_{n}\langle Q_n\rangle_{\mathrm{B}} X_n, \hat{\rho}_{\mathrm{S}}(t)\right]. \tag{2.86}$$

This is the Liouville–von Neumann equation for the relevant system, where the Hamiltonian $\hat{\mathbf{H}}_{\mathrm{S}}$ has been replaced by $\hat{\mathbf{H}}_{\mathrm{S}}+\sum_n \langle Q_n\rangle_{\mathrm{B}} X_n$. The additional term represents the expectation value of the system-bath coupling at the thermal equilibrium of the bath. So this first order solution, also called mean field approximation, will only shift the energy levels of the system, but generates no energy relaxation from the system into the bath. To calculate this effect, the next order of the perturbation expansion has to be considered.

Equation of Motion in Second–Order Perturbation Theory

To generate the equations of motion for the reduced statistical operator in second order perturbation theory with respect to the system–bath interaction, a method to restrict the operators to the relevant system has to be introduced. For this one can introduce the two projection operators $\mathcal{P}$ and $\mathcal{Q}$ defined as

$$\mathcal{P}\hat{\mathbf{O}} = \hat{\rho}_{\mathrm{B}}\mathrm{tr}_{\mathrm{B}}\left\{\hat{\mathbf{O}}\right\}, \tag{2.87}$$

where normally the bath operator is set to the equilibrium operator $\hat{\rho}_{\mathrm{B}} = \hat{\rho}_{\mathrm{B}}^{\mathrm{eq}}$ and

$$\mathcal{Q} = 1 - \mathcal{P}. \tag{2.88}$$

With this it is possible to split the Liouville–von Neumann equation Eq. (2.75) into two coupled equations

$$\begin{aligned}\mathrm{tr}_{\mathrm{B}}\left\{\mathcal{P}\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{SB}}^{I}\right\} &= \frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}^{I} \\ &= -\frac{i}{\hbar}\mathrm{tr}_{\mathrm{B}}\left\{\left[\hat{\mathbf{H}}_{\mathrm{S-B}}^{I}, \hat{\rho}_{\mathrm{B}}^{\mathrm{eq}}\hat{\rho}_{\mathrm{S}}^{I} + \mathcal{Q}\hat{\rho}_{\mathrm{SB}}^{I}\right]\right\}\end{aligned} \tag{2.89}$$

and

$$\frac{\partial}{\partial t}\mathcal{Q}\hat{\rho}_{\mathrm{SB}}^{I} = -\frac{i}{\hbar}\mathcal{Q}\left[\hat{\mathbf{H}}_{\mathrm{S-B}}^{I}, \hat{\rho}_{\mathrm{B}}^{\mathrm{eq}}\hat{\rho}_{\mathrm{S}}^{I} + \mathcal{Q}\hat{\rho}_{\mathrm{SB}}^{I}\right]. \tag{2.90}$$

These equations describe the relevant system $\hat{\rho}_{\rm S}^I$ and the irrelevant part of the density operator $\mathcal{Q}\hat{\rho}_{\rm SB}^I$.

The iterative solution of these equations generates a perturbation expansion with respect to $\hat{\mathbf{H}}_{\rm S-B}$. Neglecting the irrelevant part of the density operator $\mathcal{Q}\hat{\rho}_{\rm SB}^I$ entirely will recover the mean field expression (Eq. (2.86)) from the last section. The next order is generated by neglecting $\mathcal{Q}\hat{\rho}_{\rm SB}^I$ on the right-hand side of Eq. (2.90), which allows a formal solution for this expression:

$$\mathcal{Q}\hat{\rho}_{\rm SB}^I(t) = \mathcal{Q}\hat{\rho}_{\rm SB}^I(t_0) - \frac{i}{\hbar}\int_{t_0}^{t} d\tau\, \mathcal{Q}\left[\hat{\mathbf{H}}_{\rm S-B}^I(\tau), \hat{\rho}_{\rm B}^{\rm eq}\hat{\rho}_{\rm S}^I(\tau)\right]. \tag{2.91}$$

Inserting this into Eq. (2.89), one obtains the equation of motion for the reduced density operator in second order with respect to $\hat{\mathbf{H}}_{\rm S-B}$:

$$\begin{aligned}\frac{\partial}{\partial t}\hat{\rho}_{\rm S}^I(t) &= -\frac{i}{\hbar}{\rm tr}_{\rm B}\left\{\hat{\rho}_{\rm B}^{\rm eq}\left[\hat{\mathbf{H}}_{\rm S-B}^I(t), \hat{\rho}_{\rm S}^I(t)\right]\right\} \\ &\quad -\frac{1}{\hbar^2}\int_{t_0}^{t} d\tau\, {\rm tr}_{\rm B}\left\{\left[\hat{\mathbf{H}}_{\rm S-B}^I(t), (1-\mathcal{P})\left[\hat{\mathbf{H}}_{\rm S-B}^I(\tau), \hat{\rho}_{\rm B}^{\rm eq}\hat{\rho}_{\rm S}^I(\tau)\right]\right]\right\}. \end{aligned}\tag{2.92}$$

Reservoir Correlation Functions

Looking at Eq. (2.92) under the assumption, that the system–bath coupling Hamiltonian can be factorized as in Eq. (2.83), the first term again describes the mean field motion ((2.86))and can be written as

$${\rm tr}_{\rm B}\left\{\hat{\rho}_{\rm B}^{\rm eq}\left[\hat{\mathbf{H}}_{\rm S-B}^I(t), \hat{\rho}_{\rm S}^I(t)\right]\right\} = \sum_n \left[\langle Q_n\rangle_{\rm B} X_n, \hat{\rho}_{\rm S}(t)\right]. \tag{2.93}$$

The second term contains the second order effects in the double commutator in the integrand. After multiplying out the commutators, applying the projection operator and inserting the factorized Hamiltonian, it can be shown, that the integrand in the second term of Eq. (2.92) can be written as a sum of terms all containing functions of the type

$$C_{mn}(t) = \langle \delta Q_m(t)\Delta Q_n(0)\rangle_{\rm B} = \langle Q_m(t)Q_n(0)\rangle_{\rm B} - \langle Q_m\rangle_{\rm B} - \langle Q_n\rangle_{\rm B}, \tag{2.94}$$

which describe the fluctuations of the bath part of $\hat{\mathbf{H}}_{\rm S-B}$ around its average value. The $C_{mn}(t)$ defined in (2.94) are called reservoir correlation functions and describe the correlations between the fluctuations of the different reservoir operators at different times. These fluctuations will not change the quantum mechanical state of the bath – it will remain in thermal equilibrium. The correlation of these fluctuations normally decays after a certain correlation time τ_C.

Using the definitions of the correlation functions, the equation of motion for the reduced density operator reads

$$\frac{\partial}{\partial t}\hat{\rho}_{\rm S}^I(t) = -\frac{i}{\hbar}\sum_n \langle Q_n\rangle_{\rm B}\left[X_n, \hat{\rho}_{\rm S}^I(t)\right]$$

$$
\begin{aligned}
&- \frac{1}{\hbar^2}\sum_{m,n}\int_{t_0}^{t} d\tau \,\Big(C_{mn}(t-\tau)\left[X_m^I(t), X_n^I(\tau)\hat{\rho}_{\mathrm{S}}^I(\tau)\right] \\
&- C_{mn}^*(t-\tau)\left[X_m^I(t), \hat{\rho}_{\mathrm{S}}^I(\tau)X_n^I(\tau)\right]\Big)\,, \qquad (2.95)
\end{aligned}
$$

or in the Schrödinger picture:

$$
\begin{aligned}
\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}(t) &= -\frac{i}{\hbar}\left[\hat{\mathbf{H}}_{\mathrm{S}} + \sum_n \langle Q_n\rangle_{\mathrm{B}}\, X_n, \hat{\rho}_{\mathrm{S}}(t)\right] \\
&- \frac{1}{\hbar^2}\sum_{m,n}\int_0^{t-t_0} d\tau \,\Big(C_{mn}(\tau)\left[X_m(t), U_{\mathrm{S}}(\tau)X_n(\tau)\hat{\rho}_{\mathrm{S}}(t-\tau)U_{\mathrm{S}}^\dagger(\tau)\right] \\
&- C_{mn}^*(\tau)\left[X_m(t), U_{\mathrm{S}}(\tau)\hat{\rho}_{\mathrm{S}}(t-\tau)X_n(\tau)U_{\mathrm{S}}^\dagger(\tau)\right]\Big)\,. \qquad (2.96)
\end{aligned}
$$

This equation, which is called Quantum Master Equation (QME), shows, that the time dependence of the reduced density operator is not only generated by its actual value, but also by the history of its evolution. The length of this memory effect is determined by the decay time of the reservoir correlation functions. A discussion of this and further properties of the correlation functions can be found in the literature [61].

The Markow Approximation

Under the assumption that the environment has a time scale for the memory effects shorter than the other relevant time scales for the system the so called Markow approximation of the QME can be derived. This means that the density operator describing the system will stay approximately constant during the time span in which the reservoir correlation functions C_{mn} decay to zero. This allows one to replace the upper limit of the integral in Eq. (2.96) with ∞, because the integrand will be zero after a time interval shorter than $t - t_0$. Hence it is possible to use in Eq. (2.96) as an approximation

$$\hat{\rho}_{\mathrm{S}}(t-\tau) \approx \hat{\rho}_{\mathrm{S}}(t), \qquad (2.97)$$

and rewrite the QME in a more compact form. For this the operator

$$\Lambda_m = \sum_n \int_0^\infty d\tau\, C_{mn}(\tau) X_n^I(-\tau) \qquad (2.98)$$

is introduced and the dissipative part of Eq. (2.96) is written as

$$\left(\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}\right)_{\mathrm{diss}} = -\frac{1}{\hbar^2}\sum_m \left[X_m, \Lambda_m\hat{\rho}_{\mathrm{S}}(t) - \hat{\rho}_{\mathrm{S}}(t)\Lambda_m^\dagger\right]. \qquad (2.99)$$

Carrying out the commutator allows the definition of an effective system Hamiltonian, which contains the first order mean field terms of the QME and those parts of the commutator, which are proportional to $X_m\Lambda_m$:

$$\hat{\mathbf{H}}_{\mathrm{S}}^{\mathrm{eff}} = \hat{\mathbf{H}}_{\mathrm{S}} + \sum_m \langle Q_n\rangle_{\mathrm{B}}\, X_m - \frac{i}{\hbar}\sum_m X_m\Lambda_m. \qquad (2.100)$$

This leads to the final form of the QME in Markow approximation as

$$\begin{aligned}\frac{\partial}{\partial t}\hat{\rho}_{\mathrm{S}}(t) &= -\frac{i}{\hbar}\left(\hat{\mathbf{H}}_{\mathrm{S}}^{\mathrm{eff}}\hat{\rho}_{\mathrm{S}}(t) - \hat{\rho}_{\mathrm{S}}(t)\hat{\mathbf{H}}_{\mathrm{S}}^{\mathrm{eff}\,\dagger}\right) \\ &+ \frac{1}{\hbar^2}\sum_m \left(X_m\hat{\rho}_{\mathrm{S}}(t)\Lambda_m^\dagger + \Lambda_m\hat{\rho}_{\mathrm{S}}(t)X_m^\dagger\right). \end{aligned} \tag{2.101}$$

Model Environments and Spectral Density

The exact calculation of the correlation functions C_{mn} is practically not possible for real systems. To allow for a treatment of the system–bath interaction, models for the environment have to be introduced. One possible way of including the bath is to describe its degrees of freedom in low order, i.e. by linear coupling terms in the bath coordinates, which results in a coupling Hamiltonian of the form

$$\hat{\mathbf{H}}_{\mathrm{S-B}} = X(s)\sum_\zeta c_\zeta q_\zeta \tag{2.102}$$

where the operator X depends on the system coordinates s and represents a simplified form of the system part of the coupling. The c_ζ are the coupling constants between the system and the bath and the q_ζ are the bath coordinates, which were labeled as $\mathbf{Z}$ in Eq. (2.83). For a bath consisting of harmonic oscillators, which could describe, for instance, a solid or an only slightly perturbed liquid, this approach allows the analytical description of the bath [61].

As this analytical description is in general not possible, it is normally more convenient to introduce the so called spectral density $\mathcal{J}(\omega)$, which models the ability of the bath to exchange energy with the system at a certain frequency. This function allows the correlation functions for a harmonic oscillator bath to be written as

$$C(\omega) = 2\pi\hbar^2\omega^2\left[1 + n(\omega)\right]\left[\mathcal{J}(\omega) - \mathcal{J}(-\omega)\right] \tag{2.103}$$

with

$$n(\omega) = \frac{1}{e^{\hbar\omega/k_BT} - 1} \tag{2.104}$$

the Bose–Einstein distribution assuring the correct thermal equilibrium conditions of the correlation functions.

While these spectral densities in principle are generated from a sum of delta functions, any macroscopic bath will effectively have a continuous spectrum, being able to absorb energy at every frequency. These spectral functions can be either represented by simple models, or generated via classical molecular dynamics simulations. In this approach the correlation functions $C_{mn}(\omega)$ are calculated from the classical correlation functions $\zeta_{mn}(t) = \langle\phi_m(t)\phi_n(0)\rangle_{\mathrm{cl}}$ by relating the Fourier transform of $\zeta_{mn}(t)$ to $C_{mn}(\omega)$:

$$C_{mn}(\omega) = 2\left(1 + \exp\left\{-\frac{\hbar\omega}{k_BT}\right\}\right)^{-1}\zeta_{mn}(\omega) \tag{2.105}$$

as described in [61]. This also allows one to relate the vibrational population relaxation time T_1 of a quantum system to the correlation functions, as here the relation

$$\frac{1}{T_1} \propto \zeta(\omega) \tag{2.106}$$

holds, as described in [98] and later in Section 3.3.5.

One often used form of $\mathcal{J}(\omega)$ is given by

$$\omega^2 \mathcal{J}(\omega) = \Theta(\omega) j_0 \omega^p e^{-\omega/\omega_c}, \tag{2.107}$$

where the unit–step function $\Theta(\omega)$ guarantees that $\mathcal{J} = 0$ for $\omega < 0$, and j_0 is some normalization factor. For $p = 1$ and a cutoff frequency ω_c much higher than the relevant system frequencies, this leads to the so called Ohmic form for the spectral density, which is frequently used in modeling.

The Concept of Superoperators

To simplify the notation (e.g. of Eq. (2.96)), it is convenient to introduce the so called Liouville space, which is a linear vector space consisting of linear operators from Hilbert space. The super-operators are operators acting on the elements of this Liouville space in the same way that operators of Hilbert space act on the state vectors. The Liouville operator $\mathcal{L}$ is defined as acting on an operator as

$$\mathcal{L}\hat{\mathbf{O}} = \frac{1}{\hbar}\left[\hat{\mathbf{H}}, \hat{\mathbf{O}}\right], \tag{2.108}$$

which allows the Liouville–von Neumann equation to be written as

$$\frac{\partial}{\partial t}\hat{\rho}(t) = -i\mathcal{L}\hat{\rho}(t). \tag{2.109}$$

Introducing the relaxation super-operator $\mathcal{R}(t,\tau)$, which is defined via the dissipative part of the QME, i.e. the integrand of Eq. (2.96) and the Liouville operator for the system Hamiltonian $\mathcal{L}_\mathrm{S}$, the QME can be written in the simple form

$$\frac{\partial}{\partial t}\hat{\rho}_\mathrm{S}(t) = -i\mathcal{L}_\mathrm{S}\hat{\rho}_\mathrm{S}(t) - \int_{t_0}^{t} d\tau\, \mathcal{R}(t,\tau)\hat{\rho}_\mathrm{S}(\tau). \tag{2.110}$$

This just provides a convenient short notation, and will of course not simplify the equations themselves.

2.4 Numerical Implementation

The exact analytical solution of any quantum mechanical problem – either stationary or dynamical – beyond idealized cases like the harmonic oscillator or single hydrogen atoms,

is normally not possible. To solve the Schrödinger equations for the wavefunction or the Liouville–von Neumann equations for the density operator it is therefore necessary to introduce suitable numerical methods, allowing the use of computers to tackle the problem. The normal approach to do this is the representation of the problem in a suitable reduced basis. So in a full basis set given by $|\phi_n\rangle$ a quantum mechanical state $|\Phi\rangle$ can be written as an infinite vector of coefficients

$$|\Phi\rangle = \sum_n c_n |\phi_n\rangle \equiv \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_\infty \end{pmatrix} \tag{2.111}$$

and an operator can be represented by a matrix of infinite dimension

$$O_{mn} = \langle\phi_m| \hat{\mathbf{O}} |\phi_n\rangle \qquad n, m = 1 \ldots \infty. \tag{2.112}$$

In a reduced basis set a suitable finite subset of a full basis is chosen to reduce this to vectors and matrices of finite size. This reduces the operator equations to matrix equations and large coupled systems of differential equations for the coefficients c_n. These can be handled with standard methods from linear algebra and the theory of differential equations.

Energy Eigenfunctions

One of the most convenient representations is to use the solutions of the stationary Schrödinger equation (2.1) as a basis set. In this basis obviously the matrix representing the Hamilton operator is diagonal with the eigenenergies as the diagonal elements. So as long as there is no perturbation introduced, the time dependent Schrödinger equation can be written as

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} E_1 & 0 & \cdots & \\ 0 & E_2 & & \\ \vdots & & \ddots & \\ & & & E_n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \tag{2.113}$$

which leads to a set of uncoupled differential equations for the coefficients.

The additional convenient feature of this approach is that the dimension of the problem in principle is irrelevant. As long as the eigenfunctions are known the system can be reduced to this discretization on the 1D energy grid. The limiting factors for this representation are the fact that in higher dimensions the discrete energy levels move much closer together, requiring more and more basis functions to span a given energy interval, and of course, the fact that the eigenfunctions of the system have to be determined beforehand by other means, which quickly increases the effort when going to high dimensional systems.

2.4.1 The Fourier–Grid Method

A widely used and straightforward method to calculate the eigenfunctions of the stationary Schrödinger equation is the "Fourier–Grid Method" introduced by Marston and Balint–Kurti [99]. It is based on the representation of the wavefunction in the basis of the position operator $\hat{\mathbf{x}}$ in normal space and the momentum operator $\hat{\mathbf{p}}$ in momentum space. It relies on the fact that in the total Hamiltonian $\hat{\mathbf{H}} = \hat{\mathbf{T}} + \hat{\mathbf{V}}$ the potential energy part $\hat{\mathbf{V}}$ is diagonal in the basis of the position operator, while the kinetic energy part $\hat{\mathbf{T}}$ is diagonal in the basis of the momentum operator. As the position and the momentum picture can be transformed into each other by the Fourier transformation it is easy to derive a closed expression for the matrix elements of the total Hamiltonian in this representation. As the position and momentum operator have a continuous basis, it is then necessary to introduce a reduced basis set in the form of a discrete grid in one of the representations. (This derivation for a grid in position space is described in Appendix A.2.)

After having obtained this Hamilton matrix the eigenenergies and eigenfunctions for the system on the chosen grid can be obtained via a simple diagonalization of this matrix.

The big advantage of this straightforward method is the fact, that it calculates the eigenfunctions on a grid in position or momentum space, which later can be used for the numerical implementation of the quantum dynamics. Using the same basis set for both the stationary and the dynamical calculations eliminates any additional errors, which would be caused by transforming the results from one finite, reduced basis to another.

On the other hand the method has its limitations in the size of the systems which can be treated. The size of the grid representation of the Hamilton matrix which has to be diagonalized for a system with N grid points is $M = N \times N$. For one or two dimensional systems this represents no big problem for todays computer systems, but when going to higher dimensions the same remarks as made in Section 2.2 about the number of grid points in the potential apply: Anything above three dimensions quickly becomes close to impossible to solve due to the fact that numerical matrix diagonalization requires on the order of M^3 operations. Another technical limitation of the computer is the amount of memory needed to keep the full matrix available.

As an alternative to this method the Discrete Variable Representation (DVR) methods have to be mentioned [100]. In these the system is represented not on a regular grid but in a set of basis functions specially tailored to the problem, which normally reduces the number of grid points significantly. Often one still uses a – now irregular – grid in position space for this, reducing the number of points in regions which the wavefunction cannot reach, but in principle any basis set can be used for this. The representation in the eigenstates of the stationary Hamiltonian mentioned above is a special case of this approach. Another, completely different approach to get the eigenfunctions of a system is to use the methods from quantum dynamics discussed later in this Section and to do a so called "propagation in imaginary time", which relaxes the system to the lowest eigenstate. Together with projection methods then higher states can be constructed as well, but this

approach requires very high numerical accuracy and is limited by the finite accuracy of the computer. As the propagation with, e.g. the split operator (Section 2.4.3), scales with $N\log(N)$ even quite large systems can be handled with this approach. Further information about these methods can be found in the literature [101].

2.4.2 The Redfield Equation

The Redfield equation is obtained from a representation of the QME (Eq. (2.110)) in the basis of the eigenfunctions of the part of the system Hamiltonian $\hat{\mathbf{H}}_0$, which contains the full potential generated by the system itself (e.g. the molecular potential for a system describing a molecule). The total Hamiltonian is $\hat{\mathbf{H}}_\mathrm{S} = \hat{\mathbf{H}}_0 + \hat{\mathbf{V}}$, which includes the operator $\hat{\mathbf{V}}$ containing all parts of the system potential caused by external influences, e.g. an electric field interacting with the system.

In this basis the QME can be written for the density matrix in the form

$$\frac{\partial}{\partial t}\rho_{mn} = i\omega_{mn}\rho_{mn} - \frac{i}{\hbar}\sum_k \left(V_{nk}\rho_{km} - \rho_{nk}V_{km}\right) - \sum_{kl} R_{mn,kl}\rho_{kl}, \tag{2.114}$$

where $\omega_{mn} = (E_m - E_n)/\hbar$ are the transition frequencies between the different energy levels and V_{mn} are the matrix elements of the operator $\hat{\mathbf{V}}$. The dissipation is described by the relaxation tensor $R_{kl,mn}$. This tensor is calculated via the so called damping matrix

$$\Gamma_{kl,mn}(\omega_{mn}) = \frac{1}{\hbar^2}\Re\sum_{ij}\langle k|\, X_i\,|l\rangle\,\langle m|\,\Lambda_j\,|n\rangle\,, \tag{2.115}$$

or, using the spectral density (Eq. (2.103)) instead of the correlation function in the operator Λ_j, via

$$\Gamma_{kl,mn}(\omega_{mn}) = \sum_{ij}\langle k|\, X_i\,|l\rangle\,\langle m|\, X_j\,|n\rangle\,[1+n(\omega)]\,[\mathcal{J}(\omega) - \mathcal{J}(-\omega)]\,. \tag{2.116}$$

The imaginary parts of these terms will be neglected in the further calculations, as they do not affect the relaxation rates [61]. With this simplification, the relaxation tensor then is

$$\begin{aligned} R_{kl,mn} \;=\; & \delta_{km}\sum_j \Gamma_{lj,jn}(\omega_{nj}) + \delta_{ln}\sum_j \Gamma_{kj,jm}(\omega_{mj}) \\ & -\Gamma_{mk,ln}(\omega_{nl}) - \Gamma_{nl,km}(\omega_{mk}). \end{aligned} \tag{2.117}$$

This tetradic matrix, describing the relaxation within a set of energy eigenstates, is frequently referred to as the Redfield tensor in literature [39, 58, 59].

The application of this tensor to the density matrix results in energy dissipation for the system. The elements of a density matrix ρ_{mn} can be separated into two groups: The populations (with $m = n$) and the coherences between them ($m \neq n$). The effect of $R_{kl,mn}$ on the dynamics of the ρ_{mn} can be split into three main groups:

1. Population transfer ($k = l$, $m = n$):
   These matrix elements are given by

$$R_{kk,mm} = 2\delta_{km} \sum_j \Gamma_{kj,jm}(\omega_{mj}) - 2\Gamma_{mk,km}(\omega_{mk}) \tag{2.118}$$

   or, using the transition rate from energy level k to level m as $k_{km} = 2\Gamma_{mk,km}(\omega_{mk})$, by

$$R_{kk,mm} = 2\delta_{km} \sum_j k_{kj} - k_{mk}. \tag{2.119}$$

   The first term corresponds to the transitions from eigenstate $|\varphi_k\rangle$ into all other states $|\varphi_j\rangle$ of the system. The second term describes the transitions from all other states back into $|\varphi_k\rangle$. The forward and backward rates between the levels are connected via the principle of detailed balance:

$$k_{km} = e^{\hbar\omega_{km}/k_B T} k_{mk} \tag{2.120}$$

   which ensures that the system will move toward thermal equilibrium.

2. Coherence dephasing ($k \neq l$, $k = m$, $l = n$):
   These matrix elements are given by

$$R_{kl,kl} = \sum_j \left(\Gamma_{kj,jk}(\omega_{kj}) + \Gamma_{jk,kj}(\omega_{jk})\right) - \Gamma_{kk,ll}(0) - \Gamma_{ll,kk}(0). \tag{2.121}$$

   These terms lead to a decay of the off-diagonal elements of the density matrix, which describe the coherences (i.e. the phase relations of the different states) of the system. The loss of these coherences, also called dephasing, will lead to a system with independently populated states, as described by the density operator in Eq. (2.58).

3. All the elements of $R_{kl,mn}$ not belonging to (1) or (2):
   The interpretation of the remaining elements of the Redfield tensor is not so simple. There are several processes, which are described by them: Coherences can be moved from one pair of states to another ($\rho_{kl} \to \rho_{mn}$ via $R_{kl,mn}$), populations can be transformed into coherences ($\rho_{kk} \to \rho_{mn}$ via $R_{kk,mn}$) and coherences can move back to populations ($\rho_{kl} \to \rho_{mm}$ via $R_{kl,mm}$). These effects cause a mixing between all elements of the density matrix.

One has to note, that the elements of the Redfield tensor $R_{kl,mm}$ will become time-dependent when the system it describes includes a time-dependent component like an external field.

The Secular Approximation

In the secular approximation the terms of the Redfield tensor mentioned under (3) in the last section, which mix the elements of the density matrix, is neglected. The conditions necessary to apply this approximation can be seen from the dissipative term of Eq. (2.114) in the interaction picture:

$$\left(\frac{\partial}{\partial t}\rho^I_{mn}\right)_{\text{diss}} = -\sum_{kl} R_{mn,kl} e^{i(\omega_{mn}-\omega_{kl})(t-t_0)} \rho^I_{kl}(t). \tag{2.122}$$

The right–hand side of this term oscillates with the frequency $\omega_{mn} - \omega_{kl}$. This means that the integration of the equations of motion with a time step Δt larger than $1/(\omega_{mn} - \omega_{kl})$ results in these contributions canceling out due to destructive interference. The validity of this approximation has been shown, e.g., in [102, 103]

The contributions to $R_{kl,mn}$ which are given for cases (1) and (2) in the previous sections can never be neglected under this approximation, as for them $|\omega_{kl} - \omega_{mn}| = 0$ always holds. For those elements of the Redfield tensor listed under (3) the condition $1/(\omega_{mn} - \omega_{kl}) \ll \Delta t$ will often be true, as long as the system has no degenerate transition frequencies (e.g. in a harmonic oscillator). So the secular approximation of neglecting all parts of the dissipative part of the QME for which $|\omega_{kl} - \omega_{mn}| \neq 0$ holds, is especially suited for very anharmonic systems. In this type of system the condition $|\omega_{kl} - \omega_{mn}| = 0$ is normally only fulfilled accidentally.

2.4.3 The Split Operator

The solution of the time dependent Schrödinger equation (2.41) for a time independent operator is given by Eq. (2.45). While in most of the problems treated in this work the Hamiltonian is time dependent, for the numerical solution of the problem the differential equations are integrated by taking sufficiently small discrete time steps, during which the Hamiltonian can be considered as constant. So for a small time step Δt the time evolution operator can be written as

$$\hat{\mathbf{U}}(t, t+\Delta t) = e^{-(i/\hbar)\hat{\mathbf{H}}\Delta t}. \tag{2.123}$$

This propagation is easily done when the Hamiltonian is represented in a basis for which it is diagonal. Unfortunately this is only the case for the basis given by the eigenfunctions of the total Hamiltonian, so for a propagation in a grid representation of the wavefunction as the one generated for the "Fourier–Grid Method" (Section 2.4.1 and Appendix A.2) one has to use an additional approximation.

In the split operator approach introduced by Feit and Fleck [104] and Kosloff [105], the exponential term for the total Hamiltonian $\hat{\mathbf{H}} = \hat{\mathbf{T}} + \hat{\mathbf{V}}$ is split into a product in the form

$$e^{-(i/\hbar)\hat{\mathbf{H}}\Delta t} = e^{-(i/2\hbar)\hat{\mathbf{T}}\Delta t} e^{-(i/\hbar)\hat{\mathbf{V}}\Delta t} e^{-(i/2\hbar)\hat{\mathbf{T}}\Delta t} + \mathcal{O}(\Delta t^3). \tag{2.124}$$

The error which is introduced due to the fact that $\hat{\mathbf{T}}$ and $\hat{\mathbf{V}}$ are non-commuting operators ($[\hat{\mathbf{T}}, \hat{\mathbf{V}}] \neq 0$) scales with the cube of the time step length. From Section 2.4.1 it is known, that the potential energy operator is diagonal in a grid representation of the position space, while the kinetic energy operator has a diagonal representation in momentum space. So to propagate a wavefunction $|\Psi\rangle_x$ represented on a position grid by a single time step Δt one has to:

1. Apply a Fourier transform to bring $|\Psi\rangle_x$ into the momentum representation $|\Psi\rangle_k$.
2. Apply the – now diagonal – operator $e^{-(i/2\hbar)\hat{\mathbf{T}}\Delta t}$.
3. Apply an inverse Fourier transform to bring the result back to position space.
4. Apply the operator $e^{-(i/\hbar)\hat{\mathbf{V}}\Delta t}$.
5. Apply another Fourier transform to bring the result back to momentum space.
6. Apply the second half of the kinetic energy operator $e^{-(i/2\hbar)\hat{\mathbf{T}}\Delta t}$.
7. Do a final inverse Fourier transform to bring the result back to position space.

If the wavefunction is not required in the position representation the inverse Fourier transform at the and can be left out, which also removes the necessity for the first transformation, as the function already is in momentum space. So except for the first and last time step each propagation step requires the application of three diagonal operators (in principle a simple multiplication for each grid point), one Fourier transform and one inverse Fourier transform. For a Hamiltonian with a time dependent potential term, the potential $V(t)$ and the resulting operator $e^{-(i/\hbar)\hat{\mathbf{V}}\Delta t}$ has to be reevaluated after each time step as well. The time limiting steps here are the Fourier transforms. The most efficient algorithm is the Fast Fourier Transform (FFT) (described in Appendix A.1), which has a numerical effort of $N \log_2 N$ for N grid points. Further discussions of this method can be found in the literature [106, 107].

While the scaling of this propagation method is quite efficient and allows large grids, one still has to remember that the grid size itself has an exponential scaling behavior with respect to the degrees of freedom of the system, so going beyond three dimensions with a wavefunction set up on a full grid presents quite a challenge. In the next section the multi configuration time dependent Hartree approach will be presented, which uses a further approximation to avoid using a full grid and allow the treatment of higher dimensional problems.

2.4.4 Introduction to the Multi Configurational Time Dependent Hartree Method

The multi configurational time dependent Hartree method (MCTDH) method [108] is based on the time dependent Hartree (TDH) approximation [46, 109], which is used to avoid the

exponential scaling problem of a numerically exact propagation. In the TDH approach the wavefunction for a system with f degrees of freedom is written as

$$\Psi(x_1, \ldots, x_f, t) = a(t) \prod_{n=1}^{f} \varphi_n(x_n, t), \tag{2.125}$$

where $a(t)$ is a complex coefficient and the φ_n are the *single particle functions* (SPFs) – one dimensional wavefunctions for each degree of freedom. The whole product is called a Hartree product. The single particle functions in Eq. (2.125) each contain an arbitrary phase factor, which can be fixed with the constraint

$$\langle \varphi_n | \, \dot{\varphi}_n \rangle = 0. \tag{2.126}$$

This also fixes the norm of the φ_n, which initially are normalized to one:

$$\|\varphi_n(t)\| = 1. \tag{2.127}$$

With these conditions one can derive the TDH equations of motion for the coefficient $a(t)$ and the single particle functions $\varphi_n(t)$ as:

$$i\hbar \dot{a} = \left\langle \hat{\mathbf{H}} \right\rangle a, \tag{2.128}$$

with $\left\langle \hat{\mathbf{H}} \right\rangle = \langle \varphi_1 \ldots \varphi_f | \, \hat{\mathbf{H}} \, | \varphi_1 \ldots \varphi_f \rangle$ for the coefficient and

$$i\hbar \dot{\varphi}_n = \left(\hat{\mathbf{H}}^{(n)} - \left\langle \hat{\mathbf{H}} \right\rangle \right) \varphi_n \tag{2.129}$$

with the mean field operator $\hat{\mathbf{H}}^{(n)}$ defined as

$$\hat{\mathbf{H}}^{(n)} = \langle \varphi_1 \ldots \varphi_{n-1} \varphi_{n+1} \ldots \varphi_f | \, \hat{\mathbf{H}} \, | \varphi_1 \ldots \varphi_{n-1} \varphi_{n+1} \ldots \varphi_f \rangle \, . \tag{2.130}$$

This means that each SPF is propagated by an effective Hamiltonian depending on the mean field of all other degrees of freedom. As each SPF is only one dimensional the effort to propagate the total wavefunction is greatly reduced. The price for this is, that the correlations between the different degrees of freedom are no longer treated correctly, leading to inaccuracies for longer propagation times or strongly coupled systems. The obvious way to improve the accuracy of the method is to represent the total wavefunction not with a single Hartree product, but to take a linear combination of several configurations into account. This then leads to the MCTDH method, where several SPFs per degree of freedom can be used, to allow a better representation of the couplings between the coordinates.

The starting point for the MCTDH approximation is the representation of the wavefunction Ψ for f degrees of freedom as a linear combination of Hartree products [108]:

$$\Psi(q_1, \ldots, q_f, t) = \sum_{j_1=1}^{n_1} \cdots \sum_{j_f=1}^{n_f} A_{j_1 \ldots j_f}(t) \prod_{\kappa=1}^{f} \varphi_{j_\kappa}^{(\kappa)}(q_\kappa, t), \tag{2.131}$$

with the nuclear coordinates $q_1, \ldots, q_f$, the MCTDH expansion coefficients $A_{j_1 \ldots j_f}(t)$ and the n_κ single particle functions $\varphi_{j_\kappa}^{(\kappa)}$ for each degree of freedom. For $n_1 = \ldots = n_f = 1$ the MCTDH wavefunction reduces to the TDH case of a single Hartree product, giving the TDH approximation as a special case of the MCTDH approach. For increasing numbers of SPFs the propagation will become more accurate as the MCTDH wavefunction approaches the numerically exact one. At the same time, of course, the computational effort increases strongly with increasing values of n_κ.

As with the TDH wavefunction the definition of the MCTDH wavefunction is not unique, since the SPFs may be linearly transformed while still representing the same function. To fix this, two constraints are imposed on the SPFs – they have to be orthonormal at $t = 0$

$$\left\langle \varphi_j^{(\kappa)}(0) \middle| \, \varphi_l^{(\kappa)}(0) \right\rangle = \delta_{jl} \tag{2.132}$$

and they have to obey

$$\left\langle \varphi_j^{(\kappa)}(t) \middle| \, \dot{\varphi}_l^{(\kappa)}(t) \right\rangle = -\frac{i}{\hbar} \left\langle \varphi_j^{(\kappa)}(t) \middle| \, \hat{g}^{(\kappa)} \, \middle| \varphi_l^{(\kappa)}(t) \right\rangle , \tag{2.133}$$

where $\hat{g}^{(\kappa)}$ is a Hermitian, but otherwise arbitrary constraint operator acting only on the κ^{th} degree of freedom. To simplify the further notation the index J and configuration Φ_J are introduced as

$$A_J = A_{j_1 \ldots j_f} \quad \text{and} \quad \Phi_J = \prod_{\kappa=1}^{f} \varphi_{j_\kappa}^{(\kappa)}, \tag{2.134}$$

the projection operators on the space spanned by the SPFs of the κ^{th} degree of freedom are introduced as

$$\hat{P}^{(\kappa)} = \sum_{j=1}^{n_\kappa} \left| \varphi_j^{(\kappa)} \right\rangle \left\langle \varphi_j^{(\kappa)} \right| , \tag{2.135}$$

and the single hole function for the index $k_\kappa = l$ is introduced as

$$\Psi_l^{(\kappa)} = \sum_{j_1} \cdots \sum_{j_{\kappa-1}} \sum_{j_{\kappa+1}} \cdots \sum_{j_f} A_{j_1 \ldots j_{\kappa-1} \, l \, j_{\kappa+1} \ldots j_f} \varphi_{j_1}^{(1)} \ldots \varphi_{j_{\kappa-1}}^{(\kappa-1)} \varphi_{j_{\kappa+1}}^{(\kappa+1)} \ldots \varphi_{j_f}^{(f)}. \tag{2.136}$$

The functions (2.136) define the Hartree product from which the SPFs for the coordinate q_κ are missing.

The single hole functions define the mean field operators

$$\left\langle \hat{\mathbf{H}} \right\rangle_{jl}^{(\kappa)} = \left\langle \Psi_j^{(\kappa)} \middle| \, \hat{\mathbf{H}} \, \middle| \Psi_l^{(\kappa)} \right\rangle \tag{2.137}$$

and the density matrices

$$\rho_{jl}^{(\kappa)} = \left\langle \Psi_j^{(\kappa)} \middle| \, \Psi_l^{(\kappa)} \right\rangle . \tag{2.138}$$

$\left\langle \hat{\mathbf{H}} \right\rangle_{jl}^{(\kappa)}$ is an operator acting only on the $(\kappa)^{\text{th}}$ degree of freedom, and $\rho^{(\kappa)}$ is related to the reduced density matrix, as it describes a single degree of freedom, while the remaining

ones are removed by integrating. From the equation for the total MCTDH wavefunction (2.131) and the constraints put upon the SPFs within it, one can derive the equations of motion for the coefficients

$$i\hbar\dot{A}_J = \sum_L \langle\Psi_J|\,\hat{\mathbf{H}}\,|\Psi_L\rangle\, A_L - \sum_{\kappa=1}^{f}\sum_{l=1}^{n_\kappa} \hat{g}^{(\kappa)}_{j_\kappa l} A_{J_l^\kappa} \tag{2.139}$$

and for the SPFs themselves

$$i\hbar\dot{\varphi}^{(\kappa)} = g^{(\kappa)}\mathbf{1}_{n_\kappa}\varphi^{(\kappa)} + (1-\hat{P}^{(\kappa)})\left[(\rho^{(\kappa)})^{-1}\left\langle\hat{\mathbf{H}}\right\rangle^{(\kappa)} - g^{(\kappa)}\mathbf{1}_{n_\kappa}\right]\varphi^{(\kappa)}. \tag{2.140}$$

In these equations the SPFs are written as a vector, $\varphi^{(\kappa)} = (\varphi_1^{(\kappa)},\ldots,\varphi_{n_\kappa}^{(\kappa)})^T$, $\mathbf{1}_{n_\kappa}$ denotes the $n_\kappa \times n_\kappa$ unit matrix and the index J_l^κ denotes a composite index with the κ^{th} entry set to l. A detailed description of the derivation of the MCTDH method and its implementation for time dependent Hamiltonians is given in a recent extensive review [53, 108].

The choice of $g^{(\kappa)}$ is still arbitrary and can be used to change the final form of the equation of motion, normally it is simply set to $g^{(\kappa)} = 0$.

The reduction of the numerical effort can be estimated by an example: An f dimensional problem, represented on a grid with N points in each dimension requires N^f grid points for storage, and the handling of $(N^f) \times (N^f)$ matrices for the numerical solution. In the ideal TDH limit (decomposition into f 1d problems), this reduces to the solution of f problems with N grid points. Going to MCTDH with n SPFs per degree of freedom, one has to solve $f \cdot n$ 1d problems and handle an coefficient matrix of dimension n^f. As in general $n \ll N$ it is obvious, that the computational effort can be reduced by several orders of magnitude. At the same time, one can adjust the number of SPFs n to the desired quality of the calculation, moving from computationally cheap TDH to the more accurate MCTDH representation.

The CRS Hamiltonian and MCTDH

The MCTDH algorithm works by splitting a N–dimensional problem into a large number of low dimensional ones (preferably one dimensional). Then an approximation for the couplings between these degrees of freedom is used, which will function most effectively for problems with only weak coupling between these different DOF. If one takes a look at the CRS Hamiltonian defined in Section 2.2 by Eq. (2.37), one sees that this description of the system is quite suited for the treatment within the MCTDH approach. If the system is located in stationary reactant configuration, the total Hamiltonian will contain no couplings between the different DOF at all. Due to its construction, the kinetic terms are purely Cartesian and will introduce no interaction between the modes at all. The only terms coupling the different DOF are given by the force $f_k(\mathbf{x})$ acting on the k^{th} normal mode and the coefficient $K_{kl}(\mathbf{x})$, describing the interaction between the modes k and l. Both contributions are zero if the reaction coordinate is in its equilibrium position ($\mathbf{x} = 0$).

Within the CRS approximation, neither of these two coupling terms should get very large (if it does, it would mean that most likely the harmonic approximation used for this part of the potential is no longer valid, and this DOF should be moved into the set of reaction coordinates).

This therefore gives ideal conditions for the MCTDH method, delivering a potential which is already neatly separated into its normal modes as one dimensional DOF, with weak coupling between them. The only part of the potential which might not be suited to be treated within the MCTDH approach is the set of reaction coordinates, which by definition undergo large amplitude motions and can be strongly coupled to each other. So the normal approach will be to treat these coordinates exactly, as one low dimensional problem, coupled weakly within the MCTDH approach to all the other DOF described by the normal modes. As described in [53, 108], the multidimensional problem for the reaction surface can be handled by mode combinations, i.e. not decomposing the problem completely into 1d modes, but leaving some DOF coupled and represent them with multidimensional SPFs.

2.5 Control of Quantum Dynamics with Laser Pulses

In the field of quantum dynamics, it is one of the main goals to manipulate molecular systems directly [34, 110, 111]. Therefore a deeper theoretical understanding of methods to control the dynamics of quantum systems is necessary. The control methods are centered around the problem of finding a suitable external field, which drives the molecular system from an initial state into a desired final state. In this section the different approaches to find a form of this field appropriate for the control of a quantum mechanical system are presented shortly. Such fields are finally used to study the dynamics of protons in different molecules in Chapter 3.

2.5.1 Coupling to the Field

To allow the control of the molecular dynamics it is necessary to couple the molecular potential used to describe the system (e.g. by the CRS Hamiltonian, Eq. (2.37)) to an external electrical field, which interacts with the dipole moment of the molecule. For the usual applications in the area covered by this thesis (i.e. femtosecond spectroscopy, control of dynamics within molecules), this field is generated by a short laser pulse. Due to the high photon density effects of photon statistics can be neglected and the interaction can be treated in the so called "semiclassical dipole approximation". Within this approximation the additional interaction part of the potential is written as

$$\hat{\mathbf{H}}_{\text{field}}(\mathbf{x}, t) = \hat{\mu}(\mathbf{x}) \cdot \mathcal{E}(\mathbf{x}, t), \tag{2.141}$$

where $\hat{\mu}(\mathbf{x})$ is the dipole moment operator of the molecule and $\mathcal{E}(t)$ is the electric field part of the laser pulse. This electric field is described by

$$\mathcal{E}(\mathbf{x},t) = \sum_l \mathcal{E}_0^l \cdot s_l(t) \cdot \frac{e^{i(\mathbf{k}_l\mathbf{x}-\omega_l t)} + e^{-i(\mathbf{k}_l\mathbf{x}-\omega_l t)}}{2}, \tag{2.142}$$

where the index l numbers a sequence of laser pulses. For the l^{th} pulse, E_0^l is a vector describing the polarization and the maximum amplitude of the electric field, $s_l(t)$ is a shape function, which describes the slowly varying envelope of the pulse, ω_l is the frequency and $\mathbf{k}_l$ the wave vector of the radiation. The position dependence of the electrical field can be eliminated, if the size of the relevant system (e.g. a small molecule or a limited subsection of a larger one) is much smaller than the wavelength of the laser used. If this condition is fulfilled, the space–variable electric field can be replaced with a field constant over the extent of the molecule. As the shortest typical wavelength used normally are in the UV range, i.e. around 200 nm and the size of small molecule of the type studied in this work (i.e. system exhibiting proton dynamics and consisting of around 20 atoms) is on the order of 1 nm or smaller, the conditions for this approximation are taken as fulfilled for all systems in the following. An additional assumption necessary for the treatment of a system in condensed phase is, that the applied field does not interact significantly with the bath and does not interfere with the system bath coupling.

This leads to a coupling of the molecule to the laser pulse in the simplified, position independent form

$$\mathcal{E}(t) = \sum_l \mathcal{E}_0^l \cdot s_l(t) \cdot \cos(\omega_l t). \tag{2.143}$$

The goal of the theoretical control calculations is now, to find a set of shape functions $s_l(t)$ for the sequence of laser pulses with suitable frequencies ω_l, which drive the system from an initial state Φ_i into predefined target state Φ_f. This target state then describes a desired effect on the molecule, which one wants to observe experimentally, i.e. the breaking of a specific bond, the excitation of a specific spectroscopic state, or – treated in this work – the dynamics of a specific proton between a donor and acceptor location in a molecule.

Several ways to generate these pulses are described in the following sections.

2.5.2 Analytical Laser Pulses

The conceptually easiest approach to the problem is to generate the pulses via trial and error by hand. To do this it is necessary to calculate the vibrational eigenstates of the potential beforehand, as the transition frequencies between them will determine the necessary laser frequencies used for the control. For instance, if the goal is to drive an isomerization reaction described by a double minimum potential, as sketched in Fig. 2.3, one possible reaction path would be the pump–dump scheme developed by Paramonov and coworkers[112]

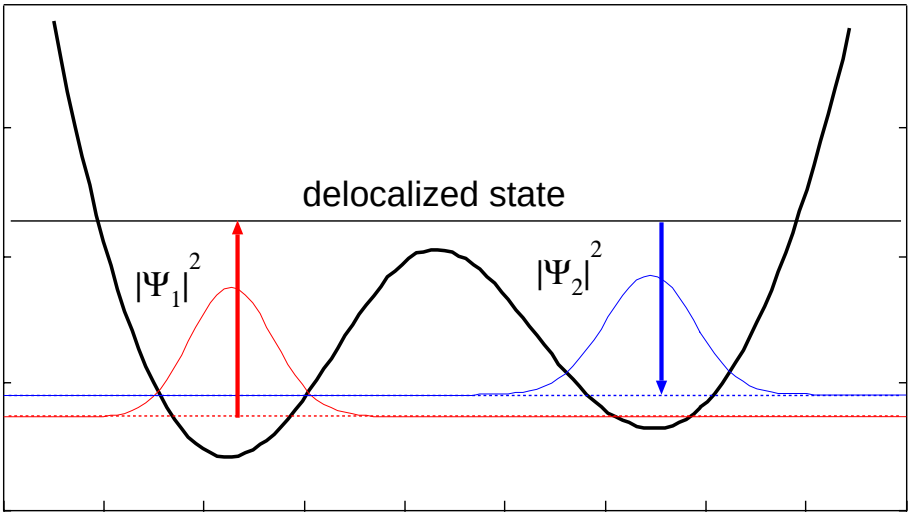

Figure 2.3: Schematic view of a pump–dump type reaction path in a double minimum potential, which could describe, e.g., a proton transfer reaction.

and Tannor and Rice [110, 113]. In this approach two short laser pulses of the form

$$\mathcal{E}_i(t) = \begin{cases} 0 & t < t_i \\ E_0^i \sin^2(\pi(t-t_i)/\tau_i)\cos(\omega_i t) & t_i \leq t \leq t_i + \tau_i \\ 0 & t > t_i + \tau_i \end{cases} , \tag{2.144}$$

where E_0^i is the amplitude, τ_i the duration, ω_i the frequency and t_i the starting time of the i^{th} pulse. The goal is to use the first pulse to transfer the population from the ground state of the system in the left well to a delocalized eigenstate above the barrier. The second pulse then depopulates this state into an eigenstate localized on the product side of the molecule. The state selectivity is reached by tuning the laser frequencies to the respective transition frequencies between the reactant and the delocalized state, and the delocalized and the product state. Before a suitable laser pulse can be constructed in this approach, it is necessary to fix a reaction path for the total transfer, which can only be done intuitively. For the greatest efficiency of the population transfer between the two states $|m\rangle$ and $|n\rangle$, one has to select pairs which have large values for the dipole matrix element $\langle m|\,\hat{\mu}\,|m\rangle$ coupling the two.

This approach of course is not limited to two pulse pump–dump schemes in the ground state, but can be extended to reaction paths involving high frequency transitions to some excited states or pathways requiring more than two pulses. The main drawback remains, as mentioned above, that the complete reaction path has to be guessed and fixed in advance, so the resulting laser pulse will normally either reproduce this exact dynamics or fail to achieve its goal altogether. This is in contrast to the genetic algorithms and optimal control methods presented later, where it is possible that the calculations find reaction paths not

thought of beforehand. The optimization of the pulses requires some experience as well, as the optimal laser frequency normally is slightly different from the exact transition frequency between initial and target energy level. This is caused by the interaction of the electric field with the diagonal elements of the Hamiltonian. More precisely, for the states $|m\rangle$ and $|n\rangle$ the dipole matrix element $\langle m|\,\hat{\mu}\,|n\rangle$ causes an interaction between the two levels, while the elements $\langle m|\,\hat{\mu}\,|m\rangle$ and $\langle n|\,\hat{\mu}\,|n\rangle$ will result in a shift of the energy of these levels. Due to the fact that $\langle m|\,\hat{\mu}\,|m\rangle \neq \langle n|\,\hat{\mu}\,|n\rangle$ this will result in a shift of the energy difference between these states. So for reaction paths containing more than two pulses the handcrafting of a good laser field can get quite tedious.

2.5.3 Genetic Algorithms

To avoid the hand optimization of a pulse sequence, it is possible to delegate this optimization work to a computer. This requires a parameter measuring the quality of the laser pulse, which in the simplest case would be the total population in the desired quantum mechanical target state, generated in dependence on all the possible parameters of the electric field. In the case of the $\sin^2$–pulses introduced in the previous section this would lead to a set of four parameters per pulse (namely the amplitude, the initial time, the duration, and the frequency). So a sequence of two pulses with a quite restrictive analytical form will already lead to a "quality–hypersurface" in eight dimensions, which has to be searched for a set of optimal parameters. Without previous knowledge of the shape of this surface, calculus based methods relying on the determination of its gradient can become quite unreliable. Schemes which follow the path of steepest ascent (hill–climbing) to find a maximum quality value only find the local maxima closest to the starting value. They also rely on the existence of well defined derivatives to generate the gradient. To avoid these requirements and limitations, the possibilities for using genetic algorithms for the optimization was tested in this work. This has been motivated by the fact, that recently in the group of Gerber [37] a genetic algorithm was successfully used to control an experimental setup. There the algorithm controlled, via a pulse shaper, the form of a femto-second laser pulse, and was able to optimize different branching ratios in a photo-chemical reaction. The experimental setup for this is sketched in Fig. 2.4.

Genetic algorithms are searching tools, based on the principle of random mutation and natural selection within a population of parameters describing the laser pulse. They differ in several ways from the more traditional optimization methods, like the gradient based methods mentioned above. The most prominent feature is, that they do require no auxiliary knowledge, like gradients, of the quality function. Another important feature is, that they do not optimize the function from a single point, but work by improving the overall quality of a population of points. In contrast to the calculus based methods a genetic algorithm will not use any deterministic method to generate a new set of parameters, but random mutations of the old parameters. This leads to a search algorithm, which is completely "blind" with respect to the shape of the quality hypersurface, and relies on

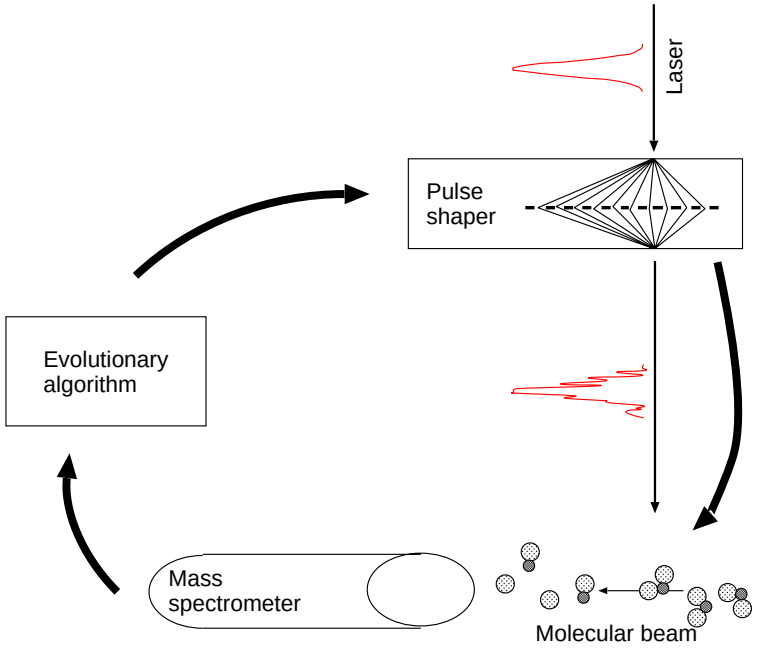

Figure 2.4: Schematic setup of a genetic algorithm in an experiment. Femtosecond laser pulses are modified by a computer controlled pulse shaper. These pulses then induce fragmentations in a molecular beam, which are recorded with a mass spectrometer. This signal is used to define the fitness of a shaped pulse and as a feedback to the algorithm.

random transitions to generate new search points. On the other hand the natural selection rules imposed on the total population of parameters assures, that there will be a general increase in the total "fitness" of the parameter population. A closer look at this approach is given in Appendix B and [114].

The genetic algorithms can be applied in two ways:

- One can define reaction paths for the dynamics by specifying only small parameter ranges, and then use the genetic algorithms only to optimize the parameters within these bounds. This still requires accurate knowledge of the system and an intuitive definition of the reaction path, but shifts the tedious work of finding the optimal set of parameters to the computer. With this approach it will also not be possible, to find new reaction pathways.

- The other approach is to allow the laser parameters to vary over the total range allowable for a physically sensible pulse. This leads to a much larger size of the parameter space, which requires a much greater effort in the search. The advantage here is, that it is not necessary to define any reaction path beforehand, and therefore in principle no prior knowledge of the system (like the transition frequencies) is

necessary. Allowing such a broad range of parameters can also generate pulses, which drive the system along a reaction pathway not considered beforehand.

The main drawback of the genetic algorithms is, that due to the randomness of the procedure it is possible to repeatedly sample the same region, which can add considerable computational overhead. Of course it is the same randomness, which prevents the algorithm from getting stuck on some local maximum of the quality surface, so this is a trade-off one has to live with.

2.5.4 Local Tracking Control

Contrary to the genetic algorithms presented in the previous section, the method of local control presented in this section and the optimal control method in the next try to calculate the field necessary for the transfer from the current to the desired state of the system analytically. The concept behind local control is conceptually simple, therefore it is presented first. In principle this approach can be based on the optimal control scheme, as shown by Fujimura in [115, 116]. The approach used here was first developed by Rabitz *et al.* [6, 117, 118, 119]. It requires the *a priori* knowledge of a reaction path the system should follow from its initial state to the desired final configuration. This path is specified as a time dependent path of a desired observable expectation value $\left\langle \hat{\mathbf{O}} \right\rangle_t = y_d(t)$. For this path a control field $E(t)$ is sought, which will exactly create the demanded dynamics of the chosen observable. If the reaction path for the observable is chosen reasonably, this can be achieved without resorting to an iterative scheme, by just using the actual position of the wavepacket at each time (therefore the term local – only the information from the current location of the wave packet is needed) to determine the necessary field for the next small time interval.

The Control Equations

The time evolution of an isolated molecular system interacting with an electrical field $\mathcal{E}(t)$ is given by the Schrödinger equation (2.41), with the Hamiltonian given by

$$\hat{\mathbf{H}}(t) = \hat{\mathbf{H}}_0 - \hat{\mu}(\mathbf{x}) \cdot \mathcal{E}(t) \tag{2.145}$$

in the semi-classical dipole approximation (cf. Section 2.5.1). The physical observables $y(t)$ measured in a laboratory are given by the expectation values of a Hermitian operator at time t:

$$y(t) = \left\langle \hat{\mathbf{O}} \right\rangle_t = \langle \psi(t) | \, \hat{\mathbf{O}} \, | \psi(t) \rangle \, . \tag{2.146}$$

From this it is possible to construct the inverse control solution for the electrical field $\mathcal{E}(t)$ by taking the time derivative of $y(t)$ and representing the time development of $\left\langle \hat{\mathbf{O}} \right\rangle$ with the Heisenberg equation of motion:

$$\frac{d \left\langle \hat{\mathbf{O}} \right\rangle}{dt} = \frac{1}{i\hbar} \left\langle \left[\hat{\mathbf{O}}, \hat{\mathbf{H}}_0 \right] \right\rangle - \frac{1}{i\hbar} \left\langle \left[\hat{\mathbf{O}}, \hat{\mu}(\mathbf{x}) \right] \right\rangle \cdot \mathcal{E}(t) + \left\langle \frac{\partial \hat{\mathbf{O}}}{\partial t} \right\rangle \tag{2.147}$$

As written, this would have to be solved for the three components of $\mathcal{E}(t)$, resulting in three operators for the x, y and z direction, but more commonly, the field is taken as being linearly polarized along a chosen direction, and only one solution for the scalar field $E(t)$ in this direction is sought. To generate the inverse control solution, this equation has to be solved for $E(t)$, while the term $d\left\langle\hat{\mathbf{O}}\right\rangle/dt$ is replaced with the time derivative of the desired trajectory dy_d/dt. Of course the relation is not invertible, when the commutator $\left[\hat{\mathbf{O}},\hat{\mu}(\mathbf{x})\right]$ is zero. To allow for this, one can continue to differentiate to higher orders, until a non–zero commutator for the i^{th} time derivative of the operator $\hat{\mathbf{O}}$ appears. The higher order equations of motion then are given recursively as

$$\frac{d^{i+1}\left\langle\hat{\mathbf{O}}\right\rangle}{dt^{i+1}} = \frac{1}{i\hbar}\left\langle\left[\hat{\mathbf{O}}_i,\hat{\mathbf{H}}_0\right]\right\rangle - \frac{1}{i\hbar}\left\langle\left[\hat{\mathbf{O}}_i,\hat{\mu}(\mathbf{x})\right]\right\rangle\cdot\mathcal{E}(t) + \left\langle\frac{\partial\hat{\mathbf{O}}_i}{\partial t}\right\rangle \tag{2.148}$$

with

$$\hat{\mathbf{O}}_{i+1} = \frac{1}{i\hbar}\left[\hat{\mathbf{O}}_i,\hat{\mathbf{H}}_0\right] - \frac{1}{i\hbar}\left[\hat{\mathbf{O}}_i,\hat{\mu}(\mathbf{x})\right]\cdot\mathcal{E}(t) + \frac{\partial\hat{\mathbf{O}}_i}{\partial t}. \tag{2.149}$$

If a suitable order with non–zero commutator $\left[\hat{\mathbf{O}}_k,\hat{\mu}(\mathbf{x})\right]$ is found at order k, the equation is solved for the field to obtain

$$\mathcal{E}(t) = \frac{-i\hbar\frac{d^{k+1}y_d(t)}{dt^{k+1}} + \left\langle\left[\hat{\mathbf{O}}_k,\hat{\mathbf{H}}_0\right]\right\rangle + i\hbar\left\langle\frac{\partial\hat{\mathbf{O}}_k}{\partial t}\right\rangle}{\left\langle\left[\hat{\mathbf{O}}_k,\hat{\mu}(\mathbf{x})\right]\right\rangle}. \tag{2.150}$$

From the equation follows, that the control field at time t depends on the wavefunction $\psi(t)$ of the system, via the expectation values of the commutators. This dependence gives a feedback from the system, and results in a control field, which not only determines the dynamics of the system, but also depends itself on it. The resulting nonlinear differential equations have to be solved numerically.

In a first step, at the initial time t_0, the control field $E(t_0)$ is generated via Eq. (2.150). Using this field, the system is propagated for a small time interval Δt to generate the new wavefunction $\psi(t_1)$, which then is used to calculate the new control field $E(t_1)$. This procedure continues until a preset final time t_N is reached.

Error Control

The implementation above assumes, that the electric field generated by the algorithm will drive the field along the desired path without deviations, i.e. the error function $e(t) = y_d(t) - \left\langle\hat{\mathbf{O}}\right\rangle_t$ is zero throughout the control period. This will normally only be possible in idealized cases, when the wavepacket stays perfectly localized and the field can instantaneously take any desired value. In reality, the desired path $y_d(t)$ and the real path described by the expectation value $\left\langle\hat{\mathbf{O}}\right\rangle_t$ will be different. This problem can be dealt with,

by adding a term to Eq. (2.150), which will asymptotically damp the error:

$$\mathcal{E}(t) = \frac{1}{\left\langle \left[\hat{\mathbf{O}}_k, \hat{\mu}(\mathbf{x})\right]\right\rangle}\left[-i\hbar\frac{d^{k+1}y_d(t)}{dt^{k+1}} + \left\langle \left[\hat{\mathbf{O}}_k, \hat{\mathbf{H}}_0\right]\right\rangle + i\hbar\left\langle\frac{\partial\hat{\mathbf{O}}_k}{\partial t}\right\rangle + \sum_{j=0}^{k}\rho_j\frac{d^j e(t)}{dt^j}\right]. \quad (2.151)$$

This is equivalent to the condition

$$\sum_{j=0}^{k}\rho_j\frac{d^j e(t)}{dt^j} = 0 \quad (2.152)$$

with $\rho_j = 1$. In this formalism the term containing the derivatives of $e(t)$ will damp the error in all differential orders exponentially to zero. The parameter ρ_j defines the speed of the decay, and can be freely chosen – high values for ρ_j will result in a faster decay of the error, but also generally in higher fields [118].

Using this approach, it is also possible to try to control multiple objectives at the same time. In a system of N different objectives, one defines N desired tracks, $e_l(t)$, $l = 1, \ldots N$ as the deviation from the l^{th} track and a set of relative weights W_l, which give the relative importance of the l^{th} observable. With this one can define a cost function

$$\mathcal{I} = \sum_{l=1}^{N} W_l\left[\sum_{j=0}^{k_l}\rho_j\frac{d^j e_l(t)}{dt^j}\right]^2 + W_{\mathcal{E}}\mathcal{E}^2(t) \quad (2.153)$$

which sums up all the deviations of the observables from their desired paths (the term $W_{\mathcal{E}}\mathcal{E}^2(t)$ is added to the functional to penalize high fields). The optimal field then minimizes this cost function at all times.

The advantages of this method are:

- it is conceptually simple,
- it is easy to implement, for equations of low order
- and, because no optimization or searching is required, it is very fast to calculate.

The most important difference to the feedback control driven by genetic algorithms presented in Section 2.5.3 is, that now some information about the actual state of the system is used. While the genetic optimization is completely blind concerning the reaction path and only sees the final state of the system, here the algorithm tries to predict the optimal field needed to get to this result from the current state of the wavefunction.

On the other hand, there are also some drawbacks:

- it is necessary to fix an *a priori* reaction path, which the system has to follow all the time; in addition, this reaction path has to be "reasonable", i.e. a pathway which can be followed without having to apply an unreasonably strong field,

- as the calculation of the control field relies on the expectation values of some operators, the method will work best for tightly localized wavepackets; it starts to break down for systems resulting in widely spread or split wavepackets, where the expectation value of the operator is no longer close to the point of maximum probability density of the wavepacket.

Especially the last limitation can be difficult to correct for, as most systems of interest are quite anharmonic, or contain barriers, which easily lead to spreading or splitting of wavepackets.

2.5.5 Optimal Control

The method of optimal control uses a variational scheme on an appropriate functional to generate a laser field which drives the quantum system from a known initial state $\phi_i(0) = \Psi_i$ at time $t = 0$ into a desired final state $\phi_f(T) = \Psi_f$ at time $t = T$. The main difference to the other methods is, that with optimal control it is not necessary to define any sort of reaction path for the system. The algorithm just requires the initial and final state, and then searches for the optimal reaction path itself.

The method, developed mainly by Rabitz and coworkers [64, 65], Fujimura *et al.* [115, 120] and Yan [36], is based on the analytic maximization of the so called "objective functional", which is generated from the total overlap of the initial wavefunction, propagated with the calculated field to the final time, with the final state. It is assumed, that the system can be described by the Schrödinger equation in semi-classical dipole approximation (cf. Section 2.5.4). A basic functional to be maximized is

$$J(\Psi_i, \Psi_f, \mathcal{E}(t)) = |\langle \phi_i(T)|\ \Psi_f\rangle|^2 - \alpha \int_0^T |\mathcal{E}(t)|^2\ dt. \tag{2.154}$$

The second term on the right hand side is added to minimize the the effect of high total energy of the laser pulse, i.e. to avoid generating theoretical pulses, which would ionize a molecule in real life. The parameter α is introduced as a penalty factor, which determines how much a high energy field will be noticed by the algorithm. This factor α is made time dependent ($\alpha = \alpha(t) = \frac{\alpha_0}{s(t)}$), allowing to shape the resulting laser pulses according to the shape function $s(t)$. When this function goes to zero the penalty factor approaches infinity and therefore the allowed laser field goes to zero [66].

To ensure, that the Schrödinger equation is fulfilled by $\phi_i(t)$ at all times, an additional term, which can be interpreted as a Lagrange type of constraint (see, e.g. [66]), is added to the functional, resulting in

$$\begin{aligned} &J(\Psi_i, \Psi_f, \mathcal{E}(t)) = \\ &\quad |\langle \phi_i(T)|\ \Psi_f\rangle|^2 - \alpha \int_0^T |\mathcal{E}(t)|^2\ dt - \\ &\quad 2\Re\left\{\langle \phi_i(T)|\ \Psi_f\rangle \int_0^T \langle \phi_f(t)|\, \frac{i}{\hbar}\left[\hat{\mathbf{H}}_0 - \hat{\mu}(\mathbf{x})\cdot\mathcal{E}(t)\right] + \frac{\partial}{\partial t}\,|\phi_i(t)\rangle\ dt\right\}. \end{aligned} \tag{2.155}$$

The variation of this functional then gives a system of coupled equations for the control field [66]:

$$\mathcal{E}(t) = -\frac{s(t)}{\hbar\alpha_0} \cdot \Im\left\{\langle\phi_i(t)|\ \phi_f(t)\rangle \langle\phi_f(t)|\,\hat{\mu}\,|\phi_i(t)\rangle\right\} \tag{2.156}$$

$$i\hbar\frac{\partial}{\partial t}\phi_i(t) = \left[\hat{\mathbf{H}}_0 - \hat{\mu}(\mathbf{x})\cdot\mathcal{E}(t)\right]\phi_i(t); \qquad \phi_i(0) = \Psi_i \tag{2.157}$$

$$i\hbar\frac{\partial}{\partial t}\phi_f(t) = \left[\hat{\mathbf{H}}_0 - \hat{\mu}(\mathbf{x})\cdot\mathcal{E}(t)\right]\phi_f(t); \qquad \phi_f(T) = \Psi_f \tag{2.158}$$

This set describes the forward propagation in time of Ψ_i and the backward propagation in time of Ψ_f with a field calculated from, and therefore coupled to, these wavefunction dynamics. It can be solved with a self consistent field method by choosing an initial field, doing the forward and backward propagations, calculate the new field and repeat the process until the field converges to a fixed value.

The big advantage of this method is, as already mentioned above, that it does only require the desired initial and final states, and no further assumptions for the reaction path have to be made. This allows the algorithm to find unexpected transition, of which no one thought before, as will be shown in Section 3.1.1. The main difference of this concept to the local control presented in Section 2.5.4 is, that no *a priory* reaction path has to be specified. Instead a global cost function is minimized, which contains no information about the actual path traveled on the potential hypersurface. In contrast to the genetic algorithms, which also need no information about the reaction path, now some analytical constraints resulting from this cost functional restrict the dynamics of the system.

This method requires significant computer resources, as it is not only iterative but also needs a forward and backward time propagation for each step. The necessary backward time propagation is also a problem in dissipative systems, where the system dynamics is not time reversible, due to the coupling to the bath. Methods for the solution of the optimal control problem resulting from the dissipative Liouville–von Neumann equation (2.110), are developed by the groups of Wilson [121], Rabitz [122] and May [123]. It should also be noted that this approach will only converge to a local maximum of the functional, which is closest in parameter–space to the initial condition. There is no random element, like the random mutations of the genetic algorithms, which could allow this analytical method to escape from one of these local maxima to a possibly better global one.