Wave packet dynamics studied by ab initio methods:

Applications to strong-field ionization of atoms and molecules

Thomas Kim Kjeldsen

Department of Physics and Astronomy
University of Århus

PhD Thesis

August 2007

Contents

Preface

This thesis is submitted to the Faculty of Science at the University of Aarhus, Denmark, in order to fulfil the requirements for obtaining the PhD degree in Physics. The studies have been carried out under the supervision of Associate Professor Lars Bojer Madsen at the Department of Physics and Astronomy, University of Aarhus, from August 2003 to July 2007.

In this thesis, I have taken the opportunity to describe some techniques that are used to solve time dependent quantum mechanical problems. This subject is usually not suited for publication as regular physics papers. The broad readership within the physics community tends to find it particularly interesting to interpret the results that can be inferred from the final solution. The detailed knowledge of *how* to obtain the solution is, on the other hand, mostly relevant only for a few number of scientists working in the field of computational physics. With this thesis I hope to shed light on various problems that computational physicists face nowadays.

Notational conventions

Atomic units ($m_e = |e| = \hbar = 1$) will be used throughout this thesis unless stated otherwise.

Vectors and matrices will generally be zero-based (c-style convention), so that the elements of an N dimensional vector are enumerated between 0 and $N-1$.

Acknowledgements

Thomas Kjeldsen, July, 2007

List of publications

1. T. K. Kjeldsen, C. Z. Bisgaard, L. B. Madsen, and H. Stapelfeldt
   Role of symmetry in strong-field ionization of molecules
   Phys. Rev. A **68**, 063407 (2003)

2. T. K. Kjeldsen and L. B. Madsen
   Strong field ionization of N_2: length and velocity gauge strong field approximation and tunnelling theory
   J. Phys. B **37**, 2033 (2004)

3. T. K. Kjeldsen, C. Z. Bisgaard, L. B. Madsen, and H. Stapelfeldt
   Influence of molecular symmetry on strong-field ionization: Studies on ethylene, benzene, fluorobenzene, and chlorofluorobenzene
   Phys. Rev. A **71**, 013418 (2005)

4. T. K. Kjeldsen and L. B. Madsen
   Strong-field ionization of diatomic molecules and companion atoms: Strong-field approximation and tunneling theory including nuclear motion
   Phys. Rev. A **71**, 023411 (2005)

5. T. K. Kjeldsen and L. B. Madsen
   Vibrational Excitation of Diatomic Molecular Ions in Strong Field Ionization of Diatomic Molecules
   Phys. Rev. Lett. **95**, 073004 (2005)

6. T. K. Kjeldsen and L. B. Madsen
   Comment on "Strong-field ionization of laser-irradiated light homonuclear diatomic molecules: A generalized strong-field approximation – linear combination of atomic orbitals model"
   Phys. Rev. A **73**, 047401 (2006)

7. T. K. Kjeldsen and L. B. Madsen
   Strong-field ionization of atoms and molecules: The two-term saddle-point method
   Phys. Rev. A **74**, 023407 (2006)

8. T. K. Kjeldsen, L. B. Madsen, and J. P. Hansen
   Ab initio studies of strong-field ionization of arbitrarily oriented $H_2{}^+$ molecules
   Phys. Rev. A **74**, 035402 (2006)

9. T. K. Kjeldsen and L. B. Madsen
   Alignment dependent above threshold ionization of molecules
   J. Phys. B **40**, 237 (2007)

10. V. N. Ostrovsky, T. K. Kjeldsen, and L. B. Madsen
    Comment on "Generalization of Keldysh's theory"
    Phys. Rev. A **75**, 027401 (2007)

11. L. A. A. Nikolopoulos, T. K. Kjeldsen, and L. B. Madsen
    Spectral and partial-wave decomposition of time-dependent wave functions on a grid: Photoelectron spectra of H and H_2^+ in electromagnetic fields
    Phys. Rev. A **75**, 063426 (2007)

12. T. K. Kjeldsen, L. A. A. Nikolopoulos, and L. B. Madsen
    Solving the m-mixing problem for the three-dimensional time-dependent Schrödinger equation by rotations: application to strong-field ionization of H_2^+
    Phys. Rev. A **75**, 063427 (2007)

13. L. A. A. Nikolopoulos, T. K. Kjeldsen, and L. B. Madsen
    Three-dimensional time-dependent Hartree-Fock approach for arbitrarily oriented molecular hydrogen in strong electromagnetic fields
    Phys. Rev. A **76**, 033402 (2007)

14. C. B. Madsen, A. S. Mouritzen, T. K. Kjeldsen, and L. B. Madsen
    Effects of orientation and alignment in high-harmonic generation and above threshold ionization
    Phys. Rev. A **76**, 035401 (2007)

15. M. Førre, S. Selstø, J. P. Hansen, T. K. Kjeldsen, and L. B. Madsen
    Molecules in intense XUV pulses: Beyond the dipole approximation in linearly and circularly polarized fields
    Phys. Rev. A **76**, 035415 (2007)

16. L. B. Madsen, L. A. A. Nikolopoulos, T. K. Kjeldsen, and J. F. Hernandez
    Extracting continuum information from $|\Psi(t)\rangle$ in time-dependent wave packet calculations
    Submitted for publication

Chapter 1

Introduction

In classical mechanics, the term *dynamics* refers to how forces act on particles and influence their motion as time passes. If all forces are known in addition to initial conditions at some instant of time, the subsequent motion can be predicted by solving Newton's equation of motion. The predictive power provided by the solution this equation of is extremely valuable for a wide range of phenomena encountered in daily life.

In the quantum mechanical world, the Newtonian description fails. There is, however, another equation of motion that in some sense determines the dynamics of particles. For non-relativistic quantum mechanics, the equation of motion is known as the time-dependent Schrödinger equation. It would be desirable to be able to predict the evolution of a quantum system with the same accuracy as is known from the classical world. The quantum world is, however, fundamentally different from the classical world, since quantum mechanics is not deterministic. More precisely, the *dynamics* governed by the Schrödinger equation *is* deterministic but the process of measurement is not. If two quantum systems are prepared identically, they are not guaranteed to give the same answer upon a certain measurement. The solution of the Schrödinger equation only provides an object that determines a probability amplitude distribution for the outcome of a measurement. Apart from the aforementioned subtlety regarding predictive power, it will still be an important task actually to solve the time-dependent Schrödinger equation. The solution of this problem, which turns out to be a much more difficult problem than its classical Newtonian counterpart, will be the main subject of the present work.

Predictive power is essentially about predicting how a system responds to some disturbance. Even more interesting is the inverse problem, which is known as quantum control. This subject concerns the question: How should a system be subjected to forces that we can control externally in order to end up in a certain final state? Quantum systems that can be controlled accurately can be used to demonstrate truly amazing phenomena that can

never be seen in a classical system. Again, it will be necessary to solve the time-dependent Schrödinger equation in order to describe the control aspect theoretically.

The present work will concern a specific class of quantum mechanical problems, namely the interaction between light, and atoms or molecules. For this situation, the Schrödinger equation can be solved approximately in some cases by perturbative methods. Such methods have been used extensively to describe the interaction between atoms and low intensity radiation fields. On the other hand when an atom is exposed to a high intensity field a series of non-perturbative phenomena may occur. In such cases the only reliable method to fully describe the dynamics is to solve the Schrödinger equation. In the present context an intense laser field is characterized by a field strength which is comparable to the Coulomb interaction between the electrons and the nuclei. The typical intensities for this to be the case are $10^{13} - 10^{15}\,\mathrm{Wcm}^{-2}$, values that are readily produced by commercial laser systems.

The time-dependent Schrödinger equation is unfortunately extremely difficult to solve for a many-particle system. The work presented in this thesis will only treat the dynamics of a single electron. Hence, the only true physical systems that can be described exactly are the hydrogen atom and the hydrogen-like ions. For many purposes it is, however, a reasonable approximation to adopt the so called single active electron model. In this model, we solve the time dependent Schrödinger equation exactly for a single electron while all other degrees of freedom are frozen and are accounted for by some mean field action. Within the single active electron model, the present one-electron description can therefore be applied as an approximation to a much wider range of systems than just the hydrogen atom.

Chapter 2

Dynamics of charged particles in electromagnetic fields

All forces encountered in atomic, molecular and optical physics are of electromagnetic nature. Common examples include the electrostatic binding between positively charged nuclei and negatively charged electrons as well as the interaction between electrons and external radiation sources. Hence, it is necessary to set up the theoretical framework that describes interaction between particles and electromagnetic fields. This chapter describes the dynamics of a charged particle in an electromagnetic field both from a classical and quantum mechanical point of view. Both cases are restricted to the non-relativistic case. Although we often imagine that the particle under consideration is the electron, this chapter treats an arbitrary particle with charge q and mass M.

2.1 The electromagnetic fields

This section briefly reviews the properties of the classical electromagnetic fields. Of particular relevance for the general subject of the thesis is the introduction of the gauge transformations.

Throughout the thesis, it is assumed that all fields can be treated classically.

The electric and magnetic fields are denoted $\boldsymbol{E}(\boldsymbol{r}, t)$ and $\boldsymbol{B}(\boldsymbol{r}, t)$, respectively. Generally, both fields are dependent on space and time. Now we consider a particle with charge q which is located at $\boldsymbol{r}$ with velocity $\dot{\boldsymbol{r}} = d\boldsymbol{r}/dt$ at time t. This particle will be subject to the Lorentz force

$$\boldsymbol{F}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = q[\boldsymbol{E}(\boldsymbol{r}, t) + \dot{\boldsymbol{r}} \times \boldsymbol{B}(\boldsymbol{r}, t)]. \tag{2.1}$$

The fields themselves are created by other charges and currents according to the Maxwell equations.

It turns out that the structure of the Maxwell equations opens the possibility to express the fields in terms of potentials. The reformulation in terms of potentials is particularly important in Lagrangian and Hamiltonian formulations of mechanics as well as in quantum mechanics. First, recall the two source-free Maxwell equations

$$\boldsymbol{\nabla} \cdot \boldsymbol{B}(\boldsymbol{r},t) \;=\; 0, \tag{2.2}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r},t) \;=\; -\frac{\partial}{\partial t}\boldsymbol{B}(\boldsymbol{r},t). \tag{2.3}$$

The Maxwell equation (2.2) allows us to introduce a vector potential $\boldsymbol{A}(\boldsymbol{r},t)$ such that

$$\boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r},t). \tag{2.4}$$

From the second source-free Maxwell equation, we now obtain

$$\boldsymbol{\nabla} \times \left(\boldsymbol{E}(\boldsymbol{r},t) + \frac{\partial}{\partial t}\boldsymbol{A}(\boldsymbol{r},t) \right) = 0. \tag{2.5}$$

Since the vector field in the parentheses is irrotational, it can be written as the negative gradient of a scalar function which we call the scalar potential and denote $\Phi(\boldsymbol{r},t)$

$$\boldsymbol{E}(\boldsymbol{r},t) + \frac{\partial}{\partial t}\boldsymbol{A}(\boldsymbol{r},t) = -\boldsymbol{\nabla}\Phi(\boldsymbol{r},t). \tag{2.6}$$

Here, we are not concerned about how to *find* the potentials. Instead we pay special attention to the ambiguity that lies in the definition of the potentials. First note that the potentials are purely mathematical quantities. The *fields*, $\boldsymbol{E}$ and $\boldsymbol{B}$, are the physical objects, since they exert forces on charged particles that can be measured. Therefore we have the freedom to adjust the potentials as long as the fields are unchanged. Such an adjustment is referred to as a gauge transformation.

In order to see how gauge transformations work in electrodynamics, we first assume that we have found some potentials Φ_0 and $\boldsymbol{A}_0$ that produce the correct fields. Now we see from Eq. (2.4) that it is the curl of $\boldsymbol{A}$ that is well defined, so that $\boldsymbol{A}$ itself is only determined up to an addition of the gradient of any scalar function. We see that such a transformation is valid, for if we let

$$\boldsymbol{A}_1(\boldsymbol{r},t) = \boldsymbol{A}_0(\boldsymbol{r},t) + \boldsymbol{\nabla}\Lambda(\boldsymbol{r},t), \tag{2.7}$$

for any function of space and time $\Lambda(\boldsymbol{r},t)$ then the magnetic field remains unchanged. The change in $\boldsymbol{A}(\boldsymbol{r},t)$ will, however, change Eq. (2.6) by an additional factor of $\boldsymbol{\nabla}\partial\Lambda(\boldsymbol{r},t)/\partial t$ on the left hand side. Fortunately, we also have the freedom to change Φ as long as Eq. (2.6) produces the correct electric field. Thus, we accompany the aforementioned change in $\boldsymbol{A}$ by the change

$$\Phi_1(\boldsymbol{r},t) = \Phi_0(\boldsymbol{r},t) - \frac{\partial\Lambda(\boldsymbol{r},t)}{\partial t}, \tag{2.8}$$

which ensures that the electric field in Eq. (2.6) remains the same.

A final point to mention is the ambiguity in choice of units for the fields. We adopt the convention from the electrostatic units (ESU) where the magnetic field and the vector potential both are a factor of c smaller than the Gaussian unit based counterparts. In the literature, there is no strict convention of how to define the atomic system of units for electromagnetic fields, i.e., whether the ESU or the Gaussian convention should be adopted. The Gaussian convention is obtained by changing $\boldsymbol{B} \to \boldsymbol{B}/c$ and $\boldsymbol{A} \to \boldsymbol{A}/c$ everywhere in this chapter.

2.2   A classical particle in an electromagnetic field

2.2.1   The Lagrangian formulation

This section formulates the dynamics of a charged particle within the framework of the Lagrangian.

We consider a particle of mass M in three spatial dimensions characterized by the Cartesian components[1] $\boldsymbol{r} = (x, y, z) = (x_1, x_2, x_3)$. We assume that there are no constraints between the coordinates, i.e., the particle is not fixed to some particular trajectory. Then we can take the Cartesian components as the generalized coordinates of the particle. The corresponding generalized velocities are equal to the usual kinetic velocities $\dot{x}_i = dx_i/dt$, and the kinetic energy reads

$$T(\dot{\boldsymbol{r}}) = \sum_{i=1}^{3} \frac{1}{2} M \dot{x}_i^2. \tag{2.9}$$

Now we let the force $\boldsymbol{F}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t)$, act on the particle and write Newton's equation of motion for each Cartesian component

$$M \ddot{x}_i = \frac{d}{dt} \frac{\partial T}{\partial \dot{x}_i} = F_i(\boldsymbol{r}, \dot{\boldsymbol{r}}, t). \tag{2.10}$$

In the Lagrangian formulation, we abandon the concept of forces and try to rewrite the Newtonian equation in terms of energy functions instead. It is, naturally, a requirement that this alternative formulation must be equivalent to Newton's equation.

First, we consider the case where the force is static and conservative. Such forces only depend on the coordinates and can be written as the gradient of some scalar function of the coordinates

$$F_i(\boldsymbol{r}) = -\frac{\partial V(\boldsymbol{r})}{\partial x_i}. \tag{2.11}$$

---

[1]Vectors will be one-based throughout this section with the three components enumerated from one to three. It would perhaps be more consistent to use the symbol $\boldsymbol{x}$ for the coordinate vector. $\boldsymbol{r}$ is, however, the standard notation.

The function $V(\boldsymbol{r})$ is called the potential corresponding to the force. Inserting this expression into Eq. (2.10) gives

$$\frac{d}{dt}\frac{\partial T(\dot{\boldsymbol{r}})}{\partial \dot{x}_i} + \frac{\partial V(\boldsymbol{r})}{\partial x_i} = 0, \tag{2.12}$$

or since $\partial T/\partial x_i = 0$ and $\partial V/\partial \dot{x}_i = 0$

$$\frac{d}{dt}\frac{\partial (T-V)}{\partial \dot{x}_i} - \frac{\partial (T-V)}{\partial x_i} = \left(\frac{d}{dt}\frac{\partial}{\partial \dot{x}_i} - \frac{\partial}{\partial x_i}\right) L(\boldsymbol{r},\dot{\boldsymbol{r}}) = 0, \tag{2.13}$$

where we introduced the *Lagrangian*

$$L(\boldsymbol{r},\dot{\boldsymbol{r}}) = T(\dot{\boldsymbol{r}}) - V(\boldsymbol{r}). \tag{2.14}$$

Equations (2.13) are called the Lagrange equations for conservative forces and are equivalent to Eq. (2.10). The whole dynamics of the particle can be determined from the coupled differential equations that appear from Eq. (2.13).

For non-conservative forces we cannot write the force as simply as in Eq. (2.11). Examples of non-conservative forces include those encountered in electrodynamics (except for the purely electrostatic case), wherefore the treatment of such forces must be properly accounted for in the present context. It turns out that if we are able to find some function $V(\boldsymbol{r},\dot{\boldsymbol{r}},t)$ such that

$$F_i(\boldsymbol{r},\dot{\boldsymbol{r}},t) = \frac{d}{dt}\frac{\partial V}{\partial \dot{x}_i} - \frac{\partial V}{\partial x_i}, \tag{2.15}$$

then Eq. (2.13) is still equivalent to Newton's equation of motion. This proposal is easily verified by direct substitution in Eq. (2.10). The difference between the non-conservative and the conservative case is that the Lagrangian now depends explicitly on time

$$L(\boldsymbol{r},\dot{\boldsymbol{r}},t) = T(\dot{\boldsymbol{r}}) - V(\boldsymbol{r},\dot{\boldsymbol{r}},t). \tag{2.16}$$

The Lagrangian for a particle of charge q in the electric and magnetic fields $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$ evidently requires a function that solves Eq. (2.15) with the Lorentz force on the left hand side, i.e.

$$F_i(\boldsymbol{r},\dot{\boldsymbol{r}},t) = qE_i(\boldsymbol{r},t) + q[\dot{\boldsymbol{r}} \times \boldsymbol{B}(\boldsymbol{r},t)]_i = \frac{d}{dt}\frac{\partial V}{\partial \dot{x}_i} - \frac{\partial V}{\partial x_i}. \tag{2.17}$$

Now we claim that the function

$$V(\boldsymbol{r},\dot{\boldsymbol{r}},t) = q[\Phi(\boldsymbol{r},t) - \dot{\boldsymbol{r}} \cdot \boldsymbol{A}(\boldsymbol{r},t)], \tag{2.18}$$

does the job to satisfy Eq. (2.17). For this choice, the right hand side of Eq. (2.17) is

$$
\begin{aligned}
\frac{d}{dt}\frac{\partial V}{\partial \dot{x}_i} - \frac{\partial V}{\partial x_i} &= -q\frac{d}{dt}A_i(\boldsymbol{r},t) - q\frac{\partial}{\partial x_i}\Phi(\boldsymbol{r},t) + q\sum_{j=1}^{3}\dot{x}_j\frac{\partial}{\partial x_i}A_j(\boldsymbol{r},t) \\
&= -q\sum_{j=1}^{3}\frac{\partial A_i}{\partial x_j}\dot{x}_j - q\frac{\partial A_i}{\partial t} - q\frac{\partial \Phi}{\partial x_i} + q\sum_{j=1}^{3}\dot{x}_j\frac{\partial A_j}{\partial x_i} \\
&= q\left(-\frac{\partial A_i}{\partial t} - \frac{\partial \Phi}{\partial x_i}\right) + q\left[\dot{\boldsymbol{r}} \times \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r},t)\right]_i \\
&= qE_i(\boldsymbol{r},t) + q\left[\dot{\boldsymbol{r}} \times \boldsymbol{B}(\boldsymbol{r},t)\right]_i .
\end{aligned}
\tag{2.19}
$$

The term containing the curl of $\boldsymbol{A}$ is verified by straightforward algebraic manipulations. Note that in the time derivative of $\boldsymbol{A}$, we must take into account that the vector potential changes partly because the particle moves and partly because $\boldsymbol{A}$ itself changes.

In conclusion, as we use the correct choice for the function V, we arrive at the Lagrangian for a charged particle in an electromagnetic field

$$
L(\boldsymbol{r},\dot{\boldsymbol{r}},t) = \frac{1}{2}M\dot{\boldsymbol{r}}^2 - q[\Phi(\boldsymbol{r},t) - \dot{\boldsymbol{r}} \cdot \boldsymbol{A}(\boldsymbol{r},t)].
\tag{2.20}
$$

2.2.2 The Hamiltonian formulation

The Hamiltonian formulation is obtained from the Lagrangian by introducing the canonical conjugate momentum

$$
p_i = \frac{\partial}{\partial \dot{x}_i}L(\boldsymbol{r},\dot{\boldsymbol{r}},t).
\tag{2.21}
$$

This relation expresses the canonical momentum $\boldsymbol{p}$ in terms of the generalized velocities $\dot{\boldsymbol{r}}$. Conversely, we can regard Eq. (2.21) as an implicit equation that allows us to express $\dot{\boldsymbol{r}}$ in terms of $\boldsymbol{p}, \boldsymbol{r}$ and t. Having established such a relationship, we define the Hamiltonian function

$$
H(\boldsymbol{r},\boldsymbol{p},t) = \sum_{i=1}^{3} p_i\dot{x}_i(\boldsymbol{r},\boldsymbol{p}) - L(\boldsymbol{r},\dot{\boldsymbol{r}}(\boldsymbol{r},\boldsymbol{p}),t).
\tag{2.22}
$$

Differentiation of this function with respect to $\boldsymbol{r}$ and $\boldsymbol{p}$ together with the Lagrange equation Eq. (2.13) gives the Hamiltonian equations

$$
\frac{d}{dt}x_i = \frac{\partial}{\partial p_i}H(\boldsymbol{r},\boldsymbol{p},t),
\tag{2.23}
$$

$$
\frac{d}{dt}p_i = -\frac{\partial}{\partial x_i}H(\boldsymbol{r},\boldsymbol{p},t).
\tag{2.24}
$$

Therefore, knowledge of the Hamiltonian function leads to a set of coupled differential equations which determine the time evolution of the coordinates of the particle. Again, these equations are equivalent to Newton's equation of motion for the correct choice of the function V in Eq. (2.15) that enters the Lagrangian.

For the case of the charged particle in an electromagnetic field, the Lagrangian in Eq. (2.20) leads to the canonical momentum

$$
\begin{aligned}
p_i &= \frac{\partial}{\partial \dot{x}_i} \left(\frac{1}{2} M \sum_{j=1}^{3} \dot{x}_j{}^2 - q[\Phi(\boldsymbol{r}, t) - \sum_{j=1}^{3} \dot{x}_j A_j(\boldsymbol{r}, t)] \right) \\
&= M \dot{x}_i + q A_i(\boldsymbol{r}, t).
\end{aligned}
\tag{2.25}
$$

It is clear that the canonical momentum is not equal to the mechanical momentum since the latter is the mass times the velocity, i.e., only the first term in Eq. (2.25). As described above, we must use Eq. (2.25) to express the generalized velocities in terms of the canonical momenta

$$
\dot{x}_i(\boldsymbol{r}, \boldsymbol{p}, t) = \frac{p_i - q A_i(\boldsymbol{r}, t)}{M},
\tag{2.26}
$$

which leads to the Hamiltonian

$$
\begin{aligned}
H(\boldsymbol{r}, \boldsymbol{p}, t) &= \sum_{i=1}^{3} p_i \dot{x}_i(\boldsymbol{r}, \boldsymbol{p}) - L(\boldsymbol{r}, \dot{\boldsymbol{r}}(\boldsymbol{r}, \boldsymbol{p}), t) \\
&= \boldsymbol{p} \cdot \frac{\boldsymbol{p} - q\boldsymbol{A}(\boldsymbol{r}, t)}{M} - \frac{1}{2} M \left(\frac{\boldsymbol{p} - q\boldsymbol{A}(\boldsymbol{r}, t)}{M} \right)^2 \\
&\quad + q \left(\Phi(\boldsymbol{r}, t) - \frac{\boldsymbol{p} - q\boldsymbol{A}(\boldsymbol{r}, t)}{M} \cdot \boldsymbol{A}(\boldsymbol{r}, t) \right) \\
&= \frac{\boldsymbol{p}^2 - 2q\boldsymbol{p} \cdot \boldsymbol{A}(\boldsymbol{r}, t) + q^2 \boldsymbol{A}^2(\boldsymbol{r}, t)}{2M} + q\Phi(\boldsymbol{r}, t) \\
&= \frac{[\boldsymbol{p} - q\boldsymbol{A}(\boldsymbol{r}, t)]^2}{2M} + q\Phi(\boldsymbol{r}, t).
\end{aligned}
\tag{2.27}
$$

It is easy to check that the Hamiltonian equations for this particular Hamilton function lead to Newton's second law with the force being equal to the Lorentz force.

2.2.3   Gauge considerations

In the previous sections we introduced the Lagrangian L and the Hamiltonian H for a charged particle in an electromagnetic field. One striking feature of both the Lagrangian, Eq. (2.20), and the Hamiltonian, Eq. (2.27), is that they do not directly refer to the physical fields, $\boldsymbol{E}$ and $\boldsymbol{B}$. Instead both functions depend on the potentials Φ and $\boldsymbol{A}$. In all derivations we have

not made any reference to a particular gauge, and, hence, we conclude that the expressions for L and H are valid in all gauges.

Let us now focus on how the Hamiltonian equations behave under gauge transformations. A gauge transformation clearly changes the Hamiltonian, and leads to changes in the Hamiltonian equations. Nevertheless, the particle dynamics, i.e. $r(t)$, can only depend on the physical fields and must therefore be gauge invariant. Hence, a gauge transformation changes the Hamiltonian equations in a very special way: they lead to invariant solutions for $r(t)$. On the contrary, there is no requirement that the solutions $p(t)$ should be invariant to gauge transformations since the canonical momentum does not represent a physical property of the particle. In fact, from Eq. (2.25) it is obvious that p is *not* gauge invariant. The true mechanical momentum is of course $M\dot{r}$ which *is* gauge invariant.

From a practical point of view, the gauge freedom can be used to simplify the Hamiltonian equations. As we shall see below, such a trick also applies in the quantum mechanical case.

2.3   Electromagnetic fields in quantum mechanics

2.3.1   The time dependent Schrödinger equation

We start the quantum mechanical discussion from the wave mechanical approach and with a wave function defined in the position basis. The dynamics of a wave function is governed by the time dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \mathcal{H}\Psi(\boldsymbol{r},t), \qquad (2.28)$$

where $\mathcal{H}$ is the Hamilton operator which is nearly identical to the classical Hamiltonian function H. The difference is that the canonical momentum in quantum mechanics is defined to be the generator of translations in position space and is therefore represented by the operator $p_j \rightarrow -i\partial/\partial x_j$ for each of the three Cartesian components [1]. Thus, from Eq. (2.27) we see that the Schrödinger equation in the presence of electromagnetic fields takes the form

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \left(\frac{[-i\boldsymbol{\nabla} - q\boldsymbol{A}(\boldsymbol{r},t)]^2}{2M} + q\Phi(\boldsymbol{r},t)\right)\Psi(\boldsymbol{r},t). \qquad (2.29)$$

It is customary to maintain the notation $\boldsymbol{p}$ for the operator $-i\boldsymbol{\nabla}$. In accordance with the classical theory, we regard $\boldsymbol{p} - q\boldsymbol{A}(\boldsymbol{r},t)$ as the operator that represents the kinetic momentum. Note that $\boldsymbol{p}$ and $\boldsymbol{A}(\boldsymbol{r},t)$ in general do not commute, so the ordering of terms is important if the square is written out.

2.3.2   Gauge transformations in quantum mechanics

As in the classical case, the quantum mechanical Hamiltonian is not gauge invariant. We now need to investigate how a gauge transformation affects the Schrödinger equation. Let us assume that the Schrödinger equation is satisfied by the wave function Ψ_0 for the potentials $(\Phi_0, \boldsymbol{A}_0)$. First we assume that the potentials are transformed as in Eqs. (2.7) and (2.8). The Schrödinger equation for the transformed Hamiltonian then reads

$$
i\frac{\partial}{\partial t}\Psi_1(\boldsymbol{r},t) = \left(\frac{[-i\boldsymbol{\nabla} - q\boldsymbol{A}_0(\boldsymbol{r},t) - q\boldsymbol{\nabla}\Lambda(\boldsymbol{r},t)]^2}{2M}\right.
$$

$$
\left. + q\Phi_0(\boldsymbol{r},t) - q\frac{\partial\Lambda(\boldsymbol{r},t)}{\partial t}\right)\Psi_1(\boldsymbol{r},t). \tag{2.30}
$$

If we take the new wave function to be related to the original as $\Psi_1(\boldsymbol{r},t) = \exp[iq\Lambda(\boldsymbol{r},t)]\Psi_0(\boldsymbol{r},t)$ we get from the left hand side

$$
i\frac{\partial}{\partial t}\Psi_1(\boldsymbol{r},t) = -q\frac{\partial\Lambda(\boldsymbol{r},t)}{\partial t}e^{iq\Lambda(\boldsymbol{r},t)}\Psi_0(\boldsymbol{r},t) + ie^{iq\Lambda(\boldsymbol{r},t)}\frac{\partial}{\partial t}\Psi_0(\boldsymbol{r},t). \tag{2.31}
$$

The first term conveniently cancels the last term in Eq. (2.30).

The right hand side requires calculation of

$$
\begin{aligned}
(-i\boldsymbol{\nabla} - q\boldsymbol{A}_0 - q\boldsymbol{\nabla}\Lambda)e^{iq\Lambda(\boldsymbol{r},t)}\Psi_0(\boldsymbol{r},t) &= q\boldsymbol{\nabla}\Lambda e^{iq\Lambda(\boldsymbol{r},t)}\Psi_0 - ie^{iq\Lambda(\boldsymbol{r},t)}\boldsymbol{\nabla}\Psi_0 \\
&\quad - qe^{iq\Lambda(\boldsymbol{r},t)}(\boldsymbol{A}_0 + \boldsymbol{\nabla}\Lambda)\Psi_0 \\
&= e^{iq\Lambda(\boldsymbol{r},t)}(-i\boldsymbol{\nabla} - q\boldsymbol{A}_0)\Psi_0(\boldsymbol{r},t).
\end{aligned}
\tag{2.32}
$$

Another application of $(-i\boldsymbol{\nabla} - q\boldsymbol{A}_0 - q\boldsymbol{\nabla}\Lambda)$ on the result above gives similarly

$$
(-i\boldsymbol{\nabla} - q\boldsymbol{A}_0 - q\boldsymbol{\nabla}\Lambda)^2 e^{iq\Lambda(\boldsymbol{r},t)}\Psi_0(\boldsymbol{r},t) = e^{iq\Lambda(\boldsymbol{r},t)}(-i\boldsymbol{\nabla} - q\boldsymbol{A}_0)^2\Psi_0(\boldsymbol{r},t).
\tag{2.33}
$$

Finally, we insert the expressions (2.31) and (2.33) into the transformed Schrödinger equation, Eq. (2.30). Then we see that we obtain the original Schrödinger equation after multiplication by $e^{-iq\Lambda}$ from the left. Conversely: If the function $\Psi_0(\boldsymbol{r},t)$ satisfies the original Schrödinger equation, then $\Psi_1(\boldsymbol{r},t) = \exp[iq\Lambda(\boldsymbol{r},t)]\Psi_0(\boldsymbol{r},t)$ solves the transformed Schrödinger equation. Now of course the key question is how observables obtained with this transformed wave function relate to observables obtained with the original wave function. It is straightforward to show that expectation values of the position and kinetic momentum operators are gauge invariant

$$
\begin{aligned}
\langle\Psi_1|\boldsymbol{r}|\Psi_1\rangle &= \langle\Psi_0|\boldsymbol{r}|\Psi_0\rangle, & \tag{2.34} \\
\langle\Psi_1|\boldsymbol{p} - q\boldsymbol{A}_1|\Psi_1\rangle &= \langle\Psi_0|\boldsymbol{p} - q\boldsymbol{A}_0|\Psi_0\rangle. & \tag{2.35}
\end{aligned}
$$

Since observables are built from combinations of the position and kinetic momentum operators (intrinsic spin being an exception), we conclude that expectation values of observables are gauge invariant.

2.3.3   The electrostatic field

Electrons are bound to nuclei primarily due to the attractive electrostatic Coulomb force, and the same force is responsible for interelectronic repulsion. Thus, the overall understanding of atomic and molecular structures is based on electrostatic interaction[2]. The whole theory outlined above becomes particularly simple when considering a purely electrostatic field. In the absence of a magnetic field, we can take the vector potential to be zero. With this obvious choice of gauge, the Schrödinger equation takes the form

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \left(\frac{\boldsymbol{p}^2}{2M} + V(\boldsymbol{r})\right)\Psi(\boldsymbol{r},t), \qquad (2.36)$$

with $V(\boldsymbol{r}) = q\Phi(\boldsymbol{r})$ being equal to the classical electrostatic potential energy. The Hamiltonian is now time independent and allows for separable solutions. Accordingly, the whole framework of first solving the time independent Schrödinger equation can be used to solve the problem. A typical example is an electron with charge $q = -1$ which sees a nuclear charge Z_{nuc} from a nucleus which is fixed in space at position $\boldsymbol{R}$. In that case, the scalar potential corresponding to the nuclear charge is $\Phi(\boldsymbol{r}) = Z_{\mathrm{nuc}}/|\boldsymbol{r} - \boldsymbol{R}|$ and $V(\boldsymbol{r}) = -Z_{\mathrm{nuc}}/|\boldsymbol{r} - \boldsymbol{R}|$.

2.3.4   Dipole gauges

We shall now discuss the interaction between atoms and electromagnetic radiation. In vacuum, the Maxwell equations admit monochromatic plane wave solutions for the electric and magnetic fields

$$\boldsymbol{E_k}(\boldsymbol{r},t) = \hat{\boldsymbol{\varepsilon}}E_{0\boldsymbol{k}}\cos(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t + \delta_{\boldsymbol{k}}), \qquad (2.37)$$

$$\boldsymbol{B_k}(\boldsymbol{r},t) = \frac{1}{c}\hat{\boldsymbol{k}} \times \boldsymbol{E_k}(\boldsymbol{r},t). \qquad (2.38)$$

$\boldsymbol{k}$ is the wave vector along which the waves propagate, ω is the (angular) frequency and c is the speed of light. The magnitude of the wave vector is related to the frequency by $\omega = kc$, and to the wavelength as $\lambda = 2\pi/k = 2\pi c/\omega$. Furthermore, the propagating waves are traverse, so that the polarization $\hat{\boldsymbol{\varepsilon}}$ is perpendicular to the propagation direction.

The fields in Eqs. (2.37) and (2.38) have infinite extension in both temporal and spatial dimensions. While such a description is appropriate for fields produced by continuous wave lasers, this situation obviously cannot correspond to an ultrashort laser pulse. Instead, a pulse of finite duration can be produced by a superposition of plane waves with different frequencies and with a well defined phase relationship between the frequency components.

---

[2]In addition to the antisymmetrization requirement and spin.

Then, for a superposition of plane waves with equal polarization and propagation directions, it is possible to obtain a field of finite duration T on the form

$$\boldsymbol{E}(\boldsymbol{r},t) = \begin{cases} \boldsymbol{E}_0 h(t) \cos(\boldsymbol{k}_0 \cdot \boldsymbol{r} - \omega_0 t + \delta_0), & 0 \leq t \leq T \\ 0, & \text{Otherwise} \end{cases}, \quad (2.39)$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \frac{1}{c}\hat{\boldsymbol{k}} \times \boldsymbol{E}(\boldsymbol{r},t). \quad\quad\quad\quad\quad\quad (2.40)$$

$h(t)$ is an envelope function which goes to zero at $t = 0$ and $t = T$ and varies somewhat slower than the oscillating cosine factor. Borrowing terminology from radio communication, we call ω_0 the carrier frequency, and the envelope $h(t)$ defines a type of amplitude modulation of the carrier wave. Common examples of envelope functions include square-, trapezoidal-, and $\sin^2$-envelopes. Regardless of the shape of the envelope function, the fields are not monochromatic anymore but have a central frequency of ω_0 and will have a bandwidth in frequency space which is inversely proportional to the pulse duration. When we refer to *the* wavelength of a field of finite duration, we usually mean the wavelength of the carrier $\lambda_0 = 2\pi c/\omega_0$.

For many purposes, the electric dipole approximation (often just referred to as the dipole approximation) is an excellent approximation. At each instant of time, the spatial variations of the field is on the order of the wavelength. For visible radiation the wavelength is several hundredths of nanometers – or on the order of ten thousand Bohr radii. On the other hand, atomic dynamics occurs on the length scale of a few Bohr radii, so it is a good approximation to regard the field as being constant in space everywhere near the atomic core. Furthermore, if the electrons move with velocities much slower than the speed of light, they interact much more weakly with the magnetic field than with the electric field. This is evident from the Lorentz force (2.1) and from the relationship between the strengths of $\boldsymbol{E}$ and $\boldsymbol{B}$, Eq. (2.40). Therefore, we neglect the magnetic field altogether. In summary, the dipole approximation only includes the simplified electric field

$$\boldsymbol{E}(t) = \boldsymbol{E}_0 h(t) \cos(\omega_0 t + \delta), \quad 0 \leq t \leq T. \quad\quad (2.41)$$

δ is called the carrier envelope phase difference and describes how the maximum of the optical cycle is displaced with respect to the maximum of the envelope. The question now arises: How can this field be expressed in terms of scalar and vector potentials. There are two particularly simple choices of potentials which give this field as $\boldsymbol{E} = -\boldsymbol{\nabla}\Phi - \partial_t \boldsymbol{A}$. Either we take

$$\Phi_{\mathrm{LG}}(\boldsymbol{r},t) = -\boldsymbol{r} \cdot \boldsymbol{E}(t), \quad\quad\quad\quad\quad\quad (2.42)$$

$$\boldsymbol{A}_{\mathrm{LG}}(\boldsymbol{r},t) = \boldsymbol{0}, \quad\quad\quad\quad\quad\quad\quad\quad (2.43)$$

or we take

$$\Phi_{\text{VG}}(\boldsymbol{r}, t) = 0, \tag{2.44}$$

$$\boldsymbol{A}_{\text{VG}}(\boldsymbol{r}, t) = -\int_0^t \boldsymbol{E}(t')dt'. \tag{2.45}$$

The gauge transformation is $\Lambda(\boldsymbol{r}, t) = \mp \boldsymbol{r} \cdot \int_0^t \boldsymbol{E}(t')dt'$ with "$-$" sign when going from the former to the latter form and "$+$" sign for the reverse transformation. The first set of potentials leads to the *length gauge* Hamiltonian

$$\mathcal{H}^{(\text{LG})} = \frac{\boldsymbol{p}^2}{2M} - q\boldsymbol{r} \cdot \boldsymbol{E}(t), \tag{2.46}$$

while the second set leads to the *velocity gauge* Hamiltonian

$$\mathcal{H}^{(\text{VG})} = \frac{[\boldsymbol{p} - q\boldsymbol{A}_{\text{VG}}(t)]^2}{2M} = \frac{\boldsymbol{p}^2}{2M} - \frac{q}{M}\boldsymbol{A}_{\text{VG}}(t) \cdot \boldsymbol{p} + \frac{q^2}{2M}\boldsymbol{A}_{\text{VG}}^2(t). \tag{2.47}$$

When obtaining the last form, we used that the velocity gauge vector potential, Eq. (2.45), is constant in space and $\boldsymbol{A}_{\text{VG}}$ and $\boldsymbol{p}$ therefore commute[3]. A further simplification can be obtained by noting that the term proportional to $\boldsymbol{A}_{\text{VG}}^2(t)$ enters the Hamiltonian as a space independent term and simply corresponds to an instantaneous change of zero point on the energy scale. Since the zero point energy can be chosen arbitrarily, we do not expect such a term to contain any physical meaning. In fact, a wave function that solves the Schrödinger equation with the Hamiltonian (2.47) will also solve the Schrödinger equation without the $\boldsymbol{A}_{\text{VG}}^2(t)$ term present – we just need to multiply it by a certain global phase factor. Such a global phase does not contain any physical significance and, hence, both solutions correspond to the same physical state. Therefore, we will often use

$$\mathcal{H}^{(\text{VG})} = \frac{\boldsymbol{p}^2}{2M} - \frac{q}{M}\boldsymbol{A}_{\text{VG}}(t) \cdot \boldsymbol{p}, \tag{2.48}$$

as an equivalent form for the velocity gauge Hamiltonian. Note that the equivalence is only valid in the electric dipole approximation since the vector potential will depend on spatial variables beyond the dipole approximation.

The length gauge has the convenient property that the kinetic and canonical momentum are equal. This is not the case for the velocity gauge – a fact that can lead to quite surprising difficulties. The problem is that after the real physical field is turned off at time T, the vector potential in the velocity gauge still has the value $\boldsymbol{A}_{\text{VG}}(T) = -\int_0^T \boldsymbol{E}(t')dt'$. The kinetic momentum is then represented by the operator $\boldsymbol{p} - q\boldsymbol{A}_{\text{VG}}(T)$, even though the field is turned off. If one accidentally assumes that the kinetic and canonical momentum operators always are equal in the absence of any field, one would

---

[3]More generally, $\boldsymbol{A}$ and $\boldsymbol{p}$ commute whenever the Coulomb gauge restriction $\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r}, t) = 0$ is fulfilled.

find a momentum expectation value that is shifted by the amount $q\boldsymbol{A}_{\mathrm{VG}}(T)$ from the correct result. As long as one is aware of this shift, it is easy to make the transformation to obtain the correct result that corresponds to the observable kinetic momentum. It turns out that in most situations we do not need to be concerned about this momentum shift. If we represent the electric field in temporal Fourier representation,

$$\tilde{\boldsymbol{E}}(\omega) = \int_0^T e^{i\omega t} \boldsymbol{E}(t) dt, \qquad (2.49)$$

we see that $\boldsymbol{A}_{\mathrm{VG}}(T)$ corresponds to the (negative) frequency component at $\omega = 0$, often called the DC component. Usually the field must represent a *travelling* wave that emerges from a laser cavity that is practically infinitely far away from the interaction region. A travelling wave cannot contain a DC component, since $\omega = 0$ implies that $\boldsymbol{k} = \boldsymbol{0}$, and the corresponding component in Eq. (2.37) is constant. This component is therefore not allowed in the superposition that creates the travelling pulse (2.40), for if it were present, the field would not vanish at $t < 0$ and $t > T$. Therefore, we conclude that a field representing a travelling wave must not contain a DC component, and accordingly the velocity gauge vector potential must vanish at the end of the pulse.

Although the situation described above is the most common, there exist certain situations where the momentum drift in the velocity gauge must be treated more carefully. First, if the particle escapes the laser focus *during* the pulse at time t, the kinetic momentum must be represented by $\boldsymbol{p} - q\boldsymbol{A}_{\mathrm{VG}}(t)$ in the velocity gauge. Second, if the field is not created by a travelling wave, it *can* contain a DC component. Imagine, for example if we place the charged particle between two capacitor plates. At time $t = 0$ we turn on a voltage over the plates and at time T, the voltage is turned off. In between $t = 0$ and T the particle is exposed to a constant field and, hence, the vector potential in the velocity gauge, Eq. (2.45), is clearly nonzero at time T.

Now, let us return to the travelling wave. If we wish to simulate a travelling wave, we will require that the field used by the theory does not contain a DC component either. Ideally, we should of course choose a field that matches experimental conditions exactly. Experimentally, it is, however, very difficult to produce a field that is fully characterized by an exact envelope and phase. Fortunately, the atomic response turns out to be largely insensitive to the exact pulse shape. The most notable exception occurs if the envelope encloses only a few optical cycles. In that case, the carrier envelope phase difference becomes important. For ordinary many-cycle pulses, however, the primary parameters that govern the dynamics are thus the peak intensity, the central wavelength, and the pulse duration. From a theoretical point of view, it is therefore acceptable to choose the most convenient envelope function as long as the aforementioned parameters are left untouched. Hence, we may as well choose a field that is guaranteed to fulfil

$\int_0^T \boldsymbol{E}(t')dt' = \boldsymbol{0}$ and then hope that such a field reflects reality rather well. As an example, we take the Gaussian envelope which is often referred to as the Fourier limited pulse since this particular envelope has the smallest possible frequency variance for a given pulse duration. Due to this property, experimentalists often try to obtain this envelope shape. For the Gaussian envelope, the field varies as

$$\boldsymbol{E}(t) = \boldsymbol{E}_0 \exp\left(-\frac{2\ln 2}{\tau^2}t^2\right)\cos(\omega_0 t), \qquad (2.50)$$

with τ being the full width at half maximum of the intensity. Such a pulse shape is, however, inconvenient if we want to solve the time dependent Schrödinger equation for a wave function that evolves from an initial state prior to the field to a final state after the pulse. Although the field strength falls off as $t \to \pm\infty$, the field never completely vanishes so that the terms "prior" and "after" the pulse become vaguely defined. Additionally, it is not obvious that the velocity gauge vector potential vanishes after the pulse. Therefore, we try to approximate the field by sine-square-like envelope

$$\boldsymbol{E}(t) = -\frac{d}{dt}\frac{\boldsymbol{E}_0}{\omega_0}\sin^2\left(\frac{\omega_0 t}{2N}\right)\cos(\omega_0 t + \delta), \quad 0 \le t \le \frac{2\pi N}{\omega_0}, \qquad (2.51)$$

and $\boldsymbol{E}(t) = \boldsymbol{0}$ otherwise. N is an integer which specifies the total number of optical cycles in the pulse and should hence be related to the value of τ in the Gaussian pulse. From Eq. (2.51), we easily read off the corresponding velocity gauge vector potential as the expression to the right of the differentiation. This vector potential clearly vanishes at the end of the pulse $T = 2\pi N/\omega_0$. In Fig. 2.1, we show an example of the fields in Eqs. (2.50) and (2.51). For better comparison, the Gaussian pulse is shifted in time so that the maxima of the two pulses coincide. The figure shows that sine-square pulse resembles the Gaussian pulse very well. The agreement is particularly good near the peak halfway through the pulse where the most important dynamics occurs, while small deviations are seen at the ends. Thus, we conclude that we can approximate the Fourier limited pulse by a finite sine-square pulse without affecting the main dynamics. In the numerical work presented in this thesis, the field will always take the form of Eq. (2.51).

If the charged particle is influenced by a total electromagnetic field which is a superposition of an electrostatic field and an external radiation field described in the dipole approximation, the full Hamiltonian finally takes the form

$$\mathcal{H}^{\text{(LG)}} = \frac{\boldsymbol{p}^2}{2M} - q\boldsymbol{r}\cdot\boldsymbol{E}(t) + V(\boldsymbol{r}), \qquad (2.52)$$

$$\mathcal{H}^{\text{(VG)}} = \frac{\boldsymbol{p}^2}{2M} - \frac{q}{M}\boldsymbol{A}_{\text{VG}}(t)\cdot\boldsymbol{p} + V(\boldsymbol{r}), \qquad (2.53)$$
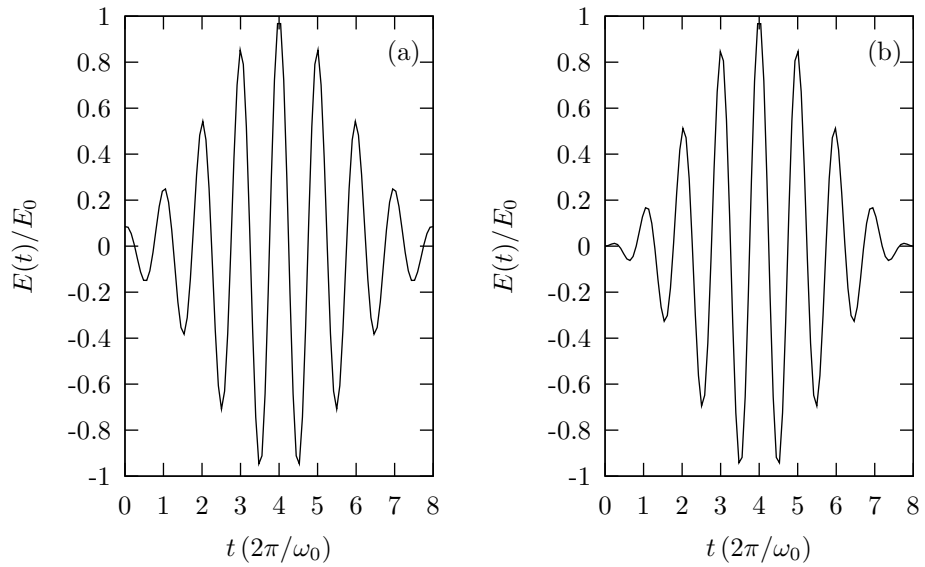
Figure 2.1: (a) Gaussian pulse envelope for τ equal to three optical cycles. (b) Sine-square pulse envelope enclosing eight cycles.

in the length and velocity gauges, respectively. As in Eq. (2.36), $V(\boldsymbol{r})$ is equal to the classical potential energy associated with the electrostatic field. As the length and velocity gauge Hamiltonians are expressed in Eqs. (2.52) and (2.53), it is not immediately obvious which form is the more convenient to use. It is generally accepted that the velocity gauge provides superior convergence properties for dynamical processes in the strong-field regime [2, 3]. On the other hand, the length gauge is often simpler to implement numerically as we shall see in chapter 5.

Chapter 3

Discretization of functions

In this chapter we discuss various methods to approximate functions by a finite set of discrete values.

3.1 Finite basis representation

The exact wave function resides in an infinitely dimensional vector space $\mathscr{H}$ which spans all square integrable functions. Furthermore, the vector space is an inner product space with the inner product defined as

$$\langle f|g\rangle = \int f^\star(x)g(x)dx. \tag{3.1}$$

This vector space is referred to as the Hilbert space. For simplicity, we only consider functions in one spatial dimension. Higher dimensions can always be treated as direct product spaces of one-dimensional function spaces.

The usual way to calculate anything within a vector space is to define a basis and to express everything, operators and wave functions, as matrix- and vector representations in terms of this basis. In order to construct a basis for the infinitely dimensional Hilbert space, we need infinitely many basis functions. Obviously, computers only have a finite amount of memory so that they cannot store an infinite expansion. Thus, any numerical solution of the time dependent Schrödinger equation must rely on an expansion of wave functions and operators in some finite basis representation. The basis must be truncated appropriately so that the essential part of the wave function is accurately described.

Having decided for a basis, we need to define a metric for the vector space spanned by the basis, i.e., a definition of how to calculate inner products that correspond to Eq. (3.1). The inner product must be chosen consistently in the sense that if we express any two functions in two different bases, their inner product must be invariant. In some cases, the inner product can be

defined simply as the ordinary dot product. A well known example of a non-simple metric is the four vectors of relativistic mechanics where the temporal component must be weighed by -1 in the inner product.

We shall use the discrete bases for the specific purpose to solve *differential* equations. Hence, for any given basis, it will be important to define how to apply coordinate and differential operators and any function of these two fundamental operators.

Some problems are constrained by certain boundary conditions. For example, if a system is known to be enclosed in a box, the wave function must vanish at the box boundaries. In fact, even if a system is not physically enclosed it is often useful to impose box boundary conditions artificially. This is of course only a matter of convenience; if we know a priori that a system is confined in some region, it can hardly matter if we impose box boundary conditions as long as the box is chosen large enough to enclose the system. The advantage of this procedure is that we can reduce the basis considerably, since we only need to include basis states which conform with the box boundary condition.

The following sections describe some of the commonly used bases which are particularly relevant for the numerical implementation used in this Thesis.

3.2   Grid representation and finite differences

In the grid representation, we use a subspace of the position basis as the basis functions. These functions are local in position space, so that each basis function is only nonzero at one specific point in space. The choice of basis therefore defines a grid in space corresponding to the location of the basis functions. For an N-point grid, it should then be sufficient to represent any function by the values on the grid points

$$f \doteq \boldsymbol{f} = \{f(x_0), f(x_1), \cdots, f(x_{N-1})\}^T. \tag{3.2}$$

The symbol "$\doteq$" denotes "represented by in some basis", with the basis being the grid in the present context. We define the metric in such a way that any integrand which is piecewise linear between the grid points must be integrated exactly. This is obtained by by defining inner products by the trapezoidal integration rule

$$\langle f|g \rangle = \int_{x_0}^{x_{N-1}} f^\star(x)g(x)dx = \sum_{i=1}^{N-1} (x_i - x_{i-1}) \frac{f^\star(x_i)g(x_i) + f^\star(x_{i-1})g(x_{i-1})}{2}, \tag{3.3}$$

where the last equality holds exactly if the product $f^\star(x)g(x)$ is linear in each subinterval. If the grid points are equally spaced by the distance Δx,

we obtain the simple integration rule

$$\langle f|g\rangle = \Delta x \left(\frac{1}{2} f^\star(x_0)g(x_0) + \sum_{i=1}^{N-2} f^\star(x_i)g(x_i) + \frac{1}{2} f^\star(x_{N-1})g(x_{N-1}) \right).$$

(3.4)

We might be tempted to use a more elaborate integration rule to define the inner product, such as the Simpsons rule. Unfortunately, such a choice leads to an inconsistent metric as we are going to show below.

In quantum mechanics we often encounter unitary operators which have the property that they conserve inner products according to

$$\langle Uf|Ug\rangle = \langle f|g\rangle,$$

(3.5)

for any two functions f and g and any unitary operator U. What will a grid representation of a unitary operator then look like? It is well known that a unitary matrix $\boldsymbol{U}$ has the properties $\boldsymbol{U}^\dagger \boldsymbol{U} = \mathbf{1}$, and hence ordinary dot products are invariant under unitary transformations

$$(\boldsymbol{U} \cdot \boldsymbol{f})^\dagger \cdot \boldsymbol{U} \cdot \boldsymbol{g} = \boldsymbol{f}^\dagger \cdot \boldsymbol{g}.$$

(3.6)

It is exactly the similarity between Eqs. (3.5) and (3.6) that forces us to define the inner product by the trapezoidal rule. The ordinary dot product only corresponds to the inner product (scaled only by a factor of Δx) for this particular choice, and unitary operators are then represented by unitary matrices. In fact, Eq. (3.4) shows that the correspondence between inner products by the trapezoidal rule and ordinary dot products are only exact if the functions vanish at the end points. If we had defined the inner product by another integration rule, it is not guaranteed that a transformation by a unitary matrix conserves inner products.

The advantage of a grid representation is that the coordinate operator x is particularly simple – it simply corresponds to a scalar multiplication in each grid point $xf(x_i) = x_i f(x_i)$. Hence, we can associate the coordinate operator with the following diagonal matrix representation

$$xf \doteq \mathrm{diag}(\boldsymbol{x}) \cdot \boldsymbol{f}, \quad \boldsymbol{x} = \{x_0, x_1, \cdots, x_{N-1}\}^T.$$

(3.7)

Generally, any function of the coordinate operator will be represented by a diagonal matrix

$$V(x)f \doteq \mathrm{diag}(\boldsymbol{V}) \cdot \boldsymbol{f}, \quad \boldsymbol{V} = \{V(x_0), V(x_1), \cdots, V(x_{N-1})\}^T.$$

(3.8)

The derivative operator is slightly more complicated. As an initial guess, we use a three-point finite difference method to approximate a derivative at the interior points $\{x_i | i \neq 0, N-1\}$

$$\left. \frac{d}{dx} f \right|_{x_i} = \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x} + \mathcal{O}(\Delta x^2 f^{(3)}),$$

(3.9)

where we assumed a uniform grid spacing Δx. We return to the values at the boundaries below. We see that the three-point finite difference method is exact only if f is a second order polynomial between x_{i-1} and x_{i+1}. There exist higher order formulas, e.g. five- and seven-point formulas, that are exact for polynomials of higher degrees. For simplicity we only consider the three-point finite difference here. From Eq. (3.9) we see that the differential operator can be represented by

$$\frac{d}{dx}f \doteq \boldsymbol{D} \cdot \boldsymbol{f}, \tag{3.10}$$

where $\boldsymbol{D}$ is a tridiagonal Toeplitz matrix with zeros on the diagonal and the value $1/(2\Delta x)$ on all the superdiagonal elements and the value $-1/(2\Delta x)$ on the subdiagonal. We might expect that the second order derivative is most simply expressed as $\boldsymbol{D} \cdot \boldsymbol{D}$. As the product of two tridiagonal matrices is a pentadiagonal matrix, this would correspond to a five-point method with an accuracy of the order $\mathcal{O}(\Delta x^2)$. It is in fact possible to obtain the same order of accuracy with just a three-point formula

$$\left.\frac{d^2}{dx^2}f\right|_{x_i} = \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{\Delta x^2} + \mathcal{O}(\Delta x^2). \tag{3.11}$$

The matrix representation of the second order derivative is then also a tridiagonal Toeplitz matrix

$$\frac{d^2}{dx^2}f \doteq \boldsymbol{D}_2 \cdot \boldsymbol{f}, \tag{3.12}$$

with the diagonal elements of $\boldsymbol{D}_2$ being $-2/\Delta x^2$ and the sub- and superdiagonal elements $1/\Delta x^2$.

Box boundary conditions $f(x_0) = f(x_{N-1}) = 0$ are imposed simply by excluding the points x_0 and x_{N-1} from the basis. The derivative operators are well defined on all the remaining interior points according to Eqs. (3.9) and (3.11) with $f(x_0) = f(x_{N-1}) = 0$.

The finite difference method has the advantage that everything is completely transparent. Differential quotients are simply approximated by the most obvious choice, namely finite differences. Furthermore, differential operators can be applied to functions at low computational cost because the matrix representations are banded. Generally, matrix-vector products scale with the dimensionality N as $\mathcal{O}(N^2)$, since it requires a sum of N products to calculate each of the N elements in the output vector. For banded matrices, the matrix-vector product scales as $\mathcal{O}(N)$, since it takes only a sum of a *fixed* number of products (three for a tridiagonal matrix) to calculate each of the N elements in the product vector.

The main drawback of the finite difference method is that the accuracy is low.

3.3 Grid-Fourier representation

As we saw in the previous section, the finite difference method suffered from the problem that derivatives are calculated with low precision. In this section, we introduce the Fourier method which calculates derivatives with superior accuracy for many functions.

As in the previous section we let f be defined on an equidistantly spaced grid with grid spacing Δx. For simplicity, we assume that the number of grid points N is even. We now define the inverse discrete Fourier transformation of the samples of f as [4]

$$\tilde{f}(k_j) = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} f(x_i)e^{-ik_j x_i}, \;\; k_j = \left\{ \begin{array}{ll} \frac{2\pi}{N\Delta x}j, & j = 0, 1, \cdots, \frac{N}{2} - 1 \\ \frac{2\pi}{N\Delta x}j - \frac{2\pi}{\Delta x}, & j = \frac{N}{2}, \cdots, N-1 \end{array} \right. ,$$

$$(3.13)$$

and the forward discrete Fourier transformation

$$f(x_i) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \tilde{f}(k_j)e^{ik_j x_i}. \tag{3.14}$$

As in Eq. (3.2), we let $\tilde{\boldsymbol{f}}$ denote a column vector that contains the values $\tilde{f}(k_j)$. Then we see that the discrete forward and inverse Fourier transformations correspond to unitary transformations

$$\tilde{\boldsymbol{f}} = \boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{f}, \quad \boldsymbol{f} = \boldsymbol{\mathcal{F}} \cdot \tilde{\boldsymbol{f}}, \tag{3.15}$$

with $\mathcal{F}_{ij} = e^{ik_j x_i}/\sqrt{N}$ being the elements of the unitary transformation matrix.

Now we regard the expression (3.14) as discrete samples of the following continuous function

$$f(x) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \tilde{f}(k_j)e^{ik_j x}. \tag{3.16}$$

Functions for which Eq. (3.16) holds exactly are called *bandwidth limited* functions. It is easy to evaluate the derivative of f on the grid points

$$\left. \frac{df}{dx} \right|_{x_i} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \tilde{f}(k_j)ik_j e^{ik_j x_i}, \tag{3.17}$$

which is just a discrete Fourier transformation on the samples $\{ik_j \tilde{f}(k_j)\}$. Using the matrix notations in (3.15) we obtain the representation of the derivative as

$$\frac{d}{dx}f \doteq \boldsymbol{\mathcal{F}} \cdot \mathrm{diag}(i\boldsymbol{k}) \cdot \boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{f}. \tag{3.18}$$

We say that $\boldsymbol{\mathcal{F}}$ defines the transformation into the representation that *diagonalizes* the derivative operator in the following sense: Assume that D is a linear operator that generates the derivative on the grid points from the function values on the grid points. Being a linear operator, it must be possible to represent D as a matrix $\boldsymbol{D}$, although we do not assume the explicit form to be known at this stage

$$\frac{d}{dx}f \doteq \boldsymbol{D} \cdot \boldsymbol{f} = \boldsymbol{\mathcal{F}} \cdot \boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{D} \cdot \boldsymbol{\mathcal{F}} \cdot \boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{f}. \tag{3.19}$$

Comparing this expression with Eq. (3.18) above, we make the identification $\boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{D} \cdot \boldsymbol{\mathcal{F}} = \mathrm{diag}(i\boldsymbol{k})$, i.e., the transformation $\boldsymbol{\mathcal{F}}$ brings $\boldsymbol{D}$ on a diagonal form.

In Fourier representation it is easy to apply any function of the derivative operator

$$T\left(\frac{d}{dx}\right)f \doteq \boldsymbol{\mathcal{F}} \cdot \mathrm{diag}(i\boldsymbol{T}) \cdot \boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{f}, \quad \boldsymbol{T} = \{T(k_0), T(k_1), \cdots, T(k_{N-1})\}. \tag{3.20}$$

Computationally, the discrete Fourier transformations can be accomplished very efficiently by the Fast Fourier Transformation (FFT) algorithm. The transformations in Eq. (3.15) are written as matrix-vector products. Such an operation has a computational complexity which scales as $\mathcal{O}(N^2)$ as mentioned in Sec. 3.2. With the Fast Fourier Transformation, it is, however, possible to do the transformation with a computational complexity that scales as $\mathcal{O}(N \log_2 N)$. In practice, the number of grid points are much larger than unity, and the complexity scaling of FFT can be regarded as being close to linear. It is no exaggeration to claim that spectral methods primarily have gained their widespread use due to the FFT.

The Fourier representation is well suited for functions that obey periodic boundary conditions. On the other hand, box boundary conditions are not fulfilled straightforwardly. In order to fulfil box boundary conditions, we try instead the following discrete sine transformation

$$\tilde{f}^{\sin}(k_j) = \sqrt{\frac{2}{N}} \sum_{i=0}^{N-1} f(x_i) \sin(k_j x_i), \quad \left\{ k_j = \frac{\pi j}{N\Delta x} \middle| j = 0, 1, .., N-1 \right\}, \tag{3.21}$$

$$f(x_i) = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} \tilde{f}^{\sin}(k_j) \sin(k_j x_i). \tag{3.22}$$

The sine transformation is evidently its own inverse with both the transformation matrices $[\mathcal{F}^{\sin}]_{ij} = \sqrt{2/N} \sin(x_i k_j)$. This transformation explicitly takes the box boundary conditions into account. Furthermore, it shares many properties with the discrete Fourier transformation, for example, the

existence of a fast $\mathcal{O}(N \log_2 N)$ algorithm. The sine transformation does, however, *not* bring the differential operator on a diagonal form. Luckily, the *second* derivative is still diagonal

$$\frac{d^2}{dx^2} f \doteq \boldsymbol{\mathcal{F}}^{\sin} \cdot \operatorname{diag}(-\boldsymbol{k}^2) \cdot \boldsymbol{\mathcal{F}}^{\sin} \cdot \boldsymbol{f}. \qquad (3.23)$$

Therefore, sine transformations can be used with advantage for problems with box boundary conditions and derivatives of even order.

3.3.1   Caveats

Having established a method to calculate derivatives from discrete samples of a function, one question naturally arises, namely the accuracy of the method. It is of course always possible to transform a set of data according to Eqs. (3.13) and (3.14). What really sets the limits when evaluating derivatives by Fourier transforms is the accuracy of the identification (3.16). This is most accurate if f can be written as a sum of trigonometric functions with the wave numbers being numerically less than $k_{\max} = \pi/\Delta x$.

Many functions are very well approximated by sums of trigonometric functions as in Eq. (3.16). This is for example the case for smooth functions which oscillate on length scales larger than the grid spacing. In such cases, the Fourier method as described above will provide any derivative with a very high accuracy. On the other hand, if a function has sharp edges or oscillates more rapidly than the grid spacing, the Fourier method may fail. The problem with rapid oscillations can be solved by choosing a denser grid. Unfortunately there seems to be no way to handle functions with sharp edges consistently. We might expect that sharp edges are unphysical and thus such functions should be excluded as wave functions on physical grounds. This is, however, not true since, e.g., eigenfunctions corresponding to singular potentials do not have continuous derivatives at the singularity. Common examples are the *s*-wave eigenfunctions of hydrogen at the position of the nucleus. Problems with singular potentials should therefore be treated with some care when using the Fourier method.

3.4   Grid-Legendre representation

In the previous section we described the grid and Fourier representation of a function f. The Fourier method is clearly most accurate if the function behaves like a sum of trigonometric functions. On the other hand, the expansion in complex exponentials is not the optimal choice if the function is a polynomial of low order. Stated alternatively: It requires an unnecessarily large expansion in sines and cosines to describe a polynomial accurately. Instead of using trigonometric functions, it is more sensible to use polynomials

as basis functions. The simplest choice of a polynomial basis is the basis of
Legendre polynomials – the subject for the present section.

If a function f is defined on the interval $x \in [-1, 1]$ and is known to
be a polynomial of order $l_{\max}$ or less, then f can be expanded in Legendre
polynomials up to order $l_{\max}$

$$f(x) = \sum_{l=0}^{l_{\max}} a_l \bar{P}_l(x), \tag{3.24}$$

where $\bar{P}_l$ are orthonormal Legendre polynomials [5]

$$\int_{-1}^{1} \bar{P}_l(x) \bar{P}_{l'}(x) dx = \delta_{ll'}. \tag{3.25}$$

As for the Fourier method, we use the term "bandwidth limited" for any
function where Eq. (3.24) is exact. The expansion coefficients a_l can be
found by multiplying Eq. (3.24) by $\bar{P}_{l'}(x)$ and integrating over x

$$\int_{-1}^{1} \bar{P}_{l'}(x) f(x) dx = a_{l'}, \tag{3.26}$$

where we used the orthogonality property (3.25). Equation (3.24) states
that f is completely determined if the $l_{\max} + 1$ expansion coefficients are
known. Suppose we now want a grid representation of f instead which
contains the same information. If we require the mapping between the grid
and the Legendre representation to be unitary, we must require that both
representations have the same number of basis functions, i.e., there must be
exactly $l_{\max}+1$ grid points. If, for example, the number of grid points exceeds
the number of expansion coefficients, two or more different functions defined
on the grid points can lead to the same Legendre expansion. In such cases
it is impossible to define a unique inverse, and, hence, the transformation
cannot be unitary.

In the case of Legendre polynomials, it is possible to define a unitary
transformation to a grid by choosing the $l_{\max} + 1$ grid points as the roots
of the polynomial $\bar{P}_{l_{\max}+1}(x)$, and make integrations according to Gauss-
Legendre quadrature weights [6]. The key feature of this particular choice
is that (3.26) is exactly reproduced by the sum

$$a_l = \int_{-1}^{1} \bar{P}_l(x) f(x) dx = \sum_{j=0}^{l_{\max}} w_j \bar{P}_l(x_j) f(x_j). \tag{3.27}$$

Choosing the ($l_{\max}+1$) order Legendre quadrature weights $\{w_j\}$ and abscissa
$\{x_j\}$ guarantees that the integral of any polynomial up to order $2l_{\max} + 1$
is exactly reproduced by a sum. Since f by assumption is a polynomial of

order less than $l_{\max} + 1$ and the order of the Legendre polynomial is at most $l_{\max}$, we see that Eq. (3.27) is exact for all l's included in the Legendre basis.

In matrix notation, we define the $(l_{\max} + 1)$-dimensional column vectors

$$\boldsymbol{a} = \{a_0, a_1, \ldots, a_{l_{\max}}\}^T, \tag{3.28}$$

$$\boldsymbol{f} = \{f(x_0), f(x_1), \ldots, f(x_{l_{\max}})\}^T. \tag{3.29}$$

The transformation from $\boldsymbol{a}$ to $\boldsymbol{f}$ is embodied in Eq. (3.24)

$$\boldsymbol{f} = \bar{\boldsymbol{P}} \cdot \boldsymbol{a}, \tag{3.30}$$

with $[\bar{P}]_{jl} = \bar{P}_l(x_j)$. The transformation from $\boldsymbol{f}$ to $\boldsymbol{a}$ is according to Eq. (3.27)

$$\boldsymbol{a} = \text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \boldsymbol{f}. \tag{3.31}$$

It may not seem obvious that a backward and forward transformation leaves either $\boldsymbol{a}$ or $\boldsymbol{f}$ unaffected, i.e., whether the following equation holds

$$\bar{\boldsymbol{P}} \cdot \text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T = \text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \bar{\boldsymbol{P}} = \boldsymbol{1}. \tag{3.32}$$

As we now show, these properties are in fact fulfilled by the particular construction of the abscissa and weights. The elements of the product matrix $\text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \bar{\boldsymbol{P}}$ are

$$\left[\text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \bar{\boldsymbol{P}}\right]_{ll'} = \sum_{j=0}^{l_{\max}} w_j \bar{P}_l(x_j) \bar{P}_{l'}(x_j) = \int_{-1}^1 \bar{P}_l(x) \bar{P}_{l'}(x) dx = \delta_{ll'}, \tag{3.33}$$

where the middle equality is true since $P_l(x)P_{l'}(x)$ is a polynomial of order $l + l' < 2l_{\max} + 1$ and therefore exactly integrated by the Gauss-Legendre quadrature. Hence, $\text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T$ is the left inverse of $\bar{\boldsymbol{P}}$ and it must therefore also be the right inverse, so that

$$\left[\bar{\boldsymbol{P}} \cdot \text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T\right]_{jj'} = \sum_{l=0}^{l_{\max}} \bar{P}_l(x_j) w_{j'} \bar{P}_l(x_{j'}) = \delta_{jj'}. \tag{3.34}$$

If Eqs. (3.30) and (3.31) are true unitary transformations, we must define inner products that are invariant under such transformations. We let $\boldsymbol{a}^{(1)}$ and $\boldsymbol{a}^{(2)}$ denote the Legendre expansion coefficients of the functions $f^{(1)}$ and $f^{(2)}$, respectively. In the Legendre representation it is most obvious to define the inner product

$$\langle f^{(1)} | f^{(2)} \rangle = \sum_{l_1=0}^{l_{\max}} \sum_{l_2=0}^{l_{\max}} \int_{-1}^1 [a_{l_1}^{(1)}]^\star a_{l_2}^{(2)} \bar{P}_{l_1}(x) \bar{P}_{l_2}(x) dx$$

$$= \sum_{l=0}^{l_{\max}} [a_l^{(1)}]^\star a_l^{(2)} = \boldsymbol{a}^{(1)\dagger} \cdot \boldsymbol{a}^{(2)}. \tag{3.35}$$

In grid representation, we calculate integrals by the Gauss-Legendre quadrature rule

$$\langle f^{(1)}|f^{(2)}\rangle = \sum_{j=0}^{l_{\max}} w_j [f^{(1)}(x_j)]^\star f^{(2)}(x_j) = \boldsymbol{f}^{(1)\dagger} \cdot \mathrm{diag}(\boldsymbol{w}) \cdot \boldsymbol{f}^{(2)}. \qquad (3.36)$$

With the particular choice of transformation, Eqs. (3.30) and (3.31), the two inner products are exactly equal.

The present case shows that it is not directly the properties of the transformation matrix that makes a transformation unitary. It is easy to see that neither $\mathrm{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T$ nor $\bar{\boldsymbol{P}}$ are unitary matrices. Instead, it is the way that inner products are defined in each representation that matters. For the Fourier method, the transformation matrix itself is unitary, but we were forced to use the trapezoidal rule as the definition of inner products in both grid and Fourier representation.

The matrices that define the transformations between Legendre and grid representation are straightforwardly calculated by stable algorithms. Unlike the discrete Fourier transformation there exists no "fast" algorithm that can do the transformations faster than the usual matrix multiplications. Thus, if a particular problem can be solved either by transformations between grid-Fourier or grid-Legendre representation (both with the same dimension), the Fourier method will usually be much faster, since transformations can be done by the fast Fourier transform.

What does a differential operator D look like in Legendre representation? It must be possible to represent such an operator as a matrix in Legendre representation, for if f is a polynomial of order $l_{\max}$, then df/dx will be a polynomial of order $l_{\max} - 1$ which also can be expanded in Legendre polynomials. As always in linear problems, it is sufficient to see how operators act on the basis functions. The derivative of any Legendre polynomial can be written as an expansion in lower order Legendre polynomials

$$\frac{d}{dx}\bar{P}_l(x) = \sum_{l'=0}^{l-1} D_{ll'} \bar{P}_{l'}(x). \qquad (3.37)$$

The coefficients $D_{ll'}$ that define the operation can in principle be found analytically. The matrix representation $\boldsymbol{D}$ of the differential operation D is evidently lower triangular with zeros on the diagonal. By the same argument, the second order derivate will be represented by a lower triangular matrix with zeros on the diagonal and the first subdiagonal, and so on for higher order derivatives.

Using the differential operator in Legendre representation, it is also possible to find the differential operator in the grid representation

$$\frac{d}{dx}f \doteq \bar{\boldsymbol{P}} \cdot \boldsymbol{D} \cdot \mathrm{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \boldsymbol{f}. \qquad (3.38)$$

Here $\mathrm{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T$ first changes from grid into Legendre representation where $\boldsymbol{D}$ can be applied. $\bar{\boldsymbol{P}}$ changes from the differentiated Legendre coefficients into the differentiated values on the grid points. The fact that $\boldsymbol{D}$ is non-diagonal, makes it difficult to apply a higher order differentiation. As an alternative method to calculate derivatives on the grid, we can construct the Lagrange interpolation polynomial of degree $l_{\max}$ from the values defined on the $l_{\max} + 1$ grid points. But since f by assumption is a polynomial of degree $l_{\max}$, the interpolating polynomial must be identical to f. Analytical differentiation of the interpolating polynomial therefore gives exactly the differential quotients evaluated at the grid points.

At this point the Legendre representation may seem useless for any practical purpose. First, there is no algorithm faster than matrix multiplication that transforms between grid and Legendre representation. Second, the derivative operator is not diagonal in Legendre representation. There is, however, one important advantage of the Legendre representation. Consider the following differential operator

$$L^2 = -\frac{d}{dx}\left(\left(1 - x^2\right) \frac{d}{dx} \right) = -\left(1 - x^2\right) \frac{d^2}{dx^2} + 2x \frac{d}{dx}. \qquad (3.39)$$

The Legendre polynomials are exactly eigenfunctions of this operator with eigenvalues $l(l+1)$

$$L^2 \bar{P}_l(x) = l(l+1)\bar{P}_l(x). \qquad (3.40)$$

The matrix representation of L^2 then reads in Legendre representation

$$L^2 \doteq \mathrm{diag}(\boldsymbol{l}^2), \quad \{l_j^2 = j(j+1) | j = 0, 1, \ldots, l_{\mathrm{lmax}}\}, \qquad (3.41)$$

and in grid representation

$$L^2 f \doteq \bar{\boldsymbol{P}} \cdot \mathrm{diag}(\boldsymbol{l}^2) \cdot \mathrm{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \boldsymbol{f}. \qquad (3.42)$$

The operator L^2 occurs in spherical coordinates as the quantum mechanical angular momentum operator. Thus, it is now evident that the Legendre representation is useful when a problem involves the angular momentum operator. This is for example the case when the Laplacian is expressed in a spherical system of coordinates.

There are many variants of the grid/Legendre representation. For example, if the function f is known to be a polynomial times a non-polynomial factor $(1 - x^2)^\alpha$, the polynomial part can be expanded in Jacobi polynomials and the grid representation is defined by the weights and abscissa corresponding to the Gauss-Jacobi quadrature rule. Another variant concerns functions that are defined on $[0, \infty)$ and fall off exponentially. In this case, the functions can be expanded in Laguerre functions and with grid representation defined by Gauss-Laguerre quadratures.

3.5   Finite element method

The grid-Fourier and grid-Legendre methods work well for trigonometric-
and polynomial functions, respectively. Derivatives can be calculated exactly
within the basis by the expressions (3.18) and (3.38). Both expressions reveal
that the derivative at a single grid point requires the knowledge of function
values on all other grid points. Correspondingly, the matrices involved in
the derivative operation are dense and the computation of derivatives at all
grid points scales quadratically with the number of points (except for the
FFT algorithm). This way of calculating derivatives may look somewhat
strange. We often think of a derivative as being a nearly *local* operation in
the sense that the derivative at a particular grid point depends on function
values near the grid point but does not depend on the behaviour very far
away. This is exactly how we defined the finite difference approximation to
the differential operator, and the banded structure of the matrix allowed the
derivatives at all grid point to be computed by a method that scales linearly
with the number of points.

The deeper lying reason for the fact that derivatives by the Fourier and
Legendre methods require all function values is that they both use globally
defined basis functions. In the literature such methods are referred to as
spectral methods. In addition to the computational complexity of deriva-
tives, the spectral methods suffer from another drawback: They are not well
suited for representing functions that vary rapidly in a small local domain
and vary slowly in other parts of space. In order to accommodate the rapidly
varying part of such a function, one should use a grid that is dense enough
to resolve these variations accurately. But for the spectral methods, it is
only possible to increase the density of grid points *globally*. For example, if
we take the Fourier method and we know that a small grid spacing Δx is
needed somewhere. Then we have to use the same dense spacing *everywhere*,
and thereby introduce a large number of unnecessary points in the regions
where the function varies slowly.

One procedure that allows for both fast computation of derivatives and
efficient handling of local variations is to use *local* basis functions. This
section describes the finite element method which is one such method. The
basis is chosen as piecewise polynomials which are partly overlapping. The
simplest case is to use normalized rooftop functions which are piecewise first

order polynomials. We define N basis functions between x_0 and x_{N-1} as

$$b_0(x) = \begin{cases} \frac{x_1-x}{x_1-x_0}, & x_0 \le x \le x_1 \\ 0 & \text{Otherwise} \end{cases} , \tag{3.43}$$

$$b_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}}, & x_{i-1} \le x \le x_i \\ \frac{x_{i+1}-x}{x_{i+1}-x_i}, & x_i \le x \le x_{i+1} \\ 0 & \text{Otherwise} \end{cases} , \tag{3.44}$$

$$b_{N-1}(x) = \begin{cases} \frac{x-x_{N-2}}{x_{N-1}-x_{N-2}}, & x_{N-2} \le x \le x_{N-1} \\ 0 & \text{Otherwise} \end{cases} . \tag{3.45}$$

Box boundary conditions are easily obtained by excluding the first and the last basis function. The i'th basis function can be represented graphically by a triangle that is defined with corners at the coordinates $(x_i, 1), (x_{i\pm 1}, 0)$ – hence the name rooftop functions. Only half of the triangles are used at the end points. There is no restriction on how the coordinates are chosen. For example, we are allowed to put many basis functions in some regions and a smaller number in other regions. Hence, this basis has exactly the desired flexibility that is absent from the global methods. Unfortunately, there is also some bad news about this local basis. First, the basis functions are not orthogonal as we introduce the sensible inner product

$$\langle b_i | b_{i'} \rangle = \int_{x_0}^{x_{N-1}} b_i(x) b_{i'}(x) dx$$
$$= \delta_{ii'} \frac{x_{i+1} - x_{i-1}}{3} + \delta_{i,i'+1} \frac{x_i - x_{i-1}}{6} + \delta_{i,i'-1} \frac{x_{i+1} - x_i}{6}. \tag{3.46}$$

Although it is always possible to orthogonalize a basis by e.g., a Gram-Schmidt orthogonalization algorithm, we will maintain the basis $\{b_i\}$. The reason for this choice becomes clear below. The second drawback is that the coordinate operator is not diagonal

$$\langle b_i | x | b_{i'} \rangle = \delta_{ii'} \int_{x_{i-1}}^{x_{i+1}} b_i(x) x b_i(x) dx + \delta_{i,i'+1} \int_{x_{i-1}}^{x_i} b_i(x) x b_{i-1}(x) dx$$
$$+ \delta_{i,i'+1} \int_{x_i}^{x_{i+1}} b_i(x) x b_{i+1}(x) dx, \tag{3.47}$$

and generally, any function of x will be represented by a tridiagonal matrix in the finite element basis. Finally, a second derivative is

$$\langle b_i | \frac{d^2}{dx^2} | b_{i'} \rangle = \int_{x_0}^{x_{N-1}} \frac{d}{dx} b_i(x) \frac{d}{dx} b_{i'}(x) dx$$
$$= -\frac{\delta_{ii'}}{x_i - x_{i-1}} - \frac{\delta_{ii'}}{x_{i+1} - x_i} + \frac{\delta_{i,i'+1}}{x_i - x_{i-1}} + \frac{\delta_{i,i'-1}}{x_{i+1} - x_i}, \tag{3.48}$$

where we assumed without proof that integration by parts applies for piece-wise differentiable functions. Now it becomes apparent why the non-orthogonal basis $\{b_i\}$ is convenient. With this basis, any operator that involves functions of the coordinate operator as well as the second derivative is represented by a tridiagonal matrix. If we had chosen to work with an orthogonalized version of the basis, we would have destroyed this property, and all matrix representations would be dense. For certain problems, such as diagonalization of a Hamiltonian, it is convenient to have both the coordinate operator and the derivative simultaneously on a tridiagonal form.

The finite element method can be extended to include higher order polynomials than first order. In such cases, coordinate- and differential operators become band-diagonal with a bandwidth larger than one. Even if we consider more than one dimension, it is possible to order the basis functions so that the matrices of the operators become banded.

A final question is: Is there a grid representation connected to the finite elements? Suppose that we have expanded a function as

$$f(x) = \sum_{i=0}^{N-1} c_i b_i(x). \tag{3.49}$$

Now, there is a unique correspondence between the function values at the points $\{x_i\}$ and the expansion coefficients $\{c_i\}$ – in fact they are equal. Thus, one could be lead to think that the finite element representation is actually identical to the grid representation. However, this is false since the coordinate operator is non-diagonal in finite element representation.

3.6 Benchmark – The time independent Schrödinger equation

As we have seen in the previous sections, there is no such thing as an optimal basis for *all* problems. This section illustrates some advantages and disadvantages of the various bases.

We benchmark the bases by solving the time independent Schrödinger equation

$$\left(-\frac{1}{2}\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\psi(x). \tag{3.50}$$

With appropriate boundary conditions, this equation can be transformed into an eigenvalue problem. We solve the problem in a discrete basis $\{e_i\}$ by defining

$$\psi(x) = \sum_{i=0}^{N-1} \psi_i e_i(x). \tag{3.51}$$

Method	$x, V(x)$	$d/dx, d^2/dx^2$	$-\frac{1}{2}\frac{d^2}{dx^2} + V(x)$	Overlap
Grid-Finite Diff.	diagonal	banded	banded	unity
Grid-Fourier	diagonal	dense	dense	unity
Fourier	dense	diagonal	dense	unity
Grid-Legendre	diagonal	dense	dense	unity
Legendre	dense	dense	dense	unity
Finite Elements	banded	banded	banded	banded

Table 3.1: Structure of matrix representations of coordinate- and differential operators in various representations

If the basis is a coordinate grid, the expansion coefficients $\{\psi_i\}$ are exactly equal to the wave function evaluated on the grid points. In the discrete basis, the Schrödinger equation reads

$$\sum_{i=0}^{N-1} \psi_i \left(-\frac{1}{2}\frac{d^2}{dx^2} + V(x) \right) e_i(x) = E \sum_{i=0}^{N-1} \psi_i e_i(x). \qquad (3.52)$$

Taking the inner product by the basis function $e_j(x)$ from the left

$$\sum_{i=0}^{N-1} \psi_i \left(-\frac{1}{2}[D_2]_{ji} + [V]_{ji} \right) = E \sum_{i=0}^{N-1} \psi_i [S]_{ji}, \qquad (3.53)$$

or equivalently

$$\left(-\frac{1}{2}\boldsymbol{D}_2 + \boldsymbol{V} - E\boldsymbol{S} \right) \cdot \boldsymbol{\psi} = 0. \qquad (3.54)$$

$[D_2]_{ji}$ and $[V]_{ji}$ are the elements of the matrix representation of the second derivative and the potential operator, respectively. $[S]_{ji} = \langle e_j | e_i \rangle$ are the elements of the overlap matrix. If the basis is orthonormal, $\boldsymbol{S}$ is simply the unit matrix, and Eq. (3.54) is a standard matrix eigenvalue- or diagonalization problem. Otherwise, Eq. (3.54) is a generalized matrix eigenvalue problem. Both types of eigenvalue problems can be solved by standard numerical routines.

The structures of the matrices in Eq. (3.54) depend on the basis. Table 3.1 summarizes the matrix structures for the bases described in the previous sections. The computational complexity for diagonalization scales as $O(N^3)$ for general dense matrices. For banded matrices, with bandwidth much smaller than the matrix dimension, the complexity scaling is only $O(N^2)$. Hence, from Table 3.1, we see that the Schrödinger equation should be solved faster with the finite difference and finite elements methods compared with the Fourier and Legendre bases. This is especially true for very large basis expansions.

n	Grid-Finite Diff.	Grid-Fourier	Finite Elements	Exact
1	-0.4987562	-0.4971154	-0.4999943	-0.5000000
2	-0.1249220	-0.1246385	-0.1249988	-0.1250000
3	-0.0555401	-0.0554484	-0.0555546	-0.0555556
4	-0.0312451	-0.0312048	-0.0312489	-0.0312500
5	-0.0199980	-0.0199768	-0.0199956	-0.0200000
6	-0.0138748	-0.0138548	-0.0138579	-0.0138685

Table 3.2: Eigenvalues for the hydrogen s-states. The results are obtained for a radial box size of 100 a.u. with approximately 1000 basis functions. The "exact" values are exact within the box. The first five eigenvalues agree with the exact eigenvalues for an infinite box to the number of significant digits given in the table [7].

3.6.1 Hydrogen radial equation

As an example of a one-dimensional Schrödinger equation, we consider the radial equation of hydrogen. This problem serves as an excellent benchmark system, since the complete set of solutions is known analytically. Furthermore, it corresponds to a physical system that has many similarities with the systems encountered later on in this thesis. We stress the importance of benchmarking on a system being as closely related to the systems under study as possible. If a particular basis performs well for an idealized model problem, it is not guaranteed to be equally good for a real world problem – we shall briefly return to this point in Sec. 10.1.1.

The differential equation for the reduced radial $u_{nl}(r)$ function is

$$\left(-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - \frac{1}{r}\right) u_{nl}(r) = E_n u_{nl}(r). \qquad (3.55)$$

The bound states fulfil the boundary conditions $u_{nl}(0) = 0$ and $u_{nl}(r \to \infty) \to 0$ which we simulate by box boundary conditions at $r = 0$ and $r = r_{\max}$. We use three different bases, namely the grid-finite difference, grid-sine-Fourier and the finite elements, since these bases can easily be modified to fulfil box boundary conditions. For the present case, we study the s-states ($l = 0$) and we take the outer boundary at $r_{\max} = 100$ a.u. which is sufficient to represent the first six bound states. For all bases we will use ≈ 1000 basis functions. For the finite difference and grid-sine-Fourier methods the number of basis functions and the box size dictate the basis, i.e., the position of the grid points. For the finite element method on the other hand, we can decide ourselves where to put the basis functions. In the present calculations, the elements are chosen densely distributed near the inner boundary near $r = 0$ and sparsely distributed at the outer boundary.

The results for the first six eigenvalues are given in Table 3.2. First we consider the lowest eigenvalue, corresponding to the exact eigenfunction $u_{10}(r) = 2re^{-r}$. The finite difference and Fourier methods are not very

accurate for this state. The problem with these two methods is that the function is very localized in the region $r < 5$ a.u. The grid spacing used here does not represent such a function very accurately. A more accurate representation would require a smaller grid spacing. But as mentioned above, the grid point density can only be made large globally, so that we must add more points everywhere. In contrast to the ground state, the higher lying excited states are less localized and vary more slowly. The relatively sparse grid representation is therefore sufficient to obtain quite good estimates for the eigenvalues. The equidistant grid methods are evidently not very well suited for problems that involve several different length scales – a uniformly dense grid would be unnecessary for the excited states while a uniformly sparse grid would be inaccurate for the ground state.

It is precisely the ability to handle different length scales that is the strength of the finite element method. From Table 3.2, we see that the finite element method can give accurate eigenvalues for both the ground state and the excited states. In order to obtain an accurate representation, the elements are chosen with a spacing of $\Delta r = 0.005$ a.u. near the origin and increasing up to $\Delta r = 1$ a.u. at the outer boundary. These values are to be compared with the uniform grid spacing of $\Delta r = 0.1$ a.u. for the finite difference and Fourier methods.

The problems with the sparse representation of the ground state can be illustrated by the eigenvectors. In all three bases, we can interpret the i'th element of the eigenvectors as the value of the wave at the point r_i. Figure 3.1 shows the wave function obtained by the three different bases. The results obtained by the equidistant grids, panels (a) and (b), clearly do not contain as detailed information near the origin as the result obtained by the finite element method.

We might expect the uniform grid to be more accurate than the finite elements at long distances, since the finite elements are spaced by a distance larger than the grid spacing. In Fig. 3.2, we show the data $\{|u_{10}(r_i)|\}$. It turns out that the finite element method is actually also the most accurate at long distances. The problem with the Fourier method is that the true function $u_{10}(r) = 2re^{-r}$ cannot be written as a finite Fourier series such as the basis expansion assumes. We may see the solution (dashed) as the best approximation we can get by the finite Fourier series. Although the error seems to be quite small $(< 10^{10})$, it can be a significant disturbing factor. For example, we could consider a hydrogen atom in the ground state which we perturb weakly. Subsequently we want to look for the changes that the perturbation induced on the wave function. The effect of the perturbation should then be significantly larger than the error that is already present in the ground state – otherwise the ground state errors could mask the effects that we are interested in. We therefore conclude that the Fourier basis is not well suited for studying atomic processes that include weak perturbations. Note that it is the exponential decay – a characteristic behaviour of all
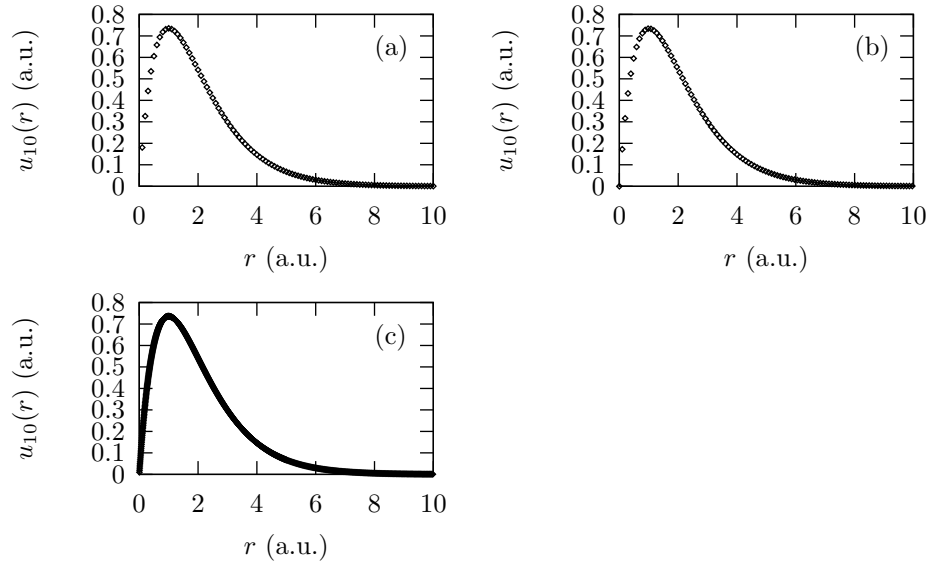
Figure 3.1: The (reduced) radial part of the 1s ground state of hydrogen. The results are obtained by diagonalization of the Hamiltonian matrix representation in the bases (a) grid-finite difference, (b) grid-sine-Fourier, and (c) finite elements.
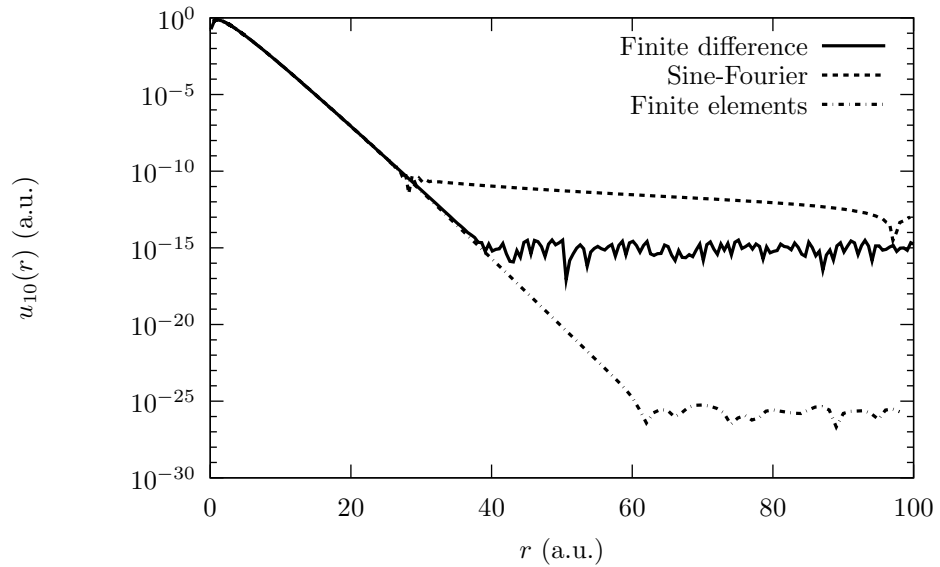

Figure 3.2: The long range part of the reduced radial wave function of the 1s ground state of hydrogen. The results are obtained by the bases: grid-finite difference (solid), grid-sine-Fourier (dashed), and finite elements (dashed-dotted).

bound atomic states – that is responsible for the error in the Fourier basis. Conversely, the Fourier-sine basis would have been an excellent choice for a particle in a box where the basis functions are the exact eigenstates.

Chapter 4

The time dependent Schrödinger equation – numerical aspects

This chapter describes general properties of the time dependent Schrödinger equation.

4.1 Time evolution operator

The time dependent Schrödinger equation is a parabolic partial differential equation which contains the first derivative in time and up to second derivatives in space

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \mathcal{H}(t)\Psi(\boldsymbol{r},t). \tag{4.1}$$

The problem of solving this type of differential equation is most often stated as an *initial value problem*, i.e., at some reference time t_0, the wave function is assumed to be known throughout space as $\Psi(\boldsymbol{r},t_0)$. The task is then to determine how such an initial wave function evolves in time. We can introduce an operator which formally solves the problem. This operator is called the time evolution operator $\mathcal{U}(t,t_0)$ and has the property

$$\Psi(\boldsymbol{r},t) = \mathcal{U}(t,t_0)\Psi(\boldsymbol{r},t_0). \tag{4.2}$$

Before turning to the explicit form of $\mathcal{U}(t,t_0)$, we first discuss some of its properties.

First, if $\Psi(\boldsymbol{r},t)$ is to be interpreted as a probability amplitude distribution, it must conform with the normalization requirement at all times

$$1 = \langle\Psi(\boldsymbol{r},t_0)|\Psi(\boldsymbol{r},t_0)\rangle, \tag{4.3}$$
$$1 = \langle\Psi(\boldsymbol{r},t)|\Psi(\boldsymbol{r},t)\rangle = \langle\Psi(\boldsymbol{r},t_0)|\mathcal{U}^\dagger(t,t_0)\mathcal{U}(t,t_0)|\Psi(\boldsymbol{r},t_0)\rangle. \tag{4.4}$$

Equations (4.3) and (4.4) must hold for *any* choice of $\Psi(\boldsymbol{r},t_0)$ and, hence, we propose the requirement

$$\mathbf{1} = \mathcal{U}^\dagger(t,t_0)\mathcal{U}(t,t_0), \tag{4.5}$$

i.e., $\mathscr{U}(t, t_0)$ must be a unitary operator.

Second, it is obvious to assume that time evolution can be done stepwise in time. If we introduce the time evolution from t_0 to t_2 via the intermediate time t_1, we thus find

$$\Psi(\boldsymbol{r}, t_2) = \mathscr{U}(t_2, t_1)\Psi(\boldsymbol{r}, t_1) = \mathscr{U}(t_2, t_1)\mathscr{U}(t_1, t_0)\Psi(\boldsymbol{r}, t_0), \qquad (4.6)$$

i.e., $\mathscr{U}(t_2, t_0) = \mathscr{U}(t_2, t_1)\mathscr{U}(t_1, t_0)$. This relation can be used recursively to decompose the total time evolution into many smaller time steps. For example, a finite time evolution operator can be constructed from an infinitely number of infinitesimal time evolution operators

$$\mathscr{U}(t + \Delta t, t) = \mathscr{U}(t + \Delta t, t + \frac{N-1}{N}\Delta t) \cdots \mathscr{U}(t + \frac{2\Delta t}{N}, t + \frac{\Delta t}{N})\mathscr{U}(t + \frac{\Delta t}{N}, t), \tag{4.7}$$

where the infinitesimal limit is obtained as the number of subdivisions N approaches infinity. Note that the time ordering of the operators on the right hand side is generally important.

We claim that the operator that propagates a wave function by an infinitely small time step dt is

$$\mathscr{U}(t + dt, t) = \mathbf{1} - i\mathcal{H}(t)dt. \tag{4.8}$$

This relation is obtained from the Schrödinger equation by approximating the left hand side by the finite difference in time and letting the time difference go to zero. The infinitesimal time evolution operator is unitary as required, since the Hamilton operator is Hermitian and terms of the order $(dt)^2$ can be neglected in the limit $dt \to 0$

$$\mathscr{U}^{\dagger}(t + dt, t)\mathscr{U}(t + dt, t) = [\mathbf{1} + i\mathcal{H}^{\dagger}(t)dt][\mathbf{1} - i\mathcal{H}(t)dt] = \mathbf{1} + [\mathcal{H}(t)dt]^2. \tag{4.9}$$

In the special case where the Hamiltonian is independent of time, $\mathcal{H}(t) = \mathcal{H}_0$, the expression (4.7) for the finite time evolution operator can be written more compactly. From Eq. (4.8), we see that all infinitesimal operators are equal for a time independent $\mathcal{H}_0$, and the time evolution operator reads

$$\mathscr{U}(t + \Delta t, t) = \lim_{N \to \infty} \mathscr{U}(t + \frac{\Delta t}{N}, t)^N = \lim_{N \to \infty} (\mathbf{1} - i\mathcal{H}_0\frac{\Delta t}{N})^N = \exp(-i\mathcal{H}_0\Delta t). \tag{4.10}$$

This time evolution operator is also easily seen to be unitary.

When the Hamiltonian is time dependent, it is unfortunately impossible to obtain a similarly elegant expression for the time evolution operator. The most obvious generalization of Eq. (4.10) would perhaps be

$$\mathscr{U}(t + \Delta t, t) \stackrel{?}{=} \exp\left(-i\int_t^{t+\Delta t} \mathcal{H}(t')dt'\right), \tag{4.11}$$
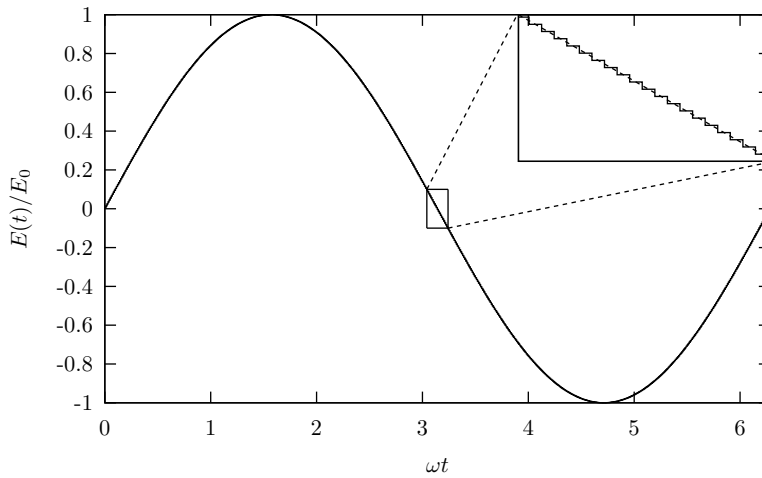
Figure 4.1: Piecewise constant approximation to a sinusoidal field. The continuously varying field is indicated by the dashed curve. The solid line is the approximation for 628 subdivisions over the period ($\omega \Delta t = 0.01$).

but this expression is generally not true due to the non-commutivity of the Hamiltonian at different instants of time. Instead the correct time evolution operator is sometimes defined by the time ordered exponential operator $T \exp(-i \int \mathcal{H}(t') dt')$, but this is just a compact notation for the time ordered product in Eq. (4.7). Application of this form turns out to be impractical and of little use in the present context.

For many purposes, it is accurate to approximate a time dependent Hamiltonian by a series of piecewise time independent Hamiltonians

$$\mathcal{H}(t) \approx \begin{cases} \mathcal{H}(\frac{\Delta t}{2}), & 0 \leq t \leq \Delta t \\ \mathcal{H}(\frac{3\Delta t}{2}), & \Delta t \leq t \leq 2\Delta t \\ \mathcal{H}(\frac{5\Delta t}{2}), & 2\Delta t \leq t \leq 3\Delta t \\ \vdots \end{cases} . \tag{4.12}$$

The temporal subdivision Δt must of course be smaller than the timescale over which the Hamiltonian changes appreciably. Consider for example a Hamiltonian which is time dependent due to an oscillating external field as in Eqs. (2.52) or (2.53). Such a Hamiltonian varies on the timescale of the optical period of the field. Figure 4.1 shows one period of a sinusoidal field. In addition to the real continuously varying field, the figure also shows a piecewise constant approximation for 628 subdivisions. Since the approximation resembles the true field very well, it is expected that the time dependent Hamiltonian is very well reproduced as in Eq. (4.12). Within each subinterval of duration Δt, we can thus use the time evolution operator for the constant Hamiltonian, Eq. (4.10). Finally, we use the composition property

to obtain the time evolution operator between t_0 and $t_0 + n\Delta t$

$$\mathscr{U}(t_0 + n\Delta t, t_0) = \exp\left[-i\mathcal{H}\left(\frac{(2n-1)\Delta t}{2}\right)\Delta t\right] \times \cdots$$
$$\times \exp\left[-i\mathcal{H}\left(\frac{3\Delta t}{2}\right)\Delta t\right] \times \exp\left[-i\mathcal{H}\left(\frac{\Delta t}{2}\right)\Delta t\right].$$
$$(4.13)$$

Note that Eq. (4.13) follows from an approximation to the *Hamiltonian*. Later on, we will encounter approximations to the time evolution operator itself which is another and more subtle type of approximation. If the Hamiltonian is truly piecewise constant, Eq. (4.13) would be the exact. Conversely, we can regard the time evolution operator as being the true time evolution operator for *some* Hamiltonian, and hence, it is guaranteed to be exactly unitary. This property is of course also easily verified by direct calculation of the Hermitian adjoint.

From a computational point of view, the time evolution operator in Eq. (4.13) turns out to be very practical. Furthermore, it can be made arbitrarily close to the true time evolution operator by choosing the time step size small enough. Therefore, in this thesis, we only use the expression in Eq. (4.13) to solve the time dependent Schrödinger equation.

4.2   The operator exponential

As we showed in the previous section, the time dependent Schrödinger equation is formally solved by successive applications of the time evolution operator on the form

$$\Psi(\boldsymbol{r}, t + \Delta t) = \exp\left[-i\mathcal{H}\left(t + \frac{\Delta t}{2}\right)\Delta t\right]\Psi(\boldsymbol{r}, t). \qquad (4.14)$$

This expression involves an operator exponential, and it is obviously necessary to know how to actually apply such an operator. Let us consider a time independent Hermitian operator $\mathcal{A}$ expressed in coordinate representation. The exponential of the operator $i\mathcal{A}$ can either be defined from the limiting procedure as in Eq. (4.10) or as the Taylor series

$$\exp(i\mathcal{A}) = \sum_{n=0}^{\infty} \frac{(i\mathcal{A})^n}{n!}. \qquad (4.15)$$

From this expression, it is clear how to act with the operator exponential on a coordinate space wave function, since a power of an operator is well defined

$$\mathcal{A}^n \Psi(\boldsymbol{r}, t) = \underbrace{\mathcal{A} \times \cdots \times \mathcal{A} \times \mathcal{A}}_{n \text{ times}} \Psi(\boldsymbol{r}, t). \qquad (4.16)$$

Although it is now clear *how* to act with the operator exponential, actually *doing* it becomes very impractical, as we need to apply the operator infinitely many times. For practical implementations of such an operator, one might think that the Taylor series can be truncated. Unfortunately such a truncation leads to a non-unitary and unstable approximation for the operator exponential as will be demonstrated in Sec. 4.3.1.

If the eigenfunctions of the Hermitian operator $\mathcal{A}$ span the space in which the wave function is defined (in the present case the coordinate space), there is a simple way to apply the operator exponential. We let $\Phi_m(\boldsymbol{r})$ denote an eigenfunction with the real eigenvalue A_m.

$$\mathcal{A}\Phi_m(\boldsymbol{r}) = A_m \Phi_m(\boldsymbol{r}). \qquad (4.17)$$

It is very easy to apply the power of the operator on the eigenfunctions

$$\mathcal{A}^n \Phi_m(\boldsymbol{r}) = (A_m)^n \Phi_m(\boldsymbol{r}), \qquad (4.18)$$

since each time $\mathcal{A}$ is applied it just brings out another factor of A_m. Therefore it is also easy to apply the exponential on each of the eigenfunctions

$$e^{i\mathcal{A}}\Phi_m(\boldsymbol{r}) = \sum_{n=0}^{\infty} \frac{(i\mathcal{A})^n}{n!}\Phi_m(\boldsymbol{r}) = \sum_{n=0}^{\infty} \frac{(iA_m)^n}{n!}\Phi_m(\boldsymbol{r}) = e^{iA_m}\Phi_m(\boldsymbol{r}), \quad (4.19)$$

where e^{iA_m} is now just a complex number. Evidently, all eigenfunctions just acquire the phase of the eigenvalue when acted on by the exponential operator. As the full set of eigenfunctions constitutes a complete set, it is possible to write the wave function in the eigenfunction expansion

$$\Psi(\boldsymbol{r},t) = \sum_m c_m(t)\Phi_m(\boldsymbol{r}). \qquad (4.20)$$

Such an expansion is often referred to as changing into the basis that diagonalizes $\mathcal{A}$. The reason for this terminology becomes more transparent in the following section when the matrix representations are introduced. We are now finally able to express the operator exponential acting on the wave function, namely

$$e^{i\mathcal{A}}\Psi(\boldsymbol{r},t) = \sum_m c_m(t)e^{iA_m}\Phi_m(\boldsymbol{r}). \qquad (4.21)$$

Although this equation somehow expresses how to apply the exponential operator, there are some points to note. First, we need to know all the eigenfunctions $\Phi_m(\boldsymbol{r})$. To obtain these often requires a separate non-trivial calculation. Second, we have to project the wave function on the eigenstates in order to find the expansion coefficients $c_m(t)$. For the time evolution operator, we have to find eigenfunctions of the instantaneous Hamiltonian,

i.e., for each time interval we must solve the time independent Schrödinger equation. Such an approach is in principle possible, but turns out to be inconvenient and computationally demanding in reality. Hence, it will be necessary to introduce approximations in order to apply the time evolution operator efficiently. This will be the subject for the following section.

4.3   Approximations to the time evolution operator

The time evolution operator can be approximated in various ways which lead to less complicated forms than solving the time independent Schrödinger equation at each time step. We will require that a reasonable approximation to the time evolution operator must be unitary.

First we consider a one-dimensional problem with a time dependent Hamiltonian on the form

$$\mathcal{H}(t) = -\frac{1}{2}\frac{d^2}{dx^2} + V(x,t). \tag{4.22}$$

As in the previous sections, we take the time step size to be small compared to the timescale of the temporal variations in the potential. The time evolution between t and $t + \Delta t$ is then, according to Eq. (4.14),

$$\Psi(x, t+\Delta t) = \mathscr{U}(t, t+\Delta t)\Psi(x,t) = \exp\left[-i\Delta t\left(\frac{p^2}{2} + V(x,\bar{t})\right)\right]\Psi(x,t), \tag{4.23}$$

where $p \equiv -i(d/dx)$ and $\bar{t} = t + \Delta t/2$. In the following two sections, we will study two methods that give approximate solutions to the equation above.

4.3.1   Crank-Nicolson method

The most straightforward way to approximate the time evolution equation (4.23) is perhaps by the first order Taylor expansion of the exponential

$$\Psi(x, t+\Delta t) = [\mathbf{1} - i\Delta t\mathcal{H}(t)]\,\Psi(x,t) + \mathcal{O}(\mathcal{H}^2\Delta t^2). \tag{4.24}$$

If we then express the wave function on an N-point equidistant grid and approximate the second order derivative by the finite difference, we expect the following matrix equation to be correct up to first order in Δt

$$\boldsymbol{\Psi}(t+\Delta t) = \left[\mathbf{1} - i\Delta t\left(-\frac{1}{2}\boldsymbol{D}_2 + \text{diag}(\boldsymbol{V}(\bar{t}))\right)\right]\cdot\boldsymbol{\Psi}(t), \tag{4.25}$$

where

$$\begin{aligned}
\boldsymbol{\Psi}(t) &= \{\Psi(x_0,t), \Psi(x_1,t), \cdots, \Psi(x_{N-1},t)\}^T, &\tag{4.26}\\
\boldsymbol{V}(\bar{t}) &= \{V(x_0,\bar{t}), V(x_1,\bar{t}), \cdots, V(x_{N-1},\bar{t})\}^T, &\tag{4.27}
\end{aligned}$$

and $\boldsymbol{D}_2$ is the tridiagonal matrix defined Eq. (3.12). Equation (4.25) is referred to as a forward time, centered space method (FTCS). This terminology originates from the fact that it is equivalent to the time dependent Schrödinger equation with a certain type of finite differences in time and space

$$i\frac{\Psi(x_i, t + \Delta t) - \Psi(x_i, t)}{\Delta t} = -\frac{1}{2}\frac{\Psi(x_{i+1}, t) - 2\Psi(x_i, t) + \Psi(x_{i-1}, t)}{\Delta x^2}$$
$$+ V(x_i, \bar{t})\Psi(x_i, t). \tag{4.28}$$

The temporal derivative uses values at t and $t + \Delta t$ and is therefore called *forward* in time finite difference. The spatial derivative, on the other hand, uses values on each side of the grid point and is called *centered* in space finite difference.

A forward in time scheme is also called an *explicit* scheme, since an equation like (4.28) explicitly states how to obtain the wave function at time $t + \Delta t$ if it is known at time t. Unfortunately, explicit schemes are often numerically unstable [4]. For simplicity, we use a time independent Hamiltonian to illustrate this instability. With an N-point grid basis, we can find N discrete energy eigenstates $\Phi_n(\{x_i\})$ with corresponding energies E_n. We can formally expand the wave function at time $t = 0$ and $t = m\Delta t$ according to Eq. (4.21)

$$\Psi(\{x_i\}, 0) = \sum_{n=0}^{N-1} c_n \Phi_n(\{x_i\}), \tag{4.29}$$

$$\Psi(\{x_i\}, m\Delta t) = \sum_{n=0}^{N-1} c_n e^{-iE_n m\Delta t} \Phi_n(\{x_i\}). \tag{4.30}$$

On the other hand, the explicit first order approximation in Eq. (4.24), leads to the expression at $t = m\Delta t$

$$\Psi^{(1)}(\{x_i\}, m\Delta t) = \sum_{n=0}^{N-1} c_n \left(1 - i\Delta t E_n\right)^m \Phi_n(\{x_i\})$$
$$= \sum_{n=0}^{N-1} c_n \left|1 - i\Delta t E_n\right|^m e^{im \arctan(-\Delta t E_n)} \Phi_n(\{x_i\}). \tag{4.31}$$

It is easy to see that the explicit first order approximation is unconditionally unstable since $|1 - i\Delta t E_n| > 1$ for all eigenvalues E_n. Hence, all the expansion coefficients blow up exponentially in time, which cannot be accepted. Accordingly, the explicit method does not conserve the norm as is required by a unitary time evolution. If the wave function is normalized at $t = 0$ according to $\sum_{n=0}^{N-1} |c_n|^2 = 1$, it will acquire the norm

$$||\Psi(m\Delta t)|| = \sum_{n=0}^{N-1} |c_n|^2 |1 - i\Delta t E_n|^{2m}, \tag{4.32}$$

at time $t = m\Delta t$.

We shall now discuss how a modification of the FTCS method can solve the problems encountered above. A quick glance at Eq. (4.28) reveals that the FTCS method is somewhat asymmetric since it evaluates the Hamiltonian only once in the time interval $[t, t + \Delta t]$, namely at the prior end point in time, t. It is quite easy to modify the FTCS to a more symmetric form which turns out to be stable and unitary. Let us in Eq. (4.28) instead use the average action of the Hamiltonian at t and $t + \Delta t$ on the right hand side

$$
\begin{aligned}
i\frac{\Psi(x_i, t + \Delta t) - \Psi(x_i, t)}{\Delta t} = & -\frac{1}{4\Delta x^2}\left[\Psi(x_{i+1}, t) - 2\Psi(x_i, t) + \Psi(x_{i-1}, t)\right] \\
& + \frac{1}{2}V(x_i, \bar{t})\Psi(x_i, t) \\
& - \frac{1}{4\Delta x^2}\left[\Psi(x_{i+1}, t + \Delta t) - 2\Psi(x_i, t + \Delta t)\right. \\
& \left. + \Psi(x_{i-1}, t + \Delta t)\right] + \frac{1}{2}V(x_i, \bar{t})\Psi(x_i, t + \Delta t).
\end{aligned}
$$

$$(4.33)$$

Now we multiply by $-i\Delta t$ and collect terms that contain $\Psi(x, t + \Delta t)$ on the left and terms with $\Psi(x, t)$ on the right. Then it is straightforward to show that Eq. (4.33) is equivalent to the matrix equation

$$
\left(\mathbf{1} + \frac{1}{2}i\Delta t\boldsymbol{H}(\bar{t})\right) \cdot \boldsymbol{\Psi}(t + \Delta t) = \left(\mathbf{1} - \frac{1}{2}i\Delta t\boldsymbol{H}(\bar{t})\right) \cdot \boldsymbol{\Psi}(t), \qquad (4.34)
$$

where $\boldsymbol{H}(\bar{t}) = -(1/2)\boldsymbol{D}_2 + \mathrm{diag}[\boldsymbol{V}(\bar{t})]$ is a tridiagonal matrix. Equation (4.34) does not explicitly state how to obtain the vector $\boldsymbol{\Psi}(t + \Delta t)$ from the known vector $\boldsymbol{\Psi}(t)$, and hence, Eq. (4.34) is called an *implicit* equation for $\boldsymbol{\Psi}(t + \Delta t)$. Even though the solution is not explicit, it is not too hard to find. First we do the explicit matrix multiplication on the right hand side to obtain

$$
\left(\mathbf{1} + \frac{1}{2}i\Delta t\boldsymbol{H}(\bar{t})\right) \cdot \boldsymbol{\Psi}(t + \Delta t) = \boldsymbol{\Psi}^{(1)}(t + \Delta t/2). \qquad (4.35)
$$

Now the right hand side is a known intermediate vector, and $\boldsymbol{\Psi}(t + \Delta t)$ is then the solution of a linear system of equations. The corresponding matrix that defines the equations is tridiagonal. Such types of linear systems of equations are easily solved by an algorithm that scales just linearly with the dimensionality as will be shown in chapter 7. This way of finding $\boldsymbol{\Psi}(t + \Delta t)$ is called the Crank-Nicolson method [4]. This method is superior to the FTCS in several aspects. First, if we rewrite Eq. (4.33), we see that the

Crank-Nicolson method is more than first order accurate in Δt

$$
\begin{aligned}
\boldsymbol{\Psi}(t + \Delta t) &= \left(\mathbf{1} + \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right)^{-1} \cdot \left(\mathbf{1} - \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right) \cdot \boldsymbol{\Psi}(t) \\
&= \left(\mathbf{1} - \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) - \frac{1}{4} \Delta t^2 \boldsymbol{H}^2(\bar{t}) + \mathcal{O}(\Delta t^3) \right) \\
&\quad \cdot \left(\mathbf{1} - \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right) \cdot \boldsymbol{\Psi}(t) \\
&= \left(\mathbf{1} - i \Delta t \boldsymbol{H}(\bar{t}) - \frac{1}{2} \Delta t^2 \boldsymbol{H}^2(\bar{t}) + \mathcal{O}(\Delta t^3) \right) \cdot \boldsymbol{\Psi}(t). \quad (4.36)
\end{aligned}
$$

The expression in the parentheses is equal to the true exponential up to *second* order in Δt. Second, the Crank-Nicolson method turns out to be much more stable than the explicit method. The Crank-Nicolson equivalent to Eq. (4.31) is

$$
\begin{aligned}
\Psi^{(\mathrm{CN})}(\{x_i\}, m\Delta t) &= \sum_{n=0}^{N-1} c_n \left(\frac{1 - i\Delta t E_n/2}{1 + i\Delta t E_n/2} \right)^m \Phi_n(\{x_i\}) \\
&= \sum_{n=0}^{N-1} c_n e^{-im\arctan(\Delta t E_n)} \Phi_n(\{x_i\}), \quad (4.37)
\end{aligned}
$$

where the last equality holds because the complex numbers in the numerator and denominator have equal modulus and opposite phase. We now see from Eq. (4.37) that no coefficient blows up exponentially. Furthermore, Eq. (4.37) shows that the Crank-Nicolson method conserves norm in the energy representation. We can also prove that the Crank-Nicolson method is unitary in grid representation. We define the time evolution matrix as the product of the two non-singular commuting matrices

$$
\boldsymbol{U}^{(\mathrm{CN})} = \left(\mathbf{1} + \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right)^{-1} \cdot \left(\mathbf{1} - \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right). \quad (4.38)
$$

Since the two matrices commute, we can define the adjoint matrix from the reverse ordering

$$
\left[\boldsymbol{U}^{(\mathrm{CN})} \right]^\dagger = \left(\mathbf{1} - \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right)^{-1} \cdot \left(\mathbf{1} + \frac{1}{2} i \Delta t \boldsymbol{H}(\bar{t}) \right). \quad (4.39)
$$

Here we used that the matrix $\boldsymbol{H}(\bar{t})$ is real and symmetric, i.e. self-adjoint, for an equidistant centered in space finite difference method. The product of Eqs. (4.38) and (4.39) clearly reveals that $\boldsymbol{U}^{(\mathrm{CN})} \cdot \left[\boldsymbol{U}^{(\mathrm{CN})} \right]^\dagger = \mathbf{1}$, i.e., the Crank-Nicolson time evolution operator is represented by a unitary matrix and hence conserves norm according to the inner product defined on the grid, as discussed in Sec. 3.2.

It is very easy to impose box boundary conditions to the Crank-Nicolson method. We simply exclude the first and last grid point from the basis. The box boundary conditions are also straightforwardly fulfilled when a first order derivative is present in the Hamiltonian – a situation where other methods may fail as we shall show below.

4.3.2 The split operator

In this subsection, we will describe another approximation to the time evolution operator. First, we note that if the two terms in the exponential in Eq. (4.23) had been ordinary numbers, it would have been possible to write the exponential as the product of two exponentials. Now we try to do a similar splitting of the exponential for the operators

$$\mathscr{U}^{(1)}(t, t + \Delta t) = \exp\left(-i\Delta t \frac{p^2}{2}\right) \exp\left(-i\Delta t V(x, \bar{t})\right) \qquad (4.40)$$

The superscript "1" denotes that the time evolution operator is accurate to first order in Δt as we show below. The motivation for the operator splitting is that we may hope that it is easier to apply the kinetic- and potential energy terms separately than it is to apply the Hamiltonian directly.

How well does Eq. (4.40) reproduce the original unsplit exponential? In order to answer this question, we consider the Taylor expansion [$T \equiv p^2/2$ and $V \equiv V(x, \bar{t})$ for notational convenience]

$$e^{-i\Delta t T} e^{-i\Delta t V} = \left(\mathbf{1} - i\Delta t T - \frac{1}{2}\Delta t^2 T^2\right)\left(\mathbf{1} - i\Delta t V - \frac{1}{2}\Delta t^2 V^2\right) + \mathcal{O}(\Delta t^3)$$

$$= \mathbf{1} - i\Delta t\,(T + V) - \frac{\Delta t^2}{2}\left(T^2 + 2TV + V^2\right) + \mathcal{O}(\Delta t^3).$$

$$(4.41)$$

The expression above is now to be compared with the Taylor expansion for the unsplit exponential

$$e^{-i\Delta t (T+V)} = \mathbf{1} - i\Delta t\,(T + V) - \frac{\Delta t^2}{2}\,(T + V)^2 + \mathcal{O}(\Delta t^3)$$

$$= \mathbf{1} - i\Delta t\,(T + V) - \frac{\Delta t^2}{2}\left(T^2 + TV + VT + V^2\right) + \mathcal{O}(\Delta t^3).$$

$$(4.42)$$

Since T and V generally do not commute, the two expressions differ by

$$e^{-i\Delta t (T+V)} = e^{-i\Delta t T} e^{-i\Delta t V} + \mathcal{O}([T, V]\Delta t^2). \qquad (4.43)$$

We can in fact eliminate the second order error in Δt by a slight modification

of the exponential splitting to the more symmetric form

$$
\begin{aligned}
\mathscr{U}^{(2)}(t, t+\Delta t) =& e^{-i\Delta t T/2} e^{-i\Delta t V} e^{-i\Delta t T/2} \\
=& \left(\mathbf{1} - i\frac{\Delta t}{2}T - \frac{\Delta t^2}{8}T^2\right) \left(\mathbf{1} - i\Delta t V - \frac{1}{2}\Delta t^2 V^2\right) \\
& \times \left(\mathbf{1} - i\frac{\Delta t}{2}T - \frac{\Delta t^2}{8}T^2\right) + \mathcal{O}(\Delta t^3) \\
=& \mathbf{1} - i\Delta t\,(T+V) - \frac{\Delta t^2}{2}\left(T^2 + TV + VT + V^2\right) + \mathcal{O}(\Delta t^3),
\end{aligned}
$$
(4.44)

which is equal to $e^{-i\Delta t(T+V)}$ up to second order in Δt.

Since T and V are Hermitian, the split operator fulfils the unitarity requirement

$$
(e^{-i\Delta t T/2} e^{-i\Delta t V} e^{-i\Delta t T/2})^\dagger e^{-i\Delta t T/2} e^{-i\Delta t V} e^{-i\Delta t T/2} = \mathbf{1}.
$$
(4.45)

Split-operator spectral algorithm

The split operator is particularly useful if we expand the wave function in the grid-Fourier basis. With these globally defined basis functions, we refer to the following algorithm as the split-operator spectral (or Fourier) method. In the grid-Fourier basis we already know from Eqs. (3.8) and (3.17) how to apply operators that are functions of either the coordinate operator or the derivative operator, provided the grid spacing is equidistant. The grid-Fourier representation of the time evolution with the split operator is then

$$
\begin{aligned}
\boldsymbol{\Psi}(t+\Delta t) =\ & \boldsymbol{\mathcal{F}} \cdot \mathrm{diag}\left(e^{-i\frac{\Delta t}{2}\frac{\boldsymbol{p}^2}{2}}\right) \cdot \boldsymbol{\mathcal{F}}^\dagger \cdot \mathrm{diag}\left(e^{-i\Delta t \boldsymbol{V}(\bar{t})}\right) \\
& \cdot \boldsymbol{\mathcal{F}} \cdot \mathrm{diag}\left(e^{-i\frac{\Delta t}{2}\frac{\boldsymbol{p}^2}{2}}\right) \cdot \boldsymbol{\mathcal{F}}^\dagger \cdot \boldsymbol{\Psi}(t),
\end{aligned}
$$
(4.46)

with $\boldsymbol{\Psi}(t)$ and $\boldsymbol{V}(\bar{t})$ defined in Eqs. (4.26) and (4.27) and

$$
\boldsymbol{p} = -i\{k_0, k_1, \cdots, k_{N-1}\}^T, \quad k_j = \begin{cases} \frac{2\pi}{N\Delta x}j, & j = 0, 1, \cdots, \frac{N}{2}-1 \\ \frac{2\pi}{N\Delta x}j - \frac{2\pi}{\Delta x}, & j = \frac{N}{2}, \cdots, N-1 \end{cases}.
$$
(4.47)

The exponentiation and power operations on the vectors are to be done elementwise.

Now it is clear why the split operator is much easier to apply than directly exponentiating the Hamiltonian. The Fourier transformation is independent of time and hence, we do not need to perform an expensive $\mathcal{O}(N^3)$ diagonalization in every time step. It is not even necessary to perform any $\mathcal{O}(N^2)$ matrix multiplication, since the discrete Fourier transformation can be done by the Fast Fourier transformation. Furthermore, operations on diagonal

matrices scale just linearly with the basis size. The overall computational complexity for the split-operator Fourier method therefore scales with the basis size as $\mathcal{O}(N \log_2 N)$.

The split-operator spectral method can be used for a great number of different problems. We do not assume any specific form of the time dependent potential. We must, however, remember that the Fourier method has some limitations for singular potentials (see Sec. 3.3.1) and for efficient representation of functions that vary on several length scales (see Sec. 3.6.1).

It is possible to use the split-operator spectral method for problems that require box boundary conditions. In that case we simply substitute the Fourier transformations by sine transformations and replace the vector $\boldsymbol{p}$ according to Eq. (3.21). This method works because the Hamiltonian only contains the *second* order derivative which is diagonalized by the sine representation. If the Hamiltonian includes a first order derivative terms, e.g. $\mathcal{H}'(t) = -iA(t)(d/dx)$, we *cannot* use a sine transformation. We easily see why this fails. The canonical momentum $p = -i(d/dx)$ is the generator of coordinate translations. Thus, we may see the operator $e^{-i\Delta t(-iA(t)(d/dx))}$ as an operator that translates a wave function by the amount $\Delta t A(t)$

$$e^{-i\Delta t(-iA(t)(d/dx))} \Psi(x,t) = \Psi(x + \Delta t A(t), t). \qquad (4.48)$$

Since the operator above must be part of the time evolution operator, we see that if the wave function satisfies box boundary conditions at time t it does not fulfil such conditions at time $t + \Delta t$; the zero points of the function are simply displaced away from the box boundaries.

On a final note, the split-operator Fourier method leads to an exactly unitary time evolution with inner product defined by the trapezoidal rule.

4.3.3   A combined split operator Crank-Nicolson method

Both the operator splitting and the Crank-Nicolson approximation are accurate up to second order in Δt. It is therefore also accurate to second order in Δt to use the following approximation

$$
\begin{aligned}
e^{-i\Delta t(T+V)} &= (1 + i\frac{\Delta t}{4}T)^{-1}(1 - i\frac{\Delta t}{4}T)e^{-iV\Delta t}(1 + i\frac{\Delta t}{4}T)^{-1}(1 - i\frac{\Delta t}{4}T) \\
&\quad + \mathcal{O}(\Delta t^3), \qquad\qquad (4.49)
\end{aligned}
$$

i.e., to split the exponential and then use the Crank-Nicolson method for the kinetic part of the Hamiltonian. This method scales linearly with the basis size, and scales therefore slightly better – by a factor of $\log_2 N$ – than the split-operator spectral method. Equation (4.49) is perhaps not so interesting. With the same order of accuracy, we could have used just a single Crank-Nicolson step for the total Hamiltonian. There are, however, certain situations where combined split-operator Crank-Nicolson method becomes

very useful. These situations involve more complicated Hamiltonians that are encountered in multidimensional problems as we shall see in chapter 5.

Generally, we can take any second order accurate split operator scheme with each of the factorized parts of the Hamiltonian being applied either on the exponential form or by the Crank-Nicolson method. Any such approximation is considered as being second order accurate.

4.3.4   The time step size

Until now, we have claimed that several approximations to the time evolution operator are second order accurate in the time step size Δt. Although the definition of second order accuracy is quite loose, it is clear that any of the approximations will converge towards the exact time evolution operator if we let $\Delta t \to 0$. But what is the effect of choosing a finite value of Δt? In order to study this question, we will again use the example of the radial equation of hydrogen as in Sec. 3.6.1. We obtain a one-dimensional Schrödinger equation with the effective time-independent Hamiltonian $\mathcal{H} = p^2/2 + l(l+1)/(2r^2) - 1/r$. If we start out in an eigenstate $u_{nl}(r)$ of this Hamiltonian at $t = 0$, the *true* time evolution will of course just amount to an energy phase factor as shown in Sec. 4.2. This means that the *correlation* function

$$\text{Corr}(t) = \langle u(r,0)|u(r,t)\rangle, \tag{4.50}$$

will have unit norm at all times for the exact time dependent wave function. If we now use an approximate scheme to obtain the time dependent function $u(r,t)$, we may encounter the situation that the overlap with the initial state becomes imperfect and, correspondingly, the norm of the correlation becomes less than unity. Hence, we can use the correlation function as a measure of how accurately $u(r,t)$ can be obtained from an approximate time evolution operator. As a technical detail, we will use the function $1 - |\text{Corr}(t)|^2$ as the measure of the accuracy instead of the correlation function itself. The former function will be zero for the exact wave function and small and positive if the approximate wave function is close to the exact.

Figure 4.2 presents the values of $1 - |\text{Corr}(t)|^2$ with the time evolution operator approximated by the split-operator Fourier method. The results shown in panel (a) are obtained for 1024 grid points while 2048 points are used in obtaining the results in panel (b). In both cases, the radial grid extends to 100 a.u. and, hence, the density of grid points is doubled in panel (b) compared to panel (a). From Fig. 4.2 we see that different choices for the time step size either lead to stable or to unstable solutions. The unstable solutions are characterized by the behaviour that the wave function gradually disappears from the ground state as time goes by. The transition between the stable and unstable regime occurs abruptly at a certain critical time step size. Additionally, the onset for the unstable regime depends on the grid spacing such that a smaller grid spacing requires a smaller time step
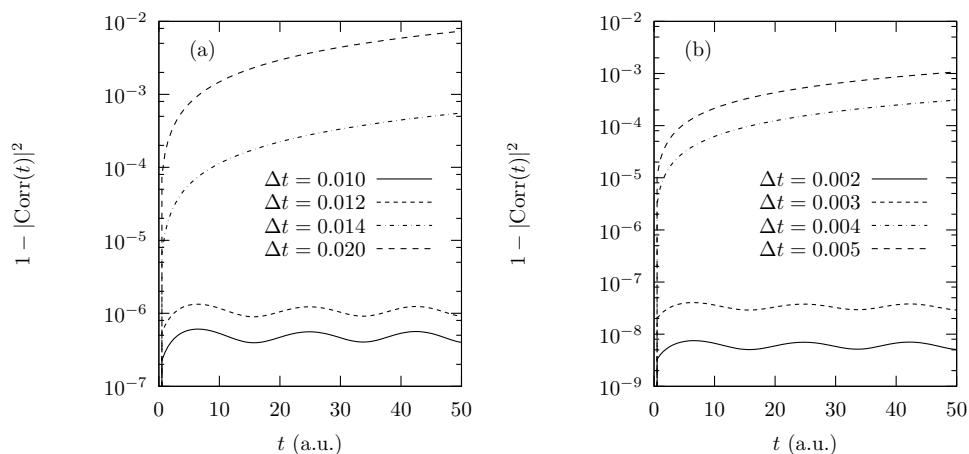
Figure 4.2: The function $1 - |\mathrm{Corr}(t)|^2$ for various time step sizes as indicated in the panels. The time dependent wave function is obtained by the split-operator Fourier method. The number of grid points is 1024 in panel (a) and 2048 in panel (b). The grid extends to $100\,\mathrm{a.u.}$ in both panels.

for a stable solution. We believe that the un-stability occurs due to the fact that the potential operator attains a large value near the Coulomb singularity near the origin. In grid representation, the potential is applied by multiplication by the set of values $\{e^{-iV(r_i)\Delta t}\}$ on the grid points. If the argument becomes too large, the resulting propagated wave function may not be bandwidth limited. Note that for the present case that a dense grid implies that grid points are located close to the Coulomb singularity at the origin. As we subsequently make a Fourier transformation on a function that is not bandwidth limited, we encounter problems known as aliasing [4], and the final outcome is quite unpredictable. It will probably be possible to set up strict quantitative criteria for how large step sizes that can be used in order to obtain stability for a given potential and grid spacing. We have simply used the empirical approach whereby the largest possible time step size is found by trial and error. We finally note that the time step size that is required for a stable propagation by a second order approximation is much smaller than typical time scales of external perturbations. For example, we can compare $\Delta t \approx 0.01\,\mathrm{a.u.}$ from Fig. 4.2 with $T_{\mathrm{period}} = 110\,\mathrm{a.u.}$ being the optical period of $800\,\mathrm{nm}$ light. Accordingly, it will be an excellent approximation to regard the Hamiltonian as being constant within Δt as proposed in Eq. (4.12).

4.3.5   Alternative approximation schemes

It is usually the second order accuracy in Δt of the split-operator and Crank-Nicolson methods that defines how the time evolution operator must be

subdivided in order to obtain a converged solution. If one could find a more accurate expression for the operator exponential, it would be possible to use a much larger time step when subdividing the time evolution operator. One method to approximate the matrix exponential is known as the *Krylov iterative* method. The detailed description of this theory can be found in Ref. [8] and will not be discussed here. We only mention that the method often allows one to calculate the approximate action of a large sparse matrix exponential on a vector in a number of steps that is much smaller than the $\mathcal{O}(N^3)$ required for diagonalization. Being an iterative procedure, the actual number of operations will depend on the requested accuracy. We have performed preliminary tests with the software routines described in Ref. [8] that seem to work well. A successful implementation is also reported in Ref. [9]. Our experience with the Krylov method is, however, still quite limited and has not been used for any of the numerical calculations presented in this thesis.

4.4   The initial state

As mentioned in Sec. 4.1, the time dependent Schrödinger equation is usually formulated as an initial value problem which requires an initial wave function. The typical situation that we will study pertains to a situation where the system evolves under a unperturbed time independent Hamiltonian $\mathcal{H}_0 = \boldsymbol{p}^2/2 + V_0(\boldsymbol{r})$ until the time t_0 where a time dependent interaction is turned on. Most often we consider a system that is initially in the lowest energy eigenstate, the ground state, of $\mathcal{H}_0$. The ground state can be found from a diagonalization procedure as described in Sec. 3.6. This requires a separate calculation and an entirely new piece of code. Furthermore, the diagonalization can be quite time consuming if the basis is large.

As long as we are only interested in the ground state, we can use an alternative method referred to as propagation in reverse imaginary time. This method relies on some time evolution scheme as described in the previous sections. If we have written a code that performs any such propagation, the same code can be used to obtain the ground state with only a minimal modification. As a purely mathematical trick, we substitute the time $t \to -i\tau$ in the time dependent Schrödinger equation. Any trial function $\Psi^{(\mathrm{Im})}(\boldsymbol{r}, 0)$ then evolves with the parameter τ under $\mathcal{H}_0$ as in Eq. (4.10)

$$\mathscr{U}^{(\mathrm{Im})}(\tau + \Delta\tau, \tau) = e^{-\mathcal{H}_0 \Delta\tau}. \tag{4.51}$$

The propagation schemes are modified accordingly, but the symbol τ is now just a real parameter that cannot be associated with a physical time. Furthermore, $\mathscr{U}^{(\mathrm{Im})}(\tau + \Delta\tau, \tau)$ is clearly non-unitary and the propagated wave function does not represent the physical state of a particle. By any of the modified propagation methods, we assume that the evolution is determined

along the real parameter axis τ such that $\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)$ is known numerically for any value of τ. At this point, it may seem quite unclear why this function will be of any use. But if we now *formally* expand the trial function in the complete set of eigenstates $\{\Phi_m\}$ of $\mathcal{H}_0$ which are unknown at this stage the formal evolution along the τ axis of the trial function becomes

$$\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau) = e^{-\mathcal{H}_0\tau}\Psi^{(\mathrm{Im})}(\boldsymbol{r},0) = \sum_m e^{-E_m\tau}c_m\Phi_m(\boldsymbol{r}). \qquad (4.52)$$

Now we see that the different components in the spectral decomposition of $\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)$ decay or explode exponentially as $\tau \to \infty$. If the spectrum contains both negative and positive eigenvalues, the component corresponding to the most negative eigenvalue will explode most rapidly and dominate the solution as τ grows large. Similarly, if the spectrum only contains positive eigenvalues, it is the component corresponding to the smallest eigenvalue that will decay at the slowest rate and dominate the solution as $\tau \to \infty$. In both cases, $\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)$ will converge towards the ground state of $\mathcal{H}_0$ after proper renormalization. An estimate for the ground state energy, can be obtained from the decay/grow rate of the wave function

$$\langle\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau+\Delta\tau)|\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau+\Delta\tau)\rangle = e^{-2E_0\Delta\tau}\langle\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)|\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)\rangle, \qquad (4.53)$$

or

$$E_0 = -\frac{1}{2\Delta\tau}\ln\left(\frac{\langle\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau+\Delta\tau)|\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau+\Delta\tau)\rangle}{\langle\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)|\Psi^{(\mathrm{Im})}(\boldsymbol{r},\tau)\rangle}\right). \qquad (4.54)$$

Equation (4.53) assumes that τ is large enough such that all components in excited states are much smaller than the ground state component. In practice, if we start from any initial state, we see that the expression on the right hand side of Eq. (4.54) converges towards E_0 as the parameter τ grows and the excited components decay.

In special cases it is possible use propagation in imaginary time to find certain excited states. If we by some clever guess can make sure that the initial trial function $\Psi^{(\mathrm{Im})}(\boldsymbol{r},0)$ is orthogonal to the ground state, the ground state term in Eq. (4.52) vanishes ($c_0 = 0$). Accordingly, the trial function converges to the first excited state, unless of course the trial wave function is also orthogonal to that state. For example, if the Hamiltonian is invariant under space inversion, we know that the spectrum supports even and odd parity eigenstates and that the lowest energy eigenstate is even. In principle, we can now find the lowest lying odd eigenstate by choosing an odd trial function which will be orthogonal to the even ground state. Unfortunately, this method can fail for a real numerical propagation. If, by even a slight numerical roundoff error, we accidentally populate the even ground state during the propagation, it will again be this even ground state that dominates at large τ. The safest way to circumvent this problem is

to explicitly restrict the whole calculation to take only the odd states into account. For example, if we use the grid-Legendre basis, Sec. 3.4, we choose only to include the odd Legendre polynomials in the basis.

It is not feasible to use imaginary time propagation to find excited states beyond the lowest state of a given symmetry. Diagonalization seems to be a better approach if the full spectrum is desired.

Chapter 5

Split operator in spherical coordinates

In this chapter we demonstrate how the dynamics of one electron can be solved by integration of the time dependent Schrödinger equation. In particular, we consider an electron which is initially bound in an atomic or molecular potential and is subject to an external radiation source.

Any realistic description of the electronic dynamics must take all three spatial dimensions into account. The most straight-forward way to represent a three dimensional wave function is by a Cartesian grid defined along each of the coordinate axes. The Cartesian grid is, however, not the optimal choice for atomic and molecular processes. For spherically symmetric potentials, it is much more natural to use spherical coordinates since it is possible to separate analytically the radial and the angular variables. Spherical coordinates are even preferred for many non-spherically symmetric problems, e.g., if an asymmetry is introduced by a linearly polarized field or a non-spherical molecular potential.

5.1   The reduced wave function in length and velocity gauge

The length and velocity gauge Hamiltonians are given by Eqs. (2.52) and (2.53) for a charged particle in an electrostatic field and an external radiation field described in the dipole approximation. For the present case, the static field is due to the atomic or molecular potential, which we shall denote $V^{(M)}(\boldsymbol{r})$. For the electron ($M = 1, q = -1$) we write Eqs. (2.52) and (2.53) in the general form

$$\mathcal{H} = -\frac{\boldsymbol{\nabla}^2}{2} - i\boldsymbol{A}(t) \cdot \boldsymbol{\nabla} + V(\boldsymbol{r}, t), \qquad (5.1)$$

with

$$\boldsymbol{A}(t) = \begin{cases} \boldsymbol{0} \\ \boldsymbol{A}_{\mathrm{VG}}(t) \end{cases}, \quad V(\boldsymbol{r},t) = \begin{cases} \boldsymbol{r} \cdot \boldsymbol{E}(t) + V^{(M)}(\boldsymbol{r}), & \text{Length} \\ V^{(M)}(\boldsymbol{r}), & \text{Velocity} \end{cases}. \quad (5.2)$$

The full wave function $\Psi(\boldsymbol{r},t)$ evolves according to the usual time dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \left(-\frac{\boldsymbol{\nabla}^2}{2} - i\boldsymbol{A}(t) \cdot \boldsymbol{\nabla} + V(\boldsymbol{r},t)\right)\Psi(\boldsymbol{r},t). \quad (5.3)$$

As mentioned above, we will express this equation in spherical coordinates $\boldsymbol{r} = (r,\theta,\phi)$. It is now convenient to factorize out a factor of r from the wave function and write the remaining "reduced" part according to

$$\Phi(r,\theta,\phi,t) = r\Psi(r,\theta,\phi,t). \quad (5.4)$$

If the true wave function Ψ is not divergent at the origin, the reduced wave function must obey the boundary condition $\Phi(0,\theta,\phi,t) = 0$. Inserting this expression into Eq. (5.3) gives a dynamical equation for the reduced wave function

$$\begin{aligned} i\frac{\partial}{\partial t}\frac{\Phi(\boldsymbol{r},t)}{r} &= \left(-\frac{\boldsymbol{\nabla}^2}{2} - i\boldsymbol{A}(t)\cdot\boldsymbol{\nabla} + V(\boldsymbol{r},t)\right)\frac{\Phi(\boldsymbol{r},t)}{r} \\ &= \left[-\frac{1}{2r}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^3} - i\boldsymbol{A}(t)\cdot\left(\frac{1}{r}\boldsymbol{\nabla} - \frac{\hat{\boldsymbol{r}}}{r^2}\right) + \frac{V(\boldsymbol{r},t)}{r}\right]\Phi(\boldsymbol{r},t), \end{aligned}$$
$$(5.5)$$

where L^2 is the usual angular momentum operator

$$L^2 = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \quad (5.6)$$

We multiply by r and obtain a Schrödinger-like equation for the reduced wave function

$$i\frac{\partial}{\partial t}\Phi(\boldsymbol{r},t) = \left[-\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} - i\boldsymbol{A}(t)\cdot\left(\boldsymbol{\nabla} - \frac{\hat{\boldsymbol{r}}}{r}\right) + V(\boldsymbol{r},t)\right]\Phi(\boldsymbol{r},t). \quad (5.7)$$

Thus, the time evolution of the reduced wave function looks much like the time evolution of the real wave function. We only need to replace the Hamiltonian by a reduced Hamiltonian. Due to this similarity, we can use all the methods developed previously to solve the time evolution for the reduced wave function just by replacing the Hamiltonian by the reduced Hamiltonian and additionally imposing box boundary conditions for the radial coordinate at the origin.

For the length and velocity gauges, the reduced Hamiltonians read

$$\mathcal{H}^{(\text{LG})} = -\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} + \boldsymbol{r}\cdot\boldsymbol{E}(t) + V^{(M)}(\boldsymbol{r}), \tag{5.8}$$

$$\mathcal{H}^{(\text{VG})} = -\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} - i\boldsymbol{A}_{\text{VG}}(t)\cdot\left(\boldsymbol{\nabla} - \frac{\hat{\boldsymbol{r}}}{r}\right) + V^{(M)}(\boldsymbol{r}). \tag{5.9}$$

Below we shall often refer to the *reduced* Hamiltonian and wave function simply as "the Hamiltonian" and "the wave function" if the meaning is clear from the context.

5.2   Choice of basis and space discretization

In the present numerical implementation, we use an equidistant grid with grid spacing Δr in the radial coordinate as the basic representation. We take the number of radial grid points to be n_r, corresponding to radial grid points distributed between $r_{\min} = r_0 = 0$ and $r_{\max} = r_{n_r-1} = (n_r - 1)\Delta r$. Box boundary conditions are imposed on the reduced wave function at both end points. The radial grid representation has the advantage that it is well suited for a description of bound states that decay exponentially as well as continuum states that oscillate asymptotically. The angular variables will be expanded in a finite number of spherical harmonics – an expansion which is often referred to as the partial wave expansion. The spherical harmonics are defined according to the standard normalization

$$\begin{aligned} Y_{lm}(\theta,\phi) &= (-1)^{\frac{m+|m|}{2}}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}P_l^{|m|}(\cos\theta)e^{im\phi} \\ &= (-1)^{\frac{m+|m|}{2}}\bar{P}_l^{|m|}(\cos\theta)\frac{e^{im\phi}}{\sqrt{2\pi}}, \end{aligned} \tag{5.10}$$

with $P_l^{|m|}(\cos\theta)$ being an associated Legendre function of $\cos\theta$. For each radial grid point we therefore use the expansion for the reduced wave function

$$\Phi(r_i,\theta,\phi,t) = \sum_{l=0}^{l_{\max}}\sum_{m=-l}^{l} f_{lm}(r_i,t)Y_{lm}(\theta,\phi), \tag{5.11}$$

with boundary conditions $f_{lm}(r_0,t) = f_{lm}(r_{n_r-1},t) = 0$. The expansion in spherical harmonics turns out to converge reasonably fast with an increasing angular quantum number for most applications. Hence, the truncation at $l_{\max}$ can be kept quite low. For a given truncation $l_{\max}$, the number of basis functions and expansion coefficients are

$$n_{Y_{lm}} = \sum_{l=0}^{l_{\max}}\sum_{m=-l}^{l} = (l_{\max}+1)^2. \tag{5.12}$$

In the spherical coordinate representation, inner products are defined between ordinary non-reduced wave functions according to

$$\langle \Psi^{(1)} | \Psi^{(2)} \rangle = \int_0^\infty \int_0^\pi \int_0^{2\pi} [\Psi^{(1)}(r,\theta,\phi)]^\star \Psi^{(2)}(r,\theta,\phi) r^2 \sin\theta d\phi d\theta dr. \quad (5.13)$$

In our discrete basis, we use the trapezoidal integration rule in the radial coordinate and an analytical evaluation of angular integrals over spherical harmonics

$$\langle \Psi^{(1)} | \Psi^{(2)} \rangle = \sum_{i=0}^{n_r-1} \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} [f_{lm}^{(1)}(r_i)]^\star f_{lm}^{(2)}(r_i) \Delta r. \quad (5.14)$$

Note that the factor r^2 disappears as we use the reduced wave function. The boundary terms at r_0 and r_{n_r-1} need not to be weighted by a factor $\frac{1}{2}$ since both values are zero when box boundary conditions are imposed on the reduced wave function.

5.3   Cylindrically symmetric problems

Some systems exhibit cylindrical (azimuthal) symmetry around some axis. Examples include an atom exposed to a linearly polarized field or a linear molecule with the molecular axis parallel to the polarization vector of a linearly polarized field. Cylindrically symmetric problems are much easier to solve than general three dimensional problems. Therefore, before turning to the general case in Sec. 5.4, we first discuss in detail how to solve the time dependent Schrödinger equation for problems that exhibit cylindrical symmetry. This simplified problem has also been studied in Refs. [3, 6, 10].

We define the coordinate axis so that the z axis coincides with the symmetry axis. The Hamiltonian is then independent of the azimuthal angle ϕ. For such a Hamiltonian, $[\mathcal{H}, L_z] = 0$ means that L_z, the z component of the angular momentum, is a constant of motion. Accordingly, the azimuthal quantum number m is a conserved quantity. If only a single m state is populated initially, no population can be transferred to other m states by the azimuthally symmetric Hamiltonian. Hence, we can write the expansion in spherical harmonics as

$$\Phi_m(r_i, \theta, \phi, t) = \sum_{l=|m|}^{l_{\max}} f_{lm}(r_i, t) \bar{P}_l^{|m|}(\cos\theta) \frac{e^{im\phi}}{\sqrt{2\pi}}. \quad (5.15)$$

The ϕ dependence is now trivial and the remaining nontrivial part of the wave function can be written as a two dimensional function

$$\tilde{\Phi}_m(r_i, \zeta, t) = \Phi_m(r_i, \theta, \phi, t) \sqrt{2\pi} e^{-im\phi} = \sum_{l=|m|}^{l_{\max}} f_{lm}(r_i, t) \bar{P}_l^{|m|}(\zeta), \quad (5.16)$$
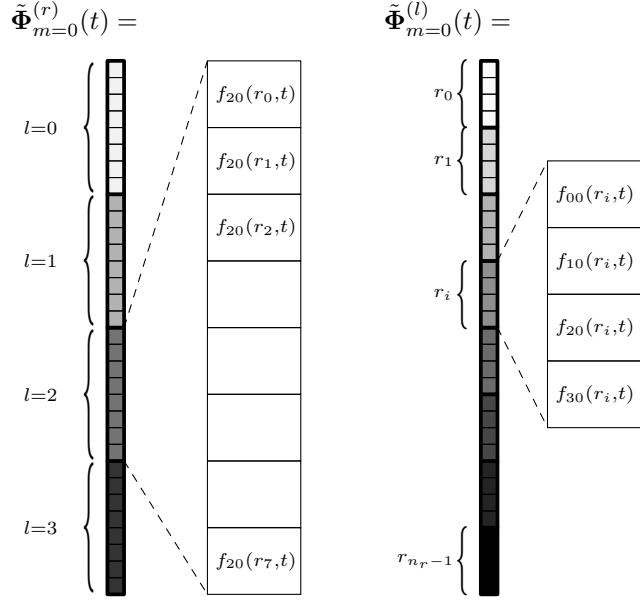
Figure 5.1: Vector representation of $\tilde{\Phi}_m(r, \zeta, t)$ for $m = 0$ in r-major and l-major ordering for $n_r = 8$ and $l_{\max} = 3$.

with the angular variable $\zeta = \cos\theta$ being defined on $-1 \leq \zeta \leq 1$. As we show below, it will be sufficient to solve the time evolution of the two-dimensional function, and hence, we need to express this function as a vector representation in the radial grid/associated Legendre basis. One complete vector representation of the two dimensional wave function is

$$
\begin{aligned}
\tilde{\Phi}_m^{(r)}(t) = \{ & f_{|m|m}(r_0, t), f_{|m|m}(r_1, t), \cdots, f_{|m|m}(r_{n_r-1}, t), \\
& f_{|m|+1,m}(r_0, t), f_{|m|+1,m}(r_1, t), \cdots, f_{|m|+1,m}(r_{n_r-1}, t), \cdots \\
& f_{l_{\max},m}(r_0, t), f_{l_{\max},m}(r_1, t), \cdots, f_{l_{\max},m}(r_{n_r-1}, t) \}^T.
\end{aligned} \tag{5.17}
$$

The superscript "r" denotes r-major ordering, i.e., the index corresponding to the radial coordinate varies most rapidly, as illustrated in Fig. 5.1. The total dimension of the vector is $n_r n_l$ with $n_l = l_{\max} + 1 - |m|$ being the angular dimension. It will also be useful to introduce the following l-major ordering as another permutation

$$
\begin{aligned}
\tilde{\Phi}_m^{(l)}(t) = \{ & f_{|m|m}(r_0, t), f_{|m|+1,m}(r_0, t), \cdots, f_{l_{\max},m}(r_0), \\
& f_{|m|m}(r_1, t), f_{|m|+1,m}(r_1, t), \cdots, f_{l_{\max},m}(r_1), \cdots \\
& f_{|m|m}(r_{n_r-1}, t), f_{|m|+1,m}(r_{n_r-1}, t), \cdots, f_{l_{\max},m}(r_{n_r-1}, t) \}^T.
\end{aligned} \tag{5.18}
$$

The transformation between Eqs. (5.17) and (5.18) is just a simple permutation. The expansion coefficient $f_{lm}(r_i, t)$ is at index $i + n_r(l - |m|)$ in the

r-major ordering and at index $(l - |m|) + n_l i$ in the l-major ordering. The two representations are shown in Fig. 5.1 for a small basis that contains 8 radial points and with $l_{\max} = 3$.

Inner products are calculated as in Eq. (5.13). The integration over the azimuthal angle is readily accomplished and the remaining part is

$$
\begin{aligned}
\langle \Psi_m^{(1)} | \Psi_m^{(2)} \rangle &= \int_0^\infty \int_{-1}^1 [\tilde{\Phi}_m^{(1)}(r_i, \zeta, t)]^\star \tilde{\Phi}_m^{(2)}(r_i, \zeta, t) d\zeta dr \\
&= \sum_{i=0}^{n_r - 1} \sum_{l=|m|}^{l_{\max}} [f_{lm}^{(1)}(r_i, t)]^\star f_{lm}^{(2)}(r_i, t) \Delta r, \qquad (5.19)
\end{aligned}
$$

where we used the orthogonality of the associated Legendre functions at the last equality. Furthermore, the trapezoidal integration rule is used in the radial direction. Note that the inner product is equal to the ordinary complex dot product between vectors on the form (5.17) or (5.18), except for the constant factor Δr. This means that an operator which is represented by a unitary matrix is guaranteed to conserve norm according to the inner product (5.19).

The cylindrically symmetric length and velocity gauge Hamiltonians take the form

$$
\mathcal{H}^{(\mathrm{LG})} = -\frac{1}{2} \frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} + rE(t)\zeta + V^{(M)}(r, \zeta), \qquad (5.20)
$$

$$
\mathcal{H}^{(\mathrm{VG})} = -\frac{1}{2} \frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} + iA_{\mathrm{VG}}(t) \left(\frac{\zeta - (1 - \zeta^2)\frac{\partial}{\partial \zeta}}{r} - \zeta \frac{\partial}{\partial r} \right) + V^{(M)}(r, \zeta). \qquad (5.21)
$$

These operators must act on the three dimensional reduced wave function $\Phi_m(r, \theta, \phi, t)$, Eq. (5.15). In order to find a dynamical equation for the two dimensional wave function, we insert the expression $\Phi_m(r, \theta, \phi, t) = \tilde{\Phi}_m(r, \zeta) e^{im\phi}/\sqrt{2\pi}$ in Eq. (5.7). In either gauge we obtain a dynamical equation for the two-dimensional wave function on the form

$$
i \frac{\partial}{\partial t} \tilde{\Phi}_m(r, \zeta, t) = \begin{cases} \mathcal{H}^{(\mathrm{LG})} \tilde{\Phi}_m(r, \zeta, t) \\ \mathcal{H}^{(\mathrm{VG})} \tilde{\Phi}_m(r, \zeta, t) \end{cases}, \qquad (5.22)
$$

with the "doubly reduced" Hamiltonians being exactly identical to (5.20) and (5.21), except that the angular momentum operator given in Eq. (5.6) is now replaced by

$$
L^2 = -\frac{\partial}{\partial \zeta} \left((1 - \zeta^2) \frac{\partial}{\partial \zeta} \right) + \frac{m^2}{1 - \zeta^2}. \qquad (5.23)
$$

Again, Eq. (5.22) has the same structure as an ordinary time dependent Schrödinger equation, and we can use the usual methods outlined in chapter 4 to solve the time evolution of $\tilde{\Phi}_m$. The following sections describe the time evolution for the length and velocity gauge Hamiltonians.

5.3.1 Length gauge split operator for cylindrically symmetric Hamiltonians

In the length gauge, we regard the Hamiltonian as a sum of three terms. We use the split-operator technique to factorize the time evolution operator. The operators can be split in several different ways; here we use the following form

$$\mathscr{U}(t+\Delta t, t) = e^{i\frac{\Delta t}{2}\frac{1}{2}\frac{\partial^2}{\partial r^2}} e^{-i\frac{\Delta t}{2}\frac{L^2}{2r^2}} e^{-i\Delta t V(r,\zeta,\bar{t})} e^{-i\frac{\Delta t}{2}\frac{L^2}{2r^2}} e^{i\frac{\Delta t}{2}\frac{1}{2}\frac{\partial^2}{\partial r^2}} + \mathcal{O}(\Delta t^3). \tag{5.24}$$

The three distinct factors will be discussed in detail below.

Radial kinetic operator

For each value of l, it is now convenient to regard each of the vectors

$$\boldsymbol{f}_{lm}(t) = \{f_{lm}(r_0, t), f_{lm}(r_1, t), \cdots, f_{lm}(r_{n_r-1}, t)\}^T, \tag{5.25}$$

as grid representations of the one-dimensional functions $f_{lm}(r, t)$. Since these functions are radial expansion functions for the reduced wave function, they must obey box boundary conditions.

The radial kinetic operator $(-1/2)\partial^2/\partial r^2$ does not act on the angular variable. It only acts *individually* on each of the one-dimensional radial functions. In the language of matrix representations, this means that the operator $(-1/2)\partial^2/\partial r^2$ is block diagonal in r-major ordering as shown in Fig. 5.2. The separation into one-dimensional problems is very convenient since we already know how to apply a kinetic energy term in one-dimensional problems.

The Crank-Nicolson method (4.38) for the kinetic term alone gives

$$e^{-i\frac{\Delta t}{2}\left(-\frac{1}{2}\frac{\partial^2}{\partial r^2}\right)} f_{lm}(r,t) \doteq \left(\boldsymbol{1} - \frac{1}{8}i\Delta t \boldsymbol{D}_2\right)^{-1} \cdot \left(\boldsymbol{1} + \frac{1}{8}i\Delta t \boldsymbol{D}_2\right) \cdot \boldsymbol{f}_{lm}(t). \tag{5.26}$$

Alternatively, we can use the Fourier spectral method with box boundary conditions

$$e^{-i\frac{\Delta t}{2}\left(-\frac{1}{2}\frac{\partial^2}{\partial r^2}\right)} f_{lm}(r,t) \doteq \boldsymbol{\mathcal{F}}^{\sin} \cdot \operatorname{diag}\left(e^{-i\frac{\Delta t}{2}\frac{k^2}{2}}\right) \cdot \boldsymbol{\mathcal{F}}^{\sin} \cdot \boldsymbol{f}_{lm}(t), \ \ k_j = \frac{\pi}{n_r \Delta r}j. \tag{5.27}$$

The total computational complexity for application of the radial kinetic operator scales as $\mathcal{O}\left(n_r \log_2(n_r) n_l\right)$, since we must perform a one-dimensional sine transform for each of the $n_l = l_{\max} + 1 - |m|$ values of l.

All results presented in this thesis are obtained with the Fourier spectral method.

$(-1/2)\partial^2/\partial r^2 \doteq$ $\tilde{\boldsymbol{\Phi}}^{(r)}_{m=0}(t) =$
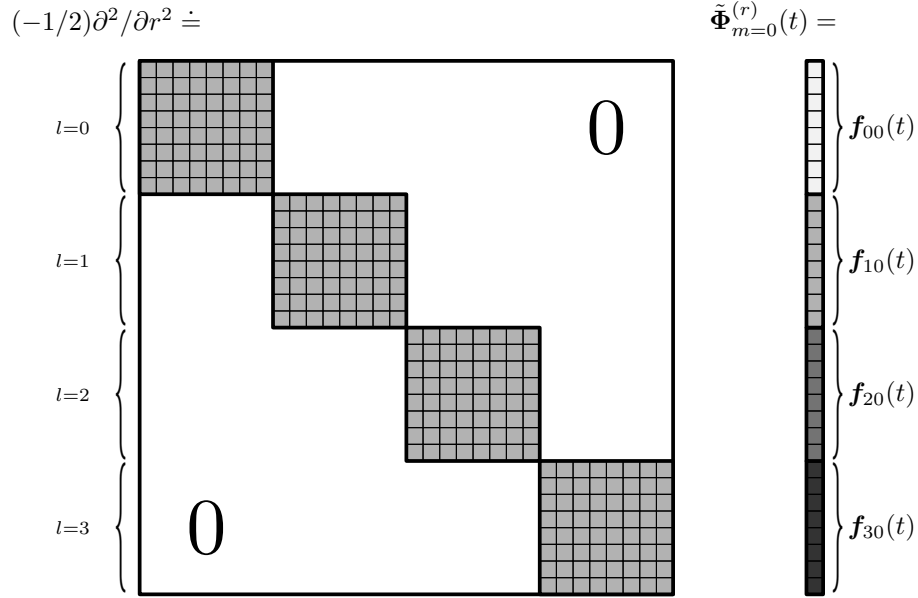

Figure 5.2: Matrix representation of $(-1/2)\partial^2/\partial r^2$ in r-major ordering. If the derivative is approximated by the three-point finite difference, the blocks will be tridiagonal. Each of the blocks are identical. The vector $\boldsymbol{\Phi}^{(r)}(t)$ is decomposed in a direct sum of one-dimensional vectors $\boldsymbol{f}_{lm}(t)$.

The centrifugal operator

The centrifugal term is particularly easy in our representation. The radial grid diagonalizes the operator r^2 in the denominator. Furthermore, L^2 is diagonal in the spherical harmonics representation with diagonal elements $l(l+1)$

$$e^{-i\frac{\Delta t}{2}\frac{L^2}{2r^2}}\tilde{\Phi}_m(\boldsymbol{r},t) \doteq \text{diag}(e^{-i\frac{\Delta t}{2}\frac{1}{2r^2}\otimes\boldsymbol{l}^2}) \cdot \tilde{\boldsymbol{\Phi}}^{(r)}_m(t) = \text{diag}(e^{-i\frac{\Delta t}{2}\boldsymbol{l}^2\otimes\frac{1}{2r^2}}) \cdot \tilde{\boldsymbol{\Phi}}^{(l)}_m(t),$$
$$(5.28)$$

with $\boldsymbol{l}^2 = \{|m|(|m|+1), (|m|+1)(|m|+2), \cdots, l_{\max}(l_{\max}+1)\}$ and $\boldsymbol{r} = \{r_0, r_1, \cdots, r_{n_r-1}\}$. The equation above is really just a very complicated way of writing that each of the expansion coefficients that constitute the vector representation $\boldsymbol{\Phi}(t)$ acquires the phase

$$f_{lm}(r_i, t) \rightarrow e^{-i\frac{\Delta t}{2}\frac{l(l+1)}{2r_i^2}} f_{lm}(r_i, t). \quad\quad\quad (5.29)$$

The potential operator in grid representation

The last operator to be applied is the potential which is a function of the coordinate operator. Hence, it would be represented by a diagonal matrix, if we represent the wave function in a grid representation in both spatial

variables, r and $\zeta = \cos\theta$. Since the radial coordinate is already represented on a grid, we will of course use the same radial grid in both representations. Thus, the transformation is only nontrivial from the spherical harmonics representation to the ζ (or θ) grid representation. The fact that we use the same radial grid makes the transformation matrix $\boldsymbol{U}^{(l,\text{grid})}$ block-diagonal with n_r subblocks corresponding to the angular transformation at each grid point.

For $m = 0$ we already showed in Sec. 3.4 that $\tilde{\Phi}_m(r_i, \zeta, t)$ has a grid representation in the angular variable. The $l_{\max} + 1$ grid points in ζ must be chosen as the roots of the $\bar{P}_{l_{\max}+1}(\zeta)$, and inner products in the grid representation must use the corresponding Gauss-Legendre quadrature rule. For each radial grid point, the unitary transformation between angular grid and Legendre representation is

$$\tilde{\Phi}_0^{(\text{grid})}(r_i, \zeta_j, t) = \sum_{l=0}^{l_{\max}} f_{l0}(r_i, t)\bar{P}_l(\zeta_j), \tag{5.30}$$

$$f_{l0}(r_i, t) = \sum_{j=0}^{l_{\max}} w_j \bar{P}_l(\zeta_j)\tilde{\Phi}_0^{(\text{grid})}(r_i, \zeta_j, t), \tag{5.31}$$

or with the matrices defined in Eqs. (3.30) and (3.31)

$$\tilde{\boldsymbol{\Phi}}_0^{(\text{grid})}(r_i, t) = \bar{\boldsymbol{P}} \cdot \boldsymbol{f}(r_i, t), \tag{5.32}$$

$$\boldsymbol{f}(r_i, t) = \text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \tilde{\boldsymbol{\Phi}}_0^{(\text{grid})}(r_i, t). \tag{5.33}$$

Note that the transformation is the same for all radial points. Below, we only consider the "i"th radial point. The procedure must be repeated for all radial grid points. With the grid representation that diagonalizes the potential, it is easy to apply the potential operator

$$e^{-i\Delta t V(r,\zeta,t)}\tilde{\Phi}_0(r_i, \zeta, t) \doteq \text{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \text{diag}(e^{-i\Delta t \boldsymbol{V}(r_i,t)}) \cdot \bar{\boldsymbol{P}} \cdot \boldsymbol{f}(r_i, t), \tag{5.34}$$

with

$$\boldsymbol{V}(r_i, t) = \{V(r_i, \zeta_0, t), V(r_i, \zeta_1, t), \cdots, V(r_i, \zeta_{l_{\max}}, t)\}. \tag{5.35}$$

It is possible to obtain a grid representation even if m is nonzero. The transformation will be unitary, if we choose grid and weights according to a certain Gauss-Jacobi quadrature rule [6][1].

For the magnetic quantum number m, the Legendre-grid transformations with matrix multiplications require $\mathcal{O}\left(n_l^2\right)$ computational operations, since the dimension of the matrices will be $n_l \times n_l$. As the transformation must be

---

[1]Note that Appendix A of Ref. [6] contains some misprints. The upper limits of the sums in Eqs. (A1a) and (A4a) should read L and $L - m$, respectively.

repeated for all radial grid points, the total number of operations scales as $\mathcal{O}\left(n_r n_l^2\right)$ when the potential is applied in angular grid representation. For most problems encountered in practice, the angular grid transformation will be much slower than applying the kinetic operator by the Fourier method. This is essentially due to the fact that the angular grid transformation can only be accomplished by matrix multiplications which scale quadratically with the number of angular basis functions. Since the angular grid transformation is the bottleneck for the overall propagation scheme, it can be useful to explore alternative methods to apply the potential operator.

The potential operator in spherical harmonics representation

A closer examination of Eq. (5.34) suggests an alternative way to apply the potential operator. In Eq. (5.34) we interpreted matrix-vector product $\bar{\boldsymbol{P}} \cdot \boldsymbol{f}(r_i, t)$ as the grid representation that brings the potential operator on a diagonal form. We can, however, equally well interpret the matrix product of the four matrices on the left as the operator $e^{-i\Delta t V(r_i, \zeta, t)}$ being expressed as a matrix representation in the basis of spherical harmonics, i.e.

$$e^{-i\Delta t V(r_i, \zeta, t)} \doteq \operatorname{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T \cdot \operatorname{diag}(e^{-i\Delta t \boldsymbol{V}(r_i, t)}) \cdot \bar{\boldsymbol{P}}. \qquad (5.36)$$

This interpretation must be valid since this matrix product states how the potential updates the coefficients $f_{lm}(r_i, t)$ in the spherical harmonics expansion to a new set of expansion coefficients which are also expressed in the spherical harmonics basis. The basic idea that will be described below is to find the spherical harmonics representation of $e^{-i\Delta t V(r_i, \zeta, t)}$ by a more direct method, namely as a matrix exponential of $\boldsymbol{V}(r_i, \zeta, t)$.

For each radial grid point r_i, we write the potential as an $n_l \times n_l$ matrix in the spherical harmonics basis with the matrix elements

$$V_{l',l}(r_i, t) = \int_{-1}^{1} \bar{P}_{l'}^{|m|}(\zeta) V(r_i, \zeta, t) \bar{P}_l^{|m|}(\zeta) d\zeta. \qquad (5.37)$$

A schematic picture of the potential matrix is given in Fig. 5.3. Since we need to apply the operator $e^{-i\Delta t V(r, \zeta, t)}$, we must now exponentiate the matrix representation of the potential. We may diagonalize the Hermitian matrix $\boldsymbol{V}(r_i, t)$ by a unitary matrix $\boldsymbol{U}$

$$\boldsymbol{V}(r_i, t) = \boldsymbol{U}_V(r_i, t) \cdot \operatorname{diag}(\boldsymbol{d}_V(r_i, t)) \cdot \boldsymbol{U}_V^\dagger(r_i, t), \qquad (5.38)$$

where $\boldsymbol{d}_V(r_i, t)$ contains the real eigenvalues of $\boldsymbol{V}(r_i, t)$. With the diagonal form of $V(r_i, \zeta, t)$ it is straightforward to apply the potential operator as

$$e^{-i\Delta t V(r, \zeta, t)}\tilde{\Phi}_m(r_i, \zeta, t) \doteq \boldsymbol{U}_V(r_i, t) \cdot \operatorname{diag}(e^{-i\Delta t \boldsymbol{d}_V(r_i, t)}) \cdot \boldsymbol{U}_V^\dagger(r_i, t) \cdot \boldsymbol{f}(r_i, t).$$
$$(5.39)$$

$$V(r,\zeta,t) \doteq \qquad\qquad\qquad\qquad \tilde{\mathbf{\Phi}}_0^{(l)}(t) =$$
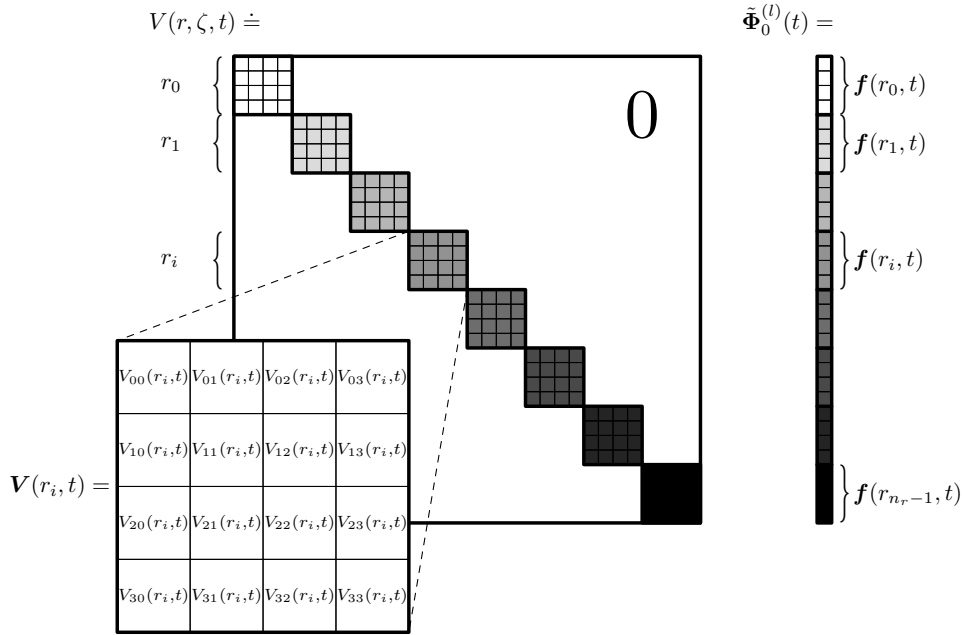

Figure 5.3: Matrix representation of the potential for the magnetic quantum number $m = 0$ and with $l_{\max} = 3$. The block diagonal structure is obtained for l-major ordering.

This general method is exactly unitary for any bandwidth limited function. The drawback is that it scales very inefficiently with the number of angular basis functions. First, we must construct the matrix elements of the potential, Eq. (5.37), in each time step. Second, the matrix must be diagonalized and at last the three matrix multiplications in Eq. (5.39) must be performed. The computational complexity involved in the diagonalization of an $n_l \times n_l$ matrix scales as $\mathcal{O}(n_l^3)$. The diagonalization must be repeated for all radial grid points which gives a total complexity of $\mathcal{O}(n_r n_l^3)$. Due to this rapid scaling, Eq. (5.39) is only generally useful for very small angular bases.

Although the potential applied by a transformation to the angular grid seems to be less complex than in the spherical harmonics representation, there are certain special cases where the latter method can be faster. In order to demonstrate such cases, we consider two examples, namely the molecular potential of the hydrogen molecular ion, $H_2{}^+$, and the scalar potential of a linearly polarized field.

The static molecular potential

In the first example we consider the hydrogen molecular ion with the two nuclei fixed in space at $\pm \boldsymbol{R}/2$. Since the molecular potential is independent of time, we need only to evaluate its exponentiated matrix representation

once and for all and subsequently store it in memory. The propagation with this term is thus reduced to a single matrix multiplication in each time step. Furthermore, the symmetry of the potential imposes some selection rules on the matrix elements, so that a full matrix multiplication can be avoided.

We will orient the molecular axis along the z-axis and express $V^{(M)}(r, \zeta)$ as

$$V^{(M)}(r,\zeta) = -\frac{1}{|\boldsymbol{r} - \frac{\boldsymbol{R}}{2}|} - \frac{1}{|\boldsymbol{r} + \frac{\boldsymbol{R}}{2}|} = -\sum_{L=0}^{\infty} \frac{r_<^L}{r_>^{L+1}} P_L(\zeta) - \sum_{L=0}^{\infty} \frac{r_<^L}{r_>^{L+1}} P_L(-\zeta)$$

$$= -2 \sum_{\substack{L=0, \\ L \, \text{Even}}}^{\infty} \sqrt{\frac{2}{2L+1}} \frac{r_<^L}{r_>^{L+1}} \bar{P}_L(\zeta), \tag{5.40}$$

where $r_<$ ($r_>$) is the smaller (larger) of $(r, R/2)$. In the last equality we have used the fact that the L'th Legendre polynomial has the parity $(-1)^L$. The matrix elements can now be written as

$$V_{l',l}^{(M)}(r_i) = -2 \sum_{L \, \text{even}} \sqrt{\frac{2}{2L+1}} \frac{r_<^L}{r_>^{L+1}} \int_{-1}^{1} \bar{P}_{l'}^{|m|}(\zeta) \bar{P}_L(\zeta) \bar{P}_l^{|m|}(\zeta) d\zeta. \tag{5.41}$$

The integral over the three spherical harmonics is known as a Gaunt coefficient and can be expressed analytically in terms of Clebsch-Gordan coefficients. Since L is even, $\bar{P}_{l'}^{|m|}$ and $\bar{P}_l^{|m|}$ must have the same parity, i.e., $V_{l',l}^{(M)}(r_i) = 0$ unless $(-1)^{l'} = (-1)^l$. More generally, this selection rule applies to all inversion symmetric potentials. Due to the selection rule we may arrange the basis functions by even and odd parity so that the matrix representation of the molecular potential becomes block-diagonal as shown in Fig. 5.4. When this matrix is exponentiated we can just consider the two blocks separately so that the exponentiated matrix preserves the same block diagonal structure. In each time step the constant exponentiated matrix will be multiplied by the $\boldsymbol{f}(r_i, t)$ vector. Instead of performing a full matrix multiplication it is only necessary to perform a number of block matrix multiplications within each of the blocks.

Except for the initial diagonalization, the complexity of this method scales as $\mathcal{O}(n_r n_l^2)$, just as the grid representation method. The present method is, however, faster by a constant factor, since it takes the parity symmetry into account. The drawback is that the large matrix $e^{-i\Delta t \boldsymbol{V}^{(M)}}$ must be stored in memory.

The time dependent field interaction

In this example, we consider a linearly polarized field with the polarization vector parallel to the z axis. The field interaction is time dependent and a number of matrix operations, equivalent to Eqs. (5.37)-(5.39), need to be performed in every time step.
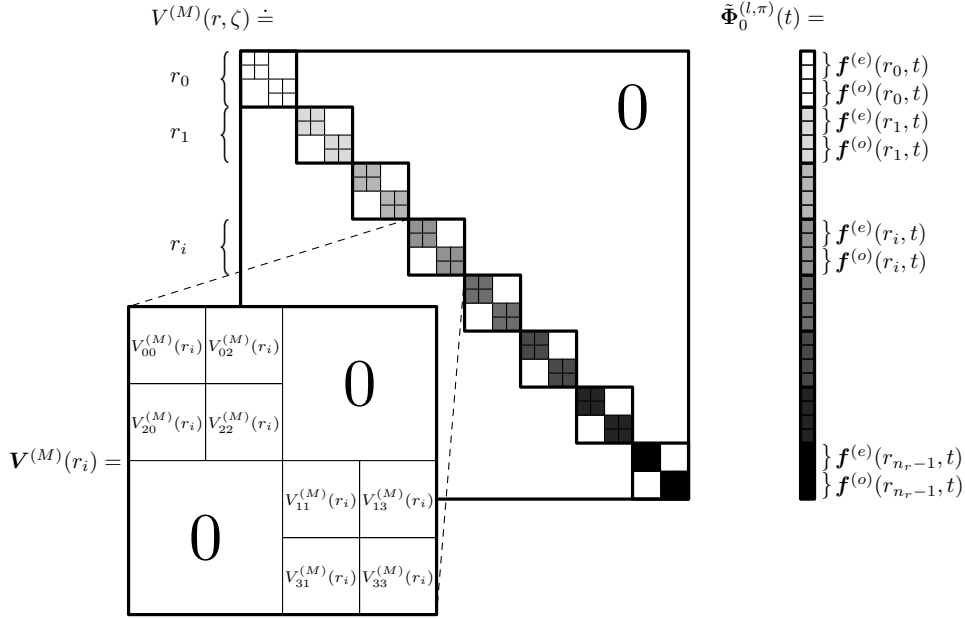
Figure 5.4: Matrix representation of the $H_2{}^+$ potential for $l_{\max} = 3$ and the magnetic quantum number $m = 0$. The ordering of basis functions is a variant of the l-major ordering with the additional ordering of even and odd l's.

The field interaction will now be written as

$$V^{(F)}(r, \zeta, t) = E(t)r\zeta = E(t)r\sqrt{\frac{2}{3}}\bar{P}_1(\zeta),\tag{5.42}$$

and the corresponding matrix elements are

$$V_{l',l}^{(F)}(r_i, t) = E(t)\sqrt{\frac{2}{3}}r_i \int_{-1}^{1} \bar{P}_{l'}^{|m|}(\zeta)\bar{P}_1(\zeta)\bar{P}_l^{|m|}(\zeta)d\zeta.\tag{5.43}$$

l and l' can only differ by one which brings the matrix representation on a tridiagonal form as shown in Fig. 5.5. Since $\boldsymbol{V}^{(F)}(r_i, t)$ is time dependent, we must in principle recalculate the matrix exponential at every time step which seems to introduce an $\mathcal{O}(n_r n_l^3)$ operation. We can avoid the construction of the matrix exponentiation from scratch by noting that the time dependent part of $\boldsymbol{V}^{(F)}(r_i, t)$ is simply a multiplicative factor

$$\boldsymbol{V}^{(F)}(r_i, t) = E(t)\sqrt{\frac{2}{3}}r_i\bar{\boldsymbol{P}}_1.\tag{5.44}$$
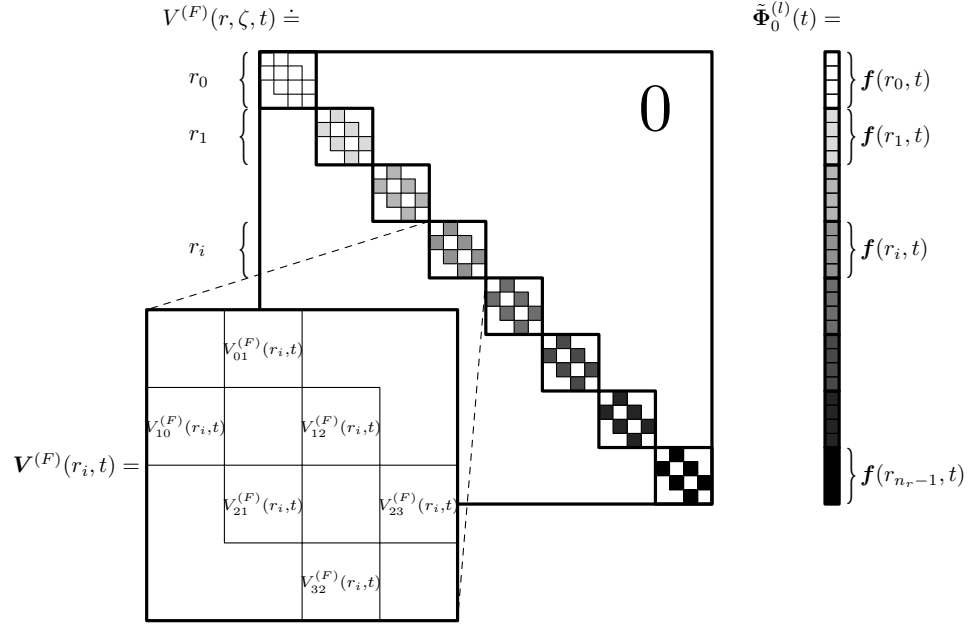
Figure 5.5: Matrix representation of $\boldsymbol{V}^{(F)}(r_i, t)$. The vector representation of the wave function must be stored in l-major ordering.

The matrix elements of $\bar{\boldsymbol{P}}_1$ are known analytically

$$
\begin{aligned}
[\bar{P}_1]_{l',l} &= \int_{-1}^{1} \bar{P}_{l'}^{|m|}(\zeta) \bar{P}_1(\zeta) \bar{P}_l^{|m|}(\zeta) d\zeta \\
&= \sqrt{\frac{(l+1)^2 - m^2}{(2l+3)(2l+1)}} \delta_{l,l'-1} + \sqrt{\frac{l^2 - m^2}{(2l+1)(2l-1)}} \delta_{l,l'+1}, \quad (5.45)
\end{aligned}
$$

and depend on neither t nor r_i. We may now simplify the matrix exponentiation, according to

$$
e^{-i\Delta t \boldsymbol{V}^{(F)}(r_i, t)} = \boldsymbol{U}^{(\bar{P}_1)} \cdot \operatorname{diag}\left(e^{-i\Delta t E(t)\sqrt{\frac{2}{3}} r_i \boldsymbol{d}^{(\bar{P}_1)}}\right) \cdot [\boldsymbol{U}^{(\bar{P}_1)}]^{\dagger}. \qquad (5.46)
$$

$\boldsymbol{U}^{(\bar{P}_1)}$ is the *constant* unitary matrix that diagonalizes $\bar{\boldsymbol{P}}_1$ to the diagonal form $\boldsymbol{d}^{(\bar{P}_1)}$. The strategy is now to calculate and store $\boldsymbol{U}^{(\bar{P}_1)}$ and $\bar{\boldsymbol{P}}_1$ once and for all. The construction of $e^{-i\Delta t \boldsymbol{V}^{(F)}(r_i, t)}$ is then reduced to an exponentiation of the diagonal elements, $e^{-i\Delta t E(t)\sqrt{\frac{2}{3}} r_i \bar{d}_j^{(\bar{P}_1)}}$, followed by two matrix multiplications. This method is in fact somewhat similar to the grid representation. In both cases we use a constant unitary transformation to bring the potential on a diagonal form, and the complexity scaling is $\mathcal{O}(n_r n_l^2)$ for both methods. The present matrix representation method has, however, one major advantage over the grid method. First we note that the potential is

represented by the Hermitian matrix $\boldsymbol{V}^{(F)}(r_i, t)$. Second, inner products in the spherical harmonics representation are equivalent to ordinary dot products. Hence, the Crank-Nicolson approximation for the exponential

$$e^{-i\Delta t \boldsymbol{V}^{(F)}(r_i,t)} = \left(1 + i\frac{\Delta t}{2}\boldsymbol{V}^{(F)}(r_i, t)\right)^{-1} \cdot \left(1 - i\frac{\Delta t}{2}\boldsymbol{V}^{(F)}(r_i, t)\right) + \mathcal{O}(\Delta t^3)$$
(5.47)

will be unitary. The Crank-Nicolson form is very simple to apply on a vector $\boldsymbol{f}(r_i, t)$, since $\boldsymbol{V}^{(F)}(r_i, t)$ is tridiagonal and thus both the inversion and multiplication scale just *linearly* with the angular dimension n_l. The total complexity scaling by the Crank-Nicolson method is just $\mathcal{O}(n_r n_l)$.

5.3.2 Velocity gauge split operator for cylindrically symmetric Hamiltonians

The cylindrically symmetric velocity gauge Hamiltonian reads [cf. Eq. (5.21)]

$$\mathcal{H}^{(\mathrm{VG})} = -\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} + iA_{\mathrm{VG}}(t)\left[\frac{1}{r}\left(\zeta - (1-\zeta^2)\frac{\partial}{\partial \zeta}\right) - \zeta\frac{\partial}{\partial r}\right] + V^{(M)}(r, \zeta).$$
(5.48)

We introduce the following short hand notation for the operators

$$T_r = -\frac{1}{2}\frac{\partial^2}{\partial r^2}, \tag{5.49}$$

$$T_l = \frac{L^2}{2r^2}, \tag{5.50}$$

$$V^{(F,1)} = -iA_{\mathrm{VG}}(t)\zeta\frac{\partial}{\partial r}, \tag{5.51}$$

$$V^{(F,2)} = iA_{\mathrm{VG}}(t)\frac{1}{r}\left(\zeta - (1-\zeta^2)\frac{\partial}{\partial \zeta}\right), \tag{5.52}$$

$$V^{(M)} = V^{(M)}(r, \zeta). \tag{5.53}$$

All of these operators are easily proven to be Hermitian. Note, however, that $V^{(F,2)}$ cannot be decomposed into two separate Hermitian operators. Only the operators T_r and $V^{(F,1)}$ commute so an attempt to split any of the other operators introduces commutation errors.

In order to apply the time evolution operator, we try the following split operator

$$\begin{aligned}
\mathscr{U}(t + \Delta t, t) &\approx e^{-i\frac{\Delta t}{2}T_r} \times e^{-i\frac{\Delta t}{2}T_l} \times e^{-i\frac{\Delta t}{2}V^{(M)}} \times e^{-i\frac{\Delta t}{2}V^{(F,2)}} \times e^{-i\Delta t V^{(F,1)}} \\
&\times e^{-i\frac{\Delta t}{2}V^{(F,2)}} \times e^{-i\frac{\Delta t}{2}V^{(M)}} \times e^{-i\frac{\Delta t}{2}T_l} \times e^{-i\frac{\Delta t}{2}T_r}.
\end{aligned} \tag{5.54}$$

The first three and the last three factors are identical to the length gauge and will be applied likewise.

The interaction $V^{(F,2)}$ can be treated much like the interaction operator in the length gauge. For each radial distance r_i, we analytically calculate the matrix elements of the interaction operator in the associated Legendre basis

$$
\begin{aligned}
[V^{(F,2)}(r_i,t)]_{l,l'} =& i\frac{A_{\mathrm{VG}}(t)}{r_i} \int_{-1}^{1} \bar{P}_l^{|m|}(\zeta) \left(\zeta - (1-\zeta^2)\frac{\partial}{\partial\zeta} \right) \bar{P}_{l'}^{|m|}(\zeta)d\zeta \\
=& i\frac{A_{\mathrm{VG}}(t)}{r_i} \left(-(l+1)\sqrt{\frac{(l+1)^2-m^2}{(2l+3)(2l+1)}}\delta_{l,l'-1} \right. \\
& \left. + l\sqrt{\frac{l^2-m^2}{(2l+1)(2l-1)}}\delta_{l,l'+1} \right).
\end{aligned}
\tag{5.55}
$$

In l-major ordering, this matrix has the same block-tridiagonal structure as shown in Fig. 5.5. The exponential can be applied either as in Eq. (5.46) or as the Crank-Nicolson approximation (5.47).

Finally we consider the interaction term $V^{(F,1)}$. This term is the most difficult interaction term to handle with the radial grid/spherical harmonics basis. The problem is that the operator is non-diagonal in both the radial and angular bases. Accordingly, neither r-major nor l-major ordering will bring the interaction operator on a block diagonal form. Any matrix-vector operation then requires an operation with a full matrix of dimension $n_r n_l$ which typically leads to a very inefficient scaling, for example the matrix-vector multiplication scales as $\mathcal{O}(n_r^2 n_l^2)$. As we have seen several times above, it is very helpful to bring either the radial- or the angular operator on a diagonal form. One idea would be to diagonalize the radial derivative by a Fourier transformation in the radial direction. But since the radial derivative is only of first order, it is impossible to use the Fourier method since box boundary conditions can only be imposed for even order derivatives as discussed in Sec. 4.3.2. A better approach is to bring the angular factor $\zeta = \cos\theta$ on diagonal form by the angular grid-Legendre representation. First, we assume that the transformation into the angular grid has been obtained. For each of the n_l angular grid points ζ_j we can now construct the following n_r-dimensional vectors

$$
\tilde{\mathbf{\Phi}}_{mj}^{(\mathrm{grid})}(t) = \{\tilde{\Phi}_m^{(\mathrm{grid})}(r_0,\zeta_j,t), \tilde{\Phi}_m^{(\mathrm{grid})}(r_1,\zeta_j,t), \cdots, \tilde{\Phi}_m^{(\mathrm{grid})}(r_{n_r-1},\zeta_j,t)\}^T.
\tag{5.56}
$$

The vectors are defined along the radial rays that emerge from the origin as shown in Fig. 5.6. We can now use a one-dimensional finite difference scheme for the interaction

$$
V^{(F,1)} \doteq -iA_{\mathrm{VG}}(t)\zeta_j \mathbf{D}.
\tag{5.57}
$$

The exponential of this tridiagonal matrix will be approximated by the Crank-Nicolson scheme which allows for box boundary conditions. The vec-
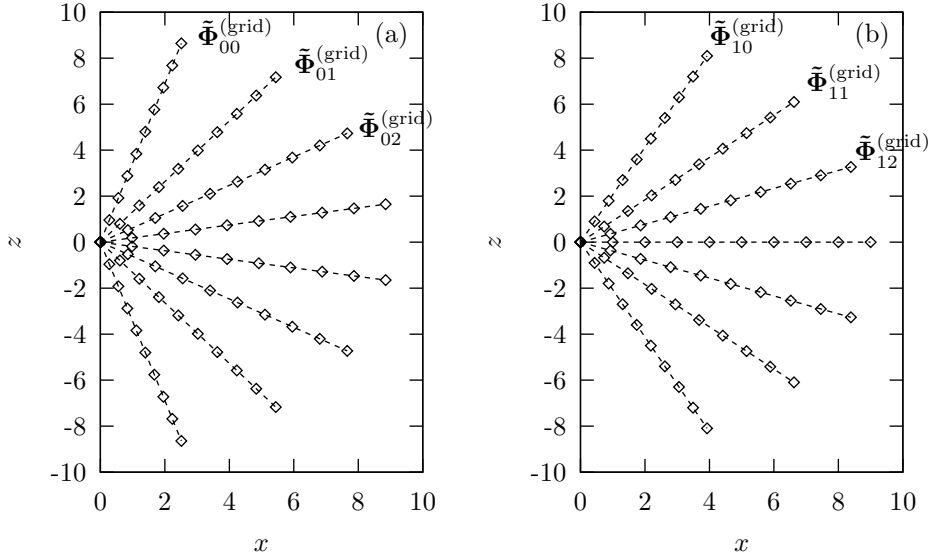
Figure 5.6: Radial/angular grid representation for $n_r = 10$ and $l_{\max} = 7$. (a) For $m = 0$ the function is defined on an eight point angular grid. (b) For $m = 1$ (and $m = -1$) the function is defined on a seven point angular grid.

tors in Eq. (5.56) are thereby updated according to

$$\tilde{\boldsymbol{\Phi}}_{mj}^{(\mathrm{grid})}(t) \rightarrow \left(\mathbf{1} + \frac{\Delta t}{2} A_{\mathrm{VG}}(t)\zeta_j \boldsymbol{D}\right)^{-1} \cdot \left(\mathbf{1} - \frac{\Delta t}{2} A_{\mathrm{VG}}(t)\zeta_j \boldsymbol{D}\right) \cdot \tilde{\boldsymbol{\Phi}}_{mj}^{(\mathrm{grid})}(t).$$
$$(5.58)$$

The updated vectors in the angular grid representation are then transformed back to spherical harmonics representation by expressions similar to Eq. (5.33), generalized to $m \neq 0$.

The transformation to angular grid representation scales as $\mathcal{O}(n_r n_l^2)$ while the Crank-Nicolson step scales as $\mathcal{O}(n_r n_l)$ so that the total scaling is much better than expressing $V^{(F,1)}$ directly in spherical harmonics representation. Furthermore, the latter representation does not allow for an efficient parallelization scheme as we shall see in Sec. 6.1.2.

5.4   Non-cylindrically symmetric problems

As we have seen in the previous section, cylindrically symmetric problems can be handled efficiently in a spherical harmonics basis. The most general three-dimensional problems are, however, not symmetric under rotation around any axis. Commonly encountered examples include linear molecules, where the molecular axis does not coincide with the polarization vector of a linearly polarized field, as well as atoms and molecules interacting with an elliptically polarized electric field.
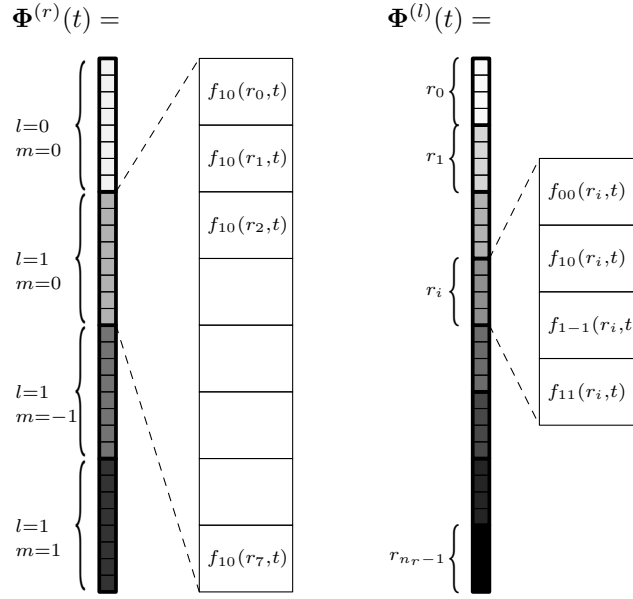
$$\mathbf{\Phi}^{(r)}(t) = \qquad\qquad\qquad \mathbf{\Phi}^{(l)}(t) =$$

Figure 5.7: Vector representation of $\Phi(\boldsymbol{r},t)$ in r-major and l-major ordering for $n_r = 8$ and $l_{\max} = 1$.

If the problem is not cylindrically symmetric, the Schrödinger equation does not allow for a separation of the ϕ dependence as in Eq. (5.15), and the reduced wave function must include all values of m, cf. Eq. (5.11). It is straightforward to extend the vector representations (5.17) and (5.18) to include all m values. The extended vector then reads in r-major ordering

$$\mathbf{\Phi}^{(r)}(t) = \{\underbrace{[-\tilde{\mathbf{\Phi}}_0^{(r)}(t)-]}_{n_r(l_{\max}+1)}, \underbrace{[-\tilde{\mathbf{\Phi}}_{-1}^{(r)}(t)-]}_{n_r l_{\max}}, \underbrace{[-\tilde{\mathbf{\Phi}}_1^{(r)}(t)-]}_{n_r l_{\max}}, \cdots, \underbrace{[-\tilde{\mathbf{\Phi}}_{l_{\max}}^{(r)}(t)-]}_{n_r}\}^T, \tag{5.59}$$

and in l-major ordering

$$\begin{aligned}
\mathbf{\Phi}^{(l)}(t) = \{&f_{00}(r_0,t), f_{10}(r_0,t), \cdots, f_{l_{\max},0}(r_0,t), f_{1-1}(r_0,t),\\
&\cdots, f_{l_{\max},-1}(r_0,t), f_{11}(r_0,t), \cdots, f_{l_{\max},l_{\max}}(r_0,t),\\
&f_{00}(r_1,t), f_{10}(r_1,t), \cdots, f_{l_{\max},l_{\max}}(r_{n_r-1},t)\}^T. \tag{5.60}
\end{aligned}$$

The vector representations are shown schematically in Fig. 5.7.

5.4.1 Split-operator for non-cylindrically symmetric potentials

As in the previous section, we will use the split-operator technique to solve the time evolution of the reduced wave function. Both the length and velocity gauge Hamiltonians contain the radial kinetic operator and the centrifugal operator. These operators enter the split-operator as $e^{-i(\Delta t/2)(-(1/2)\partial^2/\partial r^2)}$ and $e^{-i(\Delta t/2)L^2/(2r^2)}$, respectively.

The radial kinetic operator can be applied in the same way as described in Sec. 5.3.1. Again, no couplings occur between the different l, m states, and the radial kinetic operator acts independently on the one-dimensional vectors $\boldsymbol{f}_{lm}(t)$ given in Eq. (5.25). The only difference from the cylindrically symmetric case is that the number of angular states is now [cf. Eq. (5.12)] $n_{Y_{lm}} = (l_{\max} + 1)^2$ instead of $n_l = l_{\max} + 1 - |m|$. With the Fourier method the radial kinetic operator can be applied in $\mathcal{O}(n_r \log_2 n_r l_{\max}^2)$ operations.

The centrifugal operator is still diagonal in the spherical harmonics representation, and the term $e^{-i(\Delta t/2)L^2/(2r^2)}$ can be applied exactly as in Eq. (5.29).

In the length gauge, the only remaining factor in the split-operator is the potential $e^{-i\Delta t V(\boldsymbol{r},t)}$. This term introduces major difficulties when the potential is not cylindrically symmetric, but rather a function of the full three-dimensional coordinate operator. Hence, it would be represented by a diagonal matrix if the wave function were represented in a three-dimensional grid representation. As for the cylindrically symmetric case, Eq. (5.32), we try to obtain a grid representation from the l-major representation

$$\boldsymbol{\Phi}^{(\text{grid})}(t) = \boldsymbol{U}^{(l,\text{grid})} \cdot \boldsymbol{\Phi}^{(l)}(t), \qquad (5.61)$$

so that each element in the grid representation corresponds to a three dimensional grid point $\boldsymbol{r}_j = (r_j, \theta_j, \phi_j)$. If it is possible to find such a grid representation, it is trivial to apply the potential operator on each element of $\boldsymbol{\Phi}^{(\text{grid})}(t)$ as

$$e^{-i\Delta t V(\boldsymbol{r},t)}[\Phi^{(\text{grid})}(t)]_j = e^{-i\Delta t V(r_j,\theta_j,\phi_j,t)}[\Phi^{(\text{grid})}(t)]_j. \qquad (5.62)$$

As for any change of representation, we will require that the transformation corresponding to Eq. (5.61) must be unitary for some inner product in the grid representation. Since the radial coordinate is already represented on a grid, we will of course use the same radial grid in both representations. Thus, the transformation is only nontrivial from the spherical harmonics representation to the (θ, ϕ) grid representation. The fact that we use the same radial grid makes the transformation matrix $\boldsymbol{U}^{(l,\text{grid})}$ block-diagonal with n_r subblocks corresponding to the angular transformation at each grid point. Unfortunately, in the non-cylindrically symmetric case there exists no unitary transformation from a truncated spherical harmonics representation to an angular (θ, ϕ) grid representation. A number of so-called "spherical harmonics transformations" have been described in the literature [11–13]. For all these methods, there are more grid points than spherical harmonics, typically $2 - 4$ times as many. Transformations between bases with unequal number of basis functions cannot be unitary. This is easily seen since it would be possible to find two different states in the larger basis that are mapped into the same state in the smaller basis. It is then clearly impossible to make a unique transformation from the smaller basis to the larger. The
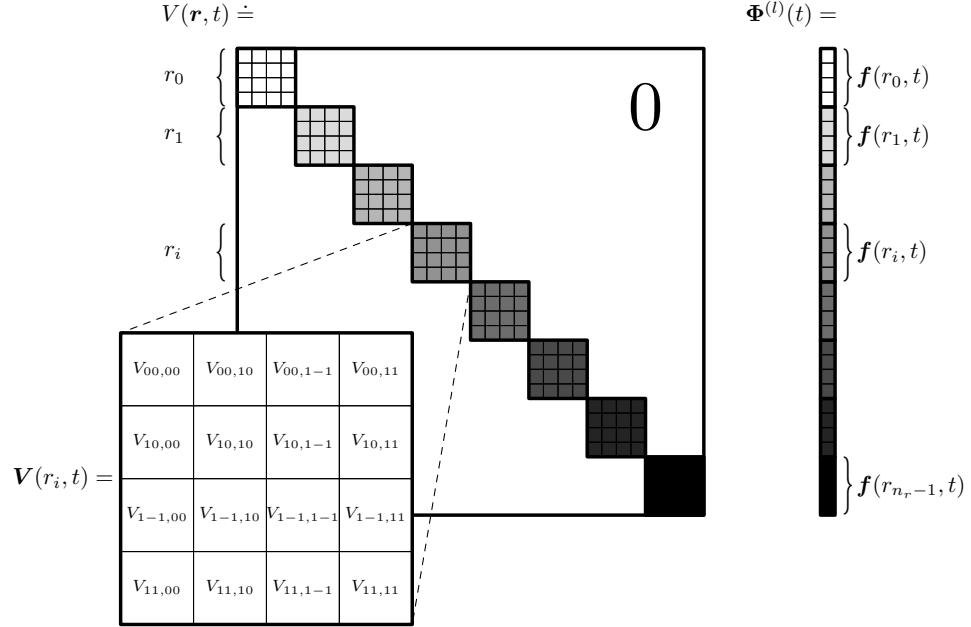
$V(\boldsymbol{r},t) \doteq$

r_0

r_1

r_i

$\boldsymbol{0}$

$\boldsymbol{\Phi}^{(l)}(t) =$

$\boldsymbol{f}(r_0,t)$

$\boldsymbol{f}(r_1,t)$

$\boldsymbol{f}(r_i,t)$

$\boldsymbol{f}(r_{n_r-1},t)$

$\boldsymbol{V}(r_i,t) =$

$V_{00,00}$	$V_{00,10}$	$V_{00,1-1}$	$V_{00,11}$
$V_{10,00}$	$V_{10,10}$	$V_{10,1-1}$	$V_{10,11}$
$V_{1-1,00}$	$V_{1-1,10}$	$V_{1-1,1-1}$	$V_{1-1,11}$
$V_{11,00}$	$V_{11,10}$	$V_{11,1-1}$	$V_{11,11}$

Figure 5.8: Matrix representation of the potential. The block diagonal structure is obtained for l-major ordering.

corresponding problem can be seen from the matrix expression in Eq. (5.61). The spherical harmonics transformations correspond to transformation matrices $\boldsymbol{U}^{(l,\text{grid})}$ with more rows than columns. The inverse matrix is then underdetermined. The spherical harmonics transformations try to define the inverse in some least square sense. However, the transformation can never be exactly unitary.

Since the potential operator cannot be diagonalized by a grid representation, the best way to apply the potential seems to be directly as a matrix representation in the spherical harmonics basis. This is just the generalization of the method described in Eqs. (5.37)-(5.39). For each radial grid point r_i, we now write the potential as an $n_{Y_{lm}} \times n_{Y_{lm}}$ matrix in the spherical harmonics basis with the matrix elements

$$V_{l'm',lm}(r_i,t) = \langle l'm'|V(r_i,\theta,\phi,t)|lm\rangle. \tag{5.63}$$

The brackets denote an integration over the angles (θ,ϕ) and $|lm\rangle \equiv Y_{lm}(\theta,\phi)$. A schematic picture of the potential matrix is given in Fig. 5.8. The exponentiation proceeds exactly as in Eqs. (5.37)-(5.39) by an intermediate diagonalization. The computational complexity involved in the diagonalization of an $n_{Y_{lm}} \times n_{Y_{lm}}$ matrix scales as in $\mathcal{O}(n_{Y_{lm}}^3)$ which results in a total complexity of $\mathcal{O}(n_r l_{\max}^6)$. If the potential is static, the diagonalization needs to be performed only once and the potential can be applied as a matrix multiplication. The computational complexity can thereby be reduced to $\mathcal{O}(n_r l_{\max}^4)$.

The inefficient scaling is probably the reason why non-cylindrically symmetric potentials have not been treated with a spherical harmonics expansion until very recently.

As we shall show in Sec. 5.5 there are certain types of non-cylindrically symmetric potentials that can be treated much more efficiently than the $\mathcal{O}(n_r l_{\max}^6)$ method. This will be the case, if the total potential can be written as a sum of potentials with each term being cylindrically symmetric around *some* (but not necessarily the same) axis.

5.4.2   Superposition of cylindrically symmetric problems

If an initial reduced wave function is an arbitrary three-dimensional function $\Phi(r, \zeta, \phi, t = 0)$ while the Hamiltonian is cylindrically symmetric, we can readily extend the method developed in Sec. 5.3. First, the initial wave function is decomposed in its m-components

$$\Phi(r, \zeta, \phi, t = 0) = \sum_{m=-l_{\max}}^{l_{\max}} \tilde{\Phi}_m(r, \zeta, t = 0) \frac{e^{im\phi}}{\sqrt{2\pi}}, \qquad (5.64)$$

with

$$\tilde{\Phi}_m(r, \zeta, t = 0) = \int_0^{2\pi} \Phi(r, \zeta, \phi, t = 0) \frac{e^{-im\phi}}{\sqrt{2\pi}} d\phi. \qquad (5.65)$$

Even though more than a single m-state must be used to describe the wave function, the Hamiltonian does not couple the different m subspaces. For example, the matrices of the potentials $V^{(M)}$ and $V^{(F)}$ must now take all m-values into account but the decoupling in m means that both matrices become block diagonal by m as shown in Fig. 5.9.

It is now possible to decompose the time evolution of the full wave function in a number of smaller cylindrically symmetric problems which evolve completely independently. The time evolution of each individual m component can be solved as $\tilde{\Phi}_m(r, \zeta, t) = \mathscr{U}(t, 0)\tilde{\Phi}_m(r, \zeta, 0)$ by, e.g., the split-operator method described in Sec. 5.3. The full solution at time t is then constructed as

$$
\begin{aligned}
\mathbf{\Phi}^{(r)}(0) &= \{[-\tilde{\mathbf{\Phi}}_0^{(r)}(0)-] \quad [-\tilde{\mathbf{\Phi}}_{-1}^{(r)}(0)-] \quad [-\tilde{\mathbf{\Phi}}_1^{(r)}(0)-] \quad \cdots\}^T \\
&\qquad\qquad \downarrow \qquad\qquad\quad \downarrow \qquad\qquad\quad \downarrow \\
\mathbf{\Phi}^{(r)}(t) &= \{[-\tilde{\mathbf{\Phi}}_0^{(r)}(t)-] \quad [-\tilde{\mathbf{\Phi}}_{-1}^{(r)}(t)-] \quad [-\tilde{\mathbf{\Phi}}_1^{(r)}(t)-] \quad \cdots\}^T
\end{aligned}
\qquad (5.66)
$$

As we have seen in Sec. 5.3, the typical computational complexity scales as $\mathcal{O}(n_r(l_{\max}+1-|m|)^2)$ for each of the cylindrically symmetric problems. The total number of computational operations for all the $2l_{\max} + 1$ values of m must be less than $\mathcal{O}\left(n_r l_{\max}^3\right)$, since this would be the operational count if the dimensionalities of all m subspaces are equal to the maximum dimension $l_{\max} + 1$ (when $m = 0$). The reduced dimensionality for higher values of $|m|$ brings the operational count down to $\sim \mathcal{O}(n_r l_{\max}^{2.7})$.
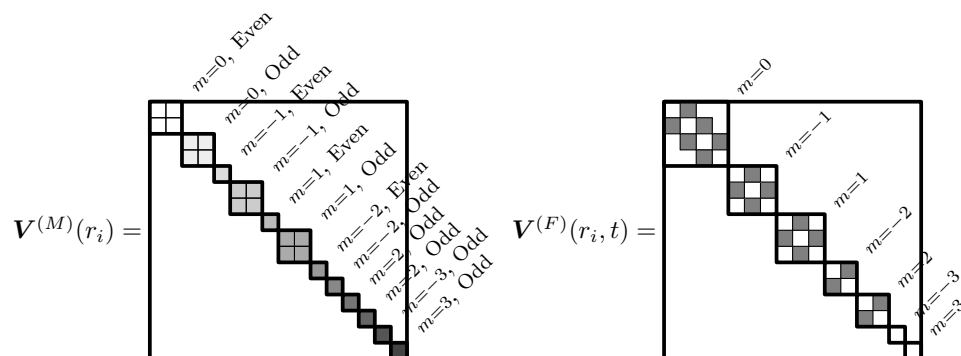
Figure 5.9: The blocks corresponding to the radial grid point r_i in Figs. 5.4 and 5.5 are extended to cover all values of m for $l_{\max} = 3$. The blocks $-m$ and $+m$ are identical for both $\boldsymbol{V}^{(M)}$ and $\boldsymbol{V}^{(F)}$. The wave function must be stored in l-major ordering (5.60) when $\boldsymbol{V}^{(F)}$ is applied, while an additional ordering by parity is required for $\boldsymbol{V}^{(M)}$.

It is perhaps not so obvious to think of a real physical system where the present method applies. One example is studied in Ref. [14] where an electron in a hydrogen atom is prepared in a $2p_{z'}$ state, i.e., a state with zero projection of the angular momentum on a z'-axis. The atom is subsequently exposed to a linearly polarized field along a z-axis that does not coincide with the z' axis. In this case, the Hamiltonian is cylindrically symmetric around the z axis but the wave function is a superposition of $m = 0, \pm 1$ states.

Although a complete decomposition in separate m subsystems is not encountered frequently, the present method forms the building blocks for a much more general situation, namely if the Hamiltonian can be decomposed in terms that are cylindrically symmetric around two or more different axes. This approach will be discussed in detail in Sec. 5.5.

5.5 Solving the m-mixing problem by rotations

The previous sections have made it clear that the presence of a non-cylindrically symmetric interaction prohibits an efficient scaling with the angular basis size. In this section, we describe a method that scales well but is restricted to the case where the interaction can be written as a sum of two interactions that are cylindrically symmetric around different axes.

5.5.1 Basic ideas and principles

We illustrate the basic ideas and principles of the method by discussing the specific example of a linear diatomic molecule in an external electromagnetic
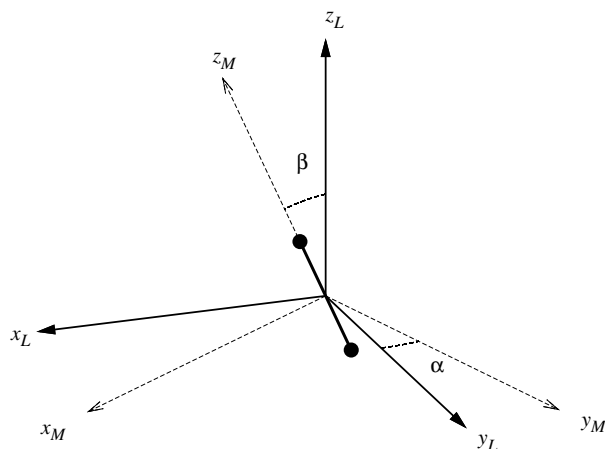
Figure 5.10: The orientation of the molecular coordinate system (M, dashed) with respect to the laboratory fixed system (L, solid). In the figure only the Euler angles α, β are relevant.

field. In Fig. 5.10 we show the coordinate systems which are relevant for the field-molecule problem. The coordinates (x_L, y_L, z_L) specify the laboratory (L) fixed coordinate system defined by the external polarization vector. We assume that the field is linearly polarized and return to the generalization to elliptically polarized light in Sec. 5.5.3. The coordinate system denoted by subscript M is the molecular fixed frame and is rotated by the Euler angles (α, β, γ) with respect to the laboratory fixed system. The rotation is accomplished by an α rotation around the z_L-axis, followed by a β rotation around the y_M-axis, and finally a γ rotation around the z_M-axis. For the case considered the only really distinct geometries are associated with the angle β. Results for different orientations due to the angle α are trivially related by a simple rotation around the z_L axis. Also the γ rotations around the molecular axis are insignificant as a consequence of the axial symmetry of the molecule.

We want to determine how the wave function of an electron is affected by the operators $V^{(M)}$ and $V^{(F)}(t)$, corresponding to the interaction with the nuclei and the field, respectively. We assume that we can treat these two operators separately, which is the case in a split-operator approach as described in Secs. 5.3.1 and 5.3.2. Our strategy is first to represent the wave function in the molecular frame and calculate the action of $V^{(M)}$. Secondly, we transform the updated wave function to the laboratory fixed frame and
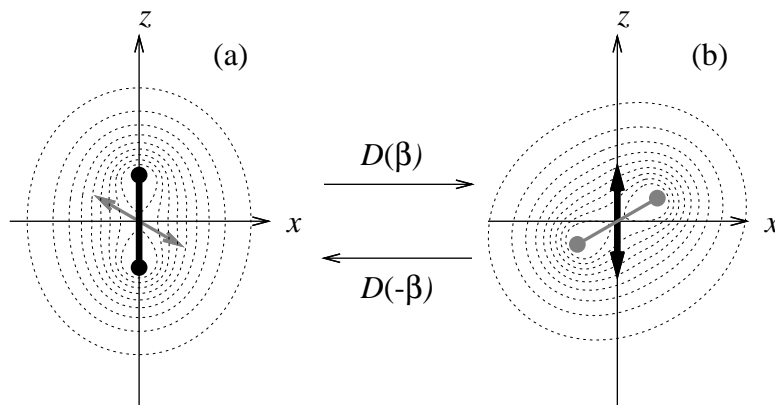
Figure 5.11: Schematic picture of the rotation operation. The contour lines indicate the field free $1\sigma_g$ ground state of $H_2{}^+$ in the xz plane. The double headed arrow shows the direction of the laser polarization vector. In (a) we calculate the action of the molecular potential and express the wave function in the molecular frame with the internuclear axis parallel to the z axis. In panel (b) we transform the wave function to the laboratory fixed system with the laser polarization parallel to the z axis in order to propagate by the field interaction. The transformation between the two frames is represented by the rotation operator D.

apply the operator $V^{(F)}$. Finally we can return to the molecular frame by the inverse rotation. These forward (β) and backward ($-\beta$) rotations of the wave function are illustrated in Fig. 5.11. The active interaction ($V^{(M)}$ or $V^{(F)}$) is marked by black and the inactive operation is gray. This propagation scheme for arbitrary orientation of the polarization axis with respect to the internuclear axis exhibits the strength of the present approach since it allows us to perform the calculations very efficiently. Whenever we apply an axially symmetric operator, we do not mix different m states provided that the wave function is expressed in the proper reference frame. Thus we can apply the operator separately on each different m state. The decoupling of different m states means effectively that we have reduced the three-dimensional problem to a number of two-dimensional problems in addition to two rotation operations.

The rotation transformation is in principle possible in all sets of coordinates and the separation in m applies to any coordinate system where the azimuthal angle ϕ is an independent variable, e.g. cylindrical, parabolic, or

spheroidal coordinates. The two unique features of the spherical representation (5.11) are that (i) the transformation matrix contains Wigner rotation functions which are known analytically and (ii) the transformation is guaranteed to be exactly unitary for functions that are bandwidth limited by a maximum $l = l_{\max}$, i.e., the population in states with $l > l_{\max}$ is zero.

5.5.2 Numerical implementation

In both the length and the velocity gauge, it is possible to write the reduced Hamiltonian as

$$\mathcal{H} = -\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{L^2}{2r^2} + V(t). \tag{5.67}$$

The operator $V(t)$ in Eq. (5.67) is the sum of the field interaction and the molecular potential

$$V(t) = V_{r,\theta_L}^{(F)}(t) + V_{r,\theta_M}^{(M)}, \tag{5.68}$$

where the subscripts denote the variables on which the operators act. θ_M is the polar angle in the molecular frame [Fig. 5.11 (a)] and θ_L the polar angle in the laboratory fixed system [Fig. 5.11 (b)]. To calculate the action of V in the propagation we make the split in accordance with Eq. (5.54)

$$e^{-iV(\bar{t})\Delta t} \approx e^{-iV^{(M)}\Delta t/2} e^{-iV^{(F)}(\bar{t})\Delta t} e^{-iV^{(M)}\Delta t/2}. \tag{5.69}$$

For each radial grid point r_i we write the wave function as a vector $\boldsymbol{f}^{(M)}(r_i,t)$ in the spherical harmonics basis, cf. Eq. (5.60), where the coefficients refer to the molecular frame. In the molecular frame, the operator $V^{(M)}$ is block diagonal by m as shown in the upper row of Fig. 5.12. The operator exponential can therefore be applied as a number of cylindrically symmetric problems as described in the previous section. After having applied the molecular potential we have obtained an intermediate vector of coefficients $\boldsymbol{f}^{(M,\mathrm{int})}(r_i,t)$ that refer to the molecular frame. These coefficients can now be transformed to the laboratory fixed frame. We relate the expansion in spherical harmonics in different frames by representation of the rotation operator in spherical harmonics, i.e., the Wigner rotation matrix $\boldsymbol{D}(\alpha,\beta,\gamma)$. The laboratory fixed expansion coefficients are then obtained as $\boldsymbol{f}^{(L,\mathrm{int})}(r_i,t) = \boldsymbol{D}(\alpha,\beta,\gamma) \cdot \boldsymbol{f}^{(M,\mathrm{int})}(r_i,t)$. This matrix multiplication is much faster than a full matrix multiplication since the rotation does not mix different l's and $\boldsymbol{D}(\alpha,\beta,\gamma)$ is consequently sparse. If the elements of the vectors are permuted so that the expansion coefficients are grouped by l instead of m, the rotation matrix becomes block-diagonal as shown in the middle row of Fig. 5.12. Note that the rotation operation is independent of the radial coordinate and we can therefore use the same rotation operation on all the vectors $\boldsymbol{f}^{(M,\mathrm{int})}(r_i,t)$ for different r's.

Having obtained the wave function in the laboratory fixed frame, we can easily apply the field interaction operator. Again, without m-couplings, the
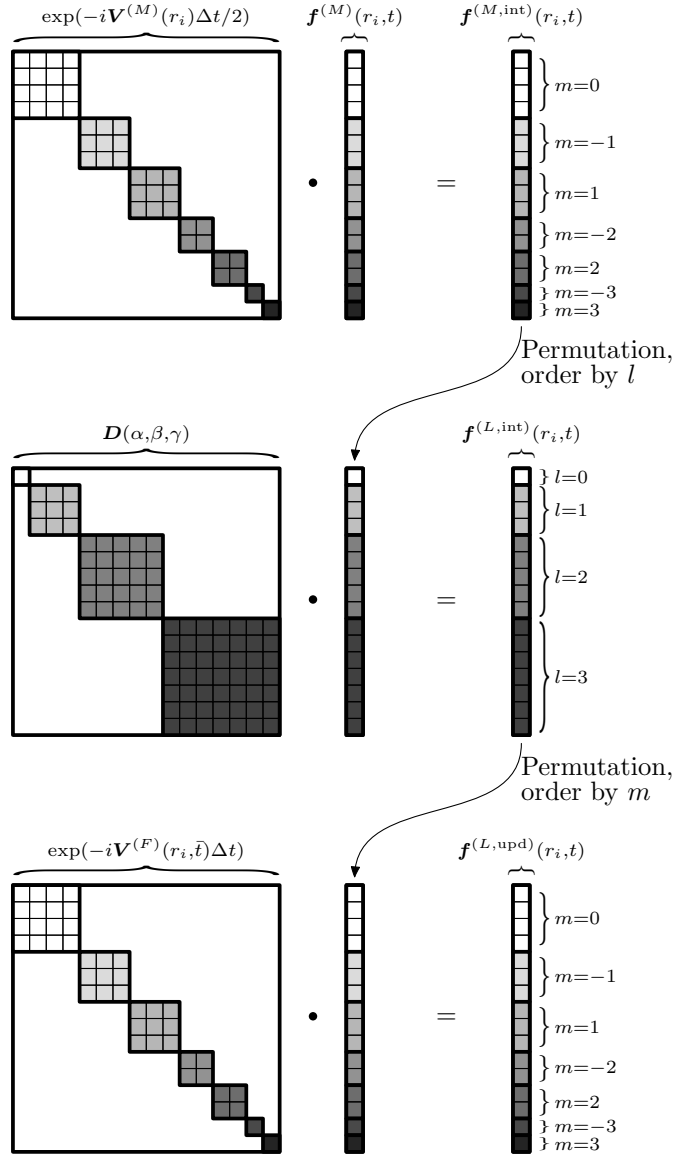
Figure 5.12: Schematic picture of the matrix representations of $e^{-iV^{(M)}\Delta t/2}$ and $e^{-iV^{(F)}(\bar{t})\Delta t}$ that enter the split-operator. For simplicity, we show the length gauge operators which do not couple $\boldsymbol{f}(r_i,t)$ and $\boldsymbol{f}(r_j,t)$ for $i \neq j$. In the upper row, the molecular potential is applied in the molecular frame, which allows for a block diagonalization by m. If $V^{(M)}$ is symmetric under inversion, the matrix can be divided further in even and odd parity blocks as in Fig. 5.4. In the middle row, the wave function is rotated into the laboratory fixed frame by the rotation matrix $\boldsymbol{D}$. The spherical harmonics representation of the rotation matrix is block diagonal by l. In the lower row, the field operator is applied in the laboratory fixed frame.

individual two-dimensional problems can be solved straightforwardly as in Sec. 5.4.2. The result is an updated vector $\boldsymbol{f}^{(L,\text{upd})}(r_i,t)$, which contains the expansion coefficients expressed in the laboratory fixed frame. Finally we return to the molecular frame by the inverse rotation transformation $\boldsymbol{f}^{(M)}(r_i,t) = \boldsymbol{D}^\dagger(\alpha,\beta,\gamma) \cdot \boldsymbol{f}^{(L)}(r_i,t)$.

The whole procedure is shown in Fig. 5.12. Note that all matrices are sparse. The total complexity scaling for all steps is $\mathcal{O}(n_r l_{\max}^{2.7})$. If we had chosen to express both operators in, e.g., the laboratory fixed frame, the molecular potential and its exponential would be represented by dense matrices that require $\mathcal{O}(n_r l_{\max}^4)$ operations for the matrix-vector multiplications.

5.5.3 General applications

We have chosen the linear molecule interacting with a linearly polarized field to illustrate our method, but a similar approach can be used in a much broader range of three-dimensional problems. For example we could consider an elliptically polarized field. In the split operator method we take the time step Δt to be small enough so that the field can be taken to be constant both in magnitude and polarization direction within the small time interval. We can therefore consider a time-dependent laboratory frame, which follows the instantaneous polarization direction. If we make the transformation from the molecular frame into the new laboratory frame, we are again able to treat the field as being linear and use a propagation scheme as discussed above. We refer to Ref. [3] for a more detailed discussion of a similar approach applied to atoms in elliptically polarized fields. Our method can also be extended to arbitrary nuclear positions. For any nuclear configuration, we can attach a coordinate system to each nucleus with a z axis from the origin to the nucleus. Then we decompose the molecular potential to a sum of nuclear potentials, each of which can be propagated with azimuthal symmetry in their own reference frame. Despite the fact that we now need rotations between the coordinate systems belonging to all of the nuclei, the total calculation is still in the same complexity class with respect to scaling in $l_{\max}$.

5.6 Summary - computational complexity

This section summarizes the computational complexity of the operations described previously in this chapter. Table 5.1 summarizes the computational complexity of the individual steps that enter the split-operator for cylindrically symmetric problems. Typical dimension used in real calculations include several thousand of radial grid points while the number of angular basis functions is much lower, often below 30. Even if the radial kinetic operator uses the FFT method, it is usually the operators that scale quadratically with the angular basis size that are most time consuming.

Operation	Method	Gauge	Complexity
T_r	Fast Fourier	LG,VG	$\mathcal{O}(n_r \log_2(n_r)n_l)$
T_r	Crank-Nicolson	LG,VG	$\mathcal{O}(n_r n_l)$
T_l	Diagonal multiplication	LG,VG	$\mathcal{O}(n_r n_l)$
$V^{(M)}(r,\zeta)$	Static matrix	LG,VG	$\mathcal{O}(n_r n_l^2)$
$V(r,\zeta,t)$	Angular grid	LG	$\mathcal{O}(n_r n_l^2)$
$V^{(F)}(r,\zeta,t)$	Matrix representation	LG	$\mathcal{O}(n_r n_l^2)$
$V^{(F)}(r,\zeta,t)$	Crank-Nicolson	LG	$\mathcal{O}(n_r n_l)$
$V^{(F,1)}$	Matrix representation	VG	$\mathcal{O}(n_r^2 n_l^2)$
$V^{(F,1)}$	Grid+Crank-Nicolson	VG	$\mathcal{O}(n_r n_l^2)$
$V^{(F,2)}$	Matrix representation	VG	$\mathcal{O}(n_r n_l^2)$
$V^{(F,2)}$	Crank-Nicolson	VG	$\mathcal{O}(n_r n_l)$

Table 5.1: The complexity of the individual steps that constitute the split-operator for cylindrically symmetric problems. n_r is the number of radial grid points and $n_l = l_{\max} + 1 - |m|$ is the number of angular basis functions.

The operational count for both the length and velocity gauges will scale as $\mathcal{O}(n_r n_l^2)$. The length gauge will, however, be faster by a constant factor, since the velocity gauge split-operator (5.54) consists of five steps that scale in this way, while the length gauge split-operator can be formulated so that only a single step scales as $\mathcal{O}(n_r n_l^2)$. The only situation where the overall propagation can be reduced to a lower complexity class is when the molecular potential is spherically symmetric, e.g., if we consider an atom. In that case, $V^{(M)}(r,\zeta) = V^{(M)}(r)$ becomes diagonal in the spherical harmonics representation and the corresponding step only requires $\mathcal{O}(n_r n_l)$ operations. If we now use the Crank-Nicolson method for the field operator in the length gauge, the overall propagation is limited by the $\mathcal{O}(n_r \log_2(n_r)n_l)$ scaling of the fast Fourier transformation. This scaling can even be improved further to $\mathcal{O}(n_r n_l)$ if also the radial kinetic operator is applied by the Crank-Nicolson scheme.

Now we turn to the case without cylindrical symmetry. The number of angular basis functions is then $(l_{\max}+1)^2$, and thus we expect that the typical scaling for operators that are non-diagonal is on the order of the $(l_{\max}+1)^2$-dimensional matrix-vector multiplication, i.e., $\mathcal{O}(l_{\max}^4)$. As shown in Sec. 5.5, it is possible to avoid such a full matrix-vector multiplication if the problem consists of two or more operators that are cylindrically symmetric around different axes. For that case, the complexity scaling is given in Table 5.2. The table shows that both in the length and velocity gauges it is possible to perform the total time evolution in a number of operations that scales only as $\mathcal{O}(n_r l_{\max}^{2.7})$. The difference between $\mathcal{O}(n_r l_{\max}^4)$ and $\mathcal{O}(n_r l_{\max}^{2.7})$ is enormous even for quite a limited number of angular momentum states.

Operation	Method	Gauge	Complexity
T_r	Fast Fourier	LG,VG	$\mathcal{O}(n_r \log_2(n_r) l_{\max}^2)$
T_r	Crank-Nicolson	LG,VG	$\mathcal{O}(n_r l_{\max}^2)$
T_l	Diagonal multiplication	LG,VG	$\mathcal{O}(n_r l_{\max}^2)$
$V^{(M)}(r, \zeta_M)$	Static matrix	LG,VG	$\mathcal{O}(n_r l_{\max}^{2.7})$
$\mathscr{D}(\alpha, \beta, \gamma)$	Static matrix	LG,VG	$\mathcal{O}(n_r l_{\max}^{2.7})$
$V^{(F)}(r, \zeta_L, t)$	Angular grid	LG	$\mathcal{O}(n_r l_{\max}^{2.7})$
$V^{(F)}(r, \zeta_L, t)$	Matrix representation	LG	$\mathcal{O}(n_r l_{\max}^{2.7})$
$V^{(F)}(r, \zeta_L, t)$	Crank-Nicolson	LG	$\mathcal{O}(n_r l_{\max}^2)$
$V_{r, \zeta_L}^{(F,1)}$	Matrix representation	VG	$\mathcal{O}(n_r^2 l_{\max}^{2.7})$
$V_{r, \zeta_L}^{(F,1)}$	Grid+Crank-Nicolson	VG	$\mathcal{O}(n_r l_{\max}^{2.7})$
$V_{r, \zeta_L}^{(F,2)}$	Matrix representation	VG	$\mathcal{O}(n_r l_{\max}^{2.7})$
$V_{r, \zeta_L}^{(F,2)}$	Crank-Nicolson	VG	$\mathcal{O}(n_r l_{\max}^2)$

Table 5.2: The complexity of the individual steps that constitute the split-operator for the class of non-cylindrically symmetric problems that consist of two operators that are cylindrically symmetric around different axes. n_r is the number of radial grid points and $l_{\max}$ defines the truncation in the spherical harmonics expansion. The total number of angular basis functions is $n_{Y_{lm}} = (l_{\max} + 1)^2$.

Chapter 6

Parallelization of the split-operator algorithm

Due to the recent development of modern microchips, the computational power available to scientists has been a rapidly increasing function of time over the last decades. Hence, one might expect that more and more complex systems can be solved numerically since the execution time of any algorithm will decrease as computers become faster. It turns out that this expectation is not strictly true.

The pure clock frequency of the central processing unit (CPU) of any new computer has reached a stagnant level in the years between 2003 and 2007. If we now consider an algorithm that executes a single instruction per clock cycle, the execution time will not be much faster when the algorithm is run on a computer from 2007 compared with a computer from 2003[1]. The reason why CPU manufacturers have not been able to increase the clock frequency is that high clock frequencies cause heat production and power consumption beyond the limit of economical feasibility. Instead of increasing the clock frequency, CPU manufacturers have resorted to make CPUs more powerful by collecting several independent computational cores within a single CPU. Although such a multicore CPU in some sense is more computationally powerful than a single core CPU, it is not guaranteed that a particular algorithm can take advantage of the additional power. The full potential of a multicore CPU can only be exploited for problems that can be *parallelized*, i.e. problems that can be divided into individual subproblems, which can be solved separately. One example of a problem that can be parallelized is the block-diagonal matrix-vector multiplication shown in Fig. 6.1. This problem can be executed as two separate problems since neither of the two subsystems needs any information from the other subsystem.

A well designed parallelization scheme requires optimal *load balancing*,

---

[1]This assumes that other factors such as loading time between memory and chip can be neglected.
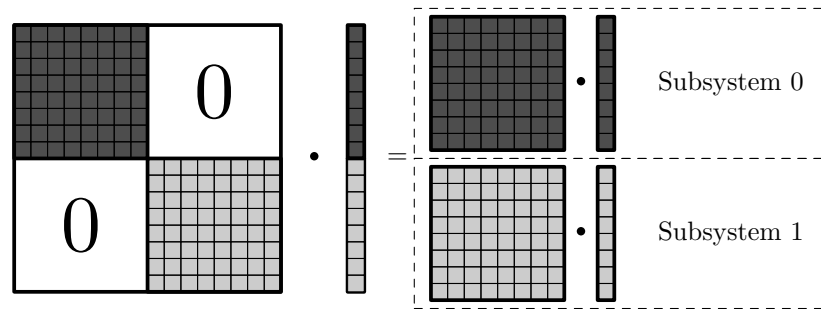
Figure 6.1: A block diagonal matrix-vector multiplication can be decomposed in independent subsystems. The matrix-vector multiplication within each subsystem can be done in parallel on different cores.

which means that the subproblems must be equally complex to solve. If, for example, the block diagonal matrix in Fig. 6.1 had consisted of two blocks of different size, it would take longer time to perform the multiplication of the larger block than the smaller block. In this case, the core assigned to the smaller problem will stay idle for some time while the other core finishes its job, and hence, the total computational power of the two cores is not fully utilized.

Parallel computation systems can be divided into two major classes, namely the shared memory and distributed memory systems. For shared memory systems, all computational cores have direct access to the same amount of memory with a large bandwidth of several GB/s and low latency of a few nanoseconds. Shared memory systems with more than a small number of cores are often highly specialized machines provided by a small number of manufacturers. Both supply of and demand for such machines are quite low and, accordingly, prices are relatively high.

In distributed memory systems, each computational core has its own amount of memory without direct access to data connected to the other cores. We often use the term "node" for one computational core and its associated memory. For a distributed memory system, the situation in Fig. 6.1 corresponds to one node that has access only to the data indicated by dark gray while the other node has access only to the light gray data. If some of the locally stored data is to be used by another node, it must be transfered by some network connection. A typical bandwidth of a network connection is around 10-100 MB/s and has a latency of several microseconds – both values are orders of magnitudes slower than direct access to the local memory. This limitation must certainly be taken into account when developing algorithms for distributed memory systems. The primary advantage of distributed memory systems is the low price. A very powerful system can be built from cheap standard components, for example a cluster of ordinary

desktop PCs connected by an ethernet network.

6.1   Split operator method – a parallel version

This section describes how the split operator method outlined in chapter 5 can be parallelized. We assume that the algorithm runs on n_p nodes in a distributed memory system.

Before we discuss the actual implementation, it is useful to introduce the terminology "distribution by r-major ordering" when the vector $\mathbf{\Phi}^{(r)}$ is distributed so that the k'th node is assigned the elements between $k \dim(\mathbf{\Phi}^{(r)})/n_p$ and $(k+1)\dim(\mathbf{\Phi}^{(r)})/n_p - 1$. Similarly, "distribution by l-major ordering" refers to the corresponding distribution of the vector $\mathbf{\Phi}^{(l)}$. Visually, the distributions can be seen as the two vectors in Fig. 5.7 being divided in n_p pieces of equal lengths and each piece being assigned to a single node (see also Fig. 6.3 below).

6.1.1   Parallelization of the length gauge algorithm

We recall the expression for the split operator approximation to the time evolution operator in the length gauge, cf. Eq. (5.24)

$$\mathscr{U}(t+\Delta t, t) = e^{-i\frac{\Delta t}{2}\left(-\frac{1}{2}\frac{\partial^2}{\partial r^2}\right)} e^{-i\frac{\Delta t}{2}\frac{L^2}{2r^2}} e^{-i\Delta t V(\mathbf{r},\bar{t})} e^{-i\frac{\Delta t}{2}\frac{L^2}{2r^2}} e^{-i\frac{\Delta t}{2}\left(-\frac{1}{2}\frac{\partial^2}{\partial r^2}\right)}.$$

(6.1)

We now wish to perform each of the three different steps in parallel on n_p different nodes. Since the centrifugal term is trivial in the r-grid/spherical harmonics basis, we only consider the parallelization of the radial kinetic operator and the potential operator.

We start out by noting that 1) the radial kinetic operator is represented by a block diagonal matrix in r-major ordering and 2) the potential is represented by a block diagonal matrix in l-major ordering. Hence, it is possible to parallelize the two steps as shown in Fig. 6.2. The load balancing will be optimal if $n_p|n_r$ and $n_p|(l_{\max}+1)^2$ so that an equal number of blocks can be assigned to each node in both steps. In Fig. 6.2, the data is distributed such that each node only knows the data within the boundaries of the dashed lines. As we first apply the radial kinetic operator, the data must initially be distributed among the nodes according to the r-major ordering. After the radial kinetic operator is applied in parallel on all nodes, the centrifugal operator must be applied. This operation can also be done in parallel since this is just a phase factor that should be multiplied on each element of the vector $\mathbf{\Phi}^{(r)}(t)$ which, clearly, can also be done locally. Next, the potential is to be applied. A parallel execution of this operation now requires that the vector representation of the wave function is distributed among the nodes in l-major ordering. In Sec. 5.2 we did not pay too much attention to the transformation between the r-major and l-major ordering – the difference
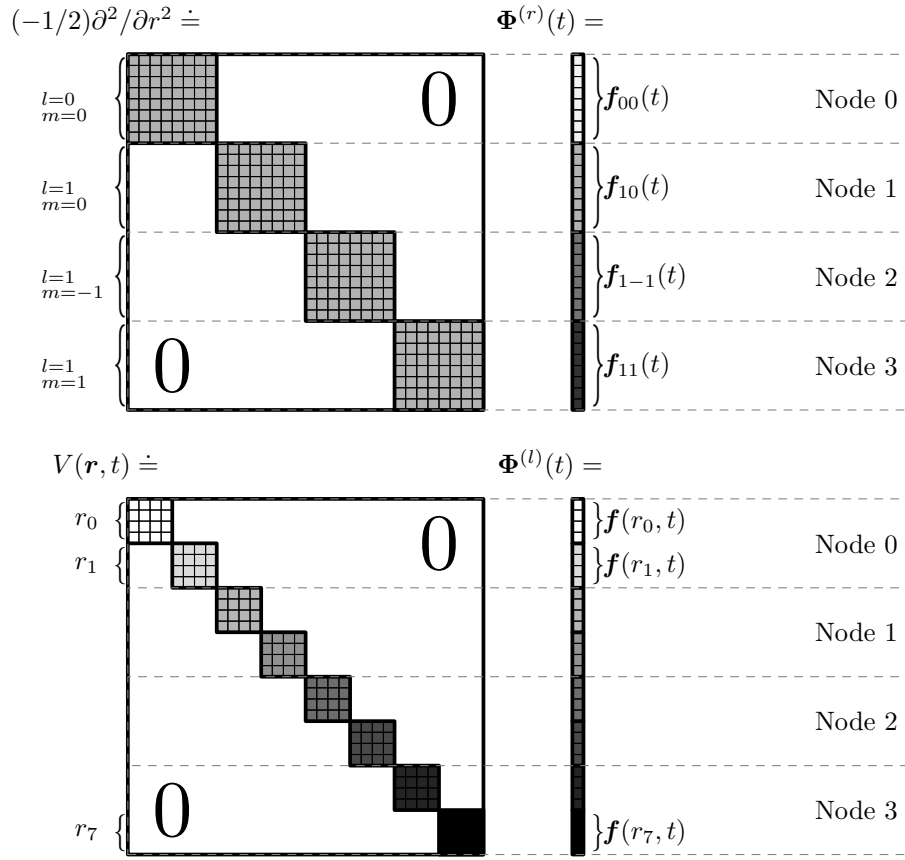
Figure 6.2: Matrix representations of the radial kinetic operator and the potential for $n_r = 8$ and $l_{\max} = 1$. Each of the operators can be applied in parallel by $n_p = 4$ nodes.

between the two forms was only a matter of permutation. In the present discussion, however, the vector is distributed among the nodes and the actual implementation of such a permutation becomes slightly tricky. As Fig. 6.3 shows, a reordering requires that nearly all numbers are redistributed among the nodes. If the dimensionality of the vectors is large, a great amount of data has to be transferred over the relatively slow network connection. A realistic example of a basis size can be $n_r = 4096$ and $l_{\max} = 31$. If all numbers are stored in 16 Bytes complex format, the total size of the wave function vector is

$$4096 \times (31 + 1)^2 \times 16\,\text{B} = 64\,\text{MB}. \tag{6.2}$$

If this amount of data is to be transferred over the present standard network with a bandwidth of 1 Gbit/s, it will take at least 0.5 s. This transfer time turns out to be a bottleneck for an efficient parallelization for bases of this
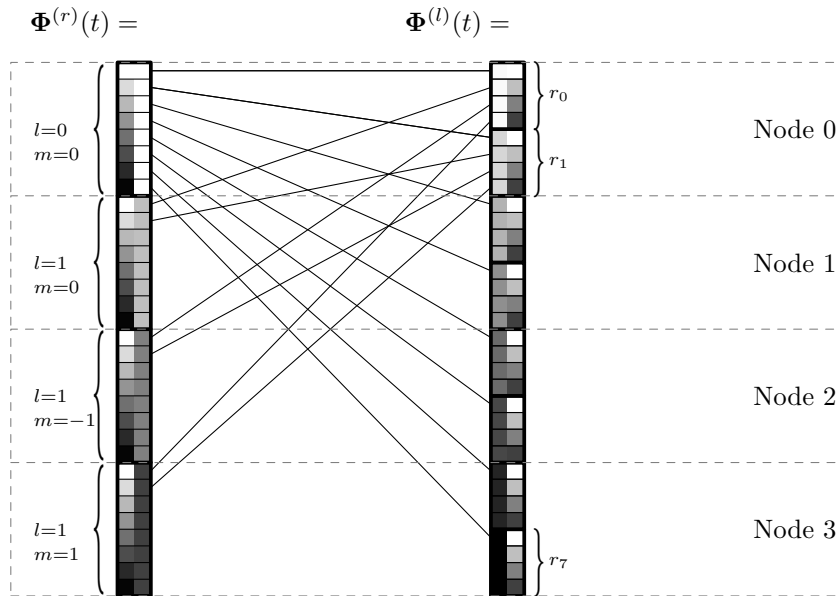
Figure 6.3: Permutation between r-major and l-major ordering when the vectors are distributed among four nodes. The first color indicates the r-index while the second color indicates the (l, m) index. The lines show the numbers that must be communicated in order to transform between the different orderings. For transparency, only the communication lines that involve node 0 are shown.

size and larger – the transfer time simply becomes significantly larger than the time it takes for the actual calculation. Hence, for large bases it is necessary to reduce the data transfer in order to exploit the advantages of parallelization on a distributed memory machine.

There is no way to reduce the amount of transferred data with the split operator *spectral* method. The reason is that derivatives are calculated by an expansion in global functions as discussed in Secs. 3.5 and 3.6. In order to construct this expansion, the values from *all* grid points are needed and, hence, the r-major distribution shown in Fig. 6.2 is strictly necessary in order to apply the radial kinetic operator. On the other hand, the potential requires an l-major distribution since it introduces couplings between all angular momentum components.

Now we propose that the amount of transfer can be greatly reduced if the second order derivative is approximated by a finite difference and the exponential operator by the Crank-Nicolson scheme. We will show that a complete redistribution can be avoided. We only need one basic distribution model, namely the l-major distribution. The only requirement is that the number of radial grid points must be much larger than the number of nodes
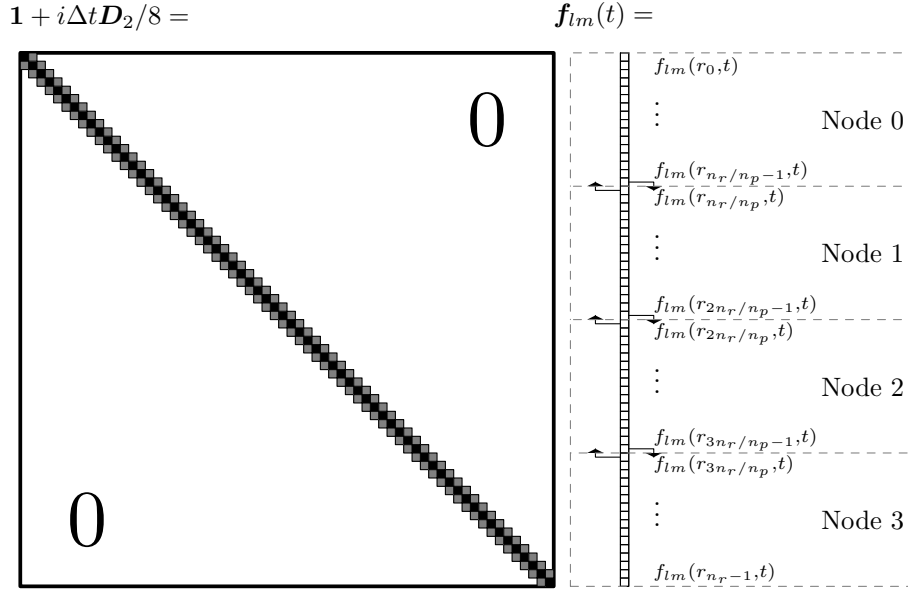
$1 + i\Delta t \boldsymbol{D}_2/8 =$ $\boldsymbol{f}_{lm}(t) =$

Figure 6.4: The explicit step of the Crank-Nicolson method. The vector $\boldsymbol{f}_{lm}(t)$ is distributed on four nodes according to l-major ordering. Communication between the nodes is only needed at the boundaries as indicated.

– a requirement that is always fulfilled in practice[2]. We recall how the radial kinetic operator is applied with a Crank-Nicolson scheme, cf. Eq. (5.26)

$$e^{-i\frac{\Delta t}{2}\left(-\frac{1}{2}\frac{\partial^2}{\partial r^2}\right)} f_{lm}(r,t) \doteq \left(\mathbf{1} - \frac{1}{8}i\Delta t \boldsymbol{D}_2\right)^{-1} \cdot \left(\mathbf{1} + \frac{1}{8}i\Delta t \boldsymbol{D}_2\right) \cdot \boldsymbol{f}_{lm}(t). \quad (6.3)$$

With the r-major distribution described above, one whole vector $\boldsymbol{f}_{lm}(t)$ is available on a single node. On the contrary, if the data are distributed by the l-major distribution, the vector is scattered among the nodes as shown in Fig. 6.4. It is easy to see that the intermediate vector obtained from the explicit step $\left(\mathbf{1} + \frac{1}{8}i\Delta t \boldsymbol{D}_2\right) \cdot \boldsymbol{f}_{lm}(t) = \boldsymbol{f}_{lm}^{(\text{int})}$ can be computed in parallel with little data transfer. Indeed, this is just a tridiagonal matrix multiplication. In order to calculate the i'th element of the product vector we only need to know the $i \pm 1$ and i'th elements of the original vector. Therefore, we only need to exchange the elements at the border between the nodes as indicated in Fig. 6.4. When these numbers are transferred, each node can calculate its own part of the product vector. The product vector itself is automatically l-major distributed by this strategy. The total number that must be transferred data for all values of l and m is only $(2n_p - 2)(l_{\max} + 1)^2$ numbers, i.e., much less than the $n_r(l_{\max} + 1)^2$ required for an l- to r-major redistribution. With the distributed intermediate vector at hand, we now

---

[2]The previous example with $n_r = 8$ and four nodes is evidently *not* a realistic example.

wish to solve the implicit system

$$\left(\mathbf{1} - \frac{1}{8} i \Delta t \boldsymbol{D}_2\right) \cdot \boldsymbol{f}_{lm}^{(\mathrm{upd})} = \boldsymbol{f}_{lm}^{(\mathrm{int})} \tag{6.4}$$

for the updated vector $\boldsymbol{f}_{lm}^{(\mathrm{upd})}$ which also has to be distributed according to l-major ordering. Such a distributed tridiagonal system of equations is, however, impossible to solve in parallel by any standard algorithm. As we show in chapter 7, the problem *can* in fact be solved in parallel for all l, m by communicating just $2 \times ((l_{\max} + 1)^2 + 1)(n_p - 1)$ numbers. Compared with the complete redistribution, the amount of transferred data is now reduced by a factor of $\sim n_r/n_p$ – a number that can easily exceed 1000 for a realistic calculation.

In summary, the split-operator method in the length gauge can be parallelized efficiently on a distributed memory system if the radial kinetic operator is approximated by the Crank-Nicolson method with the second order derivative being approximated by a three-point finite difference. On the other hand, the use of spectral methods results in poor parallelization properties.

6.1.2 Parallelization of the velocity gauge algorithm

The split operator for the velocity gauge is described in Sec. 5.3.2. We will use the same technique as for the length gauge in order to parallelize the algorithm. The basic distribution model is by l-major ordering. We only need to consider the two additional interaction operators

$$V^{(F,1)} = -i A_{\mathrm{VG}}(t) \cos\theta \frac{\partial}{\partial r}, \tag{6.5}$$

$$V^{(F,2)} = i A_{\mathrm{VG}}(t) \frac{1}{r} \left(\cos\theta + \sin\theta \frac{\partial}{\partial \theta}\right). \tag{6.6}$$

The second operator is diagonal in the radial coordinate and, hence, the matrix representation is block diagonal for each value of r_i in l-major ordering, similar to the lower panel of Fig. 6.2, and the parallelization is straightforward. As described in Sec. 5.3.2, the first operator $V^{(F,1)}$ is applied first by a transformation to an angular grid representation for each value of m. Then, the radial derivative is applied for each angular direction, cf. Eq. (5.58). For simplicity, we consider only $m = 0$ below. For each radial grid point, each node can transform from the spherical harmonics representation to the angular grid representation. This transformation is shown in the left panel of Fig. 6.5. When we have obtained the angular grid representation, the interaction is diagonal in the angular variable and, hence, the interaction can be applied for each angular grid point separately as shown in the right panel of Fig. 6.5. Since we use the Crank-Nicolson scheme, we must now
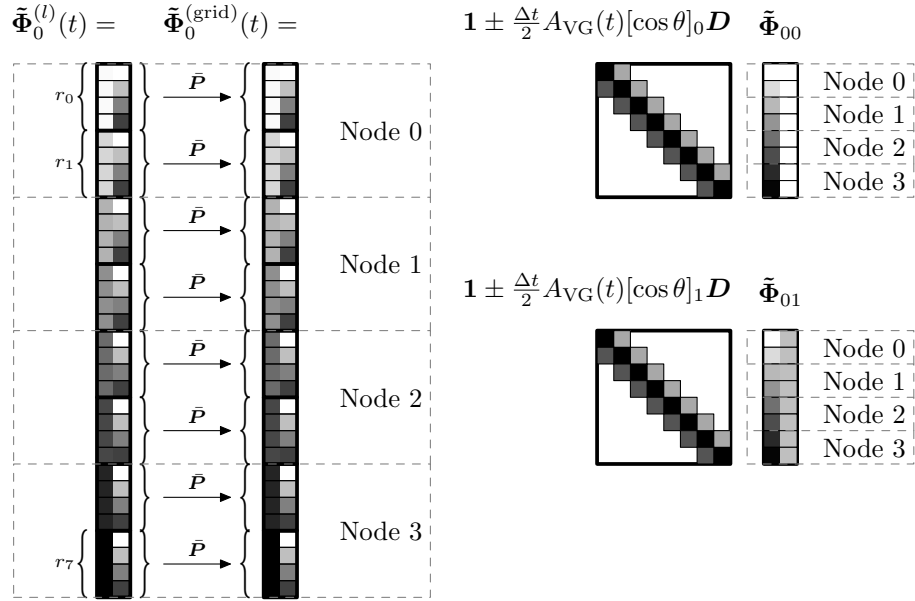
Figure 6.5: Parallel execution of $V^{(F,1)}$ by the Crank-Nicolson method. The first colors of the vector representations correspond to the radial coordinate. In spherical harmonics representation, $\tilde{\boldsymbol{\Phi}}_0^{(l)}$, the second color corresponds to the angular number l. In grid representation, $\tilde{\boldsymbol{\Phi}}_0^{(\mathrm{grid})}$, the second color corresponds to the angular grid point.

multiply and solve a number of tridiagonal matrix equations with the vector being distributed over the nodes. But we already have shown in Sec. 6.1.1 how to solve these two problems in parallel. The present method provides good load balancing and is expected to scale well as long as $n_r \gg n_p$.

Hopefully, it is now clear why it is advantageous to transform from spherical harmonics- to angular grid representation. If we stick to the spherical harmonics representation, $V^{(F,1)}$ will introduce couplings between elements whose radial *and* angular indices differ by unity. Thus, the matrix representation takes band-diagonal form with bandwidth $l_{\max} + 2$ as shown in Fig. 6.6. The matrix multiplication, which constitutes the explicit step in the Crank-Nicolson scheme can still be solved in parallel with little data transfer, even when the matrix is band-diagonal. However, the implicit step, namely the solution of the band-diagonal matrix equation, is very hard to solve in parallel. It is exactly the desired effect of the transformation into the angular grid representation to *avoid* the band-diagonal matrix equation and only solve a number of separate tridiagonal matrix equations. Fortunately, both the transformation and the solutions of the tridiagonal systems can be computed efficiently in parallel. From a mathematical point of view, we can interpret the block-diagonal transformation matrices $\mathbf{1}_{n_r} \otimes \bar{\boldsymbol{P}}$ and

$$1 \mp i\tfrac{\Delta t}{2}V^{(F,1)} = 1 \mp \tfrac{\Delta t}{2}A_{\mathrm{VG}}(t)\cos\theta\tfrac{\partial}{\partial r} \doteq \qquad\qquad \tilde{\boldsymbol{\Phi}}_0^{(l)}(t) =$$
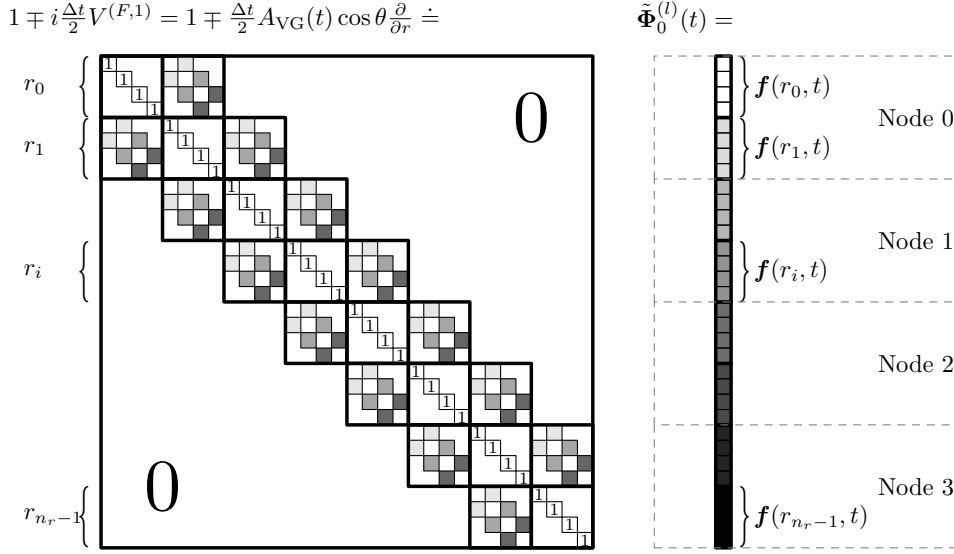

Figure 6.6: Matrix representation of $1 \mp i\tfrac{\Delta t}{2}V^{(F,1)}$ in the radial grid/spherical harmonics basis. The interaction is non-diagonal in both the radial and the angular basis and, hence, couplings occur between elements whose radial *and* angular indices differ by unity to give a band-diagonal matrix structure. In general, the bandwidth is $l_{\max} + 2$. It is difficult to solve a band-diagonal matrix equation beyond the tridiagonal in parallel.

$\mathbf{1}_{n_r} \otimes \left(\mathrm{diag}(\boldsymbol{w}) \cdot \bar{\boldsymbol{P}}^T\right)$ as the transformations that bring the band-diagonal matrix on a block-tridiagonal form.

In conclusion, the split-operator method can also be parallelized efficiently in the velocity gauge on a distributed memory system. The algorithm is slightly more complicated that the length gauge counterpart as there are two distinct factors in the velocity gauge split-operator that require communication between the nodes. It is in fact possible to reduce the number of steps that requires communication to one if we modify the split-operator (5.54) so that T_r and $V^{(F,1)}$ are treated together in a *single* Crank-Nicolson step. However, since both factors only need a small amount of data to be communicated, the actual gain may not be noticeable. In any case, we expect that the execution time should decrease with the number of nodes n_p as long as $n_r \gg n_p$.

Chapter 7

A parallel algorithm for solving tridiagonal matrix equations

This chapter describes how to solve a tridiagonal matrix equation for the unknown N-dimensional column vector $\boldsymbol{x}$

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}, \tag{7.1}$$

where $\boldsymbol{A}$ is a non-singular tridiagonal N by N matrix and $\boldsymbol{b}$ is a known N-dimensional column vector.

First, in Sec. 7.1 we recall a standard solution that works as a serial algorithm. In Sec. 7.2 we propose a new parallel algorithm. This algorithm is specifically designed to run on a distributed memory system with low interconnection bandwidth and high latency. In order to run well on such a system, the total amount of data that is to be distributed among the computational nodes must be kept low.

7.1 Solution by LU factorization

One of the standard methods that can be used to solve Eq. (7.1) is the solution by LU factorization. The tridiagonal matrix $\boldsymbol{A}$ can be LU factorized as [15]

$$\boldsymbol{A} = \boldsymbol{L} \cdot \boldsymbol{U}, \tag{7.2}$$

where $\boldsymbol{L}$ and $\boldsymbol{U}$ are lower and upper bidiagonal (Fig. 7.1) with

$$l_i = \begin{cases} \frac{A_{10}}{A_{00}}, & i = 1 \\ \frac{A_{i,i-1}}{A_{i-1,i-1} - A_{i-2,i-1} l_{i-1}}, & i > 1 \end{cases} \tag{7.3}$$

$$d_i = \begin{cases} A_{00}, & i = 0 \\ A_{ii} - \frac{A_{i,i-1} A_{i-1,i}}{d_{i-1}}, & i > 0 \end{cases} \tag{7.4}$$
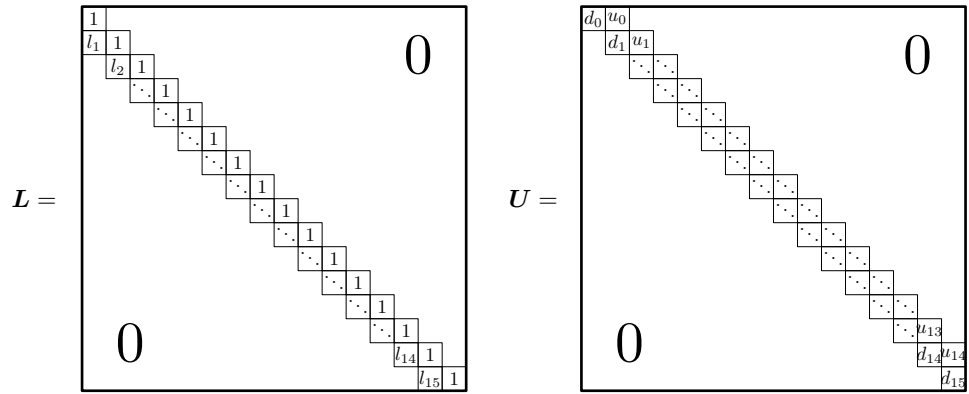
$$u_i = A_{i,i+1}. \tag{7.5}$$

Figure 7.1: The LU factorization of the 16-dimensional tridiagonal matrix $\boldsymbol{A}$ in a lower bidiagonal matrix $\boldsymbol{L}$ and an upper bidiagonal matrix $\boldsymbol{U}$.

With the LU factorization at hand it is straightforward to solve the problem as

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{L} \cdot (\boldsymbol{U} \cdot \boldsymbol{x}) = \boldsymbol{b}, \tag{7.6}$$

first by solving the intermediate system $\boldsymbol{L} \cdot \boldsymbol{y} = \boldsymbol{b}$ by forward substitution

$$\begin{aligned} y_0 &= b_0, & \text{(7.7)}\\ y_i &= b_i - l_i y_{i-1}, \quad i = 1, 2, \cdots, N-1, & \text{(7.8)} \end{aligned}$$

and then solve $\boldsymbol{U} \cdot \boldsymbol{x} = \boldsymbol{y}$ by backward substitution

$$\begin{aligned} x_{N-1} &= \frac{y_{N-1}}{d_{N-1}}, & \text{(7.9)}\\ x_i &= \frac{y_i - u_i x_{i+1}}{d_i}, \quad i = N-2, N-3, \cdots, 0. & \text{(7.10)} \end{aligned}$$

Both the forward and backward substitution are inherently serial processes. In order to calculate every new element, it is strictly necessary to know the preceding element. This fact makes the algorithm difficult to parallelize.

7.2   A parallel algorithm

This section describes how Eq. (7.1) can be solved in parallel on n_p nodes. For simplicity, we consider the situation where $n = N/n_p$ is an integer, although this requirement is not crucial for the algorithm to work. Furthermore, we assume that $\boldsymbol{b}$ is distributed over the nodes so that the k'th node holds the n dimensional vector $\boldsymbol{b}^{(k)}$, namely the elements of $\boldsymbol{b}$ between kn and $(k+1)n-1$. Similarly, we want the solution vector $\boldsymbol{x}$ to be distributed by the same pattern as vectors $\boldsymbol{x}^{(k)}$ on each node as shown in Fig. 7.2. The figures given in this section refer, for illustrative purposes, to a 16-dimensional
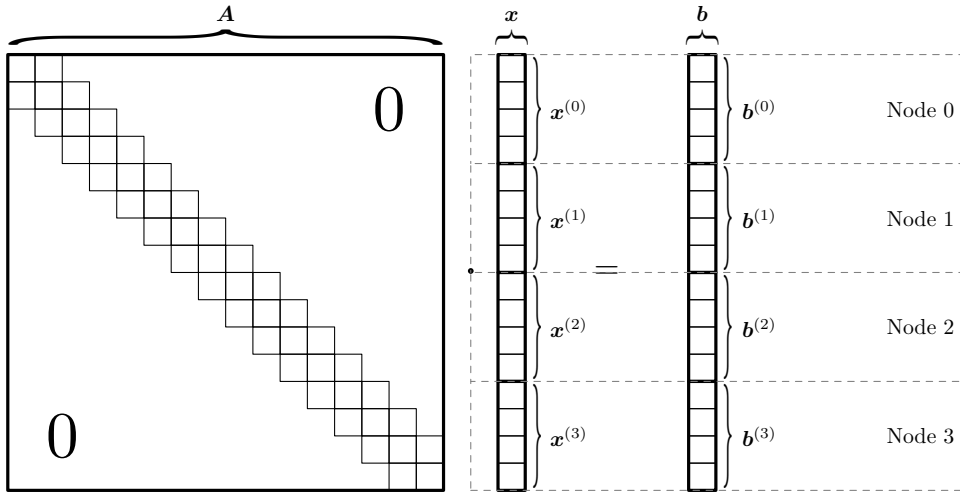
Figure 7.2: Distribution of the 16-dimensional vectors on four nodes.

system of equations distributed on four nodes. Nonetheless, the formulae will be completely general.

The parallel algorithm relies on a method that is closely related to the solution by LU factorization. We assume that all nodes know the LU factorized form of $\boldsymbol{A}$. This will typically not be a severe restriction – the LU factorization is easily obtained from $\boldsymbol{A}$ by Eqs. (7.3)-(7.5). Furthermore, for any given matrix $\boldsymbol{A}$ the LU factorization needs to be computed only once, so that the same LU factorization can be re-used if we want to solve Eq. (7.1) for many different right hand side vectors. As by the standard LU method we first solve the intermediate system $\boldsymbol{L} \cdot \boldsymbol{y} = \boldsymbol{b}$. First, we write the matrix $\boldsymbol{L}$ as

$$\boldsymbol{L} = \tilde{\boldsymbol{L}} + \Delta \boldsymbol{L}, \tag{7.11}$$

where $\tilde{\boldsymbol{L}}$ is obtained from $\boldsymbol{L}$ by replacing the numbers l_{kn}, $k = 1, 2, ..., n_p - 1$, by zeros. $\tilde{\boldsymbol{L}}$ is thereby block-diagonal

$$\tilde{\boldsymbol{L}} = \bigoplus_{k=0}^{n_p-1} \tilde{\boldsymbol{L}}^{(k)} = \mathrm{diag}(\tilde{\boldsymbol{L}}^{(0)}, \tilde{\boldsymbol{L}}^{(1)}, \cdots, \tilde{\boldsymbol{L}}^{(n_p-1)}), \tag{7.12}$$

$\tilde{\boldsymbol{L}}^{(k)}$ is an $n \times n$ block matrix which we assign to the k'th node. The numbers that are excluded from $\tilde{\boldsymbol{L}}$ are instead put in the remainder $\Delta \boldsymbol{L}$. The structures of the matrices $\tilde{\boldsymbol{L}}$ and $\Delta \boldsymbol{L}$ are shown in Fig. 7.3. $\Delta \boldsymbol{L}$ contains the numbers l_{kn} at the kn'th row and $(kn - 1)$'th column, i.e.

$$\Delta \boldsymbol{L} = \sum_{k=1}^{n_p-1} l_{kn} \boldsymbol{e}_{kn} \cdot \boldsymbol{e}_{kn-1}^T, \tag{7.13}$$

Figure 7.3: The lower bidiagonal matrix L from Fig. 7.1 can be written as the sum of the block diagonal matrix $\tilde{L}$ and the remainder ΔL.

where the unit vectors are N-dimensional column vectors.

The solution method now proceeds as follows. First, by forward substitution, we let each node solve the bidiagonal systems

$$\tilde{L}^{(k)} \cdot \tilde{y}^{(k)} = b^{(k)}, \tag{7.14}$$

$$\tilde{L}^{(k)} \cdot \Delta\tilde{y}^{(k)} = l_{kn}e_0^{(k)}, \quad k > 0, \tag{7.15}$$

where $e_0^{(k)}$ is the n-dimensional unit vector. Then we multiply the equation $L \cdot y = b$ by $\tilde{L}^{-1}$ from the left

$$\tilde{L}^{-1} \cdot L \cdot y = \left(1 + \tilde{L}^{-1} \cdot \Delta L\right) \cdot y = \tilde{L}^{-1} \cdot b. \tag{7.16}$$

Since $\tilde{L}$ is block-diagonal, the vector on the right hand side is built from the pieces $\tilde{y}^{(k)}$ that are the solutions of Eq. (7.14). Unfortunately, these pieces are not exactly equal to the desired intermediate solutions $y^{(k)}$ due to the fact that the matrix $1 + \tilde{L}^{-1} \cdot \Delta L$ is not the unit matrix. It is, however, quite close to being the unit matrix and has a structure which allows us to construct the solutions $y^{(k)}$ from the solutions of Eqs. (7.14) and (7.15). The result is given in Eqs. (7.23) and (7.24) below. The proof is as follows. First, we illustrate the structure of $\tilde{L}^{-1}\Delta L$ by calculating the m'th column

$$\tilde{L}^{-1} \cdot \Delta L \cdot e_m = \tilde{L}^{-1} \cdot \left(\sum_{k=1}^{n_p-1} l_{kn}e_{kn} \cdot e_{kn-1}^T\right) \cdot e_m = \sum_{k=1}^{n_p-1} l_{kn}\tilde{L}^{-1} \cdot e_{kn}\delta_{kn-1,m}, \tag{7.17}$$

i.e., only the columns $\{n-1, 2n-1, \ldots, (n_p-1)n-1\}$ are nonzero. We let $\Delta y^{(k)}, (k = 1, 2, \ldots, n_p-1)$ denote the nonzero $kn-1$'th column of $\tilde{L}^{-1} \cdot \Delta L$

$$\Delta y^{(k)} = \tilde{L}^{-1} \cdot \Delta L \cdot e_{kn-1} = l_{kn}\tilde{L}^{-1} \cdot e_{kn}, \tag{7.18}$$
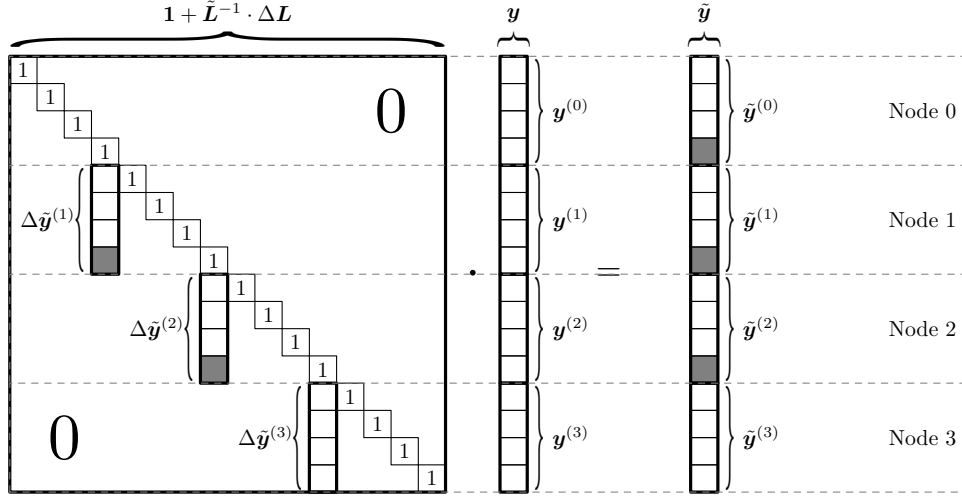
Figure 7.4: The matrix equation (7.16). All nodes calculate a part of the vector on the right hand side. The partial solutions can then be constructed locally on each of the nodes after the gray entries are broadcasted.

or

$$\tilde{\boldsymbol{L}} \cdot \Delta \boldsymbol{y}^{(k)} = l_{kn} \boldsymbol{e}_{kn}. \tag{7.19}$$

But since $\tilde{\boldsymbol{L}}$ is block diagonal, this N-dimensional equation can be reduced to n_p separate equations of dimension n (note that the blocks $\tilde{\boldsymbol{L}}^{(k)}$ all have unit determinant and are thereby invertible)

$$\tilde{\boldsymbol{L}}^{(k')} \cdot \Delta \tilde{\boldsymbol{y}}^{(k')} = \boldsymbol{0} \Leftrightarrow \Delta \tilde{\boldsymbol{y}}^{(k')} = \boldsymbol{0}, \quad k' \neq k, \tag{7.20}$$

$$\tilde{\boldsymbol{L}}^{(k)} \cdot \Delta \tilde{\boldsymbol{y}}^{(k)} = l_{kn} \boldsymbol{e}_0^{(k)}. \tag{7.21}$$

$\Delta \boldsymbol{y}^{(k)}$ is built from the separate solutions

$$\Delta \boldsymbol{y}^{(k)} = \left\{ [-\boldsymbol{0}-], [-\Delta \tilde{\boldsymbol{y}}^{(1)}-], [-\Delta \tilde{\boldsymbol{y}}^{(2)}-], \dots, [-\Delta \tilde{\boldsymbol{y}}^{(n_p-1)}-] \right\}^T$$

$$= \{ \underbrace{[-\boldsymbol{0}-], \dots, [-\boldsymbol{0}-]}_{kn}, [-\Delta \tilde{\boldsymbol{y}}^{(k)}-], \underbrace{[-\boldsymbol{0}-], \dots, [-\boldsymbol{0}-]}_{(n_p-k-1)n} \}^T. \tag{7.22}$$

In summary, the nonzero columns of $\tilde{\boldsymbol{L}} \cdot \Delta \boldsymbol{L}$ are located at the column numbers $kn-1$ and each of these columns has only nonzero elements between kn and $(k+1)n - 1$. With this result in mind, Eq. (7.16) then leads to the matrix structure as illustrated in Fig. 7.4. It is now easy to see that the intermediate solution vector fulfils

$$\boldsymbol{y}^{(0)} = \tilde{\boldsymbol{y}}^{(0)}, \tag{7.23}$$

$$\boldsymbol{y}^{(k)} = \tilde{\boldsymbol{y}}^{(k)} - \Delta \tilde{\boldsymbol{y}}^{(k)} y_{n-1}^{(k-1)}, \quad k = 1, 2, \cdots, n_p - 1. \tag{7.24}$$

The k'th node already knows the vectors $\tilde{\boldsymbol{y}}^{(k)}$ and $\Delta\tilde{\boldsymbol{y}}^{(k)}$ from Eqs. (7.14) and (7.15) and hence, it just needs the number $y_{n-1}^{(k-1)}$ in order to construct the solution. In order to obtain this missing number, we let all nodes communicate the two numbers $\tilde{y}_{n-1}^{(k)}$ and $\Delta\tilde{y}_{n-1}^{(k)}$. When these numbers are accessible to all nodes the k'th node can calculate the missing number $y_{n-1}^{(k-1)}$ as

$$
\begin{aligned}
y_{n-1}^{(0)} &= \tilde{y}_{n-1}^{(0)}, & & (7.25) \\
y_{n-1}^{(k')} &= \tilde{y}_{n-1}^{(k')} - \Delta\tilde{y}_{n-1}^{(k')}y_{n-1}^{(k'-1)}, & k' = 1, 2, \cdots, k-1. & \quad(7.26)
\end{aligned}
$$

The global ("all-to-all") communication of the numbers can in fact be reduced since the k'th node only needs the numbers $\tilde{y}_{n-1}^{(k')}$ and $\Delta\tilde{y}_{n-1}^{(k')}$ from the nodes k' where $k' < k$. The actual implementation of a global communication is, however, very simple and can be used if we are ready to accept that the amount of data being communicated will be twice as large as strictly necessary.

The second step, namely the solution of $\boldsymbol{U} \cdot \boldsymbol{x} = \boldsymbol{y}$, follows a scheme much similar to the first step. We let

$$
\boldsymbol{U} = \tilde{\boldsymbol{U}} + \Delta\boldsymbol{U}, \qquad (7.27)
$$

with $\tilde{\boldsymbol{U}}$ equal to $\boldsymbol{U}$ except that the numbers $u_{kn-1}, k = 1, 2, \cdots, n_p - 1$ are replaced by zeros. $\tilde{\boldsymbol{U}}$ is thereby block diagonal as shown in Fig. 7.5

$$
\tilde{\boldsymbol{U}} = \bigoplus_{k=0}^{n_p-1} \tilde{\boldsymbol{U}}^{(k)} = \mathrm{diag}(\tilde{\boldsymbol{U}}^{(0)}, \tilde{\boldsymbol{U}}^{(1)}, \cdots, \tilde{\boldsymbol{U}}^{(n_p-1)}). \qquad (7.28)
$$

The remainder $\Delta\boldsymbol{U} = \boldsymbol{U} - \tilde{\boldsymbol{U}}$ is

$$
\Delta\boldsymbol{U} = \sum_{k=1}^{n_p-1} u_{kn-1}\boldsymbol{e}_{kn-1} \cdot \boldsymbol{e}_{kn}^T. \qquad (7.29)
$$

By backward substitution, each node now solves the bidiagonal systems

$$
\begin{aligned}
\tilde{\boldsymbol{U}}^{(k)} \cdot \tilde{\boldsymbol{x}}^{(k)} &= \boldsymbol{y}^{(k)}, & (7.30) \\
\tilde{\boldsymbol{U}}^{(k)} \cdot \Delta\tilde{\boldsymbol{x}}^{(k)} &= u_{(k+1)n-1}\boldsymbol{e}_{n-1}^{(k)}. & (7.31)
\end{aligned}
$$

We multiply $\boldsymbol{U} \cdot \boldsymbol{x} = \boldsymbol{y}$ by $\tilde{\boldsymbol{U}}^{-1}$ from the left

$$
\tilde{\boldsymbol{U}}^{-1} \cdot \boldsymbol{U} \cdot \boldsymbol{x} = \left(1 + \tilde{\boldsymbol{U}}^{-1} \cdot \Delta\boldsymbol{U}\right) \cdot \boldsymbol{x} = \tilde{\boldsymbol{U}}^{-1} \cdot \boldsymbol{y}. \qquad (7.32)
$$

The right hand side is the solution of the block-diagonal system $\tilde{\boldsymbol{U}} \cdot \tilde{\boldsymbol{x}} = \boldsymbol{y}$ that is constructed from the separate solutions $\tilde{\boldsymbol{x}}^{(k)}$ from Eq. (7.30). The matrix structure of Eq. (7.32) is shown in Fig. 7.6. The solution vector now

Figure 7.5: The upper bidiagonal matrix U from Fig. 7.1 can be written as the sum of the block diagonal matrix $\tilde{U}$ and the remainder ΔU.

fulfils

$$
\begin{align}
\boldsymbol{x}^{(n_p-1)} &= \tilde{\boldsymbol{x}}^{(n_p-1)}, \tag{7.33}\\
\boldsymbol{x}^{(k)} &= \tilde{\boldsymbol{x}}^{(k)} - \Delta\tilde{\boldsymbol{x}}^{(k)} x_0^{(k+1)}, \quad k = n_p - 2, n_p - 3, \ldots, 0. \tag{7.34}
\end{align}
$$

i.e. the k'th node needs the number $x_0^{(k+1)}$ in order to construct the solution. Again, we let all nodes communicate two numbers, namely $\tilde{x}_0^{(k)}$ and $\Delta\tilde{x}_0^{(k)}$. When these numbers are accessible to all nodes, the k'th node can calculate the missing number $x_0^{(k+1)}$ as

$$
\begin{align}
x_0^{(n_p-1)} &= \tilde{x}_0^{(n_p-1)}, \tag{7.35}\\
x_0^{(k')} &= \tilde{x}_0^{(k')} - \Delta\tilde{x}_0^{(k')} x_0^{(k'+1)}, \quad k' = n_p - 2, n_p - 1, \ldots, k+1. \tag{7.36}
\end{align}
$$

The present algorithm is straightforwardly generalized to solution of N_{rhs} right hand sides of Eq. (7.1). In such a case, each node solves Eq. (7.14) for all the right hand side vectors and subsequently communicates all the corresponding N_{rhs} numbers $\tilde{y}_{n-1}^{(k)}$ in addition to the single number $\Delta\tilde{y}_{n-1}^{(k)}$. The solution of $U \cdot \boldsymbol{x} = \boldsymbol{y}$ can be accomplished similarly. If the communication latency between the nodes is large, it is important to note that the algorithm only requires two communications, each of $N_{\text{rhs}} + 1$ numbers.

7.2.1   Computational complexity

In this section, we consider the computational cost involved in the parallel algorithm compared with the serial algorithm. We count one operation as one scalar multiplication/division followed by one addition. The LU factorization, Eqs. (7.3)-(7.4) can be obtained in $4N$ operations, and the serial
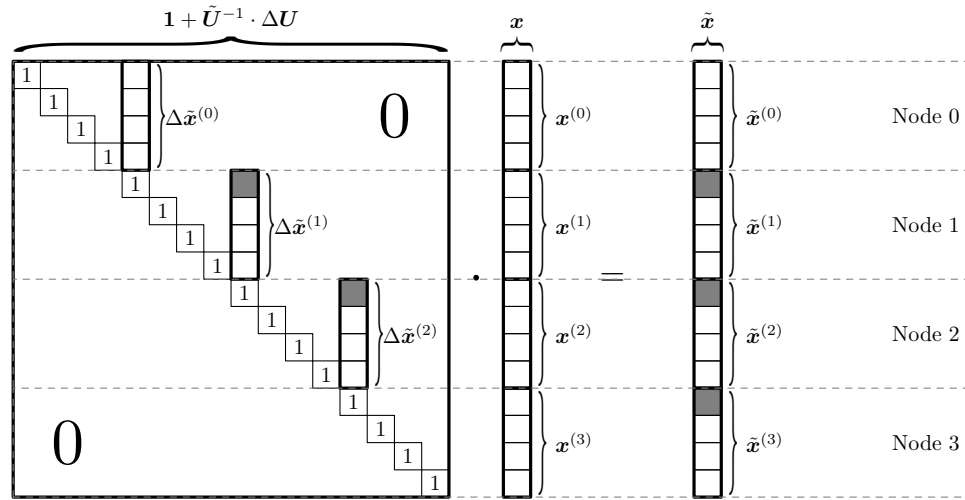
Figure 7.6: The matrix equation (7.32). All nodes calculate a part of the vector on the right hand side. The partial solutions can then be constructed locally on each of the nodes after the gray entries are broadcasted.

solution of N_{rhs} right hand sides requires $N_{\mathrm{rhs}}N$ operations for the forward substitution, Eq. (7.8), and $2N_{\mathrm{rhs}}N$ for the backward substitution (7.10), i.e., a total operational count of $4N + 3N_{\mathrm{rhs}}N$.

In the parallel algorithm, the *LU* factorization must be calculated by $4N$ operations on all nodes. Table 7.1 summarizes the number of operations involved in each step of the parallel algorithm. The total count per node is approximately $4N + (5N_{\mathrm{rhs}} + 2)N/n_p + 2N_{\mathrm{rhs}}n_p$. In addition, two global communications of $N_{\mathrm{rhs}} + 1$ numbers each are also needed.

It is now clear that the parallel algorithm is only significantly faster than the serial algorithm if $N_{\mathrm{rhs}} \gg 1$ and $N \gg n_p$. In that case, the operational count per node is approximately $5N_{\mathrm{rhs}}N/n_p$ which reflects that the local solutions of $\tilde{\boldsymbol{y}}$ and $\tilde{\boldsymbol{x}}$ for all the right hand sides are the most time consuming steps. The serial algorithm uses roughly $3N_{\mathrm{rhs}}N$ operations on a single node. For n_p nodes, the speedup employing the parallel algorithm is therefore $\sim (3/5)n_p$.

In some cases the parallel method is outperformed by the serial. Consider for example the case $N_{\mathrm{rhs}} = 1$. Each node must then perform $4N + 2n_p + 7N/n_p$ operations which is, for any choice of n_p, only slightly less than $7N$ required for the serial algorithm. The communications step furthermore slows down the parallel method. Note, however, that the serial method is only faster if the vector $\boldsymbol{b}$ is initially known entirely on a single node and the solution $\boldsymbol{x}$ is wanted only on the same single node. If, for some reason, the vector $\boldsymbol{b}$ is initially distributed on n_p nodes and the solution must be distributed likewise (cf. Fig. 7.2), the parallel method will again be faster if

	Equation	Operations
LU	(7.3)-(7.4)	$4N$
$\tilde{\boldsymbol{y}}^{(k)}$	(7.14)	$N_{\mathrm{rhs}}\frac{N}{n_p}$
$\Delta\tilde{\boldsymbol{y}}^{(k)}$	(7.15)	$\frac{N}{n_p}$
$y_{n-1}^{(k-1)}$	(7.25)-(7.26)	$< N_{\mathrm{rhs}}n_p$
$\boldsymbol{y}^{(k)}$	(7.24)	$N_{\mathrm{rhs}}\frac{N}{n_p}$
$\tilde{\boldsymbol{x}}^{(k)}$	(7.30)	$2N_{\mathrm{rhs}}\frac{N}{n_p}$
$\Delta\tilde{\boldsymbol{x}}^{(k)}$	(7.31)	$\frac{N}{n_p}$
$x_0^{(k+1)}$	(7.35)-(7.36)	$< N_{\mathrm{rhs}}n_p$
$\boldsymbol{x}^{(k)}$	(7.34)	$N_{\mathrm{rhs}}\frac{N}{n_p}$

Table 7.1: Operational counts per node involved in the parallel algorithm.

the interconnection bandwidth is low. The reason is that *all* data must initially be transferred to a single node in order to use the serial method. After the solution is obtained on this node, all the data contained in the solution must be redistributed. The total amount of data that has to be transferred is then $2N$ numbers instead of the $2n_p + 2$ numbers required for the parallel method, while the actual time spent on computational operations is roughly the same.

7.2.2 Benchmark - the one-dimensional Poisson equation

We will demonstrate that the parallel method works by considering a simple example, namely the one-dimensional Poisson equation

$$\frac{d^2}{dx^2}\Phi(x) = -\rho(x), \quad x_{\min} < x < x_{\max}, \tag{7.37}$$

where $\rho(x)$ is a known spatial charge distribution and $\Phi(x)$ is the electrostatic potential which is to be determined. We assume that the problem is subject to Dirichlet boundary conditions, so that $\Phi(x_{\min})$ and $\Phi(x_{\max})$ are known. Here we take the example $\rho = 1$, $x_{\min} = 0$, $x_{\max} = 1$, and $\Phi(x_{\min}) = \Phi(x_{\max}) = 0$. For this simple case, the Poisson equation can be integrated straightforwardly to give the solution

$$\Phi(x) = -\frac{1}{2}\left(x - \frac{1}{2}\right)^2 + \frac{1}{8}. \tag{7.38}$$

In the numerical approach, the problem is discretized on a grid as described in Sec. 3.2. The second derivative is approximated by the three-point finite difference, cf. (3.12). The electrostatic potential is then found as the solution of the tridiagonal matrix equation

$$\boldsymbol{D}_2 \cdot \boldsymbol{\Phi} = -\boldsymbol{\rho}. \tag{7.39}$$
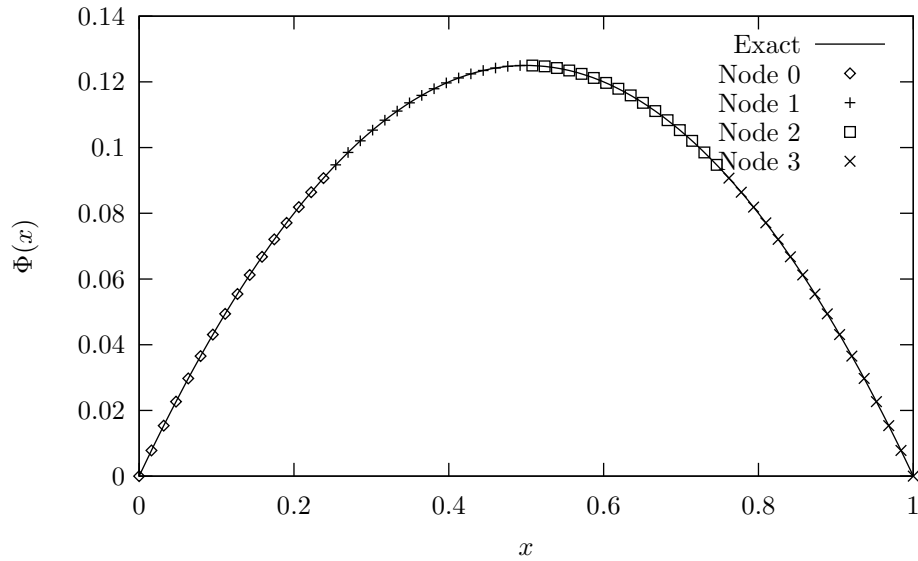
Figure 7.7: The electrostatic potential for a uniform charge distribution $\rho(x) = 1$. Both end points are grounded. The exact potential, Eq. (7.38), is indicated by the solid line. The data points mark the values obtained from a discrete calculation employing a 64-point grid with the data being distributed on four nodes as indicated.

The tridiagonal system is solved in parallel by four nodes with a 64-point grid representation. The solution is given in Fig. 7.7 which shows that the parallel method indeed produces the correct result. The data markers indicate how the solution vector is distributed among the nodes.

Now we turn to the performance of the parallel method compared with the standard serial method. We measure the performance by the speedup factor which is defined as the execution time on a single node by the serial algorithm divided by the execution time on n_p nodes by the parallel algorithm. The test is run on a 16384-dimensional tridiagonal matrix equation for 2000 right hand sides. The number of nodes is kept below 15, so that the two criteria for optimal scaling $N_{\mathrm{rhs}} \gg 1$ and $N \gg n_p$ are clearly fulfilled. We would then expect the speedup to be $\sim (3/5)n_p$.

The performance is presented in Fig. 7.8 for two different types of machines, namely a cluster of Intel Pentium and a cluster of AMD Opterons. Both clusters are connected by a gigabit network connection. The speedup measured from the Intel cluster follows very well the theoretical expectation. Amazingly, the AMD Opteron cluster is able to solve the problem even faster than the theoretical limit. One possibility for this result is that the AMD Opteron is able to use some vectorization instruction such as the Single Instruction Multiple Data stream (SIMD). A closer examination of Eqs. (7.24) and (7.34) reveals that these two steps can be vectorized. Equation (7.24)
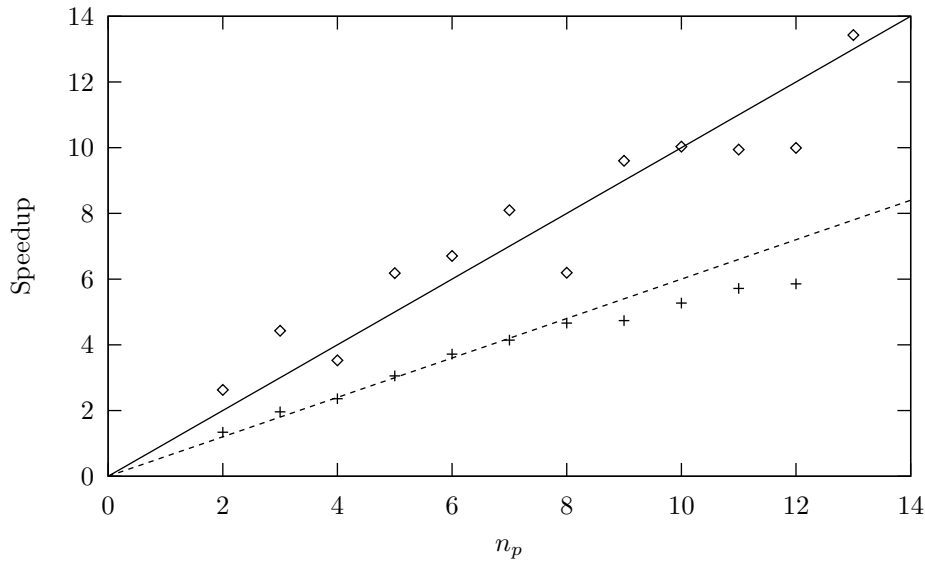
Figure 7.8: The speedup factor as a function of the number of nodes n_p for the parallel solution of 2000 right hand sides of a 16384-dimensional matrix equation. ($+$) cluster of Intel Pentium 4, 3.2 GHz and gigabit ethernet interconnection. ($\diamond$) cluster of AMD Opterons, 2.6 GHz with gigabit ethernet interconnection. The solid line indicates the linear speedup. The dashed line indicates the theoretical speedup $(3/5)n_p$.

states that each element of the vector $\boldsymbol{y}^{(k)}$ is obtained by the operation: "Multiply one element from $\Delta\boldsymbol{y}^{(k)}$ by $y_{n-1}^{(k-1)}$ and subtract the result from an element of $\tilde{\boldsymbol{y}}^{(k)}$". An ordinary scalar machine must repeat this procedure for all $n = N/n_p$ elements and, hence, we have added $2N_{\mathrm{rhs}}N/n_p$ operations to the total number of operations. The vector machine, on the other hand, can perform the *same* instruction (multiplication by $y_{n-1}^{(k-1)}$ and subtraction) on a larger set of data *simultaneously*. The time spent by execution of Eqs. (7.24) and (7.34) can thereby be reduced significantly. If these steps can be neglected entirely, the operational count per node is reduced to $3N_{\mathrm{rhs}}N/n_p$ which leads to a linear speedup $\sim 1n_p$. Note that the forward and backward substitutions, Eqs. (7.8) and (7.10) cannot benefit from vectorization since the instructions are different for the different elements.

Chapter 8

The relation between wave functions and experiments

The physical process that we wish to study in this present and in the following chapter is process of ionization induced by a strong laser field.

Before turning to the theoretical aspect, we must clarify which observables are actually measured in experiments. Therefore, we first briefly introduce the basic ideas behind a few experiments that measure strong-field ionization of atoms and molecules.

8.1   Experimental aspects of strong-field ionization

8.1.1   Ionization yield – time of flight mass spectrometer

The first experiments on strong-field ionization measured the total yield of ions or electrons produced from a gas of atoms. Such measurements use a time of flight technique as depicted in Fig. 8.1. The idea behind the time of flight method is that the ions created are accelerated in the electrostatic field between the electrodes. The positively charged ions will be accelerated to the right in the figure, if the applied voltages fulfil $V_1 > V_0$. After having escaped the extraction field through a hole or a mesh in the rightmost plate, the ions enter the field free region, moving to the right with the constant velocity gained in the field. Finally, when the ions hit the detection plate, they generate a current that can be measured. We assume that the ions are created approximately at rest somewhere within the laser focus, which is much smaller than the dimensions of the apparatus. It is then a simple matter of integrating Newton's equations of motion to find that the time when a certain ion of mass M and charge q hits the detection plate depends on the ion's mass to charge ratio as $t \propto \sqrt{M/q}$. A measurement of the time dependent current generated on the detection plate can therefore be related to the ion yield of different charge states, assuming that the mass of the
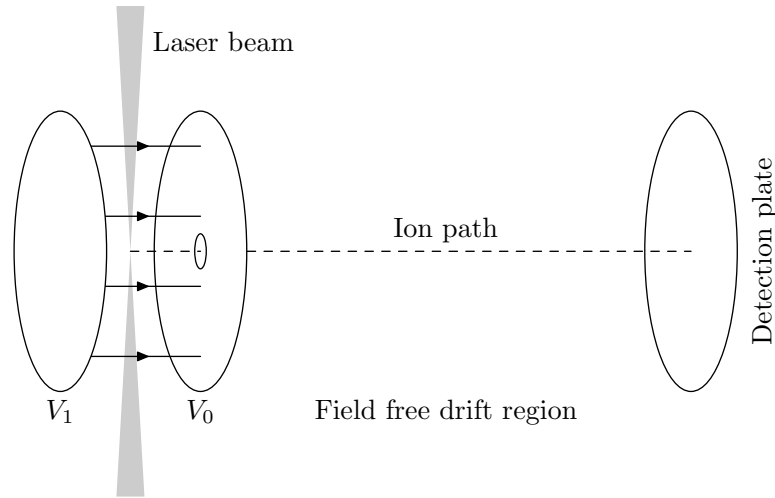
Figure 8.1: A schematic picture of a time of flight mass spectrometer.

sample gas atoms is known.

In most cases, the experiments investigate how the ion yield depends on the peak intensity of the ionizing laser. Among the phenomena of great interest observed by such an experiment is the discovery of enhanced double ionization which shows that correlated electron dynamics plays a key role in multiple ionization processes [16].

8.1.2 Above threshold ionization

Strong field ionization can occur by absorption of more photons than the minimum needed to reach the continuum threshold. Each time a photon is absorbed in the continuum, the electron gains another energy quantum from the field. The energy spectrum of the photoelectron will therefore show a series of peaks that are separated by the photon frequency. This phenomenon is referred to as above threshold ionization (ATI). In the simplest case, we might expect that an $\mathcal{N}$-photon absorption gives rise to a peak in the ATI spectrum at the position $E_\mathcal{N} = \mathcal{N}\omega - I_p$, where I_p is the ionization potential of the initial bound state and ω is the angular oscillation frequency of the field. In a very strong field, the electron in the continuum will exhibit a wiggling motion with an associated energy called the ponderomotive energy $U_p = I/(4\omega^2)$, with I being the intensity. The wiggling motion effectively rises the ionization potential such that the electron can only be freed if enough energy is absorbed from photons to overcome both the ionization potential and the ponderomotive energy. Accordingly, the $\mathcal{N}$-photon peak is shifted and occurs at the position

$$E_\mathcal{N} = \mathcal{N}\omega - I_p - U_p. \qquad (8.1)$$

Laser beam

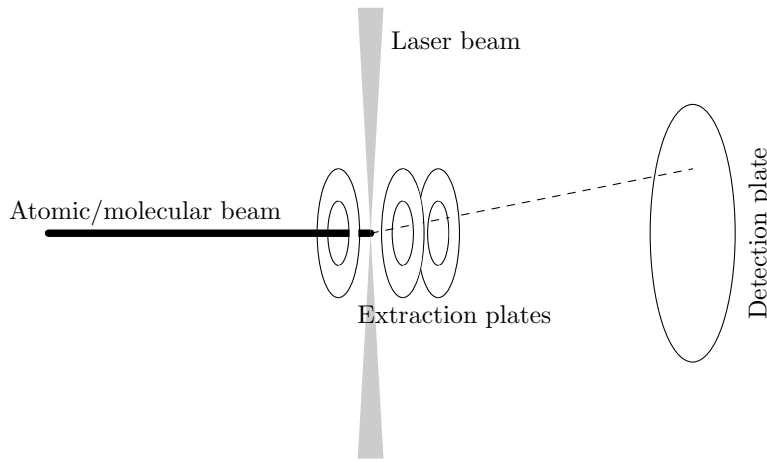Atomic/molecular beam

Extraction plates

Detection plate

Figure 8.2: A schematic picture of an experimental setup from Femtolab in Århus which uses the ion or electron imaging technique.

The number of photons must be at least $\mathcal{N}_{\min} = \lceil (I_p + U_p)/\omega \rceil$ in order to overcome the ionization threshold.

It is also possible to measure the photoelectron energy by a time of flight technique; only this time it is not necessary to apply an electrostatic extraction field. The initial velocities and energies of the electrons can be simply related to the flight time to the detector. Note that it is required that the electrons are created at a well defined initial instant of time compared to the time it takes to travel to the detector. If this is not the case, it is impossible to distinguish between slow electrons created at one time and faster electrons created at some later time. This is of course no problem if the electrons are created from pulses of femtosecond duration since flight times of, e.g., $10\,\text{eV}$ electrons over a macroscopic distance of $1\,\text{m}$ is around $0.5\,\mu\text{s}$.

Ideally, detectors should surround the interaction region so that electrons are collected from all directions. It is, however, possible to guide electrons by magnetic fields so that all electrons ejected in a hemisphere can be measured on a single plate while still maintaining the property that all electrons ejected with equal energies, but regardless of direction, arrive simultaneously [17].

8.1.3 Fully differential momentum distribution

From the late nineties and onwards, several groups have performed experiments that resolve both the direction and the energy of the outgoing electron. Such doubly differential information can be obtained in a setup as schematically depicted in Fig. 8.2. Very briefly, the experiment runs as follows. First, the sample gas is loaded into the reaction chamber. Second, the sample gas

atoms or molecules are ionized by the laser pulse. The free electrons produced are then accelerated by an electrostatic field which is created by the extraction plates held at different potentials. Finally, the electron hits a position sensitive detector which generates a light flash at the position where the electron hits. The positions of the light flashes can then be recorded by a camera and transferred to a computer for further data analysis.

In this experiment, the actual measurement is clearly the position of the electron. If the electric field which extracts the electrons is known, it is possible to relate the initial velocity or momentum of the electron to the position that is detected on the detector [18]. In Sec. 8.4.2 below, we return in more detail to the connection between initial velocity and measured image. Here we only state that the relevant quantity to calculate from a theoretical point of view is the momentum distribution of outgoing electrons.

We finally mention that the fully differential momentum distribution can also be obtained by an alternative technique, namely the cold target recoil-ion momentum spectroscopy (COLTRIMS).

8.2   Quantum mechanical theory of measurement

In the previous chapter, we let the particle be characterized by the coordinate space wave function. From the fundamental postulates of quantum mechanics, we know that the probability amplitude distribution for measuring a certain physical observable is obtained by projecting the wave function on the complete set of eigenfunction of the observable. More precisely, the possible outcome of a measurement of an observable are the eigenvalues of the operator. The probability amplitude for obtaining a particular eigenvalue from the measurement is given by the projection of the wave function on the corresponding eigenfunction. In order to calculate projections in the coordinate basis, we must be able to express the eigenfunctions of operators in this basis. For example, the coordinate representation of the energy eigenfunctions are the solutions of the time independent Schrödinger equation in the coordinate basis

$$\left(-\frac{1}{2}\boldsymbol{\nabla}^2 + V(\boldsymbol{r})\right)\Phi_E(\boldsymbol{r}) = E\Phi_E(\boldsymbol{r}). \tag{8.2}$$

Similarly, the coordinate representation of the momentum eigenfunctions are the plane waves $\phi_{\boldsymbol{k}} = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}/(2\pi)^{3/2}$ since they are the eigenfunctions of the momentum operator in the coordinate basis

$$-i\boldsymbol{\nabla}\phi_{\boldsymbol{k}}(\boldsymbol{r}) = \boldsymbol{k}\phi_{\boldsymbol{k}}(\boldsymbol{r}). \tag{8.3}$$

8.2.1   Mixed states

The theory outlined in the previous section clearly relies on the fact that the system can be described entirely by the wave function $\Psi(\boldsymbol{r}, t)$. The

probability distribution is obtained by many measurements on identically prepared systems, i.e., systems that are characterized by the same wave function. A collection of systems which are all described by the same wave function is called a pure ensemble. For example, suppose that we have a number of hydrogen atoms in the ground state at time $t = 0$. Then we expose all atoms to the same perturbation so that all atoms are described by the wave function $\Psi(\boldsymbol{r}, T)$ at some point T after the perturbation is switched off. The probability distribution after measurement of a certain one-atom observable is then simply given by projecting the state on the appropriate set of eigenstates.

In certain situations, the individual members of a collection of systems cannot be described by the same state. In that case, the ensemble is referred to as a mixed ensemble or a mixed state. If the fraction w_i of systems that comprise the ensemble is described by the wave function $\Psi^{(i)}(\boldsymbol{r}, t)$, the average value upon measurement of an observable O is now generalized to an ensemble average

$$\langle\langle O \rangle\rangle = \sum_i w_i \langle \Psi^{(i)}(\boldsymbol{r}, t) | O | \Psi^{(i)}(\boldsymbol{r}, t) \rangle. \qquad (8.4)$$

For example, the probability of measuring the momentum $\boldsymbol{k}$ from a mixed ensemble pertains to the ensemble average of the projection operator $\mathcal{P}_{\boldsymbol{k}} = |\phi_{\boldsymbol{k}}(\boldsymbol{r})\rangle\langle\phi_{\boldsymbol{k}}(\boldsymbol{r})|$. We will only consider situations where the individual members of the ensemble evolve independently. The usual situation is the following: At time $t = 0$, the distribution over the possible wave functions is known. Then an external perturbation is applied to the whole ensemble from $t = 0$ until $t = T$, and the measurement is performed at time T as an ensemble average. In that case all of the known initial wave functions $\{\Psi^{(i)}(\boldsymbol{r}, 0)\}$ evolve according to the usual time evolution due to the Hamiltonian that includes the external perturbation. At time T the individual wave functions then end up in the states $\{\Psi^{(i)}(\boldsymbol{r}, T) = \mathscr{U}(T, 0)\Psi^{(i)}(\boldsymbol{r}, 0)\}$. It is then straightforward to calculate the ensemble average at time T according to Eq. (8.4). In that way, it is sufficient to calculate the dynamics for each of the possible initial states. Then the distribution upon the desired measurement is calculated as in the pure case, utilizing the final wave functions $\Psi^{(i)}(\boldsymbol{r}, T)$. Finally, the results are averaged according to the weights.

One example which illustrates how a distribution over initial wave functions may be known is a thermally excited collection of atoms. At the temperature $\mathcal{T}$, the atoms will be excited to populate energy eigenstates according to the Boltzmann weight factor $w(E_n) = e^{-E_n/k_B \mathcal{T}}/\mathcal{Z}$, where k_B is Boltzmann's constant and $\mathcal{Z} = \sum_n e^{-E_n/k_B \mathcal{T}}$ is called the partition function. Typically, the energy difference between the ground state and the excited

states of atoms is so large that thermal excitations can be neglected[1].

8.3   Ionization

The solution of the time dependent Schrödinger equation produces a wave function which in principle contains all information about the active electron. It turns out that extracting the particular information that relates to certain experimental observables can be quite complicated. This section describes how the information that is relevant for ionization can be extracted from the wave function.

8.3.1   Fully differential photoelectron spectrum

The most complete information that can be gained from a single ionization process, is the fully differential momentum distribution of the ejected electron. We denote the probability distribution $dP/d\boldsymbol{k}$, so that $(dP/d\boldsymbol{k})\Delta\boldsymbol{k}$ is the probability of measuring the electron with the momentum vector within a three-dimensional volume element in momentum space between $\boldsymbol{k}$ and $\boldsymbol{k} + \Delta\boldsymbol{k}$.

We now assume that we have created a wave function at some time T after the end of the time dependent interaction. From this point, the time evolution is solely governed by the time-independent Hamiltonian $\mathcal{H}_0 = \boldsymbol{p}^2/2 + V(\boldsymbol{r})$. In a first naive attempt to find the momentum distribution, we can take the wave function and project it on the eigenstates of the momentum operator, i.e., the plane waves $\phi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}/(2\pi)^{3/2}$

$$\frac{dP}{d\boldsymbol{k}} \stackrel{?}{=} |\langle\phi_{\boldsymbol{k}}(\boldsymbol{r})|\Psi(\boldsymbol{r},T)\rangle|^2. \tag{8.5}$$

Note that the momentum eigenfunctions must be normalized to the three-dimensional delta function $\langle\phi_{\boldsymbol{k}'}(\boldsymbol{r})|\phi_{\boldsymbol{k}}(\boldsymbol{r})\rangle = \delta(\boldsymbol{k} - \boldsymbol{k}')$. While Eq. (8.5) surely gives some momentum distribution, it does not pertain to the actual measurement as described in Sec. 8.1.3. The point is here that we only measure on the part of the wave function which conforms with the boundary condition that it travels to the detector as time goes to infinity. Clearly, the part of the wave function that belongs to bound states does not fulfil this boundary condition. Perhaps the most obvious solution to this problem is to remove the part that belongs to the bound states by projection. Formally, if we know all bound states $\{\Phi_n|E_n < 0\}$ of the Hamiltonian, we can construct the projection on the continuum as $\mathcal{P}_C = 1 - \sum_n |\Phi_n\rangle\langle\Phi_n|$. Having only the continuum part of the wave function, we might expect that the measured

---

[1]In strong-field physics we are typically only able to resolve the gross structure of atoms. If the ground state is split in hyperfine levels, these levels *will* be populated according to the Boltzmann weight factor.

momentum distribution is obtained as

$$\frac{dP}{d\boldsymbol{k}} \stackrel{?}{=} |\langle \phi_{\boldsymbol{k}}(\boldsymbol{r})|\mathcal{P}_C \Psi(\boldsymbol{r}, T)\rangle|^2. \qquad (8.6)$$

This expression is, however, not completely correct either. One problem arises due to the fact that momentum is not a constant of motion, since $[\boldsymbol{p}, \mathcal{H}_0] = [\boldsymbol{p}, V(\boldsymbol{r})] \neq 0$. Therefore our initial task to find the momentum distribution does not even seem to be well defined; the distribution simply changes in time. The correct definition of the problem is actually: find the momentum distribution when the electron has moved so far away from the residual ion core that $V(\boldsymbol{r}) = 0$ and it can be regarded as being a free particle. We see that this formulation is in accordance with the experimental measurement. When the initial velocity distribution is related to a position on the screen, we do not specify the atomic or molecular potential from which the electron originates. We simply assume that we can regard the electron as being free. With this point in mind, we are now able to formulate the appropriate measure for the momentum distribution, namely

$$\frac{dP}{d\boldsymbol{k}} = \lim_{t\to\infty} |\langle \phi_{\boldsymbol{k}}(\boldsymbol{r})|\mathcal{P}_C \Psi(\boldsymbol{r}, T + t)\rangle|^2. \qquad (8.7)$$

In practice, the time t must be taken large enough so that slowest part of the continuum has moved out to a region in space where $V(\boldsymbol{r}) = 0$. Then, since the continuum evolves in the purely free particle Hamiltonian, momentum becomes a constant of motion, and we obtain the same distribution for any later time.

Equation (8.7) provides the correct momentum distribution but is, unfortunately, of limited use in practice in the context of in strong-field ionization. The reason is evident from Eq. (8.1). If the number of photon channels that needs to be considered is large, the electrons are ejected with a large spread in energy and momentum. This means that the continuum wave packet contains a wide range of momentum components each of which travels with a corresponding phase velocity. Accordingly, the outgoing continuum wave packet will quickly disperse and spread to a large spatial region. Hence, at the time when the slowest part that constitutes the wave packet has reached the asymptotic region where $V(\boldsymbol{r}) = 0$, the fastest part will have travelled far away. The calculation must therefore be performed in a very large basis in order to accommodate the fastest part of the continuum wave packet at all times. Of course, it would be desirable to use as small a basis as possible. In the next subsection, we describe how the momentum distribution can be obtained without a subsequent propagation. The basis then only needs to be large enough to accommodate the wave function at time T at the end of the time-dependent interaction.

Finally, note that Eq. (8.7) *can* be used if the continuum wave packet is strongly peaked around a *single* momentum which will typically be the situation encountered in one- and few-photon ionization [19]. In that case, the

whole wave packet will reach the asymptotic region nearly simultaneously with a small spread in space.

8.3.2   Momentum distribution with scattering states

In the previous section, we assumed that the time-dependent wave function is known. Therefore, it may seem surprising that differential photoelectron distributions were already studied theoretically – at least in the perturbative regime – a long time before numerical time dependent solutions became available. Such theories were based on the concepts of time independent scattering theory through the so-called differential cross section. Here, one considers, in a stationary picture, a continuous flux of incoming photons which scatter on an atom to create a continuous current of outgoing electrons. Although this picture is inadequate for short pulses and highly non-perturbative interactions, it turns out that some concepts introduced by the time independent approach are actually very useful also in the time dependent formulation.

In a time independent formulation, it is clear that the momentum distribution cannot be obtained from an expression like Eq. (8.7). Hence, such a theory must be formulated slightly differently than in Sec. 8.3.1. The trick is to find a certain continuum solution to the time-independent Schrödinger equation with the boundary condition that it behaves asymptotically as a plane wave.

In order to show the properties of the continuum solution, we first assume that the potential is short-ranged, so that $V(\boldsymbol{r}) = 0$ for $r > R$. Now we seek a solution to

$$\left(\frac{\boldsymbol{p}^2}{2} + V(\boldsymbol{r}) - E \right) \psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) = 0, \tag{8.8}$$

so that for $r \to \infty$ the only outgoing component of $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ is the plane wave $\phi_{\boldsymbol{k}}(\boldsymbol{r}) = \exp(i\boldsymbol{k} \cdot \boldsymbol{r})/(2\pi)^{3/2}$ with exactly the same energy $E = k^2/2$. We use the term scattering states for eigenstates that fulfil this boundary condition. Is is straightforward to show that a solution of the following integral equation

$$\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) = \phi_{\boldsymbol{k}}(\boldsymbol{r}) + \int_{r' < R} G_k^{(-)}(\boldsymbol{r}, \boldsymbol{r}') V(\boldsymbol{r}') \psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}') d\boldsymbol{r}', \tag{8.9}$$

which is known as the Lippmann-Schwinger equation, is also a solution to the stationary Schrödinger equation. In Eq. (8.9) $G_k^{(-)}(\boldsymbol{r}, \boldsymbol{r}')$ is a free particle Green's function which satisfies

$$\left(\boldsymbol{\nabla}_{\boldsymbol{r}}^2 + k^2 \right) G_k^{(-)}(\boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{8.10}$$

Applying $\left(\boldsymbol{\nabla}_{\boldsymbol{r}}^2 + k^2 \right)$ on both sides of Eq. (8.9) easily verifies that $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ is a solution to the Schrödinger equation with the desired energy $k^2/2$. As is

usually the case for Green's functions, Eq. (8.10) does not uniquely determine $G_k^{(-)}(\boldsymbol{r}, \boldsymbol{r}')$. On the contrary, we can exploit this underdetermination to impose the correct asymptotic boundary conditions on the solution $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ to conform with the requirements described above. Direct substitution in Eq. (8.10) shows that a valid Green's function is given by

$$G_k^{(-)}(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{(2\pi)^3} \lim_{\eta \to 0^+} \int \frac{e^{i\boldsymbol{k}' \cdot (\boldsymbol{r} - \boldsymbol{r}')}}{k^2 - k'^2 - i\eta} d\boldsymbol{k}'. \tag{8.11}$$

The integral above is straightforwardly calculated in spherical coordinates to give

$$G_k^{(-)}(\boldsymbol{r}, \boldsymbol{r}') = -\frac{1}{4\pi} \frac{e^{-ik|\boldsymbol{r} - \boldsymbol{r}'|}}{|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{8.12}$$

Thus, we can express the Lippmann-Schwinger equation (8.9) as

$$\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) = \phi_{\boldsymbol{k}}(\boldsymbol{r}) - \frac{1}{4\pi} \int_{r' < R} \frac{e^{-ik|\boldsymbol{r} - \boldsymbol{r}'|}}{|\boldsymbol{r} - \boldsymbol{r}'|} V(\boldsymbol{r}') \psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}') d\boldsymbol{r}'. \tag{8.13}$$

In order to show that this solution indeed has the desired asymptotic form, we let $r \gg R > r'$ which implies $e^{-ik|\boldsymbol{r} - \boldsymbol{r}'|} \approx e^{-ik(r - \hat{\boldsymbol{r}} \cdot \boldsymbol{r}')}$ and

$$\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) \stackrel{r \to \infty}{=} \phi_{\boldsymbol{k}}(\boldsymbol{r}) - \frac{e^{-ikr}}{r} \frac{1}{4\pi} \int_{r' < R} e^{ik\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} V(\boldsymbol{r}') \psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}') d\boldsymbol{r}'. \tag{8.14}$$

From this form it is clear that the first part is a plane wave with energy $k^2/2$ and direction $\hat{\boldsymbol{k}}$ while the second part is an incoming spherical wave which has no outgoing component that can travel to the detector. When the continuum states fulfil the asymptotic boundary condition (8.14), they are said to be normalized according to S-matrix normalization. To summarize, $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ is a continuum eigenfunction of the full Hamiltonian which satisfies the same outgoing boundary conditions as a true momentum eigenstate. Now it is clear why the scattering states are useful in perturbation theory. Together with the bound eigenstates of the unperturbed Hamiltonian, they form a complete orthogonal set [20]. In the perturbative approach we find the transition rate between eigenstates of the unperturbed Hamiltonian. In the case of photoionization, the initial state is a bound eigenstate of $\mathcal{H}_0$, typically the ground state $\Phi_{E_0}(\boldsymbol{r})$ in Eq. (8.2). Following the above discussion, it is now clear that if we take the final state as $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$, we obtain the transition rate into a state that will be detected asymptotically as a momentum eigenstate with momentum $\boldsymbol{k}$. According to the Golden Rule, the momentum probability distribution is governed by the matrix element

$$\frac{dP}{d\boldsymbol{k}} \propto |\langle \psi_{\boldsymbol{k}}^{(-)} | \boldsymbol{\epsilon} \cdot \boldsymbol{p} | \Phi_{E_0} \rangle|^2, \tag{8.15}$$

where $\boldsymbol{\epsilon}$ is the polarization vector of the incident light.

Although for simplicity we assumed that the potential is of finite range, it is possible to have scattering states with similar boundary conditions for long range potentials. Of particular interest for ionization are of course the potentials which are Coulombic in the asymptotic region.

Now we turn to the application of scattering states in the context of time dependent wave functions. As in the previous section, we assume that we have obtained the time-dependent wave function after the perturbation is turned off at time T. Being an eigenstate of the Hamiltonian immediately guarantees that $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ is orthogonal to all bound states. Hence, it is not necessary to project the bound states away from the wave function on the continuum. Furthermore, using eigenstates of the Hamiltonian ensures that the projections do not change in time

$$
\begin{aligned}
\frac{dP}{d\boldsymbol{k}} &= \lim_{t\to\infty} |\langle \psi_{\boldsymbol{k}}^{(-)} | \Psi(T+t)\rangle|^2 = \lim_{t\to\infty} |\langle \psi_{\boldsymbol{k}}^{(-)} | e^{-i\mathcal{H}_0 t}\Psi(T)\rangle|^2 \\
&= \lim_{t\to\infty} |\langle e^{i\mathcal{H}_0^\dagger t}\psi_{\boldsymbol{k}}^{(-)} | \Psi(T)\rangle|^2 = \lim_{t\to\infty} |e^{-i(k^2/2)t}\langle \psi_{\boldsymbol{k}}^{(-)} | \Psi(T)\rangle|^2 \\
&= |\langle \psi_{\boldsymbol{k}}^{(-)} | \Psi(T)\rangle|^2.
\end{aligned}
\tag{8.16}
$$

In the above equation we used the time evolution operator, Eq. (4.10), for the time-independent Hamiltonian. Compared to the corresponding expression Eq. (8.7), we see that we do not need to find the time-dependent wave function for infinite times. We simply obtain the momentum distribution from the wave function immediately after the time dependent interaction is switched off. Before getting too enthusiastic over this apparent gain, we must stress that Eq. (8.13) does not provide an explicit recipe for how to construct the scattering states. Equation (8.13) is an implicit equation that $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ has to fulfil. For simple standard problems such as the pure Coulomb potential, the scattering states are known analytically, but for more complicated potentials, it can be an extremely challenging task to find the scattering states.

8.3.3   Obtaining integral properties

As stated in Sec. 8.3.1, the fully differential momentum distribution contains the complete information that can be extracted from an experiment. Thus, it is possible to obtain quantities such as the total ionization probability and the photoelectron spectrum from the fully differential distribution. It turns out that there often exist simpler methods to obtain such integral quantities than first calculating the differential probability. For example, it is possible to calculate total ionization probabilities from numerical grid calculations by applying absorbing boundaries. Similarly, the photoelectron spectra can be obtained by projection on energy eigenstates supported by the grid. These methods will be further explained in chapter 9 for the specific implementation suited for the grid method. Here we simply demonstrate that the

probability distribution over any observable quantity in an ionization experiment can be derived from the fully differential momentum distribution.

Total ionization

Clearly, it is possible to obtain the total ionization probability from the differential probability as the three-dimensional integral over the whole momentum space

$$P_{\text{ion}} = \int \frac{dP}{d\boldsymbol{k}} d\boldsymbol{k}. \tag{8.17}$$

The total ionization probability can also be calculated as the norm of the wave function being in the continuum

$$P_{\text{ion}} = \langle \mathcal{P}_C \Psi(T) | \mathcal{P}_C \Psi(T) \rangle = 1 - \sum_{n, E_n < 0} |\langle \Phi_n | \Psi(T) \rangle|^2. \tag{8.18}$$

This latter approach requires that the set of bound states $\{\Phi_n\}$ of $\mathcal{H}_0$ is available.

Photoelectron spectrum

The photoelectron spectrum is the energy differential ionization probability, i.e., we define dP/dE so that the probability of measuring the electron with energy between E and $E + \Delta E$ is $(dP/dE)\Delta E$. Using the relation $E = k^2/2$, we readily obtain

$$
\begin{aligned}
\frac{dP}{dE} &= \left. \frac{dP}{dk} \right|_{k=\sqrt{2E}} \frac{dk}{dE} = \int_{4\pi} \left(k^2 \frac{dP}{d\boldsymbol{k}} \right) \Bigg|_{k=\sqrt{2E}} d\Omega_{\boldsymbol{k}} \frac{1}{\sqrt{2E}} \\
&= \int_{4\pi} \left. \frac{dP}{d\boldsymbol{k}} \right|_{k=\sqrt{2E}} d\Omega_{\boldsymbol{k}} \sqrt{2E}.
\end{aligned} \tag{8.19}
$$

We might think that this result can be expressed equivalently as a projection of the final wave function on a continuum energy eigenstate

$$\frac{dP}{dE} \stackrel{?}{=} |\langle \Phi_E | \Psi(T) \rangle|^2, \quad E > 0, \tag{8.20}$$

where Φ_E is normalized to the energy delta function. This relation is, however, only true in one dimension, where the continuum is referred to as a *single-channel* continuum. In higher dimensions, the continuum is degenerate so that there exists a whole set of orthogonal energy eigenstates with the same positive energy. In that case, the continuum is called a *multi-channel* continuum. Obviously, Eq. (8.20) cannot be true for the multi-channel case since it does not state *which* (or which linear combination) of the degenerate eigenfunctions that should be used. As we shall see in Sec. 9.1, a multi-channel continuum can be decomposed in a sum of single-channel continua for a spherically symmetric potential. The general treatment of a multi-channel continuum is, however, often very complex.

Angular distribution

The angular distribution $dP/d\Omega_{\boldsymbol{k}}$ reflects the angular pattern in which the electron is emitted. If a detector is located at the angles $\Omega_{\boldsymbol{k}} = (\vartheta_{\boldsymbol{k}}, \varphi_{\boldsymbol{k}})$ with an aperture covering the solid angle $\Delta\Omega_{\boldsymbol{k}}$, the probability that the electron will hit the detector is $(dP/d\Omega_{\boldsymbol{k}})\Delta\Omega_{\boldsymbol{k}}$. The angular distribution is obviously obtained by integrating the differential distribution over the magnitude of the momentum

$$\frac{dP}{d\Omega_{\boldsymbol{k}}} = \int_0^\infty \left.\frac{dP}{d\boldsymbol{k}}\right|_{\Omega_{\boldsymbol{k}}} k^2 dk. \tag{8.21}$$

8.4   Simulation of an experiment

In Sec. 8.3, we explained in detail how to obtain the electron distribution from a collection of identically prepared systems which are all exposed to the same perturbation. Such conditions are unfortunately not easily realized in experiments. In order to make a fair comparison between theory and experiment we must take the true experimental conditions into account. The purpose of the present section is to highlight some aspects of a typical experimental set-up which must be considered in a quantitative comparison with a theoretical calculation.

8.4.1   Focal volume effects

In this subsection we describe why it is difficult in the strong-field regime to expose a collection of systems to the *same* interaction.

In order to produce high intensities from a laser it is necessary to apply lenses to focus the laser beam. The most common type of laser beam is the Gaussian beam where the intensity decreases as a Gaussian with the distance from the center beam axis. The spatial intensity profile of a Gaussian beam focused to the peak intensity I_0 at the origin is given in cylindrical coordinates as

$$I(\rho, z) = \frac{I_0}{1 + z^2/z_0^2} \exp\left(-\frac{2\rho^2}{w_0^2(1 + z^2/z_0^2)}\right), \tag{8.22}$$

where w_0 is called the spot size and z_0 is called the Rayleigh range. As the name indicates, the spot size measures the radial size of the beam (in a $1/e^2$ sense) at the most focused spot. The Rayleigh range is a measure for the divergence in the z direction. As indicated in Fig. 8.3, the laser intensity varies significantly in the beam. Accordingly, atoms or molecules that are located in different regions in the laser focus will be exposed to different fields. Suppose we want to predict the outcome of a measurement of some observable when the peak laser intensity is I_0. A proper comparison now requires that we have calculated the distribution of that observable
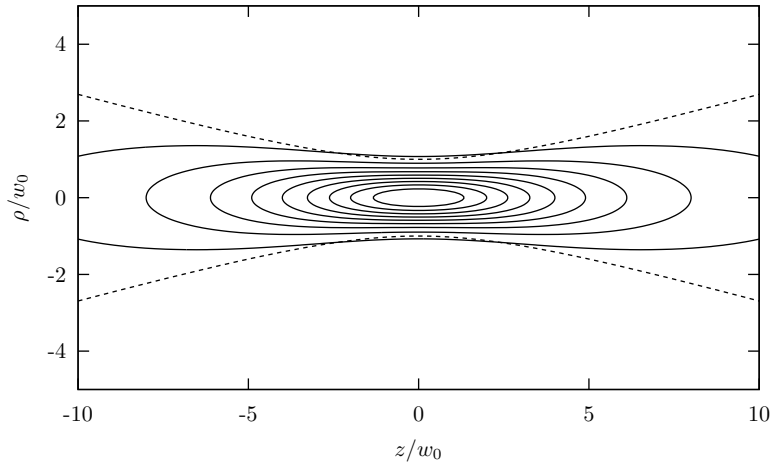
Figure 8.3: Contour plot of the intensity distribution of a Gaussian beam. The dashed lines indicate the z dependent spot size, i.e., radius in which the intensity has dropped by a factor $1/e^2$ from its value on the axis. The Rayleigh range is $z_0 = 4w_0$.

for a number of different intensities up to the peak intensity. Then we must weigh the results for each different intensity by the number of atoms or molecules being exposed to that intensity. The weighing procedure is generally straightforward, however, the details of the experimental set-up must be taken into account. The two most common types of experiments either collect particles from all space or only from a small region near the beam waist. In the first case, the aperture of the measurement apparatus is much larger than the region in the laser focus. The signal is then

$$\text{Signal} \propto 2\pi \int_0^\infty \int_{-\infty}^\infty P(I(\rho, z)) dz \rho d\rho. \tag{8.23}$$

The proportionality factor accounts for the unknown factors such as overall target density and detection efficiency. Conversely, in the second case we want a measurement device with a small aperture that only allows particles in the beam waist region to be collected. The measured signal is then

$$\text{Signal} \propto 2\pi \Delta z \int_0^\infty P(I(\rho, z = 0)) \rho d\rho, \tag{8.24}$$

when the measurement apparatus collects particles from the axial length Δz. When collecting particles from the beam waist region, it can be helpful to choose a lens set-up that produces a very long Rayleigh length and beam waist. An alternative approach which gives a similar result is to inject the sample gas can be in a narrow particle beam line that crosses the laser beam perpendicularly at the beam waist. This is the situation which is indicated in Fig. 8.2.
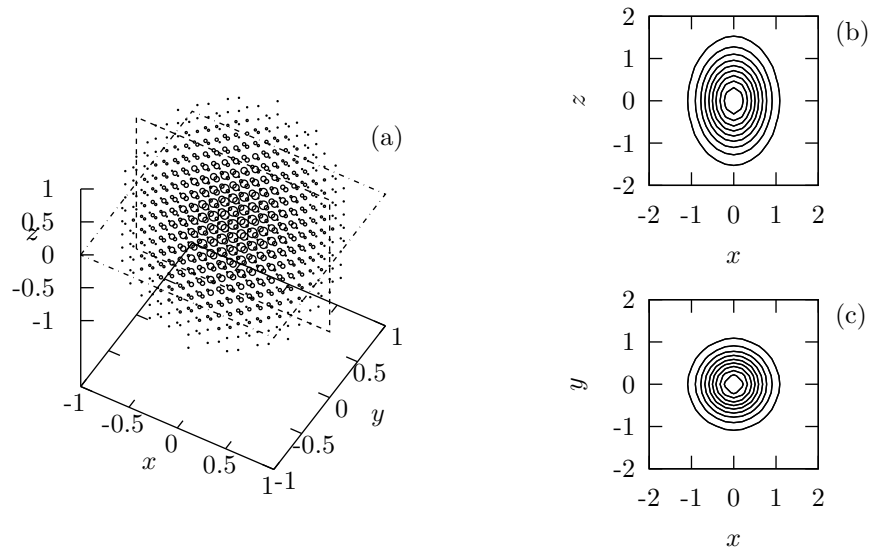
Figure 8.4: (a) A graphical representation of the anisotropic Gaussian function $f(x, y, z) = \exp[-2(x^2 + y^2) - z^2]$. The size of a data point indicates the magnitude of the function. (b) Contour plot of the function in the xz plane with the level curves separated by 0.1. (c) Contour plot of the function in the xy plane.

The typical effect that is observed when averaging over the intensity profile is some smearing out of distinct features of the spectra. This effect is often not desirable. If, for example a certain feature is predicted to occur in a spectrum for a very specific intensity, that feature may be lost in the measured signal since only a minor part of the gas is exposed to the correct intensity.

8.4.2   Abel transformation

It is in principle possible to measure and calculate the photoelectron distribution as a function of all three momentum components. One practical complication occurs, of course, when such a full distribution is to be presented graphically. The function of three variables corresponds graphically to a hypersurface in four dimensions. In order to simplify the visual presentation, a series of planar cuts can be made through the distribution. Figure 8.4 shows an example of planar cuts. We have made an attempt to visualize a function of three variables in panel (a). Clearly, it is quite difficult to read anything quantitatively from such a representation. The planar cuts in panels (b) and (c) generally do not provide the full information of the function but are much easier to read.

As an alternative visual data representation, we can project the full three-dimensional distribution on a two dimensional plane. From a mathe-

matical point of view, the projection operation is an integral transformation known as the Abel transformation. We recognize that this technique is used in, e.g., the ion imaging experiment shown in Fig. 8.2. Theoretically, we want to construct a projection in the same sense, i.e., we assume that a three-dimensional function is nonzero in a region near the origin and falls off at large distances. A projection of the function $f^{(3D)}(x, y, z)$ on a plane parallel to the xy-plane located at $z \to \infty$ is then

$$f^{(2D)}(x, y) = \int_{-\infty}^{\infty} f^{(3D)}(x, y, z)dz, \qquad (8.25)$$

and similarly for projections on planes parallel to the xz- and yz-planes. As is the case for the planar cuts, the projections will generally contain less information than the full three-dimensional function. Even simpler visualization and less information is provided by the projection on a line, e.g.,

$$f^{(1D)}(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^{(3D)}(x, y, z)dxdy. \qquad (8.26)$$

The Abel transformation is particularly interesting if a function is rotationally symmetric around the z-axis. In that case, the projection on the xz-plane *does* in fact contain the full information about the function. The rotationally symmetric function can be expressed in cylindrical coordinates as $f^{(\mathrm{cyl})}(\rho, z)$. Such a function is projected on the xz-plane as

$$f^{(2D)}(x, z) = \int_{-\infty}^{\infty} f^{(\mathrm{cyl})}(\rho(x, y), z)dy = 2 \int_{x}^{\infty} \frac{f^{(\mathrm{cyl})}(\rho, z)}{\sqrt{\rho^2 - x^2}}\rho d\rho. \qquad (8.27)$$

On the other hand, it can be proven that $f^{(\mathrm{cyl})}(\rho, z)$ can be reconstructed from the projection by an inverse Abel transformation

$$f^{(\mathrm{cyl})}(\rho, z) = -\frac{1}{\pi} \int_{\rho}^{\infty} \left.\frac{\partial f^{(2D)}(x, z)}{\partial x}\right|_{z} \frac{1}{\sqrt{x^2 - \rho^2}}dx. \qquad (8.28)$$

Since $f^{(\mathrm{cyl})}(\rho, z)$ and $f^{(2D)}(x, z)$ are an inverse pair we must conclude that $f^{(\mathrm{cyl})}(\rho, z)$ is uniquely determined from $f^{(2D)}(x, z)$. In other words: Knowledge of just the two-dimensional projection of a cylindrically symmetric distribution does in fact provide the whole information about the three-dimensional distribution. In practice, a good inverse Abel transformation requires the projection to be a smooth function. The reason is that the derivative in Eq. (8.28) can be very noisy if the projection is not sufficiently smooth. This can, in particular, be a problem if the projection is the result of an experimental measurement such as depicted in Fig. 8.2.

Chapter 9

Numerical results

We are now ready to study strong-field ionization from a theoretical point of view. First, we solve the time-dependent Schrödinger equation by the method outlined in chapter 5, and subsequently we extract the relevant information from the final solution as described in chapter 8.

The two main numerical results presented in this chapter are 1) the momentum distribution obtained after strong-field ionization of noble gas atoms (Sec. 9.2) and 2) the alignment dependent ionization of H_2^+ (Sec. 9.3). However, before we can calculate the momentum distribution we need an explicit recipe of how to obtain the scattering states – a problem which is to be studied in Sec. 9.1.

9.1 Partial wave expansion of scattering states

The time-dependent wave function is solved in a partial wave expansion, cf. Eqs. (5.4) and (5.11)

$$\Psi(\boldsymbol{r}, t) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{f_{lm}(r,t)}{r} Y_{lm}(\theta, \phi). \qquad (9.1)$$

Thus, it will be convenient if the scattering states $\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r})$ are also expanded in partial waves. Finding the scattering states is formulated in Eq. (8.8) with the appropriate boundary conditions. In this section, we describe how the scattering states can be expressed in a partial wave expansion. For simplicity, we restrict the discussion to the case of a spherically symmetric potential. If the potential is not spherically symmetric, a partial wave expansion can still be found, although the procedure becomes somewhat more complicated than shown below. For example, the partial wave expansion of the scattering states can be found for H_2^+ by the method proposed in Ref. [21].

9.1.1 Partial wave expansion for a spherically symmetric potential

As a starting point, we use the usual separation of the time-independent Schrödinger equation for a spherically symmetric Hamiltonian. We write the solutions for each partial wave at the positive energy E as

$$\psi_{Elm}(\boldsymbol{r}) = \frac{\mathscr{F}_{El}(r)}{r} Y_{lm}(\theta, \phi), \tag{9.2}$$

where $\mathscr{F}_{El}(r)$ is the solution to the radial equation

$$\left(-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V(r) \right) \mathscr{F}_{El}(r) = E\mathscr{F}_{El}(r). \tag{9.3}$$

$\mathscr{F}_{El}(r)$ is usually taken to be purely real. If $V(r)$ is the pure Coulomb potential $V(r) = -Z/r$, the radial eigenfunctions are known analytically as the regular Coulomb wave functions $\mathscr{F}_l^C(\eta, kr)$ which are normalized to match the asymptotic form

$$\mathscr{F}_l^C(\eta, kr) \stackrel{r\to\infty}{=} \sqrt{\frac{2}{\pi k}} \sin(kr - \eta \ln(2kr) - l\pi/2 + \sigma_l(k)), \tag{9.4}$$

where $k = \sqrt{2E}$, $\eta = -Z/k$ and $\sigma_l(k) = \arg[\Gamma(l+1-iZ/k)]$ are called the Coulomb parameter and Coulomb phase shift, respectively. The normalization factor in front of the sine in Eq. (9.4) ensures that ψ_{Elm} is normalized on the energy scale as

$$\langle \psi_{E'l'm'} | \psi_{Elm} \rangle = \delta(E - E')\delta_{ll'}\delta_{mm'}. \tag{9.5}$$

Energy normalization is of course useful in studies of energy differential properties such as the photoelectron spectrum. Momentum differential properties require momentum normalization

$$\langle \psi_{k'l'm'} | \psi_{klm} \rangle = \delta(k - k')\delta_{ll'}\delta_{mm'}. \tag{9.6}$$

The momentum normalized radial function is related to the energy normalized as $F_{kl}(r) = \sqrt{k}\mathscr{F}_{El}(r)$ which asymptotically approaches the sine function with the pre-factor $\sqrt{2/\pi}$, regardless of the energy. If the potential is Coulombic at asymptotic distances, $V(r) \sim -Z/r$, but contains an additional short-range potential, it is not possible to solve Eq. (9.3) analytically. In that case, a *numerical* solution is again taken purely real and often normalized as

$$\mathscr{F}_{El}^{(K)}(r) \stackrel{r\to\infty}{=} \mathscr{F}_l^C(\eta, kr) + K_l(k)\mathscr{G}_l^C(\eta, kr), \tag{9.7}$$

where $\mathscr{G}_l^C(\eta, kr)$ is an irregular energy normalized Coulomb wave function that behaves as Eq. (9.4) except that the sine is replaced by a cosine. The Coulomb parameter $\eta = -Z/k$ is now defined from the asymptotic charge.

The real number $K_l(k)$ is called the K-matrix element. Note that the K-matrix normalization in Eq. (9.7) does not fulfil Eq. (9.5) except when $K_l = 0$. Energy normalization can be obtained by multiplication of Eq. (9.7) by $(1 + K_l^2)^{-1/2}$ which leads to the asymptotic form

$$\mathscr{F}_{El}(r) \overset{r \to \infty}{=} \sqrt{\frac{2}{\pi k}} \sin(kr - \eta \ln(2kr) - l\pi/2 + \sigma_l(k) + \delta_l(k)), \qquad (9.8)$$

with the non-Coulombic phase shift $\delta_l(k) = \arctan[K_l(k)]$ being equal to zero in the Coulomb case. It is clear that none of the eigenfunctions Eq. (9.2) fulfils the correct S-matrix normalization [cf. Eq. (8.14)]. However, for a given energy E the partial wave solutions ψ_{Elm} are degenerate for the different values of l and m. Accordingly, any linear combination of the ψ_{Elm}'s will be an eigenstate of the Hamiltonian with the energy E. Now we seek the particular linear combination

$$\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{\boldsymbol{k},lm}^{(-)} \frac{\mathscr{F}_{El}(r)}{r} Y_{lm}(\theta, \phi), \qquad (9.9)$$

where the complex amplitudes $c_{\boldsymbol{k},lm}^{(-)}$ have to be chosen so that $\psi_{\boldsymbol{k}}^{(-)}$ fulfils the correct asymptotic boundary condition. The appropriate linear combination is found by comparison of Eq. (9.9) with the partial wave expansion of the (Coulomb distorted) plane wave plus an incoming spherical wave, i.e. the Coulomb generalization of Eq. (8.14). This latter expansion can be written as [22]

$$\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} e^{i(l\pi/2-\sigma_l)} \frac{\mathscr{F}_{El}^{(S)}(r)}{\sqrt{k}r} Y_{lm}^{\star}(\vartheta_{\boldsymbol{k}}, \varphi_{\boldsymbol{k}}) Y_{lm}(\theta, \phi), \qquad (9.10)$$

where the S-matrix normalized radial functions are complex and obey the asymptotic boundary condition

$$\mathscr{F}_{El}^{(S)}(r) \overset{r \to \infty}{=} \sqrt{\frac{2}{\pi k}} \frac{1}{2i} \left(e^{i[kr-\eta \ln(2kr)-l\pi/2+\sigma_l]} - S_l^{\star} e^{-i[kr-\eta \ln(2kr)-l\pi/2+\sigma_l]} \right).$$
$$(9.11)$$

Comparison of Eqs. (9.8) and (9.9) with Eqs. (9.10) and (9.11) reveals that $S_l(k) = e^{2i\delta_l(k)}$ and $c_{\boldsymbol{k},lm}^{(-)} = e^{i(l\pi/2-\sigma_l)} e^{-i\delta_l} Y_{lm}(\vartheta_{\boldsymbol{k}}, \varphi_{\boldsymbol{k}})/\sqrt{k}$. In conclusion, if for each value of l we find a real numerical solution to the radial equation and normalize it according to (9.8), then we construct the complex scattering state which obeys incoming spherical wave boundary condition as

$$\psi_{\boldsymbol{k}}^{(-)}(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} e^{i(l\pi/2-\sigma_l(k)-\delta_l(k))} \frac{F_{kl}(r)}{kr} Y_{lm}^{\star}(\vartheta_{\boldsymbol{k}}, \varphi_{\boldsymbol{k}}) Y_{lm}(\theta, \phi). \qquad (9.12)$$

Note that $F_{kl}(r)$ must be normalized on the momentum scale. The scattering states are thereby normalized to the three-dimensional momentum delta function $\langle \psi_{\mathbf{k}'}^{(-)} | \psi_{\mathbf{k}}^{(-)} \rangle = \delta(\mathbf{k} - \mathbf{k}')$. At time T, after the time dependent interaction, the momentum distribution can now be obtained from a wave function on the form (9.1) as

$$
\begin{aligned}
\frac{dP}{d\mathbf{k}} &= |\langle \psi_{\mathbf{k}}^{(-)}(\mathbf{r}) | \Psi(\mathbf{r}, T) \rangle|^2 \\
&= \left| \sum_{l=0}^{\infty} \sum_{m=-l}^{l} e^{-i(l\pi/2 - \sigma_l(k) - \delta_l(k))} \int_0^{\infty} \frac{F_{kl}(r)}{k} f_{lm}(r, T) dr\, Y_{lm}(\vartheta_{\mathbf{k}}, \varphi_{\mathbf{k}}) \right|^2 .
\end{aligned}
$$
(9.13)

If the radial expansion functions are known on a grid, the integral must be evaluated numerically from the values on the grid points.

The full momentum distribution immediately gives the energy distribution, cf. (8.19)

$$
\begin{aligned}
\frac{dP}{dE} &= \sqrt{2E} \int_0^{2\pi} \int_0^{\pi} \left. \frac{dP}{d\mathbf{k}} \right|_{k=\sqrt{2E}} \sin \vartheta_{\mathbf{k}} d\vartheta_{\mathbf{k}} d\varphi_{\mathbf{k}} \\
&= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left| \int_0^{\infty} \mathscr{F}_{El}(r) f_{lm}(r, T) dr \right|^2 ,
\end{aligned}
$$
(9.14)

where we used the orthogonality of spherical harmonics and that $\mathscr{F}_{El}(r)$ is real and energy normalized according to Eq. (9.8). Equation (9.14) can be interpreted quite simply. For example, we can imagine an experiment that measures simultaneously the magnitude and z-projection of the angular momentum as well as the energy of an electron in the continuum. Note that all three operators are compatible so that a simultaneous measurement is not forbidden by fundamental laws. One of the individual terms in Eq. (9.14) then gives the theoretical energy distribution for one of the possible results of the angular momentum measurement. Such a partial wave decomposition of an energy spectrum can often be helpful for the interpretation of few-photon ionization spectra, e.g., one-photon ionization from an s-state must result in a p-state in the continuum, two-photon ionization can populate the s and d continua, and so on.

9.1.2   Continuum in a spherical box

In Sec. 8.3, we discussed the theoretical aspects of the continuum part of the wave function. In that section, we assumed on one hand that the outgoing electron wave packet is free to evolve infinitely in space. On the other hand, the numerical solution of the time dependent Schrödinger equation, as described in chapter 5, is obtained within a spherical "box" of radius $r_{\max}$.

We argued that enclosing the system cannot influence the physical state as long as the box radius is chosen large enough so that the wave function is confined within the box boundary. It will indeed be so that the wave function obtained with two different box sizes are equal if both boxes are large enough.

Although the wave function is independent of box size, we encounter some subtleties if we express continuum projection operators in a box. If the projection operators are known analytically, such as the projections on the plane waves or the scattering states of the Coulomb problem, it is straightforward to project on these functions for any momentum. The problems arise if the projection operators are not known analytically, e.g., if the projections are to be performed on the eigenfunctions of a Hamiltonian that is not purely Coulombic. In that case, we may expect that a consistent treatment of the box requires that everything, i.e., operators and eigenfunctions must be represented within the same box as the wave function. If we solve Eq. (8.2) by diagonalization of the matrix representation of $\mathcal{H}_0$ in a basis that fulfils box boundary conditions, we only get the eigenstates that vanish at the boundary. In this way, the positive energy part of the Hamiltonian spectrum becomes discretized. This means that we cannot calculate the scattering states at any given energy or momentum that we desire. Instead, the size of the box dictates the energies at which the eigenstates are available. Furthermore, the eigenenergies that are found from Eq. (9.3) with box boundary conditions will be different for the various partial waves. This means that the degeneracy of the partial waves in the continuum is lifted and one discrete positive energy spectrum exists for the s-waves, another for the p-waves, and so on. Since the scattering states are constructed as a superposition of degenerate partial waves, they cannot be obtained from the non-degenerate box solutions of Eq. (9.3). One way to find the scattering states is to omit the boundary condition and thereby re-establish degeneracy [23]. With such an approach the desired solution of the radial equation (9.3) cannot be formulated as an eigenvalue problem and special techniques need to be applied to solve the problem.

Fortunately, it is often not necessary to construct the scattering states in order to calculate the momentum distribution from a numerically known wave function. The reason is the following. Suppose that the radial equation has been solved for each value of l with box boundary conditions. This leads to a discrete set of energy eigenvalues for both positive and negative energies. To the set of positive energies we can associate a set of discrete momenta $\{k_n^{(l)} | n = 0, 1, 2, \cdots\}$ and a corresponding set of momentum normalized real eigenfunctions $\{F_{k_n^{(l)} l}(r) | n = 0, 1, 2, \cdots\}$. For any (l, m), the following sequence of complex numbers

$$\left\{ e^{-i(l\pi/2 - \sigma_l(k_n^{(l)}) - \delta_l(k_n^{(l)}))} \int_0^\infty \frac{F_{k_n^{(l)} l}(r)}{k_n^{(l)}} f_{lm}(r, T) dr \,\middle|\, n = 0, 1, 2, \cdots \right\}, \quad (9.15)$$

will vary smoothly with the index n. By interpolation, it is then possible to find the corresponding number for any continuous value of k in between the sampled values. The momentum distribution can then be found according to Eq. (9.13) for any momentum. A similar interpolation can be applied for the energy distribution (9.14).

A final problem that arises due to the box is the normalization of the discrete continuum states. The box eigenstates are usually normalized according to the Kronecker delta for the discrete index n

$$\int_0^{r_{\max}} \mathscr{F}_{E_{n'}l}^{(\text{box})}(r) \mathscr{F}_{E_{n}l}^{(\text{box})}(r) dr = \delta_{n'n}. \tag{9.16}$$

These box eigenstates must now be properly normalized on the momentum scale. First, we perform a linear least squares fit by the parameters A and B to the form

$$\mathscr{F}_{E_{n}l}^{(\text{box})}(r) \stackrel{r \to \infty}{=} A\mathscr{F}_l^C(\eta, k_n r) + B\mathscr{G}_l^C(\eta, k_n r), \tag{9.17}$$

in the asymptotic region. K-matrix normalization (9.6) is then obtained as $\mathscr{F}_{E_{n}l}^{(K)}(r) = A^{-1}\mathscr{F}_{E_{n}l}^{(\text{box})}(r)$ and $\tan[\delta_l(k_n)] = K_l(k_n) = B/A$. Energy and momentum normalizations are related to the K-normalization as discussed in Sec. 9.1.1.

9.2 Strong-field ionization of noble gas atoms

Experiments like the one described in Sec. 8.1.3 that measure the full momentum distributions of the continuum electron have only been demonstrated recently. So far, the systems being studied have mainly been the noble gas atoms [24–26]. These new experimental results have motivated a number of theoretical studies of the same subject [27–29].

Noble gas atoms beyond helium are multi-electron systems and the exact many particle dynamics cannot be solved in a foreseeable future. In a single-active electron model we only consider the dynamics of the most loosely bound electron as we regard the inner core electrons as being strongly attached to the nucleus and not affected by the laser field. These remaining electrons are then accounted for by a static effective potential [30] which is optimized to reproduce the ground state as closely as possible. In particular, the one-electron ionization energy in the effective potential must be equal to the true ionization potential of the real atom. It is, however, not guaranteed that the atomic structure of singly excited states are well reproduced by the effective potential. If intermediate resonant states play an important role in the process under study, the effective potential model may fail.

The present numerical results try to simulate the experimental conditions in Ref. [26] as closely as possible. The systems being studied in that paper are the argon and neon atoms exposed to a strong linearly polarized field.

The quantity of interest is the one-dimensional projection of the momentum distribution onto the axis parallel to the polarization. With the laser polarized along the z-axis, the momentum component k_z is referred to as the longitudinal momentum. The laser intensity is varied between $10^{14}\,\mathrm{W/cm^2}$ and $10^{15}\,\mathrm{W/cm^2}$, and the wavelengths $400\,\mathrm{nm}$ and $800\,\mathrm{nm}$ are used. The pulse duration is around $30-45\,\mathrm{fs}$. The very brief summary of the results of Ref. [26] is that the longitudinal momentum distribution for argon at $800\,\mathrm{nm}$ is rather smoothly peaked around zero momentum. At $400\,\mathrm{nm}$ on the other hand, the momentum distribution has a minimum at zero momentum, and the distribution is dominated by peaks around $k_z = \pm(0.3-0.4)\,\mathrm{a.u.}$ For neon, the situation is almost reversed. In that case, the momentum distribution shows a local minimum at $k_z = 0\,\mathrm{a.u.}$ at $800\,\mathrm{nm}$, and two maxima around $k_z = \pm 0.1\,\mathrm{a.u.}$ At $400\,\mathrm{nm}$, at least five peaks of almost equal height are seen. One of these peaks is located at zero longitudinal momentum. All these findings are nearly independent of the intensity.

The theoretical results are obtained as follows. First we solve the time dependent Schrödinger equation with the cylindrically symmetric velocity gauge Hamiltonian. For that purpose, we have used a 16384-point radial grid that extends to $1200\,\mathrm{a.u.}$ Up to 32 angular momentum states are used to represent the angular variable. Furthermore, a time step size of $\Delta t = 6.5 \times 10^{-3}\,\mathrm{a.u.}$ is used. The pulse duration is $29\,\mathrm{fs}$. We calculate the full momentum distribution by the partial wave expansion of the scattering states as described in the previous section. The full three-dimensional distribution is then projected on the z-axis, cf. Eq. (8.26). Finally, in order to compare correctly with experiments we make the focal volume average as described in Sec. 8.4.1. In practice, this means that we must repeat the procedure above for a whole range of different intensities in order to find the final focal volume averaged signal.

Figure 9.1 presents the theoretical longitudinal momentum distributions which are to be compared with the experimental counterparts. These latter results are published in Figs. 1 and 2 of Ref. [26].

For argon, Figs 9.1 (a) and (b), we generally find good agreement with the experimental results. At $400\,\mathrm{nm}$, both the theoretical and experimental spectra show two dominant peaks around $k_z = 0.35\,\mathrm{a.u.}$ and a local minimum at zero longitudinal momentum. The most significant difference is that the theoretical distributions show several sharp peaks at higher momenta that can be associated with multiphoton ionization. These peak structures can only be seen weakly in the experimental distribution which falls off nearly monotonically as $|k_z|$ increases beyond $0.5\,\mathrm{a.u.}$ The absence of peak structures in the experimental distribution at large momenta may be due to a limited resolution in the experimental set-up. For a laser wavelength of $800\,\mathrm{nm}$, both the theoretical momentum distributions shown in panel (b) as well as the experimental distributions show maxima at $k_z = 0\,\mathrm{a.u.}$ for all intensities.
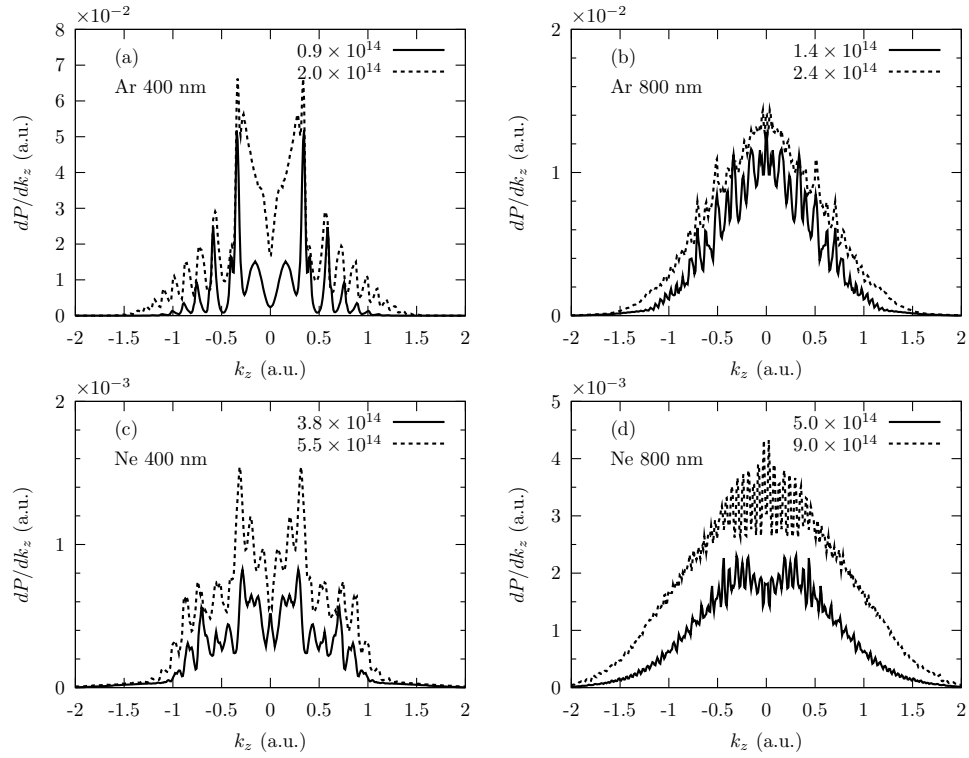
Figure 9.1: The longitudinal momentum distribution of the electron produced by strong-field ionization of the noble gas atoms argon and neon. Panels (a) and (b) show the distributions obtained for argon at the wavelengths 400 nm and 800 nm, respectively. Panels (c) and (d) show the corresponding distributions for neon, also for the wavelengths 400 nm and 800 nm. Each panel presents the results for two different peak intensities as indicated by the numbers (in units of W/cm^2). In all panels the results for the lower intensity are scaled [by 10 in panels (a), (b), and (d) and by 3 in panel (c)] for better visualization on the same scale as the higher intensity.

For neon at 400 nm, Fig. 9.1 (c), we find several peaks in the longitudinal momentum distribution for both intensities shown. Although experiment and theory do not agree perfectly on the exact location and height of the peaks, we do find both theoretically and experimentally that the peaks form a plateau-like distribution in the range $|k_z| < 1$ a.u. The neon atom exposed to a 400 nm field is the only case where such a plateau is found. In all other cases studied, the distribution falls off for $|k_z| > 0.5$ a.u. Finally, we turn to neon at 800 nm, panel (d). The theoretical momentum distribution at the peak intensity 5×10^{14} W/cm^2 shows a local minimum at zero longitudinal momentum in accordance with the experimental finding. Experimentally, however, this minimum is seen for all intensities up to 1.2×10^{15} W/cm^2. As we see from the dashed curve in Fig. 9.1 (d), the theory predicts that

the minimum disappears already at $9 \times 10^{14}\,\mathrm{W/cm^2}$. The reason for this disagreement is not well understood at the time of writing.

In summary, we find very good agreement between experiments and theory for the longitudinal momentum distributions obtained by strong-field ionization of neon and argon. The only feature that does not seem to be readily agreed upon is the local minimum at zero longitudinal momentum for neon at 800 nm. Experimentally, this minimum persists to exist for all intensities at least up to $1.2 \times 10^{15}\,\mathrm{W/cm^2}$. On the other hand, the theoretical findings predict that the minimum should disappear between the intensities 5 and $9 \times 10^{14}\,\mathrm{W/cm^2}$. Nevertheless, based on the general agreement, we conclude that a single-active electron model is quite sufficient to account for the major features seen in strong-field ionization of noble gas atoms. The agreement between theory and experiment is even obtained for an effective one-electron potential that is not particularly optimized to describe singly excited stated accurately. Hence, we may infer that intermediate resonances do not play a major role in the present studies.

9.3   Strong-field ionization of arbitrarily oriented H_2^+ molecules

In contrast to atoms, molecules do not exhibit spherical symmetry, and effects of molecular orientation with respect to the laser field are therefore present in strong-field phenomena, including ionization. While approximate theoretical models have been used to describe such alignment dependencies, *ab initio* calculations have mainly been performed for the single-electron H_2^+ molecule with internuclear axis parallel to the direction of a linearly polarized laser field [31–34]. For this particular geometry, the calculations are greatly simplified by azimuthal symmetry since the full dimensionality of the electronic problem can effectively be reduced to two dimensions as described in Sec. 5.3. The following two subsections describe some effects that arise if the molecule is not aligned parallel to the field.

9.3.1   Enhanced ionization

The subject for the present section is the phenomenon known as enhanced ionization; a term that refers to the fact that the H_2^+ molecule is more likely to ionize from a particular (large) internuclear distance. This effect was much debated in the nineties for the parallel geometry [35] while the present results now discuss the problem for non-parallel geometries.

The following numerical results are obtained with the length gauge Hamiltonian. The field frequency corresponds to a central wavelength of 800 nm ($\omega_0 = 0.057\,\mathrm{a.u.}$) and is chosen to match the Ti:sapphire lasing frequency. The pulse duration is 10 fs, corresponding to four optical cycles and the
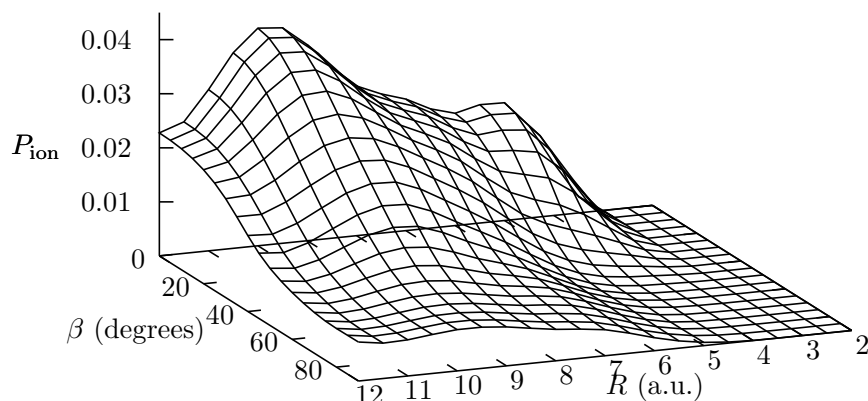
Figure 9.2: Ionization probability as function of internuclear distance R and angle β between the internuclear axis and the laser polarization. The pulse duration is 10 fs, the central wavelength 800 nm and the peak intensity 10^{14} W/cm^2.

peak intensity is 10^{14} W/cm^2. These parameters are used for all the calculations presented in this subsection. We use a radial grid consisting of 2048 points extending to a maximum distance of $r_{\max} = 120$ a.u. An absorbing boundary is applied near $r_{\max}$ in order to avoid reflections at the edge of the box. Note that we are only interested in the total ionization probability and we can afford to absorb the part of the wave function that approaches the box boundary. The angular degrees of freedom are treated in a spherical harmonics basis that includes angular momentum states up to $l_{\max} = 35$. Furthermore, a time step size of $\Delta t = 5 \times 10^{-3}$ a.u. is used for a convergent numerical solution of the time dependent Schrödinger equation.

The total ionization probability is calculated at the end of the pulse by the projection method (8.18) that only requires knowledge of the bound states. These field-free states are obtained by a diagonalization of the field-free Hamiltonian in the same basis as is used for the time-dependent calculation.

In Fig. 9.2, we present the ionization probability dependence on the internuclear distance R and the angle β between the molecular axis and the laser polarization axis. The initial state is the $1\sigma_g$ electronic ground state. In the parallel geometry $\beta = 0°$, our results are in good agreement with previously published results [31–34]. The main features are the maxima around $R = 5$ a.u. and $R = 10$ a.u. These structures of enhanced ionization were explored in detail [34, 36] and are identified as charge resonance phenomena due to a strong dipole coupling between the lowest σ_g and σ_u states.

For increasing angle β, the two peak structures gradually disappear and merge into a single broad plateau-like structure for $R = 6-10$ a.u. at $\beta = 90°$. The ionization probability simply increases with the internuclear distance since the electron becomes less tightly bound as R increases. The reason for
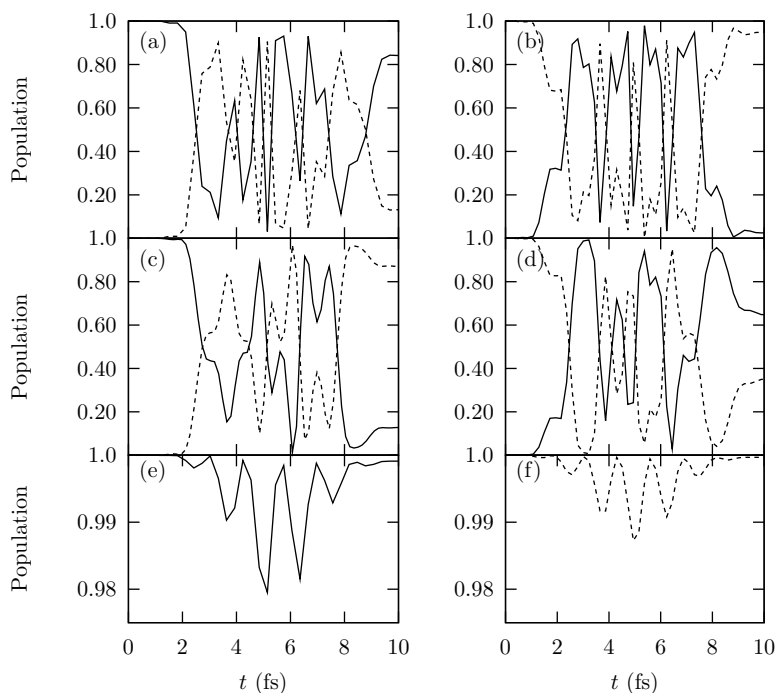
Figure 9.3: Population in the field-free states $1\sigma_g$ (solid) and $1\sigma_u$ (dashed) during the laser pulse. The initial state is the $1\sigma_g$ electronic ground state in panels (a),(c) and (e) and the $1\sigma_u$ orbital in panels (b),(d) and (f). In all panels the internuclear distance is $R = 5$ a.u. with an angle of 0° [(a) and (b)], 45° [(c) and (d)] and 90° [(e) and (f)] between the laser polarization direction and the internuclear axis. The parameters of the laser field are as in Fig. 9.2.

the lack of peak structure is explained by the symmetries of the σ_g and σ_u states. In the perpendicular geometry, the Hamiltonian is invariant under a reflection in the plane perpendicular to the molecular axis midway between the nuclei. The quantum number corresponding to the reflection symmetry in this plane is therefore conserved. Since the initial state is a σ_g state which is even under reflection, it is impossible to couple to states of odd reflection symmetry. The latter class of states includes $\sigma_u, \pi_g, \delta_u$ etc., and, hence, the resonance coupling to the $1\sigma_u$ state is excluded in the perpendicular geometry. Likewise, if one starts out in the $1\sigma_u$ state, no coupling to the even reflection states can occur. In the intermediate angles, the charge resonance coupling can be driven only by the field component parallel to the molecular axis.

We now give a more quantitative picture of the resonance coupling. Due to the fact that the $1\sigma_g$ state is more tightly bound than the $1\sigma_u$ state, we expect that coupling to the continuum is most favorable from the u state. In Fig. 9.3, we show the time evolution of the populations in the $1\sigma_g$
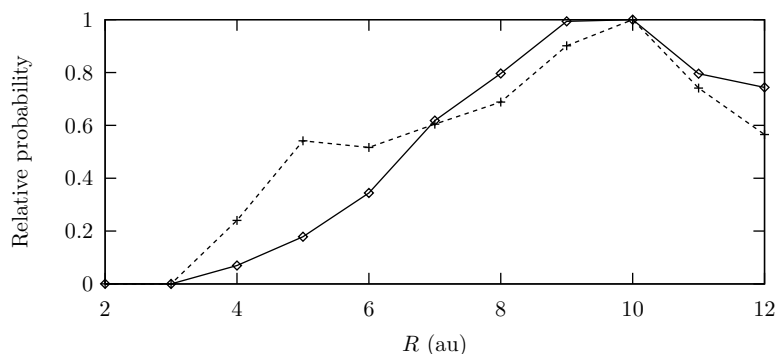
Figure 9.4: Orientation averaged ionization probability (solid) and ionization probability at parallel geometry (dashed). The data series are normalized separately. The parameters of the laser field are as in Fig. 9.2.

and $1\sigma_u$ states at different angles β between the laser polarization and the internuclear axis. We obtain the populations by a projection of the time-dependent wave function on the field-free states. In panels (a), (c) and (e), the initial state is the $1\sigma_g$ ground state corresponding to the calculations presented in Fig. 9.2, while the initial state is the $1\sigma_u$ state in panels (b), (d) and (f). Panels (a)-(d) show that the population oscillates between the states when the field has a component parallel to the molecular axis. On the contrary, panels (e) and (f) show that the population effectively stays in the initial states $1\sigma_g$ and $1\sigma_u$, respectively, when the field is perpendicular to the molecular axis. The population transfer in the parallel geometry between the $1\sigma_g$ and $1\sigma_u$ states was previously analysed by a two-level model [37]. The oscillation between the two states is characterized by a rapid Rabi frequency due to a large dipole matrix element. Since the dipole coupling between the two states is proportional to $\cos\beta$, the oscillation frequency decreases as the angle increases. At the perpendicular geometry, the coupling and the Rabi frequency are zero, and, hence, the population is trapped in the initial state.

Experimentally, the ionization rate was measured as a function of the internuclear distance by preparing a vibrational wave packet in H_2^+ to produce moving nuclei [38]. The experimental results were compared with great success with calculations in which the molecule was assumed to be aligned parallel to the field axis [31]. As we can see from Fig. 9.2, the large geometrical dependencies could possibly make a significant difference if the ionization signal is calculated for an aligned or a randomly oriented ensemble. After all, for randomly oriented molecules, the number of molecules that are aligned at an angle in $[\beta, \beta + d\beta]$ is proportional to $\sin\beta d\beta$. In Fig. 9.4, we show the R dependent orientational averaged ionization probability, which is calculated as $\langle P(R)\rangle \propto 2\pi \int_0^\pi P_{\mathrm{ion}}(R, \beta)\sin\beta d\beta$, where $P_{\mathrm{ion}}(R, \beta)$ is the ionization probability given in Fig. 9.2. Despite the large variation of $P_{\mathrm{ion}}(R, \beta)$,

the ionization probabilities surprisingly depend on the internuclear distance in quite a similar way for both random oriented as well as parallel aligned molecules. Therefore we may conclude, what was not a priori obvious, that it is indeed reasonable to compare the experimental data with calculations [31–34] for molecules aligned along the laser polarization.

9.3.2   Alignment dependent above threshold ionization

The above threshold ionization (ATI) spectrum has been studied for atoms since the early days of strong-field physics. For molecules, however, it has not been studied until recently how ATI spectra depend on *alignment*. The experimental results from Ref. [39] reported that the ATI spectrum of N$_2$ is *independent* of alignment up to an overall scaling factor. The experimental findings were supported by calculations within the framework of the molecular tunneling theory [40]. One question is now if the ATI spectrum will be independent of alignment for *any* molecule. It is of course impossible to use an *ab initio* approach to calculate the alignment dependent ATI spectra that result from strong-field ionization of N$_2$ and other molecules that might be used as experimental samples. Instead, we will simulate different molecular species by considering strong-field ionization from the $1\sigma_g$ and $1\sigma_u$ states of H$_2^+$. The $1\sigma_u$ state is in some sense an artificial initial state since the nuclei will dissociate if the electron is excited to this ungerade state. Nevertheless, we will use the $1\sigma_u$ state as a generic example of an ungerade initial state and the results may be common to other ungerade states.

We solve the time dependent Schrödinger equation as described in chapter 5. As in Sec. 9.3.1 we use a field frequency that corresponds to a central wavelength of 800 nm ($\omega_0 = 0.057$ a.u.). The pulse duration is 19 fs, corresponding to seven optical cycles. The radial grid consists of 4096 points that extend to a maximum distance of $r_{\max} = 600$ a.u. The spherical harmonics representation is truncated at $l_{\max} = 23$. Furthermore, a time step size of $\Delta t = 5 \times 10^{-3}$ a.u. is used for a convergent numerical solution of the time dependent Schrödinger equation. Here we employ the velocity gauge form of the electron-field interaction. This form has proven to be superior over the length gauge with respect to the number of angular momentum states needed to obtain convergent results.

Having found the final wave function at time T after the end of the pulse, we now encounter the problem that the scattering states of H$_2^+$ are not readily available. This means that the full momentum distribution cannot be obtained. Here we are, however, only interested in the *energy* differential distribution. Fortunately, there are simpler ways to get that information than first constructing the full momentum distribution. The most straightforward method is perhaps to find all energy eigenstates that are supported in the box. The set of discrete, box normalized energy eigenstates $\{\Phi_n^{(\mathrm{box})}\}$ can now be sorted according to the corresponding eigenenergies $\{E_n^{(\mathrm{box})}\}$. A

straightforward binning procedure can now give the energy spectrum. First, we divide the energy axis in pieces of the interval ΔE. The probability to measure the electron in the interval $[E, E + \Delta E]$ is then easily obtained as

$$\frac{dP}{dE}\Delta E = \sum_{\substack{n \\ E<E_n^{(\mathrm{box})}<E+\Delta E}} |\langle \Phi_n^{(\mathrm{box})}|\Psi(T)\rangle|^2. \qquad (9.18)$$

Yet another method to calculate the photoelectron spectrum without scattering states is proposed in Ref. [41].

Figure 9.5 presents the alignment dependent ATI spectra that result from strong-field ionization from the $1\sigma_g$ state. The nuclei are fixed at the equilibrium distance of 2 a.u. The two panels show the results for different peak intensities, namely (a) $3 \times 10^{14}\,\mathrm{W/cm}^2$ and (b) $5 \times 10^{14}\,\mathrm{W/cm}^2$. The results presented in Fig. (9.5) essentially agree with the proposal that the ATI spectrum is independent of orientation, except for an overall scaling factor. As in Sec. 9.3.1, the overall ionization decreases with increasing alignment angle. But with the present results, we are now able to present a more detailed picture, namely that an increasing angle results in a decreased ionization yield *uniformly* for all energies of the outgoing electron. For both intensities, we see the usual ATI peak structures that are separated by the photon energy, following the rule $E_\mathcal{N} = \mathcal{N}\omega - I_p - U_p$. The fact that the peaks do not occur at the same energy positions in panels (a) and (b) can be attributed to the fact that the ponderomotive energy changes from 0.66 a.u. at $3 \times 10^{14}\,\mathrm{W/cm}^2$ to 1.1 a.u. at $5 \times 10^{14}\,\mathrm{W/cm}^2$.

Now we turn to the ionization from an initial $1\sigma_u$ state. In this state, the electron is bound by 0.668 a.u., i.e., somewhat less than the binding energy of 1.103 a.u. in the $1\sigma_g$ state. Hence, an electron in the $1\sigma_u$ will ionize at a much lower intensity than used for the $1\sigma_g$ state. In Figs. 9.6 (a) and (b) we show the ATI spectra obtained with the peak intensities $5 \times 10^{13}\,\mathrm{W/cm}^2$ and $1 \times 10^{14}\,\mathrm{W/cm}^2$, respectively. Clearly, the results presented in Fig. 9.6 are completely different from Fig. 9.5 for the $1\sigma_g$ state. The difference between the spectra is certainly not just an overall scaling factor for the $1\sigma_u$ state. For the peak intensity $5 \times 10^{13}\,\mathrm{W/cm}^2$ [panel (a)], the total ionization yield is approximately constant for the alignment angles $\beta = 0°, 30°, 60°$ and much suppressed at the perpendicular alignment $\beta = 90°$. Perhaps the most interesting feature seen in Fig. 9.6 is that the ATI peaks do not coincide exactly for the different alignment angles. This effect is even more pronounced at the high peak intensity, $1 \times 10^{14}\,\mathrm{W/cm}^2$ [panel (b)]. For that intensity, the peaks that occur for $\beta = 0°$ and $30°$ are located at the same energies where the valleys occur for $\beta = 60°$ and $90°$ and vice versa. This result is quite surprising since the usual rule (8.1) for the location of the ATI peaks breaks down. The detailed interpretation of Fig. 9.6 remains to be studied and we have not yet found an explanation why the ATI peaks shift with geometry. One possibility can be that ionization occurs via an
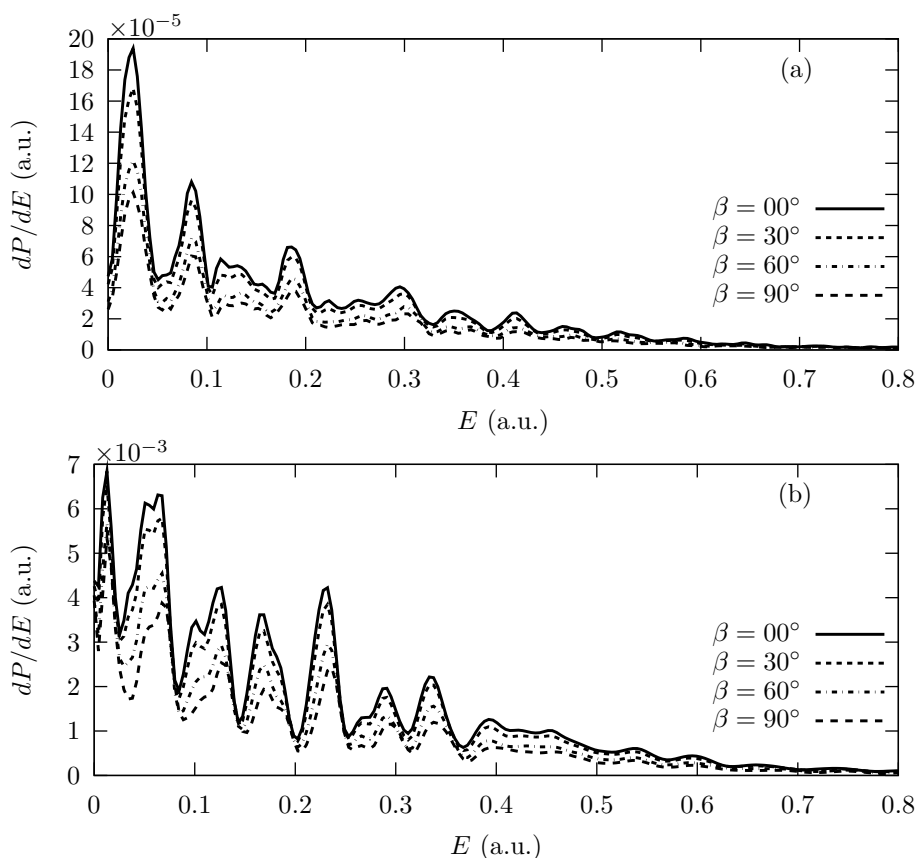
Figure 9.5: Orientation dependent ATI spectrum from the initial state $1\sigma_g$. The peak intensity is (a) $3 \times 10^{14}\,\mathrm{W/cm^2}$ and (b) $5 \times 10^{14}\,\mathrm{W/cm^2}$. The pulse duration is 19 fs in both panels.

intermediate resonance that is only accessible from particular geometries such as demonstrated in Fig. 9.3.

One may argue that the present calculation is not completely realistic due to the neglect of vibrational motion of the nuclei. It is quite questionable to believe that the fixed-nuclei approximation applies on a 19 fs timescale, since the characteristic vibrational period in the electronic ground state of H$_2{}^+$ is around 15 fs. Conventional one-photon photoelectron spectroscopy of molecules shows absorption bands that can be further resolved in structures that in turn can be associated with vibrational and rotational motion of the nuclei. These bands can have a width of several electronvolts, depending on the overlap between the initial and final vibrational wave function. If such bands are also present in the multiphoton case, each multiphoton peak will be broadened on the same scale as the photon energy and the sharp peak structures known from atomic ATI spectra may disappear. This coupling
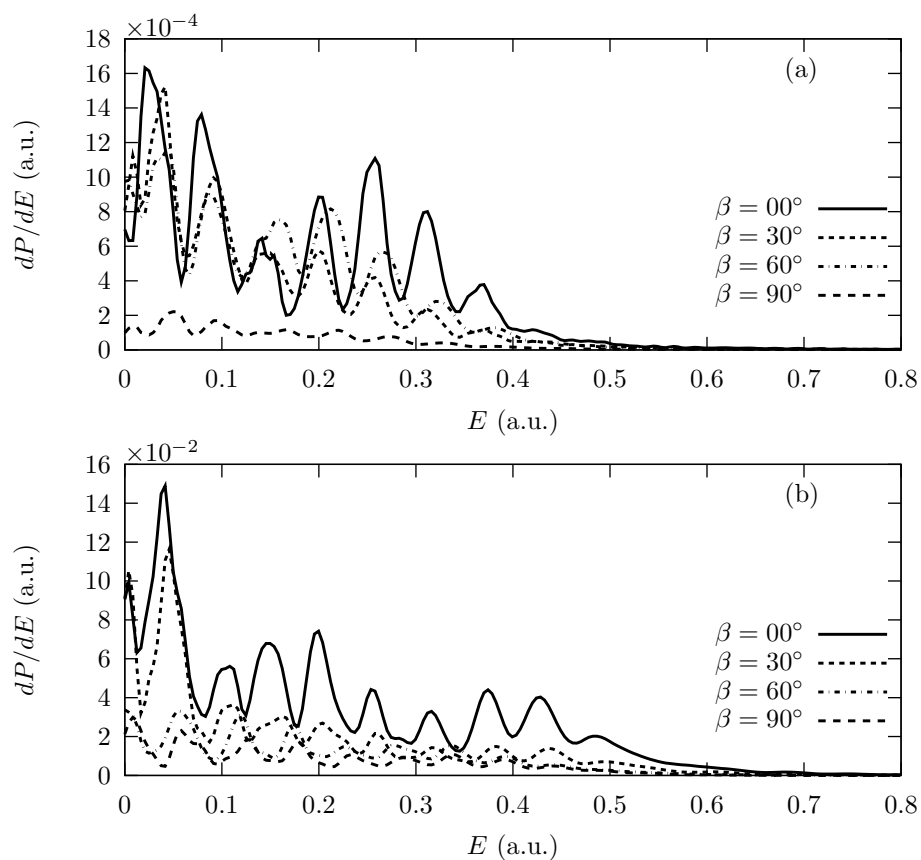
Figure 9.6: Orientation dependent ATI spectrum from the initial state $1\sigma_u$. The peak intensity is (a) $5 \times 10^{13}\,\mathrm{W/cm^2}$ and (b) $1 \times 10^{14}\,\mathrm{W/cm^2}$. The pulse duration is $19\,\mathrm{fs}$ in both panels.

between the electronic and nuclear motion is certainly an interesting study for the future.

Chapter 10

Current status and future work

With the algorithms described in chapter 5, it is in principle possible to solve any problem that concerns a single active electron that interacts with a cylindrically symmetric static potential and a linearly polarized field with arbitrary direction of the polarization vector. Even elliptically polarized fields can be studied by the rotation technique described in Sec. 5.5.3.

This final chapter briefly lists some of the problems that we may be concerned with in the future. Section 10.1 describes how the numerical solution of the single active electron Schrödinger equation can be obtained efficiently in the future. Section 10.2 concerns the Schrödinger equation for two physical systems that must be treated beyond the single active electron model.

10.1 Numerical implementations of single active electron algorithms

A number of the algorithms that have been described in chapters 5 and 6 are not implemented numerically yet. At the present, the numerical implementation includes only the Fourier spectral method for the radial kinetic energy which suffers from the poor parallelization properties as discussed in Sec. 6.1.1. It should not be too difficult to implement the Crank-Nicolson method for the radial kinetic operator. With the parallelization technique described in Sec. 6.1.1, the parallelization is expected to be improved drastically on a distributed memory system.

10.1.1 Optimization of the basis

The single center spherical harmonics representation appears to be reasonably good for many systems. Of course, it is possible to obtain convergence with a fewer number of basis functions if the basis is specifically optimized to the problem at hand. For example, a grid representation in prolate

spheroidal coordinates is close to optimal for diatomic molecules. The draw-back of a basis that is highly optimized for one particular purpose is that it is often equally far from being optimal for any other problem. It is a further advantage of the spherical harmonics representation that it has been used extensively throughout history as the basis in scattering theory. This is the reason why the partial wave expansion of scattering states has already been studied. Even though the partial wave expansion is quite complicated in the multichannel case, it can be constructed within a well established theory.

Having argued that the spherical harmonics expansion provides a reasonable representation of the angular variables, we now turn to the discretization of the radial coordinate. As already demonstrated in Sec. 3.6, the equidistant radial grid is a trade-off between accuracy and computational cost if the radial part of the wave function varies on several different length scales. This is often the situation encountered in strong-field ionization. In the initial bound state at short distances, the wave function varies on the length scale of the Bohr radius. As seen in Sec. 3.6, the grid spacing should preferably be taken somewhat smaller, e.g. ~ 0.1 a.u. The ionization is dominated by absorption of a few number of photons above the ionization threshold and the electron only gains the kinetic energy corresponding to a few number of visible or infrared photons of energies around $0.05 - 0.1$ a.u.. The maximum momentum that needs to be accounted for at large distances is then usually somewhat less than one atomic unit of momentum which can be accommodated by a grid spacing as large as $\sim \pi$ a.u. This brief estimation hints that it may be possible to reduce the number of grid points by an order of magnitude if one can take appropriate advantage of a non-equidistantly spaced radial grid. The time consumption associated with all potential operators in the split-step scheme will be reduced accordingly since the computational complexity scales linearly with the number of grid points, Tables 5.1 and 5.2. In fact the motivation for studying the finite element method is exactly the desire for a basis that can be defined non-equidistantly. We showed in Sec. 3.6.1 that this basis works well for the solution of the *stationary* Schrödinger equation. Unfortunately, it turns out that for an efficient solution of the time-dependent Schrödinger equation it is too disadvantageous that the coordinate operator is non-diagonal [cf. Eq. (3.47)]. For example, the potential operators are not block-diagonal if we use a finite element representation along the radial coordinate axis.

So far, several good reasons exist for choosing an equidistant grid. First, we have employed the discrete sine transformation in order to apply the radial kinetic operator in the diagonal representation. This transformation is only unitary for an equidistant grid. Additionally, an equidistant grid is required for a fast $\mathcal{O}(n_r \log_2(n_r))$ transformation. At the time when the numerical project was initiated, it was believed that the Fourier spectral method would lead to superior convergence properties compared with nearly any other evaluation of derivatives. This assumption relied on test cases on

idealized systems such as the harmonic oscillator. Recent studies have now indicated that the Fourier method performs particularly bad when Coulomb singularities are present. Even the Crank-Nicolson approximation with the three-point finite difference seems to be a better choice for this case [42]. In contrast to the Fourier method, it has been proposed that an extension of the Crank-Nicolson method can be used as an exactly unitary operator for a non-equidistant grid [43]. This is certainly an idea that should be considered in the future.

10.2   Beyond the single active electron model

The only true one-electron system is of course the hydrogen atom and hydrogen-like ions. For any other atomic or molecular system, the single active electron model is an approximation. While the single active electron (SAE) model successfully describes many strong-field phenomena, we cannot expect that *everything* can be explained by such an approximation. Below, we shall briefly mention two examples where the SAE model is inadequate, namely the correlated electron dynamics encountered in multielectron systems and the nuclear motion in molecules.

10.2.1   Two-electron dynamics

The range of phenomena that can occur for two electrons is much broader than for one-electron. One example is the existence of doubly excited states that are degenerate with the single-ionization continuum and lead to autoionization [44]. Entanglement is another example of a two-body phenomenon. A pair of entangled particles may show quite surprising behaviour, and furthermore, entanglement is believed to be the cornerstone in the implementation of any quantum computer.

As we have seen in Sec. 9.2, the SAE model successfully describes the single ionization dynamics of noble gas atoms. But if we are to study *double* ionization from noble gas atoms, the single active electron model will fail. In that case, the double ionization yield is known to be enhanced due to correlated dynamics of two electrons [16]. A proper description of double ionization must therefore include at least two active electrons.

It may also be necessary to include two active electrons when describing diatomic molecules. The electronic structure of many diatomic molecules, e.g. N_2, H_2, and O_2, can be approximated by two electrons outside a frozen core of inner electrons. Correlation effects in such molecular systems remain to be studied in full dimensionality. Initial attempts have tried to describe the two-electron dynamics with a reduced dimensionality so that the two electrons are confined to moving along a line [45, 46]. One might expect the two electrons to become too strongly correlated in such a model since

the electrons – in contrast to the real case – cannot pass each other without colliding.

We have already seen that a fully three-dimensional Schrödinger equation is significantly harder to solve than an effectively two-dimensional Schrödinger equation. A two-electron wave function must depend on *six* spatial variables, and the time dependent Schrödinger equation will therefore be extremely demanding to solve numerically. Even if the problem *can* be solved, it will probably be an equally difficult task to interpret the six dimensional wave function in terms of the physical observables. For example, it is already very complicated to obtain the exact two-electron scattering states for a simple system such as the helium atom.

10.2.2   Inclusion of nuclear motion

Throughout this thesis, we have considered the $H_2{}^+$ molecule exclusively as a one-electron system. It is of course an approximation to regard the nuclei as being fixed. The rotational period and vibrational timescales of $H_2{}^+$ are on the order $170\,\mathrm{fs}$ and $15\,\mathrm{fs}$, respectively. For interaction with laser pulses on the duration of a few tens of femtoseconds, the rotational motion can be considered as being frozen. On the contrary, it may not be obvious that the motion associated with the vibrational motion can be regarded as being fixed. Furthermore, if we wish to study dissociation or Coulomb explosion of the nuclei, it is of course essential to include the nuclear motion in the theory. The inclusion of nuclear degrees of freedom means that both the electron and the nuclei can absorb energy from a strong laser field. A detailed understanding of how this sharing of energy occurs has not been provided yet.

If we write down the time dependent Schrödinger equation for the $H_2{}^+$ molecule including the internuclear coordinate R as a dynamical variable, we arrive at a four dimensional partial differential equation

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r}, R, t) = \left(-\frac{1}{2\mu}\frac{\partial^2}{\partial R^2} + \frac{1}{R} + \mathcal{H}^{(e)}(t; R)\right)\Psi(\boldsymbol{r}, R, t). \tag{10.1}$$

μ is the reduced mass of the nuclei, and $\mathcal{H}^{(e)}(t; R)$ denotes the electronic Hamiltonian which is diagonal in the coordinate representation of the internuclear distance. While such an equation will certainly be at least an order of magnitude harder to solve than the three-dimensional, it may not be totally impossible within a foreseeable future. We believe that we can reuse most of the algorithms that have been developed for the fixed nuclei case. If we use a split-operator technique, we can split the full Hamiltonian such that we first propagate by $\mathcal{H}^{(e)}(t; R)$ which corresponds to a propagation of the electron while keeping the nuclei fixed. Subsequently, we propagate the nuclei by the two remaining terms with the electron being fixed. Now we assume that the internuclear coordinate is discretized on an N_R-point

grid. In the first step, we can use the usual electron propagation schemes *independently* for each of the grid points in the R coordinate as if the internuclear distance were fixed. In the second step, couplings are introduced between the partial solutions just obtained at the different internuclear distances. With this approach, the full four-dimensional equation seems to be approximately N_R times as hard to solve as the electronic problem alone. Hence, before considering the full problem, it will probably be helpful first to exploit all the optimizations that can be made for the electronic problem as discussed in Sec. 10.1.

Bibliography

[1] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley Publishing Company, Inc., 1994), rev. ed.

[2] E. Cormier and P. Lambropoulos, J. Phys. B **29**, 1667 (1996).

[3] H. G. Muller, Laser Physics **9**, 138 (1999).

[4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++* (Cambridge University Press, 2002), 2nd ed.

[5] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions* (Dover Publications, Inc., New York, 1964).

[6] M. R. Hermann and J. A. Fleck, Phys. Rev. A **38**, 6000 (1988).

[7] B. I. Schneider and N. Nygaard, Phys. Rev. E **70**, 056706 (2004).

[8] R. B. Sidje, ACM Trans. Math. Softw. **24**, 130 (1998).

[9] L.-Y. Peng and A. F. Starace, J. Chem. Phys. **125**, 154311 (2006).

[10] D. Bauer and P. Koval, Comput. Phys. Comm. **174**, 396 (2006).

[11] P. N. Swartzrauber, SIAM J. Numer. Anal. **16**, 934 (1979).

[12] D. M. Healy, D. N. Rockmore, P. J. Kostelec, and S. Moore, Journal of Fourier analysis and applications **9**, 341 (2003).

[13] R. S. Womersley and I. H. Sloan, Adv. Comput. Math. **14**, 195 (2001).

[14] T. Birkeland, M. Førre, J. P. Hansen, and S. Selstø, J. Phys. B **37**, 4205 (2004).

[15] M. Bueno and F. Dopico, BIT Numer. Math. **44**, 651 (2004).

[16] B. Walker, B. Sheehy, L. F. DiMauro, P. Agostini, K. J. Schafer, and K. C. Kulander, Phys. Rev. Lett. **73**, 1227 (1994).

[17] P. Kruit, J. Kimman, and M. J. V. der Wiel, J. Phys. B **14**, L597 (1981).

[18] A. Eppink and D. Parker, Rev. Sci. Instrum. **68**, 3477 (1997).

[19] S. Selstø, J. F. McCann, M. Førre, J. P. Hansen, and L. B. Madsen, Phys. Rev. A **73**, 033407 (2006).

[20] G. Breit and H. A. Bethe, Phys. Rev. **93**, 888 (1954).

[21] M. Brosolo and P. Decleva, Chem. Phys. **159**, 185 (1992).

[22] H. Park and R. N. Zare, J. Chem. Phys. **104**, 4554 (1996).

[23] L. A. A. Nikolopoulos, Phys. Rev. A **73**, 043408 (2006).

[24] A. Rudenko, K. Zrost, C. Schröter, V. de Jesus, B. Feuerstein, R. Moshammer, and J. Ullrich, J. Phys. B **37**, L407 (2004).

[25] C. Maharjan, A. Alnaser, I. Litvinyuk, P. Ranitovic, and C. Cocke, J. Phys. B **39**, 1955 (2006).

[26] A. S. Alnaser, C. M. Maharjan, P. Wang, and I. V. Litvinyuk, J. Phys. B **39**, L323 (2006).

[27] T. Morishita, Z. Chen, S. Watanabe, and C. Lin, Phys. Rev. A **75**, 23407 (2007).

[28] M. Wickenhauser, X. Tong, D. Arbó, J. Burgdörfer, and C. Lin, Phys. Rev. A **74**, 041402 (2006).

[29] Z. Chen, T. Morishita, A. Le, M. Wickenhauser, X. Tong, and C. Lin, Phys. Rev. A **74**, 53405 (2006).

[30] X. M. Tong and C. D. Lin, J. Phys. B **38**, 2593 (2005).

[31] L.-Y. Peng, D. Dundas, J. F. McCann, K. T. Taylor, and I. D. Williams, J. Phys. B **36**, L295 (2003).

[32] M. Vafaee and H. Sabzyan, J. Phys. B **37**, 4143 (2004).

[33] D. A. Telnov and S.-I. Chu, Phys. Rev. A **71**, 013408 (2005).

[34] A. D. Bandrauk and H. Z. Lu, Phys. Rev. A **62**, 053406 (2000).

[35] J. H. Posthumus, Rep. Prog. Phys. **67**, 623 (2004).

[36] X. Chu and S. Chu, Phys. Rev. A **63**, 013414 (2000).

[37] T. Zuo, S. Chelkowski, and A. D. Bandrauk, Phys. Rev. A **49**, 3943 (1994).

[38] G. N. Gibson, M. Li, C. Guo, and J. Neira, Phys. Rev. Lett. **79**, 2022 (1997).

[39] D. Zeidler, A. B. Bardon, A. Staudte, D. M. Villeneuve, R. Dörner, and P. B. Corkum, J. Phys. B **39**, L159 (2006).

[40] X. M. Tong, Z. X. Zhao, and C. D. Lin, Phys. Rev. A **66**, 033402 (2002).

[41] L. A. A. Nikolopoulos, T. K. Kjeldsen, and L. B. Madsen, Phys. Rev. A **75**, 063426 (2007).

[42] T. Birkeland, private communication (2007).

[43] A. K. Kazansky, J. Phys. B **31**, L579 (1998).

[44] U. Fano, Phys. Rev. **124**, 1866 (1961).

[45] M. Lein, T. Kreibich, E. K. U. Gross, and V. Engel, Phys. Rev. A **65**, 33403 (2002).

[46] N. Nguyen and A. Bandrauk, Phys. Rev. A **73**, 32708 (2006).