Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Diplom-Physiker Christian Wacker
born in Unna, Germany

Oral examination: May 4th, 2016

Image Formation in Electron Microscopes at Low Energies

Referees: Prof. Dr. Rasmus R. Schröder

Prof. Dr. Ulrich Schwarz

Abstract

In the past, the spatial resolution in electron microscopy was mainly limited by the unavoidable aberrations of round electron lenses. Nowadays, the successful realization of aberration correction allows achieving atomic resolution for electron energies as low as 20 keV. In this energy range, existing models for the image formation process must be augmented by previously neglected effects, since the commonly used paraxial approximation breaks down. Starting from quantum mechanical principles, this work will show that electron microscopic image contrast from elastic scattering can be calculated by solving a series of eigenvalue problems. Thereby, new theoretical insights into the scattering process are gained. Furthermore, the derived analytical matrix equations turn out to be suited for numerical calculations as well. Through the inclusion of backscattering effects, a unified description of the image formation process in different types of electron microscopes is obtained. Thus, for the first time, the elastic scattering channel in a scanning electron microscope (SEM) can be simulated based on the wave-mechanical description of the electron. Moreover, this work will show that the direct quadrature of the Schrödinger equation can be used to simulate electron micrographs. With the help of these new scattering solutions, the validity of the paraxial approximation is investigated in detail.

Zusammenfassung

In der Vergangenheit war die räumliche Auflösung in der Elektronenmikroskopie hauptsächlich durch die unvermeidbaren Aberrationen runder Elektronenlinsen begrenzt. Die erfolgreiche Verwirklichung der Aberrationskorrektur ermöglicht es heutzutage, für Elektronenenergien bis hinunter zu 20 keV atomare Auflösung zu erreichen. In diesem Energiebereich müssen die bestehenden Modelle für den Bildentstehungsprozess um zuvor vernachlässigte Effekte erweitert werden, da die üblicherweise verwendete paraxiale Näherung zusammenbricht. Ausgehend von quantenmechanischen Prinzipien zeigt diese Arbeit, dass der durch elastische Streuung verursachte elektronenmikroskopische Bildkontrast berechnet werden kann, indem eine Reihe von Eigenwertproblemen gelöst wird. Auf diese Weise erhält man neue theoretische Einsichten in den Streuvorgang. Es stellt sich außerdem heraus, dass sich die hergeleiteten, analytischen Matrixgleichungen auch für numerische Rechnungen eignen. Durch die Hinzunahme von Rückstreueffekten erhält man eine einheitliche Beschreibung für den Bildentstehungsprozess in unterschiedlichen Typen von Elektronenmikroskopen. Somit ist es erstmals möglich, den elastischen Streukanal im Rasterelektronenmikroskop basierend auf der wellenmechanischen Beschreibung des Elektrons zu simulieren. Darüber hinaus zeigt diese Arbeit, dass die direkte Quadratur der Schrödinger-Gleichung für die Simulation elektronenmikroskopischer Aufnahmen verwendet werden kann. Mithilfe dieser neuen Streulösungen wird die Gültigkeit der paraxialen Näherung genau untersucht.

Contents

List of Symbols and Abbreviations

a	amplitude function, page 25
A^2	propagation operator, page 41
$\underline{A}^{2,n}$	propagation matrix, page 42
a_B	Bohr radius $\frac{4\,\pi\,\varepsilon_0\,\hbar^2}{m\,e^2}$, page 15
α	atom index, page 13
B_α	B-factor, page 20
β	convergence angle, page 57
$\mathbf{b}^n$	coefficient vector for the backward-moving part of the wave, page 43
c	speed of light, page 7
γ	Lorentz factor $E/(m\,c^2)$, page 8
$\underline{D}^{2,n}$	diagonal matrix consisting of the eigenvalues of $\underline{A}^{2,n}$, page 42
Δ	three-dimensional Laplace operator, page 8
$\Delta_\perp$	two-dimensional Laplace operator, page 25
$\Delta\kappa_x, \Delta\kappa_y$	horizontal / vertical pixel size in reciprocal space, page 33
$\Delta x, \Delta y$	horizontal / vertical pixel size in direct space, page 33
Δz_n	slice thickness, page 27
δ	detector angle, page 74
E	total energy, page 7
e	elementary charge, page 7
ε_0	vacuum permittivity, page 12
$\mathcal{F}$	continuous Fourier transform, page 33
f	scattering amplitude, page 10
f^B	scattering amplitude in first Born approximation, page 11
f_α^γ	X-ray form factor, page 14
f_α^e	electronic form factor, page 13
$\mathbf{f}^n$	coefficient vector for the forward-moving part of the wave, page 43
$\hbar$	reduced Planck constant, page 7
$\mathbf{j}$	probability density current, page 8
$\mathbf{K}$	scattering vector, page 11
k	wave number of the electron $\frac{2\,\pi}{\lambda}$, page 8
$\boldsymbol{\kappa}$	two-dimensional vector in reciprocal space, page 21
$\kappa_{x,\mathrm{Ny}}, \kappa_{y,\mathrm{Ny}}$	Nyquist frequency in x- and y-direction, page 33
L	direct lattice, page 16
L'	reciprocal lattice, page 16
λ	wave length, page 8
L_x, L_y	physical width/height of the simulation area, page 33
M	total number of pixels $M_x \times M_y$, page 33
M_x, M_y	number of pixels in x- and y-direction, page 20
m	rest mass of the electron, page 7
N	total number of slices, page 28
n	slice index, page 28
$\overline{\Phi}_\alpha$	hybrid antiderivative, page 32
$\phi_{\mathbf{k}_i}$	plane wave, page 9
$\varphi^{(n)}$	total phase shift, page 29
ψ_{ex}	exit-wave, page 5
ψ_{in}	incoming wave, page 27
ψ_{ba}	backscattered wave, page 44
$\boldsymbol{\rho}$	two-dimensional vector in real space, page 21
ρ_α^e	electron charge density, page 13
$\frac{\mathrm{d}\,\sigma}{\mathrm{d}\Omega}$	differential cross section, page 10
σ_{tot}	total cross section, page 10
T	kinetic energy, page 8
ϑ	scattering angle, page 12
$\underline{U}^{2,n}$	unitary matrix whose columns are the eigenvectors of $\underline{A}^{2,n}$, page 42
V	potential energy, page 7
V_α	atomic potential, page 13
$\overline{V}_\alpha$	hybrid potential, page 32
v	reduced potential $\frac{2\,m\,\gamma}{\hbar^2}\,V$, page 9
v_α	reduced atomic potential, page 17
$v^{(n)}$	averaged reduced potential for slice n, page 27
Z	atomic number, page 12
ADF	annular dark-field, page 5
BF	bright-field, page 5

CBED convergent beam electron diffraction, page 3

FFT fast Fourier transform, page 3

FOLZ first-order Laue zone, page 18

FRC Fourier ring correlation, page 37

GPU graphics processing unit, page 4

HOLZ higher-order Laue zone, page 18

RK1 Runge-Kutta algorithm for the paraxial approximation of the Schrödinger equation, page 50

RK2 Runge-Kutta algorithm for the non-approximated Schrödinger equation, page 51

SEM scanning electron microscope, page 4

SOLZ second-order Laue zone, page 18

STEM scanning transmission electron microscope, page 3

TDS thermal diffuse scattering, page 18

TEM transmission electron microscope, page 1

ZOLZ zero-order Laue zone, page 18

Chapter 1

Introduction

Diffraction limits the resolution of every optical system. According to Abbe, the resolving power of a light microscope with parallel illumination is roughly equal to half the wavelength of the light used [1]. Therefore, it is impossible to image molecules or even atoms by using light microscopy. The electron microscope was conceived by Knoll and Ruska in the early 1930s as an instrument to overcome this limitation [2]. Because of the short wavelengths of electrons accelerated to high energies (cf. figure 1.1), the resolving power of electron microscopes quickly exceeded that of light microscopes. Scherzer, however, pointed out that the unavoidable spherical aberration of a round electron lens limits the achievable resolution [3]. Although this is not a fundamental problem, it proved to be very difficult to overcome [4]. Gabor even invented holography with the aim to correct such degraded images [5]. Only in 1998, it was demonstrated that correctors for the spherical aberration can improve the resolution of a transmission electron microscope (TEM) [6]. In the subsequent TEAM project[a], it was shown that the chromatic aberration can be corrected as well [7].

The practical resolution of an electron microscope is not only limited by electron optics, but also by the damage the electron beam inflicts on the sample. Because of the strong interactions of electrons with matter, radiation damage can hardly be avoided. Since the damage is caused by a large number of different mechanisms such as heating, ionization, knock-on damage (direct displacement of atoms) and chemical etching, its energy dependence is quite complicated [8]. Moreover, most damaging processes depend on the applied electron dose (number of electrons per area). This is very problematic, since the improved resolution of aberration-corrected electron microscopes can only be exploited by increasing the electron dose.

Novel materials, such as carbon nanotubes, fullerenes, two-dimensional boron nitride or graphene promise interesting new technologies and applications. For instance, graphene research is supported by the European Commission as a Future and Emerging Technology Flagship[b].

The study of the aforementioned materials in a TEM is challenging [9]: First of all, these materials are low-dimensional. Hence, in comparison to bulk specimens, only a smaller number of atoms can contribute to the scattering process. Moreover, these materials consist of elements with low atomic numbers and thus are weak scatterers. For these reasons, a high electron dose is required in order to obtain a usable signal-to-noise ratio. Additionally, these materials are very susceptible to knock-on damage since they are composed of light elements.

While the first TEM with a corrector for the chromatic aberration was operated at 200 kV

[a] Transmission Electron Aberration-Corrected Microscope, `http://foundry.lbl.gov/facilities/ncem/expertise.html`

[b] `http://graphene-flagship.eu/`

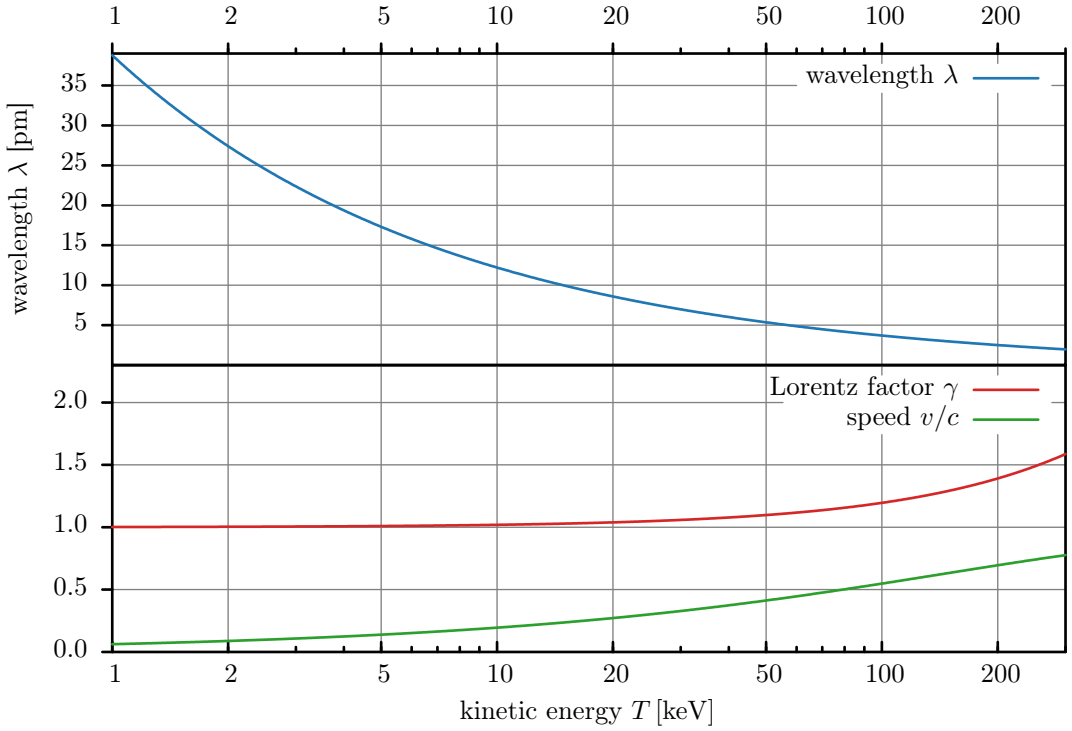

Figure 1.1: Wavelength λ, Lorentz factor γ and speed v/c of an electron as functions of its kinetic energy T.

[7], the SALVE project[c] targets acceleration voltages as low as 20 kV [10]. By this means, the knock-on damage for many carbon-based materials can be completely prevented, since a certain minimum energy is required for a direct displacement of an atom, the knock-on threshold. The precise value of this threshold depends on the binding energy of the atom. For single-layer graphene, it was found that the threshold is between 80 keV and 100 keV [9]. Since graphene is additionally an excellent electrical and thermal conductor, it is very stable under irradiation with electron energies below 80 keV [9].

The energy range between 20 keV and 80 keV is, therefore, very suitable for the investigation of these novel materials. Since a completely new energy range has thus become accessible to transmission electron microscopy, the question arises whether the existing theories of image formation and their interpretation as image simulation routines remain valid.

1.1 History of Image Simulations

Image simulations are an important tool for the interpretation of electron micrographs, since the image formation process in an electron microscope is quite complicated and very different from everyday experience. This is especially true for images with atomic resolution. A quantitative analysis of these images is in many cases only possible with the help of image simulations.

The surface of a typical object is visible because certain frequencies of the incoming light are absorbed while others are diffusely reflected. Hence, the information about the object is mainly

[c] Sub-Angstrom Low-Voltage Electron Microscopy, `http://www.salve-project.de/`

carried by the amplitude of the electromagnetic wave. In an electron microscope, however, only a small fraction of the contrast is caused by electrons absorbed inside the sample. Instead, the atoms distort the phase of the wave front, so that the image contrast is mostly generated by interference effects. Over time, different methods have been developed to optimize the information transfer in electron microscopy; this includes defocusing, phase-plates and electron holography [8].

The development of theories describing the interaction of electrons with matter started right after Davisson and Germer conducted their famous experiment in the years 1923 – 1927 [11]. With this electron-scattering experiment, they confirmed the wave-like nature of the electron. Even before the successful realization of the first electron microscope, it was recognized that the kinematic theory of scattering is insufficient to describe electron diffraction. Already in 1928, Bethe used the Schrödinger equation and what is now called Bloch waves to study dynamical scattering in crystals [12][d].

In 1957, Cowley and Moodie developed, in a series of papers starting with [14], a different description for the scattering process in transmission electron microscopy. This so-called multislice approach was derived from optical principles: A thick sample is treated as a series of thin slices. The potential in each slice is reduced to a two-dimensional phase-grating. The propagation of the wave between the phase-gratings is described by Fresnel diffraction. From this theoretical description of the scattering process, many practical computational schemes emerged. It was shown in 1961 by Fujiwara that the most important relativistic corrections can be included by using a modified Schrödinger equation [15]. In 1974, Goodman and Moodie summarized different descriptions for the scattering process in the electron microscope [16]. Moreover, they showed the relation between the different approaches, as well as the Schrödinger equation and Feynman's path integral formalism. The performance of the multislice algorithm was significantly improved by the use of the fast Fourier transform (FFT). This was independently proposed by Ishizuka and Uyeda [17] and by Bursill and Wilson [18] in 1977. While the Bloch-wave approach can be used for theoretical investigations, the multislice algorithm is only suited for numerical evaluations. Since it does not require the sample to be periodic, the multislice method can be directly applied to the study of defects, interfaces and amorphous samples.

Cowley and Spence demonstrated the possibility of simulating convergent beam electron diffraction (CBED, cf. section 5.3.2) patterns with the multislice method in 1978 [19]. The simulation of scanning transmission electron microscope (STEM, cf. section 2.1) images is computationally very intensive, since the CBED calculation must be repeated for every pixel. The first application of the multislice method for the simulation of STEM images was reported by Kirkland et al. in 1987 [20].

Backscattering effects were considered by Lynch and Moodie in 1972 [21]. A different formulation was given by Chen and Van Dyck in 1997 [22]. Very recently, a first attempt was made to implement the corresponding algorithm by Spiegelberg and Rusz [23].

1.2 Advanced Computational Algorithms

The amount of information that can be gained with the help of electron microscopy can be increased by the use of more complex image analyses and modern statistical algorithms. Like many other fields, electron microscopy thus profits from the steady increase of the available

[d] It is interesting to note, that Knoll and Ruska were not aware about the wave nature of electrons when they built the first electron microscope [13].

computing power. For example, it was recently shown that graphics processing units (GPUs) can be used to accelerate image simulations [24, 25].

By employing sophisticated processing algorithms, it has been demonstrated that the three-dimensional atomic structure of a specimen can be determined from a small number of electron micrographs [26, 27]. Furthermore, single-particle reconstruction in cryo electron microscopy could be improved by using Bayes' theorem for a distributed statistical analysis [28].

The computing power that is now available gives us the opportunity to review and reevaluate different image simulation algorithms. It is no longer necessary to use questionable approximations for a minor gain in speed. For example, the commonly used paraxial approximation is inferior to the more advanced techniques presented in chapter 5. Moreover, a direct numerical integration of the Schrödinger equation has previously been deemed too expensive; however, here we show that algorithms of this type can be used for the simulation of CBED patterns (cf. section 5.3.2).

Micrographs for scanning electron microscopes (SEMs) are typically calculated with the help of Monte-Carlo simulations based on a semiclassical model. In this work, we present the first simulation results for the elastic scattering channel based on a rigorous quantum mechanical treatment of the scattering process.

In chapter 2, a short introduction to electron microscopy will be given. Furthermore, the basic relations for electron scattering will be summarized. In chapter 3, the classical multislice algorithm will be derived from quantum mechanics. Afterward, electron scattering will be treated as an eigenvalue problem (chapter 4). In chapter 5, Runge-Kutta methods will be introduced and a detailed analysis of the paraxial approximation will be given. Finally, backscattering effects will be included in the image simulations (chapter 6).

Chapter 2

Foundations

In this chapter, the basic concepts for the description of the image formation process in an electron microscope are introduced.

Most of the derivations presented here are well-known as they appear either in introductory texts about quantum mechanics and solid-state physics or in the books by Kirkland [29], Reimer [8] and Wang [30]. Besides forming the basis for later discussions, the calculations of this chapter serve the purpose of introducing the notation.

2.1 Electron Microscopy

There is a great number of different types of electron microscopes that can be operated in various modes. At first, we take a look at the bright-field (BF) mode of a TEM (cf. figure 2.1).

In this case, the sample is illuminated by an extended electron beam that is formed by the condenser lens system. The beam is quasi-monochromatic and can be approximately described by a plane wave moving along the direction of the optical axis of the microscope. The initial kinetic energy of the electrons is determined by the applied acceleration voltage. Through its interaction with the sample, the electron wave is scattered in different directions. Thereby, the electron might suffer an energy loss. The electron wave in a plane directly behind the sample is described by the exit-wave function ψ_{ex}. With the help of the objective lens and the projective lens system, a magnified image of the exit-wave function can be created on the detector, usually a camera. A simple mathematical description for the imaging process will be discussed in section 2.6.

Alternatively, the projective lens system can be used to image the back focal plane. In this instance, much more information can be gained if the illuminating beam is focused to a small spot (CBED) [2]. This is further detailed in section 5.3.2.

In a STEM, a convergent electron beam is scanned over a certain area of the sample. Hence, a CBED pattern is formed for every scan position. The pixels of the STEM micrograph reflect the signal strengths measured by the detectors for the corresponding beam position. Roughly speaking, the BF detector measures the total beam minus the intensity that is scattered outside of its detection area. The annular dark-field (ADF) detector is complementary to the BF detector. It is a ring-shaped detector that allows the unscattered beam to pass. For a certain angular range, it thus measures the integrated intensity of the scattered electrons. Without a sample, the ADF signal vanishes.

An SEM is conceptually very similar to a STEM, but it employs much lower acceleration voltages (typically less than 30 kV). As a consequence, the electrons can penetrate only the

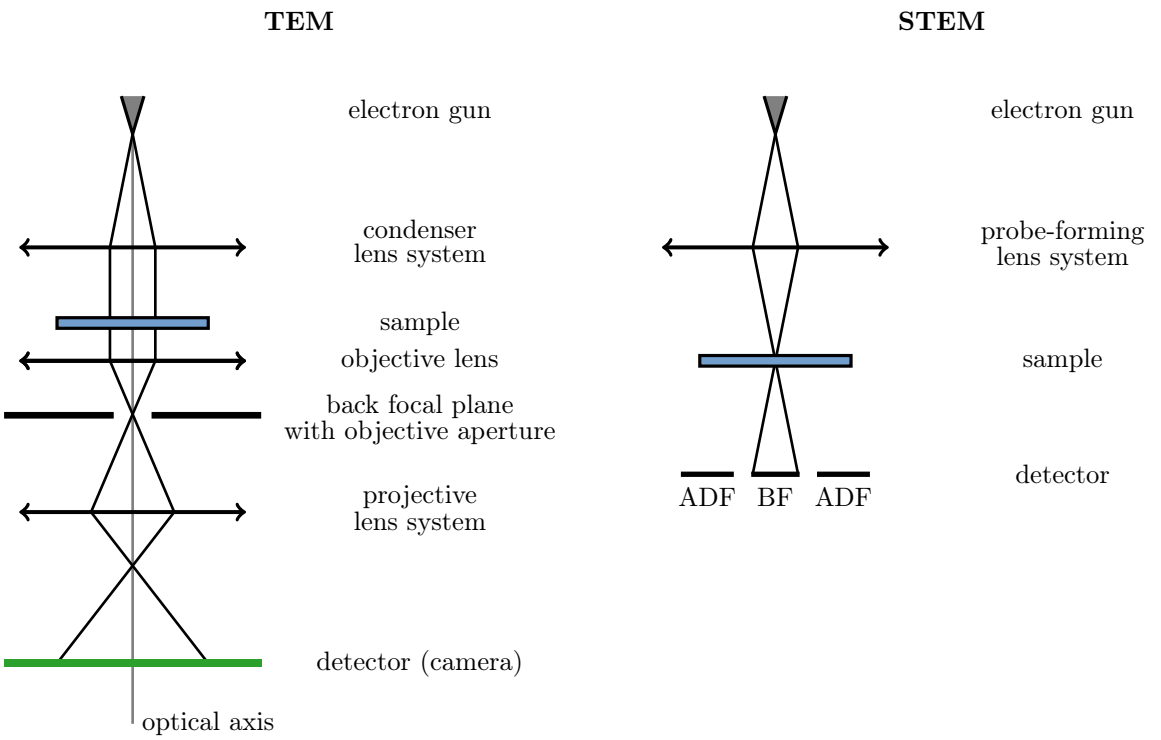

Figure 2.1: Schematics showing the most important optical components of a TEM and a STEM. Additionally, the path of the unscattered electron beam is illustrated.

surface of a sample. Hence, instead of the transmitted signal, backscattered and secondary electrons, generated by ionization processes, are analyzed.

In the main part of this thesis, we will discuss the interaction of the electron beam with a sample in great detail. Generally, these interactions can be quite complicated because many different processes are involved, e.g. elastic scattering, excitation of phonon modes, plasmon scattering or ionization. The relative importance of the different processes depends greatly on the acceleration voltage and the thickness of the sample. In this work, we will mainly discuss elastic scattering. Experimentally, inelastically scattered electrons can be excluded from the image formation process by employing an energy filter. Since the excitation energy for a phonon is below 1 eV, these processes always contribute to the final image. Hence, they must be treated by any realistic image simulation.

2.2 Equations of Motion for the Electron

If we consider only elastic scattering effects, then the sum of the kinetic energy of the incoming electron and the kinetic energy of the scattering target as a whole is the same for the initial and the final state. In other words, no inner degrees of freedom are excited. Therefore, the interaction of the electron and the target can be described using time-independent potentials. In the laboratory system (the microscope system), the momentum transfer from the electron to the target is always accompanied by a transfer of energy. Although this transfer is tiny due to the large mass ratio of the target and the electron, the kinetic energy of the electron always changes during an elastic scattering process. Nevertheless, in the following discussion this effect will be

neglected.

A very accurate description of the electron is provided by quantum electrodynamics. For our purpose, however, a much simpler theory based on the Klein-Gordon equation is sufficient. Thus, the electron will be regarded as a spin-less particle. Moreover, we will assume that the kinetic and the potential energy of the electron are small compared to its rest mass ($m \approx 511\,\mathrm{keV}/c^2$, c: speed of light).

According to the relativistic energy-momentum relation, the total energy E of a free electron is given by

$$E^2 = m^2\,c^4 + c^2\,\mathbf{p}^2. \tag{2.1}$$

Here, $\mathbf{p}$ denotes the momentum of the electron. As we consider only elastic scattering effects, the total energy is fixed and given by the acceleration voltage.

The canonical quantization procedure can be used to turn (2.1) into the free Klein-Gordon equation

$$(\mathrm{i}\,\hbar\,\partial_t)^2\,|\Psi(t)\rangle = \left(m^2\,c^4 + c^2\,\mathbf{p}^2\right)|\Psi(t)\rangle\,, \tag{2.2}$$

where $\mathbf{p}$ is now the momentum operator acting on the time-dependent state vector $|\Psi(t)\rangle$ and $\hbar$ is the reduced Planck constant. A more rigorous derivation of the Klein-Gordon equation is, for example, given in [31].

The Klein-Gordon equation (2.2) does not yet include the interaction of the electron with the electromagnetic field. In a typical TEM, the sample is placed in the magnetic field of the objective lens. This field can be regarded as homogeneous on the scale of the sample dimensions. Therefore, it is a good approximation to neglect the magnetic field of the objective lens [29]. Since only non-magnetic materials will be considered in this work, we refrain from the inclusion of the vector potential. The effects of the electrostatic potential ϕ are incorporated by using minimal coupling

$$(\mathrm{i}\,\hbar\,\partial_t + \underbrace{e\,\phi}_{-V})^2\,|\Psi(t)\rangle = \left(m^2\,c^4 + c^2\,\mathbf{p}^2\right)|\Psi(t)\rangle\,. \tag{2.3}$$

In this equation, $e > 0$ denotes the elementary charge.

Since the potential V is assumed to be weak, (2.3) can be simplified. For this purpose, we use a similar approach as presented in [32]. In a first step, the phase of the state $|\Psi(t)\rangle$ is redefined as

$$|\Psi(t)\rangle = \exp\left(\mathrm{i}\,\frac{m\,c^2}{\hbar}\,t\right)|\psi(t)\rangle\,. \tag{2.4}$$

After some calculations, we arrive at

$$\left[\mathrm{i}\,\hbar\,\partial_t - \frac{\hbar^2\,\partial_t^2}{2\,m\,c^2}\right]|\psi(t)\rangle = \left[\frac{\mathbf{p}^2}{2\,m} + V\left(1 + \frac{\mathrm{i}\,\hbar\,\partial_t}{m\,c^2}\right) - \frac{V^2}{2\,m\,c^2}\right]|\psi(t)\rangle\,. \tag{2.5}$$

In the non-relativistic limit ($c \to \infty$), the conventional Schrödinger equation is obtained

$$\mathrm{i}\,\hbar\,\partial_t\,|\psi(t)\rangle = \left[\frac{\mathbf{p}^2}{2\,m} + V\right]|\psi(t)\rangle\,. \tag{2.6}$$

For weak potentials ($|V| \ll 2\,m\,c^2$), the term $\propto V^2$ in (2.5) may be neglected

$$\left[\mathrm{i}\,\hbar\,\partial_t - \frac{\hbar^2\,\partial_t^2}{2\,m\,c^2}\right]|\psi(t)\rangle = \left[\frac{\mathbf{p}^2}{2\,m} + V\left(1 + \frac{\mathrm{i}\,\hbar\,\partial_t}{m\,c^2}\right)\right]|\psi(t)\rangle\,. \tag{2.7}$$

The time dependence of the state $|\psi(t)\rangle$ can be separated

$$|\psi(t)\rangle = \exp\left(-\frac{\mathrm{i}\,T\,t}{\hbar}\right)|\psi\rangle\,, \tag{2.8}$$

where $T = E - m\,c^2$ is the kinetic energy of the electron. In principle, an arbitrary time-dependent phase factor could be used, but only the kinetic energy ensures that the correct non-relativistic limit is obtained.

Substituting $|\psi(t)\rangle$ in (2.7) results in

$$\frac{\hbar^2\,k^2}{2\,m}\,|\psi\rangle = \left[\frac{\mathbf{p}^2}{2\,m} + \gamma\,V\right]|\psi\rangle \tag{2.9}$$

$$\text{with } k^2 = \frac{2\,m\,c^2\,T + T^2}{c^2\,\hbar^2}. \tag{2.10}$$

The first equation resembles the stationary Schrödinger equation for a non-relativistic particle. However, the potential V is scaled by the Lorentz factor

$$\gamma = \frac{E}{m\,c^2} = 1 + \frac{T}{m\,c^2}. \tag{2.11}$$

Moreover, instead of the electron energy, the relativistic dispersion relation (2.10) appears on the left-hand side of (2.9).

The energy dependence of the Lorentz factor is shown in figure 1.1. The Lorentz factor can be used to calculate the speed of the electron

$$v = \sqrt{1 - \gamma^{-2}}\,c. \tag{2.12}$$

It is sometimes useful to express (2.10) in terms of the total energy E

$$k^2 = \frac{E^2 - m^2\,c^4}{c^2\,\hbar^2}. \tag{2.13}$$

For the interpretation of k, we employ the position representation of (2.9)

$$-\frac{\hbar^2\,\Delta}{2\,m}\,\psi(\mathbf{x}) + \gamma\,V(\mathbf{x})\,\psi(\mathbf{x}) = \frac{\hbar^2\,k^2}{2\,m}\,\psi(\mathbf{x}). \tag{2.14}$$

Here, $\psi(\mathbf{x}) = \langle\mathbf{x}|\psi\rangle$ denotes the wave function and Δ is the three-dimensional Laplace operator. In free space, the potential $V(\mathbf{x})$ vanishes. A possible solution of (2.14) is then a plane wave in $\mathbf{k}$-direction

$$\psi(\mathbf{x}) \propto \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}. \tag{2.15}$$

Hence, k is the wave number of that wave

$$|\mathbf{k}| = k = \frac{2\,\pi}{\lambda}. \tag{2.16}$$

Figure 1.1 shows the wavelength λ of the electron as a function of the kinetic energy. Note, that electron microscopists often use a different definition for k that does not include the factor $2\,\pi$.

Since we will consider only the stationary equation (2.14), the divergence of the probability current density $\mathbf{j}$ must be zero

$$\nabla\cdot\mathbf{j} = 0 \tag{2.17}$$

$$\text{with } \mathbf{j} = \frac{\hbar}{m}\,\mathrm{Im}\left(\psi^*\,\nabla\,\psi\right). \tag{2.18}$$

2.3 Potential Scattering

We will now use the Schrödinger equation (2.14) to investigate the scattering solutions for a localized potential. Thus, the potential shall vanish for large distances x

$$\lim_{|\mathbf{x}|\to\infty} V(\mathbf{x}) = 0. \tag{2.19}$$

In this case, scattering solutions can be constructed from unbound states corresponding to $k^2 > 0$.

2.3.1 The Lippmann-Schwinger Equation

We want to recast the differential equation (2.14) as an integral equation. For this purpose, (2.14) is rewritten as

$$\left(\Delta + k^2\right) \psi(\mathbf{x}) = v(\mathbf{x})\, \psi(\mathbf{x}), \tag{2.20}$$

where we have conveniently defined a reduced potential

$$v(\mathbf{x}) = \frac{2\,m\,\gamma}{\hbar^2}\, V(\mathbf{x}). \tag{2.21}$$

For large x, the potential can be neglected so that the wave function fulfills a homogeneous Helmholtz equation

$$\left(\Delta + k^2\right) \psi(\mathbf{x}) = 0. \tag{2.22}$$

Possible solutions of this equation include plane waves

$$\phi_{\mathbf{k}_i}(\mathbf{x}) = \langle \mathbf{x} | \mathbf{k}_i \rangle = \frac{\mathrm{e}^{\mathrm{i}\,\mathbf{k}_i \cdot \mathbf{x}}}{(2\,\pi\,\hbar)^{3/2}}. \tag{2.23}$$

The vector $\mathbf{k}_i$ describes the direction of the wave and must have length k. Fourier synthesis allows forming an arbitrary wave packet consisting of a superposition of plane waves. Therefore, it is sufficient to only consider plane waves, although they can only be approximately realized in an experiment.

The solution of the inhomogeneous equation (2.20) can be found with the help of a Green's function G_k, which shall fulfill

$$\left(\Delta_{\mathbf{x}} + k^2\right) G_k(\mathbf{x}, \mathbf{y}) = \delta^{(3)}(\mathbf{x} - \mathbf{y}). \tag{2.24}$$

Hence, for an incoming plane wave $\phi_{\mathbf{k}_i}$ in direction $\mathbf{k}_i$, we can rewrite the scattering equation (2.20) as an integral equation

$$\psi_{\mathbf{k}_i}(\mathbf{x}) = \phi_{\mathbf{k}_i}(\mathbf{x}) + \int G_k(\mathbf{x}, \mathbf{y})\, v(\mathbf{y})\, \psi_{\mathbf{k}_i}(\mathbf{y})\, \mathrm{d}^3 y, \tag{2.25}$$

If not otherwise specified, integrals are meant to be taken over all space. The correctness of the last equation, can be proven by applying the operator $\Delta_{\mathbf{x}} + k^2$.

As we are only interested in the retarded Green's function, we find

$$G_k^{\mathrm{ret}}(\mathbf{x}, \mathbf{y}) = -\frac{\mathrm{e}^{\mathrm{i}\,k\,r}}{4\,\pi\,r} \qquad \text{with } r = |\mathbf{x} - \mathbf{y}|\,. \tag{2.26}$$

Substituting this result into (2.25), we arrive at the Lippmann-Schwinger equation

$$\psi_{\mathbf{k}_i}(\mathbf{x}) = \phi_{\mathbf{k}_i}(\mathbf{x}) - \frac{1}{4\,\pi} \int \frac{\mathrm{e}^{\mathrm{i}\,k|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x} - \mathbf{y}|}\, v(\mathbf{y})\, \psi_{\mathbf{k}_i}(\mathbf{y})\, \mathrm{d}^3 y. \tag{2.27}$$

The integral equation (2.27) is not the solution of the scattering problem since the unknown wave function $\psi_{\mathbf{k}_i}$ appears on both sides of the equation. However, this formulation of the scattering problem has the advantage that the boundary conditions are built in.

2.3.2 Scattering Amplitude and Cross Section

We assume that the potential $v(\mathbf{y})$ contributes substantially only for a region $y < a$. Thus, a denotes the extension of the potential. In the far-field ($x \gg a$), it is then possible to approximate the denominator of the integrand in (2.27): $|\mathbf{x} - \mathbf{y}| = |\mathbf{x}| = x$. Since the extension of the potential is in general large compared to the wavelength ($a \gg \lambda$), the exponential function, however, cannot be approximated by a constant. To describe the oscillations, at least the linear term in $\mathbf{y}$ is required

$$r = |\mathbf{x} - \mathbf{y}| = x\sqrt{1 - 2\,\frac{\mathbf{x}\cdot\mathbf{y}}{x^2} + \frac{y^2}{x^2}} = x\left(1 - \frac{\mathbf{x}\cdot\mathbf{y}}{x^2} + \mathcal{O}\left(\frac{y}{x}\right)^2\right). \tag{2.28}$$

Hence, we find the asymptotic representation of the wave function

$$\psi_{\mathbf{k}_i}(\mathbf{x}) \simeq \phi_{\mathbf{k}_i}(\mathbf{x}) - \frac{\mathrm{e}^{\mathrm{i}\,k\,x}}{4\,\pi\,x}\int \mathrm{e}^{-\mathrm{i}\,\mathbf{k}_f\cdot\mathbf{y}}\,v(\mathbf{y})\,\psi_{\mathbf{k}_i}(\mathbf{y})\,\mathrm{d}^3y = \frac{1}{(2\,\pi\,\hbar)^{3/2}}\left(\mathrm{e}^{\mathrm{i}\,\mathbf{k}_i\cdot\mathbf{x}} + f(\mathbf{k}_i,\mathbf{k}_f)\,\frac{\mathrm{e}^{\mathrm{i}\,k\,x}}{x}\right) \tag{2.29}$$

$$\text{with } \mathbf{k}_f = k\,\frac{\mathbf{x}}{x}. \tag{2.30}$$

The here introduced vector $\mathbf{k}_f$ describes the direction of the scattered particle. We realize that the wave function for the scattering problem is a combination of a plane wave and a spherical wave. The function

$$f(\mathbf{k}_i,\mathbf{k}_f) = -\frac{(2\,\pi\,\hbar)^{3/2}}{4\,\pi}\int \mathrm{e}^{-\mathrm{i}\,\mathbf{k}_f\cdot\mathbf{r}}\,v(\mathbf{r})\,\psi_{\mathbf{k}_i}(\mathbf{r})\,\mathrm{d}^3r \tag{2.31}$$

is called scattering amplitude. It depends on the directions of the incoming ($\mathbf{k}_i$) and the scattered particle ($\mathbf{k}_f$).

The most important quantity to describe the outcome of a scattering experiment is the differential cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$. It is defined as the area that is hit by the same number of particles as the number of particles scattered in a certain direction, i.e.

$$\mathrm{d}\sigma = \frac{\text{number of particles scattered into } \mathrm{d}\Omega \text{ per time interval}}{\text{current density of the incoming particles}} = \frac{j_f\,r^2\,\mathrm{d}\Omega}{j_i}. \tag{2.32}$$

Using the definition (2.18), we find for the probability current density of the incoming particles (plane wave $\phi_{\mathbf{k}_i}$)

$$\mathbf{j}_i = \frac{\mathbf{k}_i}{8\,\pi^3\,\hbar^2\,m}. \tag{2.33}$$

In the far field, the current density for the scattered wave (spherical wave) is given by

$$\mathbf{j}_f = -\frac{|f|^2\;k}{8\,\pi^3\,\hbar^2\,m}\,\frac{\mathbf{x}}{x^3}. \tag{2.34}$$

The incoming and scattered waves do not interfere as they pass the same point in space at different times (except near the forward direction). Therefore, the scattering cross section is given by the absolute square of the scattering amplitude

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f|^2. \tag{2.35}$$

The total cross section can be found by integrating the differential cross section over the full solid angle

$$\sigma_{\mathrm{tot}} = \int_{4\,\pi}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\Omega = \int_{4\,\pi}|f|^2\,\mathrm{d}\Omega. \tag{2.36}$$

2.3.3 Neumann Series and First Born Approximation

The Lippmann-Schwinger equation (2.27) can be used to calculate approximations for the scattering amplitude $f(\mathbf{k}_i, \mathbf{k}_f)$. For this purpose, we introduce an integral operator I:

$$I\,\psi(\mathbf{x}) = -\frac{1}{4\,\pi}\int \frac{\mathrm{e}^{\mathrm{i}\,k|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}\,v(\mathbf{y})\,\psi(\mathbf{y})\,\mathrm{d}^3y. \tag{2.37}$$

With the help of this operator, the Lippmann-Schwinger equation (2.27) can be cast into a much simpler form

$$\psi_{\mathbf{k}_i}(\mathbf{x}) = \phi_{\mathbf{k}_i}(\mathbf{x}) + I\,\psi_{\mathbf{k}_i}(\mathbf{x}). \tag{2.38}$$

If we assume, that the potential is sufficiently weak, then the geometrical series can be used to rewrite this equation

$$\psi_{\mathbf{k}_i}(\mathbf{x}) = (1-I)^{-1}\,\phi_{\mathbf{k}_i}(\mathbf{x}) = \left(\sum_{\ell=0}^{\infty} I^\ell\right)\phi_{\mathbf{k}_i}(\mathbf{x}). \tag{2.39}$$

In this context, the series is known as Neumann series or Born series.

The single terms of the series can be interpreted as repeated interactions of the electron with the scatterer: After a first interaction with the potential, the outgoing waves interfere with each other. Subsequently, these waves interact with the potential again.

If only the leading terms of the series are kept, one obtains the so-called first Born approximation

$$\psi_{\mathbf{k}_i}(\mathbf{x}) = \frac{1}{(2\,\pi\,\hbar)^{3/2}}\left(\mathrm{e}^{\mathrm{i}\,\mathbf{k}_i\cdot\mathbf{x}} - \frac{1}{4\,\pi}\int \frac{\mathrm{e}^{\mathrm{i}\,k|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}\,v(\mathbf{y})\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}_i\cdot\mathbf{y}}\mathrm{d}^3y\right). \tag{2.40}$$

The same steps that led to the scattering amplitude (2.31) can be applied here and result in the first Born approximation for the scattering amplitude

$$f^B(\mathbf{k}_i, \mathbf{k}_f) = -\frac{1}{4\,\pi}\int \mathrm{e}^{\mathrm{i}(\mathbf{k}_i-\mathbf{k}_f)\cdot\mathbf{r}}\,v(\mathbf{r})\,\mathrm{d}^3r. \tag{2.41}$$

Hence, in this approximation, the scattering amplitude is essentially the Fourier transform of the potential with respect to $\mathbf{k}_i - \mathbf{k}_f$. Thus, f^B is a function of the momentum transfer $\hbar\,\mathbf{K} = \hbar\,(\mathbf{k}_f - \mathbf{k}_i)$ only:

$$f^B(\mathbf{K}) = f^B(\mathbf{k}_i, \mathbf{k}_f). \tag{2.42}$$

For a central potential $v(\mathbf{y}) = v(y)$, the scattering amplitude in first Born approximation depends only on the magnitude of the momentum transfer $K = |\mathbf{K}|$. This can be demonstrated by performing the angular integration in (2.41). The angle between $\mathbf{K}$ and $\mathbf{r}$ will be called α

$$f^B(\mathbf{K}) = -\frac{1}{4\,\pi}\int \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\,v(r)\,\mathrm{d}^3r = -\frac{1}{2}\int_0^\infty v(r)\left\{\int_{-1}^{1}\mathrm{e}^{-\mathrm{i}\,K\,r\,\cos\alpha}\,\mathrm{d}\cos\alpha\right\}r^2\,\mathrm{d}r \tag{2.43}$$

$$\Rightarrow f^B(K) = -\int_0^\infty v(r)\,\frac{\sin K\,r}{K\,r}\,r^2\,\mathrm{d}r. \tag{2.44}$$

The angle between $\mathbf{k}_i$ and $\mathbf{k}_f$ is the scattering angle ϑ, which is given by

$$\cos\vartheta = \frac{\mathbf{k}_i \cdot \mathbf{k}_f}{k^2}. \tag{2.45}$$

We can use the scattering angle to rewrite K in terms of k and ϑ

$$K = \sqrt{(\mathbf{k}_f - \mathbf{k}_i)^2} = 2\,k\,\sin\vartheta/2. \tag{2.46}$$

Therefore, the scattering amplitude in first Born approximation and the corresponding cross section depend only on the scattering angle (and the wavelength).

The representation (2.44) reveals that the scattering amplitude in first Born approximation is real-valued. This contradicts the optical theorem [33], which relates the total scattering cross section to the imaginary part of the forward scattering amplitude

$$\sigma_{\text{tot}} = \frac{4\,\pi}{k}\,\text{Im}\,f(0). \tag{2.47}$$

Since the optical theorem is a consequence of the continuity equation for the probability, we must conclude that the first Born approximation does not conserve probability.

The expression (2.44) can be employed to derive the scattering amplitude for a screened Coulomb potential (also known as Yukawa or Wentzel potential [34])

$$V(r) = -\frac{Z\,e^2}{4\,\pi\,\varepsilon_0}\frac{1}{r}\,\mathrm{e}^{-r/R}, \tag{2.48}$$

where Z denotes the atomic number, R the screening radius and ε_0 the vacuum permittivity. We obtain the scattering amplitude

$$f^B(K) = \frac{Z\,m\,e^2\,\gamma}{2\,\pi\,\varepsilon_0\,\hbar^2\,K}\int_0^\infty \mathrm{e}^{-r/R}\,\sin K\,r\,\mathrm{d}r = \frac{Z\,m\,e^2\,\gamma}{2\,\pi\,\varepsilon_0\,\hbar^2}\frac{1}{K^2 + R^{-2}}. \tag{2.49}$$

In the limiting case $R \to \infty$, we get the amplitude for the Coulomb potential

$$f_k^B(\vartheta) = \frac{Z\,e^2}{16\,\pi\,\varepsilon_0\,T}\frac{1}{\sin^2\vartheta/2}, \tag{2.50}$$

where we have set $\gamma = 1$ and used the non-relativistic expression for the kinetic energy

$$T = \frac{\hbar^2\,k^2}{2\,m}. \tag{2.51}$$

Although this calculation is based on the first Born approximation, the absolute square of (2.50) gives the exact cross section for the Coulomb potential, i.e. the Rutherford cross section.

2.3.4 Atomic Form Factors

The screening radius is a free parameter of the Wentzel potential. By employing the statistical Thomas-Fermi model [35], a reasonable value for the screening radius can be found:

$$R = \frac{a_B}{\sqrt[3]{Z}}. \tag{2.52}$$

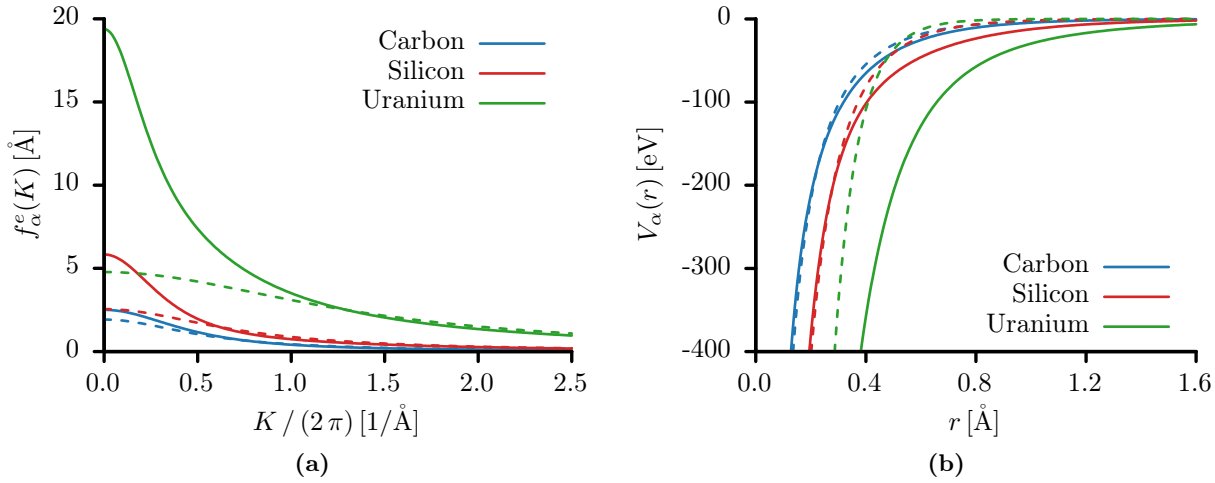

Figure 2.2: Electronic form factor (a) and the corresponding potential (b) for three elements (carbon, silicon and uranium). The simple Wentzel model (dashed, (2.48) and (2.52)) and the parametrization of Lobato and Van Dyck (solid, appendix A.2) are compared.

This can be used to calculate the elastic scattering cross section for a single atom. The result, however, is a very poor approximation (cf. figure 2.2). In order to improve it, more precise information about the electron distribution is required.

Only for the hydrogen atom, the electron wave function is known analytically. For all other atoms, numerical algorithms like the Thomas-Fermi-Dirac method or the (relativistic) Hartree-Fock method must be used. An overview is given by Kirkland [29].

Usually, so-called atomic form factors are published in place of the atomic potentials. Thus, the relation between these different quantities shall be discussed now.

The electronic form factor $f^e_\alpha(\mathbf{K})$ for an atom α is closely related to the first Born approximation of the scattering amplitude (2.41). For a given atomic potential $V_\alpha(\mathbf{r})$ the electronic form factor is defined as

$$f^e_\alpha(\mathbf{K}) = -\frac{m}{2\,\pi\,\hbar^2} \int \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\, V_\alpha(\mathbf{r})\,\mathrm{d}^3 r. \tag{2.53}$$

Note that the Lorentz factor γ is not included in this definition. Since the electronic form factor is essentially the Fourier transform of the potential, we obtain the inverse relation

$$V_\alpha(\mathbf{r}) = -\frac{\hbar^2}{4\,\pi^2\,m} \int \mathrm{e}^{\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\, f^e_\alpha(\mathbf{K})\,\mathrm{d}^3 K. \tag{2.54}$$

The potential of a free atom must be spherically symmetric. Therefore, the last equation can be reduced to a one-dimensional integral (cf. (2.44))

$$V_\alpha(r) = -\frac{\hbar^2}{\pi\,m\,r} \int_0^\infty K\,\sin(K\,r)\, f^e_\alpha(K)\,\mathrm{d}K. \tag{2.55}$$

In contrast, photons are not scattered by the atomic potential but by the electron charge density $\rho^e_\alpha(\mathbf{r})$ [36]. The charge of the nucleus is not relevant as its contribution is suppressed by the inverse of the nucleus' mass squared according to the Thomson scattering cross section. The

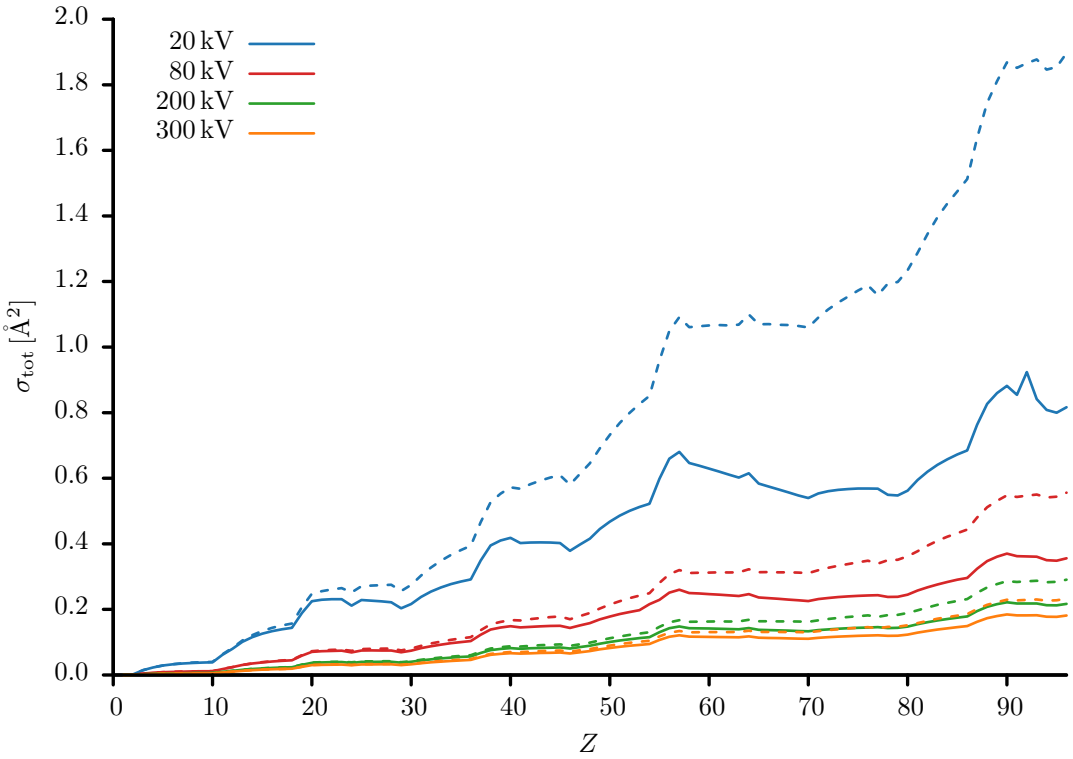

Figure 2.3: Total elastic electron scattering cross section for different energies. Shown is the cross section as published by NIST (solid, [37]) and the cross section based on the first Born approximation using the form factor parametrization of Lobato and Van Dyck (dashed, [38]).

X-ray form factor is thus defined as

$$f_\alpha^\gamma(\mathbf{K}) = -\int \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}} \rho_\alpha^e(\mathbf{r})\,\mathrm{d}^3 r \tag{2.56}$$

$$f_\alpha^\gamma(K) = -\frac{4\,\pi}{K}\int_0^\infty r\,\sin(K\,r)\,\rho_\alpha^e(r)\,\mathrm{d}r. \tag{2.57}$$

In the second line, we have exploited the spherical symmetry of the charge density.

Since an atom needs to be electrically neutral, the electron charge density fulfills

$$\int \rho_\alpha^e(\mathbf{r})\,\mathrm{d}^3 r = 4\,\pi\int_0^\infty r^2\,\rho_\alpha^e(r)\,\mathrm{d}r = -Z_\alpha\,e. \tag{2.58}$$

Together with L'Hôpital's rule, this allows the calculation of the limiting case $K \to 0$ of (2.57)

$$f_\alpha^\gamma(0) = -4\,\pi\int_0^\infty r^2\,\rho_\alpha^e(r)\,\mathrm{d}r = Z_\alpha\,e. \tag{2.59}$$

The charged density can be expressed in terms of the X-ray form factor by inverting (2.56)

$$\rho_\alpha^e(\mathbf{r}) = -\frac{1}{(2\,\pi)^3}\int \mathrm{e}^{\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\,f_\alpha^\gamma(\mathbf{K})\,\mathrm{d}^3 K. \tag{2.60}$$

A close relationship exists between the electronic form factor and the X-ray form factor. It can be derived from Poisson's equation of electrostatics

$$\Delta \frac{V_\alpha(\mathbf{r})}{e} = \frac{\rho_\alpha(\mathbf{r})}{\varepsilon_0}, \tag{2.61}$$

where ρ_α denotes the total charge density including the contributions from the nucleus $Z_\alpha\, e\, \delta^{(3)}(\mathbf{r})$. Hence, we find

$$\Delta V_\alpha(\mathbf{r}) = \frac{e}{\varepsilon_0}\left(Z_\alpha\, e\, \delta^{(3)}(\mathbf{r}) + \rho_\alpha^e(\mathbf{r})\right). \tag{2.62}$$

Using the Fourier representation of the delta function

$$\delta^{(n)}(\mathbf{r}) = \frac{1}{(2\,\pi)^n} \int \mathrm{e}^{\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\, \mathrm{d}^n K \tag{2.63}$$

and the relations (2.54) and (2.60), we can rewrite (2.62) in reciprocal space

$$f_\alpha^e(\mathbf{K}) = \frac{2}{a_B\, e\, K^2}\left(Z_\alpha\, e - f_\alpha^\gamma(\mathbf{K})\right), \tag{2.64}$$

where the Bohr radius a_B is given by

$$a_B = \frac{4\,\pi\,\varepsilon_0\,\hbar^2}{m\, e^2}. \tag{2.65}$$

The relation (2.64) is known as Bethe-Mott formula.

The Bethe-Mott formula cannot be applied directly if $K = 0$. It turns out, however, that the corresponding limit can be calculated analytically. For this purpose, we insert (2.57) into (2.64) and expand the sine function

$$\lim_{K\to 0} f_\alpha^e(K) = \lim_{K\to 0} \frac{2}{a_B\, e\, K^2}\left[Z_\alpha\, e + \frac{4\,\pi}{K}\int_0^\infty r\, \sin(K\,r)\, \rho_\alpha^e(r)\, \mathrm{d}r\right] \tag{2.66}$$

$$= \lim_{K\to 0} \frac{2}{a_B\, e\, K^2}\left[Z_\alpha\, e + 4\,\pi\int_0^\infty r^2\, \rho_\alpha^e(r)\, \mathrm{d}r - \frac{2\,\pi}{3}\, K^2 \int_0^\infty r^4\, \rho_\alpha^e(r)\, \mathrm{d}r + \mathcal{O}(K^4)\right] \tag{2.67}$$

$$\overset{(2.58)}{=} \frac{-4\,\pi}{3\, a_B\, e}\int_0^\infty r^4\, \rho_e(r)\, \mathrm{d}r \tag{2.68}$$

$$\overset{(2.58)}{=} \frac{Z_\alpha}{3\, a_B}\, \frac{4\,\pi \int_0^\infty r^4\, \rho_\alpha^e(r)\, \mathrm{d}r}{4\,\pi \int_0^\infty r^2\, \rho_\alpha^e(r)\, \mathrm{d}r} \tag{2.69}$$

$$= \frac{Z}{3\, a_B}\left\langle r^2\right\rangle. \tag{2.70}$$

Thus, the electronic form factor for $K \to 0$ is related to the mean square radius of the electron charge distribution $\left\langle r^2\right\rangle$.

As a consequence, the electronic form factor, just like the X-ray form factor, cannot diverge at the origin. Then again, there is a Coulomb singularity in the origin of the potential caused by

the atomic nucleus[a]. Therefore, the form factors can be handled more easily than the potentials, so that mostly form factors are published.

To facilitate the distribution and the further use of the calculated form factor data, so-called form factor parametrizations have been developed: The form factors are represented by certain simple functions with a small number of free parameters. These parameters are determined by a fit procedure to match the previously calculated form factor data.

The form factor parametrizations allow extrapolation of the form factor data. Therefore, it is desirable for the parametrizations to provide the correct asymptotic behavior. In this work, we will use the parametrizations by Peng et al. [39] based on Gaussian functions and the parametrizations by Lobato and Van Dyck [38] that incorporate many physical constraints. Both parametrizations are described in appendix A.

For a photon energy of 15 keV (corresponding to a wavelength of about 83 pm) the coherent scattering cross section is about $1.7 \cdot 10^{-7}\,\text{Å}^2$ for silicon and $1.5 \cdot 10^{-5}\,\text{Å}^2$ for uranium [40]. If we compare these values with the electron scattering cross sections shown in figure 2.3, we see that electrons interact much more strongly with matter than photons do. As a result, the first Born approximation, which is quite sufficient to describe X-ray diffraction, can in general not be applied to electron scattering. The comparison of the first Born approximation with the exact elastic cross section in figure 2.3 confirms that the Born approximation is only valid for weak potentials, i.e. high energies and low-Z materials.

2.4 Crystals

The structure of an idealized crystal is determined by its primitive cell. The primitive cell is a parallelepiped built from three vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$. Thus, its volume is given by the triple product

$$V_3 = |(\mathbf{a}_1 \times \mathbf{a}_2) \cdot \mathbf{a}_3| \,. \tag{2.71}$$

Translating the parallelepiped by integer multiples of these vectors fills the entire space without gaps or overlaps. The possible translation vectors can be written as

$$\mathbf{R} = \sum_{i=1}^{3} n_i \, \mathbf{a}_i, \tag{2.72}$$

where the n_i are integer numbers. This infinite set of vectors/points constitutes the three-dimensional, direct lattice L.

The atom positions need to specified only for a single primitive cell. These atoms form the so-called base of the crystal.

Periodic structures can often be conveniently described in reciprocal space. Thus, we define the base vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ of the reciprocal lattice L' by the relation

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\,\pi\,\delta_{ij}. \tag{2.73}$$

All points of the reciprocal lattice can be written as

$$\mathbf{q} = \sum_{i=1}^{3} m_i \, \mathbf{b}_i \tag{2.74}$$

[a] This is, of course, only an effective description that is valid for the length scales considered in this work.

with integer m_i.

An important equation relating the direct and reciprocal lattice is

$$\sum_{\mathbf{R}\in L} \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{R}} = V_3 \sum_{\mathbf{q}\in L'} \delta^{(3)}(\mathbf{K}-\mathbf{q}). \tag{2.75}$$

The two sums include all points of the respective lattices.

2.4.1 Kinematic Approximation

The kinematic description of electron scattering can be derived by applying the first Born approximation to the periodic potential of the crystal

$$v(\mathbf{r}) = S(\mathbf{r}) \sum_{\mathbf{R}\in L} \sum_{\substack{\alpha\in\\ \text{primitive cell}}} v_\alpha(\mathbf{r}-\mathbf{R}-\mathbf{r}_\alpha). \tag{2.76}$$

The second sum comprises all atoms of the primitive cell. The vector $\mathbf{r}_\alpha$ denotes the positions of the atoms inside the primitive cell with respect to the cell's origin. The shape function $S(\mathbf{r})$ is the characteristic function of the crystal. Note, that the crystal potential is written as a sum of independent atomic potentials, i.e. binding effects are neglected.

Substituting the potential (2.76) into (2.41), results in the following expression for the scattering amplitude

$$f^B(\mathbf{K}) = -\frac{1}{4\,\pi}\int \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\, v(\mathbf{r})\,\mathrm{d}^3r \tag{2.77}$$

$$= -\frac{1}{\sqrt{128\,\pi^5}}\,\hat{S}(\mathbf{K}) * \left[\int \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}} \sum_{\mathbf{R}\in L} \sum_{\substack{\alpha\in\\ \text{primitive cell}}} v_\alpha(\mathbf{r}-\mathbf{R}-\mathbf{r}_\alpha)\,\mathrm{d}^3r\right] \tag{2.78}$$

$$= -\frac{1}{\sqrt{128\,\pi^5}}\,\hat{S}(\mathbf{K}) * \left[\left(\sum_{\mathbf{R}\in L} \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{R}}\right) \sum_{\substack{\alpha\in\\ \text{primitive cell}}} \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}_\alpha} \int \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}}\, v_\alpha(\mathbf{r})\,\mathrm{d}^3r\right] \tag{2.79}$$

$$\overset{\substack{(2.53)\\(2.75)}}{=} -\frac{\gamma\,V_3}{\sqrt{8\,\pi^3}}\,\hat{S}(\mathbf{K}) * \left[\left(\sum_{\mathbf{q}\in L'} \delta^{(3)}(\mathbf{K}-\mathbf{q})\right) \sum_{\substack{\alpha\in\\ \text{primitive cell}}} \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}_\alpha}\, f^e_\alpha(\mathbf{K})\right] \tag{2.80}$$

$$= -\frac{\gamma\,V_3}{\sqrt{8\,\pi^3}}\,\hat{S}(\mathbf{K}) * \left[\sum_{\mathbf{q}\in L'} \delta^{(3)}(\mathbf{K}-\mathbf{q})\, g(\mathbf{K})\right] \tag{2.81}$$

$$\text{with } g(\mathbf{K}) = \sum_{\substack{\alpha\in\\ \text{primitive cell}}} \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}_\alpha} f^e_\alpha(\mathbf{K}). \tag{2.82}$$

The function $g(\mathbf{K})$ is known as structure factor. The expression $a * b$ denotes the convolution of the two functions $a(\mathbf{x})$ and $b(\mathbf{x})$

$$(a * b)(\mathbf{x}) = \int a(\mathbf{y})\, b(\mathbf{x}, -\mathbf{y})\,\mathrm{d}^3y. \tag{2.83}$$

For an infinitely extended crystal, the Fourier transformed shape function is a delta function $\hat{S}(\mathbf{K}) = \sqrt{8\,\pi^3}\,\delta^{(3)}(\mathbf{K})$ and (2.81) simplifies to

$$f^B(\mathbf{K}) = -\gamma\, V_3 \sum_{\mathbf{q}\,\in\, L'} \delta^{(3)}(\mathbf{K}-\mathbf{q})\, g(\mathbf{K}). \tag{2.84}$$

Hence, the scattering amplitude $f^B(\mathbf{K})$ is non-vanishing only if the scattering vector $\mathbf{K}$ is a vector of the reciprocal lattice.

The Ewald construction shown in figure 2.4 can be used to find the diffraction spots for a given crystal: The vector $\mathbf{k}_i$ ends at a point $\mathbf{0}$ of the reciprocal lattice. The start point of the vector $\mathbf{k}_i$ is the center of a sphere with radius k. Each intersection of the sphere with a point of the reciprocal lattice causes a diffraction spot. However, in this model, the points have no spatial extension. Thus, only very few and sharp spots would be present in the diffraction pattern.

We get a more realistic description of the diffraction process, if the Fourier transformed shape function $\hat{S}(\mathbf{K})$ of a finite-sized crystal is placed at each point of the reciprocal lattice. Since typical samples in a TEM are rather thin, the function $\hat{S}(\mathbf{K})$ often represents rod-shaped objects. An intersection of the Ewald's sphere with these extended objects is much more probable. Hence, the number of spots in the diffraction pattern is greatly increased in accordance with the experimental findings.

The points of the reciprocal lattice can be arranged in so-called Laue zones (cf. figure 2.4). All points in the plane that is perpendicular to $\mathbf{k}_i$ and contains the point $\mathbf{0}$ form the zero-order Laue zone (ZOLZ). The intersection of the Ewald's sphere with the points of the ZOLZ create the central spots of the diffraction pattern. The points of the successive planes form accordingly the first-order Laue zone (FOLZ), the second-order Laue zone (SOLZ) and so on. Together, these are called higher-order Laue zones (HOLZs). In the diffraction pattern, the points of a HOLZ appear as an isolated ring.

As we have already seen in section 2.3.4, the first Born approximation is an insufficient description for the electron scattering process. Starting with chapter 3, we will therefore discuss dynamical theories. Nevertheless, we have demonstrated in this section that general features of electron diffraction patterns can be interpreted with the help of the kinematic approximation.

2.4.2 Thermal Vibrations

Up until now, we have assumed that the atoms are arranged in a static lattice. For quantitative image simulations, it is, however, necessary to consider also the thermal vibrations of the atoms. Thus, we need a description for the so-called thermal diffuse scattering (TDS).

A simple semiclassical model – the frozen-phonon approximation – has been applied to this problem with great success. Although lattice vibrations are a time-dependent process, in this model the configuration of the crystal atoms as "seen" by a single electron is assumed to be static. This is justified, as the time the electron spends inside the crystal is much shorter than the oscillation period of the atoms. Subsequent electrons are scattered by different lattice configurations. For the calculation of the final electron micrograph, the intensities of all simulated electron waves are summed up.

Thus, electrons scattered by different lattice configurations are considered to be incoherent while the scattering with a single lattice configuration is treated as a perfectly coherent process. According to quantum mechanics, however, phonon scattering diminishes the coherence of the electron wave, since it is an inelastic scattering event. Furthermore, the time-of-flight argument is a classical one. Nevertheless, it was shown by Wang [41], that under certain conditions the frozen-phonon approximation and a correct quantum mechanical treatment yield identical results.

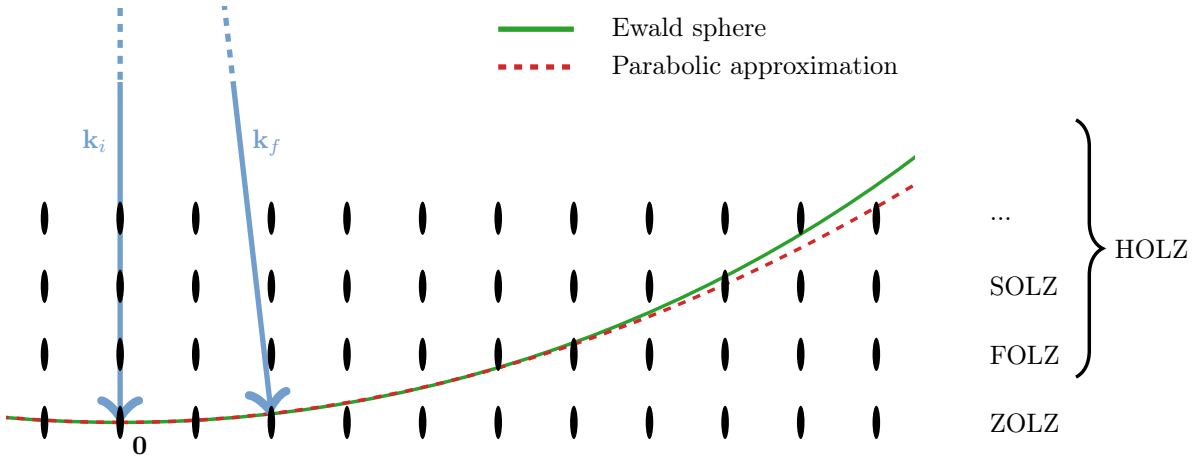

Figure 2.4: Ewald construction illustrating the formation of the electron diffraction pattern in the framework of the kinematic approximation. For an infinite crystal, a diffraction spot is present only if the Ewald's sphere intersects a point of reciprocal lattice. The finite size of the sample can be taken into account by replacing the points of the reciprocal lattice with rod-shaped objects. The parabolic approximation will be introduced in chapter 3.

As the vibrations of the atoms are considered to be independent of another, the frozen-phonon approximation is based essentially on the Einstein model.

The random displacements of the atoms destroy the symmetries of the crystals. For this reason, image simulations based on the frozen-phonon approximation are a lot more expensive in terms of computational resources. Thus, an even simpler model based on the Debye-Waller factor is often used. It can be derived by replacing the exact atom positions $\mathbf{r}_\alpha$ in (2.82) with $\mathbf{r}_\alpha + \mathbf{u}_\alpha$ where $\mathbf{u}_\alpha$ is a random displacement vector

$$g(\mathbf{K}) = \sum_{\substack{\alpha \in \\ \text{primitive cell}}} \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{r}_\alpha}\, \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{u}_\alpha} f_\alpha^e(\mathbf{K}). \tag{2.85}$$

This expression must be averaged to account for all possible displacements. As the displacements are assumed to be small, the series expansion of the exponential function can be used to rewrite the additional factor

$$\left\langle \mathrm{e}^{-\mathrm{i}\,\mathbf{K}\cdot\mathbf{u}_\alpha} \right\rangle = 1 - \mathrm{i}\,\langle \mathbf{K}\cdot\mathbf{u}_\alpha \rangle - \frac{1}{2}\left\langle (\mathbf{K}\cdot\mathbf{u}_\alpha)^2 \right\rangle + \mathcal{O}(\mathbf{K}\cdot\mathbf{u}_\alpha)^3. \tag{2.86}$$

The direction of the vector $\mathbf{u}_\alpha$ shall be completely random so that $\langle \mathbf{K}\cdot\mathbf{u}_\alpha \rangle = 0$. For the evaluation of the second-order term, we use spherical coordinates

$$\left\langle (\mathbf{K}\cdot\mathbf{u}_\alpha)^2 \right\rangle = \frac{K^2 \left\langle u_\alpha^2 \right\rangle}{4\,\pi} \int_0^{2\,\pi} \mathrm{d}\varphi \int_{-1}^{1} \cos^2\vartheta\, \mathrm{d}\cos\vartheta = \frac{K^2 \left\langle u_\alpha^2 \right\rangle}{3}. \tag{2.87}$$

Here, $\left\langle u_\alpha^2 \right\rangle$ is the mean square displacement of the atom α in three dimensions. Writing the

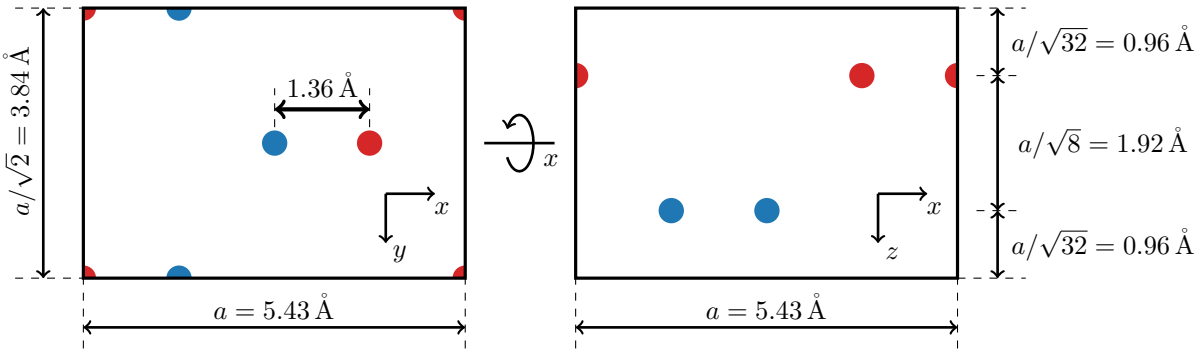

Figure 2.5: Arrangement of the silicon atoms inside the simulation box. The optical axis (z-axis) coincides with the [110]-direction of the crystal. The two atoms in the center of the xy-view (left) are approximately 1.36 Å apart. They form the so-called dumbbell structure. The xz-view (right) reveals that the two atoms are located in different z planes.

expansion (2.86) again as an exponential function, we obtain the Debye-Waller factor[b]

$$\exp\left(-B_\alpha\, K^2\right) = \exp\left(-\frac{\left\langle u_\alpha^2\right\rangle K^2}{6}\right) \approx 1 - \frac{\left\langle u_\alpha^2\right\rangle K^2}{6}. \tag{2.88}$$

For an harmonic oscillator, this expression becomes exact. The constant B_α quantifies the thermal movements of the atom α. Typical values for B_α at room temperature are about $0.003\,\text{Å}^2$ [42] causing average displacements of $\sqrt{\langle u_\alpha^2\rangle} = 0.13\,\text{Å}$. We will use a B-factor of $0.003\,\text{Å}^2$ throughout this work.

According to (2.85), the effect of the Debye-Waller factor can be included in the form factor parametrization. This fact will be exploited in appendix A. Due to the Debye-Waller factor, the intensities of the diffraction spots decrease with increasing temperature so that large scattering angles become suppressed.

2.5 Samples

In this work, silicon and graphene are utilized to investigate the properties of the different simulation algorithms.

Because of its use in the semiconductor industry, silicon is one of the best studied materials. It crystallizes in diamond structure. Thus, the lattice is face-centered cubic (fcc) with a two-atomic base. The edge length of the cubic supercell is approximately 5.43 Å.

The [110] orientation of silicon is particularly interesting because the apparent distance of the two silicon columns is only about 1.36 Å (cf. figure 2.5) and can hence be used as a benchmark for the resolution of a microscope. Figure 2.5 depicts also the minimal simulation box that respects all translational symmetries of the crystal. To ensure that all four atoms inside this box appear the same in the simulated images, they need to be placed at identical subpixel positions. Since the atoms divide the x-axis into four equidistant parts, the number of pixels in x-direction M_x must be a multiple of four. Likewise, the number of pixels in y-direction M_y needs to be a multiple of two.

[b] Note, that there are many different definitions of the Debye-Waller factor in literature with other constant factors in the exponent. Some authors call even B_α Debye-Waller factor.

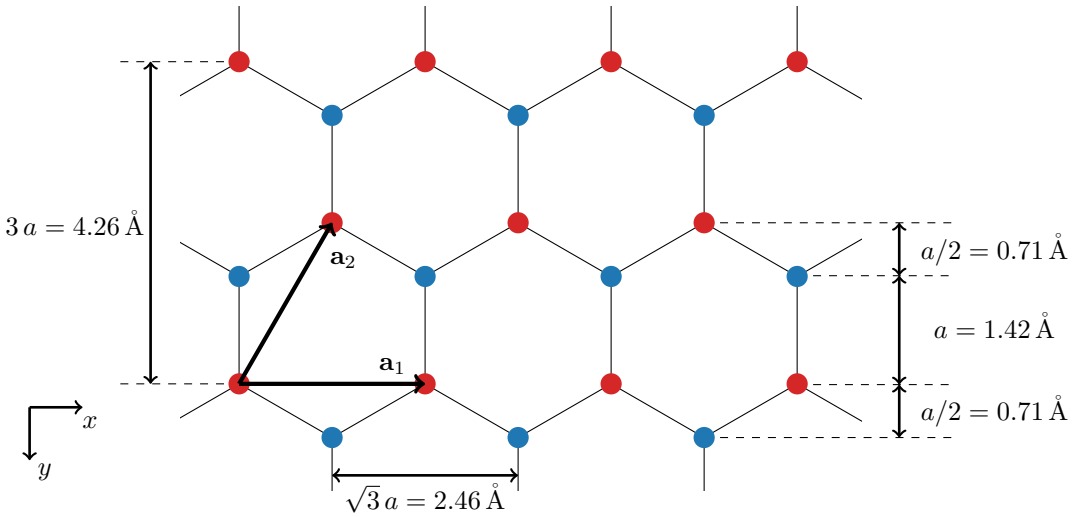

Figure 2.6: Schematic of a single graphene layer. The primitive cell is spanned by the two vector $\mathbf{a}_1$ and $\mathbf{a}_2$ that enclose an angle of 60 °. The cell contains two atoms shown in red and blue.

The second sample used for the simulations is a single layer of graphene shown in figure 2.6. Graphene is a two-dimensional crystal consisting of carbon atoms which are arranged in a hexagonal honeycomb structure. The distance between nearest neighbors is 1.42 Å. The size of the minimal simulation box is 2.46 Å × 4.26 Å. All atoms are treated in the same way during the simulation if M_x and M_y are multiples of two and six, respectively.

2.6 Optical System of a TEM

The optical system of a real microscope is very complex. For our purpose, a simple theoretical model of the optical system is nevertheless sufficient. The discussion in this section is largely based on material from [43].

Behind the objective lens, in the back focal plane of a TEM, a diffraction pattern is formed (cf. figure 2.1). The wave function in this plane is given by the Fourier transform of the exit-wave

$$\hat{\psi}_{\mathrm{ex}}(\boldsymbol{\kappa}) = \mathcal{F}\,\psi_{\mathrm{ex}}(\boldsymbol{\rho}). \tag{2.89}$$

Here, $\boldsymbol{\rho}$ and $\boldsymbol{\kappa}$ denote two-dimensional vectors in real and reciprocal space, respectively. The subsequent projective system can be used to image either the diffraction pattern or the exit-wave onto the camera. In the first case, the intensity can simply be written as $I = |\hat{\psi}_{\mathrm{ex}}|^2$. In the second case, not all frequency components of the exit-wave are identically transferred to the image plane since this transfer is affected by the aberrations of the optical system. These aberrations can be described as a sudden change of the electron wave in the back focal plane

$$\hat{\psi}_{\mathrm{im}}(\boldsymbol{\kappa}) = t(\boldsymbol{\kappa})\,\hat{\psi}_{\mathrm{ex}}(\boldsymbol{\kappa}), \tag{2.90}$$

where $\hat{\psi}_{\mathrm{im}}(\boldsymbol{\kappa})$ denotes the image wave function in reciprocal space and $t(\boldsymbol{\kappa})$ is the so-called transfer function. The intensity as detected on the camera is then given by

$$I(\boldsymbol{\rho}) = |\psi_{\mathrm{im}}(\boldsymbol{\rho})|^2 = \left|\mathcal{F}^{-1}\,\hat{\psi}_{\mathrm{im}}(\boldsymbol{\kappa})\right|^2. \tag{2.91}$$

The transfer function itself can be written as a product

$$t(\boldsymbol{\kappa}) = t_A(\boldsymbol{\kappa})\, t_L(\boldsymbol{\kappa})\, E_T(\boldsymbol{\kappa})\, E_S(\boldsymbol{\kappa}). \tag{2.92}$$

We will now discuss each of the contributing factors in turn.

The physical aperture placed in the back focal plane of the objective lens determines the maximum spatial frequency $\kappa_{\max}$ that can be transferred by the optical system. Thus, the aperture function $t_A(\boldsymbol{\kappa})$ excludes all frequency components that are larger than $\kappa_{\max}$

$$t_A(\boldsymbol{\kappa}) = \begin{cases} 1 & \text{if } \kappa < \kappa_{\max} \\ 0 & \text{else} \end{cases}. \tag{2.93}$$

Obviously, this function can be regarded as a low-pass filter for the imaging process.

The imperfections of the optical system shift the relative phases of the different frequency components. This is described by the lens transfer function

$$t_L(\boldsymbol{\kappa}) = \mathrm{e}^{-\mathrm{i}\, k\, \chi(\boldsymbol{\kappa})}, \tag{2.94}$$

where $\chi(\boldsymbol{\kappa})$ is the real-valued aberration function. In general, the aberration function can become quite complicated. However, we will consider only isotropic aberrations up to the third-order, thus neglecting effects like astigmatism. Small spatial frequencies ($\kappa \ll k$) are approximately proportional to the scattering angle ϑ

$$\kappa \approx k\, \vartheta. \tag{2.95}$$

Thus, the aberration function can be written as

$$\chi(\vartheta) = \frac{C_1}{2}\, \vartheta^2 + \frac{C_3}{4}\, \vartheta^4. \tag{2.96}$$

Here, C_1 is the defocus and C_3 is the spherical aberration coefficient. A lens with non-vanishing spherical aberration focuses electrons into different points depending on their distance to the optical axis. In a conventional electron microscope without aberration correction, the defocus is the only adjustable parameter and because of Scherzer's theorem C_3 must always be a positive number. On the other hand, in microscopes equipped with a corrector for the spherical aberration, C_3 can be controlled as well and even set to negative values.

The two functions t_A and t_L would sufficiently describe the optical system if the imaging process were perfectly coherent. As this is not the case, the transfer of high spatial frequencies is damped. This is modeled by the two remaining functions E_T and E_S.

The temporal envelope function E_T describes the damping caused by the chromatic aberration of the objective lens C_C in combination with the intrinsic energy width of the electron source ($\Delta T/T$) and the instabilities of the microscope parameters. It is given by

$$E_T(\omega) = \exp\left(-\frac{k^2\, \Delta C_1^2}{2} \left(\frac{\partial \chi}{\partial C_1}\right)^2\right) \overset{(2.96)}{=} -\frac{k^2\, \Delta C_1^2\, \vartheta^4}{8} \tag{2.97}$$

with the chromatic defocus spread

$$\Delta C_1 = C_C \sqrt{\left(\frac{\Delta T}{T}\right)^2 + \left(\frac{\Delta U}{U_0}\right)^2 + 4\left(\frac{\Delta I}{I_0}\right)^2}, \tag{2.98}$$

where $\Delta U/U_0$ denotes the instability of the high tension and $\Delta I/I_0$ the instability of the lens current. The damping is thus caused by the dependence of the aberration function χ on the defocus C_1: Electrons with slightly different energies are focused differently and the image becomes blurred. A corrector for the chromatic aberration allows setting C_C approximately to zero. Thereby, the effect of the temporal envelope is eliminated.

At last, we must consider the fact, that the electron source is not a point-like object, but it has a finite extent. Therefore, different points of the outgoing electron wave are not in perfect coherence with each other. As a consequence, the specimen is not illuminated by a plane wave, but by a cone-shaped partially coherent wave. Let the semi-angle of the cone be ϑ_S, then the envelope function due to the partial spatial coherence of the incoming wave is

$$E_S = \exp\left(-\frac{k^2\,\vartheta_S^2}{4}\left(\frac{\partial\chi}{\partial\vartheta}\right)^2\right) \overset{(2.96)}{=} \exp\left(-\frac{k^2\,\vartheta_S^2}{4}\left(C_1^2\,\vartheta^2 + 2\,C_1\,C_3\,\vartheta^4 + C_3^2\,\vartheta^6\right)\right). \tag{2.99}$$

The envelope depends on the defocus and on the spherical aberration coefficient. Thus, an aberration-corrected microscope allows the transfer of higher spatial frequencies.

For the rest of this work, the chromatic aberration coefficient C_C and the spherical aberration coefficient C_3 are assumed to be zero.

Chapter 3

The Multislice Approach

Published material – Parts of this chapter are reproduced with modifications from the publication C. Wacker and R. R. Schröder, "*Multislice algorithms revisited: Solving the Schrödinger equation numerically for imaging with electrons*", Ultramicroscopy 151 (2015), pp. 211–223, DOI: 10.1016/j.ultramic.2014.12.008, with kind permission from Elsevier.

In this chapter, the classical multislice approach for the simulation of electron micrographs is introduced. Originally, the multislice algorithm was motivated by optical principles [44]. However, the quantum mechanical derivation given here is a better starting point for a comprehensive discussion of multislice algorithms. Many concepts introduced in this chapter are of great importance for the more sophisticated simulation algorithms that will be discussed in the later chapters.

3.1 The Schrödinger Equation

Since we want to derive the multislice algorithm from quantum mechanical principles, we start from the Schrödinger equation augmented by relativistic corrections (2.14).

The coordinate system is chosen so that the z-axis and the optical axis of the microscope coincide. Moreover, the incoming electron wave shall move in the direction of increasing z values. Hence, the phase of the wave is a rapidly oscillating function of z. This motivates the following ansatz for the wave function

$$\psi(\mathbf{r}) = a(\mathbf{r})\,\mathrm{e}^{\mathrm{i}\,k\,z}. \tag{3.1}$$

The complex-valued amplitude function $a(\mathbf{r})$ varies only slowly with z. It is important to note, that (3.1) is completely general, especially, we do not assume that $\psi(\mathbf{r})$ can be approximated by a plane wave.

Substituting (3.1) into the Schrödinger equation (2.14), we arrive at

$$\partial_z^2\, a(\mathbf{r}) + 2\,\mathrm{i}\,k\,\partial_z\, a(\mathbf{r}) = -\Delta_\perp\, a(\mathbf{r}) + v(\mathbf{r})\, a(\mathbf{r}). \tag{3.2}$$

The Laplace operator Δ has been split in components perpendicular and parallel to the optical axis

$$\Delta = \Delta_\perp + \partial_z^2, \qquad \text{where } \Delta_\perp = \partial_x^2 + \partial_y^2. \tag{3.3}$$

Equation (3.2) is just a reformulation of the original Schrödinger equation. Thus, it has the same solutions, but it is better suited for numerical evaluations, making it the starting point for the discussion of the multislice algorithms.

3.1.1 The Paraxial Approximation

By neglecting the second z-derivative in (3.2), we get the paraxial approximation of the Schrödinger equation

$$2\,\mathrm{i}\,k\,\partial_z\,a(\mathbf{r}) = -\Delta_\perp\,a(\mathbf{r}) + v(\mathbf{r})\,a(\mathbf{r}). \tag{3.4}$$

The paraxial approximation is often referred to as high-energy approximation. As we will demonstrate later (cf. section 5.3.3), the high energy of the incoming electron is not sufficient to ensure the applicability of this approximation.

By using the paraxial approximation, the structure of the possible solutions changes quite significantly: Equation (3.4) is a first-order differential equation in z. By virtue of the Picard-Lindelöf theorem, this means that the function $a(\mathbf{r})$ is uniquely determined in the whole space if the values for a single xy-plane are given. However, the non-approximated equation (3.2) is a second-order differential equation. Hence, twice the amount of information is needed to uniquely specify the solution, e.g. the wave function and its first derivative for a single xy-plane. As a consequence, the paraxial approximation cannot be used to describe backscattering effects.

An electron moving in the direction $\mathbf{k}$ can be described by a plane wave

$$\psi_{\mathrm{plane}}(\mathbf{r}) \propto \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}}. \tag{3.5}$$

This is a solution of the exact Schrödinger equation (2.14). The length of the vector $\mathbf{k}$ is fixed by the kinetic energy of the electron (cf. (2.10)). The corresponding amplitude function is given by

$$a_{\mathrm{plane}}(\mathbf{r}) = \mathrm{e}^{-\mathrm{i}\,k\,z}\psi(\mathbf{r}) \propto \mathrm{e}^{-\mathrm{i}\,k\,z}\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}} = \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\,\exp\left(\mathrm{i}\,z\left[\sqrt{k^2-\kappa^2}-k\right]\right). \tag{3.6}$$

The function $a_{\mathrm{plane}}(\mathbf{r})$ is not quite a solution of (3.4). If the wave is moving mainly in z direction then $\boldsymbol{\kappa}$ is small and the square root in (3.6) can be expanded using the binomial series

$$\sqrt{1+x} = \sum_{\ell=0}^{\infty} \binom{1/2}{\ell} x^\ell. \tag{3.7}$$

This results in

$$a_{\mathrm{plane}}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\exp\left(\mathrm{i}\,k\,z\sum_{\ell=1}^{\infty}\binom{1/2}{\ell}\left(-\frac{\kappa^2}{k^2}\right)^\ell\right). \tag{3.8}$$

The leading term

$$a(\mathbf{r}) \propto \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\exp\left(-\frac{\mathrm{i}\,\kappa^2\,z}{2\,k}\right) \tag{3.9}$$

is indeed a solution of the approximated Schrödinger equation (3.4). The spherical dependence on κ in the exponent of (3.6) is replaced by a parabolic dependence in (3.9). Thus, the Ewald's sphere is distorted into a paraboloid. If κ is chosen to be greater than k, then the solution (3.6) will be exponentially damped. On the other hand, in (3.9) arbitrarily large values for κ are possible. The importance of the damped solutions for image simulations will be discussed in great detail starting with chapter 4.

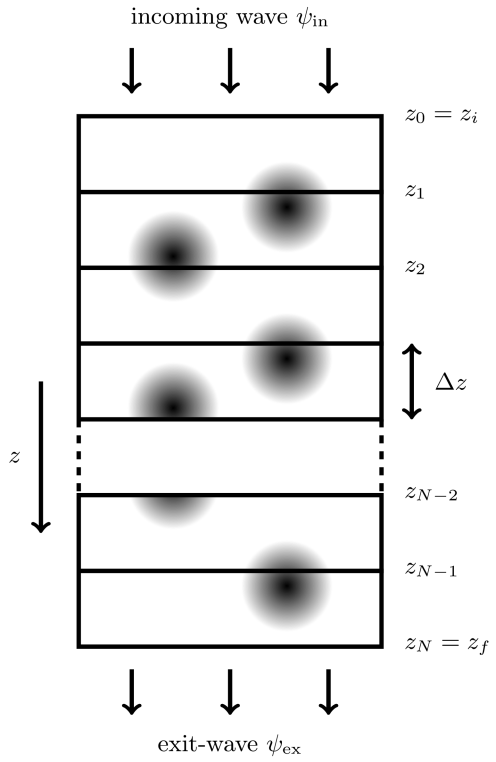

Figure 3.1: General structure of a multislice algorithm. The sample is divided into N slices of thickness Δz. The exit-wave function ψ_{ex} is calculated by propagating the electron wave successively through the slices. The circular regions mark the influence of the atomic potentials V_α. Depending on Δz a single potential can extend over several slices.

The two-dimensional Laplace operator $\Delta_\perp$ in (3.4) is Hermitian because it is essentially the square of momentum operators. Since we only consider real-valued potentials, the operator

$$H(\mathbf{r}) = \frac{1}{2\,k}\left[-\Delta_\perp + v(\mathbf{r})\right] \tag{3.10}$$

is Hermitian as well. Thus, (3.4) has the structure of a time-dependent Schrödinger equation for two spatial dimensions

$$\mathrm{i}\,\partial_z\, a(\boldsymbol{\rho}, z) = H(\boldsymbol{\rho}, z)\, a(\boldsymbol{\rho}, z). \tag{3.11}$$

The potential $v(\boldsymbol{\rho}, z)$ and therefore also the "Hamilton" operator $H(\boldsymbol{\rho}, z)$ are explicitly dependent on z. Hence, a direct solution of (3.4) requires the formalism of time-dependent quantum mechanics.

This formalism can be avoided by restricting the discussion of (3.4) for the moment to an interval $[z_n, z_{n+1}]$ on the z axis. Hence, we are looking at a slice of the three-dimensional space. The length of the interval $\Delta z_n = z_{n+1} - z_n$ shall be significantly smaller than the characteristic length scale of the potential. This allows us to replace the potential $v(\boldsymbol{\rho}, z)$ by its averaged value along the z-axis

$$v(\boldsymbol{\rho}, z) \quad \longrightarrow \quad v^{(n)}(\boldsymbol{\rho}) = \frac{1}{\Delta z_n} \int_{z_n}^{z_{n+1}} v(\boldsymbol{\rho}, z)\,\mathrm{d}z. \tag{3.12}$$

The new "Hamilton" operator $H_n(\boldsymbol{\rho})$ is independent of z, but it is a function of the slice index n

$$H_n(\boldsymbol{\rho}) = \frac{1}{2\,k}\left[-\Delta_\perp + v^{(n)}(\boldsymbol{\rho})\right]. \tag{3.13}$$

Now, if $a(\boldsymbol{\rho}, z_n)$ is known for a certain slice, then the change of the amplitude function in that slice can be expressed in terms of exponential functions

$$a(\boldsymbol{\rho}, z_{n+1}) = W(z_{n+1}, z_n)\, a(\boldsymbol{\rho}, z_n) \tag{3.14}$$

$$\text{with } W(z_{n+1}, z_n) = \exp\left(-\mathrm{i}\,\Delta z_n\, H_n\right) = \exp\left(\frac{\mathrm{i}\,\Delta z_n}{2\,k}\left[\Delta_\perp - v^{(n)}(\boldsymbol{\rho})\right]\right). \tag{3.15}$$

Since the operator H_n is Hermitian, the operator W is unitary. Hence, the total probability does not change from one slice to the next. This is an important benchmark for all numerical implementations based on the paraxial approximation.

In order to extend the formal solution ((3.14) and (3.15)) from a small interval $[z_n, z_{n+1}]$ to an arbitrary interval $[z_i, z_f]$, the latter must be divided so that each of its sub-intervals is small compared to the characteristic length scale of the potential (cf. figure 3.1). The total number of intervals shall be N. By applying (3.14) iteratively, we arrive at

$$a(\boldsymbol{\rho}, z_f) = W(z_f = z_N, z_{N-1}) \cdots W(z_2, z_1)\, W(z_1, z_i = z_0)\, a(\boldsymbol{\rho}, z_i) = \left[\prod_{n=N}^{1} W(z_n, z_{n-1})\right] a(\boldsymbol{\rho}, z_i) \tag{3.16}$$

The partitioning of the simulation volume into thin slices gives the multislice algorithm its name. Note, that the interval lengths Δz_n do not need to be the same for each slice.

The exponential function in (3.15) can be numerically evaluated by summing up a finite number of terms from the exponential series

$$\mathrm{e}^x = \sum_{\ell=0}^{\infty} \frac{x^\ell}{\ell!}. \tag{3.17}$$

Because of the unconditional convergence of this series, only a few terms are needed. In section 3.3, numerical methods for the treatment of the Laplace operator $\Delta_\perp$ will be discussed. This will then enable us to calculate numerical solutions for the approximated Schrödinger equation. We will refer to this algorithm as *Single* since it involves only a single exponential function per slice. The numerical properties of this algorithm have been previously analyzed by Cai et al. [45, 46].

The numerical evaluation of the exponential function in (3.15) can be simplified by rewriting the function into a product. However, this procedure is hampered by the fact that the operators $\Delta_\perp$ and $v^{(n)}$ do not commute. Thus, higher-order correction terms emerge from the Zassenhaus formula

$$\mathrm{e}^{t\,(X+Y)} = \mathrm{e}^{t\,X}\,\mathrm{e}^{t\,Y}\,\mathrm{e}^{-\frac{t^2}{2}\,[X,Y]}\cdots, \tag{3.18}$$

where X and Y are two linear operators and t is a complex parameter. Nevertheless, the implied Trotter formula

$$\mathrm{e}^{X+Y} = \lim_{N\to\infty}\left(\mathrm{e}^{X/N}\,\mathrm{e}^{Y/N}\right)^N \tag{3.19}$$

demonstrates that the correction terms become negligible for large numbers of slices, i.e., for small Δz_n. Under this condition, we can transform (3.15) into

$$W(z_{n+1}, z_n) = \exp\left(\frac{\mathrm{i}\,\Delta z_n}{4\,k}\,\Delta_\perp\right)\exp\left(-\frac{\mathrm{i}\,\Delta z_n}{2\,k}\,v^{(n)}\right)\exp\left(\frac{\mathrm{i}\,\Delta z_n}{4\,k}\,\Delta_\perp\right). \tag{3.20}$$

Note, that the product has been symmetrized. This is the classical multislice approach as proposed by Cowley and Moodie [44] and will hence be called *Classical*. As in the case for the *Single* algorithm, the operator W given by (3.20) is unitary.

In (3.20), the effects of the potential ($\propto v^{(n)}$) are separated from the effects of the kinetic term ($\propto \Delta_\perp$). As a consequence, the integrated potential $v^{(n)}$ causes merely a phase shift. We will see in section 3.3 that, in a similar way, an efficient evaluation of the kinetic term is possible by using an FFT. Because of this, the *Classical* algorithm is potentially much faster than the *Single* algorithm. Yet, by rewriting the exponential function as a product, the error per slice has been increased. For a fixed slice thickness Δz, the *Single* algorithm should thus deliver superior results. The numerical results calculated with these two algorithms will be compared in section 3.5.2.

Very simple simulation programs just include the effects of the potential and neglect the kinetic terms in (3.20). As there is no interaction between the different pixels left, this is called the independent column approximation.

3.2 The Potential

For the simulation of an electron micrograph, the sample potential must be known. The two multislice algorithms discussed so far utilize the averaged potential (3.12). Later in this work, algorithms will be derived that use the potential directly. In this section, we will thus look at algorithms to calculate the potential $V(\mathbf{r})$ and the phase shift[a]

$$\varphi^{(n)}(\boldsymbol{\rho}) = \int_{z_n}^{z_{n+1}} V(\boldsymbol{\rho}, z)\,\mathrm{d}z. \tag{3.21}$$

Note, that the two functions $V(\mathbf{r})$ and $\varphi^{(n)}(\boldsymbol{\rho})$ are energy-independent.

In section 2.4, we wrote the sample potential as the superposition of the free atomic potentials (2.76). This is, of course, just an approximation since at least the electrons of the outer shell will rearrange themselves to form chemical bonds. Furthermore, some electrons may be delocalized in a valence or in a conduction band. In this case, the corresponding electrons can no longer be assigned to a certain atom. As the fraction of outer electrons increase with decreasing atomic number, these effects become more important for light elements. Improved potentials calculated with the help of density functional theory have successfully been included in multislice algorithms. Although the predicted effects are small, they are indeed observable [47, 48].

Since in this work, we are mainly interested in propagation methods, we will stick to the simple independent atom model. Hence, the sample potential can be written as

$$V(\mathbf{r}) = \sum_\alpha V_\alpha(\mathbf{r} - \mathbf{r}_\alpha), \tag{3.22}$$

where the sum includes all atoms of the sample. Their positions are given by $\mathbf{r}_\alpha$. Although the atomic potentials are indexed by α, in reality they depend only on the atomic number Z, the ionization state and the Debye-Waller factor.

Numerical implementations of the multislice algorithm represent the functions $a(\boldsymbol{\rho})$ and $\varphi^n(\boldsymbol{\rho})$ as two-dimensional arrays of pixels. The evaluation of $\varphi^n(\boldsymbol{\rho})$ is computationally intensive, as, in

[a] According to (3.20), it is really the quantity $-\frac{\Delta z_n}{2\,k}\,v^{(n)} = -\frac{m\,\gamma}{k\,\hbar^2}\,\varphi^{(n)}$ that should be called phase shift. Using the definition (3.21), however, the following discussion is easier to understand.

principle, every atom contributes to every pixel in every slice. If we consider only neutral atoms, then the range of the atomic potentials is quite limited and we can introduce a cutoff distance, restricting the sum in (3.22) to the nearest atoms only.

Nevertheless, efficient evaluation algorithms are needed. In the next sections, three different approaches are introduced. At first, we will discuss the evaluation in real space.

3.2.1 Evaluation in Real Space

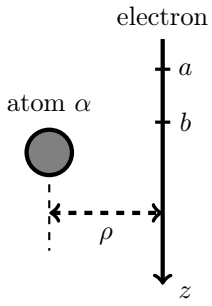

Figure 3.2: Trajectory of the incoming electron in relation to the atom α.

As we neglect bonding effects, the atomic potentials are spherically symmetric $V_\alpha(\mathbf{r}) = V_\alpha(r)$. We use the notation

$$V_\alpha(\rho, z) = V_\alpha\left(\sqrt{\rho^2 + z^2}\right). \tag{3.23}$$

For the calculation of the total phase shift (3.21), the function $\varphi_\alpha(\rho, a, b)$ is introduced which is proportional to the phase shift caused by an atom α if the incoming electron moves in a distance ρ along the z axis (cf. figure 3.2)

$$\varphi_\alpha(\rho, a, b) = \int_a^b V_\alpha(\rho, z)\,\mathrm{d}z = \left[\Phi_\alpha(\rho, b) - \Phi_\alpha(\rho, a)\right]. \tag{3.24}$$

Using this function, the total phase shift (3.21) can be expressed as the sum

$$\varphi^{(n)}(\boldsymbol{\rho}) = \sum_\alpha \int_{z_n}^{z_{n+1}} V_\alpha(|\boldsymbol{\rho} - \boldsymbol{\rho}_\alpha|, z - z_\alpha)\,\mathrm{d}z = \sum_\alpha \varphi_\alpha\left(|\boldsymbol{\rho} - \boldsymbol{\rho}_\alpha|, z_n - z_\alpha, z_{n+1} - z_\alpha\right). \tag{3.25}$$

The antiderivative Φ_α used in (3.24) is given by

$$\Phi_\alpha(\rho, z) = \int_0^z V_\alpha(\rho, z')\,\mathrm{d}z'. \tag{3.26}$$

This function depends only on two real-valued parameters. Moreover, ρ is non-negative and since $V_\alpha(\rho, z)$ is an even function of z, $\Phi_\alpha(\rho, z)$ is an odd function of z. Hence, we need to know $\Phi_\alpha(\rho, z)$ only for positive arguments. Furthermore, as the cutoff distance restricts the attainable values for z and ρ to a finite range $[0, \rho_{\max}] \times [0, z_{\max}]$, it is possible to evaluate $\Phi_\alpha(\rho, z)$ on a two-dimensional grid and to use an interpolation method for the values in between. That way, the form factor parametrizations, which typically involve computationally intensive functions (cf. appendix A), need to be evaluated only for the grid points. At the same time, the memory requirements are kept to a minimum. To prevent a loss of precision, however, the grid must be sampled at a better resolution than the potential.

The potential $V(\mathbf{r})$ can be calculated using the same methods: Instead of the function $\Phi_\alpha(\rho, z)$, the potential $V_\alpha(\rho, z)$ itself must be used for the interpolation.

Form factor parametrizations that feature the Coulomb singularity must be treated specifically. One possibility to avoid the divergences is to set the potential to a constant value below a certain radius.

3.2.2 Evaluation in Reciprocal Space

The problem caused by the Coulomb singularity can be circumvented by calculating the phase shifts in reciprocal space.

A commonly used approximation extends the integral in (3.24) to the entire z-axis

$$\varphi_\alpha(\rho, a, b) \longrightarrow \varphi_\alpha(\rho, -\infty, \infty) = \int_{-\infty}^{\infty} V_\alpha(\rho, z)\,\mathrm{d}z. \tag{3.27}$$

Since the potential decays exponentially, the error involved in this procedure can be small if the slice thickness is large enough ($\Delta z_n \gtrsim 1\,\text{Å}$, cf. figure 2.2). Moreover, the sum in (3.25) is restricted to the atoms located in the respective slice, so that every atom contributes only once to the phase shift.

According to (2.54), the atomic potential V_α in (3.27) can be written in terms of the electronic form factor $f_\alpha^e(\mathbf{K})$

$$\varphi_\alpha(\rho, -\infty, \infty) = -\frac{\hbar^2}{4\,\pi^2\,m} \int_{-\infty}^{\infty} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho} + \mathrm{i}\,K_z\,z}\, f_\alpha^e(\kappa, K_z)\,\mathrm{d}^3K\,\mathrm{d}z. \tag{3.28}$$

Changing the order of integration and using the Fourier representation of the delta function (2.63), we find

$$\varphi_\alpha(\rho, -\infty, \infty) = -\frac{\hbar^2}{2\,\pi\,m} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\, \delta(K_z)\, f_\alpha^e(\kappa, K_z)\,\mathrm{d}^3K = -\frac{\hbar^2}{2\,\pi\,m} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\, f_\alpha^e(\kappa, 0)\,\mathrm{d}^2\kappa. \tag{3.29}$$

This is the Fourier slice theorem. In this approximation, the phase shift caused by an atom is thus given by the two-dimensional Fourier transform of the electronic form factor.

Finally, we find for the total phase shift (3.25)

$$\varphi^{(n)}(\boldsymbol{\rho}) = -\frac{\hbar^2}{2\,\pi\,m} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} \left[\sum_{\alpha\,\in\,\text{slice }n} \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}_\alpha} f_\alpha^e(\kappa, 0) \right] \mathrm{d}^2\kappa. \tag{3.30}$$

Since the form factor is always finite, no divergences can occur in this expression.

The exponential function $\mathrm{e}^{-\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}_\alpha}$ is known as phase ramp. It determines the xy-position of the atom. Numerically, the two-dimensional Fourier transform ($\boldsymbol{\kappa} \to \boldsymbol{\rho}$) can be efficiently evaluated with the help of an FFT. However, the discrete Fourier transform requires that the source function (in this case, the square bracket) is periodic in $\boldsymbol{\kappa}$. Otherwise, artifacts (vertical and horizontal lines) will appear. Since the electronic form factor decays quickly, this presents no problem here. The atom positions can thus be specified with subpixel accuracy. The evaluation of the square bracket can be sped up by introducing a cutoff for κ. Then, a lookup table together with an interpolation method can be used to calculate the electronic form factor.

The requirement of a finite slice thickness in (3.27) prevents us from taking $\Delta z_n \to 0$. Only in this limiting case, however, the derivations of section 3.1.1 become exact. For instance, by the extension of the integration range, the precise z coordinate of the atoms is lost. Since CBED patterns (see section 5.3.2) contain information about the three-dimensional structure of the sample, such patterns may be simulated incorrectly if the reciprocal method is applied naively [49].

3.2.3 The Hybrid Approach

Here, we propose a new approach that extends the calculation in reciprocal space to finite slice thicknesses. As before, we represent the potential as the three-dimensional Fourier transform of the electronic form factor

$$V(\mathbf{r}) = -\frac{\hbar^2}{4\,\pi^2\,m} \sum_\alpha \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot(\boldsymbol{\rho}-\boldsymbol{\rho}_\alpha)}\,\mathrm{e}^{\mathrm{i}\,K_z(z-z_\alpha)}\,f_\alpha^e(\mathbf{K})\,\mathrm{d}^3K \tag{3.31}$$

$$= -\frac{\hbar^2}{4\,\pi^2\,m} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} \sum_\alpha \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}_\alpha}\,\overline{V}_\alpha(\kappa, z-z_\alpha)\,\mathrm{d}^2\kappa \tag{3.32}$$

$$\text{with } \overline{V}_\alpha(\kappa, z) = \int \mathrm{e}^{\mathrm{i}\,K_z\,z}\,f_\alpha^e(\kappa, K_z)\,\mathrm{d}K_z \tag{3.33}$$

We will call $\overline{V}_\alpha(\kappa, z)$ hybrid potential, as it depends on both real and reciprocal space coordinates. It is essentially the Fourier transform of the electronic form factor with respect to K_z.

If not only the potential, but also the phase shift is needed, we can insert (3.32) into (3.21)

$$\varphi^{(n)}(\rho) = -\frac{\hbar^2}{4\,\pi^2\,m} \int_{z_n}^{z_{n+1}} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} \sum_\alpha \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}_\alpha} \overline{V}_\alpha(\kappa, z-z_\alpha)\,\mathrm{d}^2\kappa\,\mathrm{d}z \tag{3.34}$$

$$= -\frac{\hbar^2}{m} \int \mathrm{e}^{\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} \sum_\alpha \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\kappa}\cdot\boldsymbol{\rho}_\alpha} \left[\,\overline{\Phi}_\alpha(\kappa, z_{n+1}-z_\alpha) - \overline{\Phi}_\alpha(\kappa, z_n - z_\alpha)\right] \mathrm{d}^2\kappa \tag{3.35}$$

$$\text{with } \overline{\Phi}_\alpha(\kappa, z) = \frac{1}{4\,\pi^2} \int_0^z \overline{V}(\kappa, z')\,\mathrm{d}z'. \tag{3.36}$$

The newly introduced function $\overline{\Phi}(\kappa, z)$ will be called hybrid antiderivative. It has the property

$$\overline{\Phi}_\alpha(\kappa, \infty) - \overline{\Phi}_\alpha(\kappa, -\infty) = \frac{1}{4\,\pi^2} \int_{-\infty}^{\infty} \overline{V}(\kappa, z')\,\mathrm{d}z' = \frac{1}{2\,\pi}\,f_\alpha^e(\kappa, 0). \tag{3.37}$$

Besides the different rules for the summations, (3.35) is thus a generalization of (3.30).

The comments regarding the evaluation of (3.30) apply accordingly to (3.32) and (3.35). As it turns out, for many common form factor parametrizations, the Fourier transform (3.33) and the integration (3.36) can be executed analytically even if the Debye-Waller factor is included. Two examples for this are presented in appendix A.

Like the commonly used reciprocal method, the hybrid approach can treat the Coulomb singularity without difficulties.

3.3 Treatment of the Laplace Operator

We have already seen, that the potential can be calculated in real or reciprocal space. The same is true for the evaluation of the Laplace operator.

It is well known, that a continuous Fourier transform turns a derivative into a multiplication. Hence, the Laplace operator applied to the amplitude function $a(\boldsymbol{\rho})$ can be written as

$$\Delta_\perp\,a(\boldsymbol{\rho}) = \mathcal{F}^{-1}\left(-\kappa^2\right)\mathcal{F}\,a(\boldsymbol{\rho}), \tag{3.38}$$

where $\mathcal{F}$ denotes the two-dimensional continuous Fourier transform. This expression can, for example, be used inside the series expansion of (3.15).

Up until now, we have treated the amplitude function a as a continuous function of $\boldsymbol{\rho}$. From now on, the discretized relations needed for the implementation of the multislice algorithm will be considered as well. As mentioned above, the function a can be represented on a two-dimensional grid with $M = M_x \times M_y$ pixels. The simulated area of the sample shall be a rectangle with the physical dimensions L_x and L_y. Hence, the sizes of the pixels in x- and y-direction are given by

$$\Delta x = \frac{L_x}{M_x}, \qquad \Delta y = \frac{L_y}{M_y}. \tag{3.39}$$

The size of the simulation area determines the lowest spatial frequencies that can be represented

$$\Delta\kappa_x = \frac{2\,\pi}{L_x}, \qquad \Delta\kappa_y = \frac{2\,\pi}{L_y}. \tag{3.40}$$

Then again, the highest spatial frequencies are the Nyquist frequencies

$$\kappa_{x,\mathrm{Ny}} = \frac{\pi\,M_x}{L_x} = \frac{\pi}{\Delta x}, \qquad \kappa_{x,\mathrm{Ny}} = \frac{\pi\,M_y}{L_y} = \frac{\pi}{\Delta y}. \tag{3.41}$$

The discrete analogue of the continuous Fourier transform used in (3.38) can be calculated with the help of an FFT. As its continuous counterpart, the discrete Fourier transform is unitary. The evaluation of a one-dimensional FFT takes $\mathcal{O}(M_x \ln M_x)$ essential arithmetic operations. Since, the two-dimensional FFT, as required by the multislice algorithm, can be decomposed into two one-dimensional FFTs, it has a computational complexity of

$$\mathcal{O}(M_x\,M_y\,(\ln M_x + \ln M_y)) = \mathcal{O}(M \ln M). \tag{3.42}$$

By the use of the discrete Fourier transform, it is implicitly assumed that all involved functions are periodic. To prevent wrap-around effects, it is thus necessary to choose a sufficiently large simulation area.

Applied to the *Classical* algorithm (3.20), the Fourier transform turns the exponential functions involving the Laplace operator into Fresnel propagators known from optics

$$\exp\left(\frac{\mathrm{i}\,\Delta z_n}{4\,k}\,\Delta_\perp\right) = \mathcal{F}^{-1}\exp\left(-\frac{\mathrm{i}\,\Delta z_n}{4\,k}\,\kappa^2\right)\mathcal{F}. \tag{3.43}$$

Hence, all exponential functions in (3.20) become pure phase factors and can be evaluated without resorting to the series expansion of the exponential function. Furthermore, this variant of the *Classical* algorithm is free from numerical instabilities and always delivers finite results.

In principle, the Fresnel propagator could be backtransformed into real space, turning (3.20) into a convolution which would take $\mathcal{O}(M^2)$ operations. However, since the FFT requires only $\mathcal{O}(M \ln M)$ operations, it is much more efficient to change back and forth between direct and reciprocal space. As the multiplication with the phase factor takes only $\mathcal{O}(M)$ operations, it does not contribute to the computational complexity.

Alternatively, the Laplace operator $\Delta_\perp$ can be discretized without taking the detour via the reciprocal space (3.38). This results in the so-called real-space algorithms [50], which will be discussed now.

The two-dimensional Laplace operator $\Delta_\perp$ is the sum of two second derivatives in orthogonal directions. Each of these derivatives can be approximated by a finite difference scheme. These

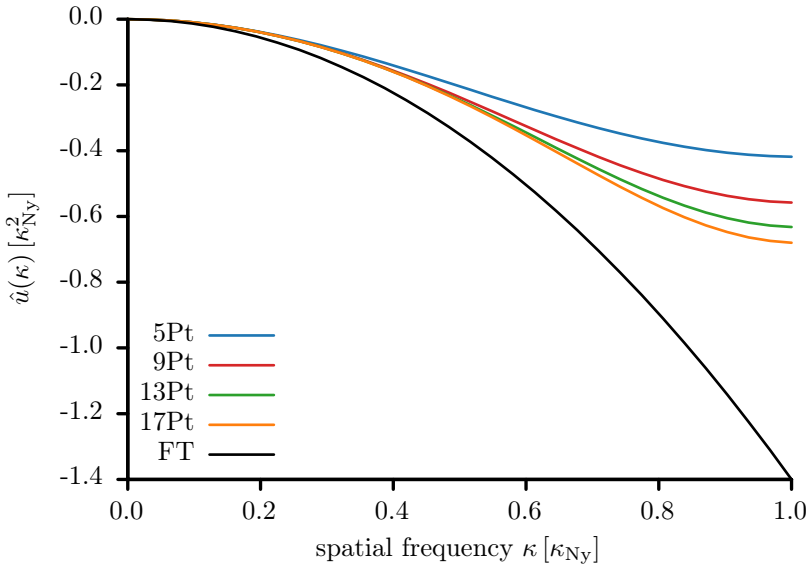

Figure 3.3: Comparison of different approximations of the Laplace operator in real and reciprocal space. The Fresnel propagator (FT) is shown in black.

schemes can be derived from Taylor expansions. In this work, only central schemes (cf. table 3.1) will be considered. For example, the two-dimensional Laplace operator approximated with nine points would read

$$\Delta_\perp a(x,y) = \frac{1}{\Delta x^2}\left[-5\,f_{i,j} + \frac{4}{3}\left(a_{i,j+1} + a_{i+1,j} + a_{i,j-1} + a_{i-1,j}\right) - \frac{1}{12}\left(a_{i,j+2} + a_{i+2,j} + a_{i,j-2} + a_{i-2,j}\right)\right] + \mathcal{O}(\Delta x^4) \tag{3.44}$$

$$\text{where } a_{i,j} = a(i\,\Delta x, j\,\Delta y). \tag{3.45}$$

For this procedure, the pixels must be quadratic ($\Delta x = \Delta y$). Otherwise, the calculated images will be distorted.

A difficulty arises at the edges of the simulation area, because the neighboring pixels are missing. This can be solved, for example, by periodic continuation of the function a. In this case, the Laplace operator is represented by a Hermitian matrix. Thus, we find that the real-space algorithms are unitary as well.

It seems that the accuracy of the finite difference schemes increases with decreasing pixel size Δx. This is, however, not correct for a numerical implementation with finite precision arithmetic, as the additional accuracy is counteracted by a loss of significance. Hence, an optimal pixel size Δx exists.

For the evaluation of the discretized Laplace operator acting on the amplitude function, only $\mathcal{O}(M)$ operations are needed. Since the amplitude function must be low-pass filtered (cf. section 3.4) an FFT must, however, be calculated nonetheless. Therefore, the computational advantage of the real-space algorithm cannot be exploited.

As we will demonstrate in section 3.5.1, the choice of a finite difference scheme has a profound influence on the calculated exit-wave function. This can be understood by analyzing these schemes in reciprocal space.

Points	0	1	2	3	4	**Accuracy**
5	-4	1				$\mathcal{O}(\Delta x^2)$
9	-5	$\frac{4}{3}$	$-\frac{1}{12}$			$\mathcal{O}(\Delta x^4)$
13	$-\frac{49}{9}$	$\frac{3}{2}$	$-\frac{3}{20}$	$\frac{1}{90}$		$\mathcal{O}(\Delta x^6)$
17	$-\frac{205}{36}$	$\frac{8}{5}$	$-\frac{1}{5}$	$\frac{8}{315}$	$-\frac{1}{560}$	$\mathcal{O}(\Delta x^8)$

Table 3.1: Coefficients of central finite difference schemes for the two-dimensional Laplace operator $\Delta_\perp$.

According to (3.38), the continuous Fourier transform turns the Laplace operator into a multiplication with $-\kappa^2$. In real space, the application of the discrete Laplace operator to a function can be read as a convolution with a kernel u. After a discrete Fourier transform, this convolution is likewise turned into a multiplication. The emerging factor, however, deviates from the wanted $-\kappa^2$ dependence (cf. figure 3.3).

For example, in one dimension the lowest-order kernel reads (cf. table 3.1)

$$u = \frac{1}{x^2} \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 & 1 \end{pmatrix}^T . \tag{3.46}$$

The results of the discrete one-dimensional Fourier transform is

$$\hat{u} = \frac{2\,\kappa_{\mathrm{Ny}}^2}{\pi^2} \left[\cos\left(\pi \, \frac{\kappa}{\kappa_{\mathrm{Ny}}} \right) - 1 \right] = -\kappa^2 \left[1 - \frac{\pi^2\,\kappa^2}{12\,\kappa_{\mathrm{Ny}}^2} + \mathcal{O}\left(\frac{\kappa^4}{\kappa_{\mathrm{Ny}}^4} \right) \right] , \tag{3.47}$$

where the Nyquist frequency κ_{Ny} is given by $\pi/\Delta x$. Hence, for a fixed κ the deviations from the correct $-\kappa^2$ dependence decrease with decreasing pixel size Δx.

3.4 Simulation Artifacts and the Low-Pass Filter

In this section, simulation artifacts that are caused by the discrete representation of the wave function and the potential will be discussed.

A first problem arises, because these two functions are defined on a rectangular grid. Thus, the maximum representable frequency depends on the direction in reciprocal space. As a result, rectangular artifacts can appear in real-space. This effect, however, can be avoided by applying an elliptical low-pass filter, enforcing a direction independent maximum frequency.

Another problem is the so-called aliasing effect. It can be explained by considering two continuous functions, both with a frequency component κ. If these two functions are multiplied, then the frequency component $2\,\kappa$ of the resulting function is non-zero. For discrete functions, the frequency $2\,\kappa$ is not representable, if it is greater than the Nyquist frequency κ_{Ny}. Because of the periodic nature of the discrete Fourier transform, the multiplication results in the wrong frequency $2\,\kappa_{\mathrm{Ny}} - 2\,\kappa$.

This problem can be solved by using a low-pass filter which removes all frequency components that are greater than $2/3$ of the Nyquist frequency. This filter must be applied before and, again, after the multiplication. This procedure removes all aliasing effects, but, because of the low-pass filter, the algorithm is no longer unitary. Hence, the intensity of the wave may decrease during the propagation.

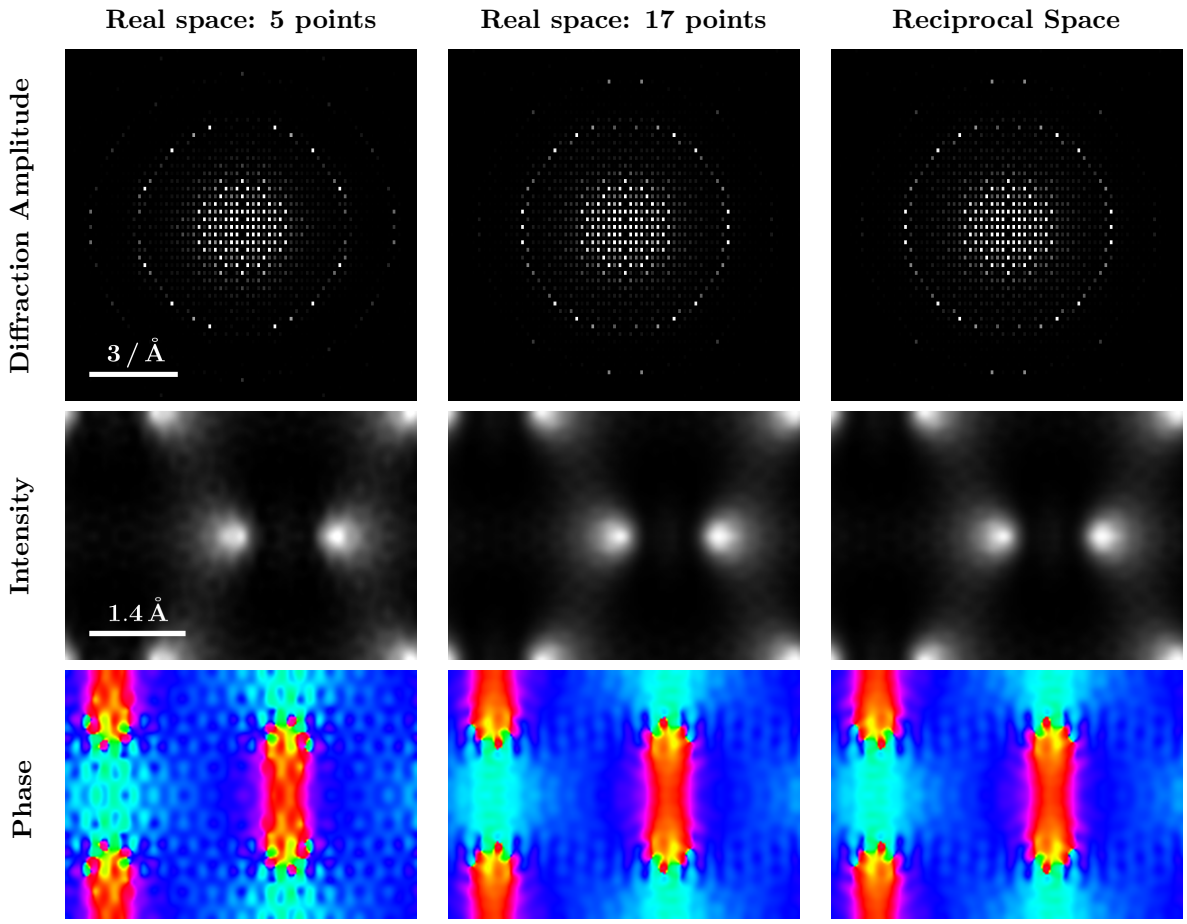

Figure 3.4: Diffraction amplitude, intensity $|\psi|^2$ and phase of the exit-wave function for three different approximations of the two-dimensional Laplace operator $\Delta_\perp$: real-space method with five points, real-space method with 17 points and the reciprocal-space method. The simulation results show 46.1 nm silicon [110] at an acceleration voltage of 80 kV. The pixel size is 0.03 Å. The diffraction patterns show the complete FOLZ ring and some faint spots of the SOLZ ring.
The high-frequency components of the exit-wave calculated in real space with five points show significant deviations. Apart from this, the images are in very good agreement.

As a consequence, not all calculated pixels contain usable information. For a quadratic simulation area, only a fraction of

$$\frac{\text{usable pixels}}{\text{total number of pixels}} = \frac{\pi \left(2/3\,\kappa_{\mathrm{Ny}}\right)^2}{4\,\kappa_{\mathrm{Ny}}^2} = \frac{\pi}{9} \approx 35\,\% \tag{3.48}$$

can be used. Still, this does not necessarily imply slower calculations, as a customized FFT can exploit the fact that a large number of pixels does not contribute.

The low-pass filter can be regarded as an elliptical aperture. As such, the filter can cause ringing artifacts. To suppress these artifacts, we propose to replace the steep filter function by a smooth Fermi function

$$f(\kappa) = \frac{1}{1 + \exp\left(\frac{\kappa/\kappa_{\mathrm{Ny}} - \mu}{T}\right)} \tag{3.49}$$

with the parameters

$$\mu = \frac{2}{3} - \frac{1}{20} \quad \text{and} \quad T = \frac{3}{200}. \tag{3.50}$$

All simulations in this work were performed using this filter function unless otherwise noted.

3.5 Numerical Results

3.5.1 Laplace Operator

To determine the impact of the different approximations of the Laplace operator on the final electron micrographs, 46.1 nm silicon [110] was simulated with the *Classical* algorithm at an acceleration voltage of 80 kV. The phase shifts were calculated with the hybrid method using the form factor parametrization by Lobato and Van Dyck (cf. appendix A.2). The slice thickness Δz was 19.2 pm, resulting in a total number of 2400 slices.

In figure 3.4, exit-waves calculated using different approximations are shown. The diffraction patterns show that the ZOLZ spots are in very good agreement for the three algorithms. However, the FOLZ ring calculated using the real-space algorithm with five points (5Pt) deviates from the other two algorithms. The intensity and especially the phase images confirm that the low spatial frequencies are the same, while the higher frequency components differ.

A powerful method to quantify the differences between two images is the Fourier ring correlation (FRC) [51]. It is defined as

$$\mathrm{FRC}(a_1, a_2;\, \kappa) = \frac{\sum_{||\tilde{\boldsymbol{\kappa}}|-\kappa|\leq\varepsilon} \hat{a}_1(\tilde{\boldsymbol{\kappa}})\, \hat{a}_2^*(\tilde{\boldsymbol{\kappa}})}{\sqrt{\left(\sum_{||\tilde{\boldsymbol{\kappa}}|-\kappa|\leq\varepsilon} |\hat{a}_1(\tilde{\boldsymbol{\kappa}})|^2\right)\left(\sum_{||\tilde{\boldsymbol{\kappa}}|-\kappa|\leq\varepsilon} |\hat{a}_2(\tilde{\boldsymbol{\kappa}})|^2\right)}}, \tag{3.51}$$

where $\hat{a}_i$ represents the Fourier transformed absolute values of the amplitude functions a_i. The summations are executed over a ring with radius κ and thickness $2\,\varepsilon$ (cf. figure 3.5). The FRC can be understood as an amplitude weighted sum of phase discrepancies.

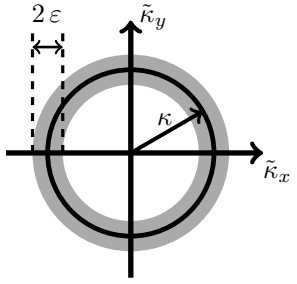

Figure 3.5: Geometry for the FRC.

Figure 3.6 compares the exit-waves calculated using the different real-space methods (5Pt, 9Pt, 13Pt and 17Pt, cf. table 3.1) with the results calculated in reciprocal space (FT). The low-frequency components, corresponding to the ZOLZ spots, always agree. Besides that, there is no clear relation between the value of the FRC and the spatial frequency κ. As already suggested by figure 3.3, the results of the real-space method converge towards the FT results.

It is also possible, to compare the real-space methods for different pixel sizes Δx. This is shown in figure 3.7. The FRCs confirm that the results of the real-space methods improve for smaller Δx (cf. (3.47)).

3.5.2 Comparison of the Single and the Classical Algorithm

We have already seen, that the *Classical* algorithm in combination with the Fourier transform always results in finite results. This is neither true for the *Single* algorithm nor for the real-space methods. Their convergence properties depend in the slice thickness Δz and the pixel size Δx.

In free space ($v = 0$), the Courant-Friedrichs-Lewy (CFL) condition [52] can be applied to the paraxial approximation of the Schrödinger equation (3.4). In this way, we find a necessary

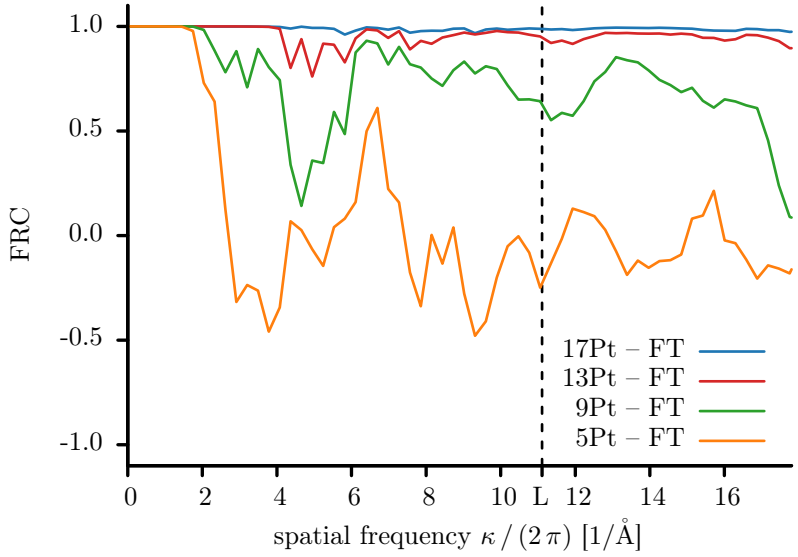

Figure 3.6: FRCs comparing the exit-waves calculated using different approximations of the two-dimensional Laplace operator $\Delta_\perp$. The results of the real-space algorithms (5Pt, 9Pt, 13Pt and 17Pt) are correlated with the result from reciprocal space (FT). The pixel size Δx is kept constant at 0.03 Å. The vertical line "L" marks the onset of the low-pass filter (11.1 Å^{-1}).
The correlations improve significantly with the number of points used for the real-space algorithm.

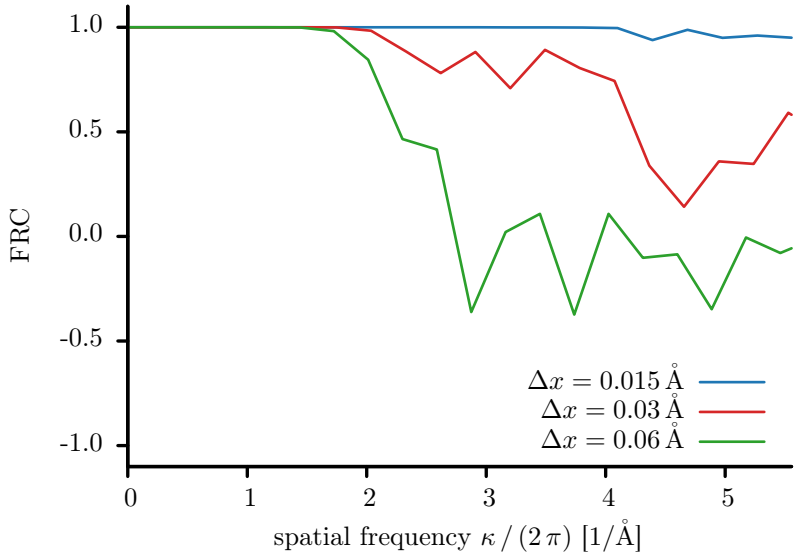

Figure 3.7: FRCs assessing the quality of the real-space algorithm 9Pt for different pixel sizes Δx. Exit-waves calculated using the real-space method with nine points and the reciprocal-space method are compared. For smaller pixel sizes, the agreement between the two results improves.

stability condition

$$\Delta z \leq 4\,k\,\Delta x^2\,C, \tag{3.52}$$

where C is a real number depending on the properties of the numerical algorithm. For explicit algorithms, as are the algorithms we have considered so far, this number is usually about one [53]. The CFL condition ensures that the resolution in z-direction is fine enough to track the spread of any disturbance. In our case, it estimates the critical Δz quite well (without the stabilizing low-pass filter):

Resolution	$4\,k\,\Delta x^2$ **[pm]**	**Empirical** Δz **[pm]**
90×64	216	$192 < \Delta z < 256$
181×128	54	$48 < \Delta z < 64$
362×256	14	$9.6 < \Delta z < 12.8$

Even if the multislice algorithms converge, the simulation results may show an unphysical dependence on the slice thickness. Therefore, Δz must be decreased until the results are independent of it. According to the discussion in section 3.1.1, the *Single* algorithm should stabilize for larger Δz than the *Classical* algorithm does. Using FSCs, it was possible to confirm this expectation. The advantage of the *Single* algorithm, however, is very small. At a resolution of 181×128 pixels, no differences can be found if Δz drops below 38.4 pm.

3.5.3 Computation Times

Finally, we look at the computation times needed by the different algorithms (cf. figure 3.8). The times were measured on a machine equipped with two Intel Xeon E5-2630 v3 processors with a total number of 16 physical cores. The presented values can give only a rough indication since the required times depend on the details of the stopping criteria for the series expansions. Furthermore, the performance of the involved FFTs improves significantly if M_x and M_y can be factorized as products of small prime numbers.

The *Classical* approach in combination with the Fourier transform is by far the fastest method as no series expansions are required. For the *Single* algorithm, however, the real-space method has distinct advantages. The required computation times depend only slightly on the number of points used for the evaluation of the Laplace operator. As the quality of the calculated results increases significantly, it would therefore be advisable to use at least 13 points. For the real-space methods, there is no significant time difference between the *Classical* and *Single* algorithms.

It is interesting to note that the computation times do not even quadruple as the number of pixels changes from 362×256 to 724×512. This underlines the need of supplementing theoretical analyses of computation times with benchmark results.

3.6 Discussion

As a consequence of the paraxial approximation, all backscattering effects have been removed from the equations. Furthermore, the high spatial frequencies are treated incorrectly as the Ewald's sphere is approximated by a paraboloid. Starting with the next chapter, these effects and improved equations will be investigated systematically.

Different methods for the calculation of the scattering potentials have been discussed. The newly proposed hybrid approach imposes no limits on the slice thickness and circumvents the problems associated with the Coulomb singularity.

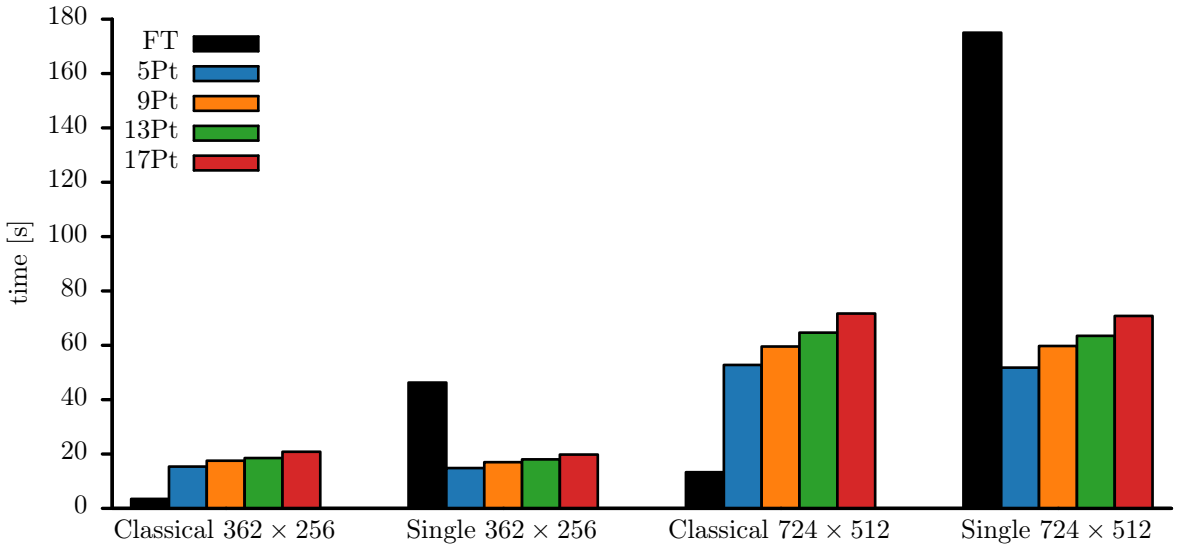

Figure 3.8: Comparison of the computation times required by the different approximations of the Laplace operator. For the measurement, 92.2 nm silicon [110] were simulated with the *Classical* algorithm at 80 kV. The slice thickness Δz of 19.2 pm results in 4800 slices.

It was shown in section 3.3, that the expansion of the Laplace operator in real space gives incorrect results, especially for large scattering angles. In section 3.5.1, we demonstrated that this is not merely a mathematical problem, but it affects the simulated micrographs as well. On the other hand, the evaluation in reciprocal space is not hampered by this problem. Moreover, in combination with the *Classical* algorithm, the evaluation in reciprocal space is even faster. For simulations with parallel illumination, the performance of the algorithms is of little concern nowadays, but it can become quite important for the simulation of STEM images, as the entire propagation procedure must be repeated for every beam position, i.e. for every pixel.

The claim that the results of the *Single* algorithm are vastly superior to that of the *Classical* algorithm [46] could not be confirmed. Our results suggest, that the differences found are caused by an incorrect treatment of the Laplace operator in real space.

Chapter 4

Electron Scattering as an Eigenvalue Problem

In the previous chapter, we showed how the classical multislice algorithm results from quantum mechanics and the paraxial approximation. To better understand the effects of this approximation, it is desirable to know the exact solutions of the Schrödinger equation. To this end, we will derive a new description of the scattering problem using as few approximations as possible. Moreover, as will be shown in chapter 6, this formulation can be turned into a practical algorithm for the calculation of backscattering effects.

It is clear, that for a realistic specimen no analytical solutions can be found. Hence, it is a necessity that the representation of the scattering problem is suitable for numerical evaluations. This will be achieved by rewriting the equations as an eigenvalue problem.

The utilization of linear algebra methods to treat electron scattering has already been proposed by Lynch and Moodie [21] for low energy electron diffraction (LEED). Their method, however, can only be applied to crystals. Furthermore, they assumed that the sample potential is constant in z, leading to severe discontinuities at the entry and the exit surface. The problems caused by this assumption will be discussed in section 6.1.

4.1 The Eigenvalue Problem

We start again from the Schrödinger equation (2.14) and rewrite it as

$$\partial_z^2 \, \psi = -\Delta_\perp \, \psi + v \, \psi - k^2 \, \psi = -A^2 \, \psi, \tag{4.1}$$

where we have used the reduced potential v as defined in (2.21). Note that at this stage, we do *not* consider A^2 to be the square of an operator A. The superscript "2" is part of the notation only. From the discussion in section 3.1.1, one can conclude that the operator A^2 is Hermitian.

The problems connected to the discretization of the two-dimensional Laplace operator $\Delta_\perp$ in real-space (cf. section 3.3) can be avoided by evaluating the Laplace operator in reciprocal space. Hence, we rewrite the operator A^2 as

$$A^2 = -\mathcal{F}^\dagger \, \kappa^2 \, \mathcal{F} - v + k^2, \tag{4.2}$$

where $\dagger$ denotes the Hermitian conjugate. The operator $\mathcal{F}$ representing the continuous Fourier transform is unitary.

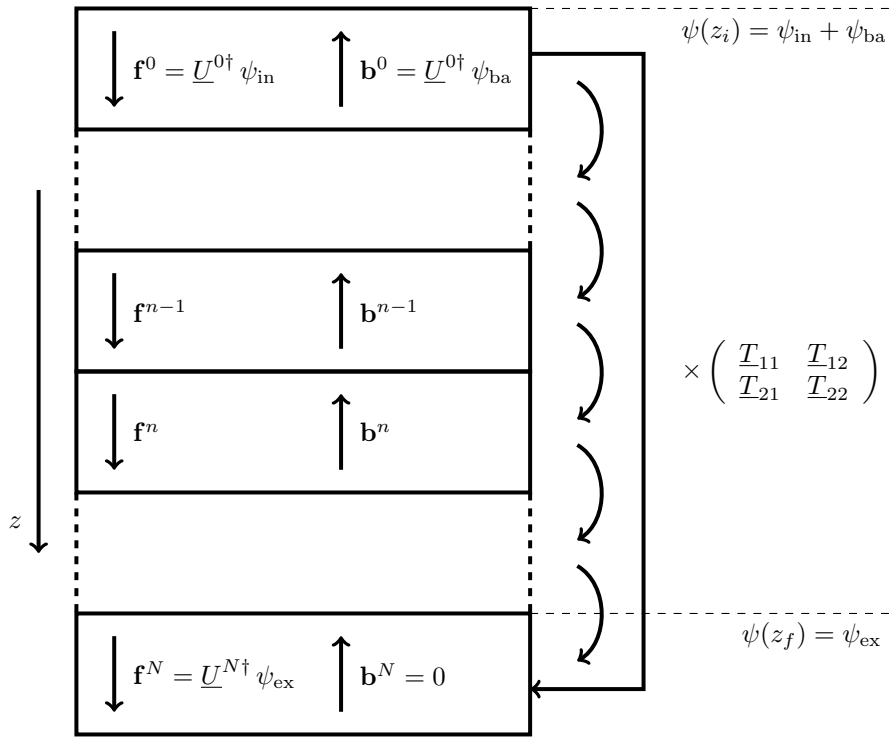

Figure 4.1: Schematic showing the definitions and relations for the coefficient vectors $\mathbf{f}^n$ and $\mathbf{b}^n$.

Since we are mainly interested in a numerical solution for the scattering problem, we can proceed by discretizing equation (4.1). As before, the potential v can be replaced by its mean value along the z-axis (3.12) if we restrict our analysis temporarily to a small interval $[z_n, z_{n+1}]$. In this case, the differential equation (4.1) no longer depends explicitly on z

$$\partial_z^2\,\psi^n = \left(\mathcal{F}^\dagger\,\kappa^2\,\mathcal{F} + v^{(n)} - k^2\right)\psi^n = -A^{2,n}\,\psi^n. \tag{4.3}$$

Instead of a single equation, however, we now have to deal with N equations, one for each slice.

Equation (4.3) can be evaluated on a finite two-dimensional grid in the xy-plane. For this purpose, the wave function $\psi^n(\boldsymbol{\rho})$ is turned into a complex-valued vector $\boldsymbol{\psi}^n$ with $M = M_x \times M_y$ entries. The operator $A^{2,n}$ is replaced in a similar way by the $M \times M$-matrix $\underline{A}^{2,n}$

$$\partial_z^2\,\boldsymbol{\psi}^n = -\underline{A}^{2,n}\,\boldsymbol{\psi}^n. \tag{4.4}$$

Hence, we have reduced the original partial differential equation (4.1) to a set of M coupled ordinary differential equations for each interval $[z_n, z_{n+1}]$.

Fortunately, these equations can be easily decoupled. As the matrix $\underline{A}^{2,n}$ is Hermitian, it follows from the spectral theorem that its eigenvalues are real. Furthermore, the matrix $\underline{A}^{2,n}$ can be diagonalized by a similarity transform using a unitary matrix $\underline{U}^n$

$$\underline{A}^{2,n} = \underline{U}^n\,\underline{D}^{2,n}\,\underline{U}^{n\dagger}. \tag{4.5}$$

Hence, $\underline{D}^{2,n}$ is a diagonal matrix with the eigenvalues of $\underline{A}^{2,n}$ on its main diagonal. The columns of $\underline{U}^n$ are given by the respective eigenvectors of $\underline{A}^{2,n}$. After the introduction of

$$\boldsymbol{\phi}^n = \underline{U}^{n\dagger}\,\boldsymbol{\psi}^n, \tag{4.6}$$

the differential equation (4.4) simplifies to

$$\partial_z^2 \boldsymbol{\phi}^n = -\underline{D}^{2,n}\, \boldsymbol{\phi}^n. \tag{4.7}$$

We proceed by analyzing these M independent equations. Each of them describes the one-dimensional motion of a particle in a constant potential. Thus, the solution of (4.7) is given by

$$\phi_m^n(z) = f_m^n\, \mathrm{e}^{\mathrm{i}\, D_m^n\, (z-z_n)} + b_m^n\, \mathrm{e}^{-\mathrm{i}\, D_m^n\, (z-z_n)} \quad \text{with } 1 \le m \le M. \tag{4.8}$$

The complex values f_m^n and b_m^n are the entries of the vectors $\mathbf{f}^n$ and $\mathbf{b}^n$ which need to be determined by the boundary conditions. To be unambiguous, the constants D_m^n shall be defined as

$$D_m^n = \begin{cases} \sqrt{D_m^{2,n}} & \text{if } D_m^{2,n} > 0 \\ \mathrm{i}\sqrt{-D_m^{2,n}} & \text{if } D_m^{2,n} < 0 \end{cases}, \tag{4.9}$$

where $D_m^{2,n}$ denotes a diagonal element of $\underline{D}^{2,n}$. Thus, a positive $D_m^{2,n}$ corresponds to an oscillating behavior. Under these circumstances, the probability current density (2.18) reveals that the terms $\propto f_m^n$ and $\propto b_m^n$ describe movements in forward and backward direction, respectively. For a negative $D_m^{2,n}$, however, the solution is a linear combination of a growing and a decaying exponential function. Hence, the probability current density vanishes. The eigenvalue spectrum will be discussed further in section 4.3.

Finally, the solution found needs to be extended to all z values. For this, we adapt the well-known transfer matrix method that has already been used by Chen and Van Dyck [22] to calculate backscattering effects.

If the Coulomb singularities caused by the atomic nuclei are removed as described in section 3.2, then the potential v is a bounded function. In this case, the wave function and its first derivative must be continuous. At each interface between two slices, they thus need to fulfill

$$\boldsymbol{\psi}^n(z_n) = \boldsymbol{\psi}^{n-1}(z_n), \qquad \boldsymbol{\psi}^{n\prime}(z_n) = \boldsymbol{\psi}^{n-1\prime}(z_n) \qquad \text{with } 1 \le n < N. \tag{4.10}$$

The resulting change in the coefficients $\mathbf{f}^n$ and $\mathbf{b}^n$ can be described with the help of a block matrix

$$\begin{pmatrix} \mathbf{f}^n \\ \mathbf{b}^n \end{pmatrix} = \begin{pmatrix} \underline{F}^n \exp\left(\mathrm{i}\,\underline{D}^{n-1}\Delta z_{n-1}\right) & \underline{B}^n \exp\left(-\mathrm{i}\,\underline{D}^{n-1}\Delta z_{n-1}\right) \\ \underline{B}^n \exp\left(\mathrm{i}\,\underline{D}^{n-1}\Delta z_{n-1}\right) & \underline{F}^n \exp\left(-\mathrm{i}\,\underline{D}^{n-1}\Delta z_{n-1}\right) \end{pmatrix} \begin{pmatrix} \mathbf{f}^{n-1} \\ \mathbf{b}^{n-1} \end{pmatrix}, \tag{4.11}$$

where "exp" denotes the matrix exponential function. Here, we have introduced the matrices

$$\underline{F}^n = \frac{1}{2}\left[\underline{U}^{n\dagger}\underline{U}^{n-1} + (\underline{D}^n)^{-1}\,\underline{U}^{n\dagger}\underline{U}^{n-1}\underline{D}^{n-1}\right], \tag{4.12}$$

$$\underline{B}^n = \frac{1}{2}\left[\underline{U}^{n\dagger}\underline{U}^{n-1} - (\underline{D}^n)^{-1}\,\underline{U}^{n\dagger}\underline{U}^{n-1}\underline{D}^{n-1}\right] = \underline{U}^{n\dagger}\underline{U}^{n-1} - \underline{F}^n. \tag{4.13}$$

By multiplying the equations (4.11) for all slices, one gets a matrix relation connecting the coefficients of the first and the last plane

$$\begin{pmatrix} \mathbf{f}^N \\ \mathbf{b}^N \end{pmatrix} = \begin{pmatrix} \underline{T}_{11} & \underline{T}_{12} \\ \underline{T}_{21} & \underline{T}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{f}^0 \\ \mathbf{b}^0 \end{pmatrix}. \tag{4.14}$$

The relationships between the coefficient vectors $\mathbf{f}^n$ and $\mathbf{b}^n$ for the different slices are also illustrated in figure 4.1.

Since the divergence of the probability current density vanishes (2.17), it can be shown that the probability current in z-direction is conserved. Gauss's theorem allows us to rewrite (2.17) in integral form

$$0 = \oint_{\partial V} \mathbf{j} \cdot \mathrm{d}^2\mathbf{f}. \tag{4.15}$$

Here, ∂V is the boundary of the volume V. This volume shall be an axis-aligned box. The discrete Fourier transform used to derive the eigenvalue problem implies periodic boundary conditions in x- and y-direction. This means that only the faces of the box parallel to the xy-plane need to be considered. The corresponding current in z-direction is given by

$$I_n = \int j_z \,\mathrm{d}x\,\mathrm{d}y \overset{(2.18)}{=} \frac{\hbar}{m} \operatorname{Im} \int \psi^* \,\partial_z \,\psi \,\mathrm{d}x\,\mathrm{d}y, \tag{4.16}$$

The integration domain is the simulation area $L_x \times L_y$. If we choose $z \in (z_n, z_{n+1})$, we find in the discrete case

$$I_n = \frac{\hbar}{m} \left[\sum_m D_m^n \,|f_m^n|^2 - \sum_m D_m^n \,|b_m^n|^2 \right], \tag{4.17}$$

where the sum only includes pixels with a positive eigenvalue. A wave moving in forward direction results in a positive contribution to the current while a wave moving in the opposite direction gives a negative contribution. The integration box in (4.15) can extend over several slices. Hence, we find that the probability current in z direction is the same throughout the sample

$$I_n = I_{n'} \quad \text{for } 0 \le n, n' \le N. \tag{4.18}$$

In this section, we derived a new formulation of the electron scattering problem. The partition of the samples into thin slices is retained from the classical multislice approach. While the classical approach propagates the electron wave through the sample directly, this new approach, however, calculates incremental updates for the transfer matrix (cf. (4.11)). This enables us to keep track of the backscattered signal as well.

The results of the current section have been obtained by applying only a minimal set of approximations: The equations were discretized in all three dimensions and thereby a finite simulation volume with periodic boundary conditions in x- and y-direction was assumed.

4.2 Boundary Conditions

The general solution of the previous section must be specialized for the situation in an electron microscope. For the interpretation of the coefficient vectors $\mathbf{b}^0$, $\mathbf{f}^0$, $\mathbf{b}^N$ and $\mathbf{f}^N$, it is necessary to undo the base change (4.6). The incoming wave is given by $\psi_{\text{in}} = \underline{U}^0 \,\mathbf{f}^0$ (cf. figure 4.1). The exit-wave function – the result of a conventional image simulation – is $\psi_{\text{ex}} = \underline{U}^N \,\mathbf{f}^N$. The corresponding signal in the opposite direction, i.e. towards the electron source, is $\psi_{\text{ba}} = \underline{U}^0 \,\mathbf{b}^0$. As no electron shall enter the sample from the side of the detector (camera), we set $\underline{U}^N \,\mathbf{b}^N = 0$. Using these boundary conditions, we find with the help of (4.14)

$$\underline{T}_{22} \,\mathbf{b}^0 = -\underline{T}_{21} \,\mathbf{f}^0, \tag{4.19}$$

$$\mathbf{f}^N = \underline{T}_{11} \,\mathbf{f}^0 + \underline{T}_{12} \,\mathbf{b}^0. \tag{4.20}$$

In principal, the first equation (4.19) can be solved for $\mathbf{b}^0$, the backscattered signal. The resulting vector can then be substituted in the second equation (4.20) to calculate $\mathbf{f}^N$ and finally the exit-wave ψ_{ex}.

Crystal	T [keV]	L_x [Å]	L_y [Å]	Without Low-Pass Filter		With Low-Pass Filter	
				M_x	M_y	M_x	M_y
Si [110]	80	5.43	3.84	183	130	390	275
Si [110]	20	5.43	3.84	89	63	189	134
Graphene	5	4.92	4.26	40	34	85	73

Table 4.1: Exemplary values for the inequality (4.24): maximum number of pixels $M_x \times M_y$ for given a kinetic energy T and a simulation area $L_x \times L_y$ to prevent the appearance of non-oscillating solutions. With the help of a low-pass filter, the effects of the non-oscillating functions can be suppressed (cf. section 5.3 and chapter 6).

Equations (4.19) and (4.20) cannot yet be used directly for the calculation of electron micrographs. As explained in section 3.4, a low-pass filter is necessary to prevent aliasing effects. If this filter is applied, however, the matrices $\underline{T}_{ij}$ become singular. Thus, (4.19) cannot be solved for $\mathbf{b}^0$. A solution circumventing this problem will be developed in chapter 6.

4.3 Analysis of the Spectrum

We have already seen that the spectrum of the matrix $\underline{A}^{2,n}$ is real-valued since the matrix is Hermitian. Depending on the sign of the eigenvalues $D_m^{2,n}$, the solution (4.8) can display either an oscillatory or an exponential behavior. For an analytical investigation, this makes little difference. The numerical treatment of the non-oscillatory functions, however, will turn out to be difficult. Hence, in this section, we will formulate a condition for the simulation parameters that ensures a positive spectrum.

According to (4.3), the entries of the matrix $\underline{A}^{2,n}$ are given by

$$A_{ij}^{2,n} = -\sum_{\ell} \mathcal{F}_{\ell i}^{*}\, \kappa_\ell^2\, \mathcal{F}_{\ell j} - v_{ij}^{(n)} + k^2. \tag{4.21}$$

For the moment, we ignore the effect of the potential $\underline{v}^{(n)}$. Since $\underline{\mathcal{F}}$ is a unitary matrix, the remaining terms can be written as

$$A_{ij}^{2,n} = \sum_{\ell} \mathcal{F}_{\ell i}^{*} \left(-\kappa_\ell^2 + k^2\right) \mathcal{F}_{\ell j}. \tag{4.22}$$

Hence, the spectrum of $\underline{A}^{2,n}$ is given by $-\kappa_\ell^2 + k^2$ where all possible values for ℓ have to be substituted.

While k^2 is determined by the acceleration voltage, the maximum value for κ^2 is given by the sum of the squared Nyquist frequencies (3.41). Thus, a positive spectrum can be guaranteed by respecting the inequality

$$k^2 > \pi^2 \left(\frac{M_x^2}{L_x^2} + \frac{M_y^2}{L_y^2} \right) = \pi^2 \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} \right). \tag{4.23}$$

Using the relativistic dispersion relation (2.10), we can rewrite the inequality (4.23) in terms of the kinetic energy of the electron

$$T^2 + 2\, m\, c^2\, T > c^2\, \hbar^2\, \pi^2 \left(\frac{M_x^2}{L_x^2} + \frac{M_y^2}{L_y^2} \right) = c^2\, \hbar^2\, \pi^2 \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} \right). \tag{4.24}$$

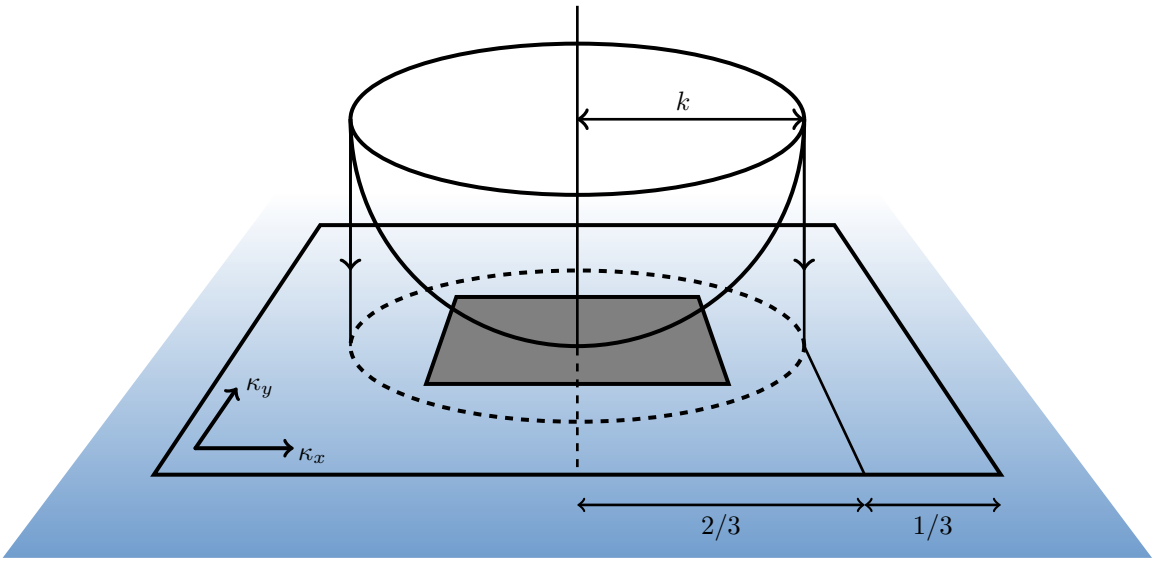

Figure 4.2: Physical interpretation of the inequality (4.23). Shown is the projection of the Ewald's sphere onto the $\kappa_x\kappa_y$-plane. The radius of the the Ewald's sphere k is determined by the kinetic energy T. Each point of the plane represents a certain transversal momentum $\hbar\,\boldsymbol{\kappa}$. Outside the circle, the "transversal kinetic energy" is greater than the available energy T. Thus, the electron wave cannot propagate in that region and the non-oscillating functions must be used to describe the wave function. If the simulation area (gray) fits into the projection of the Ewald's sphere (circle), then the non-oscillating functions can be ignored. Alternatively, a low pass filter can be used to suppress these functions so that the simulation area can be extended to the larger rectangle. This fact will be exploited in the chapters 5 and 6.

For a given kinetic energy T and a certain size of the sample $L_x \times L_y$, this inequality sets a lower limit for the pixel size (Δx and Δy). Exemplary values are given in table 4.1.

Now we continue the discussion of the potential term $v_{ij}^{(n)}$ in (4.21). Since the sample potential is attractive, all entries of $\underline{v}^{(n)}$ are negative. Consequently, by including $\underline{v}^{(n)}$ the lower limit for k set by (4.23) does not increase. Thus, (4.23) is a sufficient condition to ensure a positive spectrum.

In the non-relativistic limit, a simple physical explanation for the condition (4.23) can be given: Without a potential, the available kinetic energy is a fixed value. It is the sum of the kinetic energy in the direction of the optical axis (T_z) and in the plane perpendicular to it ($T_\perp$)

$$T = T_z + T_\perp. \tag{4.25}$$

In reciprocal space, each pixel ($\boldsymbol{\kappa}$) of that plane is associated with a certain kinetic energy

$$T_\perp = \frac{\hbar^2\,\kappa^2}{2\,m}. \tag{4.26}$$

Thus, pixels in the corners of the simulation area correspond to large transversal energies. If these energies surpass the total energy available, T_z becomes a negative number. Therefore, the wave number in z-direction becomes imaginary and the oscillatory behavior of the solution is replaced by an exponential one. As a result, the wave cannot propagate outside the limits defined by (4.23).

To put it in another way, if the rectangular simulation area in reciprocal space does not fit into the projection of the Ewald's sphere (cf. figure 4.2), non-oscillating solutions in (4.11) need to be considered.

Chapter 5

Beyond the Paraxial Approximation

Published material – Parts of this chapter are reproduced with modifications from the publication C. Wacker and R. R. Schröder, "*Multislice algorithms revisited: Solving the Schrödinger equation numerically for imaging with electrons*", Ultramicroscopy 151 (2015), pp. 211–223, DOI: 10.1016/j.ultramic.2014.12.008, with kind permission from Elsevier.

In the last chapter, a general representation of the electron scattering problem including backscattering effects was derived. As we will see in chapter 6, backscattering effects are mostly negligible for the energies possible in a TEM ($T \geq 20\,\mathrm{keV}$). Moreover, implementations of the backscattering algorithm turn out to be very expensive in terms of computational resources in comparison to the classical multislice approach. Hence, the question arises if it is possible to exclude backscattering effects without resorting to the paraxial approximation.

Two solutions to this problem will be investigated in this chapter. In section 5.1, backscattering will be excluded by using appropriate initial conditions while keeping the Schrödinger equation unchanged. Afterward, in section 5.2.2, we will look into a modification of the scattering equations, first proposed by Chen and Van Dyck [22]. In the remaining sections of this chapter, we will compare these approaches with each other and with the *Classical* multislice algorithm.

5.1 Direct Quadrature of the Schrödinger Equation

5.1.1 Preliminary Considerations

Valuable insight into the discretized scattering problem can be gained by considering only a single spatial frequency $\boldsymbol{\kappa}$. Hence, we deal with a one-dimensional problem, governed by the Schrödinger equation

$$\partial_z^2\,\psi(z) = v(z)\,\psi(z) - k^2\,\psi(z). \tag{5.1}$$

Because of the slicing procedure, the potential $v(z)$ is a piecewise constant function. To discuss a concrete situation, we look at a single potential step

$$v(z) = \begin{cases} 0 & \text{if } z < 0 \\ v_0 & \text{if } z > 0 \end{cases}. \tag{5.2}$$

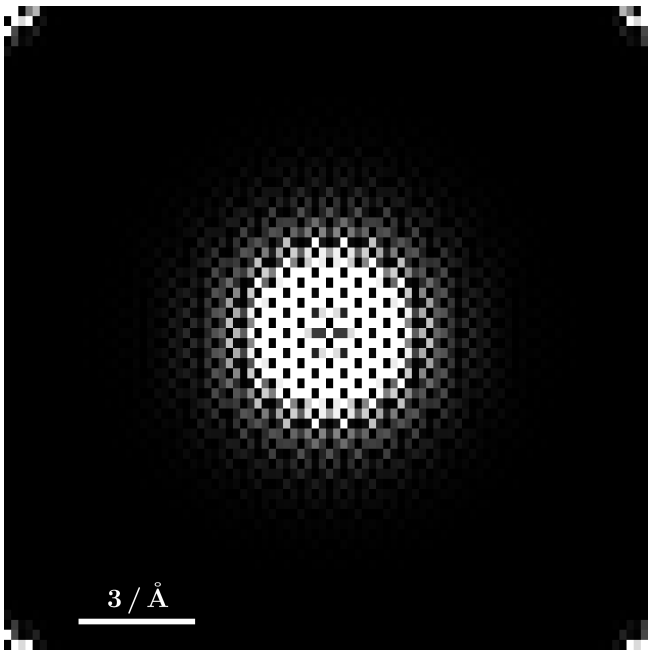

Figure 5.1: Absolute values of the diffraction amplitude after 3.84 Å of silicon [110]. The electron wave was calculated with the Runge-Kutta algorithm for the non-approximated Schrödinger equation (*RK2*, cf. section 5.1.2) without using a low-pass filter. The simulation parameters (kinetic energy: 20 kV, sample dimensions: 5.43 Å × 3.84 Å, resolution: 90 × 64, slice thickness Δz: 0.384 pm) violate the inequality (4.24) slightly, cf. table 4.1. Hence, divergences spread from the pixels representing the highest spatial frequencies κ, i.e. the corners.

Furthermore, we assume $0 < k^2 < v_0$. Thus, the region $z > 0$ is classically forbidden. In the notation of the previous chapter, this means that the eigenvalue D^2 is negative. In this case, the wave function can be written as

$$\psi(z) = \begin{cases} f\,\mathrm{e}^{\mathrm{i}\,k\,z} + b\,\mathrm{e}^{-\mathrm{i}\,k\,z} & z < 0 \\ c\,\mathrm{e}^{K\,z} + d\,\mathrm{e}^{-K\,z} & z > 0 \end{cases} \tag{5.3}$$

$$\text{with } K = \sqrt{v_0 - k^2}. \tag{5.4}$$

As a special case of (4.11), we obtain the following matrix relating the different amplitude coefficients

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\left(1 + \mathrm{i}\,\frac{k}{K}\right) & \frac{1}{2}\left(1 - \mathrm{i}\,\frac{k}{K}\right) \\ \frac{1}{2}\left(1 - \mathrm{i}\,\frac{k}{K}\right) & \frac{1}{2}\left(1 + \mathrm{i}\,\frac{k}{K}\right) \end{pmatrix} \begin{pmatrix} f \\ b \end{pmatrix}. \tag{5.5}$$

The term $b\,\mathrm{e}^{-\mathrm{i}\,k\,z}$ in (5.3) describes a left-moving wave corresponding to the backscattered signal. If we try to suppress this signal by setting $b = 0$, then the wave function grows exponentially in the forbidden region ($z > 0$) since $c \neq 0$. Thus, the expected decay of the wave function in that region can only take place, if the wave can avoid the region by being reflected at $z = 0$.

In a three-dimensional simulation, the exponential growth, as discussed here, always occurs if the inequality (4.24) is violated: In reciprocal space, the pixels corresponding to the largest spatial frequencies κ are located in the corners of the simulation area. Since the available kinetic energy for the movement in z direction T_z is reduced by the transversal momentum $\hbar\,\boldsymbol{\kappa}$, the forbidden region is located in these corners. Numerical noise can then initiate the exponential growth. As the potential v is not diagonal in reciprocal space, these divergences spread quickly

to all frequency components. In figure 5.1, it is shown that even a minimal violation of (4.24) results in an exponential growth. As expected, the effect becomes stronger for larger violations of the inequality. In conclusion, a single pixel that violates the inequality (4.24) can prevent the calculation of an electron micrograph.

Regardless of the sign of the eigenvalues, the suppression of the backscattered signal is unphysical, because every variation of the potential causes a reflection. Hence, for all slices other than the first, the backscattering coefficients usually do not vanish, even if the initial coefficients are set to zero. This means, that the mathematical solution requires the electron wave to enter the sample from the "wrong" side – the side of the camera. However, since all backscattering coefficients are insignificant, even for energies as low as 15 keV (cf. chapter 6), the error caused by this choice of initial conditions is negligible.

5.1.2 Runge-Kutta Algorithms

The Schrödinger equation as written in (3.2) can be regarded as an ordinary differential equation in z. If we assume that initial conditions for a certain z are known, we can apply a numerical integrator to find the corresponding solution. To that end, well studied algorithms such as Runge-Kutta methods and linear multistep methods can be used. An overview is given by Press et al. [54].

A common property of these algorithms is that they approximate a function of a continuous parameter by a series of discrete values. The distance between these values is called step size. We identify it with the slice thickness Δz_n. If the step size is not constant, the algorithms are called adaptive. Adaptive algorithms can utilize the available computing resources more efficiently. Yet for crystalline samples, it is advantageous to apply non-adaptive methods as these can reuse potentials that have been calculated for previous slices. Although, in principle, every numerical integrator can be utilized, we limit our discussion to the family of Runge-Kutta methods.

It is a well-known fact that a higher-order differential equation can be rewritten as a system of first-order equations. Hence, we only need to discuss equations of the form

$$\partial_z \, \mathbf{X}(z) = \underline{\tilde{H}}(z) \, \mathbf{X}(z), \tag{5.6}$$

where $\mathbf{X}(z)$ is a vector and $\underline{\tilde{H}}(z)$ is a matrix. A general Runge-Kutta algorithm with s stages can be written as

$$\mathbf{x}_k = \underline{\tilde{H}}(z_n + \gamma_k \, \Delta z_n) \left[\mathbf{X}_n + \Delta z_n \sum_{\ell=1}^{s} \alpha_{k\ell} \, \mathbf{x}_\ell \right] \quad \text{for } 1 \le k \le s, \tag{5.7}$$

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \Delta z_n \sum_{k=1}^{s} \beta_k \, \mathbf{x}_k \tag{5.8}$$

The vectors $\mathbf{x}_k$ denote intermediate results. The constants $\alpha_{k\ell}$, β_k and γ_k are real-valued and define the algorithm. They can be conveniently arranged in a Butcher tableau

$$\begin{array}{c|cccc} \gamma_1 & \alpha_{11} & \alpha_{12} & \dots & \alpha_{1s} \\ \gamma_2 & \alpha_{21} & \alpha_{22} & \dots & \alpha_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_s & \alpha_{s1} & \alpha_{s2} & \dots & \alpha_{ss} \\ \hline & \beta_1 & \beta_2 & \dots & \beta_s \end{array} . \tag{5.9}$$

A distinction is made between explicit and implicit methods: Explicit methods calculate $\mathbf{X}_{n+1}$ directly from the previous $\mathbf{X}_n$, while implicit methods need to solve an equation involving both $\mathbf{X}_n$ and $\mathbf{X}_{n+1}$. Although faster and easier to implement, explicit algorithms may suffer from numerical instabilities and require smaller step sizes. For explicit algorithms, the coefficients $\alpha_{k\ell}$ with $k \leq \ell$ must vanish.

The most famous example of an Runge-Kutta method is the original algorithm developed by Runge and Kutta [55] with four stages. Its Butcher tableau is given by

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array} . \tag{5.10}$$

This algorithm combines great speed with moderate accuracy.

The validity of this numerical approach can be verified by applying the Runge-Kutta algorithm to the paraxial approximation of the Schrödinger equation. By rewriting (3.4) in the form of (5.6), we obtain the following explicit expressions for the integration steps

$$\underline{\tilde{H}}(z) = \frac{\mathrm{i}}{2\,k}\left(\Delta_\perp - \underline{v}(z)\right), \tag{5.11}$$

$$\mathbf{x}_1 = \underline{\tilde{H}}(z_n)\,\mathbf{a}_n, \tag{5.12}$$

$$\mathbf{x}_2 = \underline{\tilde{H}}\left(z + \frac{\Delta z_n}{2}\right)\left[\mathbf{a}_n + \frac{\Delta z_n}{2}\,\mathbf{x}_1\right], \tag{5.13}$$

$$\mathbf{x}_3 = \underline{\tilde{H}}\left(z + \frac{\Delta z_n}{2}\right)\left[\mathbf{a}_n + \frac{\Delta z_n}{2}\,\mathbf{x}_2\right], \tag{5.14}$$

$$\mathbf{x}_4 = \underline{\tilde{H}}(z + \Delta z_n)\left[\mathbf{a}_n + \frac{\Delta z_n}{2}\,\mathbf{x}_3\right], \tag{5.15}$$

$$\mathbf{a}_{n+1} = \mathbf{a}_n + \Delta z_n\left[\frac{\mathbf{x}_1}{6} + \frac{\mathbf{x}_2}{3} + \frac{\mathbf{x}_3}{3} + \frac{\mathbf{x}_4}{6}\right], \tag{5.16}$$

where the vector $\mathbf{a}_n = \mathbf{X}_n$ represents the amplitude function $a(\boldsymbol{\rho}, z_n)$. The elements of the diagonal matrix $\underline{v}(z)$ are given by the potential $v(\boldsymbol{\rho}, z)$. We will call this algorithm *RK1* as it integrates the differential equation (3.4) which is first-order in z.

One step of this algorithm transfers the function $a(\boldsymbol{\rho}, z)$ from a plane z_n to the next plane z_{n+1}. Hence, this can be interpreted as a multislice algorithm. In chapter 3, however, the slicing of the sample resulted from our use of the averaged potential (3.12). Here, the potential is integrated "on the fly" and the slicing of the specimen is a consequence of the integration procedure itself.

Although the algorithm is written in matrix notation, matrices are by no means necessary for a numerical implementation of it: The matrix $\underline{v}$ is diagonal, so the corresponding multiplications can be implemented element-wise. For the Laplace operator $\Delta_\perp$ in (5.11), all the methods discussed in section 3.3 can be applied.

FRCs and difference images can be used to compare the *RK1* algorithm to the *Classical* algorithm. No significant deviations can be found if identical approximations for the two-dimensional Laplace operator are used. Thus, Runge-Kutta methods can be applied to the scattering problem of electron microscopy.

We now proceed by discussing integration procedures for the non-approximated Schrödinger equation (3.2). Since this equation is a second-order differential equation in z, it must be rewritten

as a system of first-order equations before the standard numerical integrators can be used

$$\partial_z\, a(\mathbf{r}) = b(\mathbf{r}), \tag{5.17}$$
$$\partial_z\, b(\mathbf{r}) = -\Delta_\perp\, a(\mathbf{r}) + v(\mathbf{r})\, a(\mathbf{r}) - 2\,\mathrm{i}\,k\, b(\mathbf{r}). \tag{5.18}$$

The two functions a and b must be integrated simultaneously. Thus, the above scheme ((5.7) and (5.8)) has to be applied to

$$\underline{\tilde{H}} = \begin{pmatrix} 0 & 1 \\ -\Delta_\perp + \underline{v} & -2\,\mathrm{i}\,k \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}. \tag{5.19}$$

In combination with the classical Runge-Kutta algorithm (5.10), this approach will be called *RK2* since it calculates the solution for the second-order differential equation (3.2).

So far, we have only discussed explicit integrators. As mentioned above, implicit methods are more stable and allow larger step sizes Δz_n while retaining the same accuracy. Therefore, we also looked into these implicit integration methods. A fourth-order representative is the Lobatto IIIA algorithm [56] with the Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 5/24 & 1/3 & -1/24 \\ 1 & 1/6 & 2/3 & 1/6 \\ \hline & 1/6 & 2/3 & 1/6 \end{array}. \tag{5.20}$$

Applied to (5.19), we will refer to this algorithm as *Lobatto2*. If a real-space method is used for the evaluation of the Laplace operator in (5.19), then the entries of $\underline{\tilde{H}}$ are band matrices and $\underline{\tilde{H}}$ itself is a sparsely-populated matrix. This fact can be used, to reduce the memory requirement of the algorithm significantly.

The numerical properties of the algorithms *RK2* and *Lobatto2* will be investigated in section 5.3.

5.2 Modification of the Schrödinger Equation

5.2.1 Matrix Forward Algorithm

In section 5.1.1, we demonstrated that backward scattering is an integral part of the Schrödinger equation. Hence, backscattering effects can only be excluded by modifying the differential equations. For example, this can be achieved by setting $\underline{B}^n = 0$ in (4.11). This modification was first proposed by Chen and Van Dyck [22] for a similar equation.

From (4.13), we can conclude that, in this case, the matrix $\underline{F}^n$ is given by

$$\underline{F}^n = \underline{U}^{n\dagger}\, \underline{U}^{n-1}. \tag{5.21}$$

Therefore, $\underline{F}^n$ is unitary and the matrix $\underline{T}_{11}$ of (4.14) is given by the product

$$\underline{T}_{11} = \prod_{n=N}^{1} \underline{F}^n \exp\left(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z_{n-1}\right). \tag{5.22}$$

Thus, according to (4.6), the exit-wave $\boldsymbol{\psi}_{\mathrm{ex}}$ can be written as

$$\boldsymbol{\psi}_{\mathrm{ex}} = \underline{U}^N\, \underline{T}_{11}\, \underline{U}^{0\dagger} \boldsymbol{\psi}_{\mathrm{in}}. \tag{5.23}$$

If the inequality (4.24) is fulfilled, also $\exp(\mathrm{i}\,\underline{D}^n\,\Delta z_n)$ is a unitary matrix. As already discussed in connection with the paraxial approximation in section 3.1.1, this means that the probability to find the electron in a certain xy-plane is independent of n. This property can be used to check the correctness of the simulation results.

Although it may seem that unitarity is a desirable property for a simulation algorithm, it actually is not: Neglecting the backscattered current in (4.17), i.e. $b_m^n = 0$, we find

$$I_n = \frac{\hbar}{m} \sum_m D_m^n \, |f_m^n|^2 \,. \tag{5.24}$$

Therefore, only this weighted sum of the f_m^n should be conserved, but not the probability

$$\sum_m |f_m^n|^2 \,. \tag{5.25}$$

If the inequality (4.24) is violated, some entries of the matrix $\exp(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)$ become less than one and the corresponding components of the wave vector decay quickly. Thus, the algorithm remains stable nonetheless.

This stability was obtained by modifying the underlying equations. Hence, strictly speaking, the algorithm presented here does not solve the original problem. Since backscattering effects can often be neglected (cf. chapter 6), this is, however, a useful approximation. Later, we will refer to this algorithm as *Matrix Forward*.

5.2.2 Full Algorithm

The results of the previous section can be used to derive the corrected forward-scattering algorithm as proposed by Chen and Van Dyck [22].

We start by rewriting (5.23) as

$$\boldsymbol{\psi}_{\mathrm{ex}} = \underline{U}^N \left[\prod_{n=N}^{1} \underline{U}^{n\dagger}\, \underline{U}^{n-1} \, \exp\left(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z_{n-1}\right) \right] \underline{U}^{0\dagger}\, \boldsymbol{\psi}_{\mathrm{in}} = \left[\prod_{n=N}^{1} \exp\left(\mathrm{i}\,\underline{A}^{n-1}\,\Delta z_{n-1}\right) \right] \boldsymbol{\psi}_{\mathrm{in}}, \tag{5.26}$$

where we have introduced the matrix

$$\underline{A}^n = \underline{U}^n\, \underline{D}^n\, \underline{U}^{n\dagger}. \tag{5.27}$$

This matrix can be understood as a possible square root of the matrix $\underline{A}^{2,n}$. This square root is uniquely determined by virtue of (4.9). In the next step, we expand $\underline{A}^{2,n}$

$$\exp\left(\mathrm{i}\,\underline{A}^n\,\Delta z_n\right) = \exp\left(\mathrm{i}\,\Delta z_n \sqrt{\underline{A}^{2,n}}\right) \overset{(4.3)}{=} \exp\left(\mathrm{i}\,\Delta z_n \sqrt{-\underline{\mathcal{F}}^\dagger\, \underline{\kappa}^2\, \underline{\mathcal{F}} - \underline{v}^{(n)} + k^2}\right) \tag{5.28}$$

$$= \mathrm{e}^{\mathrm{i}\,\Delta z_n\, k}\, \underline{W}(z_{n+1}, z_n) \tag{5.29}$$

$$\text{with } \underline{W}(z_{n+1}, z_n) = \exp\left(\mathrm{i}\, k\, \Delta z_n \left[\sqrt{1 - \frac{\underline{\mathcal{F}}^\dagger\, \underline{\kappa}^2\, \underline{\mathcal{F}}}{k^2} - \frac{\underline{v}^{(n)}}{k^2}} - 1 \right] \right). \tag{5.30}$$

The new matrix $\underline{W}$ takes the redefinition of the phase (3.1) into account. This allows us to rewrite (5.26) in the form of (3.16)

$$\mathbf{a}_{\mathrm{ex}} = \left[\prod_{n=N}^{1} \underline{W}(z_n, z_{n-1}) \right] \mathbf{a}_{\mathrm{in}}. \tag{5.31}$$

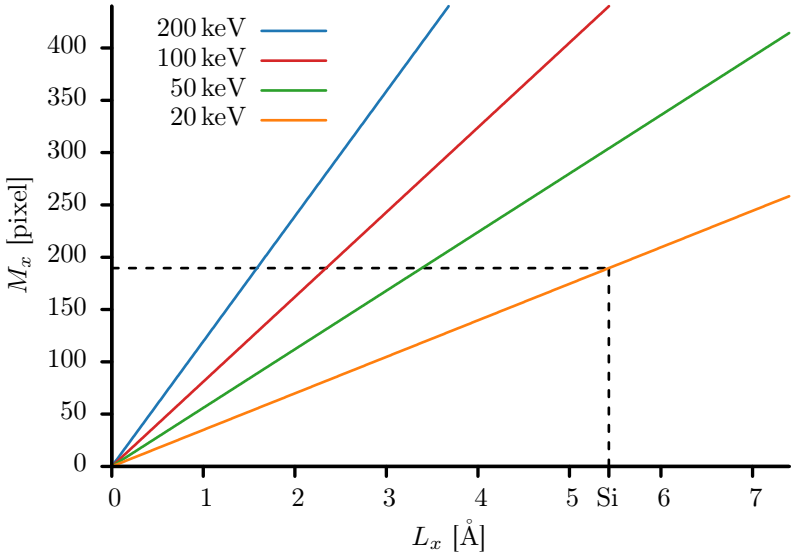

Figure 5.2: Visualization of the first inequality in (5.34) for different kinetic energies. The pixel size Δx is given by L_x/M_x. The size of the cubic supercell of silicon (5.43 Å) is marked.

In [22], the square root is further evaluated with help of the binomial series (3.7)

$$\underline{W}(z_{n+1}, z_n) = \exp\left(\mathrm{i}\, k\, \Delta z_n \sum_{\ell=1}^{\infty} \binom{1/2}{\ell} \left[-\frac{\underline{\mathcal{F}}^{\dagger}\, \underline{\kappa}^2\, \underline{\mathcal{F}}}{k^2} - \frac{\underline{v}^{(n)}}{k^2}\right]^{\ell}\right) \tag{5.32}$$

Since the binomial series only converges if $|x| \leq 1$, it cannot be used to calculate the square root of a negative number, e.g.

$$\frac{\mathrm{i}}{2} = \sqrt{1 - \frac{5}{4}} \neq \sum_{\ell=0}^{\infty} \binom{1/2}{\ell} \left(-\frac{5}{4}\right)^{\ell} = 1 + \frac{5}{8} - \frac{25}{128} + \ldots. \tag{5.33}$$

As the correct result is purely imaginary, it is obvious that the partial sums cannot even approximate it. Therefore, this algorithm can only be applied, if all eigenvalues of the matrix $\underline{A}^{2,n}$ are positive.

Note that the first term of the series expansion in (5.32) gives the paraxial approximation of the multislice algorithm (3.15).

In the following, we will refer to the algorithm of this section as *Full*. As was true for the explicit Runge-Kutta algorithms, matrices are not required for the implementation of the *Full* algorithm.

5.3 Numerical Results

The improved multislice algorithms presented in the first two sections of this chapter were implemented in Aurora (cf. appendix B).

With the exception of the *Matrix Forward* algorithm of section 5.2.1, all improved multislice algorithms show a divergent behavior if the inequality (4.24) is violated (cf. figure 5.1). The low-pass filter, that has been introduced in section 3.4 to prevent the appearance of aliasing

Algorithm	Number of Slices per Simulation Box N	Computation Times	
		376×264	752×528
Classical	40	3 s	13 s
	100	7 s	53 *s*
RK2	400	172 s	881 s
	800	575 s	1774 s
Full	40	103 s	403 s
	100	165 s	603 s

Table 5.1: Computations times for the algorithms *Classical*, *RK2* and *Full*. The measurements were taken on a system equipped with two Intel Xeon E5-2670 processors (in total: 16 physical cores) and 64 GB RAM. The simulation results were calculated for 92.2 nm of silicon [110] at an acceleration voltage of 20 kV.

effects, can be used to stabilize these algorithms, because the divergences start at high-spatial frequencies. Since the maximum transferred frequency is only ⅔ of the Nyquist frequency, the single inequality (4.24) can be replaced by the two weaker relations

$$T^2 + 2\,m\,c^2\,T > \frac{4\,c^2\,\hbar^2\,\pi^2}{9\,\Delta x^2}, \qquad T^2 + 2\,m\,c^2\,T > \frac{4\,c^2\,\hbar^2\,\pi^2}{9\,\Delta y^2}. \tag{5.34}$$

The corresponding simulation area in reciprocal space is illustrated in figure 4.2. The pixels inside the circle show no diverging behavior. However, only with the help of a low-pass filter all of these pixels can be included in the simulation. Figure 5.2 shows the inequalities (5.34) for different kinetic energies T.

Most results of this section are calculated for silicon [110] at an acceleration voltage of 20 kV. According to the relations (5.34), the minimal pixel size possible is $\Delta x = \Delta y = 2.9$ pm. This is the pixel size used for the simulations, unless otherwise noted. For the minimal simulation box (5.43 Å × 3.84 Å, cf. figure 2.5), this pixel size corresponds to a resolution of 188×132 pixels. The two-dimensional Laplace operator that appears in the *RK2* algorithm is evaluated in reciprocal space.

5.3.1 Parallel Illumination

Classical, RK2 and Full

As before, TEM brightfield images were simulated by setting the incoming wave function $\boldsymbol{\psi}_{\text{in}}$ to a plane wave in z direction. Phase shifts and potentials were calculated using the hybrid approach introduced in section 3.2.3 and the form factor parametrization of Lobato and Van Dyck ([38] and appendix A.2).

For the algorithms *Classical* and *Full,* a slice thickness Δz of 9.6 pm was chosen, resulting in 40 slices per simulation box. This value is necessary to ensure convergence of the *Full* algorithm. A smaller slice thickness of 0.96 pm was used for the *RK2* algorithm as the algorithm diverges for larger values. This is unexpected since the *RK2* approach is based on the ansatz (3.1) that removes the rapid oscillations of the wave function in z direction. The chosen slice thickness corresponds to 400 slices (integration steps) per simulation box. Since the slice thickness Δz is an unphysical parameter the simulation results must not depend on it. This was verified by repeating each simulation with a different slice thickness.

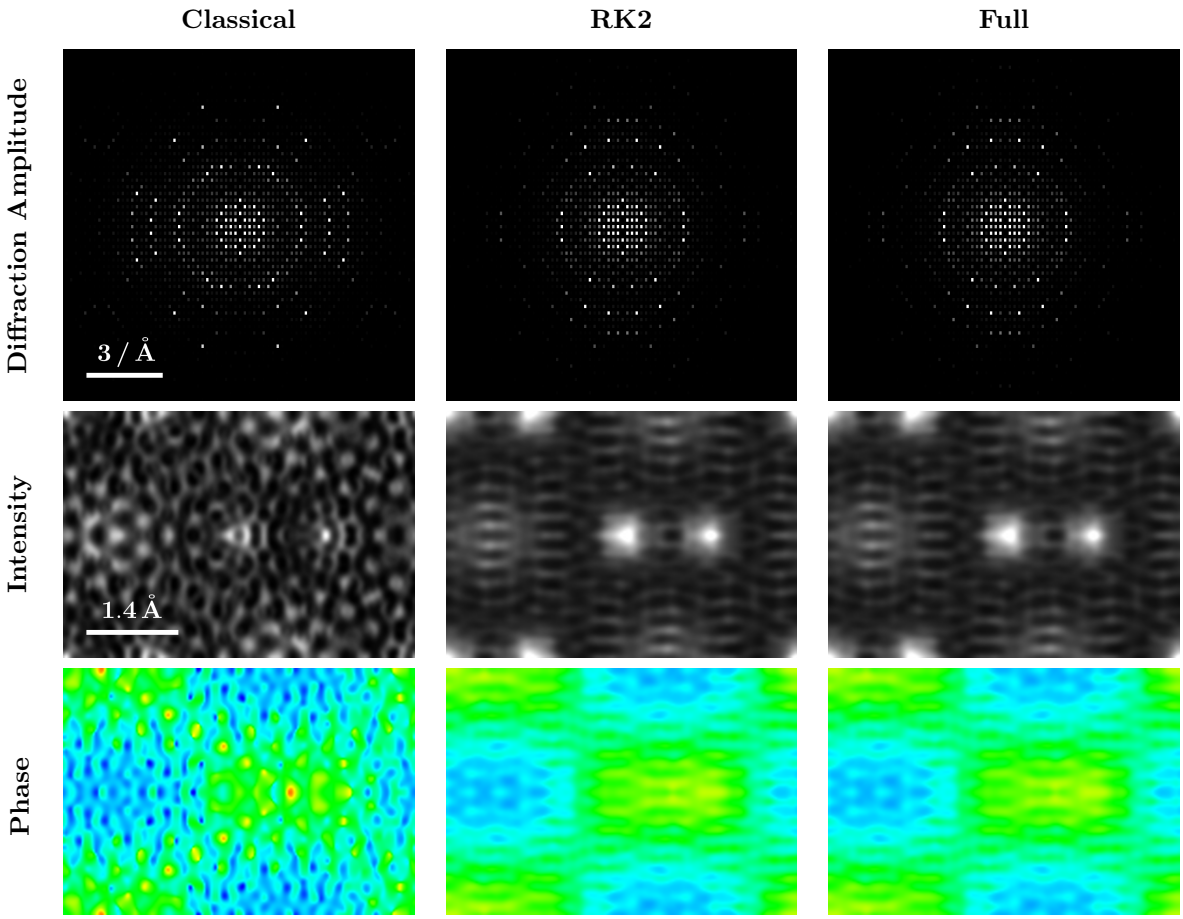

Figure 5.3: Diffraction amplitude, intensity $|\psi|^2$ and phase of the exit-wave for three different multislice algorithms: *Classical*, *RK2* and *Full*. The images were simulated for 46.1 nm silicon [110] at an acceleration voltage of 20 kV. The results from the *RK2* and *Full* algorithms are in very good agreement. For high-spatial frequencies, however, the results from the *Classical* algorithm deviate. This affects especially the HOLZ-rings in the diffraction pattern.

As can be seen in figure 5.3, the results of the *RK2* and the *Full* algorithm are in very good agreement with each other. In contrast, the result from the *Classical* algorithm shows significant deviations, especially for large spatial frequencies. These findings are confirmed by the FRCs in figure 5.4.

The differences in the diffraction patterns can be explained in large part with the help of figure 2.4. Since the *Classical* algorithm is based on the paraxial approximation, the Ewald's sphere is replaced by a paraboloid. For small scattering angles this a good approximation. For larger angles, however, this leads to wrong intersections with the rods of the reciprocal lattice. As a result, the spot intensities are incorrect and the HOLZ rings are shifted to higher spatial frequencies.

Because of the different slice thicknesses, the computation times for the three algorithms vary greatly. Exemplary computation times for the propagation step can be found in table 5.1. Note that these figures do not include the time needed for the calculation of the phase shifts and potentials. Furthermore, the time required by the *Full* algorithm depends greatly on the details of the stopping criteria for the series expansions (cf. (5.32)).

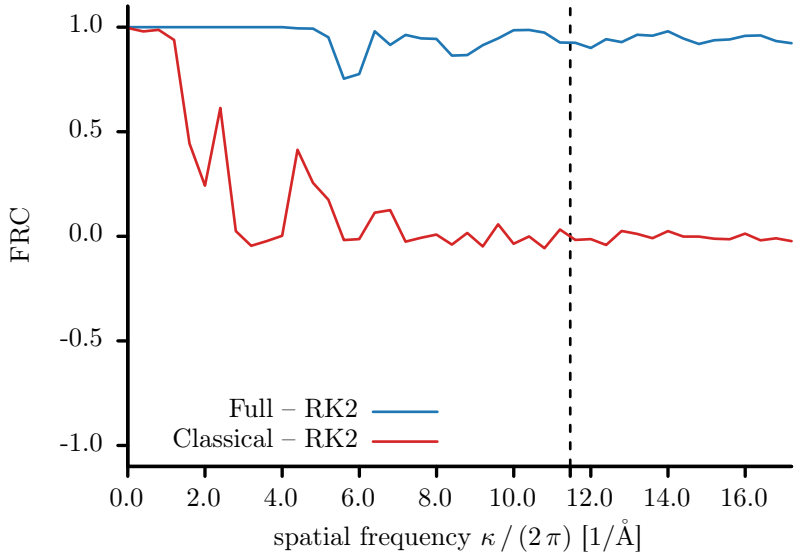

Figure 5.4: FRC comparing the results of the algorithms *Classical*, *RK2* and *Full*. Shown are the correlations for 46.1 nm of silicon [110] at an acceleration voltage 20 kV. The size of the simulation area was 5.43 Å × 3.84 Å. The dashed vertical line marks the onset of the low-pass filter. The two improved algorithms (*Full* and *RK2*) show very few deviations while the correlation of the *Classical* algorithm and the *RK2* algorithm indicates agreement only for the lowest frequencies.

Three basic image operations are used by the simulation algorithms to propagate the wave function: addition, multiplication and FFT. Of these operations, the FFT is the most costly, taking $\mathcal{O}(M \ln M)$ arithmetic operations. Thus, it defines the scaling behavior of the simulation algorithms as a whole. Moreover, the number of operations needed for the algorithms *Classical* and *RK2* depends linearly on the number of slices.

As discussed in chapter 3, the measured computation times do not always confirm the theoretical expectations. For example, the calculation time the *RK2* algorithm needs should double as the number of slices doubles. According to table 5.1, this is only the case at a resolution of 752 × 528 pixels. The deviations from the expected scaling behavior can be explained by the complicated memory hierarchy (caches) of modern computer systems and the non-uniform memory access (NUMA) in multiprocessor system. As a consequence, not only the number of operations, but also the memory access patterns are important for the performance of an algorithm.

As expected, the two improved simulation algorithms are computationally much more expensive than the *Classical* algorithm. In most cases, the *Full* algorithm outperforms the *RK2* algorithm by more than a factor of two.

Lobatto2

The *Lobatto2* algorithm is an implicit method. Therefore, (5.7) forms a system of linear equations that must be solved to obtain $\mathbf{X}_{n+1}$. For this, only real-space methods were tested. If identical approximations are used for the Laplace operator, then the results of the *RK2* and *Lobatto2* algorithms are almost identical. The enhanced stability of the implicit algorithm can, however, only be exploited, if the pixel size is significantly larger than the minimal value set by (5.34). Under these conditions, the slice thickness can be greatly increased without affecting the convergence

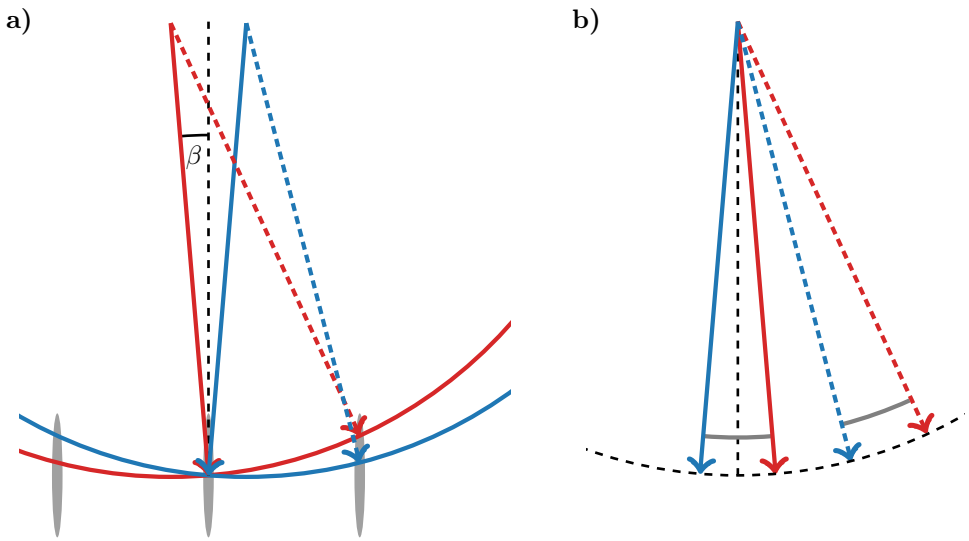

Figure 5.5: Sketch illustrating the formation of a CBED pattern. The directions of the incoming and outgoing beams are shown in reciprocal space. a) Ewald construction. The sample is illuminated by a convergent beam with semi-angle β. The red and blue solid lines represent the contours of the illumination cone. In the kinematic approximation, a beam is diffracted if the corresponding Ewald's sphere intersects a rod of the reciprocal lattice (gray). The resulting beam directions are represented by the dashed arrows. b) Directions of the outgoing beams. The unscattered beams (solid) and the scattered beams (dashed) cover certain solid angles (gray). In the back focal plane of an electron microscope, these beams appear as disks.

or the simulation results of the *Lobatto2* algorithm. Despite the reduced number of slices, the *Lobatto2* algorithm requires orders of magnitude more time than the *RK2* algorithm. Reasons for this are the higher complexity of the algorithm and the fact that the sparse linear algebra is executed only on a single core.

5.3.2 Convergent Beam Electron Diffraction

If the sample is illuminated by a convergent beam instead of a parallel beam, then a so-called convergent beam electron diffraction (CBED) pattern is formed in the back focal plane of the microscope. For thick samples, CBED patterns contain much more information about the specimen than a classical diffraction pattern since many features of the pattern are determined by dynamical scattering effects. Hence, the formation of CBED patterns depends on the full three-dimensional structure of the crystal and it is even possible to determine the space group of the sample from a single CBED pattern [2]. Therefore, CBED patterns are a most suitable test for comparing the different simulation algorithms.

Although the theory of CBED pattern formation is quite complex [57], essential features can be understood with the help of the kinematic approximation, as shown in figure 5.5. The kinematic theory, however, predicts a uniform intensity for the disks. This is only correct for very thin samples. The size of the central disks in a CBED pattern is determined by the convergence angle β. The amount of information in a CBED pattern is maximized if the disks touch each other, but do not overlap.

Typically, the multislice algorithm is used in conjunction with periodic boundary conditions. For parallel illumination, this presents no problem. However, a convergent beam breaks the

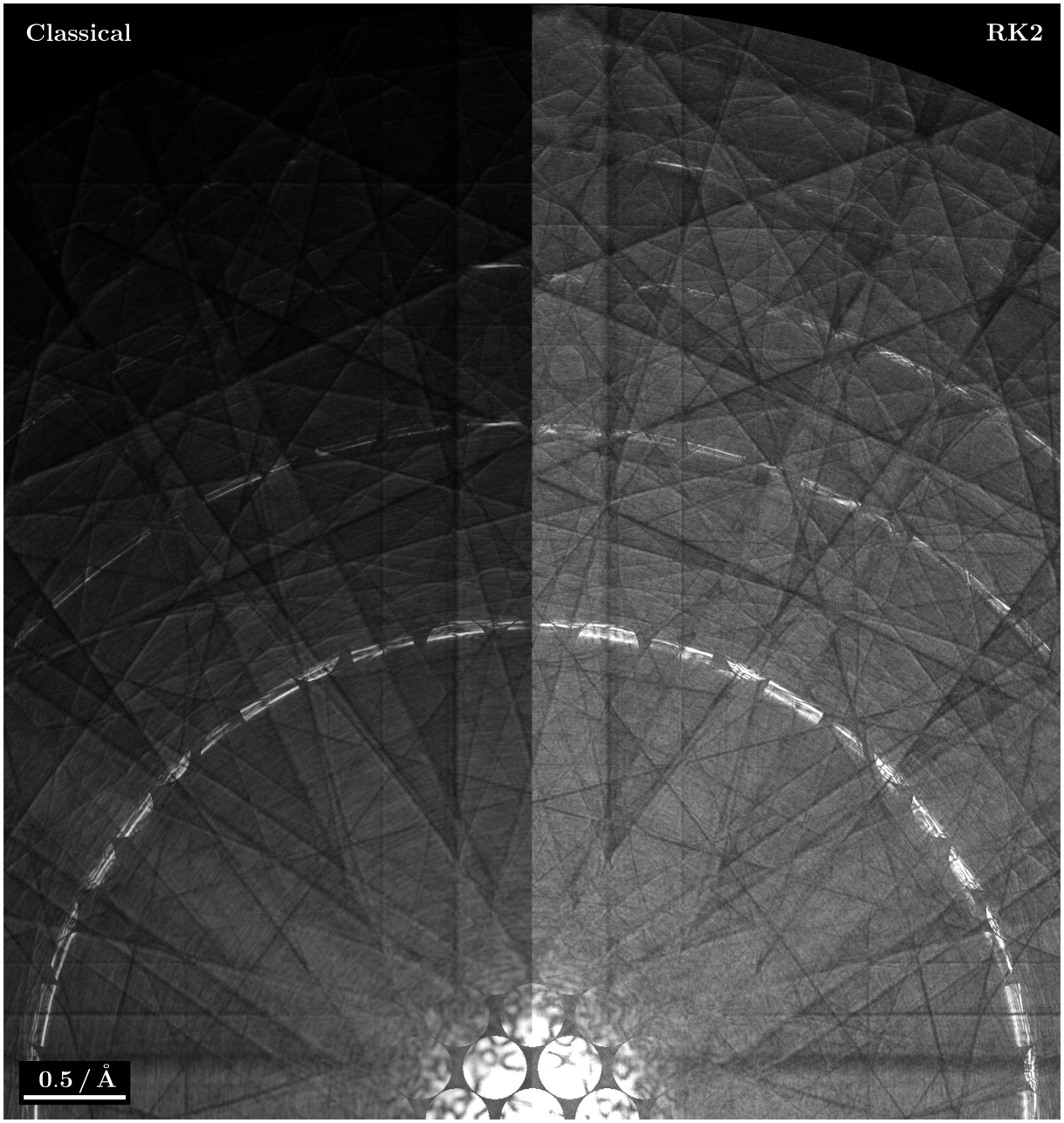

Figure 5.6: Comparison of the CBED patterns calculated using the *Classical* and the *RK2* algorithm. The intensity is represented on a logarithmic scale. Both images were post-processed in the same way. The simulated sample was 46.1 nm silicon [110] at an acceleration voltage of 20 kV. The convergence angle of the incoming beam β was 13 mrad. Broad Kikuchi bands originating from the center and several HOLZ rings can be seen.

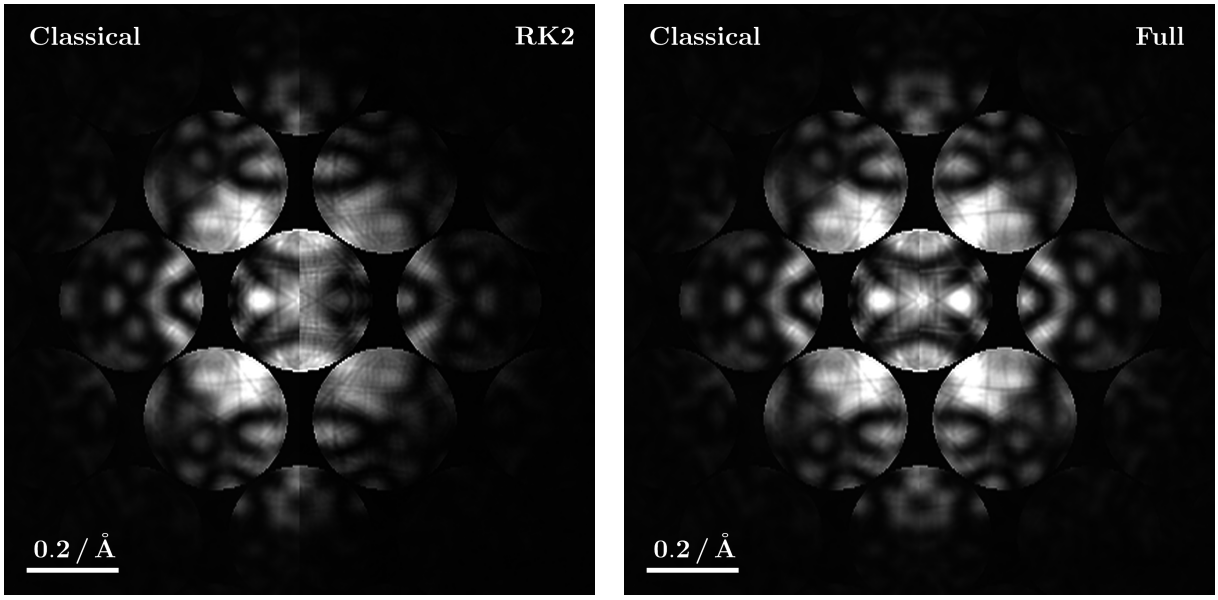

Figure 5.7: Central disks of the CBED pattern for 46.1 nm silicon [110] at an acceleration voltage of 20 kV. The results of three algorithms – *Classical*, *RK2* and *Full* – are compared. The intensity scale is linear.

translational symmetry in x- and y-direction. Thus, the simulation box must be enlarged. There is another way of looking at this: The pixel size in reciprocal space needs to be decreased to resolve the inner structure of the central disk of the CBED pattern. According to (3.40), this can only be achieved by increasing L_x and L_y.

As before, 46.1 nm silicon [110] was simulated at an acceleration voltage of 20 kV, but the size of the simulated area was enlarged to 19.5 nm × 13.8 nm. Thus, about two million atoms were included in the simulation. Likewise, the number of pixels was increased to 4752×3360. This results in a pixel size of 4.1 pm. Fifty frozen-phonon passes were used to account for TDS: Random displacements were chosen according to a Gaussian distribution whose width was determined by the B-factor. The slice thickness Δz of 1.92 pm corresponds to a number of 24 000 slices. The calculations were performed on the BioQuant computing cluster.

The results of the simulation are shown in the figures 5.6 and 5.7. According to the classical paper by Buxton et al. [57], the symmetries of the crystal, together with its zone axis orientation, determine the symmetries of the CBED pattern. This means in our case, that the central disk and the whole pattern should exhibit a 2mm-symmetry (dihedral group with four elements). Thus, the pattern should be invariant under a rotation by 180 ° and there should be two mirror symmetries. This is indeed the case, although in [57] only the paraxial approximation was discussed. As a consequence, one quadrant of the CBED pattern already contains all of the information.

Figure 5.6 compares the CBED patterns calculated with the *Classical* and the *RK2* algorithms. The broad bands originating from the center of the CBED pattern are called Kikuchi bands. For their formation, an incoherent scattering process is required. In our case, this process is provided by TDS.

The two algorithms predict different radii for the bright HOLZ rings. The mismatch increases for larger spatial frequencies. This is expected, since the *Classical* algorithm approximates the Ewald's sphere as a paraboloid. The dark lines in the two patterns are offset for the same reason.

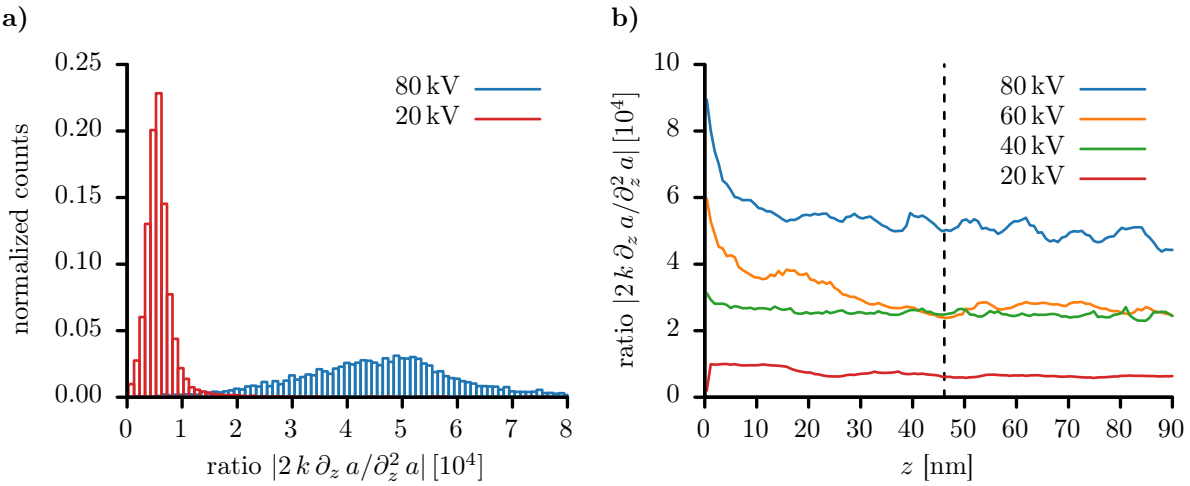

Figure 5.8: Ratios of the first and second derivatives $r = |2\,k\,\partial_z\,a/\partial_z^2\,a|$. a) Histogram showing the ratio r after 46.1 nm silicon [110] ($\Delta z = 0.48$ pm) at two different acceleration voltages. The areas of the histograms are normalized to one. b) The mean of the ratios as function of the sample thickness.

Each of these so-called HOLZ lines can be paired with a line in the central disk of the diffraction pattern. Because of the limited resolution, only some of those lines are visible in figure 5.7. Here, the central disks, calculated using the *RK2* and *Full* algorithms, are compared with the *Classical* simulation: The *RK2* results differ significantly more from the *Classical* simulation than those of the *Full* algorithm do. Furthermore, the *RK2* algorithm predicts intensity deviations for large scattering angles (cf. figure 5.6). This effect is also not confirmed by the *Full* algorithm.

5.3.3 Accumulation of Errors

We have shown that Runge-Kutta methods can be applied to find numerical solutions for the non-approximated Schrödinger equation (2.14). Now, we will use these solutions to examine the validity and applicability of the commonly used paraxial approximation of the Schrödinger equation (cf. section 3.1.1).

The paraxial approximation drops the second derivative $\partial_z^2\,a$ on the left-hand side of (3.2). This is usually justified by the statement that the contribution of this term is much smaller than the contribution of the first-derivative $2\,i\,k\,\partial_z\,a$. In general, however, the derivative of a bounded function can attain arbitrarily large values. For example, the function $\cos(\alpha\,z)$ is bounded, since $|\cos(\alpha\,z)| \leq 1$, but the maximum value of its derivative is given by α.

To check the assumption about the relative magnitude of the derivatives, we calculate the ratio $r = |2\,k\,\partial_z\,a/\partial_z^2\,a|$ for each pixel of the wave function. The histograms for these ratios after 46.1 nm of silicon [110] are shown in figure 5.8.a. They confirm that the modulus of the second derivative $|\partial_z^2\,a|$ is on average at least three orders of magnitude smaller than the modulus of the first derivative $|2\,k\,\partial_z\,a|$. As expected, the ratios r decrease at lower acceleration voltages.

Since we are dealing with differential equations, the ratios of the derivative for a single z-value are not as important as their integrated effect for the whole sample. Thus, we proceed by examining the means of the ratios as functions of z (cf. figure 5.8.b). We find, that, after a short transition region, the ratios converge to more or less constant values which depend on the acceleration voltage. If the sample is very thin ($z < 10$ nm), then the ratios display an interesting behavior: at an acceleration voltage of 80 kV, the ratios decrease starting from a very

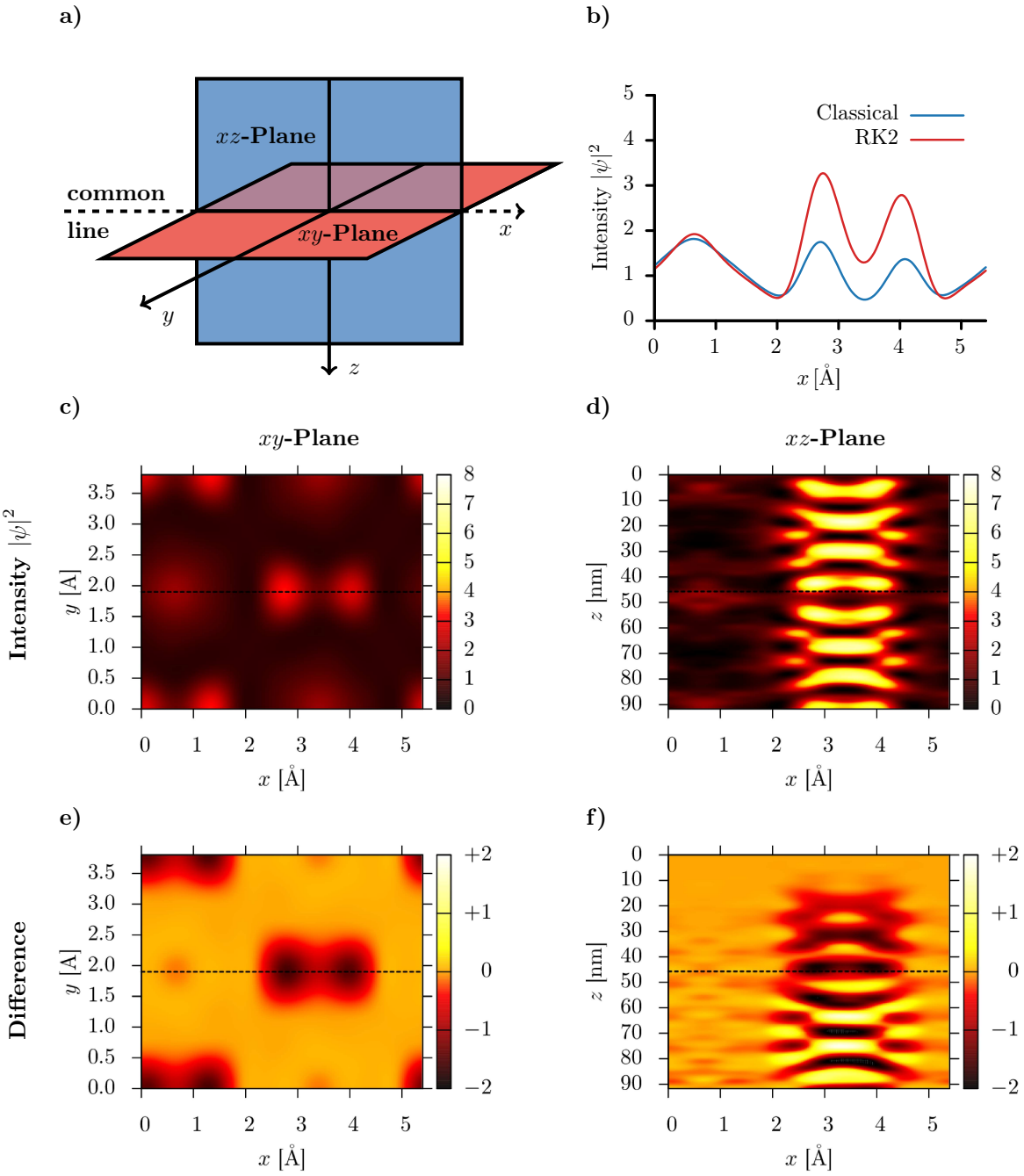

Figure 5.9: Cumulative effects of the paraxial approximation. The wave functions simulated by the *Classical* ($\Delta z = 9.6\,\text{pm}$) and the *RK2* ($\Delta z = 0.96\,\text{pm}$) algorithm are compared. The sample is silicon [110] at an acceleration voltage of 20 kV. The simulations assume an objective aperture of 80 mrad. a) Orientation of the different planes in three-dimensional space. b) Absolute square of the wave function $|\psi|^2$ along the common line of the two planes ($y = 1.94\,\text{Å}$, $z = 46.1\,\text{nm}$). c) & d) Absolute square of the wave function $|\psi|^2$ in the xy-plane and in the xz-plane as calculated by the *RK2* algorithm. e) & f) Differences to the absolute square of the wave function simulated by the *Classical* algorithm.

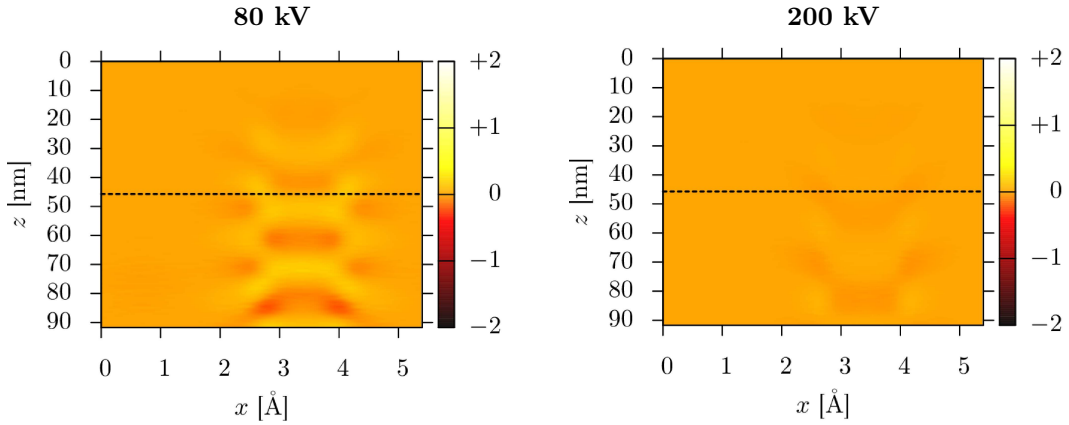

Figure 5.10: Differences in the xz-plane caused by the paraxial approximation at acceleration voltages of 80 kV and 200 kV. The corresponding image for 20 kV is shown in figure 5.9.f. Since the *RK2* algorithm converges only if the slice thickness is not too large, the slice thickness had to be adjusted accordingly. The images were calculated with a slice thickness of 0.48 pm (80 kV) and 0.384 pm (200 kV). The slice thickness for the *Classical* algorithm remained fixed at 9.6 pm.

high initial value. In contrast, at an acceleration voltage of 20 kV, the ratios start from near zero and increase, subsequently.

Despite its smallness, the second derivative has a profound influence on the wave function, as is shown in figure 5.9. This figure compares the wave functions calculated with the algorithms *Classical* and *RK2* at an acceleration voltage of 20 kV. Unlike the results shown above, an objective aperture of 80 mrad was included in the simulation. Figures 5.9.c and 5.9.d show the absolute square of the wave function calculated with the *RK2* algorithm in the xy-plane and in the xz-plane, respectively. The typical oscillatory behavior of the electron wave as a function of z is readily visible. The importance of the second derivative for the simulated images can be concluded from the images 5.9.e and 5.9.f, which show the deviations from the *Classical* calculation. Starting from zero, these deviations are steadily increasing with the sample thickness. The effect of the paraxial approximation can also be seen in figure 5.9.b. It directly compares the absolute square of the wave functions calculated with the *Classical* and the *RK2* algorithm. Thus, for thick samples, the low-spatial frequencies show deviations as well, although only high-spatial frequencies are directly affected by the paraxial approximation.

At higher acceleration voltages, the impact of the paraxial approximation is greatly reduced, as shown in figure 5.10. Nevertheless, the deviations still increase with the thickness of the sample.

All of these effects can be understood by noting that the solution of (3.2) can be expressed as a series of integrals (Picard iteration). Hence, small effects can accumulate during the propagation of the wave, if there is no physical argument demanding their cancellation. To conclude, it is not only a question of the acceleration voltage, whether the paraxial approximation can be applied, but also a question of the sample thickness.

5.4 Discussion

In this chapter, we discussed and compared different simulation algorithms that offer an improved treatment of the large scattering angles. As was already shown in section 3.3, the correct

Class	Algorithm	Equations
Paraxial approx.	*Classical* [17, 44, 60]	(3.16) and (3.20)
	Single [45]	(3.16) and (3.15)
	RK1	(5.7), (5.8), (5.10) and (5.11)
Beyond the paraxial approx.	*Full* [22]	(3.16), (5.32)
	RK2	(5.7), (5.8), (5.10) and (5.19)
	Lobatto2	(5.7), (5.8), (5.20) and (5.19)

Table 5.2: Overview of the simulation algorithms for forward scattering.

treatment of theses angles requires that the two-dimensional Laplace operator is evaluated in reciprocal space. For this reason, we combined the *Full* algorithm, as proposed by Chen and Van Dyck [22], with an FFT. In contrast, previous numerical studies evaluated the Laplace operator always in real space [45, 58]. Our derivation of the *Full* algorithm is based on the eigenvalue decomposition of the matrix $\underline{A}^{2,n}$. Thereby, we avoided the dubious operator-valued square root that was used in [22].

Furthermore, we could demonstrate that it is possible to directly integrate the non-approximated Schrödinger equation using explicit and implicit numerical methods, the Runge-Kutta algorithms. In comparison to the *Full* algorithm, the derivation of this class of algorithms requires much less mathematical machinery. Even though, for an in-depth understanding, the results from chapter 4 are needed.

The importance of the second derivative has been conjectured since 1975 [59]. The results of section 5.3 clearly show that the paraxial approximation is only applicable for thin samples and at high acceleration voltages $\gtrsim 80\,\mathrm{kV}$. Since lower acceleration voltages have become experimentally available, the simulation programs need to be improved accordingly.

The results of the *Full* and the *RK2* algorithm are essentially in good agreement, although the CBED patterns revealed some differences between these two approaches. This is not unexpected, since the algorithms are derived from different equations and boundary conditions: The *Full* algorithm neglects backscattering effects explicitly by modifying the underlying equations, while the *RK2* algorithm employs unphysical initial conditions. As a consequence, the *Full* algorithm is unitary, unlike the *RK2* algorithm. Furthermore, both algorithms fail for different reasons if the inequality (4.24) is violated: The divergence of the *RK2* algorithm is a manifest property of the Schrödinger equation for the given initial conditions. The *Full* algorithm fails because of the inappropriate use of the binomial series, even though the *Matrix Forward* solution is always well defined. A low-pass filter – originally introduced to counter aliasing artifacts – can be employed to suppress the arising divergences, if the size of the simulation area and the pixel size are chosen carefully.

All simulation algorithms discussed up until now are summarized in table 5.2. In the next chapter, backscattering effects will be included.

Chapter 6

Backscattering

The simulation results of the previous chapter are incomplete, since backscattering effects were not considered. Therefore, we will now discuss the necessary steps to turn the matrix approach of chapter 4 into a practical, numerical algorithm. An implementation of this algorithm is possible, as there are efficient numerical libraries to handle matrix algebra and eigenvalue decompositions.

6.1 Toy Problems

Before we investigate the complete three-dimensional *Matrix Multislice* algorithm, we consider a number of simpler, one-dimensional toy problems since many important properties of the backscattered signal can be studied in this simpler setting. In a first step, we will establish that the analytical and numerical solutions deliver small, but identical results for the reflection coefficient. In section 6.1.2, the violation of the inequality (4.24) will be discussed in the context of backscattering.

6.1.1 Potential Wells

Since atomic potentials are attractive, we consider the solution of the one-dimensional Schrödinger equation (5.1) for potential wells ($v(z) \leq 0$). To mimic the multislice procedure of chapter 3, the z-axis is subdivided into N intervals. On each of these intervals, the potential is replaced by its averaged value (cf. (3.12))

$$v^{(n)} = \int_{z_n}^{z_{n+1}} v(z)\,\mathrm{d}z. \tag{6.1}$$

The equations (4.11), (4.12), (4.13) and (4.9) describing the three-dimensional solution can be specialized to one dimension as well

$$\begin{pmatrix} f^n \\ b^n \end{pmatrix} = \begin{pmatrix} F^n \exp\left(\mathrm{i}\,D^{n-1}\Delta z_{n-1}\right) & B^n \exp\left(-\mathrm{i}\,D^{n-1}\Delta z_{n-1}\right) \\ B^n \exp\left(\mathrm{i}\,D^{n-1}\Delta z_{n-1}\right) & F^n \exp\left(-\mathrm{i}\,D^{n-1}\Delta z_{n-1}\right) \end{pmatrix} \begin{pmatrix} f^{n-1} \\ b^{n-1} \end{pmatrix}, \tag{6.2}$$

$$F^n = \frac{1}{2}\left(1 + \frac{D^{n-1}}{D^n}\right), \tag{6.3}$$

$$B^n = \frac{1}{2}\left(1 - \frac{D^{n-1}}{D^n}\right), \tag{6.4}$$

$$D^n = \sqrt{k^2 - v^{(n)}}. \tag{6.5}$$

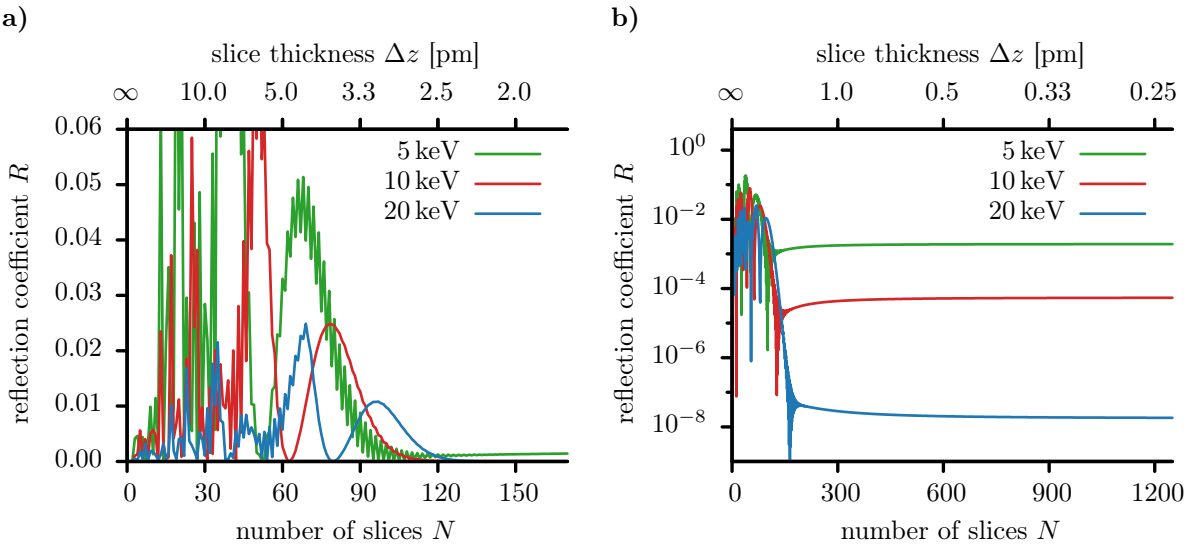

Figure 6.1: The reflection coefficient R as a function of the number of slices N for a single silicon atom placed at $z = 0$. The potential was calculated using the form factor parametrization by Peng et al. (cf. figure 6.2). The calculations were restricted to the interval $-1.5\,\text{Å} < z < 1.5\,\text{Å}$. a) For small N, arbitrary results are obtained. b) The logarithmic plot reveals that all reflection coefficients converge toward finite values greater than zero.

The matrices in (6.2) can be multiplied to find the transfer matrix connecting the first and the last slice

$$\begin{pmatrix} f^N \\ b^N \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} f^1 \\ b^1 \end{pmatrix}. \tag{6.6}$$

As described in section 4.2, electrons shall not hit the sample from the side of the camera. Using the corresponding boundary condition $b^N = 0$, we obtain for the reflection coefficient

$$R = \frac{j_\text{back}}{j_\text{in}} = \frac{|b^1|^2}{|f^1|^2} = \frac{|T_{21}|^2}{|T_{22}|^2}. \tag{6.7}$$

To avoid the problems caused by the Coulomb singularity, we utilize the form factor parametrization by Peng et al. (cf. appendix A.1). This parametrization is based on Gaussian functions so that the potential remains finite at the origin.

The described procedure can be used to investigate the dependence of the reflection coefficient on the number of slices. For this purpose, a single silicon atom was simulated. The length of the considered interval was 3 Å. The results of the calculation are shown in figure 6.1. For small slice thicknesses Δz, erratic results were obtained. The results are reliable only if the slice thickness becomes much smaller than the wavelength of the electron (e.g. $\lambda = 12\,\text{pm}$ for 10 keV). This comes as no surprise since a robust integration method for a highly oscillating function requires at least ten sampling points per oscillation.

The obtained values for the reflection coefficients are quite small and the speed of convergence is low. Therefore, the correctness of this numerical procedure was confirmed by comparing it with analytical solutions. Unfortunately, there are only very few potentials that allow an analytical treatment of the Schrödinger equation. Even for Gaussian functions, as used in the form factor parametrization by Peng et al., an analytical solution is not known. The rectangular potential well and the Pöschl-Teller potential are, however, analytically tractable and can be considered as rough approximations for the atomic potential (cf. figure 6.2).

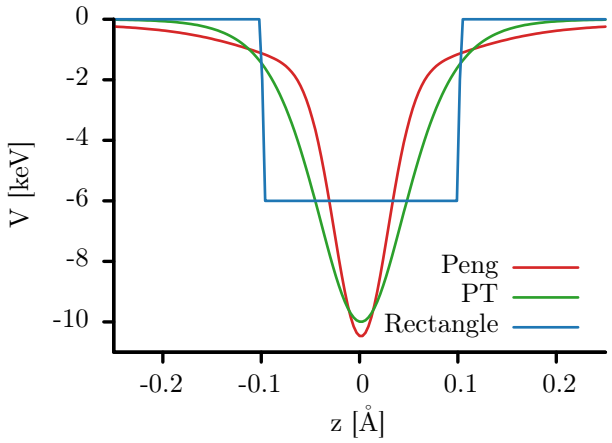

Figure 6.2: One-dimensional potentials approximating the effect of a single silicon atom at $z = 0$. Shown are the form factor parametrization of Peng et al. based on Gaussian function (cf. appendix A.1), the Pöschl-Teller (PT) potential (cf. (6.13), $a = 1/16$ Å, $V_0 = -10$ keV) and a rectangular potential well (cf. (6.8), $a = 0.1$ Å, $V_0 = -6$ keV).

The potential for a rectangular well of width $2\,a$ is given by

$$v(z) = \begin{cases} v_0 < 0 & \text{if } -a < z < a \\ 0 & \text{else} \end{cases}. \tag{6.8}$$

Since the potential is piecewise constant, the equations (6.2)–(6.5) can be applied to derive the exact solution for this problem: The potential (6.8) is subdivided into three slices ($N = 3$) of length $\Delta z = 2\,a$ each. Thus, the wave function before and after the potential well is given by (cf. (4.8))

$$\psi(z) = \begin{cases} f^1\,\mathrm{e}^{\mathrm{i}\,k(z+a)} + b^1\,\mathrm{e}^{-\mathrm{i}\,k(z+a)} & \text{if } z < -a \\ f^3\,\mathrm{e}^{\mathrm{i}\,k(z-3\,a)} + b^3\,\mathrm{e}^{-\mathrm{i}\,k(z-3\,a)} & \text{if } z > a \end{cases}. \tag{6.9}$$

The corresponding coefficients are connected by the matrix equation

$$\begin{pmatrix} f^3 \\ b^3 \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{2\,\mathrm{i}\,k\,a}\left[\cos 2\,K\,a + \frac{\mathrm{i}\,h^+}{2}\,\sin 2\,K\,a\right] & \mathrm{e}^{-2\,\mathrm{i}\,k\,a}\,\frac{\mathrm{i}\,h^-}{2}\,\sin 2\,K\,a \\ -\mathrm{e}^{2\,\mathrm{i}\,k\,a}\,\frac{\mathrm{i}\,h^-}{2}\,\sin 2\,K\,a & \mathrm{e}^{-2\,\mathrm{i}\,k\,a}\left[\cos 2\,K\,a - \frac{\mathrm{i}\,h^+}{2}\,\sin 2\,K\,a\right] \end{pmatrix} \begin{pmatrix} f^1 \\ b^1 \end{pmatrix} \tag{6.10}$$

$$\text{with } h^\pm = \frac{K}{k} \pm \frac{k}{K} \text{ and } K = \sqrt{k^2 - v_0}. \tag{6.11}$$

According to (6.7), the reflection coefficient can hence be written as

$$R = \frac{\left|h^-\,\sin 2\,K\,a\right|^2}{\left|2\cos 2\,K\,a - \mathrm{i}\,h^+\,\sin 2\,K\,a\right|^2}. \tag{6.12}$$

Potential	$1/a\,[1/\text{Å}]$	$\frac{\hbar^2 v_0}{2m}$ [keV]	Reflection Coefficient R Analytical	Numerical
Peng			–	$1.8067 \cdot 10^{-8}$
Rectangle	10	-6	0.015	–
Pöschl-Teller	14	-10	$9.849 \cdot 10^{-15}$	$9.852 \cdot 10^{-15}$
Pöschl-Teller	16	-10	$9.309 \cdot 10^{-13}$	$9.307 \cdot 10^{-13}$
Pöschl-Teller	18	-10	$3.675 \cdot 10^{-11}$	$3.676 \cdot 10^{-11}$

Table 6.1: Reflection coefficients R calculated for an electron energy of 20 keV. For the numerical procedure the range $-1.5\,\text{Å} < z < 1.5\,\text{Å}$ was subdivided into 8000 intervals.

The one-dimensional Schrödinger equation (5.1) can be regarded as the equation of motion for a parametric oscillator. Thus, the results of the numerical calculations as outlined above can only be meaningful, if the potential changes adiabatically with z. Since the rectangular potential is discontinuous, it does not fulfill this condition. A much better approximation for the atomic potential is provided the Pöschl-Teller potential that is based on a hyperbolic secant

$$v(z) = v_0 \operatorname{sech}\frac{z}{a} = \frac{2\,v_0}{\mathrm{e}^{z/a} + \mathrm{e}^{-z/a}}. \tag{6.13}$$

The resulting wave function can be expressed in terms of hypergeometric functions (§25 in [35]). The reflection coefficient is given by

$$R = 1 - \frac{\sinh^2 \pi\,k\,a}{\sinh^2 \pi\,k\,a + \cos^2\left(\frac{\pi}{2}\sqrt{1 - 4\,v_0\,a^2}\right)}. \tag{6.14}$$

The results of the different calculations are shown in table 6.1. The parameters a and v_0 were chosen to match the atomic potentials as well as possible. The analytical and numerical results for the Pöschl-Teller potential agree reasonably well. Thus, the numerical procedure can indeed be used for the determination of reflection coefficients. Even small changes in the shape of the potentials affect the reflection coefficients greatly. The coefficient for a discontinuous potential is several orders of magnitude larger, even though it covers roughly the same area.

An early attempt to estimate the magnitude of the backscattered signal was made by Lynch and Moodie [21]. Yet, they used only a single rectangular potential well, and thus their result should be regarded as unreliable.

6.1.2 Potential Barriers

In the previous section, we investigated the one-dimensional Schrödinger equation for a potential well ($v < 0$). This situation is common inside the projection of the Ewald's sphere (cf. section 4.3). Outside of the projection, potential barriers ($v > k^2$) occur. Here, we will consider only the simplest example of a rectangular barrier

$$v(z) = \begin{cases} v_0 > k^2 & \text{if } -a < z < a \\ 0 & \text{else} \end{cases}. \tag{6.15}$$

The wave function for $|z| > a$ is still given by (6.9). We can easily adapt the transfer matrix (6.10) to the new situation

$$\begin{pmatrix} f^3 \\ b^3 \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{2\,\mathrm{i}\,k\,a}\left[\cosh 2\,\kappa\,a - \frac{\mathrm{i}\,\eta^-}{2}\,\sinh 2\,\kappa\,a\right] & -\mathrm{e}^{-2\,\mathrm{i}\,k\,a}\,\frac{\mathrm{i}\,\eta^+}{2}\,\sinh 2\,\kappa\,a \\ \mathrm{e}^{2\,\mathrm{i}\,k\,a}\,\frac{\mathrm{i}\,\eta^+}{2}\,\sinh 2\,\kappa\,a & \mathrm{e}^{-2\,\mathrm{i}\,k\,a}\left[\cosh 2\,\kappa\,a + \frac{\mathrm{i}\,\eta^-}{2}\,\sinh 2\,\kappa\,a\right] \end{pmatrix} \begin{pmatrix} f^1 \\ b^1 \end{pmatrix} \tag{6.16}$$

$$\text{with } \eta^{\pm} = \frac{\kappa}{k} \pm \frac{k}{\kappa} \text{ and } \kappa = \sqrt{v_0 - k^2}. \tag{6.17}$$

As before, we suppress a left-moving wave in the region $z > a$ by setting $b^3 = 0$. Thus, the coefficients f^1 and b^1 are related by

$$b^1 = -\frac{\mathrm{e}^{4\,\mathrm{i}\,k\,a}\,\mathrm{i}\,\eta^+\,\sinh 2\,\kappa\,a}{2\,\cosh 2\,\kappa\,a + \mathrm{i}\,\eta^-\,\sinh 2\,\kappa\,a}\,f^1. \tag{6.18}$$

This equation corresponds to (4.19). For large barriers ($a \gg 1/\kappa$), the approximation

$$b^1 \approx -\mathrm{e}^{4\,\mathrm{i}\,k\,a}\,\frac{\mathrm{i}\,\eta^+}{2 + \mathrm{i}\,\eta^-}\,f^1 \tag{6.19}$$

can be used. Thus, f^1 and b^1 are of the same order of magnitude.

This result has been derived by a simple analytical calculation. The entries of the matrix in (6.16), however, become exponentially large, rendering a direct numerical treatment impossible. This problem is, once again, caused by the violation of the inequality (4.23).

6.2 The Matrix Multislice Algorithm

With the results of the toy problems in mind, we can turn the algorithm of chapter 4 into a practical simulation approach.

In section 3.4, we pointed out that discretized signals need to be low-pass filtered to eliminate aliasing effects. As discussed in chapter 5, the low-pass filter must now fulfill a second purpose: It is used to suppress the divergent solutions outside the projection of the Ewald's sphere.

The matrix $\underline{L}$ shall represent the effect of the low-pass filter in real space

$$\underline{L} = \underline{\mathcal{F}}^{\dagger}\,\underline{f}\,\underline{\mathcal{F}}. \tag{6.20}$$

Here, $\underline{f}$ is a diagonal matrix whose elements are given by the Fermi function (3.49). The vectors $\mathbf{f}^n$ and $\mathbf{b}^n$ are defined with respect to the eigenbasis of the matrix $\underline{A}^n$. Using the matrix $\underline{U}^n$, the two vectors can be transformed to real space where the low-pass filter $\underline{L}$ can be applied. The transfer equations (4.11) – (4.13) are modified accordingly

$$\begin{pmatrix} \mathbf{f}^n \\ \mathbf{b}^n \end{pmatrix} = \begin{pmatrix} \underline{F}^n\,\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\,\exp(\mathrm{i}\,\underline{D}^{n-1}\Delta z) & \underline{B}^n\,\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\,\exp(-\mathrm{i}\,\underline{D}^{n-1}\Delta z) \\ \underline{B}^n\,\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\,\exp(\mathrm{i}\,\underline{D}^{n-1}\Delta z) & \underline{F}^n\,\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\,\exp(-\mathrm{i}\,\underline{D}^{n-1}\Delta z) \end{pmatrix} \begin{pmatrix} \mathbf{f}^{n-1} \\ \mathbf{b}^{n-1} \end{pmatrix}, \tag{6.21}$$

$$\underline{F}^n = \frac{1}{2}\left[\underline{U}^{n\dagger}\,\underline{L}\,\underline{U}^{n-1} + \underline{U}^{n\dagger}\,L\,\underline{U}^n\,(\underline{D}^n)^{-1}\,\underline{U}^{n\dagger}\,\underline{L}\,\underline{U}^{n-1}\underline{D}^{n-1}\right], \tag{6.22}$$

$$\underline{B}^n = \underline{U}^{n\dagger}\,\underline{L}\,\underline{U}^{n-1} - \underline{F}^n. \tag{6.23}$$

This representation is apparently not the computationally most efficient form, but it allows tracking the insertions that have been made. Note, that the diagonal elements of the matrix $\exp\left(\mathrm{i}\,\underline{D}^{n-1}\Delta z\right)$ must be low-pass filtered as well.

As a consequence of the filtering procedure, all entries of the block matrix in (6.21) are singular matrices. Hence, the matrix T_{22} in (4.14) is singular as well, and $\mathbf{b}^0$ and $\mathbf{f}^N$ cannot be calculated using (4.19) and (4.20) directly.

The problem can be avoided by inverting (4.11) before the low-pass filter matrices are inserted. This results in

$$\begin{pmatrix}\mathbf{f}^{n-1}\\ \mathbf{b}^{n-1}\end{pmatrix} = \begin{pmatrix}\exp(-\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{F}}^n & \exp(-\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{B}}^n\\ \exp(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{B}}^n & \exp(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{F}}^n\end{pmatrix}\begin{pmatrix}\mathbf{f}^n\\ \mathbf{b}^n\end{pmatrix}, \tag{6.24}$$

where we have introduced

$$\underline{\tilde{F}}^n = \frac{1}{2}\left[\underline{U}^{n-1\dagger}\,\underline{U}^n + \left(\underline{D}^{n-1}\right)^{-1}\underline{U}^{n-1\dagger}\,\underline{U}^n\,\underline{D}^n\right], \tag{6.25}$$

$$\underline{\tilde{B}}^n = \underline{U}^{n-1\dagger}\,\underline{U}^n - \underline{\tilde{F}}^n. \tag{6.26}$$

By combining the matrices in (6.24), we obtain another description of the scattering problem

$$\begin{pmatrix}\mathbf{f}^0\\ \mathbf{b}^0\end{pmatrix} = \begin{pmatrix}\underline{S}_{11} & \underline{S}_{12}\\ \underline{S}_{21} & \underline{S}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{f}^N\\ \mathbf{b}^N\end{pmatrix}. \tag{6.27}$$

If the filter matrices $\underline{L}$ are not included, the equations (4.14) and (6.27) are, of course, equivalent.

As before, the low-pass filter matrices must be inserted. This turns (6.24)–(6.26) into

$$\begin{pmatrix}\mathbf{f}^{n-1}\\ \mathbf{b}^{n-1}\end{pmatrix} = \begin{pmatrix}\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\exp(-\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{F}}^n & \underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\exp(-\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{B}}^n\\ \underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\exp(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{B}}^n & \underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\exp(\mathrm{i}\,\underline{D}^{n-1}\,\Delta z)\,\underline{\tilde{F}}^n\end{pmatrix}\begin{pmatrix}\mathbf{f}^n\\ \mathbf{b}^n\end{pmatrix}, \tag{6.28}$$

$$\underline{\tilde{F}}^n = \frac{1}{2}\left[\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^n + \underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^{n-1}\left(\underline{D}^{n-1}\right)^{-1}\underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^n\,\underline{D}^n\right], \tag{6.29}$$

$$\underline{\tilde{B}}^n = \underline{U}^{n-1\dagger}\,\underline{L}\,\underline{U}^n - \underline{\tilde{F}}^n. \tag{6.30}$$

As in section 4.2, the boundary condition $\underline{U}^N\,\mathbf{b}^N = 0$ is imposed. Thus, (6.27) implies

$$\mathbf{b}^0 = \underline{S}_{21}\,\mathbf{f}^N. \tag{6.31}$$

In combination with (4.20)

$$\mathbf{f}^N = \underline{T}_{11}\,\mathbf{f}^0 + \underline{T}_{12}\,\mathbf{b}^0, \tag{6.32}$$

the two equations can be regarded as a single fixed-point equation for the vector $(\mathbf{f}^N, \mathbf{b}^0)$. According to the fixed-point theorem of Banach, this equation can be solved iteratively if the eigenvalues of $\underline{T}_{12}$ and $\underline{S}_{21}$ are bounded by 1. Since the reflection probabilities described by these two matrices are expected to be much smaller than the transmission probabilities, a rapid convergence of this procedure can be expected. In other words, this approach should work, as long as backscattering can be regarded as perturbation. Since the equations are not solved directly, the singular nature of the matrices has no consequences.

Sample	T [keV]	L_x [Å]	L_y [Å]	M_x	M_y	Δz [pm]
Silicon [110]	15	5.43	3.84	160	112	0.48
Silicon [110]	12.5	5.43	3.84	148	102	0.48
Silicon [110]	10.0	5.43	3.84	132	92	0.48
Silicon [110]	7.5	5.43	3.84	112	80	0.48
Silicon [110]	5	5.43	3.84	92	64	0.48
Silicon [110]	5	10.86	7.86	184	132	0.48
Graphene	5	9.84	8.52	168	144	0.5

Table 6.2: Parameters for the simulations using the *Matrix Multislice* algorithm.

6.2.1 Implementation of the Algorithm

The algorithm described has been integrated in the software suite Aurora (cf. appendix B). The matrices involved can become quite large, e.g. for the simulation of graphene (cf. table 6.2) the matrix $\underline{A}^n$ has around 585 million entries and takes about 8.9 GB of memory. The scaling of the memory requirements is quadratic in the number of pixels M. This has become a limiting factor, since the current implementation requires the three matrices involved in a multiplication to be kept in main memory.

The dependence of the computation times on the number of pixels is even worse: Typical matrix operations like multiplications or the calculation of decompositions scale with M^3. Moreover, as shown in section 6.1.1, a fairly large number of these matrices is needed to obtain reliable results. Therefore, it is advisable to exploit the parallelizability of the above algorithm.

First of all, the construction of the matrices $\underline{A}^n$ using (4.21) and the subsequent determination of the eigenvalue decompositions according to (4.5) can be run in parallel. Next, the transfer matrices (6.21) and (6.28) are calculated using the information from the decompositions. As the transfer matrices do not depend on each other, these calculations can be executed in parallel as well. To obtain the final transfer matrices ((4.14) and (6.27)), the previously calculated matrices must be multiplied. By organizing these multiplications hierarchically in a divide and conquer manner, the required wall time can be minimized.

The translational symmetry can be used to significantly speed up the calculations for crystals: If the transfer matrix for a single primitive cell is known, the matrices for thicker crystals can be obtained by calculating powers of that matrix.

It is, however, necessary to simulate the surface of the sample as well, since the potentials need to be turned on and off adiabatically (cf. section 6.1.1). For this purpose, additional vacuum slices have been included in the simulations.

This amounts to a large computational effort. For example, the simulation of silicon [110] ($L_x \times L_y = 10.86\,\text{Å} \times 7.86\,\text{Å}$) at an acceleration voltage of 5 kV took about 150 000 CPU hours on an Intel Xeon E5-2640 v3 processor.

6.3 Numerical Results

The backscattered signal was investigated for different samples at various acceleration voltages (cf. table 6.2). All calculations were performed either on the bwUniCluster[a] or the bwForCluster

[a]The bwUniCluster is funded by the Ministry of Science, Research and the Arts Baden-Württemberg and the Universities of the State of Baden-Württemberg, Germany, within the framework program bwHPC.

Figure 6.3: Simulation results for 46.8 nm silicon [110] at an acceleration voltage of 5 kV including backscattering effects. The images shown are calculated for an objective aperture of 100 mrad. a) Intensity of the wave-function $|\psi|^2$ in the detector plane. b) Corrections caused by the inclusion of backscattering. Shown are the differences with respect to the *Matrix Forward* algorithm described in section 5.2.1. The contrast of the difference image has been enhanced by a factor of 50. c) Intensity of the backscattered wave. This is the signal that would be found by a hypothetical second optical system and detector on top of the sample. The signal has been amplified by a factor of $2 \cdot 10^{13}$.

(MLS & WISO)[b].

The slice thicknesses were chosen considering the results of section 6.1.1. Using the full simulations, we verified that the final results do not depend on the number of slices.

The numbers of pixels were determined with help of (5.34). To ensure that the simulated images have the correct symmetries, the atoms must be positioned with subpixel precision (cf. section 2.5). Hence, the distances inside the primitive cell were taken into account as well.

The atomic potentials were calculated using the form factor parametrization by Lobato and Van Dyck (cf. appendix A.2).

6.3.1 Parallel Illumination

Figure 6.3 shows simulated wave functions for 46.8 nm silicon [110] and parallel illumination. The simulation box encompasses 2.88 Å before the first and after the last atom. A very low acceleration voltage of 5 kV was chosen since the results of section 6.1.1 indicated that the backscattered signal would be tiny. As can be seen in figure 6.3.b, the inclusion of the backscattering effects indeed influences the electron distribution in forward direction. In comparison to the large effects caused by the paraxial approximation (cf. chapter 5), these corrections are, however, negligible.

Figure 6.3.c shows the backscattered wave which moves in the direction of the electron gun. Clearly, this wave cannot be imaged in a real electron microscope. In the next section, the backscattered wave will be used to calculate the SEM signal in a more realistic setup. As already shown in figure 5.9.d, the spatial distribution of the intensity changes a lot while the wave is propagating through the sample. Therefore, the two atoms are not always as clearly separated, as shown in figure 6.3.

For 5 kV, only a small fraction of the electrons is scattered in backward direction. This fraction can be quantified by calculating the probability currents associated with the incoming wave I_{in} and the backscattered wave I_{ba}. We obtain $I_{\text{ba}}/I_{\text{in}} \approx 2 \cdot 10^{-13}$. Unfortunately, it is not possible to confirm this small value experimentally, since the elastically scattered electrons cannot

[b]The bwHPC initiative and the bwHPC-C5 project provided the compute services of the bwForCluster MLS & WISO (Production) at Heidelberg University and the University of Mannheim. They are funded by the Ministry of Science, Research and Arts Baden-Württemberg and the German Research Foundation (DFG).

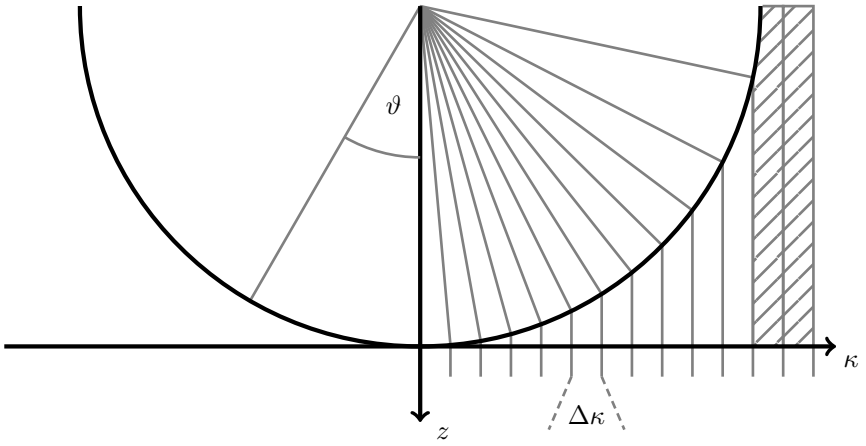

Figure 6.4: Non-uniform sampling of the scattering angle ϑ caused by the uniform sampling in the $\kappa_x\kappa_y$-plane. Shown is a cross section of the Ewald's sphere. Pixels violating the inequality (4.24) are excluded by means of a low-pass filter.

be separated from electrons that suffered a small energy loss.

A form factor parametrization and the first Born approximation can be used to calculate the fraction of backscattered electrons. This results in a cross section of about $10^{-24}\,\mathrm{m}^2$. On the other hand, the fraction $I_{\mathrm{ba}}/I_{\mathrm{in}}$ can be turned into a cross section as well, by relating it to the size of the simulation area of about $20\,\mathrm{Å}^2$. The resulting cross section of $4 \cdot 10^{-32}\,\mathrm{m}^2$ is significantly smaller than the estimated value based on the first Born approximation. The first Born approximation, however, is not reliable at low energies, according to the discussion in section 2.3, whereas the *Matrix Multislice* algorithm goes far beyond the first Born approximation.

Nevertheless, there is a problem associated with the treatment of high scattering angles: For pixels in reciprocal space corresponding to large transversal momenta, the kinetic energy available for a movement in z-direction is small so that the reflection coefficients become sizable. Therefore, these pixels are crucial for a correct calculation of the backscattered signal. These are, however, just the pixels that are damped by the Fermi filter function (3.49). Furthermore, the sampling of the different scattering angles is non-uniform, as shown in figure 6.4. Because of this, the aforementioned problem worsens. Thus, we can conclude that the calculated backscattering signal is probably too small.

The sampling problem can be mitigated by enlarging the simulation area, since, by this means, the pixel size in reciprocal space ($\Delta\kappa_x$ and $\Delta\kappa_y$) is reduced. A more ambitious solution for this problem would employ a special coordinate system for the reciprocal plane so that the sampling of the angles would become more uniform. This would, however, render the direct use of a Fourier transform impossible.

According to the Rutherford cross section (cf. (2.50))

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{Z\,e^2}{16\,\pi\,\varepsilon_0\,T}\right)^2 \frac{1}{\sin^4 \vartheta/2}, \tag{6.33}$$

the fraction of backscattered electrons ($\vartheta > \pi/2$) should scale like the inverse square of the the kinetic energy T. The Rutherford cross section is special, as it can be factored into an angular-dependent and an energy-dependent term. The scaling with T^{-2} is a universal feature of

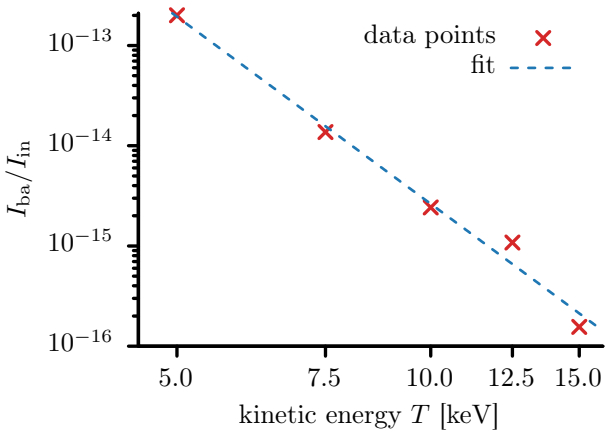

Figure 6.5: Double-logarithmic plot showing the dependence of the backscattered current I_{ba} on the kinetic energy of the incoming electron T. The simulation box contains four silicon atoms.

$1/r$-interactions and, hence, of electrodynamics, but it is, in general, modified by an additional angular dependence. Since the form factor parametrization by Lobato and Van Dyck incorporates the correct asymptotic scaling for the atomic potentials, the corresponding cross section for backward scattering calculated with the help of the first Born approximation shows the same T^{-2} dependence.

The numerical results calculated with the *Matrix Multislice* algorithm for a simulation box containing four silicon atoms are shown in figure 6.5. A logarithmic fit reveals that the data can indeed be described by a power function, but with an exponent of about -6.2. Thus, we must conclude that at least one of the two calculations, i.e. the first Born approximation or the *Matrix Multislice* algorithm, does not treat the backscattering correctly.

6.3.2 Scanning Mode

Since it is not possible to image the backscattered wave, we will now use the information contained in the transfer matrices to calculate SEM images. Due to the fact that STEM images can be calculated using the same methods, we will discuss them here as well.

Also the conventional multislice algorithms, as described in chapters 3 and 5, can be used for the simulation of STEM images. As for the simulation of CBED patterns, the parallel illumination has to be replaced by a convergent beam. For every beam position, i.e. for every pixel, the propagation step of the multislice algorithm must be repeated. Although the potential calculations can be reused, the required computational effort increases many times over.

Using the *Matrix Multislice* approach, however, all probe positions can be calculated at the same time with little additional effort: Once the transfer matrices are determined, the vectors $\mathbf{f}^0$, $\mathbf{f}^N$ and $\mathbf{b}^0$ in (6.31) and (6.32) are replaced by matrices. Each column of these matrices represents a single position of the electron probe. The same iterative procedure as described in section 6.2 is then used to solve these equations.

The idealized detector geometry shown in figure 6.6 was used for the simulations. The STEM signal is registered by an ADF-detector that accepts all electrons with scattering angles between δ and $\pi/2$. Likewise, the detector for the SEM-signal registers all electrons scattered in backward direction. Both detectors are integrating, this means all directional information is lost.

Large parts of the formalism introduced in section 2.6 for the description of the optical

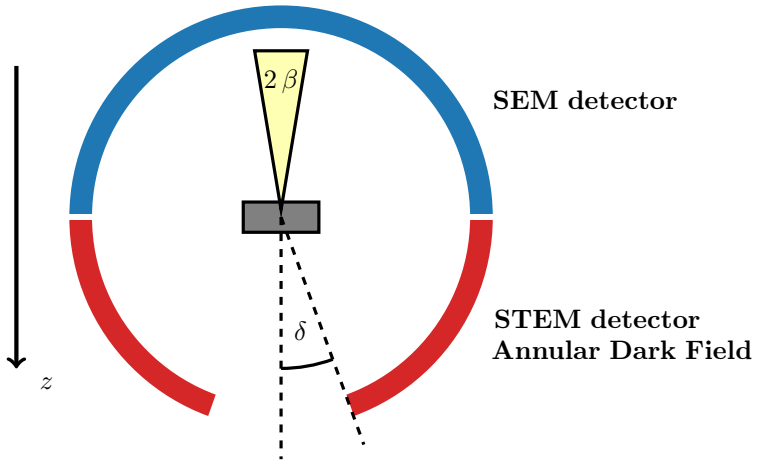

Figure 6.6: Geometry of the idealized detectors used to simulate scanning microscopes.

system can also be utilized to describe the formation of the STEM probe. Traditionally, the achievable resolution was limited by the spherical aberration of the probe forming lens. In an aberration-corrected microscope, however, the size of the probe is essentially determined by diffraction effects. The illumination aperture controls the convergence angle β of the incoming beam. If the aperture is illuminated by a parallel beam, the intensity of the electron probe will resemble an Airy pattern and the maximum resolution $r_{\max}$ can be estimated with the help of the Rayleigh criterion [1]

$$r_{\max} \approx 0.61 \, \frac{\lambda}{\beta}. \tag{6.34}$$

As already explained in section 5.3.2, a convergent probe is in conflict with the periodic boundary conditions assumed. This problem can be solved by increasing L_x and L_y. Because of the limited computational resources (main memory), the simulation area could only be quadrupled to $10.86\,\text{Å} \times 7.68\,\text{Å}$. Thus, the thickness of the silicon sample had to be reduced to 3.84 nm. For a convergence angle of 100 mrad, this thickness prevents interactions of the electron probe with itself due to wrap-around effects. The acceptance angle of the ADF detector δ is identical to the convergence angle. The simulation results are shown in figure 6.7.

By simulating the two-dimensional material graphene, the problems associated with the thickness of the sample can be avoided. The carbon atoms are placed at $z = 0$ while the simulation box extends from $z = -2\,\text{Å}$ to $2\,\text{Å}$.

The graphene lattice has a three-fold rotational symmetry about every atom. For parallel illumination, this symmetry is, as expected, present in the simulation results (cf. first row of figure 6.8). A close inspection of the STEM and SEM images, however, reveals that the symmetry is absent (cf. second row of figure 6.8). This is caused by the poor representation of the circular aperture in reciprocal space, as illustrated in figure 6.9: Because of the large pixel size in reciprocal space (corresponding to a small simulation area), the aperture appears distorted and has ragged edges. As a consequence, the circular symmetry of the Airy pattern in real space is lost. This explains the failure of the STEM and SEM images to completely reproduce the three-fold rotational symmetry of the graphene crystal.

The simulation of the SEM images is of course far from complete, as two important contribu-

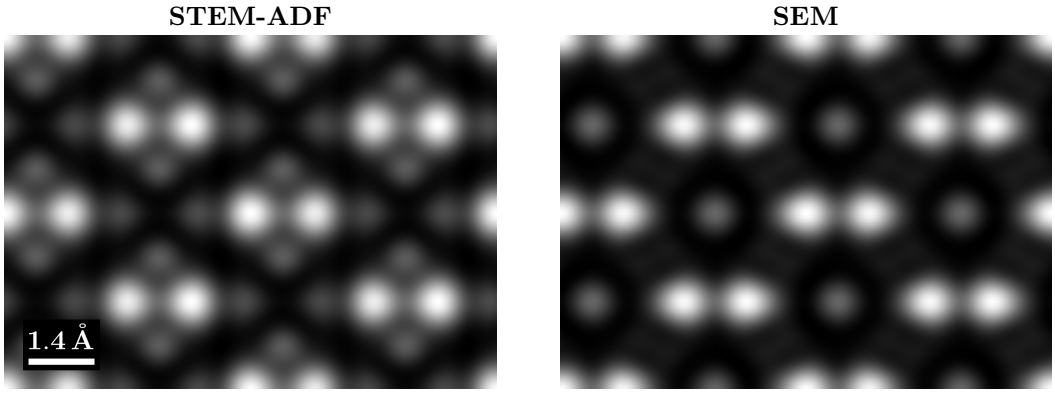

Figure 6.7: STEM and SEM simulation results for 3.84 nm silicon [110] at an acceleration voltage of 5 kV. The convergence angle β and the STEM detector angle δ were set to 100 mrad. The intensity of the SEM signal has been amplified by a factor of about $3.3 \cdot 10^6$.

tions are missing: No inelastic scattering processes have been considered, although it is known, that a large fraction of the detected electrons has undergone multiple inelastic scattering events [61]. Besides, the secondary electrons, generated as side products of inelastic scattering processes, are not considered. Modern instruments, however, can distinguish secondary from backscattered electrons. In the near future, it will become even possible to record SEM images using only the elastically backscattered electrons. This will allow a direct experimental comparison with the simulation results.

6.4 Discussion

The *Matrix Multislice* algorithm has been derived from the Schrödinger equation by using the correct boundary conditions for the situation in an electron microscope. In contrast to the other simulation algorithms discussed in this work, this approach uses only a minimal set of approximations and, thus, requires a high computational effort. The algorithm is analytically well defined. Nevertheless, an additional rule is needed for the numerical treatment of the exponentially growing matrix elements. Here, we excluded the corresponding pixels by using a low-pass filter.

The iterative approach used to solve the final equations works, as long as backscattering can be treated as a perturbation.

Due to the low-pass filter, most matrix elements vanish. This fact can be exploited by future implementations to decrease the amount of computational resources required by the matrix algorithm: Roughly $\pi/9$ of the pixels contain meaningful information. Thus, the storage requirements for the matrices can be reduced to $\pi^2/81 \approx 12\,\%$, and the number of operations to $\pi^3/729 \approx 4\,\%$.

Since the *Matrix Multislice* algorithm involves only linear algebra operations, it stands to reason that the simulation can be considerably accelerated by offloading the computations to a GPU.

The derivation in chapter 4 implicitly assumed that the electron wave has perfect temporal coherence. Because of this, the electron wave scattered at the bottom of the sample can interfere with the electron wave at the top. This may not be correct for large sample thicknesses. A typical field emission gun (FEG) has an energy width of $\Delta E = 0.4\,\mathrm{eV}$. Hence, the coherence

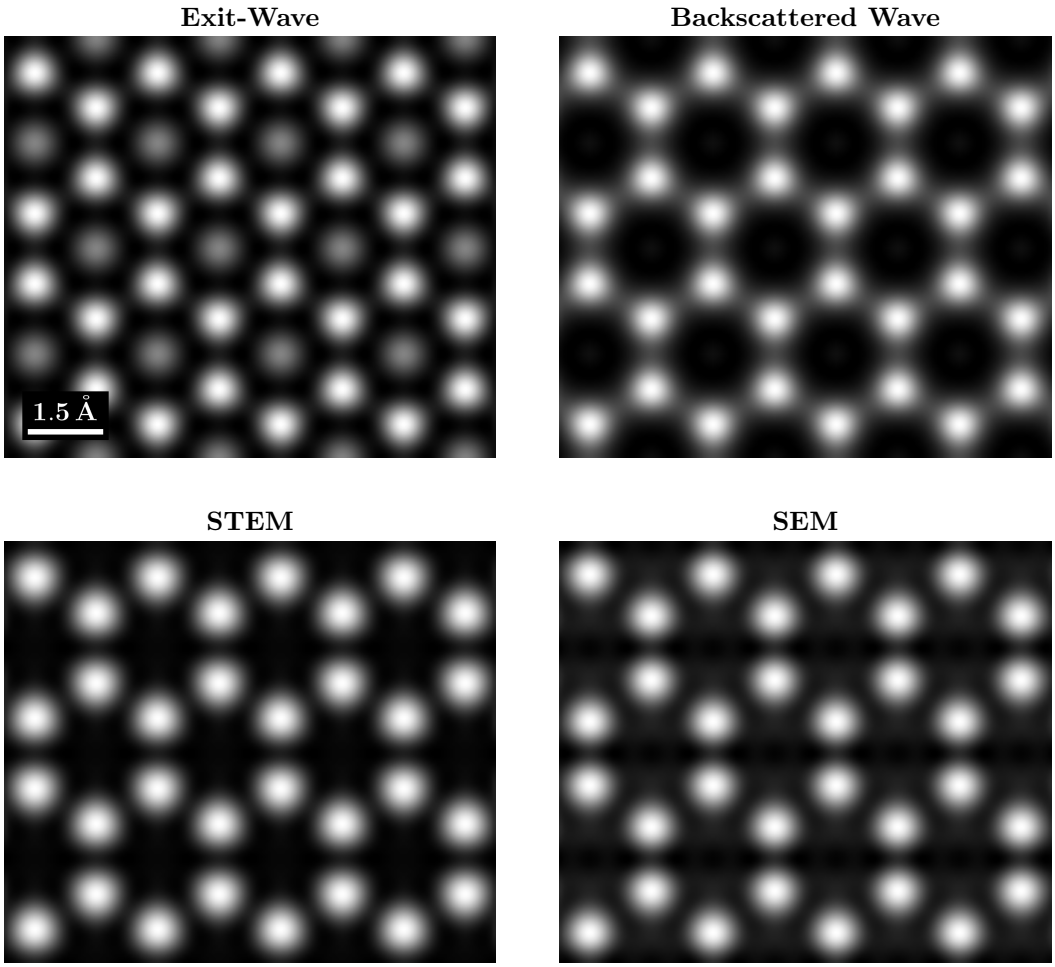

Figure 6.8: Simulation results for graphene at an acceleration voltage of 5 kV. The top row shows the absolute square of the exit-wave and the backscattered wave for parallel illumination. The wave functions have been filtered to simulate an aperture of 150 mrad. The contrast of the backscattered signal has been enhanced by a factor of about $3.7 \cdot 10^{12}$. The STEM image shows the signal of an ADF detector whose acceptance angle matches the convergence angle of the probe ($\delta = \beta = 120$ mrad). The contrast of the SEM image was increased by a factor of about 10^7.

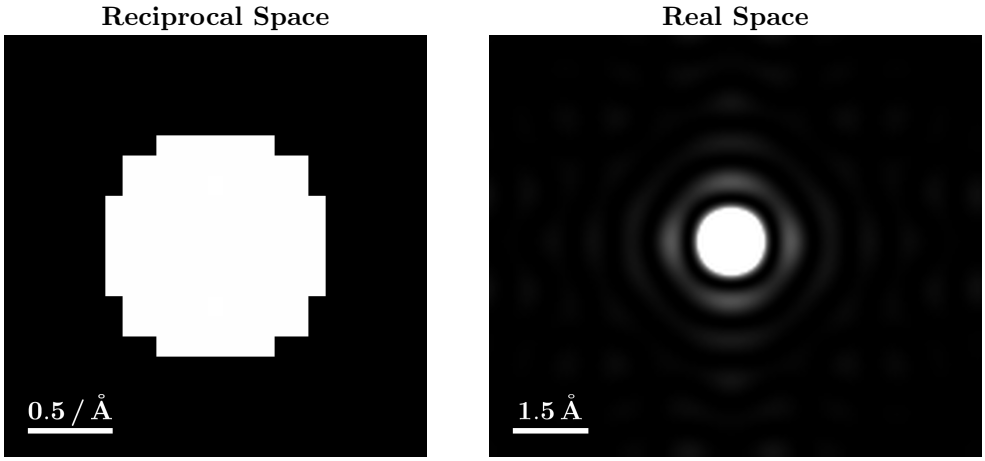

Figure 6.9: Electron probe for the graphene simulation in reciprocal and real space. The convergence angle is 120 mrad. The reciprocal space image shows only the innermost 25×21 pixels. The contrast of the real space image has been adjusted to enhance the visibility of the first "ring" of the Airy pattern.

time is in the order of femtoseconds. An electron accelerated by 10 kV moves at approximately 20 % of the speed of light. Therefore, its coherence length is

$$\ell = \frac{0.2\, c\, \hbar}{\Delta E} \approx 100\,\mathrm{nm}. \tag{6.35}$$

As the electron moves forward and backward through the sample, this limits the maximum sample thickness that can be simulated with this method to about 50 nm.

The *Matrix Multislice* algorithm allows the inclusion of backscattering effects in TEM image simulations. Even for unrealistically low electron energies, we found that the resulting corrections are negligible, so that the tremendous computational effort cannot be justified. This insight suggests, that future research should concentrate on improving the description of inelastic scattering processes.

On the other hand, the presented *Matrix Multislice* approach is the first description of the image formation process in an SEM which is not based on a semiclassical model. By this means, interference and resonance effects can be considered.

Chapter 7

Conclusions and Outlook

In this work, a unified description for different elastic imaging modalities in electron microscopes was developed (cf. chapter 4). The same formalism can be applied to the image formation in a TEM, a STEM or an SEM. In a certain sense, this new description can be regarded as a generalization of the classical multislice approach since it allows us to keep track of the electron scattered in backward direction as well.

Large Scattering Angles

Traditionally, most image simulations are based on the paraxial approximation (cf. section 3.1.1). Here, we could show, that a direct quadrature of the non-approximated Schrödinger equation is also possible (*RK2* algorithm, cf. section 5.1.2). Using the matrix formalism of chapter 4, we gave a new, rigorous derivation of the improved multislice algorithm, first suggested by Chen and Van Dyck (*Full* algorithm, cf. section 5.2.2).

The availability of improved solutions for the Schrödinger equation allowed us to study the effects of the paraxial approximation systematically (cf. chapter 5.3). We found that this approximation is only valid for thin samples and high energies ($\gtrsim 80\,\mathrm{keV}$). State-of-the-art TEMs, like the high-resolution low-voltage SALVE microscope, should be able to confirm these findings in the near future. Consequently, the simulation software currently in use must be adapted for low energies. The simulated CBED patterns revealed minor differences between the two approaches investigated, i.e. the *RK2* and the *Full* algorithm (cf. section 5.3.2). As both are based on sound approximations, it is not possible to predict for which approach the results will be in better agreement with the experimental findings.

The approximation used for the discretized Laplace operator is essential for a correct treatment of the large scattering angles (cf. section 3.3). The comparison of the different approximations showed that the reciprocal-space method is, in terms of quality, superior to real-space methods (cf. section 3.5.1). In view of the available computing power, the use of the real-space method should therefore be discouraged.

Backscattering and Transmission Electron Microscopy

The large impact of the paraxial approximation on the simulated TEM images for electron energies below 80 keV, triggered the investigation of backscattering effects. Since the description of the scattering process derived in chapter 4 is based on matrices, it can be used for numerical calculations as well (cf. section 6.2). This is facilitated by the availability of efficient libraries for linear algebra operations.

	Algorithm	Analytical	Numerical	Section
forward only	*RK2*	divergent	divergent	5.1
	Matrix Forward	stable	stable	5.2.1
	Full	divergent	divergent	5.2.2
forward & backward	*Matrix Multislice*	stable	divergent	6.1.2

Table 7.1: Stability of the different simulation algorithms if the inequality (4.24) is violated. Listed are only algorithms that are *not* based on the paraxial approximation. The effect of the low-pass filter has not been considered.

We demonstrated that it is feasible to include backscattering effects in TEM simulations (*Matrix Multislice* algorithm, cf. section 6.3.1). The backscattered current, however, is significantly smaller than expected. Moreover, it does not exhibit the T^{-2} scaling behavior as implied by the Rutherford formula. These discrepancies are most probably a consequence of the insufficient representation of the scattering angles near 90 °, which is due to the small size of the simulated area.

The calculated corrections for the exit-wave caused by backscattering effects are found to be negligible, especially in comparison to the adverse effects caused by the paraxial approximation. In retrospect, this was already indicated by the similarity of the *Full* and *RK2* results: Although both algorithms exclude backscattering effects by different means, the calculated exit-waves mostly agree.

Scanning Electron Microscopy

We have shown that the *Matrix Multislice* approach can be used for the simulation of SEM images (cf. section 6.3.2). All previous simulations for the SEM were based on semiclassical models and Monte-Carlo methods. The algorithm introduced here is the first treatment of the image formation process in an SEM that considers the wave nature of the electron. Thereby, interference and resonance effects are included.

To be competitive with the commonly used Monte-Carlo simulations, the *Matrix Multislice* algorithm must, however, be extended by a model for inelastically scattered and secondary electrons.

Quite recently, it was demonstrated that the energy resolution of backscattered electron detectors in SEMs can be significantly improved by means of aberration-correction [62]. Using these new microscopes, it will become possible to analyze only elastically scattered electrons. This will be an appealing opportunity to validate the image simulation algorithm. Furthermore, these experiments might indicate how inelastic scattering events need to be treated.

The Eigenvalue Spectrum

The spectrum of the matrix $\underline{A}^{2,n}$ (cf. section 4.3) has consequences for all simulation algorithms that go beyond the paraxial approximation. The effects are summarized in table 7.1. Only the *Matrix Multislice* algorithm, which includes backscattering effects, gives a consistent description, free of divergences. Nevertheless, without a rule on how to treat the exponentially large matrix elements, the numerical implementation fails as soon as a negative eigenvalue appears (violation of the inequality (4.24)). In this work, we employed a low-pass filter to suppress the corresponding matrix elements (cf. sections 5.3 and 6.2).

The backscattered signal can be excluded by choosing appropriate initial conditions. This has been used for the direct quadrature of the Schrödinger equation (*RK2* algorithm). In this case, however, even the analytical solution of the Schrödinger equation becomes divergent if the spectrum comprises negative eigenvalues (cf. section 5.1.1).

A modification of the scattering equations can remove the divergences (*Matrix Forward* algorithm, cf. section 5.2.1). Even then, great care must be taken not to reintroduce them (*Full* algorithm, cf. section 5.2.2).

Time-Dependent Simulations on a Three-Dimensional Grid

Most problems of the *Matrix Multislice* algorithm and particularly the sampling problem are caused by the unequal treatment of the three coordinate axes, since right from the beginning the z-coordinate is separated. A numerical solution of the Schrödinger equation on a three-dimensional grid could solve these problems altogether. In this case, the electron would be modeled as a wave packet, whose time evolution would need to be tracked for a finite number of steps. A simple estimate shows that the available computing power is insufficient for the simulation of a bulk specimen, but the simulation of a single sheet of graphene for very low energies ($\leq 3\,\mathrm{keV}$) might be possible even nowadays.

Next Steps for the Inclusion of Inelastic Scattering

Considering the small effect of the backscattered electrons and the tremendous computational effort that is necessary to include them in TEM simulations, it seems to be more worthwhile for future work to include inelastic scattering effects. At an acceleration voltage of 80 kV, the ratio of the total inelastic and the total elastic cross section is given by [63, 64]

$$\frac{\sigma_{\mathrm{inel}}}{\sigma_{\mathrm{el}}} = \frac{20}{Z}. \tag{7.1}$$

For low-Z materials inelastic scattering is hence at least as important as elastic scattering.

For a realistic description of inelastic scattering, however, not only the wave function, but also its coherence properties must be included in the simulation approach. A method to accomplish this for the TEM was proposed by Dinges et al. [65]. It is based on the mutual coherence function, known from optics [1]. The implementation of the *Matrix Multislice* algorithm presented here proves that available computing resources are sufficient for a direct realization of the mutual coherence function approach, which shall be analyzed in the near future.

An Experimental Verification of the New Simulation Approach

An interesting experimental setup to check the validity of our numerical backscattering algorithm and, especially, the treatment of the large scattering angles is reflection high-energy electron diffraction (RHEED): An electron beam strikes the surface of the sample at a very small angle. Afterward, the reflected electron wave creates an interference pattern that can be detected by a camera. This method can be used for detailed surface studies of a sample. The experimental situation can be reproduced in a simulation by choosing a z-axis perpendicular to the surface. In this case, the incoming as well as the reflected beams correspond to very large scattering angles.

Appendix A

Form Factor Parametrizations

A.1 Gaussian Functions

The form factor parametrization by Doyle and Turner [66], as well as the parametrization by Peng et al. [39] are based on a sum of Gaussian functions. Hence, the electronic form factor (2.53) is given by

$$f_\alpha^e(K) = \sum_{i=1}^{n} a_i \, \exp\left(-\frac{b_i}{16\,\pi^2}\,K^2\right), \tag{A.1}$$

where a_i and b_i are parameters that must be determined by the fit procedure. The number of Gaussian functions n is four for Doyle and Turner and five for Peng et al. Although these functions do not provide the correct physical behavior for very small and very large K, they have advantageous analytical properties and can be evaluated quite quickly. The effect of the Debye-Waller factor (2.88) can be incorporated easily by substituting $b_i \to b_i + 16\,\pi^2\,B_\alpha$.

According to (2.70), the mean square radius of the electron charge distribution is essentially the sum of the coefficients a_i

$$\left\langle r^2 \right\rangle = \frac{3\,a_B}{Z} \sum_{i=1}^{n} a_i. \tag{A.2}$$

The atomic potential (2.54) is given by a sum of Gaussian functions as well

$$V_\alpha(r) = -\frac{16\,\pi^{5/2}\,\hbar^2}{m} \sum_{i=1}^{n} \frac{a_i}{b_i^{3/2}} \exp\left(-\frac{4\,\pi^2}{b_i}\,r^2\right). \tag{A.3}$$

The antiderivative (3.26) can be written as

$$\Phi_\alpha(\rho, z) = \frac{4\,\pi^2\,\hbar^2}{m} \sum_{i=1}^{n} \frac{a_i}{b_i} \, \exp\left(-\frac{4\,\pi^2}{b_i}\,\rho^2\right) \operatorname{erf}\left(\frac{2\,\pi}{\sqrt{b_i}}\,z\right), \tag{A.4}$$

where $\operatorname{erf}(x)$ denotes the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2}\,\mathrm{d}t. \tag{A.5}$$

The hybrid potential (3.33) is given by

$$\overline{V}_\alpha(\kappa, z) = 4\sqrt{\pi^3}\sum_{i=1}^{n}\frac{a_i}{\sqrt{b_i}}\exp\left(-b_i\frac{\kappa^2}{16\,\pi^2}\right)\exp\left(-\frac{4\,\pi^2\,z^2}{b_i}\right). \tag{A.6}$$

By integrating the last line, we obtain the hybrid antiderivative (3.36)

$$\overline{\Phi}_\alpha(\kappa, z) = \sum_{i=1}^{n}\frac{a_i}{4\,\pi}\exp\left(-b_i\frac{\kappa^2}{16\,\pi^2}\right)\operatorname{erf}\left(\frac{2\,\pi\,z}{\sqrt{b_i}}\right). \tag{A.7}$$

A.2 Parametrization Based on Physical Constraints

The form factor parametrization by Lobato and Van Dyck [38] is based on physical constraints. For the electronic form factor (2.53) they propose the following expression

$$f^e_\alpha(K) = \sum_{i=1}^{5} a_i\,\frac{2 + b_i\,\frac{K^2}{4\,\pi^2}}{\left(1 + b_i\,\frac{K^2}{4\,\pi^2}\right)^2}. \tag{A.8}$$

As above, a_i and b_i are the form factor parameters.

The corresponding potential (2.54) provides the $1/r$-singularity

$$V_\alpha(r) = -\frac{2\,\pi^2\,\hbar^2}{m}\sum_{i=1}^{5}\frac{a_i}{b_i^{3/2}}\left(\pi + \frac{\sqrt{b_i}}{r}\right)\exp\left(-\frac{2\,\pi\,r}{\sqrt{b_i}}\right). \tag{A.9}$$

While an analytical expression for the antiderivative $\Phi_\alpha(\rho, z)$ is unknown, the projected potential (3.27) can be written as

$$\int_{-\infty}^{\infty} V_\alpha(\rho, z)\,\mathrm{d}z = -\frac{4\,\pi^2\,\hbar^2}{m}\sum_{i=1}^{5}\frac{a_i}{b_i}\left[K_0\left(\frac{2\,\pi\,\rho}{\sqrt{b_i}}\right) + \frac{\pi\,\rho}{\sqrt{b_i}}\,K_1\left(\frac{2\,\pi\,\rho}{\sqrt{b_i}}\right)\right], \tag{A.10}$$

where K_0 and K_1 denote the modified Bessel functions of the second kind with order zero and one, respectively.

The mean square radius of the charge distribution is given by

$$\left\langle r^2\right\rangle = \frac{6\,a_B}{Z}\sum_{i=1}^{n} a_i. \tag{A.11}$$

The electronic form factor (A.8) can be supplemented with a Debye-Waller factor (2.88). In this case, we obtain for the corresponding hybrid potential (3.33)

$$\begin{aligned}\overline{V}_\alpha(\kappa, z) = 4\,\pi^3\,\mathrm{e}^{-B_\alpha\,\kappa^2}\sum_{i=1}^{n}\frac{a_i}{b_i^2\,A_i^2}\Bigg\{&4\,\sqrt{B_\alpha\,\pi^3}\,\mathrm{e}^{-\frac{z^2}{4\,B_\alpha}} + \frac{\mathrm{e}^{A_i^2\,B_\alpha}}{2\,A_i}\times\\ &\times\Bigg[\mathrm{e}^{A_i\,z}\left(b_i\,\kappa^2 + 6\,\pi^2 - 4\,\pi^2\,A_i^2\,B_\alpha - 2\,\pi^2\,A_i\,z\right)\operatorname{erfc}\left(A_i\sqrt{B_\alpha} + \frac{z}{2\,\sqrt{B_\alpha}}\right)\\ &+ \mathrm{e}^{-A_i\,z}\left(b_i\,\kappa^2 + 6\,\pi^2 - 4\,\pi^2\,A_i^2\,B_\alpha + 2\,\pi^2\,A_i\,z\right)\operatorname{erfc}\left(A_i\sqrt{B_\alpha} - \frac{z}{2\sqrt{B_\alpha}}\right)\Bigg]\Bigg\}\end{aligned} \tag{A.12}$$

$$\text{with } A_i^2 = \kappa^2 + \frac{4\,\pi^2}{b_i}. \tag{A.13}$$

Here, $\operatorname{erfc}(x)$ is the complementary error function

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \mathrm{e}^{-t^2} \mathrm{d}t. \tag{A.14}$$

The formula (A.12) simplifies considerably for the limiting case $B_\alpha \to 0$

$$\overline{V}_\alpha(\boldsymbol{\kappa}, z) = 4\,\pi^3 \sum_{i=1}^{n} \frac{a_i\,\mathrm{e}^{-A_i|z|}}{b_i^2\,A_i^3} \left[b_i\,\kappa^2 + 6\,\pi^2 + 2\,\pi^2\,A_i\,|z|\right]. \tag{A.15}$$

By integrating the hybrid potential (A.12), we find the hybrid antiderivative (3.36)

$$\begin{aligned}\overline{\Phi}_\alpha(\kappa, z) = \pi\,\mathrm{e}^{-B_\alpha\,\kappa^2} \sum_{i=1}^{n} \frac{a_i}{b_i^2\,A_i^4} \Bigg\{ \left(b_i\,\kappa^2 + 8\,\pi^2\right) \operatorname{erf}\left(\frac{z}{2\,\sqrt{B_\alpha}}\right) + \frac{\mathrm{e}^{A_i^2\,B_\alpha}}{2} \times \\ \times \Bigg[\mathrm{e}^{A_i\,z} \left(b_i\,\kappa^2 + 8\,\pi^2 - 4\,\pi^2\,A_i^2\,B_\alpha - 2\,\pi^2\,A_i\,z\right) \operatorname{erfc}\left(A_i\sqrt{B_\alpha} + \frac{z}{2\sqrt{B_\alpha}}\right) \\ - \mathrm{e}^{-A_i\,z} \left(b_i\,\kappa^2 + 8\,\pi^2 - 4\,\pi^2\,A_i^2\,B_\alpha + 2\,\pi^2\,A_i\,z\right) \operatorname{erfc}\left(A_i\sqrt{B_\alpha} - \frac{z}{2\sqrt{B_\alpha}}\right)\Bigg]\Bigg\}. \end{aligned} \tag{A.16}$$

Once again, the expression becomes much simpler for the limiting case $B_\alpha \to 0$

$$\overline{\Phi}_\alpha(\kappa, z) = \sum_{i=1}^{n} \frac{\pi\,a_i}{b_i^2\,A_i^4} \frac{z}{|z|} \left\{b_i\,\kappa^2 + 8\,\pi^2 - \mathrm{e}^{-A_i\,|z|} \left(b_i\,\kappa^2 + 8\,\pi^2 + 2\,\pi^2\,A_i\,|z|\right)\right\}. \tag{A.17}$$

Appendix B

Implementation Details

A software suite, named Aurora (the Roman goddess of dawn, cf. Eos[a]), was developed to assess and compare the different image simulation algorithms presented in this work. All numerical results shown were calculated with the help of this software. Some aspects of Aurora's design and structure are discussed in this appendix.

Aurora is written in C++14 and can be compiled for Microsoft Windows and GNU/Linux. To avoid code duplication, most of the algorithms reside inside a shared library. Unit tests using Google Test[b] ensure the correctness of the basic building blocks. The algorithms are accessible from a command line interface and partially from a graphical user interface (GUI) based on QT[c].

Internally, the calculations can be performed either with 32-bit (single precision) or 64-bit (double precision) floating-point numbers. While for the forward-scattering algorithms single precision is mostly sufficient, for backscattering additional precision is required.

The complex-valued wave functions are represented as two-dimensional arrays of complex numbers. To facilitate the evaluation of the two-dimensional Laplace operator in real space, the corresponding buffers can be padded with pixels that replicate the values from the opposite side of the buffer. To increase the speed and the readability of the algorithms, expression templates are used. As most algebraic operations act on every pixel independently, these operations can be executed in parallel. The wave buffers can be serialized for data exchange. To maximize the compatibility, they can be written into a TIFF file[d] with two layers (32-bit floating-point numbers). One layer is used for the real, the other layer for the imaginary part of the wave function. Three custom tags are used to save the physical dimensions of the buffer (65 000 and 65 001) and the acceleration voltage (65 002). The command-line arguments are stored in the description tag.

For the calculation of the FFTs, two different backends can be used: FFTW[e] [67] is one of the fastest implementations of the FFT algorithm, but it is a large library that is not available on every system. KissFFT[f] is a much smaller library that could be directly integrated into the source code.

Besides the simple Wentzel model, five different form factor parametrizations have been implemented:

[a] `http://project.het.physik.tu-dortmund.de/eos/`
[b] `https://github.com/google/googletest/`
[c] `http://www.qt.io/`
[d] TIFF specification: `https://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf`
[e] `http://www.fftw.org/`
[f] `http://kissfft.sourceforge.net/`

- Doyle and Turner [66],
- Peng et al. [39],
- Weickenmeier and Kohl [68],
- Kirkland [29],
- Lobato and Van Dyck [38].

Phase shifts and potentials can be calculated as needed (modes named *Jit*) or beforehand. While the former method requires much less memory, the latter is faster for periodic potentials (crystals). If the phase shifts and potentials are calculated in real space (cf. section 3.2), the calculations can be accelerated by evaluating them on a supersampled two-dimensional grid and using bilinear interpolation to obtain the final results (modes named *Opt*).

The calculation of the potentials and phase shifts is embarrassingly parallel. Hence, OpenMP is used to exploit this fact.

Crystal structures can be algorithmically generated. Additionally, sample descriptions in the pdb (protein database) format[g] and a custom format can be processed. The latter circumvents the many limitations of the pdb format.

Forward Scattering The forward scattering algorithms described in this work are accessible from the *emule* (elastic multislice) client. It drives the conventional TEM / CBED and STEM simulations.

All possible combinations of the Laplace operators, propagation methods and form factors are available. The thermal movements of the atoms can be considered by using either a Debye-Waller factor or the frozen-phonon approximation.

The implicit methods presented in section 5.1 are implemented using the sparse linear algebra module of the Eigen library[h] [69]. The equations are solved with the biconjugate gradient stabilized solver.

Backward Scattering The algorithm for backward scattering is implemented in the *matrix* client. In general, the Eigen library is used for the dense linear algebra operations. Matrix multiplications and the eigenvalue decompositions are, however, delegated to the Intel Math Kernel Library (MKL)[i] that provides BLAS[j] and LAPACK[k] functionality. As the matrices are quite large, only a minimal working set is kept in memory. All data, that is not required immediately, is transferred to disk storage. Although disk accesses are quite slow, this procedure does not affect the computation times very much, since the algorithms execute $\mathcal{O}(M^3)$ operations while the memory bandwidth requirements scale only as $\mathcal{O}(M^2)$.

[g] `http://www.wwpdb.org/documentation/file-format`
[h] `http://eigen.tuxfamily.org/`
[i] `https://software.intel.com/en-us/intel-mkl/`
[j] `http://www.netlib.org/blas/`
[k] `http://www.netlib.org/lapack/`

Bibliography

[1] M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. 7th. Cambridge University Press, 1999. ISBN: 978-0-521-64222-4.

[2] D. B. Williams and C. B. Carter. *Transmission Electron Microscopy: A Textbook for Material Science*. 2nd. Springer New York, 2009. ISBN: 978-0-387-76500-6.

[3] O. Scherzer. "Über einige Fehler von Elektronenlinsen." In: *Zeitschrift für Physik* 101.9–10 (1936), pp. 593–603. DOI: 10.1007/BF01349606.

[4] P. W. Hawkes. "Aberration correction past and present." In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 367.1903 (2009), pp. 3637–3664. DOI: 10.1098/rsta.2009.0004.

[5] D. Gabor. "A New Microscopic Principle." In: *Nature* 161.4098 (1948), pp. 777–778. DOI: 10.1038/161777a0.

[6] M. Haider, S. Uhlemann, E. Schwan, H. Rose, B. Kabius, and K. Urban. "Electron microscopy image enhanced." In: *Nature* 392 (1998), pp. 768–769. DOI: 10.1038/33823.

[7] B. Kabius, P. Hartel, M. Haider, H. Müller, S. Uhlemann, U. Loebau, J. Zach, and H. Rose. "First application of Cc-corrected imaging for high-resolution and energy-filtered TEM." In: *Journal of Electron Microscopy* 58.3 (2009), pp. 147–155. DOI: 10.1093/jmicro/dfp021.

[8] L. Reimer and H. Kohl. *Transmission Electron Microscopy: Physics of Image Formation*. 5th. Springer New York, 2008. ISBN: 978-0-387-40093-8.

[9] J. C. Meyer, F. Eder, S. Kurasch, V. Skakalova, J. Kotakoski, H. J. Park, S. Roth, A. Chuvilin, S. Eyhusen, G. Benner, A. V. Krasheninnikov, and U. Kaiser. "Accurate Measurement of Electron Beam Induced Displacement Cross Sections for Single-Layer Graphene." In: *Physical Review Letters* 108.19 (2012), p. 196102. DOI: 10.1103/PhysRevLett.108.196102.

[10] U. Kaiser, J. Biskupek, J. C. Meyer, J. Leschner, L. Lechner, H. Rose, M. Stöger-Pollach, A. N. Khlobystov, P. Hartel, H. Müller, M. Haider, S. Eyhusen, and G. Benner. "Transmission electron microscopy at 20 kV for imaging and spectroscopy." In: *Ultramicroscopy* 111.8 (2011), pp. 1239–1246. DOI: 10.1016/j.ultramic.2011.03.012.

[11] C. Davisson and L. H. Germer. "Diffraction of Electrons by a Crystal of Nickel." In: *Physical Review* 30.6 (1927), pp. 705–740. DOI: 10.1103/PhysRev.30.705.

[12] H. Bethe. "Theorie der Beugung von Elektronen an Kristallen." In: *Annalen der Physik* 392.17 (1928), pp. 55–129. DOI: 10.1002/andp.19283921704.

[13] E. Ruska. "The Development of the Electron Microscope and of Electron Microscopy." In: *Nobel Lectures, Physics 1981-1990*. Ed. by T. Frängsmyr. World Scientific Publishing Co. Singapore, 1993, pp. 355–380. ISBN: 981-02-0729-8.

[14] J. M. Cowley and A. F. Moodie. "Fourier Images: I – The Point Source." In: *Proceedings of the Physical Society, Section B* 70 (1957), pp. 486–496. DOI: 10.1088/0370-1301/70/5/305.

[15] K. Fujiwara. "Relativistic Dynamical Theory of Electron Diffraction." In: *Journal of the Physical Society of Japan* 16.11 (1961), pp. 2226–2238. DOI: 10.1143/JPSJ.16.2226.

[16] P. Goodman and A. F. Moodie. "Numerical Evaluations of N-beam Wave Functions in Electron Scattering by the Multi-slice Method." In: *Acta Crystallographica Section A* 30.2 (1974), pp. 280–290. DOI: 10.1107/S056773947400057X.

[17] K. Ishizuka and N. Uyeda. "A New Theoretical and Practical Approach to the Multislice Method." In: *Acta Crystallographica Section A* 33 (1977), pp. 740–749. DOI: 10.1107/S0567739477001879.

[18] L. A. Bursill and A. R. Wilson. "Electron-Optical Imaging of the Hollandite Structure at 3 Å Resolution." In: *Acta Crystallographica Section A* 33 (1977), pp. 672–676. DOI: 10.1107/S0567739477001661.

[19] J. Cowley and J. Spence. "Innovative Imaging and Microdiffraction in STEM." In: *Ultramicroscopy* 3 (1978), pp. 433–438. DOI: 10.1016/S0304-3991(78)80068-0.

[20] E. J. Kirkland, R. F. Loane, and J. Silcox. "Simulation of Annular Dark Field STEM Images Using a Modified Multislice Method." In: *Ultramicroscopy* 23.1 (1987), pp. 77–96. DOI: 10.1016/0304-3991(87)90229-4.

[21] D. F. Lynch and A. F. Moodie. "Numerical evaluation of low energy electron diffraction intensities." In: *Surface Science* 32.2 (1972), pp. 422–438. DOI: 10.1016/0039-6028(72)90170-7.

[22] J. H. Chen and D. Van Dyck. "Accurate multislice theory for elastic electron scattering in transmission electron microscopy." In: *Ultramicroscopy* 70.1–2 (1997), pp. 29–44. DOI: 10.1016/S0304-3991(97)00071-5.

[23] J. Spiegelberg and J. Rusz. "A multislice theory of electron scattering in crystals including backscattering and inelastic effects." In: *Ultramicroscopy* 159, Part 1 (2015), pp. 11–18. DOI: 10.1016/j.ultramic.2015.07.002.

[24] W. Van den Broek, X. Jiang, and C. T. Koch. "FDES, a GPU-based multislice algorithm with increased efficiency of the computation of the projected potential." In: *Ultramicroscopy* 158 (2015), pp. 89–97. DOI: 10.1016/j.ultramic.2015.07.005.

[25] I. Lobato and D. Van Dyck. "MULTEM: A new multislice program to perform accurate and fast electron diffraction and imaging simulations using Graphics Processing Units with CUDA." In: *Ultramicroscopy* 156 (2015), pp. 9–17. DOI: 10.1016/j.ultramic.2015.04.016.

[26] S. Van Aert, K. J. Batenburg, M. D. Rossell, R. Erni, and G. Van Tendeloo. "Three-dimensional atomic imaging of crystalline nanoparticles." In: *Nature* 470.7334 (2011), pp. 374–377. DOI: 10.1038/nature09741.

[27] W. Van den Broek and C. T. Koch. "Method for Retrieval of the Three-Dimensional Object Potential by Inversion of Dynamical Electron Scattering." In: *Physical Review Letters* 109 (2012), p. 245502. DOI: 10.1103/PhysRevLett.109.245502.

[28] S. H. Scheres. "A Bayesian View on Cryo-EM Structure Determination." In: *Journal of Molecular Biology* 415.2 (2012), pp. 406–418. DOI: 10.1016/j.jmb.2011.11.010.

[29] E. J. Kirkland. *Advanced Computing in Electron Microscopy*. 2nd. Springer New York, 2010. ISBN: 978-1-4419-6532-5.

[30] Z. L. Wang. *Elastic and Inelastic Scattering in Electron Diffraction and Imaging*. Plenum Press New York, 1995. ISBN: 0-306-44929-3.

[31] S. Weinberg. *The Quantum Theory of Fields, Volume 1. Foundations*. 4th. Cambridge University Press, 2010. ISBN: 978-0-521-67053-1.

[32] H. Müller. "A Coherence Funtion Approach to Image Simulation." PhD thesis. TU Darmstadt, 2000. URL: http://d-nb.info/96052942X/34.

[33] S. Weinberg. *Lectures on Quantum Mechanics*. 1st. Cambridge University Press, 2013. ISBN: 978-1-107-02872-2.

[34] G. Wentzel. "Zwei Bemerkungen über die Zerstreuung korpuskularer Strahlen als Beugungserscheinung." In: *Zeitschrift für Physik* 40.8 (1926), pp. 590–593. DOI: 10.1007/BF01390457.

[35] L. D. Landau and E. M. Lifschitz. *Lehrbuch der Theoretischen Physik – Band III. Quantenmechanik*. 9th. Harri Deutsch Frankfurt am Main, 2007. ISBN: 978-3-8171-1328-6.

[36] L. D. Landau and E. M. Lifschitz. *Lehrbuch der Theoretisehen Physik – Band VIII. Elektrodynamik der Kontinua*. 5th. Akademie-Verlag Berlin, 1990. ISBN: 3-05500-072-2.

[37] A. Jablonski, F. Salvat, and C. J. Powel. *NIST Electron Elastic-Scattering Cross-Section, Database, Version 3.2, SRD 64*. National Institute of Standards and Technology Gaithersburg, MD, 2010.

[38] I. Lobato and D. Van Dyck. "An accurate parameterization for scattering factors, electron densities and electrostatic potentials for neutral atoms that obey all physical constraints." In: *Acta Crystallographica Section A* 70.6 (2014), pp. 636–649. DOI: 10.1107/S205327331401643X.

[39] L.-M. Peng, G. Ren, S. L. Dudarev, and M. J. Whelan. "Robust Parameterization of Elastic and Absorptive Electron Atomic Scattering Factors." In: *Acta Crystallographica Section A* 52 (1996), pp. 257–276. DOI: 10.1107/S0108767395014371.

[40] M. J. Berger, J. H. Hubbell, S. M. Seltzer, J. Chang, J. S. Coursey, R. Sukumar, D. S. Zucker, and K. Olsen. *XCOM: Photon Cross Section Database (version 1.5)*. 2010. URL: http://physics.nist.gov/xcom.

[41] Z. L. Wang. "The 'Frozen-Lattice' Approach for Incoherent Phonon Excitation in Electron Scattering. How Accurate Is It?" In: *Acta Crystallographica Section A* 54.4 (1998), pp. 460–467. DOI: 10.1107/S0108767398001457.

[42] V. F. Sears and S. A. Shelley. "Debye-Waller Factor for Elemental Crystals." In: *Acta Crystallographica Section A* 47.4 (1991), pp. 441–446. DOI: 10.1107/S0108767391002970.

[43] R. Erni. *Aberration-Corrected Imaging in Transmission Electron Microscopy*. Imperial College Press London, 2010. ISBN: 978-1-84816-536-6.

[44] J. M. Cowley and A. F. Moodie. "The scattering of electrons by atoms and crystals. I. A new theoretical approach." In: *Acta Crystallographica* 10 (1957), pp. 609–619. DOI: 10.1107/S0365110X57002194.

[45] C. Y. Cai and J. H. Chen. "An accurate multislice method for low-energy transmission electron microscopy." In: *Micron* 43.2–3 (2012), pp. 374–379. DOI: 10.1016/j.micron.2011.09.018.

[46] C. Y. Cai, S. J. Zeng, H. R. Liu, and Q. B. Yang. "Computational comparison of the conventional multislice method and the real space multislice method for simulating exit wavefunctions." In: *Micron* 40.3 (2009), pp. 313–319. DOI: 10.1016/j.micron.2008.11.003.

[47] J. Ciston, J. S. Kim, S. J. Haigh, A. I. Kirkland, and L. Marks. "Optimized conditions for imaging the effects of bonding charge density in electron microscopy." In: *Ultramicroscopy* 111.7 (2011), pp. 901–911. DOI: 10.1016/j.ultramic.2010.12.003.

[48] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Groß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser. "Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy." In: *Nature Materials* 10.3 (2011), pp. 209–215. DOI: 10.1038/nmat2941.

[49] A. Chuvilin and U. Kaiser. "On the peculiarities of CBED pattern formation revealed by multislice simulation." In: *Ultramicroscopy* 104.1 (2005), pp. 73–82. DOI: 10.1016/j.ultramic.2005.03.003.

[50] D. Van Dyck. "Image Calculations in High-Resolution Electron Microscopy: Problems, Progress, and Prospects." In: *Advances in Electronics and Electron Physics* 65 (1985), pp. 295–355. DOI: 10.1016/S0065-2539(08)60880-X.

[51] G. J. Jensen, ed. *Methods in Enzymology, Volume 482, Cryo-EM, Part B, 3-D Reconstruction*. Academic Press San Diego, 2010. ISBN: 978-0-12-384991-5.

[52] R. Courant, K. Friedrichs, and H. Lewy. "Über die partiellen Differenzengleichungen der mathematischen Physik." In: *Mathematische Annalen* 100.1 (1928), pp. 32–74. DOI: 10.1007/BF01448839.

[53] D. W. Zingg. "Comparison of High-Accuracy Finite-Difference Methods for Linear Wave Propagation." In: *SIAM Journal on Scientific Computing* 22.2 (2000), pp. 476–502. DOI: 10.1137/S1064827599350320.

[54] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes, The Art of Scientific Computing*. 3rd. Cambridge University Press, 2007. ISBN: 978-0-521-88407-5.

[55] J. Stoer and R. Bulirsch. *Numerische Mathematik 2*. Springer Berlin Heidelberg New York, 2005. ISBN: 978-3-540-23777-8. DOI: 10.1007/b137272.

[56] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. Wiley, 2003. ISBN: 978-0-471-96758-3.

[57] B. F. Buxton, J. A. Eades, J. W. Steeds, and G. M. Rackham. "The Symmetry of Electron Diffraction Zone Axis Patterns." In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 281.1301 (1976), pp. 171–194. DOI: 10.1098/rsta.1976.0024.

[58] W. Q. Ming and J. H. Chen. "Validities of three multislice algorithms for quantitative low-energy transmission electron microscopy." In: *Ultramicroscopy* 134 (2013), pp. 135–143. DOI: 10.1016/j.ultramic.2013.04.013.

[59] A. L. Lewis, R. B. Hammond, and R. E. Villagrana. "The Importance of Second-order Partial Derivatives in the Theory of High-Energy-Electron Diffraction from Imperfect Crystals." In: *Acta Crystallographica Section A* 31.2 (1975), pp. 221–227. DOI: 10.1107/S0567739475000447.

[60] D. Van Dyck. "Fast computational procedures for the simulation of structure images in complex or disordered crystals: a new approach." In: *Journal of Microscopy* 119.1 (1980), pp. 141–152. DOI: 10.1111/j.1365-2818.1980.tb04084.x.

[61] L. Reimer. *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis*. 2nd. Springer Berlin Heidelberg New York, 1998. ISBN: 3-540-63976-4.

[62] M. Steigerwald, C. Hendrich, D. Preikszas, and K. Schubert. "A Mirror-Corrected Scanning Electron Microscope." In: *Frontiers of Characterization and Metrology for Nanoelectronics: 2013*. Ed. by E. Secula and D. Seiler. 2013, pp. 51–55. URL: http://www.nist.gov/pml/div683/conference/upload/2013_color.pdf.

[63] R. F. Egerton. "Measurement of Inelastic/Elastic Scattering Ratio for Fast Electrons and Its Use in the Study of Radiation Damage." In: *physica status solidi (a)* 37.2 (1976), pp. 663–668. DOI: 10.1002/pssa.2210370237.

[64] L. Reimer and M. Ross-Messemer. "Contrast in the electron spectroscopic imaging mode of a TEM." In: *Journal of Microscopy* 159.2 (1990), pp. 143–160. DOI: 10.1111/j.1365-2818.1990.tb04772.x.

[65] C. Dinges, A. Berger, and H. Rose. "Simulation of TEM images considering phonon and electronic excitations." In: *Ultramicroscopy* 60.1 (1995), pp. 49–70. DOI: 10.1016/0304-3991(95)00041-X.

[66] P. A. Doyle and P. S. Turner. "Relativistic Hartree-Fock X-ray and Electron Scattering Factors." In: *Acta Crystallographica Section A* 24.3 (1968), pp. 390–397. DOI: 10.1107/S0567739468000756.

[67] M. Frigo and S. G. Johnson. "The Design and Implementation of FFTW3." In: *Proceedings of the IEEE* 93.2 (2005), pp. 216–231. DOI: 10.1109/JPROC.2004.840301.

[68] A. Weickenmeier and H. Kohl. "Computation of Absorptive Form Factors for High-Energy Electron Diffraction." In: *Acta Crystallographica Section A* 47.5 (1991), pp. 590–597. DOI: 10.1107/S0108767391004804.

[69] G. Guennebaud, B. Jacob, et al. *Eigen v3*. 2010. URL: http://eigen.tuxfamily.org.

Journal Publications

- C. Wacker and R. R. Schröder. "Multislice algorithms revisited: Solving the Schrödinger equation numerically for imaging with electrons". In: *Ultramicroscopy* 151 (2015), pp. 211–223, DOI: 10.1016/j.ultramic.2014.12.008
- S. Hettler, J. Wagner, M. Dries, M. Oster, C. Wacker, R. R. Schröder, and D. Gerthsen. "On the role of inelastic scattering in phase-plate transmission electron microscopy". In: *Ultramicroscopy* 155 (2015), pp. 27–41, DOI: 10.1016/j.ultramic.2015.04.001
- C. Wacker and R. R. Schröder. "Image Formation in Electron Microscopes: Backscattering as an Eigenvalue Problem" – in preparation

Conference Presentations

- C. Wacker. "TEM Image Simulations". Talk. 4th SALVE Summer Workshop (2014), Hirschegg, Austria
- C. Wacker and R. R. Schröder. "Numerical Treatment of the Full, Non-Approximated Schrödinger Equation at Low Energies". Talk. IMC2014, Prague, Czech Republic
- C. Wacker. "Image Formation in Electron Microscopes – Backscattering as an Eigenvalue Problem". Talk. 5th SALVE Summer Workshop (2015), Hirschegg, Austria
- C. Wacker and R. R. Schröder. "Including backscattered electron in the simulation of TEM images". Poster. MC2015, Göttingen, Germany
- S. Hettler, M. Dries, T. Schulze, M. Oster C. Wacker, R. R. Schröder, and D. Gerthsen. "High-resolution transmission electron microscopy with Zach phase plate". Poster. MC2015, Göttingen, Germany
- M. Oster, C. Wacker, and R. R. Schröder. "Automatic Zernike phase plate alignment use to eliminate cut-on artifacts". Poster. MC2015, Göttingen, Germany

Acknowledgments

First of all, I would like to thank Prof. Dr. Rasmus Schröder for giving me the opportunity to work on this interesting topic and for supervising my research. His ideas for new applications constantly challenged the image simulations and have been a great motivation. I want to thank Prof. Dr. Ulrich Schwarz for his offer to prepare the second report for my thesis.

I gratefully acknowledge the Studienstiftung des Deutschen Volkes for a PhD scholarship. The financial support of the Deutsche Gesellschaft für Elektronenmikroskopie (DGE) made it possible to present my work at the International Microscopy Congress (IMC 2014).

I would like to thank Prof. Dr. Ute Kaiser for inviting me to the SALVE summer workshops. The members of the electron microscopy group of Ulm University always created a friendly atmosphere. The discussions at the workshops inspired many improvements for the simulation algorithms. I especially enjoyed the exchange of ideas with Dr. Ralf Hambach and Dr. Wouter Van den Broek.

I am very grateful to all present and former members of the CryoEM group for all the help, discussions and distractions. Certainly, this thesis would not have been possible without their support. The discussions with Marco Oster were always very fruitful and allowed the solution of numerous problems. I am indebted to Anne Kast for her continuous and extensive support. Her assistance significantly improved the manuscript. I would like to thank Dr. Ira Mang for all her advice and useful help. My thanks go to Dr. Götz Hofhaus for many interesting (and sometimes quite heated) discussions. I want to thank Jörg Eisele, Dr. Irene Wacker und Johan Zeelen for proofreading the manuscript and their valuable suggestions.

Finally, I would like to thank my parents for their constant and unlimited support through all these years.