

ANTONIUS DORDA

A novel numerical scheme for the Wigner transport equation and its application to resonant tunneling diodes

MASTER THESIS

For obtaining the academic degree
Diplom-Ingenieur

Master Programme of
Technical Physics



Graz University of Technology

Supervisor:

Ao. Univ.-Prof. Dipl.-Ing. Dr. Ferdinand Schürer
Institute of Theoretical and Computational Physics
Graz University of Technology

Graz, November 2012

Meinen Eltern

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

EIDESSTÄTTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am

.....
(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

.....
date

.....
(signature)

Abstract

In analogy to the reconstruct-evolve-average (REA) algorithm of finite volume methods, a novel discretization scheme for the pseudo-differential operator in the Wigner transport equation (WTE) is developed. For this purpose, the Wigner function is expressed in a local basis given by piecewise polynomials. Properly chosen grid cells are defined to arrive at a cell-averaged formulation which ensures the conservation of the particle density on arbitrary grids. As a result, the scheme allows for a non-equidistant and highly flexible choice of the grid. The advection term of the WTE is solved by making use of a well-approved weighted essentially non oscillatory (WENO) scheme. To evaluate the performance of the developed algorithm, a resonant tunneling diode is considered as test case and the obtained current voltage characteristics are compared with those of a non-equilibrium Green's function (NEGF) reference calculation. A remarkably good agreement of the results could be achieved. Furthermore, by making extensive use of the ability to apply non-equidistant grids, rapidly varying and large-valued oscillation patterns in the stationary Wigner function were discovered in the case of bias voltages where resonant tunneling processes occur. To examine the dynamics of these oscillations and the whole Wigner function, a simple time-dependent situation is considered in the last part of this work.

Key words: Quantum mechanics in phase space, Wigner transport equation, quantum transport, resonant tunneling diode, finite volume methods, REA algorithm, non-equilibrium Green's function.

Eine neue numerische Methode für die Wignergleichung und deren Anwendung auf resonante Tunnelnioden

Kurzfassung

In Anlehnung an die Konzepte des REA (reconstruct-evolve-average) Algorithmus von finiten Volumen Methoden wird in dieser Arbeit ein neuartiges Diskretisierungsverfahren für den pseudo-differentiellen Operator der Wignergleichung entwickelt. Im Konkreten wird die Wignerfunktion in eine lokal definierte Basis, gegeben durch stückweise Polynome entwickelt. Durch Einführung geeigneter Gitterzellen und Mittelung der Wignergleichung über diese Zellen ist es möglich, die Erhaltung der Teilchendichte für beliebig gewählte Gitter sicherzustellen. Daraus ergibt sich der wesentliche Vorteil der entwickelten Methode, nämlich die beliebig gute Auflösung von Teilgebieten des Phasenraums sowie die Größe des Grundgebietes in einfacher Weise an das konkrete physikalische Problem anpassen zu können. Zur Diskretisierung des Advektionsoperators in der Wignergleichung kommt ein bewährtes WENO (weighted essentially non oscillatory) Verfahren zur Anwendung. Der damit erhaltene Algorithmus wurde dann auf das Beispiel der resonanten Tunnelnioden angewendet und die berechnete Strom-Spannungs-Charakteristik mit Referenzwerten verglichen. Als Referenz dienten Simulationen, die mit der Technik der Nichtgleichgewichts-Greenfunktionen gewonnen wurden. Die Ergebnisse ließen eine sehr gute Übereinstimmung erkennen. Um alle Oszillationsmuster der berechneten Wignerfunktion hinreichend genau aufzulösen, scheint ein nicht äquidistant gewähltes Gitter unabdingbar zu sein. Aus physikalischer Sicht ist die Entdeckung markant ausgebildeter und überaus scharfer Oszillationsmuster in den berechneten, stationären Wignerfunktionen in denjenigen Fällen, in denen resonante Tunnelprozesse auftreten, besonders interessant. Im letzten Teil dieser Arbeit wurde ein einfaches zeitabhängiges Problem simuliert und die detaillierte Dynamik der Wignerfunktion analysiert.

Schlagwörter: Quantenmechanik im Phasenraum, Wignergleichung, Quantentransport, resonante Tunnelnioden, finite Volumen Methoden, REA Algorithmus, Nichtgleichgewichts-Greenfunktionen.

Contents

Abstract	i
Kurzfassung	iii
Introduction	1
1 Quantum mechanics in phase space	5
1.1 Density operator and the Wigner transformation	5
1.2 Equation of motion in phase space	11
2 Wigner transport equation	19
2.1 Single band, parabolic approximation	19
2.2 Properties of the pseudo-differential operator	23
3 Numerical methods for the Wigner transport equation	27
3.1 Methods based on the application of discrete Fourier transformation	27
3.1.1 Operator Splitting	28
3.1.2 Discrete Fourier Transformation	30
3.2 Discretization of the pseudo-differential operator based on a piecewise polynomial approximation of the Wigner function	33
3.2.1 Continuous approximation	34
3.2.2 Implementation of sine and cosine integrals	43
3.2.3 Application of finite volume methods	49
3.2.4 Numerical study on the accuracy of the Fourier transform of a piecewise polynomial approximation of a given function	83
3.3 Numerical schemes for the advection term and time-stepping methods	89
4 Non-equilibrium Green's function technique	101
4.1 Calculation of steady state properties	101
4.1.1 Coherent transport regime	105
4.2 Simulation of Resonant Tunneling Diodes	109
5 Simulation of Resonant Tunneling Diodes in phase space	115
5.1 Numerical tests on the time evolution to steady states	115
5.2 Numerical study on the accuracy and convergence of the algorithm	124
5.3 Transient response simulation	140
Conclusion	145

Acknowledgments	147
6 Appendix	149
6.1 Singularities in the Wigner function for the square potential barrier . .	149
Bibliography	157

Introduction

Nowadays, the most common approaches for the simulation of charge transport in semiconductor devices are drift-diffusion and hydrodynamic models. Such macroscopic descriptions of particle transport enable a rather quick computation and therefore, an industrial application. But, these models are only applicable as long as the device dimensions are large enough to justify the definition of macroscopic quantities like mobilities or diffusion coefficients. Depending on the particular application, such macroscopic transport models are applicable to devices with characteristic lengths in the range of micrometers or even several hundred nanometers. For smaller devices a microscopic description is necessary, usually based on the semi-classical Boltzmann transport equation (BTE). The BTE relies on the classical equations of motion stemming from the Liouville equation, whose characteristic lines are given by Newton's laws [1] [2] [3]. The quantum mechanical behaviour of the electrons in a crystal is then incorporated by replacing the free-particle dispersion relation $E = p^2/2m$ by the one obtained from band structure calculations and by adjusting the collision operator. On the one hand, one has to account for the appropriate particle statistics in the collision operator to arrive at the correct equilibrium distribution (Fermi-Dirac in the case of electrons) and on the other hand, the scattering matrix elements have to be determined by quantum mechanical calculations. Such a semi-classical description may enable an accurate simulation of the charge transport through various small-scaled devices. But, one has to bear in mind that the underlying equations of motion are classical and the electrons are thus treated as localized particles without accounting for their wave nature. Disregarding their wave-like propagation may cause on the one hand, errors in the results as soon as the spatial variations in the device are comparable to the wave length of electrons, causing for instance (quasi-bound) quantized states or interference effects. On the other hand, qualitative wrong results may be obtained as soon as tunneling effects become prominent, thus rendering a semi-classical description useless. A prototypical example for a semiconductor device whose behaviour is dominated by quantum mechanical effects is the resonant tunneling diode (RTD). This particular device type has been extensively used as a benchmark problem to evaluate the performance of different numerical methods for quantum transport models, see e.g [4] [5] [6] [7] [8], and will serve as a test case in this work as well.

In order to arrive at a quantum transport model suited for the application in device simulations, two fundamental demands must be fulfilled: On the one hand, in order to focus solely on the physics inside the device and to neglect the detailed time evolution of the electrons inside the contacts, which is not of interest of course, one has to devise a transport model describing an open quantum system. The considered spatial domain is then cut off at a certain point and appropriate boundary conditions

are imposed to model the inflow and outflow of charge carriers. On the other hand, it must be possible to include elastic and inelastic scattering mechanisms, between electrons and phonons, for instance, to enable an accurate description of the device behaviour at room temperature. A direct application of Schrödinger's equation for the purpose of device simulations is of limited use since the inclusion of scattering mechanisms is problematic. The problem arises due to the fact that inelastic scattering transitions require a more localized description of the electron states, constituted out of a superposition of different eigenstates [4]. An appropriate framework for such a description is the density operator and the corresponding von Neumann equation. But, Frensley [4] pointed out that it is problematic to impose appropriate boundary conditions for the density operator since only for the case of a sufficiently large damping inside the device a stable solution may be obtained. To circumvent this problem one can make use of the Wigner approach [1] [2] [9] [10] by transforming the density operator in the position space $\rho(x', x'', t) = \langle x' | \hat{\rho}(t) | x'' \rangle$ into a quasi-distribution function $f(p, x, t)$ in the phase space. The Wigner function $f(p, x, t)$ enables an analogous phase space description as in the case of the Boltzmann distribution function with the same marginals $n(x, t)$ and $n(p, t)$ when integrating $f(p, x, t)$ with respect to p and x , respectively. However, $f(p, x, t)$ does not share all properties of a probability distribution since it can take on negative values and is thus termed quasi-distribution function. Due to the analogy to the BTE one can make use of the same boundary conditions and clearly distinguish the incoming part from the outgoing part of the distribution $f(p, x, t)$.

In Chap. 1 the Wigner transformation is introduced together with discussing central properties of the Wigner function $f(p, x, t)$ before deriving Moyal's equation [9], the governing equation for the time evolution of $f(p, x, t)$ in the phase space. In addition, the \star -genvalue equation [9] is listed which constitutes together with Moyal's equation an autonomous formulation of quantum mechanics in phase space. After discussing the theory in general, Chap. 2 focuses on the derivation of the Wigner transport equation (WTE) from Moyal's equation for the case of a single, parabolic band and one spatial dimension. As we will see, the time evolution equation for $f(p, x, t)$ then contains an advection term as in the case of the BTE, but the electrostatic potential $V(x)$ enters in a complicated, non-local way represented by a pseudo-differential operator. When discussing central properties of this pseudo-differential operator, the connection of the WTE to the continuity equation is shown. Chapter 3 presents different numerical methods to treat the WTE. To begin with, the common approach to evaluate the pseudo-differential operator, based on the application of the discrete Fourier transformation is outlined. Its disadvantages are discussed in order to motivate the development of the discretization methods of the pseudo-differential operator built upon a piecewise polynomial approximation of the Wigner function. The first such method is termed continuous approximation to highlight its central feature. The description of the method is done in a fairly chronologic way, analogous to the original development of the method in the course of this thesis. We believe that such a description facilitates to understand the central points of the discretization approach. Due to the fact that several disadvantages are associated with the constraint of a

continuous approximation of the Wigner function, a second method without this restriction was developed. The method makes use of concepts known from finite volume methods and constitutes in principle a direct application of the REA algorithm [11]. The analytical calculations are presented in detail and general aspects are discussed to highlight the extensibility of the algorithm to higher-order approximations. All of the numerical calculations in the following chapters are performed with this second method. The chapter concludes with different numerical schemes for the advection term together with time-stepping methods. In Chap. 4 the non-equilibrium Green's function (NEGF) technique, a different approach to treat open quantum systems, is presented. The NEGF method has proven itself in the calculation of steady state properties of semiconductor devices and serves in this work as a reference solution for the Wigner function calculations. After stating the governing equations some numerical results are presented to estimate the accuracy of the calculated reference case. Finally, Chap. 5 presents numerical results obtained with the developed method for the WTE. In all numerical calculations a RTD is considered as test case. Before comparing the WTE and NEGF results quantitatively, the time evolution of the Wigner function to a steady state is analyzed in detail. In the case of under-resolved grids, a peculiar non-smooth time evolution may be observed with a sudden buildup of errors. The origin of this problems together with its manifestations are analyzed by comparing simulations with different grid spacings. After this preliminary considerations the influence of the various grid parameters on the accuracy of the obtained steady state $I(V)$ curves for the RTD are examined. A very good agreement between the Wigner and NEGF calculations is demonstrated for large enough grids, as well as the convergence of the results when refining the grid spacings. Finally, as an example for a time-dependent problem, the large-signal transient response of a RTD is considered.

1 Quantum mechanics in phase space

In Sec. 1.1 it is shown how to map the density operator onto the Wigner function. After introducing the Wigner transformation, it is demonstrated that the quantities of interest, the particle density, the momentum distribution and the current density are given by the same expressions as in classical physics. The last part focuses on the peculiarity of the Wigner function to be not positive-semidefinite in some cases. A simple example is considered to illustrate that interferences can lead to negative values of the Wigner function in certain regions of the phase space. In Sec. 1.2 the Moyal equation, representing the equation of motion for the Wigner distribution is derived by applying the Wigner mapping to the von Neumann equation. After mentioning interesting aspects of a formulation of quantum mechanics in phase space, an integral expression of the Moyal bracket is introduced for the later use in the Wigner transport equation.

1.1 Density operator and the Wigner transformation

Before discussing the Wigner function we recap some basic properties of the density operator. The wave function known from Schrödinger's formulation of quantum mechanics enables to describe pure states. In order to describe a general, mixed ensemble of states the density operator is the appropriate framework [12]. For a certain wave function $|\psi\rangle$ the density operator is given by [10]

$$\hat{\rho} = |\psi\rangle\langle\psi|, \quad (1.1)$$

written as the outer product of the ket $|\psi\rangle$ and the bra vector $\langle\psi|$. If $|\psi\rangle$ is expanded in a complete basis $|m\rangle$, such that

$$|\psi\rangle = \sum_{m=0}^{\infty} \psi_m |m\rangle, \quad (1.2)$$

the density operator can be rewritten as [10]

$$\hat{\rho} = \sum_{m,n=0}^{\infty} \rho_{m,n} |m\rangle\langle n|, \quad (1.3)$$

where the coefficients $\rho_{m,n} = \psi_m \psi_n^*$ determine the density matrix in the chosen basis [10]. Here we introduced the density matrix starting from a certain, pure state $|\psi\rangle$ but in the case of a mixed state such a connection to a single wave function does

not exist. In general, the diagonal elements $\rho_{m,m}$ of the density matrix determine the probability P_m to find the quantum system in the particular state $|m\rangle$ and the off-diagonal elements are sometimes termed coherences [10] [13]. Since the single probabilities P_m must add up to one the trace of the density operator, i.e. the sum of the diagonal elements of the density matrix is always equal to one

$$\begin{aligned}\mathrm{Tr}(\hat{\rho}) &= \sum_{n'=0}^{\infty} \langle n' | \hat{\rho} | n' \rangle \\ &= 1.\end{aligned}\tag{1.4}$$

For the square of the density operator the following equation holds true [13]

$$\mathrm{Tr}(\hat{\rho}^2) \leq 1,\tag{1.5}$$

referred to as the degree of coherence. The equality sign holds only for pure states and mixed states exhibit a degree of coherence smaller than one.

The expectation value of an operator \hat{A} is given in the framework of the density operator by [10]

$$\langle \hat{A} \rangle = \mathrm{Tr}(\hat{A}\hat{\rho}).\tag{1.6}$$

The time evolution of the density operator is determined by the von Neumann equation [10]

$$i\hbar\partial_t\hat{\rho} = [\hat{H}, \hat{\rho}],\tag{1.7}$$

where $[\hat{H}, \hat{\rho}] = \hat{H}\hat{\rho} - \hat{\rho}\hat{H}$ is the commutator with the Hamilton operator. The von Neumann equation will serve as a starting point in the next section to derive the equation of motion for the Wigner function.

Up to now we used a discrete basis to write down the density matrix, in coordinate representation the density operator is given in an analogous way by [2]

$$\rho(\mathbf{r}', \mathbf{r}'', t) = \langle \mathbf{r}' | \hat{\rho}(t) | \mathbf{r}'' \rangle, \quad \mathbf{r}', \mathbf{r}'' \in \mathbb{R}^3,\tag{1.8}$$

where we included the time dependence of ρ explicitly. Again, the diagonal elements of ρ determine the probability to find the quantum system in the particular state. Since $\mathbf{r}', \mathbf{r}''$ are continuous variables this corresponds to a probability density to be specific and directly represents the particle density at position \mathbf{r} [1]

$$n(\mathbf{r}, t) = \rho(\mathbf{r}, \mathbf{r}, t).\tag{1.9}$$

The off-diagonal elements represent again the coherences and determine the spatial correlations of the quantum system [4]. A second quantity of interest for device simulations is the current density $\mathbf{j}(\mathbf{r}, t)$ of electrons, given by [12] [1]

$$\mathbf{j}(\mathbf{r}, t) = -\frac{i\hbar q}{2m^*} (\nabla_{\mathbf{r}'} - \nabla_{\mathbf{r}''}) \rho(\mathbf{r}', \mathbf{r}'', t) \Big|_{\mathbf{r}'=\mathbf{r}''},\tag{1.10}$$

with $q = -e$, where $e > 0$ represents the elementary charge, and m^* the effective mass of the considered particles.

Now, to derive from the density matrix $\rho(\mathbf{r}', \mathbf{r}'', t)$ a (quasi) distribution function $f(\mathbf{p}, \mathbf{r}, t)$ in phase space, we perform a change of variables as a first step and Fourier transform then one of the spatial coordinates to the corresponding momentum coordinate. For this purpose we introduce the central coordinate \mathbf{r} and the difference coordinate ξ [10]

$$\begin{aligned}\mathbf{r} &= \frac{1}{2}(\mathbf{r}' + \mathbf{r}''), \\ \xi &= \mathbf{r}' - \mathbf{r}'',\end{aligned}\tag{1.11}$$

for which the density matrix is then given by [10]

$$\begin{aligned}\tilde{\rho}(\mathbf{r}, \xi, t) &= \rho(\mathbf{r}', \mathbf{r}'', t) \\ &= \rho\left(\mathbf{r} + \frac{1}{2}\xi, \mathbf{r} - \frac{1}{2}\xi, t\right).\end{aligned}\tag{1.12}$$

Upon Fourier transformation with respect to the difference coordinate ξ we arrive at the Wigner distribution $f(\mathbf{p}, \mathbf{r}, t)$ [10]

$$f(\mathbf{p}, \mathbf{r}, t) = \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \tilde{\rho}(\mathbf{r}, \xi, t) \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\xi.\tag{1.13}$$

The Wigner transformation can also be written directly in terms of the density operator as [10]

$$f(\mathbf{p}, \mathbf{r}, t) = \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \left\langle \mathbf{r} + \frac{1}{2}\xi \left| \hat{\rho}(t) \right| \mathbf{r} - \frac{1}{2}\xi \right\rangle \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\xi.\tag{1.14}$$

In general, the Wigner transformation maps an operator onto a c-number phase space function. For the case of scalar, Hermitian operators the corresponding phase space functions are real, as one can easily show:

$$\begin{aligned}f^\dagger(\mathbf{p}, \mathbf{r}, t) &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \left[\left\langle \mathbf{r} + \frac{1}{2}\xi \left| \hat{\rho}(t) \right| \mathbf{r} - \frac{1}{2}\xi \right\rangle \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) \right]^\dagger d\xi \\ &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \left\langle \mathbf{r} - \frac{1}{2}\xi \left| \hat{\rho}^\dagger(t) \right| \mathbf{r} + \frac{1}{2}\xi \right\rangle \exp\left(i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\xi \\ &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \left\langle \mathbf{r} + \frac{1}{2}\xi' \left| \hat{\rho}^\dagger(t) \right| \mathbf{r} - \frac{1}{2}\xi' \right\rangle \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi'\right) d\xi' \\ &= f(\mathbf{p}, \mathbf{r}, t),\end{aligned}\tag{1.15}$$

when substituting $\xi' = -\xi$. For the situation that a Hermitian $\hat{\rho}$ consists of a matrix of operators $\hat{\rho}_{i,j}$, one can only deduce that the corresponding $f(\mathbf{p}, \mathbf{r}, t)$ has the form of a Hermitian matrix as well. This would be, for instance, the case when describing a two-band transport where $\hat{\rho}_{0,0}$ and $\hat{\rho}_{1,1}$ describe the occupancy of the lower and upper band, respectively, and $\hat{\rho}_{0,1} = (\hat{\rho}_{1,0})^*$ the coupling between the two bands. Then, the diagonal elements $f_{i,i}(\mathbf{p}, \mathbf{r}, t)$ would be real as well but the off-diagonal elements

$f_{0,1}(\mathbf{p}, \mathbf{r}, t) = [f_{1,0}(\mathbf{p}, \mathbf{r}, t)]^*$ could take on complex values. Since we restrict ourselves in this work to the single band case, $f(\mathbf{p}, \mathbf{r}, t)$ is viewed to be a real-valued function.

The marginal distributions of $f(\mathbf{p}, \mathbf{r}, t)$ are obtained upon integration with respect to \mathbf{p} and \mathbf{r} . In the first case we arrive at [10]

$$\begin{aligned}
 f(\mathbf{r}, t) &= \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} \\
 &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \left\langle \mathbf{r} + \frac{1}{2}\xi \left| \hat{\rho}(t) \right| \mathbf{r} - \frac{1}{2}\xi \right\rangle \int_{\mathbb{R}^3} \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\mathbf{p} d\xi \\
 &= \int_{\mathbb{R}^3} \left\langle \mathbf{r} + \frac{1}{2}\xi \left| \hat{\rho}(t) \right| \mathbf{r} - \frac{1}{2}\xi \right\rangle \delta(\xi) d\xi \\
 &= \langle \mathbf{r} | \hat{\rho}(t) | \mathbf{r} \rangle \\
 &= n(\mathbf{r}, t),
 \end{aligned} \tag{1.16}$$

when making use of the relation

$$\delta(\mathbf{x}) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \exp(i\mathbf{k} \cdot \mathbf{x}) d\mathbf{k}, \quad \mathbf{x}, \mathbf{k} \in \mathbb{R}^N. \tag{1.17}$$

To calculate the second marginal we return to Eq. (1.11) to rewrite the integrals as [10]

$$\begin{aligned}
 f(\mathbf{p}, t) &= \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{r} \\
 &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \left\langle \mathbf{r} + \frac{1}{2}\xi \left| \hat{\rho}(t) \right| \mathbf{r} - \frac{1}{2}\xi \right\rangle \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\mathbf{r} d\xi \\
 &= \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \langle \mathbf{r}' | \hat{\rho}(t) | \mathbf{r}'' \rangle \exp\left[-i\frac{\mathbf{p}}{\hbar} \cdot (\mathbf{r}' - \mathbf{r}'')\right] d\mathbf{r}' d\mathbf{r}'' .
 \end{aligned} \tag{1.18}$$

With the knowledge of the expression for the momentum eigenstate $|\mathbf{p}\rangle$ in position space [12]

$$\langle \mathbf{r} | \mathbf{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(i\frac{\mathbf{p}}{\hbar} \cdot \mathbf{r}\right), \tag{1.19}$$

as well as the completeness relation

$$\mathbb{1} = \int_{\mathbb{R}^3} |\mathbf{r}\rangle \langle \mathbf{r}| d\mathbf{r}, \tag{1.20}$$

equation (1.18) can be related to the momentum distribution [10]

$$\begin{aligned}
 f(\mathbf{p}, t) &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \langle \mathbf{p} | \mathbf{r}' \rangle \langle \mathbf{r}' | \hat{\rho}(t) | \mathbf{r}'' \rangle \langle \mathbf{r}'' | \mathbf{p} \rangle d\mathbf{r}' d\mathbf{r}'' \\
 &= \langle \mathbf{p} | \hat{\rho}(t) | \mathbf{p} \rangle .
 \end{aligned} \tag{1.21}$$

The two marginals $f(\mathbf{r}, t)$ and $f(\mathbf{p}, t)$ were just related to the two desired quantities, the position distribution (i.e. particle density) and the momentum distribution, respectively, in analogy to a classical distribution function $f_{cl.}(\mathbf{p}, \mathbf{r}, t)$. To arrive at an

expression connecting the current density $\mathbf{j}(\mathbf{r}, t)$ with $f(\mathbf{p}, \mathbf{r}, t)$, we rewrite Eq. (1.10) as a first step by making use of Eqs. (1.11) and (1.12) as

$$\mathbf{j}(\mathbf{r}, t) = -\frac{i\hbar q}{m^*} \nabla_{\xi} \tilde{\rho}(\mathbf{r}, \xi, t) \Big|_{\xi=0}. \quad (1.22)$$

The inverse relation of Eq. (1.13) is given by

$$\tilde{\rho}(\mathbf{r}, \xi, t) = \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) \exp\left(i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\mathbf{p}, \quad (1.23)$$

so that Eq. (1.22) can be expressed as

$$\begin{aligned} \mathbf{j}(\mathbf{r}, t) &= -\frac{i\hbar q}{m^*} \left[\nabla_{\xi} \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) \exp\left(i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\mathbf{p} \right] \Big|_{\xi=0} \\ &= \frac{q}{m^*} \left[\int_{\mathbb{R}^3} \mathbf{p} f(\mathbf{p}, \mathbf{r}, t) \exp\left(i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\mathbf{p} \right] \Big|_{\xi=0} \\ &= \frac{q}{m^*} \int_{\mathbb{R}^3} \mathbf{p} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p}. \end{aligned} \quad (1.24)$$

This expression is as well in complete analogy to the case of the classical Boltzmann phase space distribution.

As a final point, we inspect the peculiar property of the Wigner distribution function to take on negative values, in some cases at least. For this we state without derivation the trace product rule [10]

$$\text{Tr}(\hat{\rho}_1 \hat{\rho}_2) = (2\pi\hbar)^3 \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f_{\hat{\rho}_1}(\mathbf{p}, \mathbf{r}, t) f_{\hat{\rho}_2}(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} d\mathbf{r}. \quad (1.25)$$

In the special case of two pure, orthogonal states $\hat{\rho}_1 = |\psi_1\rangle\langle\psi_1|$ and $\hat{\rho}_2 = |\psi_2\rangle\langle\psi_2|$ the left-hand side of the equation above vanishes:

$$\begin{aligned} \text{Tr}(\hat{\rho}_1 \hat{\rho}_2) &= \text{Tr}(|\psi_1\rangle\langle\psi_1| |\psi_2\rangle\langle\psi_2|) \\ &= |\langle\psi_1|\psi_2\rangle|^2 \\ &= 0. \end{aligned} \quad (1.26)$$

Such orthogonal states could be for instance the eigenfunctions of the harmonic oscillator or any other bound states. In order that the right-hand side of Eq. (1.25) vanishes as well, it must be possible that the Wigner functions $f_{\hat{\rho}_1}(\mathbf{p}, \mathbf{r}, t)$ and/or $f_{\hat{\rho}_2}(\mathbf{p}, \mathbf{r}, t)$ can take on negative values, leading to cancellations in the phase space integral [10]. The possibility of a Wigner function which is positive everywhere is therefore not excluded, but the occurrence of negative values in some cases hampers to view the Wigner function as a true probability distribution.

To become more familiar with the negativity of the Wigner function we consider two simple examples of a coherent and an incoherent mixture of two plane waves traveling in opposite directions. The two plane waves are chosen to be given by

$$\langle \mathbf{r} | \pm \mathbf{k} \rangle = \exp(\pm i \mathbf{k} \cdot \mathbf{r}). \quad (1.27)$$

The density matrix for the incoherent case is labeled mixed and determined by

$$\hat{\rho}_{mix.} = \frac{1}{2}(|\mathbf{k}\rangle\langle\mathbf{k}| + |-\mathbf{k}\rangle\langle-\mathbf{k}|). \quad (1.28)$$

For the coherent case we choose the entangled state $|\psi\rangle = \frac{1}{\sqrt{2}}(|\mathbf{k}\rangle + |-\mathbf{k}\rangle)$ and evaluate the corresponding density matrix $|\psi\rangle\langle\psi|$ to arrive at

$$\hat{\rho}_{coh.} = \frac{1}{2}(|\mathbf{k}\rangle\langle\mathbf{k}| + |-\mathbf{k}\rangle\langle-\mathbf{k}| + |-\mathbf{k}\rangle\langle\mathbf{k}| + |\mathbf{k}\rangle\langle-\mathbf{k}|). \quad (1.29)$$

It is obvious to see that $\hat{\rho}_{mix.}$ is diagonal when calculating the corresponding 2×2 density matrix with the states $|\pm \mathbf{k}\rangle$, whereas $\hat{\rho}_{coh.}$ exhibits nonzero off-diagonal elements, i.e. coherences. In coordinate representation the density matrices can be written as

$$\begin{aligned} \langle \mathbf{r}' | \hat{\rho}_{mix.} | \mathbf{r}'' \rangle &= \frac{1}{2} \{ \exp[i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r}'')] + \exp[-i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r}'')] \}, \\ \tilde{\rho}_{mix.}(\mathbf{r}, \xi) &= \cos(\mathbf{k} \cdot \xi), \end{aligned} \quad (1.30)$$

as well as

$$\begin{aligned} \langle \mathbf{r}' | \hat{\rho}_{coh.} | \mathbf{r}'' \rangle &= \frac{1}{2} \{ \exp[i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r}'')] + \exp[-i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r}'')] \\ &\quad + \exp[i\mathbf{k} \cdot (\mathbf{r}' + \mathbf{r}'')] + \exp[-i\mathbf{k} \cdot (\mathbf{r}' + \mathbf{r}'')] \}, \\ \tilde{\rho}_{coh.}(\mathbf{r}, \xi) &= \cos(\mathbf{k} \cdot \xi) + \cos(2\mathbf{k} \cdot \mathbf{r}). \end{aligned} \quad (1.31)$$

The corresponding Wigner functions can then be calculated by making use of Eq. (1.13) and are given by

$$f_{mix.}(\mathbf{p}, \mathbf{r}) = \frac{1}{2} [\delta(\mathbf{p} - \hbar\mathbf{k}) + \delta(\mathbf{p} + \hbar\mathbf{k})] \quad (1.32)$$

and by

$$f_{coh.}(\mathbf{p}, \mathbf{r}) = \frac{1}{2} [\delta(\mathbf{p} - \hbar\mathbf{k}) + \delta(\mathbf{p} + \hbar\mathbf{k})] + \cos(2\mathbf{k} \cdot \mathbf{r})\delta(\mathbf{p}). \quad (1.33)$$

As one can see, the Wigner function for the mixed case is strictly positive and exhibits two delta function peaks at the corresponding momenta of the two plane waves. With respect to \mathbf{r} the function is constant. In contrast, the Wigner function for the coherent case displays another delta function at $\mathbf{p} = 0$, modulated along \mathbf{r} by a cosine function. Clearly, $f_{coh.}(\mathbf{p}, \mathbf{r})$ can take on negative values. Upon integration of the Wigner functions with respect to the momentum coordinate one obtains the following particle densities:

$$n_{mix.}(\mathbf{r}) = 1 \quad (1.34)$$

and

$$n_{coh.}(\mathbf{r}) = 1 + \cos(2\mathbf{k} \cdot \mathbf{r}). \quad (1.35)$$

Obviously, the particle density for the mixed state has a constant value of 1 since the single plane waves do not interact with each other, whereas the particle density for the coherent state oscillates and takes on values between 0 and 2. This stationary oscillation pattern is caused by the interference of the two plane waves traveling in opposite directions. The results enable to draw the preliminary conclusion that interference effects manifest itself in phase space as oscillating regions, causing the Wigner function to become negative-valued as well.

To sum up, the central results of this section are that it is possible to map the density operator onto a phase space distribution $f(\mathbf{p}, \mathbf{r}, t)$ by making use of the Wigner transformation as defined in Eq. (1.14). The Wigner function $f(\mathbf{p}, \mathbf{r}, t)$ obtained by this is real-valued (single band case) and the marginals as well as the first moment with respect to p result in the same quantities as in the case of the classical Boltzmann probability distribution. A distinct feature of the Wigner function is the possibility to take on negative values, hampering to view it as a real probability distribution. The Wigner function is thus also termed quasi-distribution function. The quantities of interest for device simulations are the particle density $n(\mathbf{r}, t)$ and the current density $\mathbf{j}(\mathbf{r}, t)$, determined by the zeroth and first moment of $f(\mathbf{p}, \mathbf{r}, t)$ with respect to \mathbf{p} :

$$\begin{aligned} n(\mathbf{r}, t) &= \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p}, \\ \mathbf{j}(\mathbf{r}, t) &= \frac{q}{m^*} \int_{\mathbb{R}^3} \mathbf{p} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p}. \end{aligned} \quad (1.36)$$

1.2 Equation of motion in phase space

In Eq. (1.14) it was shown how to transform the density operator $\hat{\rho}$ to the Wigner function $f(\mathbf{p}, \mathbf{r}, t)$. This mapping can be extended to arbitrary operators \hat{A} by the Weyl-Wigner correspondence [10]

$$A(\mathbf{p}, \mathbf{r}, t) = \int_{\mathbb{R}^3} \left\langle \mathbf{r} + \frac{1}{2}\boldsymbol{\xi} \left| \hat{A}(t) \right| \mathbf{r} - \frac{1}{2}\boldsymbol{\xi} \right\rangle \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \boldsymbol{\xi}\right) d\boldsymbol{\xi}. \quad (1.37)$$

In this way Hilbert space operators can be mapped onto c-number phase space functions $A(\mathbf{p}, \mathbf{r}, t)$. As before, in the case of scalar, Hermitian operators the corresponding phase space functions are real-valued. The expectation value of an operator,

$$\langle \hat{A} \rangle = \text{Tr} \left(\hat{A} \hat{\rho} \right), \quad (1.38)$$

can be expressed in terms of phase space functions as [10]

$$\langle \hat{A} \rangle = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} A(\mathbf{p}, \mathbf{r}, t) f(\mathbf{p}, \mathbf{r}, t) d\mathbf{r} d\mathbf{p}, \quad (1.39)$$

when making use of the trace product rule Eq. (1.25). The equation is again formally equivalent to classical statistical physics [10].

A central component which has not been discussed yet is the equation of motion for $f(\mathbf{p}, \mathbf{r}, t)$. In the case of the density matrix, the von Neumann equation

$$i\hbar\partial_t\hat{\rho} = [\hat{H}, \hat{\rho}] \quad (1.40)$$

determines the time evolution of $\hat{\rho}$. To find the corresponding equation in phase space we now apply a Wigner transformation to both sides of Eq. (1.40). The result of this operation on the left-hand side is simply the time derivative of the Wigner function $i\hbar\partial_t f(\mathbf{p}, \mathbf{r}, t)$. The result of the right-hand side is more complicated to evaluate. On the one hand, we need to know the phase space function $H(\mathbf{p}, \mathbf{r}, t)$ to the Hamilton operator $\hat{H}(t)$. This can be calculated in general when making use of Eq. (1.37) and for the special case of a Hamiltonian of the form $\hat{H} = T(\hat{\mathbf{p}}) + U(\hat{\mathbf{r}}, t)$, i.e. without products of non-commuting operators, the corresponding phase space function is simply given by replacing the operators by their eigenvalues, $H(\mathbf{p}, \mathbf{r}, t) = T(\mathbf{p}) + U(\mathbf{r}, t)$ [9]. On the other hand, we need to know the result of applying the Wigner transformation to the product of two operators. This is calculated in the following.

We now consider the general case of an operator \hat{Z} given as the product of two other operators, i.e. by $\hat{Z} = \hat{X}\hat{Y}$, and calculate the corresponding phase space function $Z(\mathbf{p}, \mathbf{r})$ in terms of $X(\mathbf{p}, \mathbf{r})$ and $Y(\mathbf{p}, \mathbf{r})$ by evaluating Eq. (1.37). The temporal variable t is suppressed for the sake of clarity and the derivation listed below is in close analogy to the one given in [13]. The coordinate representation of \hat{Z} can be written as

$$\langle \mathbf{r}' | \hat{Z} | \mathbf{r}'' \rangle = \int_{\mathbb{R}^3} X(\mathbf{r}', \mathbf{s}) Y(\mathbf{s}, \mathbf{r}'') d\mathbf{s}, \quad (1.41)$$

which results upon inserting in Eq. (1.37) in

$$Z(\mathbf{p}, \mathbf{r}) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} X\left(\mathbf{r} + \frac{1}{2}\xi, \mathbf{s}\right) Y\left(\mathbf{s}, \mathbf{r} - \frac{1}{2}\xi\right) \exp\left(-i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\xi d\mathbf{s}. \quad (1.42)$$

To express the operators \hat{X} and \hat{Y} in terms of one spatial and one momentum coordinate as well, we make use of the inverse relation of Eq. (1.37)

$$A\left(\mathbf{r} + \frac{1}{2}\xi, \mathbf{r} - \frac{1}{2}\xi\right) = \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} A(\mathbf{p}, \mathbf{r}) \exp\left(i\frac{\mathbf{p}}{\hbar} \cdot \xi\right) d\mathbf{p}. \quad (1.43)$$

This enables us to rewrite the two terms in Eq. (1.42) as

$$X\left(\mathbf{r} + \frac{1}{2}\xi, \mathbf{s}\right) = \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} X\left(\mathbf{p}', \frac{\mathbf{r} + \mathbf{s}}{2} + \frac{\xi}{4}\right) \exp\left[i\frac{\mathbf{p}'}{\hbar} \cdot \left(\mathbf{r} - \mathbf{s} + \frac{\xi}{2}\right)\right] d\mathbf{p}' \quad (1.44)$$

and

$$Y\left(\mathbf{s}, \mathbf{r} - \frac{1}{2}\xi\right) = \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} Y\left(\mathbf{p}'', \frac{\mathbf{r} + \mathbf{s}}{2} - \frac{\xi}{4}\right) \exp\left[-i\frac{\mathbf{p}''}{\hbar} \cdot \left(\mathbf{r} - \mathbf{s} - \frac{\xi}{2}\right)\right] d\mathbf{p}'', \quad (1.45)$$

when identifying the appropriate central and difference coordinates. Inserting this into Eq. (1.42) results in

$$\begin{aligned}
 Z(\mathbf{p}, \mathbf{r}) &= \frac{1}{(2\pi\hbar)^6} \iiint\!\!\!\int X \left[\mathbf{p}', \mathbf{r} - \frac{1}{2} \left(\mathbf{r} - \mathbf{s} - \frac{\xi}{2} \right) \right] Y \left[\mathbf{p}'', \mathbf{r} - \frac{1}{2} \left(\mathbf{r} - \mathbf{s} + \frac{\xi}{2} \right) \right] \\
 &\quad \times \exp \left[i \frac{\mathbf{p}'}{\hbar} \cdot \left(\mathbf{r} - \mathbf{s} + \frac{\xi}{2} \right) \right] \exp \left[-i \frac{\mathbf{p}''}{\hbar} \cdot \left(\mathbf{r} - \mathbf{s} - \frac{\xi}{2} \right) \right] \\
 &\quad \times \exp \left(-i \frac{\mathbf{p}}{\hbar} \cdot \xi \right) d\mathbf{p}' d\mathbf{p}'' d\xi d\mathbf{s}. \tag{1.46}
 \end{aligned}$$

It is convenient to perform the following change of variables:

$$\begin{aligned}
 \eta' &= \mathbf{r} - \mathbf{s} - \frac{\xi}{2}, \\
 \eta'' &= - \left(\mathbf{r} - \mathbf{s} + \frac{\xi}{2} \right). \tag{1.47}
 \end{aligned}$$

Since the value of the corresponding Jacobian is given by one, we arrive at the expression

$$\begin{aligned}
 Z(\mathbf{p}, \mathbf{r}) &= \frac{1}{(2\pi\hbar)^6} \iiint\!\!\!\int X \left(\mathbf{p}', \mathbf{r} - \frac{1}{2}\eta' \right) Y \left(\mathbf{p}'', \mathbf{r} + \frac{1}{2}\eta'' \right) \exp \left[i \frac{\mathbf{p}}{\hbar} \cdot (\eta' + \eta'') \right] \\
 &\quad \times \exp \left(-i \frac{\mathbf{p}'}{\hbar} \cdot \eta'' \right) \exp \left(-i \frac{\mathbf{p}''}{\hbar} \cdot \eta' \right) d\mathbf{p}' d\mathbf{p}'' d\eta' d\eta''. \tag{1.48}
 \end{aligned}$$

As a next step we introduce the so called Bopp shifts [2] [9]

$$f(\mathbf{x} + \mathbf{a}) = \exp(\mathbf{a} \cdot \nabla_{\mathbf{x}}) f(\mathbf{x}), \tag{1.49}$$

which can be seen as a Taylor series expansion to infinite order. In addition, we introduce the directional nabla operators $\overleftarrow{\nabla}$ and $\overrightarrow{\nabla}$ which are meant to act to the left and to the right, respectively. With this in hands it is possible to rewrite the terms of the integrand in the first line of Eq. (1.48) as

$$\begin{aligned}
 I_a &= X \left(\mathbf{p}', \mathbf{r} - \frac{1}{2}\eta' \right) Y \left(\mathbf{p}'', \mathbf{r} + \frac{1}{2}\eta'' \right) \exp \left[i \frac{\mathbf{p}}{\hbar} \cdot (\eta' + \eta'') \right] \\
 &= X(\mathbf{p}', \mathbf{r}) \exp \left(-\overleftarrow{\nabla}_{\mathbf{r}} \cdot \frac{\eta'}{2} \right) \exp \left(\frac{\eta''}{2} \cdot \overrightarrow{\nabla}_{\mathbf{r}} \right) Y(\mathbf{p}'', \mathbf{r}) \exp \left[i \frac{\mathbf{p}}{\hbar} \cdot (\eta' + \eta'') \right] \\
 &= X(\mathbf{p}', \mathbf{r}) \exp \left(i \frac{\mathbf{p}}{\hbar} \cdot \eta'' \right) \exp \left(-i \frac{\hbar}{2} \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}} \right) \exp \left(i \frac{\hbar}{2} \overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} \right) \\
 &\quad \times \exp \left(i \frac{\mathbf{p}}{\hbar} \cdot \eta' \right) Y(\mathbf{p}'', \mathbf{r}), \tag{1.50}
 \end{aligned}$$

where we made use of the Bopp shifts in the second and in the third line. As one can see, the first two terms depend solely on the integration variables \mathbf{p}' and η'' but not on \mathbf{p}'' and η' , and vice versa in the case of the last two terms. The exponentials in

the second line of Eq. (1.48) exhibit an analogous dependence, so that it is possible to split up the quadruple integral into a product of two double integrals in such a way that

$$Z(\mathbf{p}, \mathbf{r}) = \frac{1}{(2\pi\hbar)^3} \iint X(\mathbf{p}', \mathbf{r}) \exp\left(i\frac{\mathbf{p} - \mathbf{p}'}{\hbar} \cdot \boldsymbol{\eta}''\right) d\mathbf{p}' d\boldsymbol{\eta}'' \exp\left[i\frac{\hbar}{2} \left(\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} - \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}\right)\right] \\ \times \frac{1}{(2\pi\hbar)^3} \iint Y(\mathbf{p}'', \mathbf{r}) \exp\left(i\frac{\mathbf{p} - \mathbf{p}''}{\hbar} \cdot \boldsymbol{\eta}'\right) d\mathbf{p}'' d\boldsymbol{\eta}' . \quad (1.51)$$

When recalling the representation of the delta function listed in Eq. (1.17), it is straightforward to identify the first integral expression as

$$\frac{1}{(2\pi\hbar)^3} \iint X(\mathbf{p}', \mathbf{r}) \exp\left(i\frac{\mathbf{p} - \mathbf{p}'}{\hbar} \cdot \boldsymbol{\eta}''\right) d\mathbf{p}' d\boldsymbol{\eta}'' = \int X(\mathbf{p}', \mathbf{r}) \delta(\mathbf{p} - \mathbf{p}') d\mathbf{p}' \\ = X(\mathbf{p}, \mathbf{r}) \quad (1.52)$$

and in the same manner the second integral expression as $Y(\mathbf{p}, \mathbf{r})$. The exponential operators in between $X(\mathbf{p}, \mathbf{r})$ and $Y(\mathbf{p}, \mathbf{r})$ are usually collected in the symbol [9]

$$\star = \exp\left[i\frac{\hbar}{2} \left(\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} - \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}\right)\right] , \quad (1.53)$$

enabling to write the final result in the compact notation

$$Z(\mathbf{p}, \mathbf{r}) = X(\mathbf{p}, \mathbf{r}) \star Y(\mathbf{p}, \mathbf{r}) . \quad (1.54)$$

With this in hands, we are now able to write down the phase space equivalent to the von Neumann equation. When comparing the definition of the Wigner transformation and the Wigner-Weyl correspondence, Eqs. (1.14) and (1.37), respectively, one notices a difference of a factor of $(2\pi\hbar)^3$. But, since the inverse relation Eq. (1.43) is involved in the derivation of the star product, the factors balance each other and one finds that the product $\hat{H}\hat{\rho}$ is mapped onto

$$\hat{H}(t)\hat{\rho}(t) \mapsto H(\mathbf{p}, \mathbf{r}, t) \star f(\mathbf{p}, \mathbf{r}, t) . \quad (1.55)$$

Therefore, upon applying a Wigner transformation to the von Neumann equation, (1.40), one finds the phase space analogue, Moyal's equation [9],

$$\partial_t f(\mathbf{p}, \mathbf{r}, t) = \{\{H(\mathbf{p}, \mathbf{r}, t), f(\mathbf{p}, \mathbf{r}, t)\}\} , \quad (1.56)$$

with the Moyal bracket defined as [9]

$$\{\{H(\mathbf{p}, \mathbf{r}, t), f(\mathbf{p}, \mathbf{r}, t)\}\} = \frac{H(\mathbf{p}, \mathbf{r}, t) \star f(\mathbf{p}, \mathbf{r}, t) - f(\mathbf{p}, \mathbf{r}, t) \star H(\mathbf{p}, \mathbf{r}, t)}{i\hbar} . \quad (1.57)$$

In addition to the time evolution equation (1.56) and in analogy to the eigenvalue equations of Hilbert space quantum mechanics, one finds the following system of equations for static Wigner functions [9]

$$H(\mathbf{p}, \mathbf{r}, t) \star f(\mathbf{p}, \mathbf{r}, t) = Ef(\mathbf{p}, \mathbf{r}, t) , \\ f(\mathbf{p}, \mathbf{r}, t) \star H(\mathbf{p}, \mathbf{r}, t) = Ef(\mathbf{p}, \mathbf{r}, t) . \quad (1.58)$$

Moyal's equation (1.56) together with the \star -genvalue [9] equations (1.58) completely determine possible Wigner functions and thus constitute an autonomous formulation of quantum mechanics in phase space [9].

The star product \star is a central concept in this formulation and incorporates the non-commutativity of Hilbert space operators. To evaluate the expressions explicitly, various techniques are helpful. For instance, one can make use of the Bopp shifts as defined in Eq. (1.49) to obtain [9]

$$H(\mathbf{p}, \mathbf{r}, t) \star f(\mathbf{p}, \mathbf{r}, t) = H\left(\mathbf{p} - i\frac{\hbar}{2}\overleftarrow{\nabla}_{\mathbf{r}}, \mathbf{r} + i\frac{\hbar}{2}\overrightarrow{\nabla}_{\mathbf{p}}, t\right) f(\mathbf{p}, \mathbf{r}, t). \quad (1.59)$$

When inserting the particular expression for $H(\mathbf{p}, \mathbf{r}, t)$ one may end up with a simple differential equation. A prototypical example for an eigenvalue problem in phase space is the harmonic oscillator, as in Hilbert space quantum mechanics. In this case applying the Bopp shifts to the \star -genvalue equations (1.58) results in a system of differential equations for $f(\mathbf{p}, \mathbf{r}, t)$, whose solution involves Laguerre polynomials, as outlined in detail in [10] or [9]. Another useful technique is to apply a Taylor expansion or also called gradient expansion [14] to the star product:

$$\star = 1 + i\frac{\hbar}{2}\left(\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} - \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}\right) - \frac{\hbar^2}{8}\left(\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} - \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}\right)^2 + \mathcal{O}(\hbar^3). \quad (1.60)$$

This is of particular use when considering the classical limit $\hbar \rightarrow 0$, or if one is interested in calculating the leading order quantum corrections for transport equations. The leading order terms with respect to \hbar of the Moyal bracket are

$$\begin{aligned} \{\{H, f\}\} &= \frac{1}{i\hbar} \left[Hf - fH + i\frac{\hbar}{2}H\left(\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} - \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}\right)f \right. \\ &\quad \left. - i\frac{\hbar}{2}f\left(\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}} - \overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}\right)H + \mathcal{O}(\hbar^2) \right] \\ &= H\overleftarrow{\nabla}_{\mathbf{r}} \cdot \overrightarrow{\nabla}_{\mathbf{p}}f - H\overleftarrow{\nabla}_{\mathbf{p}} \cdot \overrightarrow{\nabla}_{\mathbf{r}}f + \mathcal{O}(\hbar), \end{aligned} \quad (1.61)$$

when viewing H and f again as scalars and not as matrices. As a result, the Moyal bracket reduces in the classical limit to the Poisson bracket [9]

$$\begin{aligned} \lim_{\hbar \rightarrow 0} \{\{H, f\}\} &= \{H, f\} \\ &= (\nabla_{\mathbf{r}}H) \cdot (\nabla_{\mathbf{p}}f) - (\nabla_{\mathbf{p}}H) \cdot (\nabla_{\mathbf{r}}f). \end{aligned} \quad (1.62)$$

Quantum mechanics in phase space is the natural framework to calculate quantum corrections to classical statistical physics due to the analogous descriptions. In both cases a probabilistic description for an ensemble of particles is used (or also for a single particle alone in the case of quantum mechanics) and the measurable quantities are given by formally equivalent expressions [10].

Nevertheless, the aim of this work is not the consideration of the classical limit but a full quantum mechanical description of the charge transport in tunneling devices.

Due to the fact that strong spatial variations of the potential $U(\mathbf{r}, t)$ are present in such cases, a gradient expansion as outlined in Eq. (1.60) cannot be applied and the full star product has to be considered. For this purpose it is convenient to introduce an integral representation of the star product. For simplicity and the later use of the expressions in the next chapter we restrict ourselves to the one-dimensional case $p, x \in \mathbb{R}$. For the part of Moyal's equation (1.56) containing the potential $U(x, t)$ we need to consider

$$\begin{aligned} U(x, t) \star f(p, x, t) &= U(x, t) \exp \left[\frac{i\hbar}{2} \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) \right] f(p, x, t) \\ &= U(x, t) \exp \left(\frac{i\hbar}{2} \overleftarrow{\partial}_x \overrightarrow{\partial}_p \right) f(p, x, t). \end{aligned} \quad (1.63)$$

We now replace $f(p, x, t)$ by the equivalent expression

$$\begin{aligned} f(p, x, t) &= \int_{-\infty}^{\infty} f(p', x, t) \delta(p - p') dp' \\ &= \frac{1}{2\pi\hbar} \iint f(p', x, t) \exp \left(i \frac{p - p'}{\hbar} \eta' \right) dp' d\eta', \end{aligned} \quad (1.64)$$

so that we find for Eq. (1.63) when applying two times the Bopp shifts:

$$\begin{aligned} U(x, t) \star f(p, x, t) &= \frac{1}{2\pi\hbar} \iint U(x, t) \exp \left(\frac{i\hbar}{2} \overleftarrow{\partial}_x \overrightarrow{\partial}_p \right) \\ &\quad \times f(p', x, t) \exp \left(i \frac{p - p'}{\hbar} \eta' \right) dp' d\eta' \\ &= \frac{1}{2\pi\hbar} \iint U(x, t) \exp \left(i \frac{p - p'}{\hbar} \eta' - \frac{\eta'}{2} \overleftarrow{\partial}_x \right) f(p', x, t) dp' d\eta' \\ &= \frac{1}{2\pi\hbar} \iint U \left(x - \frac{\eta'}{2}, t \right) f(p', x, t) \exp \left(i \frac{p - p'}{\hbar} \eta' \right) dp' d\eta'. \end{aligned} \quad (1.65)$$

As a result, the Moyal bracket for the potential term of the Hamiltonian can be written in integral representation as

$$\begin{aligned} \{ \{ U(x, t), f(p, x, t) \} \} &= \frac{i}{2\pi\hbar^2} \iint \left[U \left(x + \frac{\eta'}{2}, t \right) - U \left(x - \frac{\eta'}{2}, t \right) \right] f(p', x, t) \\ &\quad \times \exp \left(i \frac{p - p'}{\hbar} \eta' \right) dp' d\eta'. \end{aligned} \quad (1.66)$$

We will return to this equation in the next chapter in the context of the pseudo-differential operator.

As a side note: In Eq. (1.54) we derived in general how the product $\hat{Z} = \hat{X}\hat{Y}$ of two operators is mapped onto the corresponding phase space function $Z(p, x) = X(p, x) \star$

$Y(p, x)$. The star product \star is a central concept and enables a quick calculation of the phase space representation of more general Hamiltonians. A common problem for device physics is to incorporate a spatially dependent effective mass $m^*(x)$. In this case, the kinetic part of the Hamilton operator exhibits a dependence on the position coordinate and is usually written in a symmetrically ordered form as [15] [16]

$$T(\hat{p}, \hat{x}) = \hat{p} \frac{1}{2m^*(\hat{x})} \hat{p}. \quad (1.67)$$

To arrive at the corresponding phase space function $T(p, x)$, we can make direct use of the previous result $Z(p, x) = X(p, x) \star Y(p, x)$ by replacing the operators \hat{p} and \hat{x} by their phase space functions, which are simply their eigenvalues p and x , respectively, and by linking the single terms with the star product. This results in [16]

$$\begin{aligned} T(p, x) &= p \star \frac{1}{2m^*(x)} \star p \\ &= \frac{p^2}{2m^*(x)} + \frac{\hbar^2}{8} \partial_x^2 \frac{1}{m^*(x)}. \end{aligned} \quad (1.68)$$

As one can see, the second term depends solely on x and can be seen as a correction to the potential $U(x, t)$. The evaluation of the Moyal bracket for the first term is more involved due to the simultaneous dependence on p and x . The resulting integral expressions can be found for instance in [17].

2 Wigner transport equation

After the general considerations in the last chapter, we rewrite Moyal's equation in Sec. 2.1 for the special case of a single band in the parabolic band approximation to arrive at the so-called Wigner transport equation (WTE). All calculations are done for the one-dimensional case. After a brief discussion of the appropriate boundary conditions, the pseudo-differential operator of the WTE is considered in more detail in Sec. 2.2.

2.1 Single band, parabolic approximation

To derive the Wigner transport equation for an electron in a single band and in one dimension, we start with the Moyal equation, see Eqs. (1.56), (1.57) and (1.53),

$$\begin{aligned} \partial_t f(p, x, t) &= \{\{H(p, x, t), f(p, x, t)\}\} \\ &= \frac{H(p, x, t) \star f(p, x, t) - f(p, x, t) \star H(p, x, t)}{i\hbar}, \quad x, p \in \mathbb{R}. \end{aligned} \quad (2.1)$$

The Hamiltonian is chosen to be of the form $H(p, x, t) = E(p) + qV(x, t)$, represented as the sum of the kinetic energy $E(p)$ and the product of the electrostatic potential $V(x, t)$ times the electron charge $q = -e$ with $e > 0$. In principle, either E and V could depend on both coordinates x and p . This would be the case in more complicated situations as for instance when taking a spatially dependent effective mass or magnetic fields into account. In this work we restrict ourselves to the simplest case and neglect such effects.

In [2] it is shown that in the case of a single-band, such that H and f are simply real functions and no matrices, one can rewrite the exponential functions of the star-product in terms of a single sine function:

$$\begin{aligned} \{\{H, f\}\} &= \frac{1}{i\hbar} \left\{ H \exp \left[\frac{i\hbar}{2} \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) \right] f - f \exp \left[\frac{i\hbar}{2} \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) \right] H \right\} \\ &= \frac{1}{i\hbar} H \left\{ \exp \left[\frac{i\hbar}{2} \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) \right] - \exp \left[-\frac{i\hbar}{2} \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) \right] \right\} f \\ &= \frac{2}{\hbar} H \sin \left[\frac{\hbar}{2} \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_p - \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) \right] f. \end{aligned} \quad (2.2)$$

From this equation it is evident that the action of the Moyal bracket on $f(p, x, t)$ is purely real. For the special choice of the Hamiltonian $H(p, x, t) = E(p) + qV(x, t)$ the

equation can be simplified further to

$$\{\{H, f\}\} = \frac{2}{\hbar} \left[E(p) \sin \left(-\frac{\hbar}{2} \overleftarrow{\partial}_p \overrightarrow{\partial}_x \right) f + qV(x, t) \sin \left(\frac{\hbar}{2} \overleftarrow{\partial}_x \overrightarrow{\partial}_p \right) f \right]. \quad (2.3)$$

The term containing the potential will be treated later. For the kinetic term we perform a gradient expansion, i.e. use the Taylor-series of the sine function to examine the leading-order terms in more detail

$$\{\{E, f\}\} = -\frac{\partial E}{\partial p} \frac{\partial f}{\partial x} + \frac{\hbar^2}{24} \frac{\partial^3 E}{\partial p^3} \frac{\partial^3 f}{\partial x^3} + \mathcal{O}(\hbar^5). \quad (2.4)$$

The zeroth-order contribution in Eq. (2.4) with respect to \hbar is $-\partial_p E \partial_x f$, corresponding to an advection of $f(p, x, t)$ with propagation speed $\partial_p E(p)$. The higher-order terms describe possible additional effects, e.g. dispersion in the case of the second-order term with respect to \hbar , compare e.g. [18, Sec. 11.1.2]. In the case of a parabolic dispersion relation $E(p)$, as for instance for free electrons, all higher derivatives of $E(p)$ vanish such that the Moyal bracket reduces to the advection term only, with the propagation speed given by $\partial_p E = p/m$. But is this still valid if the band structure is only approximated by a parabolic dispersion relation for states close to the conduction band edge? In the case of the semi-classical Boltzmann transport equation this question is easy to be responded to since no non-local terms appear in the equation. Only the slope of the dispersion relation at a certain momentum p in the phase space contributes to the temporal evolution of the distribution $f(p, x, t)$ at this momentum. Due to the non-local nature of the Moyal equation it is not clear if it suffices to consider only the advection term. In principle, the whole band structure influences the time evolution of the distribution function at a certain phase space element.

To have a closer look at this, we consider the simple case of a cosine-shaped band structure

$$E(p) = E_C + \frac{E_{BW}}{2} \left[1 - \cos \left(\frac{a}{\hbar} p \right) \right] \quad (2.5)$$

with E_C being the energy of the conduction band edge, E_{BW} the band width and a the lattice constant. For small values of p , such that $\frac{a}{\hbar} p \ll 1$, we retrieve the parabolic band approximation and can consider the prefactors as an effective mass m^* :

$$E(p) \approx E_C + \frac{1}{2} \left[\frac{E_{BW}}{2} \left(\frac{a}{\hbar} \right)^2 \right] p^2 = E_C + \frac{p^2}{2m^*} \quad (2.6)$$

Inserting the cosine-shaped band structure into Eq. (2.4) and considering the limit of small p , we obtain for the zeroth-order term

$$\begin{aligned} \partial_p E(p) &= \frac{E_{BW}}{2} \frac{a}{\hbar} \sin \left(\frac{a}{\hbar} p \right) \\ &= \frac{E_{BW}}{2} \left(\frac{a}{\hbar} \right)^2 p \left[1 - \frac{1}{3!} \left(\frac{a}{\hbar} p \right)^2 + \mathcal{O} \left(\left(\frac{a}{\hbar} p \right)^4 \right) \right] \\ &\approx \frac{p}{m^*}. \end{aligned} \quad (2.7)$$

One can see, that no additional assumptions must be made and as soon as the parabolic approximation well describes the real band structure also the ordinary advection is a good approximation, at least for the zeroth-order term.

To see if the terms of higher order give any significant contribution, we consider the second-order term of Eq. (2.4):

$$\begin{aligned}
 \frac{\hbar^2}{24} \partial_{ppp} E(p) &= -\frac{a\hbar E_{BW}}{24} \frac{(a)^2}{2} \sin\left(\frac{a}{\hbar} p\right) \\
 &= -\frac{a^2}{24} \frac{p}{m^*} \left[1 - \frac{1}{3!} \left(\frac{a}{\hbar} p\right)^2 + \mathcal{O}\left(\left(\frac{a}{\hbar} p\right)^4\right) \right] \\
 &\approx -\frac{a^2}{24} \frac{p}{m^*}.
 \end{aligned} \tag{2.8}$$

The additional prefactor is in general very small, namely of $\mathcal{O}(10^{-20} \text{ m}^2)$. Only in the case of a rapidly varying $f(p, x, t)$ with respect to the x -variable one could expect a significant contribution, since $\partial_{xxx} f(p, x, t)$ could take on very large values. But even if $f(p, x, t)$ would vary on the length scale of nanometers, the action of the second-order term should be at least of $\mathcal{O}(100)$ smaller than that of the zeroth-order term. Therefore, the higher-order terms can be ignored and the time evolution of electrons in a cosine-shaped band structure around E_C is well described by the ordinary advection term. Since the dispersion relation $E(p)$ is in general a very smooth function, so that a Taylor expansion around the band minimum converges rapidly, considering solely the advection term seems to be a valid procedure also for more complicated band structures.

The potential term of the Hamiltonian is more complicated to handle and requires to consider the full Moyal bracket $\{\{qV(x, t), f(p, x, t)\}\}$ in order to model the quantum transport in tunneling devices appropriately. For this purpose, we make use of the integral expression of the Moyal bracket given in Eq. (1.66) and set $U(x, t) = qV(x, t)$. We follow the notation used in [1] and denote the integral expression

$$\begin{aligned}
 (\Theta_{\hbar}[V]f)(p, x, t) &= \frac{im^*}{2\pi\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{[V(x + \frac{\eta'}{2}, t) - V(x - \frac{\eta'}{2}, t)]}{\hbar} f(p', x, t) \\
 &\quad \times \exp\left(i\frac{p-p'}{\hbar}\eta'\right) dp' d\eta'
 \end{aligned} \tag{2.9}$$

as pseudo-differential operator. With this the WTE for an arbitrary potential $qV(x, t)$, obtained from Eq. (2.1), can be written as

$$\partial_t f(p, x, t) = -\frac{p}{m^*} \partial_x f(p, x, t) + \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t). \tag{2.10}$$

In the following we will denote the first and the second term on the right-hand side as advection and drift term, respectively. The advection term is completely analogous to the semi-classical BTE but the drift term is more involved. In the classical case the drift term is a local expression and constitutes solely of a force $F(x, t)$ times the

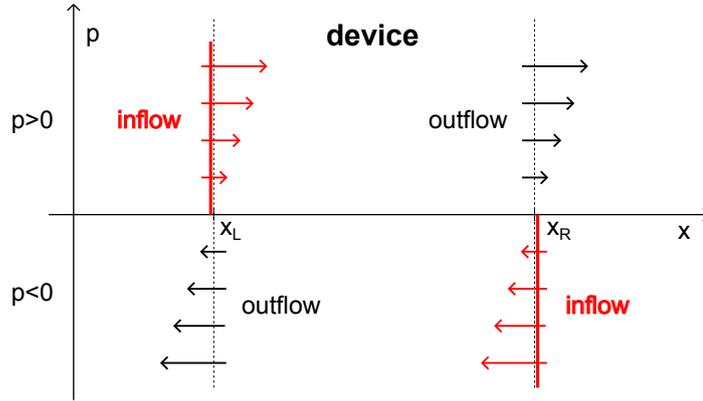


Figure 2.1: Phase space illustration of the inflow and outflow boundary conditions. The red regions represent the arguments of the momentum distribution inside the contacts which are then advected into the device. The lengths of the arrows indicate the propagation speeds in the different phase space regions.

momentum derivative of the distribution function $\partial_p f(p, x, t)$ [1]. In the case of the WTE, the whole potential $V(x', t)$ with $x' \in \mathbb{R}$ influences the time evolution of the quasi-distribution function $f(p, x, t)$ at a certain position x .

As mentioned in the introduction, an important aspect of using the Wigner function for describing quantum transport is the possibility to model contacts in a convenient way. In the trivial case that the potential is constant everywhere the action of the pseudo-differential operator vanishes and the time evolution of $f(p, x, t)$ is completely determined by advection. The analytical solution of Eq. (2.10) for the case $(\Theta_h[V]f)(p, x, t) = 0$ is given by

$$f(p, x, t) = f\left(p, x - \frac{p}{m^*}(t - t'), t'\right). \quad (2.11)$$

Therefore, the parts of $f(p, x, t)$ in phase space with $p > 0$ travel from smaller to larger values of x and for $p < 0$ the situation is vice versa. Suppose that one is interested in reducing the considered x domain to a finite region by introducing contacts at the positions x_L and x_R , with $x_L < x_R$. The appropriate boundary conditions are then to specify the inflow from the contacts at x_L for $p > 0$ and at x_R for $p < 0$, see also Fig. 2.1. The phase space regions at x_L for $p < 0$ and at x_R for $p > 0$ correspond to the outflow of $f(p, x, t)$. It should be noted that scattering processes are not considered in this work and the outflow is the only mechanism that enables a relaxation of the solution in the course of time. Therefore, for fixed boundary conditions and given an arbitrary initial condition, $f(p, x, t)$ will converge to a stationary distribution for $t \rightarrow \infty$, completely determined by the particular boundary conditions. It is worth noting that the specification of one boundary condition for each value of p is in accordance with the order of the derivative with respect to x in Eq. (2.10) [4].

Now, to model the more interesting situation with $(\Theta_h[V]f)(p, x, t) \neq 0$, one usually applies the same boundary conditions, preferably in regions where the action of the

pseudo-differential operator becomes negligible. This treatment is not exact, of course, and one introduces an error in the simulations when considering a too small x domain. This point will be considered in more detail in Chap. 5 when examining the numerical results. Despite of this source of error, it is important to note that the inflow and outflow boundary conditions enable to clearly distinguish the incoming and outgoing parts of the Wigner distribution and thus to appropriately model contacts.

2.2 Properties of the pseudo-differential operator

The pseudo-differential operator $\Theta_h[V]$ introduced in Eq. (2.9) can also be expressed in the following form

$$\begin{aligned} (\Theta_h[V]f)(p, x, t) &= \frac{im^*}{2\pi\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right]}{\hbar} f(p', x, t) \\ &\quad \times \exp\left(i\frac{p-p'}{\hbar}\eta'\right) dp' d\eta' \\ &= im^* \int_{-\infty}^{\infty} \frac{\left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right]}{\hbar} \tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta', \end{aligned} \quad (2.12)$$

where $\tilde{f}(\eta', x, t)$ labels the Fourier transform of $f(p', x, t)$ and is given by

$$\tilde{f}(\eta', x, t) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} f(p', x, t) \exp\left(-i\frac{p'}{\hbar}\eta'\right) dp'. \quad (2.13)$$

To illustrate the analogy with a purely differential operator we consider briefly the classical operator $\mathcal{L}_p = -F\partial_p$ with a force F , for which we find

$$\begin{aligned} (\mathcal{L}_p f)(p, x, t) &= -F\partial_p f(p, x, t) \\ &= -F\partial_p \int_{-\infty}^{\infty} \tilde{f}(\eta, x, t) \exp\left(i\frac{p}{\hbar}\eta\right) d\eta \\ &= -F \int_{-\infty}^{\infty} i\frac{\eta}{\hbar} \tilde{f}(\eta, x, t) \exp\left(i\frac{p}{\hbar}\eta\right) d\eta, \end{aligned} \quad (2.14)$$

when making use of the inverse expression of Eq. (2.13). Therefore, the differential operator \mathcal{L}_p acts in Fourier transformed space as a simple multiplication on $\tilde{f}(\eta, x, t)$. This is very similar to the result obtained in Eq. (2.12), making it reasonable to introduce the notation pseudo-differential operator. The multiplier

$$(\delta V)_h(\eta', x, t) = \frac{im^*}{\hbar} \left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right] \quad (2.15)$$

is called the symbol of the pseudo-differential operator [1].

As can be seen from Eq. (2.12), the potential acts in a non-local way on the Wigner function $f(p, x, t)$. Since in general $f(\eta', x, t) \rightarrow 0$ for $\eta' \rightarrow \pm\infty$ (corresponding to $f(p, x, t)$ having a finite width with respect to the p -variable), the influence of the potential on the Wigner function decays with increasing distance from the actual position x . This is expected, of course, and is sometimes used in practical computations to introduce a finite correlation length r_C corresponding to a cut-off in the integrals [4]. The non-locality is typical for a quantum mechanical description and, from a numerical point of view, makes actual computations much more complicated than in the case of the semi-classical BTE. The question on how to treat the WTE numerically is discussed in more detail in the following section.

In Sec. 1.1 it was shown how to arrive at measurable quantities like the particle density $n(x, t)$ and the current density $j(x, t)$, see Eq. (1.36), by calculating the first moments of $f(p, x, t)$ with respect to p . One fundamental property of the WTE is that the continuity equation can be retrieved by integrating it with respect to p . To show this, we consider at first only the pseudo-differential operator Eq. (2.12):

$$\begin{aligned}
 \int_{-\infty}^{\infty} (\Theta_{\hbar}[V]f)(p, x, t) dp &= \frac{im^*}{2\pi\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{[V(x + \frac{\eta'}{2}, t) - V(x - \frac{\eta'}{2}, t)]}{\hbar} f(p', x, t) \\
 &\quad \times \exp\left(-i\frac{p'}{\hbar}\eta'\right) \left(\int_{-\infty}^{\infty} \exp\left(i\frac{p}{\hbar}\eta'\right) dp\right) dp' d\eta' \\
 &= \frac{im^*}{2\pi\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{[V(x + \frac{\eta'}{2}, t) - V(x - \frac{\eta'}{2}, t)]}{\hbar} f(p', x, t) \\
 &\quad \times \exp\left(-i\frac{p'}{\hbar}\eta'\right) 2\pi\hbar\delta(\eta') dp' d\eta' \\
 &= \frac{im^*}{\hbar} \left[\lim_{u' \rightarrow x^+} V(u', t) - \lim_{u' \rightarrow x^-} V(u', t) \right] \int_{-\infty}^{\infty} f(p', x, t) dp'.
 \end{aligned} \tag{2.16}$$

In the case of a continuous potential $V(x, t)$, it is obvious that the potential term in Eq. (2.16) vanishes. But also if the potential has a finite number of steps the contribution can be neglected. If we assume that $V(x, t)$ has N discontinuities at the positions x_n and each of them having a finite difference between the left and right limit, we can write Eq. (2.16) as a sum

$$\int_{-\infty}^{\infty} (\Theta_{\hbar}[V]f)(p, x, t) dp = \sum_{n=1}^N c_n \Theta(x = x_n), \tag{2.17}$$

where each of the coefficients c_n takes on finite values. The function $\Theta(x = x_n)$ stands for

$$\Theta(x = x_n) = \begin{cases} 1 & \text{if } x = x_n \\ 0 & \text{elsewhere} \end{cases}. \tag{2.18}$$

As a result, when integrating the whole WTE (2.10) with respect to p , we arrive at

$$\partial_t n(x, t) + \frac{1}{q} \partial_x j(x, t) = \frac{q}{m^*} \sum_{n=1}^N c_n \Theta(x = x_n), \quad (2.19)$$

with

$$n(x, t) = \int_{-\infty}^{\infty} f(p, x, t) dp,$$

$$j(x, t) = q \int_{-\infty}^{\infty} \frac{p}{m^*} f(p, x, t) dp.$$

Eq. (2.19) seems to contain an additional source term on the right-hand side but in the sense of the Lebesgue measure it is obvious that this term does not give any contribution as soon as one integrates the equation over an arbitrary interval $C_j = [x_{j-1/2}, x_{j+1/2}]$. It is reasonable to deal with an integral representation since $n(x, t)$ represents the particle density at point x which is not directly accessible and one can only measure the number of particles contained in a certain interval. When we introduce

$$\bar{n}(x_j, t) = \frac{1}{\Delta x_j} \int_{C_j} n(x, t) dx, \quad (2.20)$$

we can write the integral form of Eq. (2.19) as

$$q \partial_t \bar{n}(x_j, t) + \frac{1}{\Delta x_j} [j(x_{j+1/2}, t) - j(x_{j-1/2}, t)] = 0, \quad (2.21)$$

which is exactly the well-known continuity equation. It is, therefore, valid to say that any solution of the WTE (2.10) inherently fulfills the continuity equation as well, which may be written in differential form as

$$q \partial_t n(x, t) + \partial_x j(x, t) = 0. \quad (2.22)$$

As an additional note, it is known from Schrödinger's formulation of quantum mechanics that even infinitely high steps or singularities in the potential $V(x, t)$ can be considered. To be consistent with the continuity equation one has to take into account that the current is not well defined at such points because the wave function is allowed to have kinks at singular points of $V(x, t)$, see e.g. [12]. Since, in this work, we will restrict ourselves to cases where $V(x, t)$ stays finite, this question is not treated in more detail.

For the later use in Chap. 3 we show that the outcome of the operator acting on the distribution function, $(\Theta_{\hbar}[V]f)(p, x, t)$, is real. At first, consider the Fourier transform of the Wigner function, defined by Eq. (2.13). Since $f(p, x, t)$ is real it is obvious that

$$\tilde{f}(-\eta, x, t) = \left(\tilde{f}(\eta, x, t) \right)^*. \quad (2.23)$$

The same holds true for the symbol of the pseudo-differential operator

$$(\delta V)_\hbar(-\eta, x, t) = ((\delta V)_\hbar(\eta, x, t))^* , \quad (2.24)$$

as can be seen from Eq. (2.15). Combining these two equations yields

$$(\delta V)_\hbar(-\eta, x, t)\tilde{f}(-\eta, x, t) = \left((\delta V)_\hbar(\eta, x, t)\tilde{f}(\eta, x, t) \right)^* , \quad (2.25)$$

so that the pseudo-differential operator can be rewritten as

$$\begin{aligned} (\Theta_\hbar[V]f)(p, x, t) &= \int_{-\infty}^{\infty} (\delta V)_\hbar(\eta', x, t)\tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \\ &= \int_0^{\infty} (\delta V)_\hbar(\eta', x, t)\tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \\ &\quad - \int_{\infty}^0 (\delta V)_\hbar(-\eta', x, t)\tilde{f}(-\eta', x, t) \exp\left(-i\frac{p}{\hbar}\eta'\right) d\eta' \\ &= 2\Re \left\{ \int_0^{\infty} (\delta V)_\hbar(\eta', x, t)\tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \right\} . \end{aligned} \quad (2.26)$$

3 Numerical methods for the Wigner transport equation

Before focusing on the numerical schemes developed in this work, Sec. 3.1 gives a quick overview of other methods used in literature. In the common approach, a discrete Fourier transformation is applied to calculate the action of the pseudo-differential operator and the most important aspects of one method of this kind are listed as example. For this purpose and also for later use in this work, the concept of operator splitting is introduced. The constraints associated with making use of a discrete Fourier transformation are discussed to highlight the central problems and to motivate the introduction of the new algorithms. The new schemes developed in this thesis are presented in detail in Sec. 3.2. Section 3.3 deals with different numerical schemes to treat the advection term in the Wigner transport equation. The reconstruct-evolve-average (REA) algorithm is outlined and the simple upwind as well as slope limiter methods are briefly discussed. The WENO5 scheme, a weighted essentially non oscillatory scheme with an accuracy up to fifth order and which was used for the actual simulations, is introduced thereafter. This scheme exhibits the peculiarity to be unstable in combination with a forward Euler method and an appropriate time stepper is therefore given as well. Finally, possible time steppers for the drift term, i.e. the pseudo-differential operator, are discussed as well.

3.1 Methods based on the application of discrete Fourier transformation

In his pioneering work [4], Frensley devised a discretization of the WTE based on a discrete Fourier transformation to evaluate the potential term. Instead of Fourier transforming the distribution function $f(p', x, t)$ as in Eq. (2.12), his approach consists of calculating the Fourier transform $(\delta\tilde{V})_h(p - p', x, t)$ of the symbol of the pseudo-differential operator. When introducing

$$(\delta\tilde{V})_h(p - p', x, t) = \int_{-\infty}^{\infty} (\delta V)_h(\eta', x, t) \exp\left(i\frac{p - p'}{\hbar}\eta'\right) d\eta', \quad (3.1)$$

the pseudo-differential operator, Eq. (2.12), may be rewritten as

$$\begin{aligned} (\Theta_h[V]f)(p, x, t) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\delta V)_h(\eta', x, t) f(p', x, t) \exp\left(i\frac{p-p'}{\hbar}\eta'\right) dp' d\eta' \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} (\delta\tilde{V})_h(p-p', x, t) f(p', x, t) dp', \end{aligned} \quad (3.2)$$

resulting in a convolution integral with respect to p' . Upon discretization, the integral in Eq. (3.2) is replaced by a summation and one ends up with a matrix expression for the action of the pseudo-differential operator. To evaluate Eq. (3.1) for different values of $(p-p')$ a discrete Fourier transformation is used. To discretize the advection term of the WTE, Frensley used a conventional upwind scheme.

With this scheme, Frensley demonstrated for the case of a resonant tunneling diode (RTD) that the negative differential resistance region in the $I(V)$ curve could be reproduced and not only stationary but fully time-dependent simulations could be achieved. Despite of this success, quantitative comparison of the obtained $I(V)$ curve with reference solutions showed differences of several tens of percents. Subsequent investigations by other groups could improve the results to some degree by using higher-order approximations for the advection term, see e.g. [5], [19].

A recently developed scheme based on an adaptive spectral element method [20] produces very accurate results for the case of a Gauss wave packet scattering at a Gaussian potential barrier. The method uses a Gauss-Chebyshev interpolation for the p -variable and Gauss-Lobatto collocation points for the x -variable combined with a fast Fourier transform. Since the authors consider only smooth potential shapes and rather short time evolutions (compared to what is needed to reach an approximately stationary state), it is hard to assess how well the algorithm would perform when simulating RTDs.

Another approach based on a combination of operator splitting and discrete Fourier transformation was proposed by Frosali and Morandi (see e.g. [21], [22]). On the next pages the method will be described in more detail after giving a general introduction to operator splitting. Splitting methods are widely applicable and are also included in the algorithms developed in the following sections. Furthermore, Morandi's method was actually used in this thesis to perform the first simulations and served as a reference for the first results obtained with the new algorithms - for not too sharp potential shapes at least.

3.1.1 Operator Splitting

The technique of operator splitting is applicable to various problems. Consider an operator \mathcal{L} acting on $f(p, x, t)$ which may be separated into two parts, so that

$$\begin{aligned} \partial_t f &= \mathcal{L}f \\ &= \mathcal{L}_1 f + \mathcal{L}_2 f. \end{aligned} \quad (3.3)$$

In the following we consider only time-independent operators which do not act on the t -variable, enabling it to write the formal solution to Eq. (3.3) as

$$\begin{aligned} f(t) &= \exp(\mathcal{L}t) f(t_0) \\ &= \exp((\mathcal{L}_1 + \mathcal{L}_2)t) f(t_0) \\ &= \exp(\mathcal{L}_1 t) \exp(\mathcal{L}_2 t) f(t_0) + \mathcal{O}([\mathcal{L}_1 t, \mathcal{L}_2 t]) f(t_0). \end{aligned} \quad (3.4)$$

All the calculations outlined here can be extended in a straight-forward manner to time-dependent operators by replacing each term $\mathcal{L}_i t$ by $\int_{t_0}^t \mathcal{L}_i(\tau) d\tau$, with the effect that the single terms in the derivations grow in size but the principle results obtained are the same. All exponentials of operators are defined via a series expansion and the error term in the last line can be found when comparing the following two expansions:

$$\exp((\mathcal{L}_1 + \mathcal{L}_2)t) = 1 + (\mathcal{L}_1 + \mathcal{L}_2)t + \frac{(\mathcal{L}_1 + \mathcal{L}_2)^2}{2} t^2 + \mathcal{O}(t^3) \quad (3.5)$$

and

$$\begin{aligned} \exp(\mathcal{L}_1 t) \exp(\mathcal{L}_2 t) &= \left[1 + \mathcal{L}_1 t + \frac{\mathcal{L}_1^2}{2} t^2 + \mathcal{O}(t^3) \right] \left[1 + \mathcal{L}_2 t + \frac{\mathcal{L}_2^2}{2} t^2 + \mathcal{O}(t^3) \right] \\ &= 1 + (\mathcal{L}_1 + \mathcal{L}_2)t + \frac{1}{2} (\mathcal{L}_1^2 + \mathcal{L}_2^2) t^2 + \mathcal{L}_1 \mathcal{L}_2 t^2 + \mathcal{O}(t^3) \\ &= 1 + (\mathcal{L}_1 + \mathcal{L}_2)t + \frac{(\mathcal{L}_1 + \mathcal{L}_2)^2}{2} t^2 + \frac{1}{2} (\mathcal{L}_1 \mathcal{L}_2 - \mathcal{L}_2 \mathcal{L}_1) t^2 + \mathcal{O}(t^3) \end{aligned} \quad (3.6)$$

This implies that if \mathcal{L}_1 and \mathcal{L}_2 commute, the two expansions give the same result and no error is made when applying the operators in Eq. (3.4) sequentially, one after the other on $f(t_0)$. Otherwise, one encounters an error proportional to the commutator $[\mathcal{L}_1, \mathcal{L}_2]$ and t^2 . This is an example of a first-order splitting, called Lie-Trotter splitting [23].

Higher-order methods can be derived by splitting up the exponential in line two of Eq. (3.4) in a different way. A commonly used method of second-order is Strang splitting [23], [18], which can be written as

$$\begin{aligned} f(t) &= \exp((\mathcal{L}_1 + \mathcal{L}_2)t) f(t_0) \\ &= \exp\left(\mathcal{L}_1 \frac{t}{2}\right) \exp(\mathcal{L}_2 t) \exp\left(\mathcal{L}_1 \frac{t}{2}\right) f(t_0) + \mathcal{O}(t^3) f(t_0), \end{aligned} \quad (3.7)$$

when using the formal solution again.

In practice, a numerical or even analytical method is used, which is specialized on the particular operator to calculate the time evolution of $f(t)$. Let L_1 and L_2 stand for the matrices obtained by discretization of the operators \mathcal{L}_1 and \mathcal{L}_2 , respectively. In the case of Lie-Trotter splitting and a simple forward Euler time stepping scheme, $f(t_n)$ is advanced to $f(t_{n+1})$ by two steps:

$$\begin{aligned} f^{n+1/2} &= (\mathbb{1} + \Delta t L_1) f^n, \\ f^{n+1} &= (\mathbb{1} + \Delta t L_2) f^{n+1/2}. \end{aligned} \quad (3.8)$$

In the case of Strang splitting three steps are needed to evolve the function in question by Δt . Of course, an overall second-order accuracy can only be achieved if all methods for the sub-steps are at least second-order accurate in time. [18]

To see if a splitting error occurs for the advection operator

$$\mathcal{L}_A f(p, x, t) = -\frac{p}{m^*} \partial_x f(p, x, t) \quad (3.9)$$

and the drift operator

$$\mathcal{L}_D f(p, x, t) = \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) \quad (3.10)$$

in the WTE, we evaluate their commutator

$$\begin{aligned} [\mathcal{L}_A, \mathcal{L}_D] f &= -\frac{pq}{(m^*)^2} [\partial_x, \Theta_h[V]] f \\ &= -\frac{pq}{(m^*)^2} (\Theta_h[\partial_x V]f + \Theta_h[V]\partial_x f - \Theta_h[V]\partial_x f) \\ &= -\frac{pq}{(m^*)^2} \Theta_h[\partial_x V]f. \end{aligned} \quad (3.11)$$

The action of the pseudo-differential operator $\Theta_h[V]f$ vanishes for the case $V(x) = \text{const.}$, compare Eq. (2.12). One thus finds that no splitting error is encountered in the case of a constant electric field $\partial_x V(x) = \text{const.}$.

3.1.2 Discrete Fourier Transformation

As seen in the preceding section, the Fourier transformed operator $\mathcal{F}\{\Theta_h[V]f\}(\eta, x, t)$ acts as a simple multiplier on $\tilde{f}(\eta, x, t)$. This means that one is able to calculate the action of the potential by Fourier transforming $f(p, x, t)$, evaluating the multiplication in the (η, x) -space and inverse Fourier transforming it to the (p, x) -space. Additionally, one has to use an operator splitting scheme to decouple the drift and the advection term, so that one has to consider instead of the full WTE

$$\partial_t f(p, x, t) = (\mathcal{L}_A + \mathcal{L}_D)f(p, x, t) \quad (3.12)$$

only the pseudo-differential operator of the drift term in one sub-step $t_n \rightarrow t_{n+1/2}$:

$$\partial_t f(p, x, t) = \frac{q}{m^*} (\Theta_h[V]f)(p, x, t). \quad (3.13)$$

If we insert for the right hand side the expression found in Eq. (2.12) and apply a Fourier transformation we arrive at

$$\begin{aligned}
 \partial_t \tilde{f}(\eta, x, t) &= \frac{iq}{2\pi\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{[V(x + \frac{\eta'}{2}, t) - V(x - \frac{\eta'}{2}, t)]}{\hbar} \tilde{f}(\eta', x, t) \\
 &\quad \times \exp\left[i\frac{p}{\hbar}(\eta' - \eta)\right] d\eta' dp \\
 &= \frac{iq}{2\pi\hbar} \int_{-\infty}^{\infty} \frac{[V(x + \frac{\eta'}{2}, t) - V(x - \frac{\eta'}{2}, t)]}{\hbar} \tilde{f}(\eta', x, t) 2\pi\hbar \delta(\eta' - \eta) d\eta' \\
 &= \frac{iq}{\hbar} [V(x + \frac{\eta}{2}, t) - V(x - \frac{\eta}{2}, t)] \tilde{f}(\eta, x, t). \tag{3.14}
 \end{aligned}$$

The multiplier on the right-hand side of this equation is again, except for the prefactor m^*/q , the symbol of the pseudo-differential operator, Eq. (2.15). One can solve Eq. (3.14) by some standard ODE solver or also analytically, when the potential is static or at least considered to be constant during each time step Δt by

$$\tilde{f}(\eta, x, t_{n+1/2}) = \tilde{f}(\eta, x, t_n) \exp\left(\frac{iq}{\hbar} [V(x + \frac{\eta}{2}, t_n) - V(x - \frac{\eta}{2}, t_n)] \Delta t\right). \tag{3.15}$$

After applying the inverse Fourier transformation the solution $f(p, x, t_{n+1/2})$ is obtained and one sub-step of the time evolution complete (advection being the other one).

In practice one cannot calculate the continuous Fourier transform $\tilde{f}(\eta, x, t)$ exactly. Usually one restricts the calculations to a finite set of p_i values and then applies a discrete Fourier transformation on the pointwise given function $f(p_i, x, t)$ to arrive at the discrete function $\tilde{f}(\eta_k, x, t)$. Well developed algorithms exist to calculate the discrete Fourier transformation, called fast Fourier transform (FFT). Details on FFT can be found in various books, e.g. [24] and are not recapped here, only the following fundamental connection between the p_i and the η_k grid.

For the FFT routines to be applicable, one needs an equidistant grid of N_p values so that

$$p_i = p_1 + (i - 1)\Delta p, \quad i = 1, \dots, N_p, \quad \Delta p \in \mathbb{R}, \tag{3.16}$$

where the integer variable N_p has to be a power of two (also other FFT variants exist but each of them only for special values of N_p). The values of η_k may then be written as [24]

$$\eta_k = \frac{2\pi\hbar}{N_p\Delta p} k, \quad k = -\frac{N_p}{2} + 1, \dots, \frac{N_p}{2}. \tag{3.17}$$

As we can see, the spacings of the two grids are inversely proportional:

$$\Delta p \Delta \eta = \frac{2\pi\hbar}{N_p}. \tag{3.18}$$

Frenshley [4] showed that with this choice of $\Delta\eta$, Δp and N_p the particle density conserving property of the pseudo-differential operator is preserved so that the action of the potential vanishes when summing the discretized pseudo-differential operator over all values of p_i . As pointed out in Sec. 2.2, this property is needed to be consistent with the continuity equation.

Practical simulations with this algorithm showed good results for not too sharp potential shapes, enabling to resolve all features and interference patterns of the Wigner function. But as soon as thinner potential barriers were considered, in order to simulate tunneling phenomena, serious problems were encountered. It seems to be a highly demanding task to perform simulations of resonant tunneling diodes with this method. This is particularly due to the constraints listed in the following.

- The grid for the p variable has to be chosen strictly equidistantly. One immediately notices in practice that the Wigner function is by all means not as smooth and well behaved as the Boltzmann probability distribution. As soon as interesting quantum phenomena like interferences occur, $f(p, x, t)$ forms heavily oscillating patterns regarding the p variable, which may be in the case of tunneling diodes or RTDs some orders of magnitude smaller than the extension of $f(p, x, t)$ with respect to p . The oscillations are especially located around $p = 0$ (in the case of small bias voltages), whereas the distribution usually stays smooth for larger values of p . Therefore, it would be desirable to use a non-equidistant grid, enabling to choose the grid spacing around $p = 0$ fine enough to resolve the oscillations at all but still maintaining a reasonable number of total grid points ($N_p \approx 1000$) by increasing the spacing for larger p values.
- Sharp potentials are problematic to resolve. A perfect square potential would always produce errors in the solution due to Gibb's phenomenon [25] but even if one argues that no perfect square potentials are present in nature as well and one considers only smooth potential shapes, a very high η resolution would be needed as soon as interesting cases, e.g. tunneling devices, are treated.
- The coupling of the x and p variables of the grid causes the following difficulties. Since in practical computations one usually needs to choose $\Delta\eta = 2\Delta x$, Eq. (3.18) imposes constraints on the x and p variable. This results in a bad scaling of the computation time as the resolution is increased. For instance, to double the resolution in the position variable one needs, on the one hand, twice as many x points, of course, but on the other hand, Eq. (3.18) requires to double $N_p\Delta p$ as well. In order to maintain the chosen value of Δp it is, therefore, needed to choose N_p twice as large, yielding a total factor of four for the grid points. [20]

3.2 Discretization of the pseudo-differential operator based on a piecewise polynomial approximation of the Wigner function

In order to allow for a non-equidistant grid spacing in the p variable, numerical methods using a piecewise approximation of the Wigner function are developed in the following. Therefore, $f(p, x, t)$ is written as a sum of polynomials with each of them defined locally on one grid cell and being zero elsewhere. This is very similar to what is done in finite element methods or in the REA-algorithm for finite volume schemes, as outlined in Sec. 3.2.3 in more detail. The approach results in a matrix expression for the action of the pseudo-differential operator. Since the matrix elements involve oscillatory integrals, a piecewise polynomial approximation is chosen for the potential shape as well and the integrals are evaluated analytically - as far as possible at least. In order to achieve the conservation of the particle density, i.e. to be consistent with the continuity equation, the equations are averaged over each grid cell. This results in more complicated expressions for the matrix elements but enables us to conserve the particle density for arbitrarily chosen grid spacings for the p variable.

The outlined approach of a piecewise approximation of the Wigner function was used in both of the developed algorithms. At the beginning it seemed reasonable to demand a continuous approximation for $f(p, x, t)$ but this constraint was dropped for the second developed algorithm after considering the square barrier as a first benchmark problem. On the one hand, the analytical solution itself, see Sec. ??, revealed that the Wigner function exhibits a discontinuity and singularity at $p = 0$, making it obsolete to demand a continuous solution. On the other hand, a continuous approximation in combination with a cell averaged formulation seemed to be disadvantageous from a numerical point of view, since, besides increasing the computational time due to a matrix expression not only for the drift term but for the advection term as well, the algorithm produced spurious oscillations in some cases. These oscillations were especially present in the intermediate regime, between a well resolved solution and a solution in which certain fine-scaled oscillations were averaged out. Therefore, a second algorithm without a continuous approximation and direct use of cell averaged values of $f(p, x, t)$ was developed, in close analogy to the finite volume methods for hyperbolic problems [11].

In Sec. 3.2.1 a brief description of the developed algorithm using a continuous approximation is given, by outlining the basic concepts. The equations for the matrix elements are not derived in detail since all of the final simulations of RTDs were done with the second algorithm. In Sec. 3.2.2 a central part of the algorithms is discussed, namely the evaluation of the sine and cosine integrals, appearing in all the expressions for the matrix elements of the discretized pseudo-differential operator. With this in hands, Sec. 3.2.3 focuses on the derivation of the expressions for the second algorithm. Sec. 3.2.4 concludes with numerical tests examining the accuracy of the Fourier transform of a piecewise polynomial function.

3.2.1 Continuous approximation

At first, we will calculate solely the action of the potential operator, Eq. (2.12), when considering the following representation of the distribution function:

$$f(p', x, t) = \sum_{i=0}^{N_p} P^i(p', x, t). \quad (3.19)$$

The polynomials are defined by

$$P^i(p', x, t) := \begin{cases} a_i + b_i(p' - p_i) & \text{if } p_i \leq p' < p_{i+1}, \\ 0 & \text{elsewhere,} \end{cases} \quad (3.20)$$

with

$$\begin{aligned} a_i &= f(p_i, x, t), \\ b_i &= \frac{f(p_{i+1}, x, t) - f(p_i, x, t)}{\Delta p_i}. \end{aligned} \quad (3.21)$$

The discretization of the p variable due to the grid constants $\Delta p_l > 0$ is given by

$$p_i = p_0 + \sum_{l=0}^{i-1} \Delta p_l, \quad p_0 \in \mathbb{R}, \quad i = 1, \dots, N_p + 1. \quad (3.22)$$

For the values at the endpoints of the finite grid we demand

$$\begin{aligned} f(p_0, x, t) &= 0, \\ f(p_{N_p+1}, x, t) &= 0. \end{aligned} \quad (3.23)$$

Eq. (3.23) together with Eq. (3.21) ensures a continuous interpolation function $f(p', x, t)$, which is defined for $p' \in \mathbb{R}$ and completely determined by the values $f(p_i, x, t)$, $i = 1, \dots, N_p$. See also Fig. 3.1 for an illustration. The function clearly has the necessary asymptotic behaviour, namely that $f(p', x, t) \rightarrow 0$ sufficiently fast for $p' \rightarrow \pm\infty$. The x and t coordinates will be restricted to discrete values as well but for the moment they enter just as parameters.

For evaluating the action of the potential we need the Fourier transform of Eq. (3.19),

$$\begin{aligned} \tilde{f}(\eta', x, t) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \sum_{i=0}^{N_p} P^i(p', x, t) \exp\left(-i\frac{p'}{\hbar}\eta'\right) dp' \\ &= \frac{1}{2\pi\hbar} \sum_{i=0}^{N_p} \int_{p_i}^{p_{i+1}} [a_i + b_i(p' - p_i)] \exp\left(-i\frac{p'}{\hbar}\eta'\right) dp', \end{aligned} \quad (3.24)$$

which can be split up into a sum of integrals over finite intervals. The individual integrals are straightforward to calculate and after connecting the coefficients a_i and

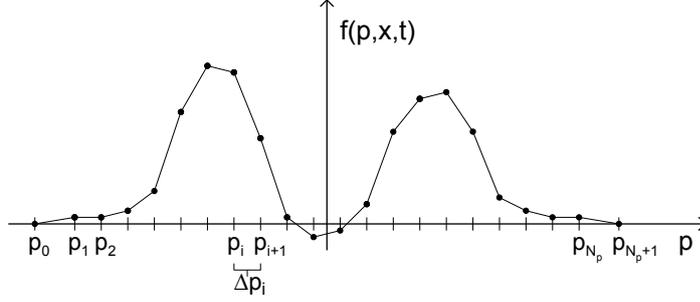


Figure 3.1: Schematic representation of the continuous piecewise linear approximation of $f(p, x, t)$ as defined in Eqs. (3.19) - (3.23).

b_i via Eq. (3.21) to the values $f(p_i, x, t)$ at the grid points and regrouping the terms, one may write the result as

$$\tilde{f}(\eta', x, t) = \sum_{i=1}^{N_p} C_i(\eta') f(p_i, x, t), \quad (3.25)$$

with the coefficients $C_i(\eta')$ given by

$$C_i(\eta') = \frac{1}{2\pi\hbar} \left(\frac{\hbar}{\eta'} \right)^2 \exp\left(-i\frac{p_i}{\hbar}\eta'\right) \times \left\{ \frac{1}{\Delta p_i} \left[1 - \exp\left(-i\frac{\Delta p_i}{\hbar}\eta'\right) \right] + \frac{1}{\Delta p_{i-1}} \left[1 - \exp\left(i\frac{\Delta p_{i-1}}{\hbar}\eta'\right) \right] \right\}. \quad (3.26)$$

Inserting Eq. (3.25) into the expression for the pseudo-differential operator, Eq. (2.12), together with evaluating the expression for certain values $p = p_m$, results in the following matrix expression

$$\frac{q}{m^*} (\Theta_{\hbar}[V]f)(p_m, x, t) = \sum_{i=1}^{N_p} D_{mi}^0(x, t) f(p_i, x, t), \quad (3.27)$$

with the matrix elements given by

$$D_{mi}^0(x, t) = \frac{iq}{\hbar} \int_{-\infty}^{\infty} \left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right] C_i(\eta') \exp\left(i\frac{p_m}{\hbar}\eta'\right) d\eta'. \quad (3.28)$$

From Eq. (3.26) it is obvious that

$$C_i(-\eta') = (C_i(\eta'))^*, \quad (3.29)$$

so that we can apply the same calculational steps as in Eq. (2.26) to obtain

$$D_{mi}^0(x, t) = 2\Re \left\{ \frac{iq}{\hbar} \int_0^{\infty} \left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right] C_i(\eta') \exp\left(i\frac{p_m}{\hbar}\eta'\right) d\eta' \right\}. \quad (3.30)$$

Using this fact we can finally write the matrix elements as

$$D_{mi}^0(x, t) = \frac{q}{\pi\hbar} \int_0^\infty \left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right] \left(\frac{1}{\eta'}\right)^2 [A_{m,i}(\eta') - A_{m,i-1}(\eta')] d\eta', \quad (3.31)$$

when defining

$$A_{m,i}(\eta') = \frac{\hbar}{\Delta p_i} \left[\sin\left(\frac{p_m - p_{i+1}}{\hbar} \eta'\right) - \sin\left(\frac{p_m - p_i}{\hbar} \eta'\right) \right]. \quad (3.32)$$

In principle, one could already use Eq. (3.31) to calculate the action of an arbitrary potential $V(x, t)$ on the Wigner function and perform simulations with that. But practical tests showed that a numerical evaluation of the integrals is a highly demanding task and seems to be unfeasible due to the poor convergence. The problem arises because of the oscillating nature of the integrand. To circumvent this problem, also the potential $V(x, t)$ is approximated by a piecewise polynomial function enabling to directly calculate the integrals. In the following we will consider only the simplest case, a single box-like barrier. Further approximation variants will only be discussed for the second algorithm in Sec. 3.2.3.

When considering a discontinuous potential, one needs to check the convergence of the integral in Eq. (3.31). In principle, problems could arise due to the factor $1/\eta'^2$ in the integrand at discontinuous positions x , since

$$\lim_{\eta' \rightarrow 0} \left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right] \quad (3.33)$$

takes on non-zero values. To have a closer look at this problem, we Taylor expand $A_{m,i}(\eta')$ up to first order,

$$\begin{aligned} A_{m,i}(\eta') &= \frac{\hbar}{\Delta p_i} \left[\frac{p_m - p_{i+1}}{\hbar} \eta' - \frac{p_m - p_i}{\hbar} \eta' \right] + \mathcal{O}(\eta'^3) \\ &= -\eta' + \mathcal{O}(\eta'^3), \end{aligned} \quad (3.34)$$

so that we find for the limit

$$\begin{aligned} \lim_{\eta' \rightarrow 0} \left(\frac{1}{\eta'}\right)^2 [A_{m,i}(\eta') - A_{m,i-1}(\eta')] &= \lim_{\eta' \rightarrow 0} \left(\frac{1}{\eta'}\right)^2 [-\eta' + \mathcal{O}(\eta'^3) + \eta' + \mathcal{O}(\eta'^3)] \\ &= \lim_{\eta' \rightarrow 0} \mathcal{O}(\eta') \\ &= 0. \end{aligned} \quad (3.35)$$

This result tells us, that the integral Eq. (3.31) exists for any bounded potential shape, independent of the particular values of the left and right sided limit of $V(x)$ at a certain point x (Eq. (3.33)). Using this fact and that $A_{m,i}(\eta')$ is an odd function,

$$A_{m,i}(-\eta') = -A_{m,i}(\eta'), \quad (3.36)$$

we can write the matrix elements also in the following form

$$D_{mi}^0(x, t) = \frac{q}{\pi\hbar} \int_{-\infty}^{\infty} V\left(x + \frac{\eta'}{2}, t\right) \left(\frac{1}{\eta'}\right)^2 [A_{m,i}(\eta') - A_{m,i-1}(\eta')] d\eta'. \quad (3.37)$$

Eq. (3.37) is equivalent to Eq. (3.31) but may be easier to evaluate for potential shapes with compact support. It is now straightforward to consider the case of a single box-shaped barrier of height V_0 and length $x_1 - x_0$ with $x_1 > x_0$, so that

$$V\left(x + \frac{\eta'}{2}\right) = V_0 \Theta[2(x_0 - x) \leq \eta' \leq 2(x_1 - x)]. \quad (3.38)$$

The Θ -function used here denotes an abbreviation for the product of two Heaviside step functions $H(x)$

$$\Theta(a \leq x \leq b) = H(x - a)H(b - x). \quad (3.39)$$

Eq. (3.37) then simplifies to

$$D_{mi}^0(x, t) = \frac{qV_0}{\pi\hbar} \int_{2(x_0-x)}^{2(x_1-x)} \left(\frac{1}{\eta'}\right)^2 [A_{m,i}(\eta') - A_{m,i-1}(\eta')] d\eta'. \quad (3.40)$$

The integrals to be solved are of the form

$$\begin{aligned} I &= \int_a^b \frac{1}{x^2} \sin(kx) dx \\ &= -\frac{1}{x} \sin(kx) \Big|_a^b + k \int_a^b \frac{1}{x} \cos(kx) dx. \end{aligned} \quad (3.41)$$

The last term is a so called cosine integral [26] [27] and cannot be solved analytically. But, libraries are available with ready-made routines to evaluate the integral at wanted values, which will be discussed in more detail in the next section.

Tests using the drift matrix of Eq. (3.40) showed promising results. For comparison, a trapezoidal barrier was implemented as well and one could reproduce results achieved with the previously presented approach based on operator splitting and FFT. Furthermore, one could also simulate tunneling currents for very sharp potentials which posed a problem for the FFT method. But, the algorithm at the present stage has the downside that the number of particles is not exactly conserved. This is discussed in the following in more detail and an approach is presented, based on integrating the equations over grid cells, to arrive at a conservative scheme. Since the integrating involves many calculational steps and the algorithm presented in this section was not used anymore for the final simulations of the RTDs, the steps will only be described schematically together with discussing crucial aspects and problems that arise.

For the moment, we only consider the drift term, i.e. the part of the WTE containing the pseudo-differential operator:

$$\partial_t f(p, x, t) = \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t). \quad (3.42)$$

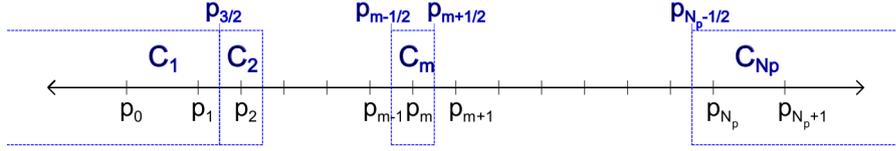


Figure 3.2: Schematic illustration of the p_m grid together with the grid cells C_m and the intermediate points $p_{m+1/2}$ for the continuous piecewise linear approximation of $f(p, x, t)$.

Up to now we simply evaluated the equation pointwise

$$\partial_t f(p_m, x, t) = \frac{q}{m^*} (\Theta_h[V]f)(p_m, x, t), \quad (3.43)$$

which directly enabled to update the discrete set of values of $f(p_m, x, t)$ when using some time stepping method as for instance in the simplest case a forward Euler scheme: $\partial_t f(p_m, x, t_n) = \frac{1}{\Delta t} [f(p_m, x, t_{n+1}) - f(p_m, x, t_n)]$. From Sec. 2.2 we know the result of integrating Eq. (3.42) with respect to p , which is for the left-hand side the rate of change of the particle density,

$$\int_{-\infty}^{\infty} \partial_t f(p, x, t) dp = \partial_t n(x, t), \quad (3.44)$$

and the integration of the right-hand side yields zero for any function $f(p, x, t)$

$$\int_{-\infty}^{\infty} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) dp = 0. \quad (3.45)$$

For the discrete form Eq. (3.43) it is straightforward to approximate the particle density by

$$n(x, t) \approx \sum_{m=1}^{N_p} f(p_m, x, t) \frac{\Delta p_m + \Delta p_{m-1}}{2}, \quad (3.46)$$

or some similar expression. But, the problem is that the pointwise evaluation of $(\Theta_h[V]f)(p, x, t)$ results in general in a nonzero value for the summation of the right-hand side of Eq. (3.43),

$$\sum_{m=1}^{N_p} \frac{q}{m^*} (\Theta_h[V]f)(p_m, x, t) \frac{\Delta p_m + \Delta p_{m-1}}{2} \neq 0. \quad (3.47)$$

As a consequence of this, the temporal change of $n(x, t)$ due to the pseudo-differential operator alone is nonzero which in turn violates the conservation of the number of particles.

Now, to arrive at a conservative scheme we use instead of the pointwise evaluation an integration over grid cells to discretize Eq. (3.42). In order to make use of the

property stated in Eq. (3.45), we need to define grid cells in such a way that the whole range of $p \in \mathbb{R}$ is covered. One possible choice is to define grid cells by

$$\begin{aligned} C_m &= (p_{m-1/2}, p_{m+1/2}), \quad m = 2, \dots, N_p - 1, \\ C_1 &= (-\infty, p_{3/2}), \\ C_{N_p} &= (p_{N_p-1/2}, \infty), \end{aligned} \quad (3.48)$$

with

$$p_{m-1/2} = \frac{p_{m-1} + p_m}{2}, \quad m = 2, \dots, N_p, \quad (3.49)$$

see also Fig. 3.2. Integrating Eq. (3.42) over each of the N_p intervals C_m yields a system of equations:

$$\int_{C_m} \partial_t f(p, x, t) dp = \int_{C_m} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) dp, \quad m = 1, \dots, N_p. \quad (3.50)$$

It is evident that a summation over all integration intervals C_m corresponds to an integration over all $p \in \mathbb{R}$:

$$\sum_{m=1}^{N_p} \int_{C_m} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) dp = \int_{-\infty}^{\infty} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) dp. \quad (3.51)$$

If we now insert into Eq. (3.27) the derived expression for the drift matrix Eq. (3.37) without the restriction of p to discrete values p_m , we obtain

$$\begin{aligned} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) &= \frac{q}{\pi \hbar} \sum_{i=1}^{N_p} f(p_i, x, t) \\ &\quad \times \int_{-\infty}^{\infty} V\left(x + \frac{\eta'}{2}, t\right) \left(\frac{1}{\eta'}\right)^2 [A_i(p, \eta') - A_{i-1}(p, \eta')] d\eta', \end{aligned} \quad (3.52)$$

with

$$A_i(p, \eta') = \frac{\hbar}{\Delta p_i} \left[\sin\left(\frac{p - p_{i+1}}{\hbar} \eta'\right) - \sin\left(\frac{p - p_i}{\hbar} \eta'\right) \right]. \quad (3.53)$$

This allows us to rewrite Eq. (3.50) in the form

$$\int_{C_m} \partial_t f(p, x, t) dp = \sum_{i=1}^{N_p} D'_{m,i} f(p_i, x, t), \quad m = 1, \dots, N_p, \quad (3.54)$$

by defining new, cell-integrated drift matrix elements:

$$\begin{aligned} D'_{m,i}(x, t) &= \frac{q}{\pi \hbar} \int_{-\infty}^{\infty} V\left(x + \frac{\eta'}{2}, t\right) \left(\frac{1}{\eta'}\right)^2 \int_{C_m} [A_i(p, \eta') - A_{i-1}(p, \eta')] dp d\eta', \\ &\quad m, i = 1, \dots, N_p. \end{aligned} \quad (3.55)$$

As can be seen from Eq. (3.51), a summation over m corresponds to integrating the pseudo-differential operator over \mathbb{R} , resulting in a value of zero as recaped in Eq. (3.45). Therefore, the column sum of the drift matrix vanishes per construction

$$\sum_{m=1}^{N_p} D'_{mi}(x, t) = 0, \quad \forall i, \quad (3.56)$$

which in turn ensures the conservation of the particle density, given by the exact relation

$$n(x, t) = \sum_{m=1}^{N_p} \int_{C_m} f(p, x, t) dp. \quad (3.57)$$

Two crucial aspects need to be mentioned here, namely

- It is straightforward how to evaluate Eq. (3.55) for $m = 2, \dots, N_p - 1$ but integrating the equation over C_1 and C_{N_p} is problematic due to the half-bounded integration intervals.
- Eq. (3.54) is only an equation for the time-dependence of the cell integrals but not for the point-wise given values $f(p_m, x, t)$, which we are interested in.

In the case of the bounded intervals C_m , $m = 2, \dots, N_p - 1$, one can simply consider one sine function of $A_i(p, \eta')$ at a time and integrate over p , yielding a cosine function and an additional factor $1/\eta'$. After choosing some approximation for the potential, the remaining integration with respect to η' in Eq. (3.55) can be done by parts. It is not possible to apply this strategy to the half-bounded intervals due to the indefinite expressions, but performing the integration of the function $A_i(p, \eta')$, i.e. two sine functions at a time, is feasible, as demonstrated in the following for C_1 :

$$\begin{aligned} \int_{C_1} A_i(p, \eta') dp &= \int_{-\infty}^{p_{3/2}} \frac{\hbar}{\Delta p_i} \left[\sin\left(\frac{p - p_{i+1}}{\hbar} \eta'\right) - \sin\left(\frac{p - p_i}{\hbar} \eta'\right) \right] dp \\ &= \frac{\hbar}{\Delta p_i} \left[\int_{-\infty}^{p_{3/2} - p_{i+1}} \sin\left(\frac{q'}{\hbar} \eta'\right) dq' - \int_{-\infty}^{p_{3/2} - p_i} \sin\left(\frac{q''}{\hbar} \eta'\right) dq'' \right] \\ &= \frac{\hbar}{\Delta p_i} \int_{p_{3/2} - p_i}^{p_{3/2} - p_{i+1}} \sin\left(\frac{q'}{\hbar} \eta'\right) dq' \\ &= \frac{\hbar}{\Delta p_i} \left(-\frac{\hbar}{\eta'} \right) \left[\cos\left(\frac{p_{3/2} - p_{i+1}}{\hbar} \eta'\right) - \cos\left(\frac{p_{3/2} - p_i}{\hbar} \eta'\right) \right] \end{aligned} \quad (3.58)$$

where we used the substitutions $q' = p - p_{i+1}$ and $q'' = p - p_i$ in the second line. As can be seen, the half-bounded integrations cancel each other and only the value $p_{3/2}$ of the cell boundary gives a contribution.

In order to obtain from Eq. (3.54) the time dependence of the pointwise given values $f(p_m, x, t)$, we also assume for $f(p, x, t)$ on the left-hand side the same piecewise polynomial approximation, Eqs. (3.19) to (3.21). Evaluating the integrations then yields a system of equations for the N_p values $\partial_t f(p_m, x, t)$. In the case of an interior interval C_m , $m \neq \{1, N_p\}$, the integration may be done as follows:

$$\begin{aligned}
 \int_{C_m} \partial_t f(p, x, t) dp &= \int_{p_{m-1/2}}^{p_{m+1/2}} \partial_t \sum_{i=0}^{N_p} P^i(p, x, t) dp \\
 &= \int_{p_{m-1/2}}^{p_m} \partial_t P^{m-1}(p, x, t) dp + \int_{p_m}^{p_{m+1/2}} \partial_t P^m(p, x, t) dp \\
 &= \int_{p_{m-1/2}}^{p_m} \left[\dot{f}_{m-1} + \frac{\dot{f}_m - \dot{f}_{m-1}}{\Delta p_{m-1}} (p - p_{m-1}) \right] dp \\
 &\quad + \int_{p_m}^{p_{m+1/2}} \left[\dot{f}_m + \frac{\dot{f}_{m+1} - \dot{f}_m}{\Delta p_m} (p - p_m) \right] dp \\
 &= \frac{\Delta p_{m-1}}{8} \dot{f}_{m-1} + \frac{3}{8} (\Delta p_{m-1} + \Delta p_m) \dot{f}_m + \frac{\Delta p_m}{8} \dot{f}_{m+1}, \quad (3.59)
 \end{aligned}$$

where we used in the second and third line the definitions of the polynomials, Eqs. (3.19) to (3.21), together with the abbreviation $\dot{f}_m = \partial_t f(p_m, x, t)$. The calculation for the first and last interval can be done in the same way, when noting that all of the polynomials $P^i(p, x, t)$ are zero outside (p_0, p_{N_p+1}) so that the half-bounded integrals automatically reduce to bounded ones. The results are

$$\begin{aligned}
 \int_{C_1} \partial_t f(p, x, t) dp &= \left(\frac{\Delta p_0}{2} + \frac{3}{8} \Delta p_1 \right) \dot{f}_1 + \frac{\Delta p_1}{8} \dot{f}_2, \\
 \int_{C_{N_p}} \partial_t f(p, x, t) dp &= \frac{\Delta p_{N_p-1}}{8} \dot{f}_{N_p-1} + \left(\frac{\Delta p_{N_p}}{2} + \frac{3}{8} \Delta p_{N_p-1} \right) \dot{f}_{N_p}. \quad (3.60)
 \end{aligned}$$

Inserting this results into Eq. (3.54) enables us to write the equations for the time evolution in matrix form

$$M_p^0 \dot{\mathbf{f}}(x, t) = D'(x, t) \mathbf{f}(x, t), \quad (3.61)$$

where $\mathbf{f}(x, t)$ stands for the column vector obtained by grouping all $f(p_i, x, t)$, $i = 1, \dots, N_p$ and $D'(x, t) = (D'_{m,i}(x, t))$ stands for the matrix with its elements defined by Eq. (3.55). The matrix elements of M_p^0 are given by the coefficients appearing in Eqs. (3.59) and (3.60), obviously yielding a tridiagonal system of equations. The matrix was labeled this way since M_p^0 allows us to calculate the zeroth moment of the

piecewise linear function $f(p, x, t)$ by taking the column sum and multiplying it with $\mathbf{f}(x, t)$, compare Eq. (3.57). An explicit expression determining the time evolution of $\mathbf{f}(x, t)$ may now be obtained by inversion of M_p^0 , so that

$$\begin{aligned}\dot{\mathbf{f}}(x, t) &= (M_p^0)^{-1} D'(x, t) \mathbf{f}(x, t) \\ &= D(x, t) \mathbf{f}(x, t).\end{aligned}\tag{3.62}$$

The new drift matrix $D(x, t) = (M_p^0)^{-1} D'(x, t)$ is not labeled with a prime anymore since it is now in its final form which can be directly used for implementing the algorithm. One may note that the matrix M_p^0 does not depend on x , so that the inversion has to be carried out only once and is therefore not critical from a numerical point of view. But, a different problem arises due to the inversion. As mentioned before, the present algorithm has the tendency to create spurious oscillations when the p grid is chosen in an intermediate regime, between a well resolved solution and one in which short-scaled oscillations are averaged out. This can be understood when noting that the Eqs. (3.59) and (3.60) directly correspond to the problem of fitting values for the cell integrals to a piecewise polynomial function. The values of the cell integrals on the left-hand side of Eqs. (3.59) and (3.60) are directly given via the expression for the drift matrix (Eq. (3.54)), when considering a certain time t and a certain set of functional values $f(p_i, x, t)$. The remaining problem is to find the proper piecewise polynomial function $\partial_t \sum_{i=0}^{N_p} P^i(p, x, t)$ which has on each of the intervals C_m the same cell integrals. Due to the choice of the polynomials given by Eqs. (3.20), (3.21) and (3.23), the N_p values for the cell integrals uniquely determine the N_p values of $\dot{f}(p_m, x, t)$, represented by the inversion of M_p^0 in Eq. (3.62). In practice, the set of cell integrals shows an oscillating behaviour on different length scales. A coupled interpolation scheme as represented above is therefore not a good choice and one unavoidably encounters an overshooting of the interpolation function, creating spurious oscillations. Furthermore, the cell integrals at the endpoints usually take on very large values (note the half-bounded integration interval) which does not influence solely the values of $\dot{f}(p_m, x, t)$ at p_1 and p_{N_p} but on the whole domain. As can be seen from Eq. (3.60), this influence may be reduced by choosing Δp_0 and Δp_{N_p} large compared to the other values of Δp_m but one has to be careful in order to introduce not too large numerical errors.

It was tried to circumvent these problems by making use of different integration intervals C_m . Since the intervals are free to be chosen one may devise a set of $N_p - 1$ integration intervals, leaving one degree of freedom to find the N_p values of $\dot{f}(p_m, x, t)$. It is then possible to optimize in the remaining, one-dimensional subspace for the least oscillating interpolation function. In [24] different variation principles like minimum curvature or minimum total variation methods are described, which, upon optimization using Lagrange multipliers for instance, may lead to a linear set of equations. This set of equations can then be written in matrix form as well, replacing M_p^0 in Eq. (3.61). Such a strategy using a minimum total variation method was tested in practice but the problems described above could not be eliminated.

It is therefore believed, that the problem of spurious oscillations arises in particular due to the constraint of a continuous interpolation function. This and the fact that even analytical solutions for the Wigner function may show discontinuities, see Sec. 6.1 in the appendix, motivated to redo the calculations for a piecewise polynomial approximation of $f(p, x, t)$ without the constraint of continuity. It is then possible to directly apply the ideas of finite volume methods and work with a discrete set of values $f(p_m, x, t)$ which already represent cell averages. The approach will be described in detail in the following sections, after some general remarks on exponential integrals.

One point left to be addressed is how to include the advection term of the WTE. In order to be consistent, the same approximation of $f(p, x, t)$ as a continuous piecewise polynomial function was inserted in the expression for the advection term as well. One then replaces the averaging of Eq. (3.50) by integrating the whole WTE over the intervals C_m ,

$$\int_{C_m} \partial_t f(p, x, t) dp = \int_{C_m} (\mathcal{L}_A + \mathcal{L}_D) f(p, x, t) dp, \quad m = 1, \dots, N_p. \quad (3.63)$$

The integration over $\mathcal{L}_A f(p, x, t)$, as defined in Eq. (3.9), is replaced by an integration over $\frac{-1}{m^*} \partial_x \sum_{i=0}^{N_p} p P^i(p, x, t)$ and can be done in a completely analogous way as described above. The new matrix obtained by this takes on a tridiagonal form as well and may be labeled M_p^1 since it corresponds to the first moment. This results in

$$M_p^0 \partial_t \mathbf{f}(x, t) = -\frac{1}{m^*} M_p^1 \partial_x \mathbf{f}(x, t) + D'(x, t) \mathbf{f}(x, t). \quad (3.64)$$

After multiplying again with the inverse of M_p^0 one ends up with

$$\partial_t \mathbf{f}(x, t) = -\frac{1}{m^*} A \partial_x \mathbf{f}(x, t) + D(x, t) \mathbf{f}(x, t), \quad (3.65)$$

by defining

$$A = (M_p^0)^{-1} M_p^1. \quad (3.66)$$

The advection matrix A is not sparse as a result of the inversion process. It therefore couples different values of $f(p_m, x, t)$ when being advected along x , solely due to the particular choice of the continuous, piecewise approximation. Having to deal with a full matrix multiplication for the advection term as well is disadvantageous from a numerical point of view, since it drastically slows down calculations for larger systems.

3.2.2 Implementation of sine and cosine integrals

The calculations in the previous section demonstrated the appearance of a cosine integral, see Eq. (3.41), when calculating the drift matrix elements for the simplest case, a square potential barrier without cell-averaging the equations. It is a general feature of the piecewise polynomial approximation that one has to deal with cosine

and sine integrals. Since they will appear throughout the upcoming calculations we have a closer look on the question, how one actually evaluates the expressions in practice.

The usual strategy to tackle such integrals is to use a combination of a series expansion and complex continued fraction, with a certain crossover point [24]. Library routines are based on the standard definitions of the cosine, sine and exponential integrals, which read

$$\begin{aligned} \text{Ci}(x) &= - \int_x^\infty \frac{\cos t}{t} dt \\ &= \gamma + \ln x + \int_0^x \frac{\cos t - 1}{t} dt, \end{aligned} \quad (3.67)$$

$$\text{Si}(x) = \int_0^x \frac{\sin t}{t} dt \quad (3.68)$$

and

$$E_1(x) = \int_x^\infty \frac{\exp(-t)}{t} dt, \quad x \geq 0. \quad (3.69)$$

All of them are defined for complex arguments as well and γ is the Euler–Mascheroni constant [24] [26]. In our case we use the built in Matlab functions, whereby $\text{Ci}(x)$ takes on complex values for $x < 0$. For this reason and in order to make use of an interconnection between the sine, cosine and exponential integral listed at the end of this section, all formulas are expressed in terms of $\text{Ci}(|x|)$.

In the following different integral expressions are listed which will be encountered in the calculations of the drift matrix elements and it is shown how to link them to the standard definitions of the cosine and sine integrals. We will start with the two basic forms of a cosine and a sine integral, labeled by I_c and I_s . After that we consider two more cosine integrals labeled by I'_c and I''_c for which a careful evaluation requires a number of steps due to the modulus involved in the integral. To present them at this point should help in later calculations to keep track of the main steps. The first expression we consider is the cosine integral

$$I_c = \int_{ka}^{kb} \frac{1}{t} \cos(t) dt, \quad (3.70)$$

with a and b representing two distances and k a particular wave number. The value of I_c does only depend on the absolute values of ka and kb , but not on their sign.

This can be seen from the following two expressions,

$$\int_{ka}^{kb} \frac{1}{t} \cos(t) dt = \int_{-ka}^{-kb} \frac{1}{t'} \cos(t') dt' \quad (3.71)$$

and

$$P.V. \int_{-|ka|}^{|ka|} \frac{1}{t} \cos(t) dt = 0, \quad (3.72)$$

where the principal value integral vanishes since the integrand is odd. Using this fact, one can relate I_c in the following way to $\text{Ci}(x)$

$$\begin{aligned} I_c &= \int_{|ka|}^{|kb|} \frac{1}{t} \cos(t) dt \\ &= \int_{|ka|}^{\infty} \frac{1}{t} \cos(t) dt + \int_{\infty}^{|kb|} \frac{1}{t} \cos(t) dt \\ &= \text{Ci}(|kb|) - \text{Ci}(|ka|). \end{aligned} \quad (3.73)$$

The analogous situation for a sine integral is straightforward to evaluate and one finds

$$\begin{aligned} I_s &= \int_{ka}^{kb} \frac{1}{t} \sin(t) dt \\ &= \text{Si}(kb) - \text{Si}(ka). \end{aligned} \quad (3.74)$$

When performing a cell averaging with respect to the x variable one needs to integrate the sine and cosine integrals once more. We therefore need their primitive integrals, given by [26] [28]

$$\int \text{Ci}(z) dz = z \text{Ci}(z) - \sin(z), \quad (3.75)$$

$$\int \text{Si}(z) dz = z \text{Si}(z) + \cos(z). \quad (3.76)$$

To formula for the sine integral can be applied again in a straightforward manner but we need to have a closer look at the expression for the cosine integral due to the modulus of the argument.

One of the expressions we will encounter is of the form

$$I'_c = \int_{C_j} -k \text{Ci} [|2k(x_b - x)|] dx, \quad (3.77)$$

where the integration interval C_j stands for $(x_{j-1/2}, x_{j+1/2})$ and we assume that $x_{j+1/2} > x_{j-1/2}$. To get rid of the modulus we make the following three case differentiations. At first we consider $(x_b - x_{j+1/2}) > 0$ and $(x_b - x_{j-1/2}) > 0$:

$$\begin{aligned}
 I'_{c,1} &= \int_{C_j} -k \text{Ci} [2|k|(x_b - x)] dx \\
 &= \int_{2|k|(x_b - x_{j-1/2})}^{2|k|(x_b - x_{j+1/2})} \frac{k}{2|k|} \text{Ci}(u) du \\
 &= \frac{1}{2} \text{sgn}(k) [u \text{Ci}(u) - \sin(u)] \Big|_{2|k|(x_b - x_{j-1/2})}^{2|k|(x_b - x_{j+1/2})} \\
 &= \frac{1}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})}, \tag{3.78}
 \end{aligned}$$

where we used the substitution $u = 2|k|(x_b - x)$ in the second line. For the case $(x_b - x_{j+1/2}) < 0$ and $(x_b - x_{j-1/2}) < 0$ one finds in an analogous manner:

$$\begin{aligned}
 I'_{c,2} &= \int_{C_j} -k \text{Ci} [2|k|(x - x_b)] dx \\
 &= \int_{-2|k|(x_b - x_{j-1/2})}^{-2|k|(x_b - x_{j+1/2})} -\frac{k}{2|k|} \text{Ci}(u) du \\
 &= -\frac{1}{2} \text{sgn}(k) [u \text{Ci}(u) - \sin(u)] \Big|_{-2|k|(x_b - x_{j-1/2})}^{-2|k|(x_b - x_{j+1/2})} \\
 &= \frac{1}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})}. \tag{3.79}
 \end{aligned}$$

Finally we look at the case $(x_b - x_{j+1/2}) < 0$ and $(x_b - x_{j-1/2}) > 0$. Since $\text{Ci}(z = 0)$ is not defined we introduce an infinitesimal quantity ε :

$$\begin{aligned}
 I'_{c,3} &= \int_{C_j} -k \text{Ci} [2|k|(x_b - x)] dx \\
 &= \lim_{\varepsilon \rightarrow 0^+} -k \left[\int_{x_{j-1/2}}^{x_b - \varepsilon} \text{Ci} [2|k|(x_b - x)] dx + \int_{x_b + \varepsilon}^{x_{j+1/2}} \text{Ci} [2|k|(x - x_b)] dx \right] \\
 &= \frac{1}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\
 &\quad + \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k[x_b - (x_b + \varepsilon)]}^{2k[x_b - (x_b - \varepsilon)]}. \tag{3.80}
 \end{aligned}$$

When performing the limit $\varepsilon \rightarrow 0^+$ it is clear that the sine term vanishes and for the

other one we obtain

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} 2k\varepsilon \text{Ci}(|2k\varepsilon|) &= \lim_{\varepsilon \rightarrow 0^+} \left[2k\varepsilon(\gamma + \ln(|2k\varepsilon|)) + 2k\varepsilon \int_0^{|2k\varepsilon|} \frac{\cos t - 1}{t} dt \right] \\ &= 0, \end{aligned} \quad (3.81)$$

since $x \ln x = 0$ for $x \rightarrow 0^+$, as can be checked e.g. by using L'Hospital's Rule, and the other terms vanish trivially. Therefore, one can rewrite I'_c in general to

$$\begin{aligned} I'_c &= - \int_{x_{j-1/2}}^{x_{j+1/2}} k \text{Ci} [|2k(x_b - x)|] dx \\ &= \frac{1}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})}. \end{aligned} \quad (3.82)$$

The second, similar expression is of the form

$$I''_c = \int_{C_j} -kx \text{Ci} [|2k(x_b - x)|] dx. \quad (3.83)$$

At first we consider again $(x_b - x_{j+1/2}) > 0$ and $(x_b - x_{j-1/2}) > 0$:

$$\begin{aligned} I''_{c,1} &= \int_{C_j} -kx \text{Ci} [2|k|(x_b - x)] dx \\ &= \frac{k}{2|k|} \int_{2|k|(x_b - x_{j-1/2})}^{2|k|(x_b - x_{j+1/2})} \left(x_b - \frac{u}{2|k|} \right) \text{Ci}(u) du \\ &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\ &\quad - \frac{\text{sgn}(k)}{4|k|} \left[\frac{1}{2} u^2 \text{Ci}(u) \Big|_{2|k|(x_b - x_{j-1/2})}^{2|k|(x_b - x_{j+1/2})} - \frac{1}{2} \int_{2|k|(x_b - x_{j-1/2})}^{2|k|(x_b - x_{j+1/2})} u \cos(u) du \right] \\ &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\ &\quad + \frac{1}{8k} [-u^2 \text{Ci}(|u|) + u \sin(u) + \cos(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})}, \end{aligned} \quad (3.84)$$

when substituting again in the second line $u = 2|k|(x_b - x)$, using Eq. (3.78) in the third line and integrating by parts in line three and four. For the case $(x_b - x_{j+1/2}) < 0$

and $(x_b - x_{j-1/2}) < 0$ we find:

$$\begin{aligned}
 I''_{c,2} &= \int_{C_j} -kx \text{Ci}[2|k|(x - x_b)] dx \\
 &= -\frac{k}{2|k|} \int_{-2|k|(x_b - x_{j-1/2})}^{-2|k|(x_b - x_{j+1/2})} \left(x_b + \frac{u}{2|k|}\right) \text{Ci}(u) du \\
 &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\
 &\quad - \frac{\text{sgn}(k)}{4|k|} \left[\frac{1}{2} u^2 \text{Ci}(u) \Big|_{-2|k|(x_b - x_{j-1/2})}^{-2|k|(x_b - x_{j+1/2})} - \frac{1}{2} \int_{-2|k|(x_b - x_{j-1/2})}^{-2|k|(x_b - x_{j+1/2})} u \cos(u) du \right] \\
 &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\
 &\quad + \frac{1}{8k} [-u^2 \text{Ci}(|u|) + u \sin(u) + \cos(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})}, \tag{3.85}
 \end{aligned}$$

For the last case $(x_b - x_{j+1/2}) < 0$ and $(x_b - x_{j-1/2}) > 0$ we introduce once more an infinitesimal quantity ε :

$$\begin{aligned}
 I''_{c,3} &= \int_{C_j} -kx \text{Ci}[2|k|(x_b - x)] dx \\
 &= \lim_{\varepsilon \rightarrow 0^+} -k \left[\int_{x_{j-1/2}}^{x_b - \varepsilon} x \text{Ci}[2|k|(x_b - x)] dx + \int_{x_b + \varepsilon}^{x_{j+1/2}} x \text{Ci}[2|k|(x - x_b)] dx \right] \\
 &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\
 &\quad + \frac{1}{8k} [-u^2 \text{Ci}(|u|) + u \sin(u) + \cos(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\
 &\quad + \lim_{\varepsilon \rightarrow 0^+} \frac{1}{8k} [-u^2 \text{Ci}(|u|) + u \sin(u) + \cos(u)] \Big|_{-2k\varepsilon}^{2k\varepsilon}, \tag{3.86}
 \end{aligned}$$

where one can make use of Eq. (3.81) to see that the last term vanishes when taking the limit $\varepsilon \rightarrow 0^+$, so that one finds in general

$$\begin{aligned}
 I''_c &= - \int_{x_{j-1/2}}^{x_{j+1/2}} kx \text{Ci}[|2k(x_b - x)|] dx \\
 &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})} \\
 &\quad + \frac{1}{8k} [-u^2 \text{Ci}(|u|) + u \sin(u) + \cos(u)] \Big|_{2k(x_b - x_{j-1/2})}^{2k(x_b - x_{j+1/2})}. \tag{3.87}
 \end{aligned}$$

As mentioned earlier, one is able to interconnect the cosine, sine and exponential integrals. For purely imaginary arguments of the exponential integral they are related by [24] [27]

$$E_1(ix) = i \left(\text{Si}(x) - \frac{\pi}{2} \right) - \text{Ci}(x), \quad x > 0. \quad (3.88)$$

This enables us to express $\text{Ci}(|x|)$ in terms of the real part of $E_1(i|x|)$ and by noting that the sine integral is an odd function, as can be seen from the definition Eq. (3.68), we are able to deduce

$$\begin{aligned} \text{Si}(x) &= \text{sgn}(x) \left(\Im \{E_1(i|x|)\} + \frac{\pi}{2} \right), \\ \text{Ci}(|x|) &= -\Re \{E_1(i|x|)\}. \end{aligned} \quad (3.89)$$

The relations are useful for actual computations since on the one hand, tests showed that the evaluation of the exponential integral routine is about one to two orders of magnitude faster than that for the sine or cosine integral and on the other hand, in the formulas that we will deal with in the later chapters sine and cosine integrals have to be evaluated for the same arguments. Therefore, all formulas, even though if written in terms of sine and cosine integrals to keep things clear are calculated in actual computations via the connection to the exponential integral. Practical tests showed that no significant error is encountered when doing so and the results agree to machine precision.

3.2.3 Application of finite volume methods

In Sec. 3.2.1 we experienced that several problems were associated with the inversion of the matrix M_p^0 , which came into play in order to obtain an explicit update formula for the discrete values of the Wigner function. In this section we will develop a method based on similar ideas but one in which the discrete values of the Wigner function already represent cell averages, so that one directly arrives at an explicit update formula without the need to solve a system of equations. In fact, this corresponds to a direct application of the ideas of finite volume methods.

Usually, finite volume methods are applied to deal with conservation laws of hyperbolic type which may be cast into a flux conservative form [18] [11] [24]. Writing the equations in a flux conservative form has the great advantage that every approximate, numerical solution automatically fulfills the conservation law in question and is thereby a valid physical solution. A higher precision of the numerical solution may then be obtained by using some higher order approximation for the fluxes at the cell boundaries. We will have a closer look at this in Sec. 3.3 when treating the advection term of the Wigner transport equation. An example which may be cast into a flux conservative form, relevant for device simulations, is the semi-classical Boltzmann transport equation. In the case of the WTE a flux conservative form cannot be found but one can still rewrite the equation as a conservation law by averaging it over grid cells. When choosing the grid cells properly, such that the set of grid cells covers the

whole real p axis, one can make sure that the number of particles is exactly conserved independent of the specific choice of grid points.

In the following we will average the Wigner transport equation over grid cells $C_{m,j}$. For this we define

$$\begin{aligned} C_m &= (p_{m-1/2}, p_{m+1/2}), \quad m = 1, \dots, N_p, \\ C_j &= (x_{j-1/2}, x_{j+1/2}), \quad j = 1, \dots, N_x, \\ C_{m,j} &= C_m \times C_j. \end{aligned} \quad (3.90)$$

The $N_p \times N_x$ cells $C_{m,j}$ define the compact domain on which the approximate quasi-distribution function $f(p, x, t)$ is allowed to take on non-zero values. In order to construct a conservative method we need in addition the half-bounded intervals to $\pm\infty$, which we label by

$$\begin{aligned} C_0 &= (-\infty, p_{1/2}), \\ C_{N_p+1} &= (p_{N_p+1/2}, \infty), \\ C_{0,j} &= C_0 \times C_j, \quad j = 1, \dots, N_x, \\ C_{N_p+1,j} &= C_{N_p+1} \times C_j, \quad j = 1, \dots, N_x. \end{aligned} \quad (3.91)$$

The grid for the x variable is chosen to be equidistant in our case but all derivations of the drift matrix elements can be directly adopted to a non-equidistant x grid as well. No constraints are assumed for the grid spacing of the p variable, enabling a highly flexible and adaptable grid for different physical situations. In particular, the cell boundaries for the x and p grid are defined by

$$x_{j-1/2} = x_{1/2} + (j-1)\Delta x, \quad \Delta x \in \mathbb{R}, \quad j = 1, \dots, N_x + 1 \quad (3.92)$$

and

$$p_{m-1/2} = p_{1/2} + \sum_{l=1}^{m-1} \Delta p_l, \quad p_{1/2} \in \mathbb{R}, \quad \Delta p_l \in \mathbb{R}, \quad m = 2, \dots, N_p + 1. \quad (3.93)$$

The central points of the grid cells are labeled by integer indices and given by

$$x_j = \frac{x_{j-1/2} + x_{j+1/2}}{2}, \quad j = 1, \dots, N_x \quad (3.94)$$

and

$$p_m = \frac{p_{m-1/2} + p_{m+1/2}}{2}, \quad m = 1, \dots, N_p. \quad (3.95)$$

For a schematic outline of the p grid see Fig. 3.3. The cell average of the Wigner function over the grid cell $C_{m,j}$ is given by

$$F_{m,j}(t) = \frac{1}{\Delta p_m \Delta x} \iint_{C_{m,j}} f(p, x, t) dp dx, \quad m = 1, \dots, N_p, \quad j = 1, \dots, N_x, \quad (3.96)$$

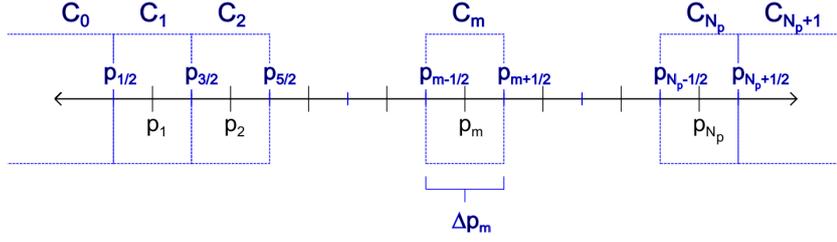


Figure 3.3: Schematic representation of the p grid and the grid cells C_m for the piecewise polynomial approximation of $f(p, x, t)$ based on finite volume methods. Note that the spacing $\Delta p_m = p_{m+1/2} - p_{m-1/2}$ is now defined as the difference between two intermediate grid points.

where we used a capital letter for the cell average to distinguish it from the continuous function $f(p, x, t)$. To arrive at a conservative scheme we average now the whole Wigner transport equation Eq. (2.10) over grid cells:

$$\frac{1}{\Delta p_m \Delta x} \iint_{C_{m,j}} \partial_t f(p, x, t) dp dx = \frac{1}{\Delta p_m \Delta x} \iint_{C_{m,j}} \left[-\frac{p}{m^*} \partial_x f(p, x, t) + \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) \right] dp dx. \quad (3.97)$$

The left-sided term can be directly identified as the the time derivative of the cell average, $\partial_t F_{m,j}(t)$. For the moment, we rewrite the first term on the right-hand side, the advection term, as

$$\begin{aligned} \frac{1}{\Delta p_m \Delta x} \iint_{C_{m,j}} -\frac{p}{m^*} \partial_x f(p, x, t) dp dx &= \frac{1}{\Delta p_m \Delta x} \int_{C_m} -\frac{p}{m^*} [f(p, x_{j+1/2}, t) \\ &\quad - f(p, x_{j-1/2}, t)] dp \\ &= -\frac{1}{\Delta x} [\hat{h}_{m,j+1/2}(t) - \hat{h}_{m,j-1/2}(t)], \end{aligned} \quad (3.98)$$

with

$$\hat{h}_{m,j+1/2}(t) = \frac{1}{\Delta p_m} \int_{C_m} \frac{p}{m^*} f(p, x_{j+1/2}, t) dp. \quad (3.99)$$

The quantity $\hat{h}_{m,j+1/2}(t)$ herein labels the flux at the cell boundary $x_{j+1/2}$, averaged over all values of p within the interval C_m . Later on, in Sec. 3.3, we will see how to express $\hat{h}_{m,j+1/2}$ in terms of the cell averages $F_{m,j}(t)$ when discussing discretization methods for flux conservative equations. For simplicity we leave out the advection term in the following equations in this section and focus on the temporal change of $F_{m,j}(t)$ due to the pseudo-differential operator.

In order to express the last term in Eq. (3.97) in terms of the cell averages, we approximate $f(p, x, t)$ by a piecewise polynomial function whose coefficients depend

linearly on the set of cell averages $F_{m,j}(t)$. In the following, $f(p, x, t)$ does not label the exact solution of the WTE anymore but stands for a function defined for continuous values of p and x , reconstructed out of the set of discrete values $F_{m,j}(t)$. In principle, this is a direct application of Godunov's method or REA algorithm, known from the finite volume methods developed for hyperbolic systems [11]. The abbreviation REA stands for reconstruct - evolve - average. This is exactly what we will do in the following: at first we reconstruct a function $f(p, x, t)$ for continuous arguments out of the set of cell averages, then, to evolve the function in time we calculate the action of the pseudo-differential operator on $f(p, x, t)$ exactly (for a given potential) and at last, we average the result over grid cells to obtain the new cell averages. As a side note: At the moment we obtain instead of the new value of $F_{m,j}(t + \Delta t)$ the rate of temporal change, $\partial_t F_{m,j}(t)$. Since we choose the reconstruction function $f(p, x, t)$ in such a way that it depends linearly on the set of $F_{m,j}(t)$, all three steps of the REA algorithm will finally condense into a single matrix expression

$$\partial_t F_{m,j}(t) = \sum_{i=1}^{N_p} D_{m,i,j}(t) F_{i,j}(t), \quad (3.100)$$

with D termed drift matrix in this work.

General considerations In this part of the present section we state the approach to a piecewise polynomial approximation of $f(p, x, t)$ in general, for the purpose of pointing out that a polynomial approximation to any order is feasible. Two crucial aspects will be discussed: Firstly, the existence of the integrals and secondly the problem of cell averaging over half-bounded intervals. The first point is meant to elucidate that the integrals determining the drift matrix elements exist for any bounded potential shape, even though the single terms in the integrals show divergent behaviour. The order of these singularities increases with the order of the polynomial approximation but there is an easy way to show the existence of all of this integrals. The second point is crucial for the calculation of the drift matrix elements for the boundaries, i.e. the first and last row of D . In Sec. 3.2.1 we saw that one could evaluate the half-bounded integrals by regrouping the terms and using a substitution but this approach cannot be used in general. For instance the first-order approximation presented later poses a problem. But again, an easy way exists to evaluate the half-bounded expressions in general, closely related to the calculation for the first point.

To reconstruct a function $f(p, x, t)$ defined for continuous arguments p and x , given the set of cell averages $F_{m,j}(t)$, we choose the following piecewise polynomial approximation

$$f(p, x, t) = \sum_{m=1}^{N_p} \sum_{j=1}^{N_x} P_{m,j}^n(p, x, t), \quad (3.101)$$

with the polynomials of order n defined by

$$P_{m,j}^n(p, x, t) := \begin{cases} \sum_{\nu=0}^n a_{m,j,\nu}(t)(p - p_m)^\nu & \text{if } (p, x) \in C_{m,j}, \quad m \neq \{0, N_p + 1\}, \\ 0 & \text{elsewhere.} \end{cases} \quad (3.102)$$

Herein the order of the polynomial refers only to the p variable, with respect to the x variable we always use a piecewise constant approximation. The particular coefficients $a_{m,j,\nu}(t)$ are not specified at this point and the accuracy of the algorithm depends on a smart choice of the coefficients. The only necessary constraint is that the coefficients have to be chosen in a way such that Eq. (3.96) is always fulfilled, i.e. that the reconstructed function $f(p, x, t)$ is consistent with the cell averages $F_{m,j}(t)$. Plugging Eqs. (3.101) and (3.102) into Eq. (3.96) results in the requirement

$$F_{m,j}(t) = \frac{1}{\Delta p_m} \int_{p_{m-1/2}}^{p_{m+1/2}} \sum_{\nu=0}^n a_{m,j,\nu}(t)(p - p_m)^\nu dp, \quad m = 1, \dots, N_p, \quad j = 1, \dots, N_x. \quad (3.103)$$

In addition, we demand that the coefficients $a_{m,j,\nu}(t)$ depend linearly on the set of cell averages $F_{m,j}(t)$ in order to finally arrive at a linear expression of the form of Eq. (3.100) but this requirement is not essential, of course.

Choosing a piecewise constant approximation with respect to x greatly simplifies the cell averaging over the x variable. Otherwise one could not obtain a drift matrix expression of the form of Eq. (3.100) with a set of N_x two-dimensional matrices but would have to deal with one four-dimensional tensor instead. This neglecting of the coupling of $F_{m,j}(t)$'s with different values of j , i.e. along the x axis, seems to be reasonable since one can see also from the definition of the pseudo-differential operator, Eq. (2.12), that the potential couples different $f(p, x, t)$ and $f(p', x, t)$ but always at the same point x .

To calculate the action of the pseudo-differential operator on the piecewise polynomial function $f(p, x, t)$ we calculate the Fourier transformed function $\tilde{f}(\eta, x, t)$ first. Inserting the definition of the polynomials into Eq. (2.13) yields

$$\begin{aligned} \tilde{f}(\eta, x, t) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \sum_{m=1}^{N_p} \sum_{j=1}^{N_x} P_{m,j}^n(p, x, t) \exp\left(-i\frac{p}{\hbar}\eta\right) dp \\ &= \frac{1}{2\pi\hbar} \sum_{m=1}^{N_p} \sum_{j=1}^{N_x} \sum_{\nu=0}^n a_{m,j,\nu}(t) \Theta(x_{j-1/2} \leq x < x_{j+1/2}) \\ &\quad \times \int_{p_{m-1/2}}^{p_{m+1/2}} (p - p_m)^\nu \exp\left(-i\frac{p}{\hbar}\eta\right) dp, \end{aligned} \quad (3.104)$$

with the abbreviation

$$\Theta(x_{j-1/2} \leq x < x_{j+1/2}) = H(x - x_{j-1/2})H(x_{j+1/2} - x) \quad (3.105)$$

defined as the product of two Heaviside step functions. For calculating the action of the pseudo-differential operator one has to insert the Fourier transform $\tilde{f}(\eta, x, t)$ in

Eq. (2.12). To ensure the existence of all integrals its is important to know the limit $\eta \rightarrow 0$ of $\tilde{f}(\eta, x, t)$, which can be readily evaluated from the general form Eq. (3.104) to

$$\tilde{f}(\eta = 0, x, t) = \frac{1}{2\pi\hbar} \sum_{m=1}^{N_p} \sum_{j=1}^{N_x} \sum_{\nu=0}^n \Theta(x_{j-1/2} \leq x < x_{j+1/2}) a_{m,j,\nu}(t) \int_{p_{m-1/2}}^{p_{m+1/2}} (p - p_m)^\nu dp, \quad (3.106)$$

which is a finite, real number.

Now, to calculate the action of the pseudo-differential operator we make use of the fact that $(\Theta_{\hbar}[V]f)(p, x, t)$ is always real and insert $\tilde{f}(\eta, x, t)$ into Eq. (2.26)

$$\begin{aligned} (\Theta_{\hbar}[V]f)(p, x, t) &= 2\Re \left\{ \int_0^{\infty} \frac{im^*}{\hbar} \left[V\left(x + \frac{\eta'}{2}, t\right) - V\left(x - \frac{\eta'}{2}, t\right) \right] \tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \right\} \\ &= 2\Re \left\{ \int_0^{\infty} \frac{im^*}{\hbar} V\left(x + \frac{\eta'}{2}, t\right) \tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \right\} \\ &\quad - 2\Re \left\{ \int_0^{\infty} -\frac{im^*}{\hbar} V\left(x + \frac{\eta''}{2}, t\right) \tilde{f}(-\eta'', x, t) \exp\left(-i\frac{p}{\hbar}\eta''\right) d\eta'' \right\} \\ &= 2\Re \left\{ \int_0^{\infty} \frac{im^*}{\hbar} V\left(x + \frac{\eta'}{2}, t\right) \tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \right\} \\ &\quad - 2\Re \left\{ \int_0^{\infty} \frac{m^*}{\hbar} V\left(x + \frac{\eta''}{2}, t\right) \left[\tilde{f}(\eta'', x, t) \exp\left(i\frac{p}{\hbar}\eta''\right) \right]^* d\eta'' \right\} \\ &= 2\Re \left\{ \int_{-\infty}^{\infty} \frac{im^*}{\hbar} V\left(x + \frac{\eta'}{2}, t\right) \tilde{f}(\eta', x, t) \exp\left(i\frac{p}{\hbar}\eta'\right) d\eta' \right\}, \quad (3.107) \end{aligned}$$

where we used in the second line a simple substitution $\eta'' = -\eta'$, in the third line the property $\tilde{f}(-\eta, x, t) = [\tilde{f}(\eta, x, t)]^*$ and in the last line the fact, that the real part of a complex number and the real part of its complex conjugate are the same. Here we already made implicitly use of the fact that $\tilde{f}(\eta, x, t)$ results in a finite, real number for $\eta \rightarrow 0$, since the splitting of the integral into two parts in the second line is only allowed when both parts exist separately.

Equation (3.107) facilitates to calculate the action of a given potential. As mentioned in Sec. 3.2.1, a numerical evaluation of the integral for an arbitrary potential is unfeasible due to the oscillatory integrand. We thus choose a piecewise polynomial approximation for $V(x, t)$ as well and calculate the integrals exactly. The cosine and sine integrals appearing therein are evaluated by using some library routine. In this work, we consider at the most a piecewise linear potential shape and write $V(x, t)$ therefore as

$$V(x, t) = \sum_{k=0}^{N_V} p_k^V(x, t), \quad (3.108)$$

with

$$p_k^V(x, t) := \begin{cases} v_k^0(t) + v_k^1(t)x & \text{if } x_k^V \leq x < x_{k+1}^V, \\ 0 & \text{elsewhere.} \end{cases} \quad (3.109)$$

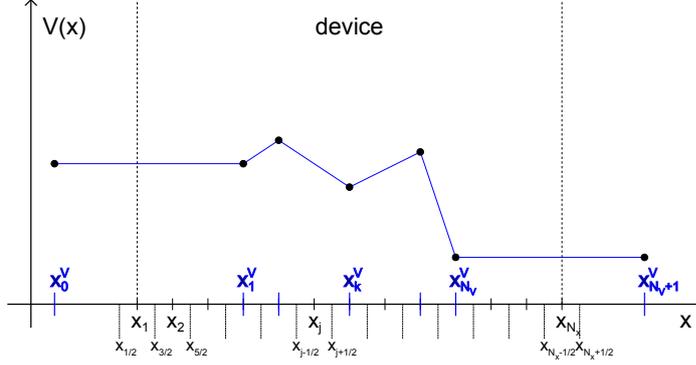


Figure 3.4: Schematic drawing of the piecewise linear approximation of the potential $V(x, t)$ as defined in Eqs. (3.108) and (3.109), here for the special case of a continuous approximation. As indicated in the figure, the nodes of the potential x_k^V , $k = 1, \dots, N_V$, are chosen as a subset of the grid points x_j , $j = 1, \dots, N_x$, for the Wigner function. The two points x_0^V and $x_{N_V+1}^V$ are located outside of the considered x domain.

The grid points for the potential are chosen as a subset of the grid points for the Wigner function, i.e. they are given by

$$\{x_k^V | k = 1, \dots, N_V\} \subseteq \{x_j | j = 1, \dots, N_x\}, \quad (3.110)$$

with $N_V \leq N_x$. Furthermore, two additional grid points outside the device are introduced, x_0^V and $x_{N_V+1}^V$, which we will let go to $\pm\infty$ in the final equations to model the situation of semi-infinite leads under bias. An example for a continuous potential approximated in this manner is given in Fig. 3.4. Inserting the potential as defined in Eqs. (3.108) and (3.109) into Eq. (3.107), yields

$$\begin{aligned} (\Theta_{\hbar}[V]f)(p, x, t) = & 2\Re \left\{ \sum_{k=0}^{N_V} \int_{2(x_k^V-x)}^{2(x_{k+1}^V-x)} \frac{im^*}{\hbar} \left[v_k^0(t) + v_k^1(t) \left(x + \frac{\eta'}{2} \right) \right] \right. \\ & \left. \times \tilde{f}(\eta', x, t) \exp \left(i \frac{p}{\hbar} \eta' \right) d\eta' \right\}, \end{aligned} \quad (3.111)$$

which can be directly evaluated for a specific approximation of $f(p', x, t)$ and therefore a given expression for $\tilde{f}(\eta', x, t)$. To explicitly calculate the integrals in the following parts of this section, we consider the simpler case of a single, linear potential segment which is nonzero only on one interval (x_1^V, x_2^V) . We introduce the abbreviations $x_a = x_1^V$ and $x_b = x_2^V$ to label these two potential nodes, see also Fig. 3.5. Furthermore, we consider a time-independent potential and drop the index k so that Eq. (3.111) reduces to the more compact form

$$(\Theta_{\hbar}[V]f)(p, x, t) = 2\Re \left\{ \frac{im^*}{\hbar} \int_{2(x_a-x)}^{2(x_b-x)} \left[v^0 + v^1 \left(x + \frac{\eta'}{2} \right) \right] \tilde{f}(\eta', x, t) \exp \left(i \frac{p}{\hbar} \eta' \right) d\eta' \right\}. \quad (3.112)$$

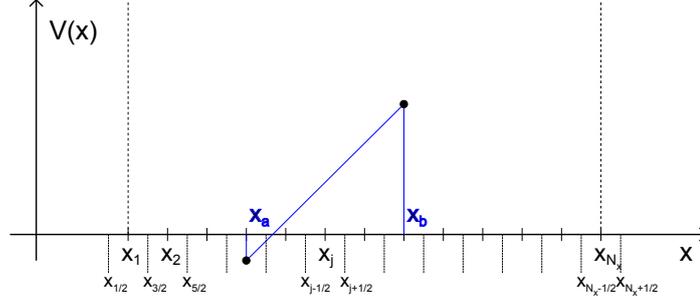


Figure 3.5: Analogous plot to Fig. 3.4 for the special case of a single, piecewise linear potential segment, which has nonzero values only on the interval (x_a, x_b) .

This reduction to a single potential element is introduced for the sake of clarity of the upcoming calculations and it should be clear, that the knowledge of the action of the pseudo-differential operator for this single potential segment suffices to construct any other, more complicated piecewise linear potential shape as stated in Eqs. (3.108) - (3.109).

We now focus on the cell averaging process to arrive at a conservative scheme. Therefore, we return to Eq. (3.97) but with the advection term left out for simplicity. For the case of the interior grid cells $C_{m,j}$, with interior referring to the p variable, we can use exactly the same expression as stated in Eq. (3.97)

$$\partial_t F_{m,j}(t) = \frac{1}{\Delta p_m \Delta x} \iint_{C_{m,j}} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) dp dx, \quad m = 2, \dots, N_p - 1, \quad j = 1, \dots, N_x. \quad (3.113)$$

For the boundary cells we need an additional step and start with the following integral expressions, written as example at the lower boundary:

$$\iint_{C_{1,j} \cup C_{0,j}} \partial_t f(p, x, t) dp dx = \iint_{C_{1,j} \cup C_{0,j}} \frac{q}{m^*} (\Theta_h[V]f)(p, x, t) dp dx, \quad j = 1, \dots, N_x, \quad (3.114)$$

which result in half-bounded integrals with respect to p as can be seen from the definition of the intervals in Eqs. (3.90) and (3.91). In order to evaluate the left-hand side of Eq. (3.114) we use the piecewise polynomial approximation of $f(p, x, t)$ and recall that the Wigner function is assumed to be zero outside $(p_{1/2}, p_{N_p+1/2})$, compare Eqs. (3.101) and (3.102). Therefore, the left-hand side reduces to

$$\begin{aligned} \iint_{C_{1,j} \cup C_{0,j}} \partial_t f(p, x, t) dp dx &= \iint_{C_{1,j}} \partial_t f(p, x, t) dp dx \\ &= \Delta p_1 \Delta x \partial_t F_{1,j}(t). \end{aligned} \quad (3.115)$$

For the right-hand side of Eq. (3.114) we have to consider the full, half-bounded interval $C_{1,j} \cup C_{0,j}$ since the action of the pseudo-differential operator does not vanish

outside the considered (p, x) -domain. One can see from the definition of the pseudo-differential operator, Eq. (2.12), that as soon as $f(p', x, t)$ is nonzero in some part of $p' \in \mathbb{R}$ the pseudo-differential operator takes on nonzero values for any value of p , in principle, only that the factor $\exp\left(i\frac{(p-p')}{\hbar}\eta'\right)$ leads to a phase cancellation for large differences of $p - p'$ and thus to a decay of the action of the pseudo-differential operator. On the whole, we obtain the following expressions for the temporal change of $F_{1,j}(t)$ and $F_{N_p,j}(t)$:

$$\begin{aligned}\partial_t F_{1,j}(t) &= \frac{1}{\Delta p_1 \Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{p_{3/2}} \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) dp dx, \quad j = 1, \dots, N_x, \\ \partial_t F_{N_p,j}(t) &= \frac{1}{\Delta p_{N_p} \Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{p_{N_p-1/2}}^{\infty} \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) dp dx, \quad j = 1, \dots, N_x.\end{aligned}\tag{3.116}$$

When comparing Eqs. (3.113) and (3.116) one can notice that in the case of the interior grid cells the temporal change of $F_{m,j}(t)$ is determined by averaging the pseudo-differential operator over the corresponding interval $C_{m,j}$, but in contrast, in the case of the exterior grid cells the temporal change of $F_{1,j}(t)$ and $F_{N_p,j}(t)$ is determined by an integration of the pseudo-differential operator over the semi-infinite intervals to $\pm\infty$. This is, of course, an approximation and furthermore a source of error for practical computations. In order to keep this error small one has to choose the p -domain on which $f(p, x, t)$ is allowed to take on nonzero values large enough. The advantage of the present formulation is that all the action of the pseudo-differential operator outside the p -domain is captured by the rate of change of $F_{1,j}$ and $F_{N_p,j}$, without directly affecting the other cell averages inside the domain as in the case of the continuous approximation presented in Sec. 3.2.1. Therefore, one is able to adjust the domain size required for a particular simulation by monitoring the values of the endpoints, for instance by increasing the size of $(p_{1/2}, p_{N_p+1/2})$ until the values of $F_{1,j}$ and $F_{N_p,j}$ stay below a certain value.

With the cell-averaged time evolution equations as stated in Eqs. (3.113) and (3.116), the conservation of the particle density in each interval of position space,

$$n(x_j, t) = \sum_{m=1}^{N_p} F_{m,j}(t) \Delta p_m,\tag{3.117}$$

is always ensured independent of the particular polynomial approximation of $f(p, x, t)$ and the potential shape. We recall the result from Sec. 2.2 that the action of the

pseudo-differential operator vanishes when integrating over all p , so that we find

$$\begin{aligned}
 \partial_t \sum_{m=1}^{N_p} F_{m,j}(t) \Delta p_m &= \sum_{m=2}^{N_p-1} \frac{1}{\Delta x} \iint_{C_{m,j}} \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) dp dx \\
 &\quad + \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{p_{3/2}} \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) dp dx \\
 &\quad + \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{p_{N_p-1/2}}^{\infty} \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) dp dx \\
 &= \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} \frac{q}{m^*} (\Theta_{\hbar}[V]f)(p, x, t) dp dx \\
 &= 0.
 \end{aligned} \tag{3.118}$$

Due to the fact that the particle density $n(x_j, t)$ is unaffected by the action of the pseudo-differential operator, the continuity equation for $n(x_j, t)$ is fulfilled.

By inserting Eq. (3.107) into Eq. (3.113) we find the following general, cell averaged expression valid for any polynomial approximation as defined in Eqs. (3.101) and (3.102)

$$\begin{aligned}
 \partial_t F_{m,j}(t) &= \frac{1}{\Delta p_m \Delta x} \Re \left\{ \frac{2iq}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{p_{m-1/2}}^{p_{m+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x, t) \right. \\
 &\quad \left. \times \exp \left(i \frac{p}{\hbar} \eta' \right) d\eta' dp dx \right\} \\
 &= \frac{1}{\Delta p_m \Delta x} \Re \left\{ \frac{2iq}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x_j, t) \right. \\
 &\quad \left. \times \int_{p_{m-1/2}}^{p_{m+1/2}} \exp \left(i \frac{p}{\hbar} \eta' \right) dp d\eta' dx \right\} \\
 &= \frac{1}{\Delta p_m \Delta x} \Re \left\{ \frac{2iq}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x_j, t) \right. \\
 &\quad \left. \times \left(-i \frac{\hbar}{\eta'} \right) \left[\exp \left(i \frac{p_{m+1/2}}{\hbar} \eta' \right) - \exp \left(i \frac{p_{m-1/2}}{\hbar} \eta' \right) \right] d\eta' dx \right\}, \\
 &\quad m = 2, \dots, N_p - 1, \quad j = 1, \dots, N_x,
 \end{aligned} \tag{3.119}$$

where we used in the second line on the one hand, that the exponential function is the only p -dependent quantity and on the other hand, that $\tilde{f}(\eta', x, t) = \tilde{f}(\eta', x_j, t)$ for $x \in (x_{j-1/2}, x_{j+1/2})$ since $f(p, x, t)$ is assumed to be a piecewise constant function with respect to x . The critical point is now to evaluate the same expression for the

boundary terms $m = \{1, N_p\}$. In the case of $m = 1$ we need to calculate

$$\begin{aligned} \partial_t F_{1,j}(t) &= \frac{1}{\Delta p_1 \Delta x} \Re \left\{ \frac{2iq}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x_j, t) \right. \\ &\quad \left. \times \int_{-\infty}^{p_{3/2}} \exp \left(i \frac{p}{\hbar} \eta' \right) dp d\eta' dx \right\}, \quad j = 1, \dots, N_x. \end{aligned} \quad (3.120)$$

The calculation of the integral over p cannot be done in a straightforward manner anymore. But, one can make use of the theory of distributions, see [29] [30], to link the integral to the Fourier transform of the Heaviside step function, given by [31]

$$\begin{aligned} \tilde{H}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(t) \exp(-i\omega t) dt \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{i\omega} + \sqrt{\frac{\pi}{2}} \delta(\omega). \end{aligned} \quad (3.121)$$

With this in hands we find for the integration with respect to p

$$\begin{aligned} \int_{-\infty}^{p_{3/2}} \exp \left(i \frac{p}{\hbar} \eta' \right) dp &= \hbar \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) \int_0^{\infty} \exp(-ik' \eta') dk' \\ &= \hbar \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) \left[\frac{1}{i\eta'} + \pi \delta(\eta') \right] \\ &= -i \frac{\hbar}{\eta'} \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) + \hbar \pi \delta(\eta'), \end{aligned} \quad (3.122)$$

where we used the substitution $k' = -\frac{p}{\hbar} + \frac{p_{3/2}}{\hbar}$ in the first line. After insertion in Eq. (3.120) we can write the result as

$$\begin{aligned} \partial_t F_{1,j}(t) &= \frac{1}{\Delta p_1 \Delta x} \Re \left\{ \frac{2iq}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x_j, t) \right. \\ &\quad \left. \times \left[-i \frac{\hbar}{\eta'} \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) + \hbar \pi \delta(\eta') \right] d\eta' dx \right\} \\ &= \frac{1}{\Delta p_1 \Delta x} \Re \left\{ \frac{2iq}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x_j, t) \right. \\ &\quad \left. \times \left(-i \frac{\hbar}{\eta'} \right) \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) d\eta' dx \right\} \\ &\quad + \frac{1}{\Delta p_1 \Delta x} \Re \left\{ 2iq\pi \int_{x_{j-1/2}}^{x_{j+1/2}} V(x, t) \tilde{f}(0, x_j, t) dx \right\} \\ &= \frac{1}{\Delta p_1 \Delta x} \Re \left\{ \frac{2q}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{-\infty}^{\infty} V \left(x + \frac{\eta'}{2}, t \right) \tilde{f}(\eta', x_j, t) \frac{\hbar}{\eta'} \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) d\eta' dx \right\}, \\ &\quad j = 1, \dots, N_x, \end{aligned} \quad (3.123)$$

where we used the previous result, Eq. (3.106), that $\tilde{f}(0, x_j, t)$ is a finite, real number, so that the last term in the second line vanishes. When repeating the calculation for $m = N_p$ an analogous result is obtained, the only changes are that $p_{3/2}$ is replaced by $p_{N_p-1/2}$ in Eq. (3.123) and an overall minus sign occurs since the half-bounded interval $(p_{N_p-1/2}, \infty)$ is now extended to ∞ and not $-\infty$. As can be seen from Eq. (3.123) the integration over the half-bounded interval results in a very similar expression like that for the interior points, Eq. (3.119). The only difference is that just one of the interval endpoints, namely the finite one, contributes and the term for the other endpoint drops out of the equation. This result facilitates the calculation of the drift matrix elements for a particular approximation remarkable. First of all, one does not need to do the integration over the semi-infinite intervals for each, new polynomial approximation order again, which may be tedious and complicated to do in some cases, and in the second place, since Eqs. (3.119) and (3.123) contain terms of the same form one effectively needs to carry out only one integration and the result can be used for all of the matrix elements.

After this general considerations we will calculate the integrals for a certain approximation of $f(p, x, t)$ and $V(x, t)$ explicitly. In both cases we choose a piecewise linear approximation. We thus write the first-order polynomials determining $f(p, x, t)$ by Eq. (3.101),

$$f(p, x, t) = \sum_{m=1}^{N_p} \sum_{j=1}^{N_x} P_{m,j}^1(p, x, t), \quad (3.124)$$

as

$$P_{m,j}^1(p, x, t) := \begin{cases} F_{m,j}(t) + \sigma_{m,j}(t)(p - p_m) & \text{if } (p, x) \in C_{m,j}, m \neq \{0, N_p + 1\}, \\ 0 & \text{elsewhere.} \end{cases} \quad (3.125)$$

It is obvious to see that an averaging of $f(p, x, t)$ over the grid cells $C_{m,j}$ results in $F_{m,j}(t)$, so that the constraint Eq. (3.103) is always fulfilled independent of the particular choice of the slopes $\sigma_{m,j}(t)$. In order to finally arrive at a linear set of equations as stated earlier in Eq. (3.100), we choose $\sigma_{m,j}(t)$ to depend linearly on the cell averages $F_{m,j}(t)$. In particular we determine $\sigma_{m,j}(t)$ by central finite differences at the interior grid points and by one-sided finite differences at the boundaries, such that

$$\begin{aligned} \sigma_{m,j}(t) &= \frac{F_{m+1,j}(t) - F_{m-1,j}(t)}{p_{m+1} - p_{m-1}}, & m = 2, \dots, N_p - 1, j = 1, \dots, N_x, \\ \sigma_{1,j}(t) &= \frac{F_{2,j}(t) - F_{1,j}(t)}{p_2 - p_1}, & j = 1, \dots, N_x, \\ \sigma_{N_p,j}(t) &= \frac{F_{N_p,j}(t) - F_{N_p-1,j}(t)}{p_{N_p} - p_{N_p-1}}, & j = 1, \dots, N_x, \end{aligned} \quad (3.126)$$

see also Fig. 3.6. A linear dependence is not obligatory at all, it just simplifies computations in practice since a simple and quick time stepping scheme can be used

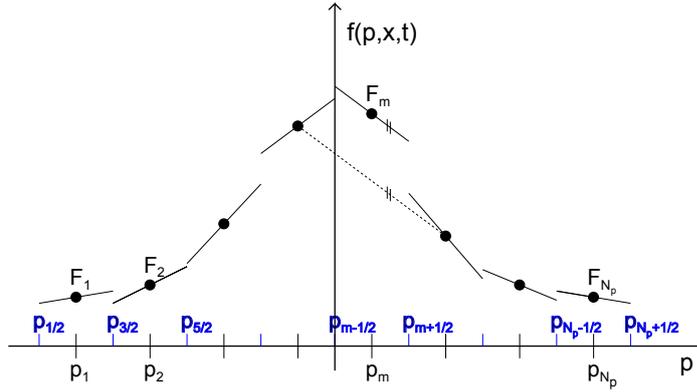


Figure 3.6: Illustration of the non-continuous, piecewise linear reconstruction of $f(p, x, t)$ out of the set of cell averages $F_{m,j}$ and with slopes as defined in Eq. (3.126). Indicated are the grid points p_m as well as the cell boundaries $p_{m+1/2}$.

and it furthermore reduces memory requirements because the linear set of equations can be rewritten as one matrix acting on the set of cell averages $F_{m,j}(t)$. But, one could also use a non-linear, more sophisticated method to determine the slopes, for instance with the purpose to suppress spurious oscillations. Various approaches exist for flux conservative methods which rely for example on different slope limiters, as can be found in [11] and is also briefly mentioned in Sec. 3.3. The slope limiters pursue the strategy to use a high-order approximation of the slopes in regions where the solution is well resolved and limit the slopes in badly resolved regions to avoid an overshooting of the reconstructed function. This strategy may also help to reduce spurious oscillations in the treatment of the WTE but was not examined in this work. The problem is that no general criterion as for total variation diminishing methods exists, see [18] [11], since the WTE inherently produces oscillations. Therefore, it is hard to distinguish real, physical oscillations from numerical ones. Because of this we think that applying such a strategy would require intensive investigations and is beyond the scope of the present work. But, in order not to restrict the calculations of the matrix elements to a special choice of $\sigma_{m,j}(t)$ we will do the calculations for both $F_{m,j}(t)$ and $\sigma_{m,j}(t)$ separately, yielding a different drift matrix for each of them. Only then, in the final step we make use of Eq. (3.126) to combine the two matrices to a single drift matrix and to arrive at a form as stated in Eq. (3.100).

The potential $V(x, t)$ is approximated by a piecewise linear polynomial as well. The general form was already stated in Eqs. (3.108) - (3.109). But, to increase the readability of the upcoming calculations we consider the case of a single, linear potential element for which the expression for the pseudo-differential operator was written down explicitly in Eq. (3.112). It is straightforward to make use of this expression for the cell-averaged formulas Eqs. (3.119) and (3.123).

Finally, to sum up the results of the last part of this section the cell averaged equations for this specific choice of $f(p, x, t)$ and $V(x, t)$ are listed. The following

equations serve as the starting point for the detailed calculations in the next parts of this section:

$$\begin{aligned}
 \partial_t F_{m,j}(t) &= \frac{1}{\Delta p_m \Delta x} \frac{2q}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left[v^0 + v^1 \left(x + \frac{\eta'}{2} \right) \right] \\
 &\quad \times \Re \left\{ \tilde{f}(\eta', x_j, t) \frac{\hbar}{\eta'} \left[\exp \left(i \frac{p_{m+1/2}}{\hbar} \eta' \right) - \exp \left(i \frac{p_{m-1/2}}{\hbar} \eta' \right) \right] \right\} d\eta' dx, \\
 m &= 2, \dots, N_p - 1, \quad j = 1, \dots, N_x, \\
 \partial_t F_{1,j}(t) &= \frac{1}{\Delta p_1 \Delta x} \frac{2q}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left[v^0 + v^1 \left(x + \frac{\eta'}{2} \right) \right] \\
 &\quad \times \Re \left\{ \tilde{f}(\eta', x_j, t) \frac{\hbar}{\eta'} \exp \left(i \frac{p_{3/2}}{\hbar} \eta' \right) \right\} d\eta' dx, \quad j = 1, \dots, N_x, \\
 \partial_t F_{N_p,j}(t) &= \frac{1}{\Delta p_{N_p} \Delta x} \frac{2q}{\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left[v^0 + v^1 \left(x + \frac{\eta'}{2} \right) \right] \\
 &\quad \times \Re \left\{ -\tilde{f}(\eta', x_j, t) \frac{\hbar}{\eta'} \exp \left(i \frac{p_{N_p-1/2}}{\hbar} \eta' \right) \right\} d\eta' dx, \quad j = 1, \dots, N_x,
 \end{aligned} \tag{3.127}$$

with the Fourier transform $\tilde{f}(\eta', x_j, t)$ on the interval $(x_{j-1/2}, x_{j+1/2})$ given by

$$\tilde{f}(\eta', x_j, t) = \frac{1}{2\pi\hbar} \sum_{i=1}^{N_p} \int_{p_{i-1/2}}^{p_{i+1/2}} [F_{i,j}(t) + \sigma_{i,j}(t)(p' - p_i)] \exp \left(-i \frac{p'}{\hbar} \eta' \right) dp', \tag{3.128}$$

according to Eq. (3.104). Both $V(x, t)$ and $f(p, x_j, t)$ consist of a constant and a linear part with respect to x and p , respectively. Each of the four combinations will be treated separately in the next parts of this section to finally arrive at an expression containing the four drift matrices

$$\begin{aligned}
 \partial_t F_{m,j}(t) &= \sum_{i=1}^{N_p} \left[D_{m,i,j}^{F,v^0} + D_{m,i,j}^{F,v^1} \right] F_{i,j}(t) + \sum_{i=1}^{N_p} \left[D_{m,i,j}^{\sigma,v^0} + D_{m,i,j}^{\sigma,v^1} \right] \sigma_{i,j}(t), \\
 m &= 1, \dots, N_p, \quad j = 1, \dots, N_x,
 \end{aligned} \tag{3.129}$$

with the matrix elements directly determined by the Eqs. (3.127) and (3.128). By the use of Eq. (3.126), connecting the slopes with the cell averages, the four matrices may then be merged into a single drift matrix

$$\partial_t F_{m,j}(t) = \sum_{i=1}^{N_p} D_{m,i,j} F_{i,j}(t), \quad m = 1, \dots, N_p, \quad j = 1, \dots, N_x. \tag{3.130}$$

The drift matrix, as do the four single matrices, obeys the property

$$\sum_{m=1}^{N_p} \Delta p_m D_{m,i,j} = 0, \quad i = 1, \dots, N_p, \quad j = 1, \dots, N_x, \tag{3.131}$$

which in turn ensures the conservation of the number of particles (comp. Eq. (3.117)). This property can also be directly seen from Eq. (3.127) when looking at the factors ± 1 in front of the exponential terms containing the cell boundaries $p_{m+1/2}$.

Constant potential segment, constant part of the polynomial for the Wigner function: We start with the simplest case, a piecewise constant approximation of $f(p, x, t)$ and $V(x, t)$. The first step is to calculate the Fourier transform of $f(p, x, t)$ which is simply given by

$$\begin{aligned}\tilde{f}(\eta', x_j, t) &= \frac{1}{2\pi\hbar} \sum_{i=1}^{N_p} \int_{p_{i-1/2}}^{p_{i+1/2}} F_{i,j}(t) \exp\left(-i\frac{p'}{\hbar}\eta'\right) dp' \\ &= \frac{1}{2\pi\hbar} \sum_{i=1}^{N_p} F_{i,j}(t) \left(i\frac{\hbar}{\eta'}\right) \left[\exp\left(-i\frac{p_{i+1/2}}{\hbar}\eta'\right) - \exp\left(-i\frac{p_{i-1/2}}{\hbar}\eta'\right)\right].\end{aligned}\tag{3.132}$$

Inserting this into Eq. (3.127) results in the following expression for the interior grid points

$$\begin{aligned}\partial_t F_{m,j}(t) &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi\hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \Re \left\{ \sum_{i=1}^{N_p} F_{i,j}(t) \frac{i}{\eta'^2} \left[\exp\left(-i\frac{p_{i+1/2}}{\hbar}\eta'\right) \right. \right. \\ &\quad \left. \left. - \exp\left(-i\frac{p_{i-1/2}}{\hbar}\eta'\right) \right] \left[\exp\left(i\frac{p_{m+1/2}}{\hbar}\eta'\right) - \exp\left(i\frac{p_{m-1/2}}{\hbar}\eta'\right) \right] \right\} d\eta' dx, \\ & m = 2, \dots, N_p - 1, \quad j = 1, \dots, N_x.\end{aligned}\tag{3.133}$$

For the boundary terms $m = 1$ and $m = N_p$ similar expressions are readily obtained. It is convenient to define terms $I_{m+1/2, i+1/2, j}^{F, v^0}$ containing all the integrals

$$I_{m+1/2, i+1/2, j}^{F, v^0} = - \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \frac{1}{\eta'^2} \sin\left(\frac{p_{m+1/2} - p_{i+1/2}}{\hbar}\eta'\right) d\eta' dx, \tag{3.134}$$

which result from taking the real part of the product of two exponential functions times i . This enables us to rewrite Eq. (3.133) in matrix form

$$\partial_t F_{m,j}(t) = \sum_{i=1}^{N_p} D_{m,i,j}^{F, v^0} F_{i,j}(t), \quad m = 1, \dots, N_p, \quad j = 1, \dots, N_x, \tag{3.135}$$

with the matrix elements given by

$$\begin{aligned}
 D_{m,i,j}^{F,v^0} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \left[I_{m+1/2,i+1/2,j}^{F,v^0} - I_{m+1/2,i-1/2,j}^{F,v^0} - \left(I_{m-1/2,i+1/2,j}^{F,v^0} - I_{m-1/2,i-1/2,j}^{F,v^0} \right) \right], \\
 &\quad m = 2, \dots, N_p - 1, \\
 D_{1,i,j}^{F,v^0} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \left(I_{3/2,i+1/2,j}^{F,v^0} - I_{3/2,i-1/2,j}^{F,v^0} \right), \\
 D_{N_p,i,j}^{F,v^0} &= -\frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \left(I_{N_p-1/2,i+1/2,j}^{F,v^0} - I_{N_p-1/2,i-1/2,j}^{F,v^0} \right),
 \end{aligned} \tag{3.136}$$

for $i = 1, \dots, N_p$, $j = 1, \dots, N_x$. As we will see in the following, the structure of all drift matrices in Eq. (3.129) is the same, only the terms $I_{m+1/2,i+1/2,j}^{F,v^0}$ change for a different approximation of $f(p, x, t)$ and $V(x, t)$. Due to the particular structure, a summation of the product of the matrix elements times Δp_m with respect to m , as done in Eq. (3.131), results in a telescoping sum whereby the remaining terms $m = 2$ and $m = N_p - 1$ cancel each other with the terms $m = 1$ and $m = N_p$, respectively, resulting on the whole in a value of zero.

Now, to explicitly calculate $I_{m+1/2,i+1/2,j}^{F,v^0}$ we use the abbreviation

$$k_{m,i} = \frac{p_{m+1/2} - p_{i+1/2}}{\hbar} \tag{3.137}$$

and start with the integral over η' :

$$\begin{aligned}
 I_{m+1/2,i+1/2,j}^{F,v^0} &= - \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \frac{1}{\eta'^2} \sin(k_{m,i}\eta') d\eta' dx \\
 &= \int_{x_{j-1/2}}^{x_{j+1/2}} \left[\frac{1}{\eta'} \sin(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} - k_{m,i} \int_{2(x_a-x)}^{2(x_b-x)} \frac{1}{\eta'} \cos(k_{m,i}\eta') d\eta' \right] dx \\
 &= \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{1}{\eta'} \sin(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} dx \\
 &\quad - \int_{x_{j-1/2}}^{x_{j+1/2}} k_{m,i} \{ \text{Ci} [|2k_{m,i}(x_b-x)|] - \text{Ci} [|2k_{m,i}(x_a-x)|] \} dx,
 \end{aligned} \tag{3.138}$$

where we integrated by parts in the second line and used the result of Eqs. (3.70) - (3.73) in the last line. To perform the integration with respect to x we begin with the sine integral

$$\begin{aligned}
 \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{1}{\eta'} \sin(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} dx &= -\frac{1}{2} \int_{2k_{m,i}(x_b-x_{j-1/2})}^{2k_{m,i}(x_b-x_{j+1/2})} \frac{1}{u'} \sin(u') du' \\
 &\quad + \frac{1}{2} \int_{2k_{m,i}(x_a-x_{j-1/2})}^{2k_{m,i}(x_a-x_{j+1/2})} \frac{1}{u''} \sin(u'') du'' \\
 &= -\frac{1}{2} \left[\text{Si}(u') \Big|_{2k_{m,i}(x_b-x_{j-1/2})}^{2k_{m,i}(x_b-x_{j+1/2})} - \text{Si}(u'') \Big|_{2k_{m,i}(x_a-x_{j-1/2})}^{2k_{m,i}(x_a-x_{j+1/2})} \right],
 \end{aligned} \tag{3.139}$$

where we used the substitutions $u' = 2k_{m,i}(x_b - x)$ and $u'' = 2k_{m,i}(x_a - x)$ and the definition of the sine integral Eq. (3.74). The integration of the cosine integral with respect to x is more involved due to the modulus of the argument. The calculation was done in detail in Sec. 3.2.2 and we can directly make use of the result from Eq. (3.82). When conflating the results from the sine and the cosine integral it is practical to introduce the function

$$\mathcal{T}^{F,v^0}(u) = \frac{1}{2} [u\text{Ci}(|u|) - \text{Si}(u) - \sin(u)] , \quad (3.140)$$

containing all the trigonometric integrals and functions, as well as the integration intervals

$$\begin{aligned} C_{m,i,j}^{x_a} &= (2k_{m,i}(x_a - x_{j-1/2}), 2k_{m,i}(x_a - x_{j+1/2})) , \\ C_{m,i,j}^{x_b} &= (2k_{m,i}(x_b - x_{j-1/2}), 2k_{m,i}(x_b - x_{j+1/2})) , \end{aligned} \quad (3.141)$$

to write the final result in compact notation as

$$I_{m+1/2,i+1/2,j}^{F,v^0} = \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_a}} . \quad (3.142)$$

Constant potential segment, linear part of the polynomial for the Wigner function: We now consider again a constant potential segment but for $f(p, x, t)$ solely the part of Eq. (3.128) containing $\sigma_{i,j}(t)$, i.e. the first-order correction. For the Fourier transform $\tilde{f}(\eta', x_j, t)$ we then find

$$\begin{aligned} \tilde{f}(\eta', x_j, t) &= \frac{1}{2\pi\hbar} \sum_{i=1}^{N_p} \int_{p_{i-1/2}}^{p_{i+1/2}} \sigma_{i,j}(t)(p' - p_i) \exp\left(-i\frac{p'}{\hbar}\eta'\right) dp' \\ &= \frac{1}{2\pi\hbar} \sum_{i=1}^{N_p} \sigma_{i,j}(t) \left(i\frac{\hbar}{\eta'}\right) \left\{ (p_{i+1/2} - p_i) \exp\left(-i\frac{p_{i+1/2}}{\hbar}\eta'\right) \right. \\ &\quad \left. - (p_{i-1/2} - p_i) \exp\left(-i\frac{p_{i-1/2}}{\hbar}\eta'\right) \right. \\ &\quad \left. - \left(i\frac{\hbar}{\eta'}\right) \left[\exp\left(-i\frac{p_{i+1/2}}{\hbar}\eta'\right) - \exp\left(-i\frac{p_{i-1/2}}{\hbar}\eta'\right) \right] \right\} \\ &= \frac{1}{2\pi\hbar} \sum_{i=1}^{N_p} \sigma_{i,j}(t) \left\{ \frac{\Delta p_i}{2} \left(i\frac{\hbar}{\eta'}\right) \left[\exp\left(-i\frac{p_{i+1/2}}{\hbar}\eta'\right) + \exp\left(-i\frac{p_{i-1/2}}{\hbar}\eta'\right) \right] \right. \\ &\quad \left. + \left(\frac{\hbar}{\eta'}\right)^2 \left[\exp\left(-i\frac{p_{i+1/2}}{\hbar}\eta'\right) - \exp\left(-i\frac{p_{i-1/2}}{\hbar}\eta'\right) \right] \right\} , \end{aligned} \quad (3.143)$$

when integrating by parts and making use of $\Delta p_i = p_{i+1/2} - p_{i-1/2}$. After inserting the expression for $\tilde{f}(\eta', x_j, t)$ into Eq. (3.127) for the interior grid points we arrive at

$$\begin{aligned}
 \partial_t F_{m,j}(t) &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \sum_{i=1}^{N_p} \sigma_{i,j}(t) \\
 &\quad \times \Re \left\{ \left\{ \frac{\Delta p_i}{2} \frac{i}{\eta'^2} \left[\exp\left(-i \frac{p_{i+1/2}}{\hbar} \eta'\right) + \exp\left(-i \frac{p_{i-1/2}}{\hbar} \eta'\right) \right] \right. \right. \\
 &\quad \left. \left. + \frac{\hbar}{\eta'^3} \left[\exp\left(-i \frac{p_{i+1/2}}{\hbar} \eta'\right) - \exp\left(-i \frac{p_{i-1/2}}{\hbar} \eta'\right) \right] \right\} \right. \\
 &\quad \left. \times \left[\exp\left(i \frac{p_{m+1/2}}{\hbar} \eta'\right) - \exp\left(i \frac{p_{m-1/2}}{\hbar} \eta'\right) \right] \right\} d\eta' dx, \\
 &\quad m = 2, \dots, N_p - 1, \quad j = 1, \dots, N_x.
 \end{aligned} \tag{3.144}$$

Again we evaluate the real part of the product of the exponential functions with $p_{m\pm 1/2}$ and $p_{i\pm 1/2}$ and summarize all the integrals in the two terms $I_{m+1/2,i^+,j}^{\sigma,v^0}$ and $I_{m+1/2,i^-,j}^{\sigma,v^0}$ given by

$$\begin{aligned}
 I_{m+1/2,i^+,j}^{\sigma,v^0} &= \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left[-\frac{\Delta p_i}{2} \frac{1}{\eta'^2} \sin\left(\frac{p_{m+1/2} - p_{i+1/2}}{\hbar} \eta'\right) \right. \\
 &\quad \left. + \frac{\hbar}{\eta'^3} \cos\left(\frac{p_{m+1/2} - p_{i+1/2}}{\hbar} \eta'\right) \right] d\eta' dx
 \end{aligned} \tag{3.145}$$

and

$$\begin{aligned}
 I_{m+1/2,i^-,j}^{\sigma,v^0} &= - \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left[-\frac{\Delta p_i}{2} \frac{1}{\eta'^2} \sin\left(\frac{p_{m+1/2} - p_{i-1/2}}{\hbar} \eta'\right) \right. \\
 &\quad \left. - \frac{\hbar}{\eta'^3} \cos\left(\frac{p_{m+1/2} - p_{i-1/2}}{\hbar} \eta'\right) \right] d\eta' dx.
 \end{aligned} \tag{3.146}$$

In the case of the zeroth-order approximation it suffices to introduce one term $I_{m+1/2,i+1/2,j}^{F,v^0}$ depending solely on the values of the cell boundaries but now we need two terms to account for the explicit i dependence. This dependence on the particular grid cell arises due to the factor $(p' - p_i)$ in Eq. (3.143). When calculating the integrals in $I_{m+1/2,i^+,j}^{\sigma,v^0}$ one can make use of the result obtained for $I_{m+1/2,i+1/2,j}^{F,v^0}$ due to the same

form of the integrals, so that one arrives at

$$\begin{aligned}
 I_{m+1/2,i^+,j}^{\sigma,v^0} &= \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left[-\frac{\Delta p_i}{2} \frac{1}{\eta'^2} \sin(k_{m,i}\eta') + \frac{\hbar}{\eta'^3} \cos(k_{m,i}\eta') \right] d\eta' dx \\
 &= \frac{\Delta p_i}{2} I_{m+1/2,i+1/2,j}^{F,v^0} - \frac{\hbar}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} \left[\frac{1}{\eta'^2} \cos(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} \right. \\
 &\quad \left. + k_{m,i} \int_{2(x_a-x)}^{2(x_b-x)} \frac{1}{\eta'^2} \sin(k_{m,i}\eta') d\eta' \right] dx \\
 &= \left(\frac{\Delta p_i}{2} + \frac{\hbar}{2} k_{m,i} \right) I_{m+1/2,i+1/2,j}^{F,v^0} - \frac{\hbar}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{1}{\eta'^2} \cos(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} dx,
 \end{aligned} \tag{3.147}$$

with the abbreviation $k_{m,i}$ as defined by Eq. (3.137). To integrate the last term we apply the more compact notation with the intervals $C_{m,i,j}^{x_a}$ and $C_{m,i,j}^{x_b}$, see Eq. (3.141):

$$\begin{aligned}
 -\frac{\hbar}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{1}{\eta'^2} \cos(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} dx &= \frac{\hbar k_{m,i}}{4} \left[\int_{C_{m,i,j}^{x_b}} \frac{1}{u'^2} \cos(u') du' \right. \\
 &\quad \left. - \int_{C_{m,i,j}^{x_a}} \frac{1}{u''^2} \cos(u'') du'' \right],
 \end{aligned} \tag{3.148}$$

when substituting $u' = 2k_{m,i}(x_b - x)$ and $u'' = 2k_{m,i}(x_a - x)$. In the case of $C_{m,i,j}^{x_b}$ we find by taking Eq. (3.74) into account

$$\begin{aligned}
 \frac{\hbar k_{m,i}}{4} \int_{C_{m,i,j}^{x_b}} \frac{1}{u'^2} \cos(u') du' &= \frac{\hbar k_{m,i}}{4} \left[-\frac{1}{u'} \cos(u') \Big|_{C_{m,i,j}^{x_b}} - \int_{C_{m,i,j}^{x_b}} \frac{1}{u'} \sin(u') du' \right] \\
 &= -\frac{\hbar k_{m,i}}{4} \left[\frac{1}{u'} \cos(u') \Big|_{C_{m,i,j}^{x_b}} + \text{Si}(u') \Big|_{C_{m,i,j}^{x_b}} \right],
 \end{aligned} \tag{3.149}$$

so that it is convenient to introduce the new function

$$\mathcal{T}^{\sigma,v^0}(u) = \frac{1}{4} \left[\frac{1}{u} \cos(u) + \text{Si}(u) \right]. \tag{3.150}$$

Before writing down the result for $I_{m+1/2,i^+,j}^{\sigma,v^0}$ we perform the following reshaping of the prefactor of $I_{m+1/2,i+1/2,j}^{F,v^0}$ in Eq. (3.147)

$$\begin{aligned}
 \frac{\Delta p_i}{2} + \frac{\hbar}{2} k_{m,i} &= p_{i+1/2} - p_i + \frac{1}{2}(p_{m+1/2} - p_{i+1/2}) \\
 &= \frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i,
 \end{aligned} \tag{3.151}$$

to finally arrive at

$$I_{m+1/2,i^+,j}^{\sigma,v^0} = \left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \left[\mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_a}} \right] - (p_{m+1/2} - p_{i+1/2}) \left[\mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_a}} \right]. \quad (3.152)$$

In the case of $I_{m+1/2,i^-,j}^{\sigma,v^0}$ we find nearly the same terms as in Eq. (3.147), only that a minus sign appears in front of $\frac{\Delta p_i}{2}$ and $k_{m,i}$ is replaced by $k_{m,i-1}$. As done in Eq. (3.151) we reshape the prefactors

$$\begin{aligned} -\frac{\Delta p_i}{2} + \frac{\hbar}{2} k_{m,i-1} &= p_{i-1/2} - p_i + \frac{1}{2}(p_{m+1/2} - p_{i-1/2}) \\ &= \frac{p_{m+1/2} + p_{i-1/2}}{2} - p_i \end{aligned} \quad (3.153)$$

and since the other calculational steps are identical we can directly write down the result as

$$I_{m+1/2,i^-,j}^{\sigma,v^0} = \left(\frac{p_{m+1/2} + p_{i-1/2}}{2} - p_i \right) \left[\mathcal{T}^{F,v^0}(u)|_{C_{m,i-1,j}^{x_b}} - \mathcal{T}^{F,v^0}(u)|_{C_{m,i-1,j}^{x_a}} \right] - (p_{m+1/2} - p_{i-1/2}) \left[\mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i-1,j}^{x_b}} - \mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i-1,j}^{x_a}} \right]. \quad (3.154)$$

With this in hands we can use Eq. (3.144) and the analogous ones for $m = 1$ and $m = N_p$ to express the matrix elements corresponding to Eq. (3.129) in terms of $I_{m+1/2,i^+,j}^{\sigma,v^0}$ and $I_{m+1/2,i^-,j}^{\sigma,v^0}$:

$$\begin{aligned} D_{m,i,j}^{\sigma,v^0} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \left[I_{m+1/2,i^+,j}^{\sigma,v^0} - I_{m+1/2,i^-,j}^{\sigma,v^0} - \left(I_{m-1/2,i^+,j}^{\sigma,v^0} - I_{m-1/2,i^-,j}^{\sigma,v^0} \right) \right], \\ &\quad m = 2, \dots, N_p - 1, \\ D_{1,i,j}^{\sigma,v^0} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \left(I_{3/2,i^+,j}^{\sigma,v^0} - I_{3/2,i^-,j}^{\sigma,v^0} \right), \\ D_{N_p,i,j}^{\sigma,v^0} &= -\frac{\hbar}{\Delta p_m \Delta x} \frac{qv^0}{\pi \hbar} \left(I_{N_p-1/2,i^+,j}^{\sigma,v^0} - I_{N_p-1/2,i^-,j}^{\sigma,v^0} \right), \end{aligned} \quad (3.155)$$

with (i, j) taking on the values $i = 1, \dots, N_p$, $j = 1, \dots, N_x$.

Linear potential segment, constant part of the polynomial for the Wigner function: To include a linear potential element we start again with the simplest case, a piecewise constant approximation of $f(p, x, t)$, for which the Fourier transform is

known from Eq. (3.132). We now consider the following part of Eq. (3.127)

$$\begin{aligned} \partial_t F_{m,j}(t) &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left(x + \frac{\eta'}{2}\right) \Re \left\{ \sum_{i=1}^{N_p} F_{i,j}(t) \frac{i}{\eta'^2} \left[\exp\left(-i \frac{p_{i+1/2}}{\hbar} \eta'\right) \right. \right. \\ &\quad \left. \left. - \exp\left(-i \frac{p_{i-1/2}}{\hbar} \eta'\right) \right] \left[\exp\left(i \frac{p_{m+1/2}}{\hbar} \eta'\right) - \exp\left(i \frac{p_{m-1/2}}{\hbar} \eta'\right) \right] \right\} d\eta' dx, \\ &\quad m = 2, \dots, N_p - 1, \quad j = 1, \dots, N_x. \end{aligned} \quad (3.156)$$

Once more, we collect all the integrals in the expression

$$I_{m+1/2, i+1/2, j}^{F, v^1} = - \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left(x + \frac{\eta'}{2}\right) \frac{1}{\eta'^2} \sin\left(\frac{p_{m+1/2} - p_{i+1/2}}{\hbar} \eta'\right) d\eta' dx \quad (3.157)$$

and can make direct use of the results from the previous part of this section, *Constant potential segment, constant part of the polynomial for the Wigner function*, to perform the integration with respect to η' :

$$\begin{aligned} I_{m+1/2, i+1/2, j}^{F, v^1} &= \int_{x_{j-1/2}}^{x_{j+1/2}} x \left\{ \frac{1}{\eta'} \sin(k_{m,i} \eta') \Big|_{2(x_a-x)}^{2(x_b-x)} \right. \\ &\quad \left. - k_{m,i} \{ \text{Ci} [|2k_{m,i}(x_b - x)|] - \text{Ci} [|2k_{m,i}(x_a - x)|] \} \right\} dx \\ &\quad - \frac{1}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} \{ \text{Si} [2k_{m,i}(x_b - x)] - \text{Si} [2k_{m,i}(x_a - x)] \} dx, \end{aligned} \quad (3.158)$$

where we used for the term proportional to x the result of Eq. (3.138) and the other term can be directly identified as a sine integral, Eq. (3.74). For the integration with respect to x we split up the integral in the three distinct parts and consider in each one only the terms containing x_b , so that the first integral, which we need to solve, is

$$\begin{aligned} I_1 &= \int_{x_{j-1/2}}^{x_{j+1/2}} x \frac{1}{2(x_b - x)} \sin(2k_{m,i}(x_b - x)) dx \\ &= -\frac{1}{2} \int_{C_{m,i,j}^{x_b}} \left(x_b - \frac{u}{2k_{m,i}}\right) \frac{1}{u} \sin(u) du \\ &= -\frac{x_b}{2} \int_{C_{m,i,j}^{x_b}} \frac{1}{u} \sin(u) du + \frac{1}{4k_{m,i}} \int_{C_{m,i,j}^{x_b}} \sin(u) du \\ &= -\frac{1}{2} \left[x_b \text{Si}(u) \Big|_{C_{m,i,j}^{x_b}} + \frac{1}{2k_{m,i}} \cos(u) \Big|_{C_{m,i,j}^{x_b}} \right], \end{aligned} \quad (3.159)$$

when substituting with $u = 2k_{m,i}(x_b - x)$. The second part with the cosine integral as integrand was already evaluated in Sec. 3.2.2 in all detail and we recap here only

the result from Eq. (3.87):

$$\begin{aligned}
 I_2 &= - \int_{x_{j-1/2}}^{x_{j+1/2}} k_{m,i} x \text{Ci} [|2k_{m,i}(x_b - x)|] dx \\
 &= \frac{x_b}{2} [u \text{Ci}(|u|) - \sin(u)] \Big|_{C_{m,i,j}^{x_b}} + \frac{1}{8k_{m,i}} [-u^2 \text{Ci}(|u|) + u \sin(u) + \cos(u)] \Big|_{C_{m,i,j}^{x_b}} .
 \end{aligned} \tag{3.160}$$

For the last part we find

$$\begin{aligned}
 I_3 &= -\frac{1}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} \text{Si} [2k_{m,i}(x_b - x)] dx \\
 &= \frac{1}{4k_{m,i}} \int_{C_{m,i,j}^{x_b}} \text{Si}(u) du \\
 &= \frac{1}{4k_{m,i}} [u \text{Si}(u) + \cos(u)] \Big|_{C_{m,i,j}^{x_b}} ,
 \end{aligned} \tag{3.161}$$

when substituting again $u = 2k_{m,i}(x_b - x)$ and making use of the anti-derivative of the sine integral, Eq. (3.76). Summing up all three parts yields

$$\begin{aligned}
 \sum_{i=1}^3 I_i &= \frac{x_b}{2} [u \text{Ci}(|u|) - \text{Si}(u) - \sin(u)] \Big|_{C_{m,i,j}^{x_b}} \\
 &\quad + \frac{1}{8k_{m,i}} [-u^2 \text{Ci}(|u|) + 2u \text{Si}(u) + u \sin(u) + \cos(u)] \Big|_{C_{m,i,j}^{x_b}} ,
 \end{aligned} \tag{3.162}$$

where the first terms with x_b as prefactor correspond to $\mathcal{T}^{F,v^0}(u)$, Eq. (3.140), and the new trigonometric terms may be summarized in the function

$$\mathcal{T}^{F,v^1}(u) = \frac{1}{8} [-u^2 \text{Ci}(|u|) + 2u \text{Si}(u) + u \sin(u) + \cos(u)] . \tag{3.163}$$

With these definitions we finally arrive at

$$\begin{aligned}
 I_{m+1/2,i+1/2,j}^{F,v^1} &= x_b \mathcal{T}^{F,v^0}(u) \Big|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{F,v^0}(u) \Big|_{C_{m,i,j}^{x_a}} \\
 &\quad + \frac{1}{k_{m,i}} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_a}} \right] ,
 \end{aligned} \tag{3.164}$$

which determine the matrix elements corresponding to Eq. (3.129) via the relation

$$\begin{aligned}
 D_{m,i,j}^{F,v^1} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \left[I_{m+1/2,i+1/2,j}^{F,v^1} - I_{m+1/2,i-1/2,j}^{F,v^1} - \left(I_{m-1/2,i+1/2,j}^{F,v^1} - I_{m-1/2,i-1/2,j}^{F,v^1} \right) \right] , \\
 &\quad m = 2, \dots, N_p - 1 , \\
 D_{1,i,j}^{F,v^1} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \left(I_{3/2,i+1/2,j}^{F,v^1} - I_{3/2,i-1/2,j}^{F,v^1} \right) , \\
 D_{N_p,i,j}^{F,v^1} &= -\frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \left(I_{N_p-1/2,i+1/2,j}^{F,v^1} - I_{N_p-1/2,i-1/2,j}^{F,v^1} \right) ,
 \end{aligned} \tag{3.165}$$

with (i, j) taking on the values $i = 1, \dots, N_p$, $j = 1, \dots, N_x$.

Linear potential segment, linear part of the polynomial for the Wigner function:

The Fourier transform $\tilde{f}(\eta', x_j, t)$ is already known from Eq. (3.143) and we now consider the following part of Eq. (3.127)

$$\begin{aligned}
 \partial_t F_{m,j}(t) &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \sum_{i=1}^{N_p} \sigma_{i,j}(t) \left(x + \frac{\eta'}{2}\right) \\
 &\quad \times \Re \left\{ \left[\frac{\Delta p_i}{2} \frac{i}{\eta'^2} \left[\exp\left(-i \frac{p_{i+1/2}}{\hbar} \eta'\right) + \exp\left(-i \frac{p_{i-1/2}}{\hbar} \eta'\right) \right] \right. \right. \\
 &\quad \left. \left. + \frac{\hbar}{\eta'^3} \left[\exp\left(-i \frac{p_{i+1/2}}{\hbar} \eta'\right) - \exp\left(-i \frac{p_{i-1/2}}{\hbar} \eta'\right) \right] \right] \right\} \\
 &\quad \times \left[\exp\left(i \frac{p_{m+1/2}}{\hbar} \eta'\right) - \exp\left(i \frac{p_{m-1/2}}{\hbar} \eta'\right) \right] \Bigg\} d\eta' dx, \\
 m &= 2, \dots, N_p - 1, \quad j = 1, \dots, N_x,
 \end{aligned} \tag{3.166}$$

so that it is convenient to define

$$\begin{aligned}
 I_{m+1/2, i^+, j}^{\sigma, v^1} &= \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left(x + \frac{\eta'}{2}\right) \left[-\frac{\Delta p_i}{2} \frac{1}{\eta'^2} \sin\left(\frac{p_{m+1/2} - p_{i+1/2}}{\hbar} \eta'\right) \right. \\
 &\quad \left. + \frac{\hbar}{\eta'^3} \cos\left(\frac{p_{m+1/2} - p_{i+1/2}}{\hbar} \eta'\right) \right] d\eta' dx
 \end{aligned} \tag{3.167}$$

and

$$\begin{aligned}
 I_{m+1/2, i^-, j}^{\sigma, v^1} &= - \int_{x_{j-1/2}}^{x_{j+1/2}} \int_{2(x_a-x)}^{2(x_b-x)} \left(x + \frac{\eta'}{2}\right) \left[-\frac{\Delta p_i}{2} \frac{1}{\eta'^2} \sin\left(\frac{p_{m+1/2} - p_{i-1/2}}{\hbar} \eta'\right) \right. \\
 &\quad \left. - \frac{\hbar}{\eta'^3} \cos\left(\frac{p_{m+1/2} - p_{i-1/2}}{\hbar} \eta'\right) \right] d\eta' dx.
 \end{aligned} \tag{3.168}$$

We perform the explicit calculation for $I_{m+1/2, i^+, j}^{\sigma, v^1}$ and start with the integration with respect to η' . To make use of previous results it is advisable to integrate the cosine

term by parts as a first step,

$$\begin{aligned}
 I_0 &= \int_{2(x_a-x)}^{2(x_b-x)} \left(x + \frac{\eta'}{2} \right) \frac{\hbar}{\eta'^3} \cos(k_{m,i}\eta') d\eta' \\
 &= x \frac{\hbar}{2} \left[-\frac{1}{\eta'^2} \cos(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} - k_{m,i} \int_{2(x_a-x)}^{2(x_b-x)} \frac{1}{\eta'^2} \sin(k_{m,i}\eta') d\eta' \right] \\
 &\quad + \frac{\hbar}{2} \left[-\frac{1}{\eta'} \cos(k_{m,i}\eta') \Big|_{2(x_a-x)}^{2(x_b-x)} - k_{m,i} \int_{2(x_a-x)}^{2(x_b-x)} \frac{1}{\eta'} \sin(k_{m,i}\eta') d\eta' \right] \\
 &= -\frac{\hbar}{2} k_{m,i} \int_{2(x_a-x)}^{2(x_b-x)} \left(x + \frac{\eta'}{2} \right) \frac{1}{\eta'^2} \sin(k_{m,i}\eta') d\eta' \\
 &\quad - \frac{\hbar}{2} \left[x \frac{1}{\eta'^2} \cos(k_{m,i}\eta') + \frac{1}{\eta'} \cos(k_{m,i}\eta') + \frac{k_{m,i}}{2} \text{Si}(k_{m,i}\eta') \right] \Big|_{2(x_a-x)}^{2(x_b-x)}. \quad (3.169)
 \end{aligned}$$

The first term appears in Eq. (3.167) as well and the integrals with respect to η' and x were already solved before in Eq. (3.157) when calculating $I_{m+1/2,i+1/2,j}^{F,v^1}$. The term in Eq. (3.169) with the sine integral was integrated for I_3 in Eq. (3.161), so that just two integrals are left to be solved which we will do explicitly only for the upper bound $2(x_b - x)$,

$$\begin{aligned}
 I_4 &= -\frac{\hbar}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} x \frac{1}{[2(x_b-x)]^2} \cos[2k_{m,i}(x_b-x)] dx \\
 &= \frac{\hbar}{4} \int_{C_{m,i,j}^{x_b}} \left(k_{m,i}x_b - \frac{u}{2} \right) \frac{1}{u^2} \cos(u) du \\
 &= \frac{\hbar}{4} \left\{ k_{m,i}x_b \left[-\frac{1}{u} \cos(u) - \text{Si}(u) \right] \Big|_{C_{m,i,j}^{x_b}} - \frac{1}{2} \text{Ci}(|u|) \Big|_{C_{m,i,j}^{x_b}} \right\} \quad (3.170)
 \end{aligned}$$

and

$$\begin{aligned}
 I_5 &= -\frac{\hbar}{2} \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{1}{2(x_b-x)} \cos[2k_{m,i}(x_b-x)] dx \\
 &= \frac{\hbar}{4} \text{Ci}(|u|) \Big|_{C_{m,i,j}^{x_b}}, \quad (3.171)
 \end{aligned}$$

where we substituted in both cases $u = 2k_{m,i}(x_b - x)$ and used the definitions of the sine and cosine integrals from Sec. 3.2.2. Summing up the terms I_3 with the proper prefactor, I_4 and I_5 yields

$$\begin{aligned}
 \frac{\hbar k_{m,i}}{2} I_3 + I_4 + I_5 &= -\hbar k_{m,i} x_b \frac{1}{4} \left[\frac{1}{u} \cos(u) + \text{Si}(u) \right] \Big|_{C_{m,i,j}^{x_b}} \\
 &\quad + \hbar \frac{1}{8} [\text{Ci}(|u|) + u \text{Si}(u) + \cos(u)] \Big|_{C_{m,i,j}^{x_b}}, \quad (3.172)
 \end{aligned}$$

with the trigonometric functions in the first square bracket already defined as the function $\mathcal{T}^{\sigma,v^0}(u)$ and the new terms in the second bracket make it obvious to introduce

$$\mathcal{T}^{\sigma,v^1}(u) = \frac{1}{8} [\text{Ci}(|u|) + u\text{Si}(u) + \cos(u)] . \quad (3.173)$$

With this in hands we can now write the result for $I_{m+1/2,i^+,j}^{\sigma,v^1}$ in the following form

$$\begin{aligned} I_{m+1/2,i^+,j}^{\sigma,v^1} &= \left(\frac{\Delta p_i}{2} + \frac{\hbar}{2} k_{m,i} \right) I_{m+1/2,i+1/2,j}^{F,v^1} + \hbar \left[\mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i,j}^{x_a}} \right] \\ &\quad - \hbar k_{m,i} \left[x_b \mathcal{T}^{\sigma,v^0}(u) \Big|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{\sigma,v^0}(u) \Big|_{C_{m,i,j}^{x_a}} \right] , \end{aligned} \quad (3.174)$$

and furthermore, when using Eq. (3.151) to reshape the prefactor as well as Eq. (3.164) for $I_{m+1/2,i+1/2,j}^{F,v^1}$, we arrive at

$$\begin{aligned} I_{m+1/2,i^+,j}^{\sigma,v^1} &= \left(\frac{p_{m+1/2} + p_{i+1/2} - p_i}{2} \right) \left\{ x_b \mathcal{T}^{F,v^0}(u) \Big|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{F,v^0}(u) \Big|_{C_{m,i,j}^{x_a}} \right. \\ &\quad \left. + \frac{1}{k_{m,i}} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_a}} \right] \right\} \\ &\quad - \hbar k_{m,i} \left[x_b \mathcal{T}^{\sigma,v^0}(u) \Big|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{\sigma,v^0}(u) \Big|_{C_{m,i,j}^{x_a}} \right] \\ &\quad + \hbar \left[\mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i,j}^{x_a}} \right] . \end{aligned} \quad (3.175)$$

In the case of $I_{m+1/2,i^-,j}^{\sigma,v^1}$ we have a change of sign in the prefactor $\frac{\Delta p_i}{2}$, so that we have to use Eq. (3.153) to reshape the prefactor and for the rest of the calculation we can use the same results. Consequently we finally find

$$\begin{aligned} I_{m+1/2,i^-,j}^{\sigma,v^1} &= \left(\frac{p_{m+1/2} + p_{i-1/2} - p_i}{2} \right) \left\{ x_b \mathcal{T}^{F,v^0}(u) \Big|_{C_{m,i-1,j}^{x_b}} - x_a \mathcal{T}^{F,v^0}(u) \Big|_{C_{m,i-1,j}^{x_a}} \right. \\ &\quad \left. + \frac{1}{k_{m,i-1}} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i-1,j}^{x_b}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i-1,j}^{x_a}} \right] \right\} \\ &\quad - \hbar k_{m,i-1} \left[x_b \mathcal{T}^{\sigma,v^0}(u) \Big|_{C_{m,i-1,j}^{x_b}} - x_a \mathcal{T}^{\sigma,v^0}(u) \Big|_{C_{m,i-1,j}^{x_a}} \right] \\ &\quad + \hbar \left[\mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i-1,j}^{x_b}} - \mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i-1,j}^{x_a}} \right] . \end{aligned} \quad (3.176)$$

The matrix elements of Eq. (3.129) and corresponding to Eq. (3.166) are determined by the relation

$$\begin{aligned} D_{m,i,j}^{\sigma,v^1} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \left[I_{m+1/2,i^+,j}^{\sigma,v^1} - I_{m+1/2,i^-,j}^{\sigma,v^1} - \left(I_{m-1/2,i^+,j}^{\sigma,v^1} - I_{m-1/2,i^-,j}^{\sigma,v^1} \right) \right] , \\ &\quad m = 2, \dots, N_p - 1 , \\ D_{1,i,j}^{\sigma,v^1} &= \frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \left(I_{3/2,i^+,j}^{\sigma,v^1} - I_{3/2,i^-,j}^{\sigma,v^1} \right) , \\ D_{N_p,i,j}^{\sigma,v^1} &= -\frac{\hbar}{\Delta p_m \Delta x} \frac{qv^1}{\pi \hbar} \left(I_{N_p-1/2,i^+,j}^{\sigma,v^1} - I_{N_p-1/2,i^-,j}^{\sigma,v^1} \right) , \end{aligned} \quad (3.177)$$

with (i, j) taking on the values $i = 1, \dots, N_p$, $j = 1, \dots, N_x$.

Summary of the equations, diagonal elements and inclusion of particular potential shapes: The following list gives an overview of the trigonometric functions defined in the calculations on the last pages, with the labeling $F/\sigma, v^0/v^1$ referring to their appearance in a certain approximation:

$$\begin{aligned}
 \mathcal{T}^{F,v^0}(u) &= \frac{1}{2} [u\text{Ci}(|u|) - \text{Si}(u) - \sin(u)] , \\
 \mathcal{T}^{\sigma,v^0}(u) &= \frac{1}{4} \left[\frac{1}{u} \cos(u) + \text{Si}(u) \right] , \\
 \mathcal{T}^{F,v^1}(u) &= \frac{1}{8} [-u^2\text{Ci}(|u|) + 2u\text{Si}(u) + u\sin(u) + \cos(u)] , \\
 \mathcal{T}^{\sigma,v^1}(u) &= \frac{1}{8} [\text{Ci}(|u|) + u\text{Si}(u) + \cos(u)] .
 \end{aligned} \tag{3.178}$$

These functions enable us to write the terms $I_{m+1/2,i+1/2,j}^{F,v^0/1}$ and $I_{m+1/2,i^+,j}^{\sigma,v^0/1}$ in compact notation as

$$\begin{aligned}
 I_{m+1/2,i+1/2,j}^{F,v^0} &= \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_a}} , \\
 I_{m+1/2,i+1/2,j}^{F,v^1} &= x_b \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_a}} \\
 &\quad + \frac{1}{k_{m,i}} \left[\mathcal{T}^{F,v^1}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^1}(u)|_{C_{m,i,j}^{x_a}} \right] , \\
 I_{m+1/2,i^+,j}^{\sigma,v^0} &= \left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \left[\mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_a}} \right] \\
 &\quad - \hbar k_{m,i} \left[\mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_a}} \right] , \\
 I_{m+1/2,i^+,j}^{\sigma,v^1} &= \left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \left\{ x_b \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{F,v^0}(u)|_{C_{m,i,j}^{x_a}} \right. \\
 &\quad \left. + \frac{1}{k_{m,i}} \left[\mathcal{T}^{F,v^1}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{F,v^1}(u)|_{C_{m,i,j}^{x_a}} \right] \right\} \\
 &\quad - \hbar k_{m,i} \left[x_b \mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_b}} - x_a \mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_a}} \right] \\
 &\quad + \hbar \left[\mathcal{T}^{\sigma,v^1}(u)|_{C_{m,i,j}^{x_b}} - \mathcal{T}^{\sigma,v^1}(u)|_{C_{m,i,j}^{x_a}} \right] ,
 \end{aligned} \tag{3.179}$$

with the intervals defined by

$$\begin{aligned}
 C_{m,i,j}^{x_a} &= (2k_{m,i}(x_a - x_{j-1/2}), 2k_{m,i}(x_a - x_{j+1/2})) , \\
 C_{m,i,j}^{x_b} &= (2k_{m,i}(x_b - x_{j-1/2}), 2k_{m,i}(x_b - x_{j+1/2})) ,
 \end{aligned} \tag{3.180}$$

and the abbreviation

$$k_{m,i} = \frac{p_{m+1/2} - p_{i+1/2}}{\hbar} . \tag{3.181}$$

The terms for i^- are listed explicitly in the corresponding parts of this section and are not recapped here, but can be readily obtained when replacing $p_{i+1/2}$ by $p_{i-1/2}$.

When looking at Eq. (3.178) one may notice that the functions have a well defined parity which depends only on the index $v^{0/1}$. In the case of v^0 all functions are odd. This fact, together with the way the functions enter the terms $I_{m+1/2,i+1/2,j}^{F,v^0}$ and $I_{m+1/2,i^+,j}^{\sigma,v^0}$ as a direct difference without any prefactors, make it obvious to see that the action of a square potential barrier vanishes when the cell-point x_j is placed exactly in the middle of (x_a, x_b) . When using the abbreviations d_1 and d_2 for the two distances

$$\begin{aligned} d_1 &= (x_a - x_{j-1/2}) = -(x_b - x_{j+1/2}), \\ d_2 &= (x_a - x_{j+1/2}) = -(x_b - x_{j-1/2}), \end{aligned} \quad (3.182)$$

it is apparent that any expression of the form

$$\mathcal{T}^{F/\sigma,v^0}(u) \Big|_{d_1}^{d_2} - \mathcal{T}^{F/\sigma,v^0}(u) \Big|_{-d_2}^{-d_1} = 0 \quad (3.183)$$

vanishes for an odd function $\mathcal{T}^{F/\sigma,v^0}(u)$. This fact can be used to check the consistency of the derived equations since the property can be directly seen from the definition of the pseudo-differential operator, Eq. (2.12). It is interesting to note that the action of the pseudo-differential operator vanishes in general when x coincides with the center of a symmetric potential shape.

For the implementation in practice it is important to analyze the behaviour of the terms in Eq. (3.179) for small arguments since singular functions are present in some cases. Earlier, for Eq. (3.107) we showed that all the integral expressions exist since $\tilde{f}(\eta, x, t)$ stays finite for $\eta \rightarrow 0$ (see Eq. (3.106)). Due to this it is clear that the matrix elements $D_{m,i,j}^{F/\sigma,v^0/v^1}$ are well defined in all cases. But despite of this, one may encounter very large values for the single terms in Eq. (3.179) which are subtracted from each other. In particular we will examine the case $k_{m,m} = 0$ since a value of zero does definitely occur in the present formalism. The nodes for the potential defined in Eqs. (3.108) - (3.109) are chosen to coincide with grid points x_j of the Wigner function. Therefore, since the potential nodes x_a and x_b (or more general x_k^V) are defined on grid points and the cell boundaries $x_{j-1/2}$ on intermediate points, the terms $x_{a/b} - x_{j\pm 1/2}$ are never zero but take on a minimum value of $\Delta x/2$. One could assume that numerical problems due to subtractive cancellation are encountered for small values of Δx but actual simulations did not reveal any problems, not even in the case of tests done with very small values of Δx .

We begin with examining the behaviour of the functions $\mathcal{T}^{F/\sigma,v^0/v^1}(u)$ in the limit $u \rightarrow 0$. From Eq. (3.81) we know that

$$\lim_{u \rightarrow 0} u \text{Ci}(|u|) = 0, \quad (3.184)$$

so that we find

$$\begin{aligned} \lim_{u \rightarrow 0} \mathcal{T}^{F,v^0}(u) &= 0, \\ \lim_{u \rightarrow 0} \mathcal{T}^{F,v^1}(u) &= \frac{1}{8} \end{aligned} \quad (3.185)$$

and the other two functions, $\mathcal{T}^{\sigma,v^0}(u)$ and $\mathcal{T}^{\sigma,v^1}(u)$ diverge for $u \rightarrow 0$. Therefore, the expression for $I_{m+1/2,i+1/2,j}^{F,v^0}$ does not cause any problems but the other ones have to be looked at in more detail. In the case of $I_{m+1/2,i+1/2,j}^{F,v^1}$ the term with $1/k_{m,i}$ needs to be examined, see Eq. (3.179). When using k instead of $k_{m,i}$ together with d_1 and d_2 labeling two distances, we find

$$\begin{aligned} \lim_{k \rightarrow 0} \frac{1}{k} \mathcal{T}^{F,v^1}(u) \Big|_{kd_1}^{kd_2} &= \lim_{k \rightarrow 0} \frac{1}{8k} [-u^2 \text{Ci}(|u|) + 2u \text{Si}(u) + u \sin(u) + \cos(u)] \Big|_{kd_1}^{kd_2} \\ &= \lim_{k \rightarrow 0} \frac{1}{8k} [\cos(kd_2) - \cos(kd_1)] \\ &= \lim_{k \rightarrow 0} \frac{1}{8k} \left[1 - \frac{(kd_2)^2}{2} - 1 + \frac{(kd_1)^2}{2} + \mathcal{O}(k^4) \right] \\ &= 0. \end{aligned} \quad (3.186)$$

As a result, for both terms determining the matrix elements of $D_{m,i,j}^{F,v^0/v^1}$ we find for $m = i$:

$$\begin{aligned} I_{m+1/2,m+1/2,j}^{F,v^0} &= 0, \\ I_{m+1/2,m+1/2,j}^{F,v^1} &= 0. \end{aligned} \quad (3.187)$$

Furthermore, since in the expression for the matrix elements $D_{m,i,j}^{F,v^0/v^1}$ the terms $I_{m+1/2,i-1/2,j}^{F,v^0/v^1}$ and $I_{m-1/2,i+1/2,j}^{F,v^0/v^1}$, which are odd functions with respect to $k_{m,i}$, enter with the same sign (comp. Eqs. (3.136) and (3.165)). They cancel each other for $m = i$ and one obtains for the diagonal elements of the drift matrices for $F_{m,j}$

$$\begin{aligned} D_{m,m,j}^{F,v^0} &= 0, \\ D_{m,m,j}^{F,v^1} &= 0. \end{aligned} \quad (3.188)$$

As we will see shortly this is not true in the case of the drift matrices for $\sigma_{m,j}$, so that

$$\begin{aligned} D_{m,m,j}^{\sigma,v^0} &\neq 0, & x_j &\neq \frac{x_a + x_b}{2}, \\ D_{m,m,j}^{\sigma,v^0} &= 0, & x_j &= \frac{x_a + x_b}{2}, \\ D_{m,m,j}^{\sigma,v^1} &\neq 0, \end{aligned} \quad (3.189)$$

where the dependence of $D_{m,m,j}^{\sigma,v^0}$ on x_j is a consequence of Eq. (3.183). This property was mentioned here since it is to some extent an analogy to classical physics described by the Boltzmann transport equation. In the BTE the drift term consists of the derivative

$$\mathcal{L}_D^{\text{cl}} f(p, x, t) \propto \partial_x V(x, t) \partial_p f(p, x, t), \quad (3.190)$$

so that solely the slope of $f(p, x, t)$ at a certain point (p, x) is important for the action of the potential but not the magnitude of $f(p, x, t)$ itself.

In the case of the term $I_{m+1/2, i^+, j}^{\sigma, v^0}$, the expression containing $\mathcal{T}^{F, v^0}(u)$ is unproblematic due to Eq. (3.185) and for the other part we examine the expression

$$\begin{aligned} \lim_{k \rightarrow 0} k \mathcal{T}^{\sigma, v^0}(u) \Big|_{kd_1}^{kd_2} &= \lim_{k \rightarrow 0} \frac{k}{4} \left[\frac{1}{u} \cos(u) + \text{Si}(u) \right] \Big|_{kd_1}^{kd_2} \\ &= \lim_{k \rightarrow 0} \frac{k}{4} \left[\frac{1}{kd_2} \cos(kd_2) - \frac{1}{kd_1} \cos(kd_1) \right] \\ &= \frac{1}{4} \left[\frac{1}{d_2} - \frac{1}{d_1} \right]. \end{aligned} \quad (3.191)$$

Therefore, we obtain

$$\begin{aligned} I_{m+1/2, m^+, j}^{\sigma, v^0} &= I_{m-1/2, m^-, j}^{\sigma, v^0} \\ &= -\frac{\hbar}{8} \left[\left(\frac{1}{x_b - x_{j+1/2}} - \frac{1}{x_b - x_{j-1/2}} \right) - \left(\frac{1}{x_a - x_{j+1/2}} - \frac{1}{x_a - x_{j-1/2}} \right) \right]. \end{aligned} \quad (3.192)$$

From Eq. (3.155) one can see that the terms $I_{m+1/2, i^+, j}^{\sigma, v^0}$ and $I_{m-1/2, i^-, j}^{\sigma, v^0}$ enter with the same sign, so that they add up in the case of $D_{m, m, j}^{\sigma, v^0}$ and yield a nonzero value for the diagonal elements of the matrix.

Concerning $I_{m+1/2, i^+, j}^{\sigma, v^1}$ we can make use of the previous results for most of the terms. As shown on the last pages $\mathcal{T}^{F, v^0}(u)$ vanishes in the limit $u \rightarrow 0$, $\frac{1}{k_{m, i}} \mathcal{T}^{F, v^1}(u) \Big|_{C_{m, i, j}^{x_a/b}}$ vanishes as well and the result for $k_{m, i} \mathcal{T}^{\sigma, v^0}(u) \Big|_{C_{m, i, j}^{x_a/b}}$ is known from Eq. (3.191). Therefore, only one term remains which may be evaluated by considering the following expression

$$\begin{aligned} \lim_{k \rightarrow 0} \mathcal{T}^{\sigma, v^1}(u) \Big|_{kd_1}^{kd_2} &= \frac{1}{8} \lim_{k \rightarrow 0} [\text{Ci}(|u|) + u \text{Si}(u) + \cos(u)] \Big|_{kd_1}^{kd_2} \\ &= \frac{1}{8} \lim_{k \rightarrow 0} [\text{Ci}(|kd_2|) - \text{Ci}(|kd_1|)] \\ &= \frac{1}{8} \lim_{k \rightarrow 0} \int_{|kd_1|}^{|kd_2|} \frac{1}{t} \cos(t) dt \\ &= \frac{1}{8} \lim_{k \rightarrow 0} \int_{|d_1|}^{|d_2|} \frac{1}{u} \cos(|k|u) du \\ &= \frac{1}{8} [\ln |d_2| - \ln |d_1|], \end{aligned} \quad (3.193)$$

when substituting $t = |k|u$. With this in hands we can write the result as

$$\begin{aligned}
 I_{m+1/2,m^+,j}^{\sigma,v^1} &= I_{m-1/2,m^-,j}^{\sigma,v^1} \\
 &= -\frac{\hbar}{8} \left[x_b \left(\frac{1}{x_b - x_{j+1/2}} - \frac{1}{x_b - x_{j-1/2}} \right) - x_a \left(\frac{1}{x_a - x_{j+1/2}} - \frac{1}{x_a - x_{j-1/2}} \right) \right] \\
 &\quad + \frac{\hbar}{8} \left[\ln \left(\frac{|x_b - x_{j+1/2}|}{|x_b - x_{j-1/2}|} \right) - \ln \left(\frac{|x_a - x_{j+1/2}|}{|x_a - x_{j-1/2}|} \right) \right], \tag{3.194}
 \end{aligned}$$

yielding nonzero values for the diagonal elements of $D_{m,m,j}^{\sigma,v^1}$. The results from Eqs. (3.192) and (3.194) are important in practice when constructing the drift matrices and have to be explicitly included for the cases $k_{m,m} = 0$.

Now, we derive the expressions for particular potential shapes. For the purpose of simulating devices we need to know on the one hand, how to include a bias voltage and on the other hand, how to implement more complicated, composite potential shapes. This can be done in a simple way by combining different constant and linear potential segments in an appropriate way, whereby we make direct use of the derived expressions for $I_{m+1/2,i+1/2,j}^{F,v^{0/1}}$ and $I_{m+1/2,i^{\pm},j}^{\sigma,v^{0/1}}$. For this purpose it is convenient to slightly alter the definition of the drift matrices, comp. Eq. (3.179), by including the factors v^0 and v^1 for the potential in the terms $I_{m+1/2,i+1/2,j}^F$ and $I_{m+1/2,i^{\pm},j}^\sigma$, so that

$$\begin{aligned}
 I_{m+1/2,i+1/2,j}^F &= \frac{qv^0}{\pi\hbar} I_{m+1/2,i+1/2,j}^{F,v^0} + \frac{qv^1}{\pi\hbar} I_{m+1/2,i+1/2,j}^{F,v^1}, \\
 I_{m+1/2,i^+,j}^\sigma &= \frac{qv^0}{\pi\hbar} I_{m+1/2,i^+,j}^{\sigma,v^0} + \frac{qv^1}{\pi\hbar} I_{m+1/2,i^+,j}^{\sigma,v^1} \tag{3.195}
 \end{aligned}$$

and in an analogous way for $I_{m+1/2,i^-,j}^\sigma$. The drift matrices for $F_{m,j}$ and $\sigma_{m,j}$ are then given by

$$\begin{aligned}
 D_{m,i,j}^F &= \frac{\hbar}{\Delta p_m \Delta x} \left[I_{m+1/2,i+1/2,j}^F - I_{m+1/2,i-1/2,j}^F - \left(I_{m-1/2,i+1/2,j}^F - I_{m-1/2,i-1/2,j}^F \right) \right], \\
 &\quad m = 2, \dots, N_p - 1, \\
 D_{1,i,j}^F &= \frac{\hbar}{\Delta p_m \Delta x} \left(I_{3/2,i+1/2,j}^F - I_{3/2,i-1/2,j}^F \right), \\
 D_{N_p,i,j}^F &= -\frac{\hbar}{\Delta p_m \Delta x} \left(I_{N_p-1/2,i+1/2,j}^F - I_{N_p-1/2,i-1/2,j}^F \right), \tag{3.196}
 \end{aligned}$$

and

$$\begin{aligned}
 D_{m,i,j}^\sigma &= \frac{\hbar}{\Delta p_m \Delta x} \left[I_{m+1/2,i^+,j}^\sigma - I_{m+1/2,i^-,j}^\sigma - \left(I_{m-1/2,i^+,j}^\sigma - I_{m-1/2,i^-,j}^\sigma \right) \right], \\
 &\quad m = 2, \dots, N_p - 1, \\
 D_{1,i,j}^\sigma &= \frac{\hbar}{\Delta p_m \Delta x} \left(I_{3/2,i^+,j}^\sigma - I_{3/2,i^-,j}^\sigma \right), \\
 D_{N_p,i,j}^\sigma &= -\frac{\hbar}{\Delta p_m \Delta x} \left(I_{N_p-1/2,i^+,j}^{\sigma,v^1} - I_{N_p-1/2,i^-,j}^{\sigma,v^1} \right), \tag{3.197}
 \end{aligned}$$

where in each of the equations (i, j) takes on the values $i = 1, \dots, N_p$, $j = 1, \dots, N_x$.

To include a bias voltage, we consider the following potential shape

$$V(x) = V_{DS}\Theta(x_c < x \leq x_0) + V_{DS}\frac{x_L - x}{x_L - x_0}\Theta(x_0 < x \leq x_L), \quad (3.198)$$

with the function Θ as defined in Eq. (3.39). The positions x_0 and x_L herein label the device-endpoints, or at least the positions after which the potential is considered to be constant and x_c is some point inside the contact on the left-hand side which we let go to $-\infty$ in the final expressions. The symbol V_{DS} stands for Drain-Source voltage and corresponds to the difference in applied voltages at both ends of the device. As one can see, we consider the simplest case of a strictly linear potential drop along the device. Of course, this is only an approximation and is definitely not true anymore as soon as one aims for a self-consistent solution of the equations by coupling the WTE to Poisson's equation. Still, one could use the algorithm presented here but would have to use a more complicated, compound potential shape by including a series of additional potential nodes. In this work we consider only the simpler, linear case. With $V(x)$ as chosen above, the factors v^0 and v^1 are

$$\begin{aligned} v^0 &= V_{DS}, \\ v^1 &= 0 \end{aligned} \quad (3.199)$$

in the region $x_c < x \leq x_0$, and

$$\begin{aligned} v^0 &= \frac{x_L}{x_L - x_0}V_{DS}, \\ v^1 &= \frac{-1}{x_L - x_0}V_{DS} \end{aligned} \quad (3.200)$$

in the region $x_0 < x \leq x_L$. In the case of the drift matrix for $F_{m,j}$ we thus need to evaluate the following expression

$$\begin{aligned} I_{m+1/2, i+1/2, j}^F &= \frac{qV_{DS}}{\pi\hbar} \left\{ \mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_0}} - \mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_c}} \right. \\ &\quad \left. + \frac{x_L}{x_L - x_0} \left[\mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_L}} - \mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_0}} \right] \right\} \\ &\quad - \frac{1}{x_L - x_0} \frac{qV_{DS}}{\pi\hbar} \left\{ x_L \mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_L}} - x_0 \mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_0}} \right. \\ &\quad \left. + \frac{1}{k_{m, i}} \left[\mathcal{T}^{F, v^1}(u) \Big|_{C_{m, i, j}^{x_L}} - \mathcal{T}^{F, v^1}(u) \Big|_{C_{m, i, j}^{x_0}} \right] \right\}, \end{aligned} \quad (3.201)$$

when combining Eqs. (3.195) and (3.179). As one can see, the three terms $\mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_0}}$ cancel each other, as well as the two terms $\mathcal{T}^{F, v^0}(u) \Big|_{C_{m, i, j}^{x_L}}$. The latter vanish due to the choice of the potential shape since $V(x_L) = 0$. The former do because $V(x)$ is continuous and no step exists between the regions $x_c < x < x_0$ and $x_0 < x \leq x_L$. This

is a general feature and one can observe for any continuous, piecewise linear potential function that all terms $\mathcal{T}^{F/\sigma, v^0}$ drop out of the equations and only those with v^1 as index remain. As a result, the expression reduces to

$$I_{m+1/2, i+1/2, j}^F = \frac{qV_{DS}}{\pi\hbar} \left\{ -\mathcal{T}^{F, v^0}(u)|_{C_{m, i, j}^{x_c}} - \frac{1}{x_L - x_0} \frac{1}{k_{m, i}} \left[\mathcal{T}^{F, v^1}(u)|_{C_{m, i, j}^{x_L}} - \mathcal{T}^{F, v^1}(u)|_{C_{m, i, j}^{x_0}} \right] \right\}. \quad (3.202)$$

To model the situation of semi-infinite leads we now focus on the limiting process $x_c \rightarrow -\infty$. From the definition of $\mathcal{T}^{F, v^0}(u)$, comp. e.g. Eq. (3.178), it is not apparent how to evaluate this limit, but when we go one step back to Eq. (3.138) and undo the integration with respect to x , we can readily evaluate the limit to

$$\begin{aligned} \lim_{x_c \rightarrow -\infty} \mathcal{T}^{F, v^0}(u)|_{C_{m, i, j}^{x_c}} &= \lim_{x_c \rightarrow -\infty} \left\{ \int_{x_{j-1/2}}^{x_{j+1/2}} \frac{1}{2(x_c - x)} \sin(2k_{m, i}(x_c - x)) dx \right. \\ &\quad \left. - \int_{x_{j-1/2}}^{x_{j+1/2}} k_{m, i} \text{Ci}[|2k_{m, i}(x_c - x)|] dx \right\} \\ &= 0, \end{aligned} \quad (3.203)$$

since the integrand vanishes in both terms. The asymptotic behaviour $\text{Ci}(x) = 0$ for $x \rightarrow \infty$ can be seen directly from the definition of the cosine integral, as listed in Sec. 3.2.2. Finally, the expression for a bias voltage can be written in the compact form

$$I_{m+1/2, i+1/2, j}^F = -\frac{qV_{DS}}{\pi\hbar} \frac{1}{x_L - x_0} \frac{1}{k_{m, i}} \left[\mathcal{T}^{F, v^1}(u)|_{C_{m, i, j}^{x_L}} - \mathcal{T}^{F, v^1}(u)|_{C_{m, i, j}^{x_0}} \right]. \quad (3.204)$$

In the case of the drift matrix for $\sigma_{m, j}$ we need to perform the analogous calculation for $I_{m+1/2, i^+, j}^\sigma$. One can see from Eq. (3.179) that the single trigonometric functions $\mathcal{T}^{F/\sigma, v^0}(u)$ enter in completely the same manner, namely that for each term in $I_{m+1/2, i^+, j}^{\sigma, v^0}$ a corresponding term in $I_{m+1/2, i^+, j}^{\sigma, v^1}$ exists with the same prefactors except for an additional $x_{a/b}$. Therefore, all the terms with v^0 as index cancel each other, as it was the case for $I_{m+1/2, i+1/2, j}^F$, and one arrives at the similar expression

$$\begin{aligned} I_{m+1/2, i^+, j}^\sigma &= \frac{qV_{DS}}{\pi\hbar} \left\{ -\left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \mathcal{T}^{F, v^0}(u)|_{C_{m, i, j}^{x_c}} + \hbar k_{m, i} \mathcal{T}^{\sigma, v^0}(u)|_{C_{m, i, j}^{x_c}} \right\} \\ &\quad - \frac{qV_{DS}}{\pi\hbar} \frac{1}{x_L - x_0} \left\{ \hbar \left[\mathcal{T}^{\sigma, v^1}(u)|_{C_{m, i, j}^{x_L}} - \mathcal{T}^{\sigma, v^1}(u)|_{C_{m, i, j}^{x_0}} \right] \right. \\ &\quad \left. + \left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \frac{1}{k_{m, i}} \left[\mathcal{T}^{F, v^1}(u)|_{C_{m, i, j}^{x_L}} - \mathcal{T}^{F, v^1}(u)|_{C_{m, i, j}^{x_0}} \right] \right\}, \end{aligned} \quad (3.205)$$

where the limit $x_c \rightarrow -\infty$ still has to be carried out. The limit for $\mathcal{T}^{F,v^0}(u)$ is known from Eq. (3.203) and the other one can be calculated as follows

$$\begin{aligned} \lim_{x_c \rightarrow -\infty} \mathcal{T}^{\sigma,v^0}(u)|_{C_{m,i,j}^{x_c}} &= \lim_{x_c \rightarrow -\infty} \frac{1}{4} \left[\frac{1}{u} \cos(u) + \text{Si}(u) \right] \Big|_{C_{m,i,j}^{x_c}} \\ &= -\frac{1}{4} \left[\frac{\pi}{2} - \frac{\pi}{2} \right] \\ &= 0, \end{aligned} \quad (3.206)$$

because of the factor $\frac{1}{u}$ in the case of the first term and for the sine integral it is known that [28]

$$\lim_{x \rightarrow \infty} \text{Si}(x) = \frac{\pi}{2}. \quad (3.207)$$

As a result, both terms for x_c in Eq. (3.205) vanish and we finally arrive at

$$\begin{aligned} I_{m+1/2,i^+,j}^{\sigma} &= -\frac{qV_{DS}}{\pi\hbar} \frac{1}{x_L - x_0} \left\{ \hbar \left[\mathcal{T}^{\sigma,v^1}(u)|_{C_{m,i,j}^{x_L}} - \mathcal{T}^{\sigma,v^1}(u)|_{C_{m,i,j}^{x_0}} \right] \right. \\ &\quad \left. + \left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \frac{1}{k_{m,i}} \left[\mathcal{T}^{F,v^1}(u)|_{C_{m,i,j}^{x_L}} - \mathcal{T}^{F,v^1}(u)|_{C_{m,i,j}^{x_0}} \right] \right\}. \end{aligned} \quad (3.208)$$

The results of Eqs. (3.204) and (3.208) tell us how to include a bias voltage and the derivation of the equations demonstrated that for every continuous potential shape only trigonometric functions $\mathcal{T}^{F/\sigma,v^1}(u)$ with v^1 as index enter the expressions. When replacing $-V_{DS}$ by $V_{x_L} - V_{x_0}$ and using the linearity of the pseudo-differential operator, it seems obvious that the terms $I_{m+1/2,i+1/2,j}^F$ and $I_{m+1/2,i^+,j}^{\sigma}$ for a compound potential shape, consisting of several terms, should be given by a simple sum of expressions like those of Eqs. (3.204) and (3.208), each weighted with a prefactor $\frac{V_{k+1} - V_k}{x_{k+1} - x_k}$ and x_0 and x_L replaced by the corresponding positions x_k and x_{k+1} . To illustrate this point, we consider as an example for a simple compound shape a single barrier of height V_0 ,

$$V(x) = \begin{cases} V_0 \frac{x - x_1}{x_2 - x_1} & \text{if } x_1 < x \leq x_2, \\ V_0 & \text{if } x_2 < x \leq x_3, \\ V_0 \frac{x_4 - x}{x_4 - x_3} & \text{if } x_3 < x \leq x_4, \end{cases} \quad (3.209)$$

where we assume $x_1 < x_2 < x_3 < x_4$. To do the calculation for this potential one could repeat the same steps as before, by identifying v^0 and v^1 in each region and then adding up the different terms containing the functions $\mathcal{T}^{F/\sigma,v^0/v^1}(u)$, or, equally well, one can make use of the results for the bias voltage and rewrite $V(x)$ for the single barrier as

$$V(x) = V_l(x) + V_r(x), \quad (3.210)$$

with

$$V_l(x) = V_0 \frac{x - x_1}{x_2 - x_1} \Theta(x_1 < x \leq x_2) + V_0 H(x - x_2) \quad (3.211)$$

and

$$V_r(x) = -V_0 \frac{x - x_3}{x_4 - x_3} \Theta(x_3 < x \leq x_4) - V_0 H(x - x_4), \quad (3.212)$$

where $H(x)$ stands for the Heaviside step function and Θ is defined as in Eq. (3.39). Using the previous results we directly find for the term $I_{m+1/2,i+1/2,j}^F$

$$I_{m+1/2,i+1/2,j}^F = \frac{qV_0}{\pi\hbar} \frac{1}{k_{m,i}} \left\{ \frac{1}{x_2 - x_1} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_2}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_1}} \right] - \frac{1}{x_4 - x_3} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_4}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_3}} \right] \right\}, \quad (3.213)$$

and an analogous result for $I_{m+1/2,i^+,j}^\sigma$. For the case of a general, piecewise linear potential shape, consisting of N_V potential nodes x_k^V with values V_k and for which we assume the potential to be constant in the semi-infinite regions $x < x_1^V$ and $x > x_{N_V}^V$, so that

$$V(x) = \sum_{k=1}^{N_V-1} \left[V_k + \frac{V_{k+1} - V_k}{x_{k+1}^V - x_k^V} (x - x_k^V) \right] \Theta(x_k^V < x \leq x_{k+1}^V) + V_1 H(x_1^V - x) + V_{N_V} H(x - x_{N_V}^V), \quad (3.214)$$

the previous results make it obvious to deduce the general expressions

$$I_{m+1/2,i+1/2,j}^F = \frac{q}{\pi\hbar} \frac{1}{k_{m,i}} \sum_{k=1}^{N_V-1} \frac{V_{k+1} - V_k}{x_{k+1}^V - x_k^V} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_{k+1}^V}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_k^V}} \right],$$

$$I_{m+1/2,i^+,j}^\sigma = \frac{q}{\pi\hbar} \sum_{k=1}^{N_V-1} \frac{V_{k+1} - V_k}{x_{k+1}^V - x_k^V} \left\{ \hbar \left[\mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i,j}^{x_{k+1}^V}} - \mathcal{T}^{\sigma,v^1}(u) \Big|_{C_{m,i,j}^{x_k^V}} \right] + \left(\frac{p_{m+1/2} + p_{i+1/2}}{2} - p_i \right) \frac{1}{k_{m,i}} \left[\mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_{k+1}^V}} - \mathcal{T}^{F,v^1}(u) \Big|_{C_{m,i,j}^{x_k^V}} \right] \right\}. \quad (3.215)$$

The corresponding one for $I_{m+1/2,i^-,j}^\sigma$ is readily obtained by replacing $p_{i+1/2}$ by $p_{i-1/2}$ and leaving p_i the same. With this in hands it is possible to approximate any desired potential shape by a piecewise linear function as defined in Eq. (3.214) and calculate the drift matrices for the action of the pseudo-differential operator via Eqs. (3.215), (3.196) and (3.197). To recap, the nodes of the potential are chosen as a subset of the grid points for the Wigner function

$$\{x_k^V | k = 1, \dots, N_V\} \subseteq \{x_j | j = 1, \dots, N_x\},$$

and in order to keep the computation time for the drift matrices small it is advisable to choose N_V as small as possible, of course. In the simulations for the resonant tunneling diodes, which consist of a double barrier structure under bias as discussed in the upcoming chapters, it suffices to consider a rather small number of potential nodes. Since we assume a strictly linear bias voltage and for each of the barriers a shape as defined in Eq. (3.209), the total number of potential nodes necessary to be included reduces to ten.

3.2.4 Numerical study on the accuracy of the Fourier transform of a piecewise polynomial approximation of a given function

In this section some preliminary numerical results for the Fourier transform of a piecewise polynomial approximation of a given reference function are presented. The results are supposed to give an insight into the limitations as well as the performance of the approximation approach. To test the approximations we choose a modulated Gaussian distribution as reference function,

$$f_{ref.}(p) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(p-\mu)^2}{2\sigma^2}\right] \cos(\lambda p), \quad (3.216)$$

for which the Fourier transform is given by

$$\tilde{f}_{ref.}(\eta) = \frac{1}{4\pi} \left\{ \exp\left[-\frac{\sigma^2}{2}(\eta-\lambda)^2 - i\mu(\eta-\lambda)\right] + \exp\left[-\frac{\sigma^2}{2}(\eta+\lambda)^2 - i\mu(\eta+\lambda)\right] \right\}. \quad (3.217)$$

In particular, the values are set to $\mu = 0.25$, $\sigma = 0.05$ and $\lambda = 8\pi$. The p -grid is given by the same definition as stated earlier,

$$p_{m-1/2} = p_{1/2} + \sum_{l=1}^{m-1} \Delta p_l, \quad p_{1/2} \in \mathbb{R}, \quad \Delta p_l \in \mathbb{R}, \quad m = 2, \dots, N_p + 1,$$

whereby we choose Δp_l to be the same in certain regions, i.e. determine the grid by $\Delta p^0 \in \mathbb{R}$, $\Delta p^1 = 2\Delta p^0$, $\Delta p^2 = 4\Delta p^0$, $\Delta p^3 = 16\Delta p^0$ and

$$\begin{aligned} \Delta p_l &= \Delta p^0 & p_{m-1/2} &\in (-0.1 - \Delta p^0, 0.1 + \Delta p^0), \\ \Delta p_l &= \Delta p^1 & |p_{m-1/2}| &\in (0.1 + \Delta p^0, 0.2 + \Delta p^1), \\ \Delta p_l &= \Delta p^2 & |p_{m-1/2}| &\in (0.2 + \Delta p^1, 0.7 + \Delta p^2), \\ \Delta p_l &= \Delta p^3 & |p_{m-1/2}| &\in (0.7 + \Delta p^2, 2 + \Delta p^3), \end{aligned} \quad (3.218)$$

so that the grid is symmetric with respect to $p = 0$ and has an extension to a value of at least ± 2 . The grid spacings Δp^i are inversely proportional to the total number of grid points N_p and the calculations were performed for

$$N_p = 102, 200, 400, 600, 798, 998, 1198, 1398, 1596, 1796, 1996. \quad (3.219)$$

To evaluate the piecewise constant and piecewise linear approximation, the cell averages are calculated from the reference function as defined in Eq. (3.216)

$$F_m = \frac{1}{\Delta p_m} \int_{p_{m-1/2}}^{p_{m+1/2}} f_{ref.}(p) dp, \quad m = 1, \dots, N_p. \quad (3.220)$$

From the knowledge of these cell averages, a piecewise polynomial approximation $f(p)$, as defined in Eqs. (3.124) - (3.125), is used to reconstruct the function $f_{ref.}(p)$. For the case of the piecewise constant approximation the slopes σ_m are set to zero

and for the piecewise linear approximation the σ_m are determined by Eq. (3.126), i.e. by centered slopes for the interior grid points. The Fourier transformed function $\tilde{f}(\eta)$ is then calculated as defined in Eq. (3.128) and by making use of Eqs. (3.132) and (3.143), together with setting $\hbar = 1$.

Results of the tests for the cases $N_p = 102$ and $N_p = 200$ are depicted in the Figs. (3.7) and (3.8), respectively. Figure 3.7 shows a good agreement of $\tilde{f}(\eta)$ and $\tilde{f}_{ref.}(\eta)$ only for values of η around zero and a significant deviation for larger values of η . This behaviour is apparent for both approximations, whereas it is less pronounced in the case of the piecewise linear approximation. With an increasing number of grid points N_p the deviation is drastically reduced and as one can see from Fig. 3.8, that already for the case $N_p = 200$, the spurious oscillations in the Fourier transform $\tilde{f}(\eta)$ are remarkably extenuated.

It is important to be aware of the creation of spurious oscillations in the Fourier transform $\tilde{f}(\eta)$ in the case of a too large spacing of the p -grid. In principle, due to the reconstruction of a continuous function out of the cell averages, one is able to compute the Fourier transform $\tilde{f}(\eta)$ for arbitrarily large values of η even on coarse p -grids. But, in order to arrive at reliable values one has to make sure that the spacing of the p_m is fine enough for possibly desired large values of η . As one can see from the definition of the pseudo-differential operator Eq. (2.12), the variable η corresponds to a length and more specifically, to the long-range interaction of the potential $V(x)$ with the Wigner function. Therefore, in practical simulations the choice of the grid spacing for the p variable has to be adjusted to the device length. As we will see later, in order not to introduce too large errors due to the inflow/outflow boundary conditions originating from classical transport theory, one is required to include not only the device but also some part of the contacts as well in the simulation domain. Usually, one would expect the results to improve when increasing the simulated length of the contacts. But this is only true, in general, if one adjusts the p grid as well when changing the size of the x domain. A certain p spacing may suffice for a rather short contact length but introduce spurious oscillations when increasing the contact length without refining the p grid. This inverse proportionality of the p and the η variable is typical for a Fourier transformation and is very obvious in the case of the discrete Fourier transformation, compare e.g. [24] or Eq. (3.18).

In order to analyze the accuracy of the approximation method, the Fourier transform of the approximated function is now compared with the Fourier transform of the reference function for larger values of N_p , as listed in (3.219). To estimate the error, the 2-norm and max norm difference are used, defined by [11]

$$\|\Delta f\|_2 = \left(\Delta\eta \sum_{i=1}^{N_\eta} |\tilde{f}(\eta_i) - \tilde{f}_{ref.}(\eta_i)|^2 \right)^{\frac{1}{2}} \quad (3.221)$$

and

$$\|\Delta f\|_\infty = \max_{i=1, \dots, N_\eta} |\tilde{f}(\eta_i) - \tilde{f}_{ref.}(\eta_i)|. \quad (3.222)$$

The formulas were evaluated for an equidistant set of $N_\eta = 6000$ points in the interval

$\eta_i \in (-300, 300)$ and the results are depicted in Fig. 3.9. A log-log plot is chosen to gain information about the order of convergence of each approximation. The curves for both approximations exhibit a kink at low values of N_p but can then be well approximated by a straight line in this illustration. The kink is related to the creation of strong spurious oscillations for the case $N_p = 102$. It is apparent that the piecewise linear approximation is superior for all considered values of N_p and exhibits a higher order of convergence. To quantify this, we calculate the order of convergence in analogy to the definition stated in [15] by

$$\mathcal{O}_{N_p} = \frac{\log \left(\frac{\|\Delta f\|_2^{N_p}}{\|\Delta f\|_2^{N'_p}} \right)}{\log \left(\frac{N'_p}{N_p} \right)}, \quad (3.223)$$

whereby we choose for N'_p a value listed in Eq. (3.219) which is approximately two times N_p . The results of this calculation can be found in Tab. 3.1. As one can see, the order of convergence of the piecewise constant approximation can be quantified to be about two and the piecewise linear approximation displays an order of convergence close to four. The results are not very intuitive and further analytic investigation would be needed to explain the particular order of convergence. However, this aspect was not analyzed in more detail due to the fact that the device simulations presented in the later sections do not reveal this particular and apparent order of convergence anymore but instead, the situation becomes more involved due to the dynamics of the time evolution of $f(p, x, t)$. More detailed convergence studies for the full WTE are presented in Chap. 5 and the present section serves mainly as a supplement to illustrate the performance and limitations associated with the piecewise polynomial approximation, e.g. the problem of spurious oscillations, as well as the utility of devising higher-order approximations.

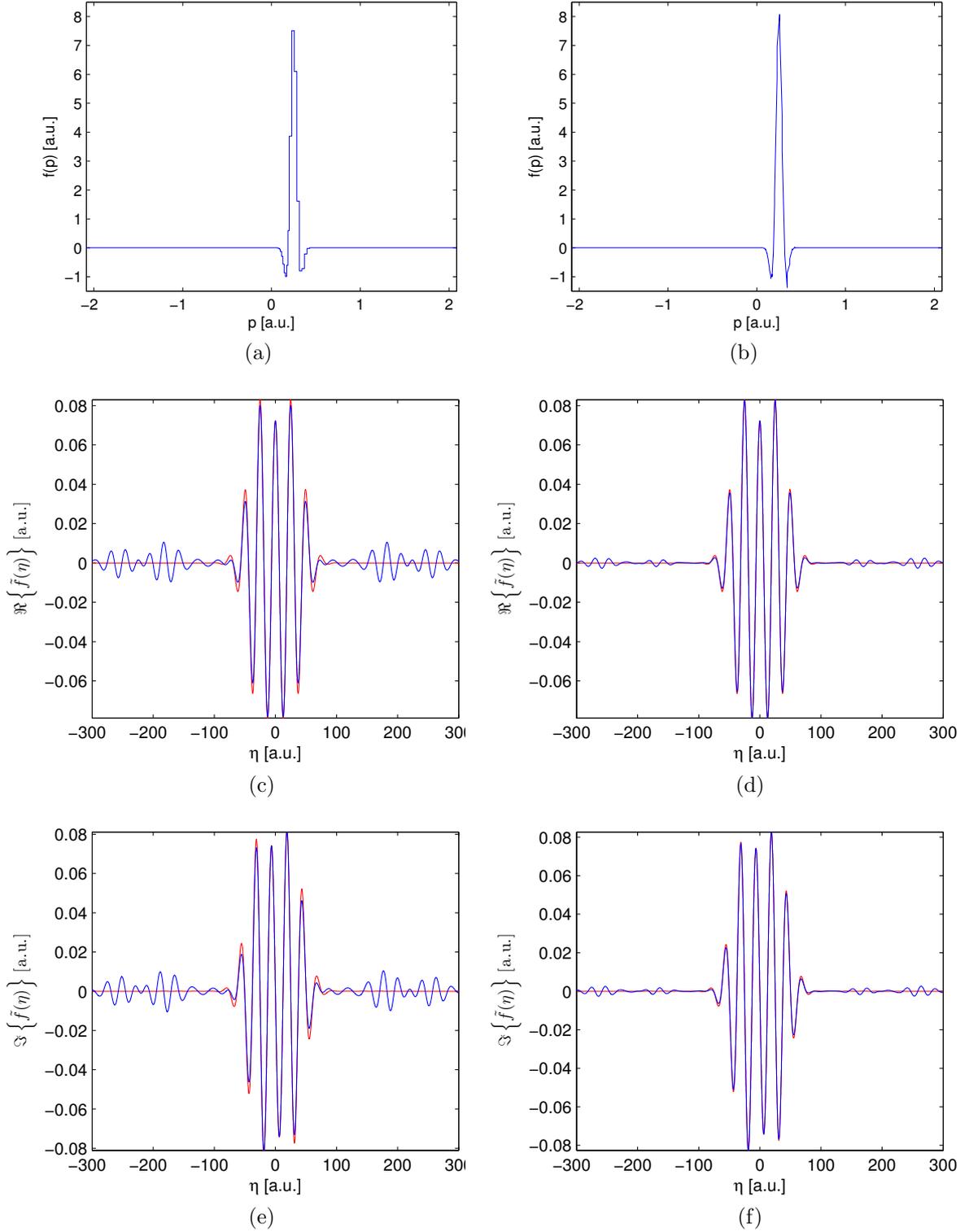


Figure 3.7: The plots in (a) and (b) illustrate $f(p)$ for the piecewise constant and the piecewise linear approximation, respectively, for $N_p = 102$. In (c) $\Re\{\tilde{f}(\eta)\}$ and in (e) $\Im\{\tilde{f}(\eta)\}$ are depicted for the piecewise constant approximation in blue. For the case of the piecewise linear approximation, $\Re\{\tilde{f}(\eta)\}$ is shown in (d) and $\Im\{\tilde{f}(\eta)\}$ is shown in (f) in blue. For comparison, $f_{ref.}(\eta)$ is plotted in (c), (d), (e) and (f) in red colour.

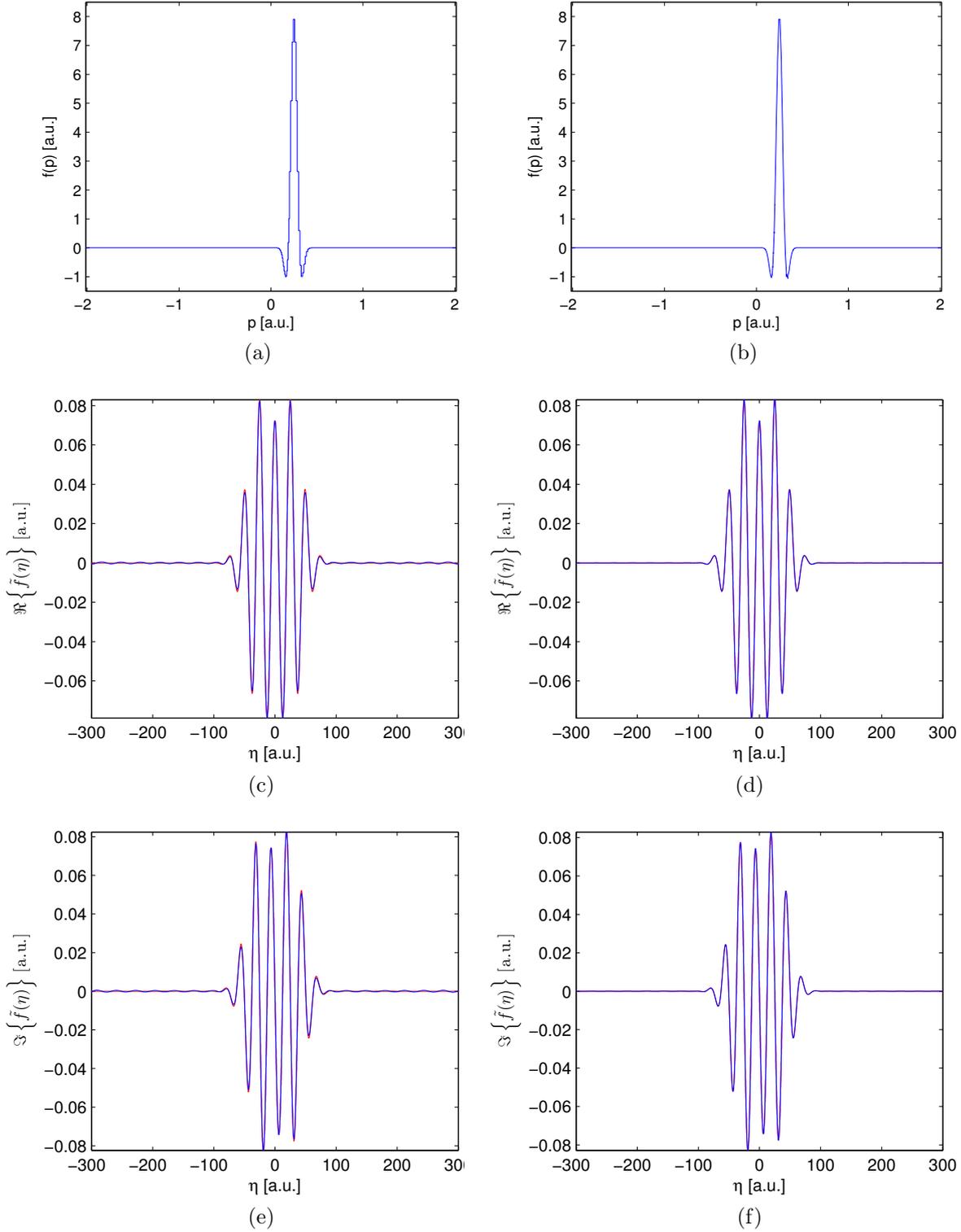


Figure 3.8: The plots in (a) and (b) illustrate $f(p)$ for the piecewise constant and the piecewise linear approximation, respectively, for $N_p = 200$. In (c) $\mathfrak{R}\{\tilde{f}(\eta)\}$ and in (e) $\mathfrak{S}\{\tilde{f}(\eta)\}$ are depicted for the piecewise constant approximation in blue. For the case of the piecewise linear approximation, $\mathfrak{R}\{\tilde{f}(\eta)\}$ is shown in (d) and $\mathfrak{S}\{\tilde{f}(\eta)\}$ is shown in (f) in blue. For comparison, $f_{ref}(\eta)$ is plotted in (c), (d), (e) and (f) in red colour.

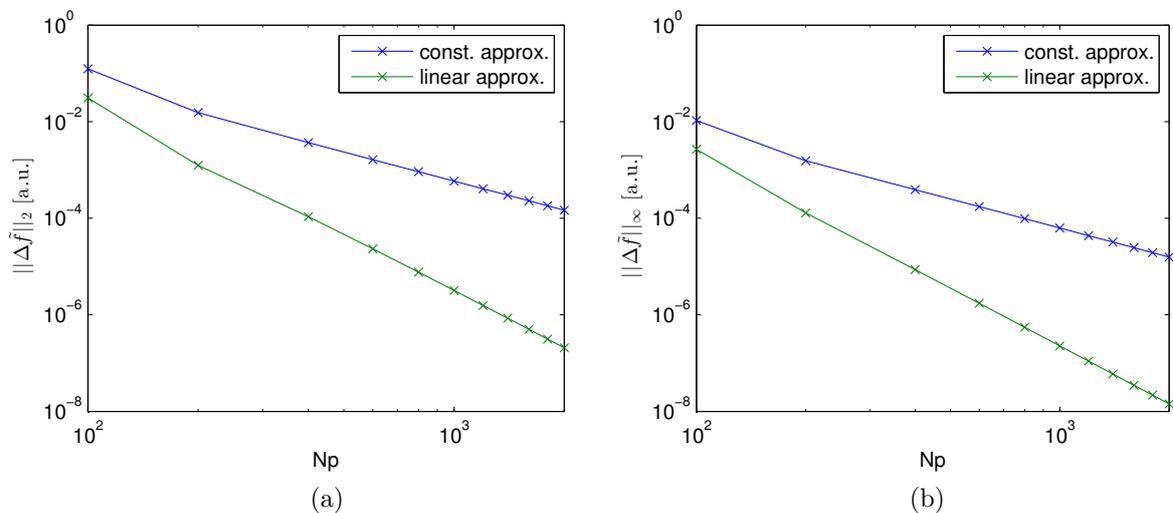


Figure 3.9: Double logarithmic plot of the 2-norm (a) and the the max norm (b) difference of $\tilde{f}(\eta)$ with respect to $\tilde{f}_{ref.}(\eta)$, see Eqs. (3.221) and (3.222), as a function of the grid points N_p . The results are shown for the piecewise constant as well as for the piecewise linear approximation.

Table 3.1: Order of the convergence determined by Eq. (3.223) for the 2-norm difference as depicted in Fig. 3.9, listed for the piecewise constant and the piecewise linear approximation.

N_p	const. approx.	linear approx.
102	3.11	4.78
200	2.05	3.54
400	2.01	3.82
798	2.00	3.93

3.3 Numerical schemes for the advection term and time-stepping methods

In the present section we focus on the advection term and consider the following sub-problem of the Wigner transport equation, Eq. (2.10):

$$\partial_t f(p, x, t) = -\frac{p}{m^*} \partial_x f(p, x, t). \quad (3.224)$$

In Eqs. (3.97) and (3.98) we already averaged the WTE over grid cells to arrive at the expression

$$\partial_t F_{m,j}(t) = -\frac{1}{\Delta x} \left[\hat{h}_{m,j+1/2}(t) - \hat{h}_{m,j-1/2}(t) \right], \quad m = 1, \dots, N_p, j = 1, \dots, N_x, \quad (3.225)$$

with the flux at the right-sided boundary of grid cell $C_{m,j}$ given by Eq. (3.99):

$$\hat{h}_{m,j+1/2}(t) = \frac{1}{\Delta p_m} \int_{p_{m-1/2}}^{p_{m+1/2}} \frac{p}{m^*} f(p, x_{j+1/2}, t) dp. \quad (3.226)$$

In the derivation of the expressions for the drift matrices we assumed a specific form for $f(p, x, t)$, namely to be piecewise constant with respect to x and the dependence with respect to p to be given by some higher-order piecewise approximation. As discussed earlier, this procedure seems to be appropriate since the pseudo-differential operator itself acts solely on the momentum variable. Now we have completely the opposite situation. The partial derivative in the advection term acts only on x , so that it is convenient to assume that $f(p, x, t)$ is piecewise constant with respect to p in this case. With respect to the x variable we then use some higher order approximation to determine the values of the flux at the cell boundaries $x_{j\pm 1/2}$. The assumption that $f(p, x, t)$ is constant on the interval $(p_{m-1/2}, p_{m+1/2})$ greatly simplifies Eq. (3.226) and one arrives at

$$\hat{h}_{m,j+1/2}(t) = \frac{p_m}{m^*} f(p_m, x_{j+1/2}, t), \quad (3.227)$$

with p_m representing the cell center of $(p_{m-1/2}, p_{m+1/2})$. One can see that the momentum index m enters only as a parameter in Eqs. (3.225) and (3.227). Consequently, the equations (3.225) decouple with respect to m and the problem reduces to solving a set of N_p one-dimensional advection equations. In practice one knows the set of cell averages $F_{m,j}(t)$ at a certain time t_n and wants to advance forward in time by a time step Δt to calculate the new cell averages at time t_{n+1} . Integration of Eq. (3.225) with respect to the temporal variable results in [11]

$$F_{m,j}(t_{n+1}) - F_{m,j}(t_n) = -\frac{1}{\Delta x} \int_{t_n}^{t_{n+1}} \left[\hat{h}_{m,j+1/2}(t) - \hat{h}_{m,j-1/2}(t) \right] dt. \quad (3.228)$$

For the right-hand side of the equation we introduce the numerical flux at the cell boundaries [11]

$$\hat{H}_{m,j+1/2}^n \approx \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \frac{p_m}{m^*} f(p_m, x_{j+1/2}, t) dt, \quad (3.229)$$

so that we arrive at the following equation

$$F_{m,j}^{n+1} = F_{m,j}^n - \frac{\Delta t}{\Delta x} \left(\hat{H}_{m,j+1/2}^n - \hat{H}_{m,j-1/2}^n \right), \quad (3.230)$$

where we used the compact notation to label a certain time t_n or t_{n+1} . The form of Eq. (3.230) expresses the important feature, namely that the temporal change of $F_{m,j}^n$ is completely determined by the flux through the cell boundaries. This automatically preserves the number of particles, when viewing $F_{m,j}^n$ as some particle density. For the case of N_x cell values for a certain momentum p_m , the total amount of particles in this region can only change due to the flux at the domain boundaries, as can be seen when calculating the sum over all j

$$\sum_{j=1}^{N_x} F_{m,j}^{n+1} = \sum_{j=1}^{N_x} F_{m,j}^n - \frac{\Delta t}{\Delta x} \left(\hat{H}_{m,N_x+1/2}^n - \hat{H}_{m,1/2}^n \right). \quad (3.231)$$

For the case that the fluxes at the boundaries vanish, the sum of $F_{m,j}^n$ with respect to j stays constant over time, independent of the particular approximation of the numerical fluxes $\hat{H}_{m,j+1/2}^n$. The equations stated so far do not enable a direct implementation since Eq. (3.229) is only a formal expression and we have to rely on approximations to arrive at an explicit scheme. The key point is to construct an appropriate method to calculate the average flux through the cell boundaries $x_{j\pm 1/2}$ from the knowledge of the cell averages $F_{m,j}^n$. Various approaches exist and the accuracy of a numerical result for a certain problem depends on the applicability of the particular approach to the problem, but, for every approximation it is guaranteed that the underlying conservation law is fulfilled. Extensive literature exists for the numerical treatment of such hyperbolic equations, as described for instance in [18] or [11], and we will present here only a brief overview of the subject. In the following we will outline the REA-algorithm, start with its simplest variant to derive the upwind scheme and then look at slope limiter methods. Finally, we discuss the very efficient WENO (weighted essentially non oscillatory) methods, which rely on similar ideas as the slope limiter schemes.

The abbreviation REA stands for reconstruct-evolve-average and describes a procedure to approximately calculate the numerical fluxes $\hat{H}_{m,j+1/2}^n$ [11]:

- In the reconstruction step one chooses for $f(p_m, x, t_n)$ a piecewise polynomial approximation with respect to x . This makes it possible to retain a function defined for a continuous argument x from the knowledge of the discrete set of cell averages $F_{m,j}^n$.
- This function may then be evolved exactly or approximately in time to obtain after a time step Δt the function $f(p_m, x, t_{n+1})$.
- Upon averaging $f(p_m, x, t_{n+1})$ over grid cells one obtains the new cell averages $F_{m,j}^{n+1}$.

In the case of the advection equation, one can use the exact solution for the second step, given by

$$f(p_m, x, t_{n+1}) = f(p_m, x - \frac{p_m}{m^*} \Delta t, t_n), \quad (3.232)$$

which can be readily checked when replacing t_{n+1} by t as well as Δt by $t - t_n$ and inserting it into Eq. (3.224). Typical for hyperbolic problems is the propagation along characteristics, which are simply straight lines in our case of the the linear advection equation with constant coefficients:

$$x - \frac{p_m}{m^*} t = \text{const.} \quad (3.233)$$

One can see that for $p_m > 0$ the solution propagates from smaller to larger values of x and for $p_m < 0$ vice versa. It is therefore natural to apply boundary conditions at x_1 for $p_m > 0$ and at x_{N_x} for $p_m < 0$, when concerning N_x grid points with the ordering $x_1 < x_2 < \dots < x_{N_x}$. This boundary conditions then completely determine the flux into the device and the stationary solution one may attain.

The knowledge of the characteristics helps us to identify the domain of dependence of the underlying PDE. In the case of $p_m > 0$ for instance, we know that $f(p_m, x_{j+1/2}, t + \tau)$, $\tau > 0$, depends only on values of $f(p_m, x, t)$ with $x \leq x_{j+1/2}$. It is crucial to know the domain of dependence of the PDE to construct the numerical domain of dependence properly. This is expressed by the famous CFL condition (Courant, Friedrichs and Lewy, see [18], [24], [11]), which states that it is a necessary condition for the stability of a numerical method that the numerical domain of dependence contains the physical domain of dependence of the underlying PDE [11]. In particular, when we want to express the flux $\hat{H}_{m,j+1/2}^n$ solely in terms of the values of cell averages at neighbouring cells $F_{m,j+1}^n$ and $F_{m,j}^n$, we need to limit the time step by

$$\nu = \frac{p_m}{m^*} \frac{\Delta t}{\Delta x} \leq 1, \quad (3.234)$$

where ν is called the CFL- or Courant-number [11]. The limitation $\nu \leq 1$ ensures that the solution cannot propagate over distances longer than one grid cell during the time step Δt . In the following we will always, except for the WENO schemes, consider $\nu \leq 1$.

For the particular case of a piecewise constant reconstruction we thus find

$$\begin{aligned} f(p_m, x_{j+1/2}, t_n + \tau) &= F_{m,j}^n, & \text{for } p_m > 0, \quad 0 \leq \tau \leq \Delta t, \\ f(p_m, x_{j+1/2}, t_n + \tau) &= F_{m,j+1}^n, & \text{for } p_m < 0, \quad 0 \leq \tau \leq \Delta t. \end{aligned} \quad (3.235)$$

Upon insertion into Eq. (3.229) we arrive at

$$\begin{aligned} \hat{H}_{m,j+1/2}^n &= \frac{p_m}{m^*} F_{m,j}^n, & \text{for } p_m > 0, \\ \hat{H}_{m,j+1/2}^n &= \frac{p_m}{m^*} F_{m,j+1}^n, & \text{for } p_m < 0, \end{aligned} \quad (3.236)$$

which constitutes together with Eq. (3.230) an explicit update formula for the cell averages $F_{m,j}$, known as the upwind method in literature [18], [24], [11]. The reconstruction and evolution step of the REA-algorithm are hereby contained in Eq. (3.235) and the averaging in Eq. (3.230) since we already averaged over x in the derivation. Finite volume methods are often similar to finite difference methods. The exact same expression for the upwind method could also be obtained when directly replacing the partial derivatives in Eq. (3.224) by one-sided finite differences, which are first-order accurate. The upwind method is conditionally stable for $\nu \leq 1$ and produces the best results for ν close to one. Especially for lower values of ν a strong smearing of the solution function can be observed since the leading-order error term is proportional to Δx^2 and one can show that this corresponds to a dissipative error [11]. In the present case of the WTE this is an unfavorable situation since to fulfill the condition $\nu \leq 1$ for all N_p advection equations we have to adjust Δt to the largest value $p_{max} = \max |p_m|$. For the case of p_{max} , ν can be chosen fairly close to one but especially for the interior equations with small values of p_m , ν may be very small and one thus introduces significant dissipation. This is a problem, of course, since especially at the lower values of p_m the quasi-distribution function $f(p, x, t)$ takes on large values. Despite of this, the upwind method was used as a first starting point for practical simulations done in this work. The smearing might not produce very accurate results but it avoids to some degree that short-scaled oscillations become too pronounced and definitely does not introduce additional, unphysical oscillations. This aided in the development of the polynomial approximation method for the pseudo-differential operator.

To develop higher-order methods one can use higher-order polynomial reconstructions for $f(p_m, x, t)$ and repeat the steps of the REA-algorithm. We now consider a piecewise linear reconstruction of the form [11]

$$f(p_m, x, t_n) = F_{m,j}^n + \hat{\sigma}_{m,j}^n(x - x_j), \quad \text{for } x_{j-1/2} < x \leq x_{j+1/2}, \quad j = 1, \dots, N_x. \quad (3.237)$$

When restricting ourselves again to the case $\nu \leq 1$ and using the knowledge of the propagation of the solution along the characteristics Eq. (3.233), one can make use of geometrical considerations for instance to derive the following update formulas for the cell averages [11]

$$F_{m,j}^{n+1} = F_{m,j}^n - \nu (F_{m,j}^n - F_{m,j-1}^n) - \frac{1}{2}\nu(1 - \nu)\Delta x (\hat{\sigma}_{m,j}^n - \hat{\sigma}_{m,j-1}^n), \quad \text{for } p_m > 0, \quad (3.238)$$

and

$$F_{m,j}^{n+1} = F_{m,j}^n - \nu (F_{m,j+1}^n - F_{m,j}^n) + \frac{1}{2}\nu(1 + \nu)\Delta x (\hat{\sigma}_{m,j+1}^n - \hat{\sigma}_{m,j}^n), \quad \text{for } p_m < 0, \quad (3.239)$$

corresponding to the fluxes

$$\begin{aligned} \hat{H}_{m,j+1/2}^n &= \frac{p_m}{m^*} \left[F_{m,j}^n + \frac{1}{2} \left(\Delta x - \frac{p_m}{m^*} \Delta t \right) \hat{\sigma}_{m,j}^n \right], \quad \text{for } p_m > 0, \\ \hat{H}_{m,j+1/2}^n &= \frac{p_m}{m^*} \left[F_{m,j+1}^n - \frac{1}{2} \left(\Delta x + \frac{p_m}{m^*} \Delta t \right) \hat{\sigma}_{m,j+1}^n \right], \quad \text{for } p_m < 0. \end{aligned} \quad (3.240)$$

Now, the particular choice of the slopes strongly influences the performance of the algorithm. A famous choice is to use downwind slopes ($p_m > 0$)

$$\hat{\sigma}_{m,j}^n = \frac{F_{m,j+1}^n - F_{m,j}^n}{\Delta x}, \quad (3.241)$$

resulting in the Lax-Wendroff method, which is known to be second-order accurate [11]. One could also use upwind or centered slopes which may be better suited for certain situations, but all of these approaches suffer from the drawback that the methods are not total variation diminishing (TVD), meaning that spurious oscillations may arise. The TVD concept was extensively used to derive various high-resolution methods, based on flux or slope-limiter approaches, see [11]. In short, the total variation serves as a measure of the degree of oscillations in the solution function. For a discrete set of values it is defined by [11]

$$\text{TV}(F_m^n) = \sum_{j=2}^{N_x} |F_{m,j}^n - F_{m,j-1}^n|. \quad (3.242)$$

From the analytical solution of the advection equation, Eq. (3.232), it is known that the true solution simply propagates along x without changing its shape. The total variation thus stays constant. It is advisable to impose such a demand also for the numerical methods. Therefore, one defines the TVD property by the following requirement: “A two-level method is called total variation diminishing (TVD) if, for any set of data F_m^n , the values F_m^{n+1} computed by the method satisfy” [11]

$$\text{TV}(F_m^{n+1}) \leq \text{TV}(F_m^n). \quad (3.243)$$

Having a numerical method for which it is guaranteed that the total variation of the solution function cannot increase guarantees that no unphysical oscillations are produced by the method. Various slope limiter approaches exist which fulfill this property. Their common strategy is to choose the slope at a particular position depending on the smoothness of the function. The slope may be chosen as one of the values of the upwind-, downwind- and centered slopes or as a combination of those. Two examples are the minmod-scheme and the MC limiter. The minmod-scheme determines the slope via [11]

$$\hat{\sigma}_{m,j}^n = \text{minmod} \left(\frac{F_{m,j+1}^n - F_{m,j}^n}{\Delta x}, \frac{F_{m,j}^n - F_{m,j-1}^n}{\Delta x} \right), \quad (3.244)$$

with the minmod function defined by

$$\text{minmod}(a, b) = \begin{cases} a & \text{if } |a| < |b|, \quad ab > 0, \\ b & \text{if } |a| > |b|, \quad ab > 0, \\ 0 & \text{if } ab \leq 0. \end{cases} \quad (3.245)$$

Therefore, for the case that the upwind- and downwind slope have the same sign the method always picks the one smaller in magnitude. For the case that the two

slopes have opposite signs or one of them is zero the slope is simply set to zero and a piecewise constant approximation is retained. This is particularly useful for discontinuous solutions since choosing a slope at a discontinuous position, via Eq. (3.241) for instance, results in a reconstruction function which overshoots the values of the set of cell averages and thus results in spurious oscillations. In our case of the simulations performed with the WTE no truly discontinuous solutions are present, however, due to the strong spatial variations, when simulating RTDs, problems with spurious oscillations arise when applying the Lax-Wendroff method or related ones. A very similar scheme which makes use of all three slopes is the MC limiter, standing for monotonized central-difference limiter and determining the slopes by [11]

$$\hat{\sigma}_{m,j}^n = \text{minmod} \left(\frac{F_{m,j+1}^n - F_{m,j-1}^n}{2\Delta x}, 2\frac{F_{m,j+1}^n - F_{m,j}^n}{\Delta x}, 2\frac{F_{m,j}^n - F_{m,j-1}^n}{\Delta x} \right), \quad (3.246)$$

with the minmod function for three arguments defined in an analogous way. The factor of two in front of the one-sided slopes might seem confusing but in literature it is shown that the TVD property is still fulfilled [11]. In addition, the factor of two favours the choice of centered slopes in regions where the solution function is smooth. Again, a value of zero is assigned as soon as the single slopes differ in sign or one of them is zero. A slope equal to zero corresponds, of course, to a piecewise constant approximation of the function. This is a central feature of all high-resolution methods. Whenever the solution is smooth and well resolved they apply a high-order method for better accuracy but in badly resolved regions or regions with discontinuities they switch to a lower order approximation to avoid the creation of oscillations. In this work, the minmod-scheme and MC-limiter were tested in practice and the results obtained seemed to be fairly accurate. The dissipation effects of the upwind scheme were virtually eliminated, whereby the results produced with the MC-limiter appeared to be more accurate than the results of the minmod-scheme. However, a considerable change of the Wigner function could be observed when reducing the grid constant Δx . Especially the computed charge density $n(x)$ inside the double-barrier changed quite significantly, affecting the whole behaviour of the device. This indicated that the methods had problems with resolving the strong spatial variations of the Wigner function close to the barriers on moderate grid sizes. In order to avoid extremely fine grids which cause too long CPU times, a higher order method was sought for. Since the problem of strong spatial variations of the distribution function is a common problem in device simulations due to steep doping profiles, see e.g. [3], well developed schemes exist that can cope with such cases, known as ENO or WENO methods.

Essentially non oscillatory (ENO) and weighted essentially non oscillatory (WENO) methods are based on similar ideas as the high-resolution methods, namely to use a high-order approximation in regions where the solution is smooth and to reduce the order near discontinuities or badly resolved regions. ENO schemes make use of a set of stencils to construct the flux at the cell boundaries. A smoothness indicator is used to pick the one stencil with the smallest amount of oscillations. WENO schemes work in a similar way, only that they do not pick a single stencil but use a convex combination of all stencils. This enables one to increase the order of the method in

smooth regions whereas maintaining the property to handle discontinuities. The first such methods were developed in [32] and further ones can be found in [33] or [34]. In this work we will focus on the WENO5 scheme as presented in [33] or [3], where we use in the following a notation based on the latter one. The method and especially its stability properties were extensively analyzed in [35].

The WENO5 scheme directly approximates the flux at the cell boundaries with three different stencils and we therefore return to the very beginning, to Eqs. (3.224) - (3.227), and write the flux at time t_n as

$$\hat{h}_{m,j+1/2}^n = \omega_1^n h_{m,j+1/2}^{n,(1)} + \omega_2^n h_{m,j+1/2}^{n,(2)} + \omega_3^n h_{m,j+1/2}^{n,(3)}. \quad (3.247)$$

For the case $p_m > 0$ the three single fluxes are determined by

$$\begin{aligned} h_{m,j+1/2}^{n,(1)} &= \frac{p_m}{m^*} \left(\frac{1}{3} F_{m,j-2}^n - \frac{7}{6} F_{m,j-1}^n + \frac{11}{6} F_{m,j}^n \right), \\ h_{m,j+1/2}^{n,(2)} &= \frac{p_m}{m^*} \left(-\frac{1}{6} F_{m,j-1}^n + \frac{5}{6} F_{m,j}^n + \frac{1}{3} F_{m,j+1}^n \right), \\ h_{m,j+1/2}^{n,(3)} &= \frac{p_m}{m^*} \left(\frac{1}{3} F_{m,j}^n + \frac{5}{6} F_{m,j+1}^n - \frac{1}{6} F_{m,j+2}^n \right). \end{aligned} \quad (3.248)$$

The single weights ω_i^n are normalized by the equation

$$\omega_i^n = \frac{\tilde{\omega}_i^n}{\sum_{l=1}^3 \tilde{\omega}_l^n}, \quad i = 1, 2, 3, \quad (3.249)$$

with

$$\tilde{\omega}_i^n = \frac{\gamma_i}{(\varepsilon + \beta_i^n)^2}, \quad i = 1, 2, 3, \quad (3.250)$$

where the values of γ_i are fixed by $\gamma_1 = \frac{1}{10}$, $\gamma_2 = \frac{3}{5}$ and $\gamma_3 = \frac{3}{10}$, $\varepsilon = 10^{-6}$ is a small quantity to avoid that the denominator takes on a value of zero and the smoothness indicators β_i^n are determined by the following equations

$$\begin{aligned} \beta_1^n &= \left(\frac{p_m}{m^*} \right)^2 \left[\frac{13}{12} (F_{m,j-2}^n - 2F_{m,j-1}^n + F_{m,j}^n)^2 + \frac{1}{4} (F_{m,j-2}^n - 4F_{m,j-1}^n + 3F_{m,j}^n)^2 \right], \\ \beta_2^n &= \left(\frac{p_m}{m^*} \right)^2 \left[\frac{13}{12} (F_{m,j-1}^n - 2F_{m,j}^n + F_{m,j+1}^n)^2 + \frac{1}{4} (F_{m,j-1}^n - F_{m,j+1}^n)^2 \right], \\ \beta_3^n &= \left(\frac{p_m}{m^*} \right)^2 \left[\frac{13}{12} (F_{m,j}^n - 2F_{m,j+1}^n + F_{m,j+2}^n)^2 + \frac{1}{4} (3F_{m,j}^n - 4F_{m,j+1}^n + F_{m,j+2}^n)^2 \right]. \end{aligned} \quad (3.251)$$

For completeness we also list here the expressions for $h_{m,j+1/2}^{n,(i)}$ and β_i^n for the case

$p_m < 0$:

$$\begin{aligned}
 h_{m,j+1/2}^{n,(1)} &= \frac{p_m}{m^*} \left(\frac{1}{3} F_{m,j+3}^n - \frac{7}{6} F_{m,j+2}^n + \frac{11}{6} F_{m,j+1}^n \right), \\
 h_{m,j+1/2}^{n,(2)} &= \frac{p_m}{m^*} \left(-\frac{1}{6} F_{m,j+2}^n + \frac{5}{6} F_{m,j+1}^n + \frac{1}{3} F_{m,j}^n \right), \\
 h_{m,j+1/2}^{n,(3)} &= \frac{p_m}{m^*} \left(\frac{1}{3} F_{m,j+1}^n + \frac{5}{6} F_{m,j}^n - \frac{1}{6} F_{m,j-1}^n \right),
 \end{aligned} \tag{3.252}$$

and

$$\begin{aligned}
 \beta_1^n &= \left(\frac{p_m}{m^*} \right)^2 \left[\frac{13}{12} (F_{m,j+3}^n - 2F_{m,j+2}^n + F_{m,j+1}^n)^2 + \frac{1}{4} (F_{m,j+3}^n - 4F_{m,j+2}^n + 3F_{m,j+1}^n)^2 \right], \\
 \beta_2^n &= \left(\frac{p_m}{m^*} \right)^2 \left[\frac{13}{12} (F_{m,j+2}^n - 2F_{m,j+1}^n + F_{m,j}^n)^2 + \frac{1}{4} (F_{m,j+2}^n - F_{m,j}^n)^2 \right], \\
 \beta_3^n &= \left(\frac{p_m}{m^*} \right)^2 \left[\frac{13}{12} (F_{m,j+1}^n - 2F_{m,j}^n + F_{m,j-1}^n)^2 + \frac{1}{4} (3F_{m,j+1}^n - 4F_{m,j}^n + F_{m,j-1}^n)^2 \right].
 \end{aligned} \tag{3.253}$$

The formulas can be directly used for the interior grid points but at the boundaries we need to make use of ghost cells [11] to specify values for $F_{m,j}^n$ at the positions $j = -2, -1, 0$ and $j = N_x + 1, N_x + 2, N_x + 3$. On the left contact at $j = 1$ we know from the characteristics that in the phase space region with $p_m > 0$ the quasi-distribution function propagates to the right, i.e. into the device and in the region $p_m < 0$ it propagates to the left, i.e. leaves the device. To model the inflow from the contacts into the device we set the exterior points for $j < 1$, $p_m > 0$ to fixed values according to a certain, desired distribution function in the contacts. In our case of simulating quantum devices this is the Fermi-Dirac distribution. Since we consider a one-dimensional problem we need to integrate the three-dimensional Fermi-Dirac distribution over the transverse momenta to arrive at a distribution function for the longitudinal momentum alone. The appropriate one-dimensional distribution may be found in [4] or [36] and is given by

$$f_{l,r}(p) = \frac{m^*}{2\pi^2 \hbar^2 \beta} \ln \left\{ 1 + \exp \left[-\beta \left(\frac{p^2}{2m^*} + qV_{l,r} - \mu_{l,r} \right) \right] \right\}, \tag{3.254}$$

$V_{l,r}$ and $\mu_{l,r}$ label the voltage and the chemical potential in the left or right contact, respectively and $\beta = \frac{1}{k_B T}$ incorporates the thermal energy. An additional factor of $\frac{1}{2\pi}$ is included in the equation above since our derivations are based on a slightly different definition of the Wigner transformation than the one used in [4]. To assign a value to a certain $F_{m,j}^n$ we average over the particular grid cell. To model the outflow properly, without causing reflections at the contacts we apply the boundary conditions for ohmic contacts as described in [3] and simply copy the last values of $F_{m,1}^n$ to the

exterior points, so that we have on the whole for the left-sided contact

$$F_{m,j}^n = \begin{cases} \frac{1}{\Delta p_m} \int_{p_{m-1/2}}^{p_{m+1/2}} f_l(p) dp & \text{if } p_m > 0, \\ F_{m,1}^n & \text{if } p_m < 0, \end{cases} \quad (3.255)$$

for $j = 0, -1, -2$. The ghost cells for the contact on the right-hand side are given by the analogous expression

$$F_{m,j}^n = \begin{cases} F_{m,N_x}^n & \text{if } p_m > 0, \\ \frac{1}{\Delta p_m} \int_{p_{m-1/2}}^{p_{m+1/2}} f_r(p) dp & \text{if } p_m < 0, \end{cases} \quad (3.256)$$

for $j = N_x + 1, N_x + 2, N_x + 3$. Having the equations stated to determine the fluxes at the cell boundaries, we now return to Eq. (3.225) and to the problem of devising an appropriate time stepping method. Before, in the case of the upwind and the high-resolution methods we applied in principle a forward Euler scheme to calculate the update from $F_{m,j}^n$ to $F_{m,j}^{n+1}$. In the case of the WENO5 scheme this is not possible anymore since the overall scheme is unstable. In practice it is not immediately apparent that the method is unstable since no sudden exponential growth of the solution can be observed, as a consequence of the nonlinear flux calculation, but a rather chaotic behaviour with oscillations being created, smoothed out again or even amplified. One can only notice that it is reduced with decreasing Δt but never eliminated. Wang [35] analyzed the stability properties of the WENO5 scheme coupled with different time-stepping methods in great detail and due to his results we decided to use a third-order Runge Kutta method, termed SSP(3,3) in his work and those of others, to perform the time step. In the SSP(3,3) method three sub-steps are needed to advance one complete time step and the abbreviation SSP stands for the strong stability preserving property [35]. The particular Runge Kutta coefficients for the SSP(3,3) algorithm are given in [35] in terms of a Butcher tableaux [23]. For our purposes it suffices to consider the differential equation

$$\dot{y} = L\{y\} \quad (3.257)$$

with a time-independent operator L acting on a function y . For this case the three sub-steps of the SSP(3,3) algorithm needed to advance y from time t_n to t_{n+1} are

$$\begin{aligned} k_1 &= L\{y^n\}, \\ k_2 &= L\{y^n + \Delta t k_1\}, \\ k_3 &= L\left\{y^n + \frac{\Delta t}{4}(k_1 + k_2)\right\}, \\ y^{n+1} &= y^n + \Delta t \left(\frac{1}{6}k_1 + \frac{1}{6}k_2 + \frac{2}{3}k_3\right). \end{aligned} \quad (3.258)$$

To combine this scheme with the WENO5 method it is convenient to introduce the functions $y^{n'}$ and $y^{n''}$ at the intermediate times $t_{n'}$ and $t_{n''}$ and rewrite the individual

steps as

$$\begin{aligned}
 y^{n'} &= y^n + \Delta t L\{y^n\}, \\
 y^{n''} &= y^n + \frac{\Delta t}{4} \left(L\{y^n\} + L\{y^{n'}\} \right) \\
 &= \frac{1}{4} \left(3y^n + y^{n'} + \Delta t L\{y^{n'}\} \right), \\
 y^{n+1} &= y^n + \Delta t \left(\frac{1}{6} L\{y^n\} + \frac{1}{6} L\{y^{n'}\} + \frac{2}{3} L\{y^{n''}\} \right) \\
 &= \frac{1}{3} \left(y^n + 2y^{n''} + 2\Delta t L\{y^{n''}\} \right). \tag{3.259}
 \end{aligned}$$

To apply the SSP(3,3) stepper to the advection term of the WTE, one can rewrite Eq. (3.225) in the following form

$$\partial_t F_{m,j}(t) = L^{m,j} \{F_{m,1}(t), \dots, F_{m,N_x}(t)\}, \tag{3.260}$$

when defining the operator

$$L^{m,j} \{F_{m,1}(t_n), \dots, F_{m,N_x}(t_n)\} = -\frac{1}{\Delta x} \left[\hat{h}_{m,j+1/2}^n - \hat{h}_{m,j-1/2}^n \right]. \tag{3.261}$$

The three sub-steps of Eq. (3.259) can then be written in terms of the fluxes as

$$\begin{aligned}
 F_{m,j}^{n'} &= F_{m,j}^n - \frac{\Delta t}{\Delta x} \left[\hat{h}_{m,j+1/2}^n - \hat{h}_{m,j-1/2}^n \right], \\
 F_{m,j}^{n''} &= \frac{1}{4} \left(3F_{m,j}^n + F_{m,j}^{n'} \right) - \frac{1}{4} \frac{\Delta t}{\Delta x} \left[\hat{h}_{m,j+1/2}^{n'} - \hat{h}_{m,j-1/2}^{n'} \right], \\
 F_{m,j}^{n+1} &= \frac{1}{3} \left(F_{m,j}^n + 2F_{m,j}^{n''} \right) - \frac{2}{3} \frac{\Delta t}{\Delta x} \left[\hat{h}_{m,j+1/2}^{n''} - \hat{h}_{m,j-1/2}^{n''} \right], \tag{3.262}
 \end{aligned}$$

where the individual fluxes $\hat{h}_{m,j+1/2}^n$, $\hat{h}_{m,j+1/2}^{n'}$ and $\hat{h}_{m,j+1/2}^{n''}$ are determined by the WENO5 scheme, Eqs. (3.247) - (3.253), out of the set of cell averages $F_{m,j}^n$, $F_{m,j}^{n'}$ and $F_{m,j}^{n''}$, respectively.

In addition, one can identify the overall flux to advance from time t_n to t_{n+1} to be given by

$$\hat{H}_{m,j+1/2}^n = \frac{1}{6} \hat{h}_{m,j+1/2}^n + \frac{1}{6} \hat{h}_{m,j+1/2}^{n'} + \frac{2}{3} \hat{h}_{m,j+1/2}^{n''}, \tag{3.263}$$

when comparing Eq. (3.230) and the first expression for y^{n+1} in Eq. (3.259). The numerical flux $\hat{H}_{m,j+1/2}^n$ is in practical computations a better estimator for the current density than to make use of a discretized expression of Eq. (1.36), i.e. to calculate the current density by

$$j(x_j, t_n) = q \sum_{m=1}^{N_p} \frac{p_m}{m^*} F_{m,j}^n \Delta p_m. \tag{3.264}$$

The problem is that a current density calculated by Eq. (3.264) does not necessarily fulfill the requirement that its divergence vanishes in the stationary case (see

continuity equation Eq. (2.22)). Especially when strong spatial variations of the quasi-distribution function are present, as for the case of tunneling devices, the current density shows strong deviations from an even, stationary solution. Much better results were achieved when calculating the current density on the cell boundaries $x_{j+1/2}$ by summing up the numerical fluxes $\hat{H}_{m,j+1/2}^n$ since these fluxes truly determine the change of the cell averages from time t_n to time t_{n+1} . An additional point has to be taken into account: One may notice from Eqs. (3.262) and (3.263) that the numerical flux is not independent of the time step in the present case but contains terms proportional to Δt and Δt^2 . To suppress this dependence to some degree it is convenient to determine the current density at a certain time t_n by averaging over two adjacent time steps, i.e. by

$$j(x_{j+1/2}, t_n) = q \sum_{m=1}^{N_p} \Delta p_m \frac{\hat{H}_{m,j+1/2}^n + \hat{H}_{m,j+1/2}^{n-1}}{2}. \quad (3.265)$$

In general, when dealing with explicit time-stepping schemes one has to make sure to use a small enough time step for the scheme to be stable. In [35] one can find that the WENO5 scheme together with the SSP(3,3) stepper is stable for CFL numbers of $\nu \leq 1.43$ for smooth solutions and in the case of discontinuities one has to restrict the Courant number to $\nu \leq 1$. For the simulations of RTDs the latter condition is the appropriate one and in the computations done in this work, the CFL number was restricted to $\nu \approx 0.9$ because values too close to one caused problems and enhanced oscillations to some degree.

Now, to use a dedicated time stepping method for both terms in the WTE we apply an operator splitting scheme as described earlier in Sec. 3.1.1. In particular a Strang-splitting scheme is reasonable to achieve an overall second-order accuracy. It was found advisable to solve the advection term in the two $\frac{\Delta t}{2}$ steps of the Strang-splitting, i.e. in place of \mathcal{L}_1 in Eq. (3.7), in order to gain a factor of two for the CFL condition. In addition, in a calculation of multiple time steps one Δt drift step is followed by two $\frac{\Delta t}{2}$ advection steps, so that the two adjacent advection steps enable us to calculate the current density by Eq. (3.265).

Time stepping methods for the drift term have not been discussed yet. The calculations in the last section revealed that it is possible to rewrite the discretized pseudo-differential operator in matrix form as (see Eq. 3.130)

$$\partial_t F_{m,j}(t) = \sum_{i=1}^{N_p} D_{m,i,j} F_{i,j}(t), \quad m = 1, \dots, N_p, \quad j = 1, \dots, N_x. \quad (3.266)$$

The formal solution to this equation is given by

$$F_{m,j}^{n+1} = \sum_{i=1}^{N_p} \exp(\Delta t D_{m,i,j}) F_{i,j}^n, \quad m = 1, \dots, N_p, \quad j = 1, \dots, N_x. \quad (3.267)$$

One could use for instance a series expansion of the exponential function or also apply a Runge Kutta method directly to Eq. (3.266) to arrive at an update formula of desired

order. The strategy chosen in this work was to use the built in Matlab function *expm* which uses the scaling and squaring method [37] to calculate the exponential of a matrix with high accuracy. Due to the convincing performance of the *expm* routine our own routines based on series expansions were discarded and the *expm* Matlab function used in all of the simulations. Since we were mainly interested in reaching stationary solutions in order to validate the developed algorithm against a NEGF reference solution, the *expm* routine had to be called only once in a particular calculation. Therefore, the routine did not cause any severe extra computational cost even on large grids. But, if one would be interested in simulating fully time-dependent potentials $V(x, t)$ other time stepping strategies may be better suited. A comparison of the results achieved with *expm* and those by a series expansion of Eq. (3.267) revealed, that a fourth-order expansion produced practically identical results, a first-order expansion showed instabilities and a second-order expansion seemed to work quite well in most test cases. The series expansion method was not tested thoroughly against the *expm* routine but a second-order expansion or also a Runge Kutta method of second order seems to be a good starting point to treat fully time-dependent potentials. The extra cost in each time step can be estimated to be at least about a factor of two for the case of a second-order method since two matrix multiplications need to be carried out for each position x_j instead of one multiplication as in Eq. (3.267). The extra computational cost needed to update the matrix elements of $D_{m,i,j}$ is strongly dependent on the particular situation. The simple case of a time-dependent bias voltage together with a fixed potential distribution inside the device is directly manageable and one can make use of the linearity of the equations, i.e. that in a matrix $D_{m,i,j}$ for the bias voltage alone the particular value of V_{DS} enters solely as a prefactor. Time-dependent bias voltages may be of interest to analyze the frequency-dependent behaviour of a device.

4 Non-equilibrium Green's function technique

In the Keldysh formalism, the technique of Green's functions known from equilibrium many-body physics is extended to the non-equilibrium case [14]. The formulas for expectation values bear close analogy to the equilibrium case but one has to deal with a set of equations for four different Green's functions. To apply the formalism to the field of device simulations, the fully time-dependent treatment is unfeasible, at the moment, and one usually restricts the calculations to the stationary case. For steady state situations the set of equations can be simplified and two Green's functions suffice to describe all the physics and calculate all measurable quantities. Extensive literature on the subject exists and an elementary introduction to the non-equilibrium Green's functions (NEGF) technique for the purpose of device simulations can be found in the textbooks of Datta [36], [38]. In this chapter we will give only a fairly brief introduction to the subject, since on the one hand, the NEGF-formalism is not the main topic of this work and served primarily as a reference solution for the Wigner function calculations and on the other hand, the case of fully coherent transport through a RTD, as considered in the following, is treated in many textbooks as a basic example and also has been examined in various publications in all detail, see e.g. [15], [6].

4.1 Calculation of steady state properties

In the stationary case, the retarded and the lesser Green's function G^R and $G^<$, respectively, enable to calculate all quantities of interest. The retarded Green's function G^R specifies the coherent propagation of electrons inside the device and furthermore, is closely related to the local density of states $\rho(x, E)$ inside the device [38]. The occupancy of these states, in particular, resulting from the balance of the incoming and outgoing electrons from the left- and right-sided contacts in combination with scattering events, can be quantified by the lesser Green's function. Thus $G^<$ determines the charge density $n(x)$ and the current $j(x)$ [8] [38].

In analogy to the theory of Green's functions in the context of the solution of differential equations, we introduce the retarded and advanced Green's functions as the solutions to the time-independent Schrödinger equation with a delta function as source term

$$(E - H' \pm i\eta) G(x, x', E) = \delta(x - x'), \quad x, x' \in \mathbb{R}. \quad (4.1)$$

Here, H' stands for the specific Hamilton operator and the infinitesimal quantity η was added to uniquely determine the solution, whereby $+i\eta$ and $-i\eta$ correspond to the retarded and advanced Green's function, respectively. Only in the final solution one considers the limit $\eta \rightarrow 0^+$. The problem with the definition of Green's function as in Eq. (4.1) is that one has to consider the infinite domain $x, x' \in \mathbb{R}$. Datta [38] showed in detail how to truncate the domain by rewriting the equation in terms of a Green's function for the device alone, one for the exterior region and a coupling function to finally arrive at an expression for an open quantum system

$$[E - H - \Sigma^R(x, x', E)] G^R(x, x', E) = \delta(x - x'), \quad x, x' \in (x_1, x_N). \quad (4.2)$$

Here we introduced a new quantity Σ^R known as the retarded self-energy. For the advanced functions the equation is completely analogous and every label R is replaced by an A . Now, H represents the Hamiltonian of the isolated device and Σ^R incorporates the interaction of the device region with the contacts. When assuming that the single leads are independent of each other, one can split up the self-energy in the single contributions

$$\Sigma^R = \sum_{\alpha=1}^2 \Sigma_{\alpha}^R. \quad (4.3)$$

The specific Σ_{α}^R are then given by [38]

$$\Sigma_{\alpha}^R = \tau_{\alpha}^{\dagger} g_{\alpha}^R \tau_{\alpha}, \quad (4.4)$$

with g_{α}^R representing Green's function of lead α and τ_{α} the coupling between the particular lead and the device. Due to the fact that g_{α}^R is given by a simple analytical expression for the case of non-interacting, semi-infinite leads and in addition, that τ_{α} couples only adjacent points of g_{α}^R and G^R , i.e. the surface terms, the Σ_{α}^R can be calculated exactly and are thus known functions. Therefore, Eq. (4.2) together with Eq. (4.3) completely determines $G^R(x, x', E)$ and the problem of finding Green's function on an infinite domain reduces to calculating $G^R(x, x', E)$ on a compact domain, i.e. for points x, x' inside the device, whereby the original Hamiltonian is replaced by the effective one, $H + \Sigma^R$ [38]. The reduction to a compact domain is essential since it enables to approximate Eq. (4.2) by a matrix expression when introducing grid points

$$x_j = x_1 + (j - 1)\Delta x, \quad x_1, \Delta x \in \mathbb{R}, \quad j = 1, \dots, N, \quad (4.5)$$

and making use of a finite differences (FD) or finite element (FE) approximation [15] to replace the operator H by a matrix. Datta [38] introduces a tight-binding Hamiltonian to rewrite H in matrix form, which results in the same matrix elements as in the case of the simplest FD-approximation, as presented in the next section. The matrices obtained by discretizing Eq. (4.2) are of finite size $N \times N$ and thus directly enable a numerical calculation of G^R . Upon discretization, the delta function $\delta(x - x')$ is replaced by the identity matrix $\mathbb{1}$ such that G^R can be calculated by a matrix inversion [38]

$$G^R = (E\mathbb{1} - H - \Sigma^R)^{-1}. \quad (4.6)$$

The advanced Green's function G^A and self-energy Σ^A may be obtained by calculating the Hermitian conjugate

$$\begin{aligned} G^A &= (G^R)^\dagger, \\ \Sigma^A &= (\Sigma^R)^\dagger. \end{aligned} \quad (4.7)$$

It is convenient to introduce the functions

$$\begin{aligned} \Gamma_\alpha &= i(\Sigma_\alpha^R - \Sigma_\alpha^A), \\ \Gamma &= \sum_{\alpha=1}^2 \Gamma_\alpha, \end{aligned} \quad (4.8)$$

representing the coupling of the device to the leads $\alpha = 1, 2$ and the spectral function A [38]

$$\begin{aligned} A &= i(G^R - G^A) \\ &= G^R \Gamma G^A \\ &= G^A \Gamma G^R. \end{aligned} \quad (4.9)$$

The spectral function enables to calculate the local density of states at point x_j out of the knowledge of the j -th diagonal element of A

$$\rho(x_j, E) = \frac{1}{2\pi} A_{jj}(E), \quad (4.10)$$

as well as the total density of states inside the device when taking the trace of A , i.e. the sum of the diagonal elements of A

$$D(E) = \frac{1}{2\pi} \text{Tr}[A(E)]. \quad (4.11)$$

In equilibrium the local density of states together with the Fermi-Dirac distribution suffices to calculate the charge density $n(x)$ at a certain position. In the non-equilibrium case we need a second set of functions referred to as the lesser and greater Green's functions and self-energies, also known as correlation and scattering functions, respectively [38]. For the case of equilibrium situations spatial correlations are described in the framework of the density matrix [4] [38]. One can generalize this concept to non-equilibrium situations when considering correlations in space and time by introducing a function $G^<(x, x', t, t')$, depending on two spatial and two time coordinates. For the steady state case no dependence on absolute time is present and it suffices to consider only the time difference $t - t'$. Upon Fourier transformation $t - t'$ may be replaced by a frequency or energy dependence. To calculate the particle density $n(x)$ out of $G^<$ we need to consider the limit $x = x', t = t'$ which may be written in terms of an integration with respect to E as [38] [39]

$$n(x) = \frac{-i}{2\pi} \int G^<(x, x, E) dE. \quad (4.12)$$

The second quantity of interest for device simulations, the current density $j(x)$ is given by [39]

$$j(x) = -\frac{\hbar q}{2m^*} \int \frac{1}{2\pi} [(\nabla - \nabla')G^<(x, x', E)]|_{x'=x} dE. \quad (4.13)$$

In the Keldysh formalism a kinetic equation is derived which puts the different Green's functions and self energies in relation. For the steady state case the kinetic equation reduces to [38]

$$\begin{aligned} G^< &= G^R \Sigma^< G^A, \\ G^> &= G^R \Sigma^> G^A. \end{aligned} \quad (4.14)$$

The functions $\Sigma^<$ and $\Sigma^>$ are termed lesser and greater self-energies, respectively, or scattering functions [38] [39]. Datta [38] points out that in analogy to $G^<$ representing a correlation function for electrons, $G^>$ can be interpreted as a hole correlation function.

Up to now we introduced self-energies only to account for the interaction of electrons inside the device with the contacts. But the concept is far more general and allows one to include particle-particle interactions such as electron-phonon, electron-electron or electron-impurity scattering mechanisms in the same framework. The only difference is that one cannot write down an analytic expression for the self-energies in the case of particle-particle interactions anymore and one has to rely on a perturbative expansion instead. The lowest-order approximations for electron-electron and electron-phonon interactions may be found in the textbook of Datta [38]. In general, the self-energies may be written as a sum of the different contributions

$$\begin{aligned} \Sigma^R &= \Sigma_\phi^R + \sum_{\alpha=1}^2 \Sigma_\alpha^R, \\ \Sigma^< &= \Sigma_\phi^< + \sum_{\alpha=1}^2 \Sigma_\alpha^<, \\ \Sigma^> &= \Sigma_\phi^> + \sum_{\alpha=1}^2 \Sigma_\alpha^>, \end{aligned} \quad (4.15)$$

where the index ϕ labels self-energies due to phase-breaking processes, i.e. due to scattering events [38]. These self-energies depend in general on the occupancy of the particular states and therefore on $G^<$ and $G^>$, with the consequence that one needs to solve the equations iteratively. The self-energies due to the contacts are related to Green's functions of the leads, which may be rewritten as

$$\begin{aligned} (\Sigma_\alpha^R)_{i,j} &= (\tau_\alpha^\dagger g_\alpha^R \tau_\alpha)_{i,j} \\ &= t^2 (g_\alpha^R)_{\alpha_i, \alpha_j}, \end{aligned} \quad (4.16)$$

where the index α_i, α_j labels positions in the leads adjacent to the positions i, j inside the device and t represents the particular matrix elements of the coupling matrix τ_α . Furthermore, Datta [38] lists the following explicit expressions for the case of semi-infinite, non-interacting leads

$$(\Sigma_\alpha^R)_{i,j} = -t \sum_{m \in \alpha} \chi_m(\alpha_i) \exp(ik_m \Delta x) \chi_m(\alpha_j), \quad (4.17)$$

with

$$t = \frac{\hbar^2}{2m^* \Delta x^2} \quad (4.18)$$

and χ_m and k_m labeling the transverse mode-profile and wave number in mode m of contact α , respectively. The lesser and greater self-energies due to the contacts are then given by [38]

$$\begin{aligned} \Sigma_\alpha^<(E) &= if_\alpha(E)\Gamma_\alpha(E), \\ \Sigma_\alpha^>(E) &= -i[1 - f_\alpha(E)]\Gamma_\alpha(E), \end{aligned} \quad (4.19)$$

with $f_\alpha(E)$ representing the Fermi-Dirac distribution in contact α .

Overall, the solution procedure consists of the following steps [38]:

1. Calculate the contributions to the self-energies as listed in Eq. (4.15). The contact self-energies can be calculated directly with Eqs. (4.17) and (4.19) but the self-energies due to scattering events depend on $G^<$ and $G^>$ and need to be updated after every iteration step.
2. From the knowledge of the retarded self-energy, the retarded and advanced Green's function are obtained by Eqs. (4.6) and (4.7).
3. The kinetic equation Eq. (4.14) enables us to calculate the lesser and greater Green's functions. With these it is possible to update the scattering terms in the self-energies and the steps 1 - 3 have to be repeated until convergence is reached.
4. Finally, one can calculate the particle density $n(x)$ and the current density $j(x)$.

For more details we refer to the textbook of Datta [38, Sec. 8.5].

In place of the current density one can also calculate the terminal currents via the expression [15] [38] [7]

$$I_\alpha = -i \frac{q}{2\pi\hbar} \int \text{Tr} [\Sigma_\alpha^<(E)A(E) - \Gamma_\alpha(E)G^<(E)] dE. \quad (4.20)$$

4.1.1 Coherent transport regime

In the following we will state the governing equations to describe coherent transport without scattering effects in detail. In this special case the equations simplify further

since the self-energies are known exactly and an iterative calculation procedure is not required. For the case of a single mode m in the leads, Eq. (4.17) reduces to [15] [6]

$$\begin{aligned} (\Sigma_1^R)_{i,j} &= -t \exp(ik_1 \Delta x) \delta_{i,1} \delta_{j,1}, \\ (\Sigma_2^R)_{i,j} &= -t \exp(ik_2 \Delta x) \delta_{i,N} \delta_{j,N}, \end{aligned} \quad (4.21)$$

for $i, j = 1, \dots, N$ and with the wave number for each contact given by

$$k_\alpha = \frac{\sqrt{2m^*(E - v^\alpha)}}{\hbar}. \quad (4.22)$$

Here we assumed that the potential in each of the contacts has a constant value v^α . Inserting Eq. (4.21) into Eq. (4.8) results in

$$\begin{aligned} (\Gamma_1)_{i,j} &= 2t \sin(k_1 \Delta x) \delta_{i,1} \delta_{j,1}, \\ (\Gamma_2)_{i,j} &= 2t \sin(k_2 \Delta x) \delta_{i,N} \delta_{j,N}, \end{aligned} \quad (4.23)$$

which directly determines the lesser and greater self-energies for each contact by Eq. (4.19). When neglecting scattering effects all phase-breaking terms Σ_ϕ^R , $\Sigma_\phi^<$ and $\Sigma_\phi^>$ in Eq. (4.15) are zero and the self-energies are thus completely known.

For the device Hamiltonian H we assume a constant effective mass m^* , so that it may be written in the position space as

$$H = -\frac{\hbar^2}{2m^*} \partial_x^2 + qV(x). \quad (4.24)$$

To discretize the operator we use the following finite differences approximation for the kinetic term [15]

$$\partial_x^2 f(x) \Big|_{x=x_j} \approx \frac{f(x_{j-1}) - 2f(x_j) + f(x_{j+1}))}{\Delta x^2}, \quad (4.25)$$

and for the potential we simply evaluate $V(x)$ at positions x_j to arrive at the matrix elements

$$(H)_{i,j} = qV(x_j) \delta_{i,j} - t (\delta_{i,j-1} - 2\delta_{i,j} + \delta_{i,j+1}), \quad (4.26)$$

with t as defined in Eq. (4.18). The equations stated so far already enable us to calculate the matrix G^R for different energies, see Eq. (4.6). Due to the fact that all self-energies are known from the start, one can make use of the kinetic equation to express the quantities $n(x)$ and I_α directly in terms of G^R and Γ_α without the need to explicitly calculate $G^<$. This is useful, since as we will see below, only a small number of matrix elements of G^R is actually needed to calculate the quantities of interest and one thus does not need to calculate the full inverse in Eq. (4.6).

To express the particle density $n(x)$ in terms of G^R and Γ_α , one can insert in Eq. (4.12) the kinetic equation (4.14) and the definition of the lesser self-energy Eq. (4.19).

For the discrete set of points x_j the particle density is given by [7]

$$\begin{aligned} n(x_j) &= \frac{1}{2\pi\Delta x} \int \{G^R(E) [f_1(E)\Gamma_1(E) + f_2(E)\Gamma_2(E)] G^A(E)\}_{j,j} dE \\ &= \frac{1}{2\pi\Delta x} \int \sum_{\alpha=1}^2 \{G^R(E)\Gamma_{\alpha}(E)[G^R(E)]^{\dagger}\}_{j,j} f_{\alpha}(E) dE. \end{aligned} \quad (4.27)$$

When inserting the explicit expression for Γ_{α} as given in Eq. (4.23) one can find

$$\begin{aligned} \{G^R(E)\Gamma_1(E)[G^R(E)]^{\dagger}\}_{j,j} &= 2t \sin [k_1(E)\Delta x] |[G^R(E)]_{j,1}|^2, \\ \{G^R(E)\Gamma_2(E)[G^R(E)]^{\dagger}\}_{j,j} &= 2t \sin [k_2(E)\Delta x] |[G^R(E)]_{j,N}|^2, \end{aligned} \quad (4.28)$$

so that it suffices to know the first and last column of G^R to calculate the particle density inside the device.

In the case of coherent transport it is possible to rewrite the expression for the terminal current Eq. (4.20) in terms of a transmission function. We therefore make use of the relation of the spectral function A to G^R , G^A and Γ , Eq. (4.9), as well as Eqs. (4.14) and (4.19) to find for the case $\alpha = 1$

$$\begin{aligned} \Sigma_1^< A - \Gamma_1 G^< &= i f_1 \Gamma_1 G^R (\Gamma_1 + \Gamma_2) G^A - \Gamma_1 G^R (i f_1 \Gamma_1 + i f_2 \Gamma_2) G^A \\ &= i (f_1 - f_2) \Gamma_1 G^R \Gamma_2 G^A, \end{aligned} \quad (4.29)$$

where we suppressed the explicit energy dependency for the sake of clarity. Inserting this into Eq. (4.20) makes it obvious to relate the following term to the so-called transmission function

$$T(E) = \text{Tr}[\Gamma_1(E)G^R(E)\Gamma_2(E)G^A(E)], \quad (4.30)$$

so that the current is given by

$$I_1 = \frac{q}{2\pi\hbar} \int T(E)[f_1(E) - f_2(E)] dE. \quad (4.31)$$

For the particular Γ_{α} matrices stated above, the expression for the transmission function reduces to

$$T(E) = 4t^2 \sin [k_1(E)\Delta x] \sin [k_2(E)\Delta x] |[G^R(E)]_{1,N}|^2. \quad (4.32)$$

As can be seen, the knowledge of the single matrix element $[G^R(E)]_{1,N}$ linking the first and last position in the device is enough to determine the current.

As mentioned already in Sec. 3.3, when we discussed the boundary conditions for the solution of the Wigner transport equation, one has to consider for $f_{\alpha}(E)$ the one-dimensional Fermi-Dirac distribution, see also Eq. (3.254). Written in terms of the energy, the appropriate distribution is [36] [7] [40]

$$f_{\alpha}(E) = \frac{m^*}{\pi\hbar^2\beta} \ln \{1 + \exp [-\beta (E - \mu_{\alpha})]\}, \quad (4.33)$$

with μ_α representing the chemical potential in the specific contact and β standing for the thermal energy.

Two crucial numerical aspects still need to be discussed: On the one hand, how to perform the matrix inversion in Eq. (4.6) effectively and on the other hand, how to calculate the integrals with respect to E of Eqs. (4.27) and (4.31). In principle, numerical routines exist to calculate the full inverse matrix as written down in Eq. (4.6). But, due to the fact that in the special case of coherent transport only the first and last row of G^R contribute to the values of $n(x_j)$ and I_1 , more effective strategies can be used. The problem of finding G^R may be written in the form

$$\sum_{k=1}^N (E\mathbb{1} - H - \Sigma^R)_{i,k} (G^R)_{k,j} = \delta_{i,j}, \quad i, j = 1, \dots, N, \quad (4.34)$$

representing a system of N^2 equations. As one can see, the index j enters as an external parameter such that it is possible to simply leave out the terms $j = 2, \dots, N-1$, which are not of interest, and consider only the $2N$ equations

$$\sum_{k=1}^N (E\mathbb{1} - H - \Sigma^R)_{i,k} (G^R)_{k,j} = \delta_{i,j}, \quad i = 1, \dots, N, \quad j = 1, N. \quad (4.35)$$

The remaining problem can then be solved by some standard routine for linear systems of equations to obtain the desired values of $(G^R)_{i,1/N}$. This procedure has to be performed for multiple values of E to numerically calculate the integrals in Eqs. (4.27) and (4.31). Due to the fact that the integrands are not smooth in general but may show sharp peaks, a sophisticated method preferably with an included error estimation and control is required to obtain reliable values for $n(x_j)$ and I_1 . The sharp features are very prominent in the case of resonant tunneling diodes and one can see that the transmission function $T(E)$ exhibits narrow peaks at the energy of the resonant state inside the well. A good overview and comparison of different numerical integrators can be found in [39] and the results encouraged us to make use of a double adaptive quadrature routine. In particular the routine *coteda* developed by Espelid [41] and available online as Matlab code was used for all NEGF simulations in this work. The routine applies both a five point and a nine point closed Newton-Cote rule in combination with a locally double adaptive strategy. Double adaptive means in this context that on the one hand, for a given set of sampling points of the function in question, different Newton-Cote rules are used to perform the integration, in which the algorithm changes from lower-order to higher-order approximations together with monitoring the error estimate and, on the other hand, if the estimated error did not fulfill a required tolerance, new sampling points are added. The subdivision of the integration interval is done in the case of the *coteda*-routine by bisection so that the five point and nine point Newton-Cote rules can make use of the same set of sampling points which increases the efficiency of the algorithm. [41]

4.2 Simulation of Resonant Tunneling Diodes

In the following, simulations with the NEGF technique are presented for the same resonant tunneling diode, consisting of an $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}/\text{Al}_x\text{Ga}_{1-x}\text{As}$ heterostructure, as used in the next chapter for the WTE calculations. The heterostructure itself is discussed in more detail in Sec. 5.2 and here we simply consider the corresponding potential shape $qV(x)$ as depicted in Fig. 4.1. Herein the height of the barriers is fixed to a value of 0.27 eV , as well as the width by $5a$, the length of the slopes of the barriers by a and the spacing of the barriers by $8a$, expressed in terms of the lattice constant $a = 0.565\text{ nm}$ [4] [6]. The source-drain voltage V_{DS} across the device is varied and for each situation the steady state is calculated, so that one is able to determine a $I(V)$ curve for the particular device. Characteristic for a RTD is a region with negative differential resistance.

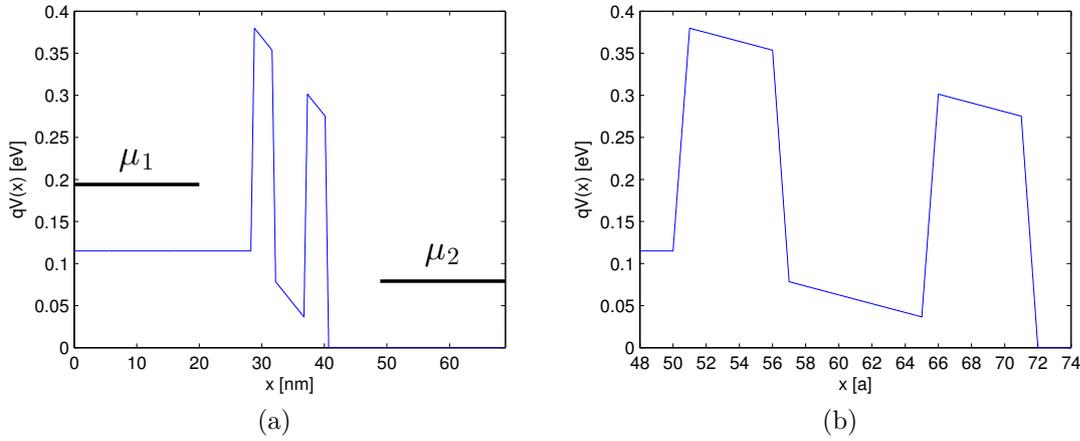


Figure 4.1: Illustration of the considered potential $qV(x)$ of a resonant tunneling diode for a barrier height of 0.27 eV (a). Depicted is $qV(x)$ together with the chemical potentials $\mu_1 = 0.19\text{ eV}$ and $\mu_2 = 0.08\text{ eV}$ at each contact for the case of a bias voltage of $V_{DS} = 0.115\text{ V}$. In (b) a detailed plot of the double barriers is shown and the position coordinate is expressed in terms of the lattice constant $a = 0.565\text{ nm}$.

Furthermore, we assume for the simulation a spatially constant effective mass given by the value for GaAs, $m^* = 0.067 m_e$ [6], a temperature of $T = 300\text{ K}$ and a donor density inside the contacts of $N_D = 2 \times 10^{18}\text{ cm}^{-3}$ [6]. From this one can calculate the corresponding chemical potentials μ_α in each contact by finding the root of the equation [4]

$$\frac{N_D}{N_C} = \mathcal{F}_{1/2}[\beta(\mu_\alpha - v^\alpha)], \quad (4.36)$$

with the potential energy v^α at the contact α , $\beta = \frac{1}{k_B T}$, $\mathcal{F}_{1/2}$ denoting the Fermi-Dirac

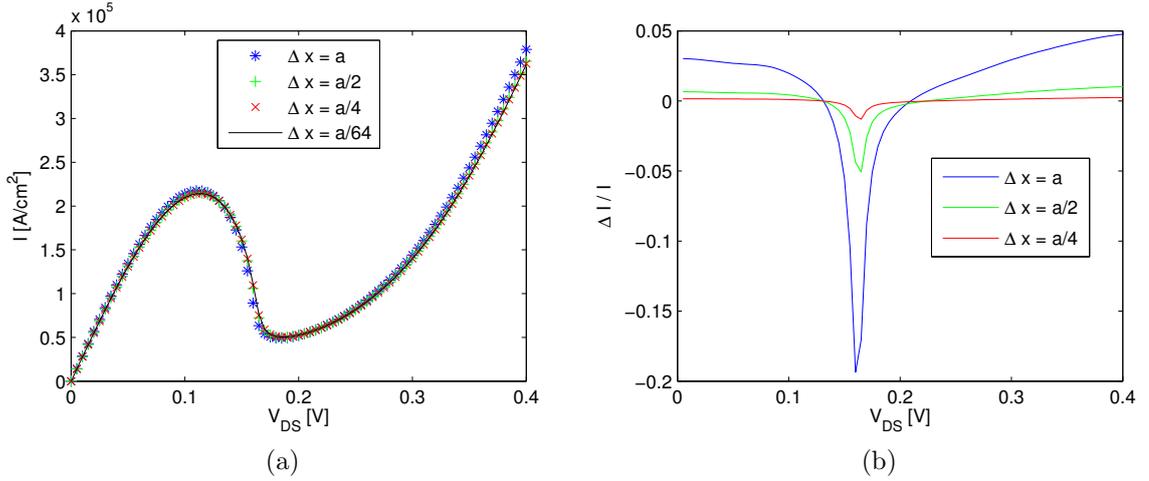


Figure 4.2: Calculated $I(V)$ curves of a RTD as illustrated in Fig. 4.1 for various grid spacings Δx (a). The case $\Delta x = a/64$ is considered to be a reference solution for the simulation on coarser grids. In (b) a plot of the relative error is shown.

integral of order $\frac{1}{2}$ [4] [42] and N_C given by [4]

$$N_C = 2 \left(\frac{m^*}{2\pi\hbar^2\beta} \right)^{\frac{3}{2}}. \quad (4.37)$$

In Fig. 4.1 the calculated chemical potentials $\mu_1 = 0.19 \text{ eV}$ and $\mu_2 = 0.08 \text{ eV}$ for the case $V_{DS} = 0.115 \text{ V}$ are shown as well.

In Fig. 4.2 the calculated $I(V)$ curves for different grid spacings Δx are shown, where the current was calculated by making use of Eq. (4.31). The characteristic feature of a RTD, the negative differential resistance is obvious to see. The peak value is taken on at a voltage of about $V_{DS} = 0.115 \text{ V}$ and the valley is located at approximately $V_{DS} = 0.185 \text{ V}$. The peak corresponds to the situation, where the incoming electron distribution from contact 1 shows the best possible energy alignment with the quasi-bound well-state, so that resonant tunneling processes are enhanced. In contrast, at the valley of the $I(V)$ curve the conduction band edge $v^1 = qV_{DS}$ exceeds the energy of the well-state. This causes that the incoming electrons from contact 1 cannot undergo a resonant tunneling process anymore and therefore, a reduction of the current. A further increase of V_{DS} results in an increase of the current due to the enhanced tunneling probability and the stronger acceleration of electrons.

To understand the accessibility of the well-state to electrons from the contacts, it is illustrative to compare the Fermi-distributions $f_\alpha(E)$ of each contact and the transmission function $T(E)$, as depicted in Fig. 4.3. For the resonant case $V_{DS} = 0.115 \text{ V}$ one can see that $T(E)$ contains a sharp spike, corresponding to the resonant tunneling process through the well-state and in addition that $f_1(E)$ exhibits large values at the energy of that spike. In contrast, in the case of $V_{DS} = 0.185 \text{ V}$ the sharp

spike in $T(E)$ has vanished, i.e. the well-state has become inaccessible and the overlap of $T(E)$ and $f_1(E)$ is much smaller. In both cases $T(E)$ shows an oscillating behaviour at higher energies as well. Due to the width of these peaks one can assume that they do not correspond to resonances with quasi-bound well-states but to interference effects. But, this high-energy behaviour is not of concern for the range of bias voltages discussed here and will not be looked at in more detail.

From the numerical point of view, the following parameters are crucial for accurate results. The upper limit and the error tolerance for the energy integration must be chosen appropriately, as well as the particular grid spacing and included length of the contacts. To perform the energy integration it was decided after various tests to choose an upper limit of $E_{max} = 1$ eV and a relative tolerance of 10^{-7} for the *coteda*-routine. From Fig. 4.3 it should be apparent that the choice of E_{max} is sufficient since the Fermi-distributions $f_\alpha(E)$ exhibit significant values only at much lower energies than E_{max} . The contact length L_c was varied in different simulations and the results revealed only a minor dependence on the particular value of L_c , as a consequence of the analytically evaluated self-energies for the contacts. The differences in the corresponding $I(V)$ curves were orders of magnitude smaller than those obtained for different values of Δx . Therefore, the particular value of the contact length was fixed to $L_c = 50 a$, for the purpose of calculating $n(x)$ on a sufficiently large interval for the later comparison with the WTE calculations, and kept the same for all simulations. To estimate the accuracy of the results obtained with a particular grid spacing, the order of convergence of the algorithm has been calculated. In Fig. 4.2 (a) the $I(V)$ curves for three relatively coarse grid spacings and a fairly fine grid spacing of $\Delta x = a/64$ are presented. The latter one serves as a reference solution. Figure 4.2 (b) displays a plot of the relative difference $\Delta I/I$ for the three coarse grid spacings compared to the case $\Delta x = a/64$. It is apparent that the error declines rapidly. To quantify this, the 2-norm error and the order of convergence are listed in Tab. 4.1. From this test it is obvious to ascribe the algorithm a convergence of second order with respect to Δx .

Table 4.1: Estimation of the order of convergence $\mathcal{O}_{\Delta x}$ out of the simulated $I(V)$ curves as shown in Fig. 4.2. Compared is the 2-norm error as defined in Eq. (3.221) and the order of convergence determined by the logarithm to the base of 2 of the quotient of two subsequent 2-norm errors, Eq. (3.223).

Δx	$\ \Delta I\ _2$	$\mathcal{O}_{\Delta x}$
a	4334	2.20
$a/2$	945	2.05
$a/4$	228	

Since we are not only interested in the current but also in the particle density, analogous convergence tests were performed for $n(x)$ for the cases $V_{DS} = 0.115$ V and $V_{DS} = 0.185$ V and the results are presented in Fig. 4.4 and Tab. 4.2. Again, the order of convergence can be very well estimated by two. In principle this is just a verification of an expected result since the finite difference approximation, Eq. (4.25),

used in discretizing the Hamiltonian is second-order accurate. We can assume that the results obtained with $\Delta x = a/64$ are well converged and can be viewed as an exact solution for the purpose of evaluating the Wigner function calculations.

Table 4.2: Estimation of the order of convergence $\mathcal{O}_{\Delta x}$ from the $n(x)$ -calculations as illustrated in Fig. 4.4. The 2-norm error and order of convergence are determined in the same manner as described for Tab. 4.1.

	$V_{DS} = 0.115 \text{ V}$		$V_{DS} = 0.185 \text{ V}$	
Δx	$\ \Delta n(x)\ _2$	$\mathcal{O}_{\Delta x}$	$\ \Delta n(x)\ _2$	$\mathcal{O}_{\Delta x}$
a	8.13E16	2.05	6.76E16	2.06
$a/2$	1.97E16	2.02	1.62E16	2.02
$a/4$	0.49E16		0.40E16	

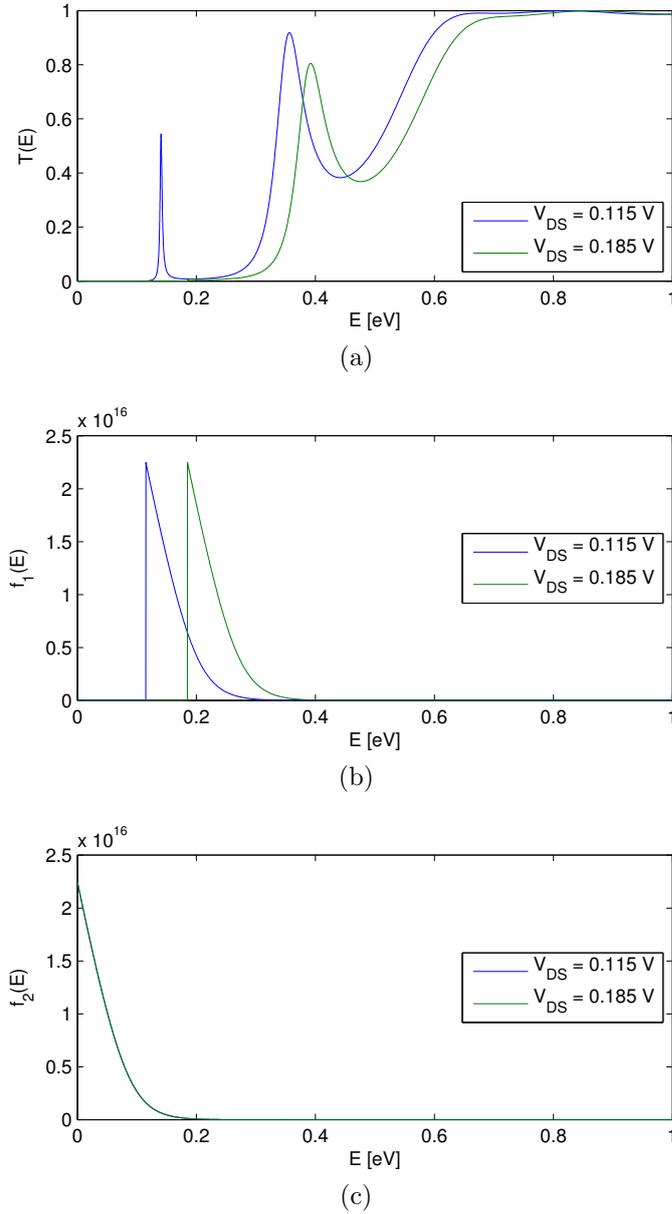


Figure 4.3: Transmission function $T(E)$ (a) and Fermi distributions $f_{1/2}(E)$ of the left (b) and right (c) contact for the peak voltage $V_{DS} = 0.115$ V and the valley voltage $V_{DS} = 0.185$ V. The single parameters are chosen as described in the text and shown are the results for the calculation with $\Delta x = a/64$. The particular values of the chemical potentials are for the case $V_{DS} = 0.115$ V: $\mu_1 = 0.194$ eV, $\mu_2 = 0.079$ eV and for $V_{DS} = 0.185$ V: $\mu_1 = 0.264$ eV and $\mu_2 = 0.079$ eV.

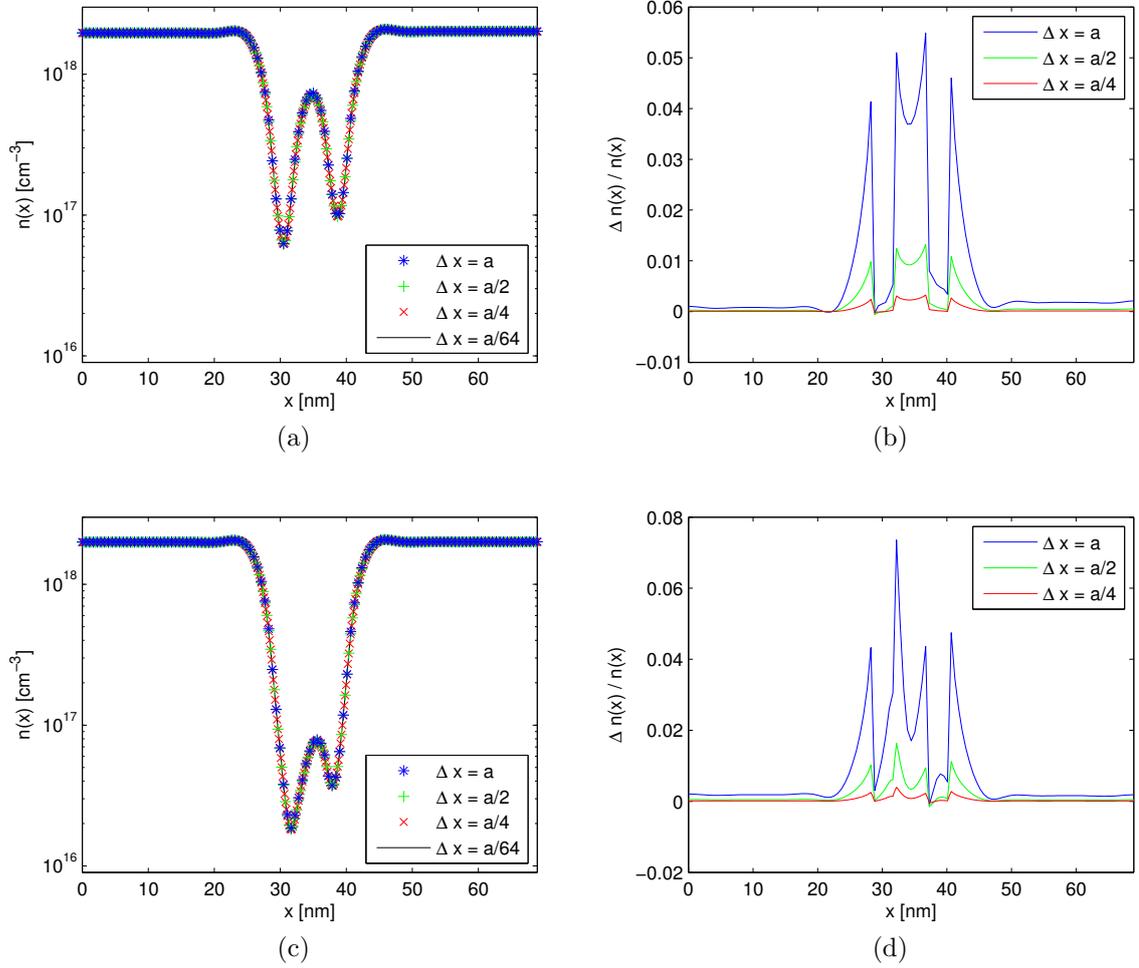


Figure 4.4: Calculated particle density $n(x)$ of a RTD for the cases $V_{DS} = 0.115$ V in (a) and $V_{DS} = 0.185$ V in (c) for various grid spacings Δx . The case $\Delta x = a/64$ is considered to be a reference solution for the simulation on coarser grids. The relative errors are shown in (b) for $V_{DS} = 0.115$ V and in (d) for $V_{DS} = 0.185$ V.

5 Simulation of Resonant Tunneling Diodes in phase space

5.1 Numerical tests on the time evolution to steady states

The Wigner transport equation (WTE) is a time evolution equation for the quasi-distribution function $f(p, x, t)$. It can be used to simulate the carrier transport in the steady state. One starts with an arbitrary initial distribution for $f(p, x, t)$ and proceeds until a stationary situation is reached. Of course, a smart choice of the initial condition may reduce the time needed to reach the desired steady state. But, the problem is not as easy as stated so far since an inappropriate choice of the (p, x) -grid may cause large errors and lead to unphysical results, if a stationary situation is reached at all. The problem arises due to the fact that $f(p, x, t)$ may show heavily oscillating regions for steady states where tunneling and coherence phenomena are prominent. This is exactly the case for resonant tunneling diodes. One needs to carefully make sure that a chosen grid is appropriate for the simulated, physical situation. If not, one encounters that in the progress of the time evolution to the stationary state large errors build up as soon as the oscillations become too short scaled to be resolved by the chosen grid. These errors are then not confined to the particular region but may spread out in the phase space and influence the overall behaviour of the device. To become familiar with this peculiar behaviour of the algorithm and to understand its origin and manifestations, various time evolutions are analyzed in the following. In addition, the piecewise constant and the piecewise linear approximation of the Wigner function are compared to see to what degree a higher-order method may improve the situation.

In all simulations performed on resonant tunneling diodes, the spacing of the p -grid was the crucial factor deciding whether a smooth convergent or oscillating behaviour was observed. The spacing of the x -grid was comparatively unproblematic. Therefore, we will compare in the following different p -grid spacings for two different values of Δx . For simplicity, all physical constants are set to one and we make use of the same p -grid (Eq. (3.218)) as applied before in Sec. 3.2.4 when studying the Fourier transform of a piecewise polynomial approximation. The x grid is chosen to be equidistant and takes on the values 0 and 300 at the contact on the left-hand side and the right-hand side, respectively. The potential $V(x)$ is assumed to be constant for values $x < 100$ and $x > 150$ and exhibits a double barrier shape superimposed with a linear potential drop in the region $100 \leq x \leq 150$, as depicted in Fig. 5.5 (c). The linear potential drop is

assumed to result from a bias voltage of magnitude $V_{DS} = 0.05$ a.u. and the barriers have a height of 0.15 a.u.. The barrier width and the well width have an equal value of 3 a.u. and a finite slope over a distance of 1 a.u. is included for the barriers. For the boundary condition, not the usual Fermi-Dirac distribution but a displaced Gaussian distribution is chosen. The center of the Gaussian is set to $\mu = +(-)0.32$ a.u. of the left (right) contact and the width to $\sigma = 0.05$ a.u. (see e.g. Eq. (3.216) with $\lambda = 0$). This has the advantage that the boundary distribution does not exhibit significant values around $p = 0$ which in turn accelerates the time evolution to a steady state. On the whole, the parameters of the potential $V(x)$ and the boundary conditions are chosen in a way such that resonant tunneling processes are prominent and $f(p, x, t)$ thus exhibits heavily oscillating regions in phase space.

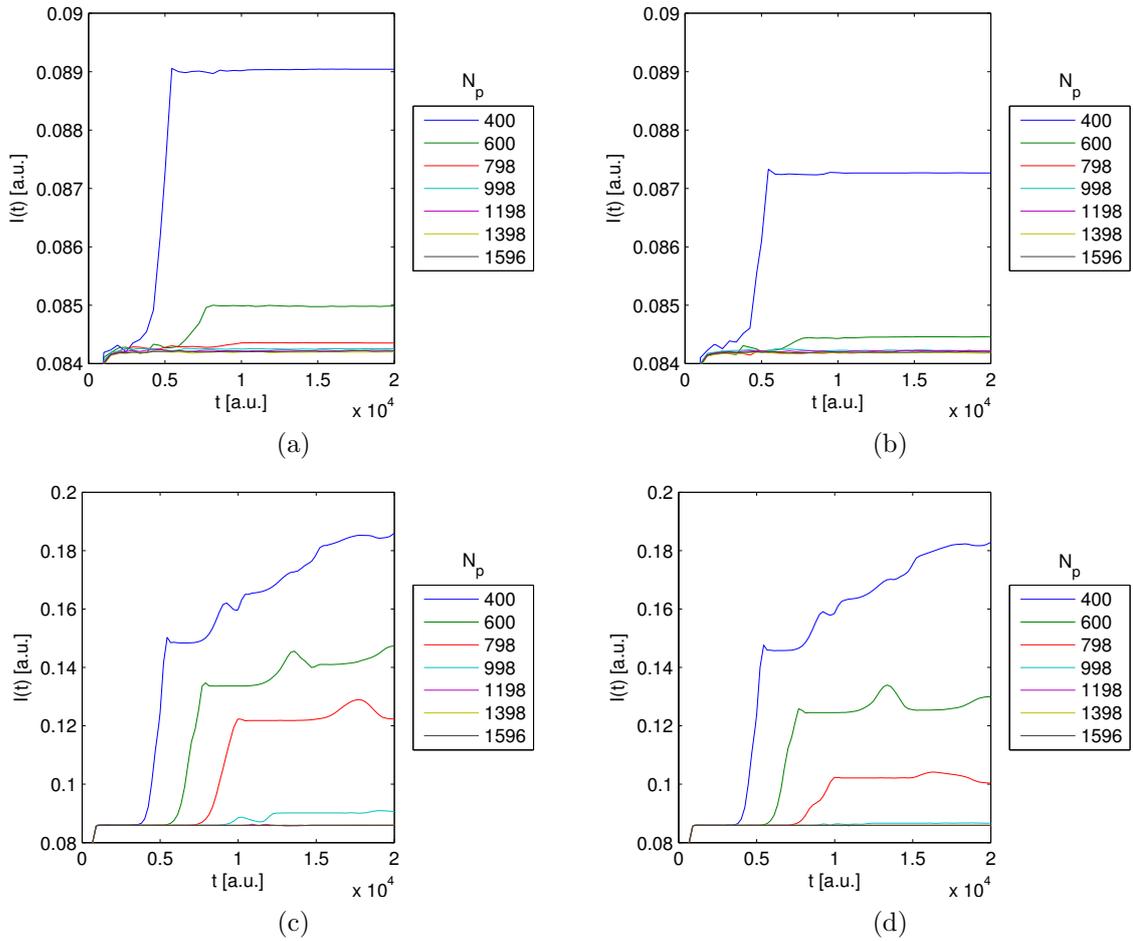


Figure 5.1: Time evolution of the current $I(t)$ calculated with the piecewise constant approximation of the Wigner function for the cases (a) $N_x = 301$ and (c) $N_x = 601$, as well as with the piecewise linear approximation for (b) $N_x = 301$ and (d) $N_x = 601$, each for different values of N_p . The other parameters are chosen as described in the text.

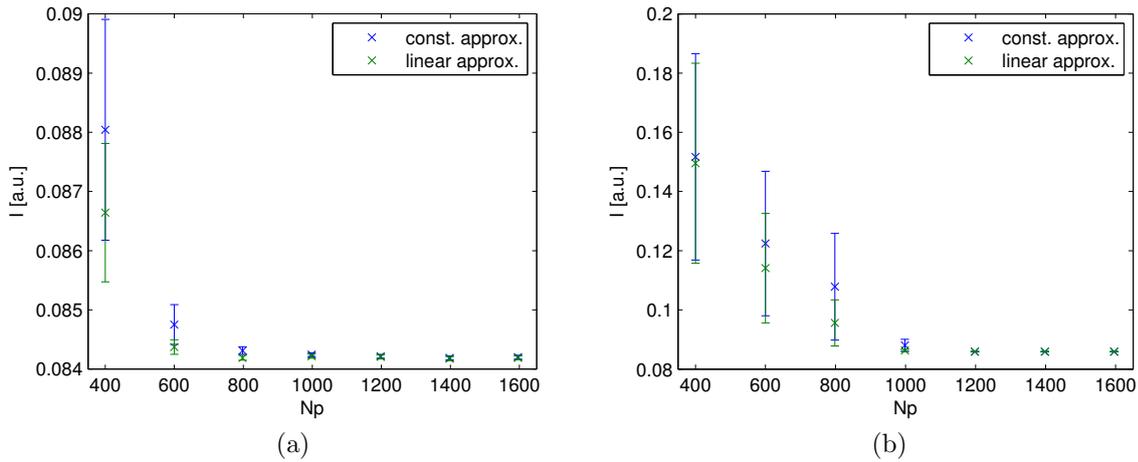


Figure 5.2: Time-averaged current calculated from the simulations presented in Fig. 5.1. Depicted are the averages and the standard deviations of $I(t)$ for the time interval $0.1 \times 10^4 \text{ a.u.} \leq t \leq 2 \times 10^4 \text{ a.u.}$. The piecewise constant and linear approximations are compared with each other in (a) for $N_x = 301$ and (b) for $N_x = 601$.

In Fig. 5.1 the temporal change of the current $I(t)$ is depicted for various p -grid spacings. As initial condition $f(p, x, t=0)$ the Gaussian distribution used as boundary condition is continued from the contact on the left to the position $x = 100 \text{ a.u.}$ and from the right to $x = 150 \text{ a.u.}$. With this choice the time needed to arrive at an approximately stationary situation reduces to about $t = 10^3 \text{ a.u.}$, whereby the simulations are performed to times up to $t = 2 \times 10^4 \text{ a.u.}$. This enables us to examine if a stable stationary state is reached. As one can see in Fig. 5.1, in the case of the lower values $N_p \lesssim 1000$ no smooth convergent behaviour is observed but the $I(t)$ curves display a kink at certain times. For small times, i.e. $1000 \text{ a.u.} \lesssim t \lesssim 3000 \text{ a.u.}$, all simulations show values for I close to the true stationary value, but then, if the number of grid points N_p is chosen too small a rapid change is observed at a certain time followed by an oscillating behaviour in most cases. To understand this behaviour it is elucidating to look at the time evolution of $f(p, x, t)$ in the phase space. During the simulations snapshots of $f(p, x, t)$ were taken at the times $t = 1 \text{ k}, 2 \text{ k}, 5 \text{ k}, 10 \text{ k}, 15 \text{ k}, 20 \text{ k}$, with k representing a factor of 10^3 a.u. . As one can see from Fig. 5.1 (c), the two simulations for the piecewise constant approximation with $N_x = 601$ and $N_p = 400$ on the one hand and $N_p = 798$ on the other hand exhibit a jump in $I(t)$ just before one of these times. Therefore, we examine the two cases as examples in more detail. In Fig. 5.3 one can find illustrations of $f(p, x, t)$, $n(x)$ and $j(x)$ for the case $N_p = 400$ and the two times $t = 2 \text{ k}, 5 \text{ k}$, and in Fig. 5.4 the same is shown for the case $N_p = 798$ and the two times $t = 5 \text{ k}, 10 \text{ k}$.

Before looking at the detailed numerical problems we shortly describe the physical situation to be able to properly read the phase space plots of $f(p, x, t)$ and take for

that purpose Fig. 5.3 (a) as an example. All regions in phase space with $p > 0$ a.u. propagate from the left to the right, i.e. to larger x values and for $p < 0$ a.u. the situation is the other way around. Therefore, on the contact on the left-hand side the part of the phase space with $p > 0$ a.u. flows into the device and the part with $p < 0$ a.u. flows out of the device. On the right-hand side the region $p > 0$ a.u. corresponds to the outflow and $p < 0$ a.u. to the inflow. As mentioned before, Gaussian distributions are used for the boundary conditions and as one can see, the Gaussian shape is maintained in good approximation in the regions $0 \text{ a.u.} < x < 50 \text{ a.u.}$ for $p > 0$ a.u. and $200 \text{ a.u.} < x < 300 \text{ a.u.}$ for $p < 0$ a.u.. After that the interaction with

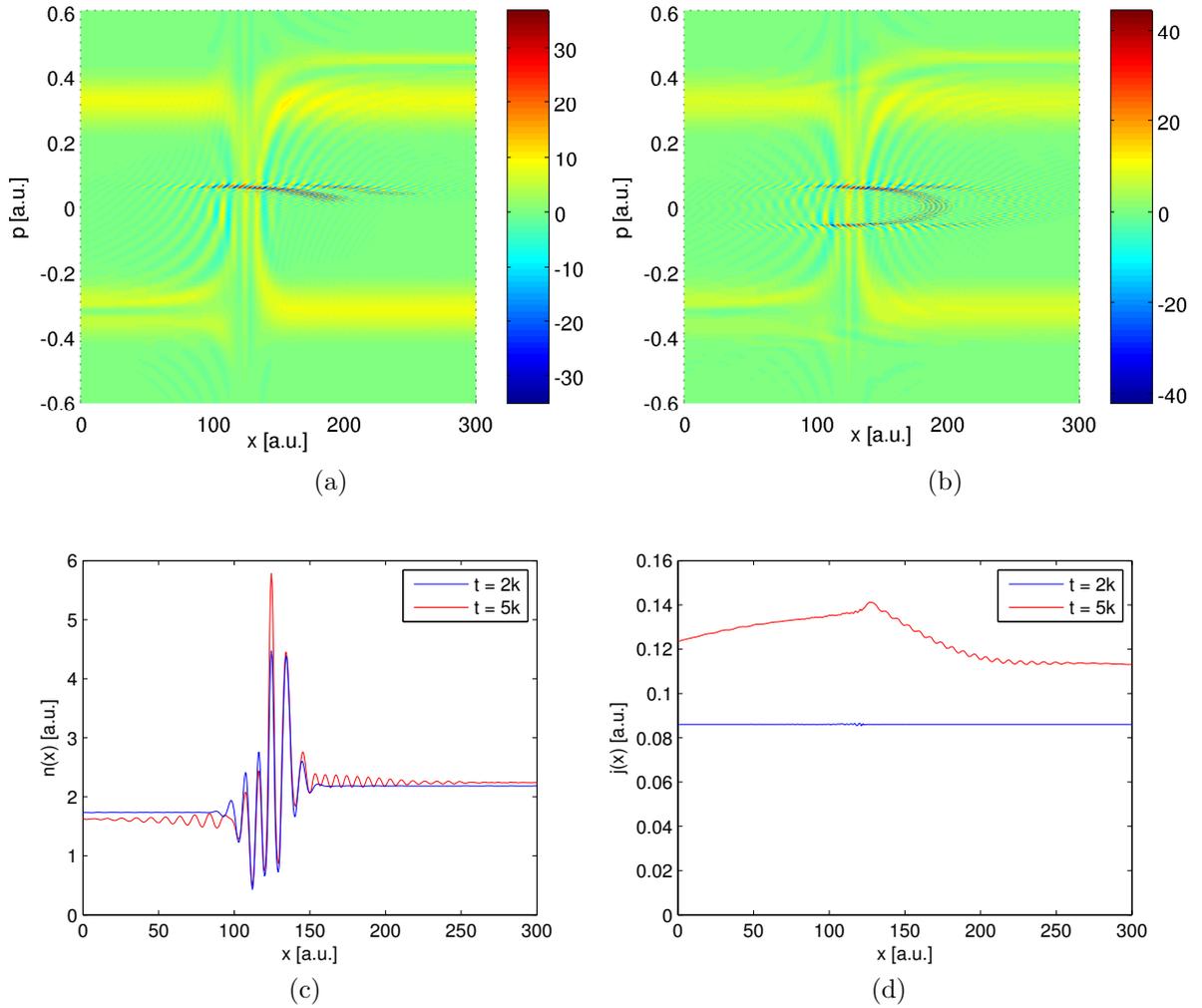


Figure 5.3: Illustrations of $f(p, x, t)$, $n(x)$ and $j(x)$ for the two times $t = 2k, 5k$, resulting from a simulation using the piecewise constant approximation as well as $N_x = 601$ and $N_p = 400$ grid points. A phase space plot of $f(p, x, t)$ is shown in (a) for $t = 2k$ and in (b) for $t = 5k$. In (c) $n(x)$ and in (d) $j(x)$ are displayed for the two times.

the potential becomes significant and strong deviations from the Gaussian shape can be observed. The double barrier is centered at $x \approx 125$ a.u. and one can see strong interference patterns in this region. A part of the distribution tunnels through the first barrier and occupies the quasi-bound well state, represented in this plot by the vertical yellow stripe at $x \approx 125$ a.u. with an approximate extension of -0.5 a.u. $\lesssim p \lesssim 0.5$ a.u.. The other part of the distribution is reflected off the barrier and propagates back to the left-sided contact. As one can see, the reflected beam does not obey a Gaussian shape anymore and seems to be split into two parts. This is because the part of the incoming distribution with momenta $p \approx 0.32$ a.u. is able to undergo a resonant tunneling process and has therefore high probability to tunnel through the double barrier. Therefore, the corresponding momenta are extenuated in the reflected distribution. On the right-hand side of the barrier one can see the outgoing beam after the resonant tunneling process, which is accelerated by the bias voltage to higher momenta of approximately 0.4 a.u. $\lesssim p \lesssim 0.5$ a.u.. In addition to this overall behaviour strong oscillating patterns can be observed. These correspond to interference and coherence phenomena. On the one hand, rather slowly varying wave patterns occur throughout the phase space and on the other hand, a narrow and sharply peaked stripe of oscillations is formed in the region 0 a.u. $\lesssim p \lesssim 0.1$ a.u., starting out in the region of the double barrier and extending with time to larger values of x . It is exactly those stripe of oscillations which causes the numerical problems.

In Fig. 5.5 the stationary solution obtained with a fine enough grid is presented. In (a) the section of phase space with -0.6 a.u. $< p < 0.6$ a.u. and 0 a.u. $< x < 300$ a.u. and in (b) a smaller section focusing on the stripe of oscillations is depicted. One can see on the one hand, that the oscillations vary on very short scales and on the other hand, that the amplitude of these oscillations is quite large and exceeds the values of $f(p, x, t)$ in the rest of the phase space. In addition to that, another problematic feature for numerical simulations is that for larger values of x the variation of $f(p, x, t)$ with respect to p becomes faster. One reason for that is a sort of shear movement: As one can see, the wave lines of the oscillations are roughly vertical in the region of the double barrier but turn more and more in a horizontal direction with increasing values of x and exhibit at $x = 300$ a.u. an angle of about 45° . This observation is consistent with the fact that those parts of $f(p, x, t)$ with larger absolute values of p move faster in phase space, which might serve as an explanation for the observed behaviour even though the detailed dynamics are certainly more involved. Another reason is in this case that the stripe of oscillations splits up in two parts for larger values of x which are shifted along x relative to each other. On the whole, one can say that the stripe of oscillations formed by $f(p, x, t)$ at low values of p requires a very fine p -grid in this region and in addition, the oscillation pattern is not confined to a region close to the barriers but spreads out along x together with becoming more problematic to resolve with increasing values of x .

We now return to the two badly-resolved cases illustrated in the Figs. 5.3 and 5.4. In Fig. 5.3 (a), i.e. at the lower time $t = 2$ k, one can see that the smooth features of $f(p, x, t)$ are already nearly fully formed (comp. e.g. Fig. 5.5 (a)) but the stripe of oscillations is just starting to build up. In the progress of this formation the oscillations

expand to larger x values and at a certain position, in this case at $x \approx 175$ a.u., the spacing of the p -grid does not suffice anymore to resolve the oscillating behaviour properly, so that errors build up and the oscillations evolve in a very different manner than in the case of a well-resolved simulation. Fig. 5.3 (b) displays $f(p, x, t)$ at the later time $t = 5$ k and one is able to observe that the propagation of the oscillation pattern has reversed its direction and returned to the region of the double barrier. As a consequence of this process the oscillating behaviour of the whole phase space domain increases and one can see an effect in the values of the measurable quantities $n(x)$ and $j(x)$. In the case of $n(x)$ additional, spurious oscillations are formed and for the case of $j(x)$ one observes strong spatial variations and the current density begins to rise, starting in the middle region of the barriers and then expanding to the rest of the device. Clearly, the results obtained are erroneous. Not only because a well-resolved solution displays qualitative different solutions for $n(x)$ and $j(x)$ but also since unphysical results might be produced. This was very apparent for a simulated test case in which the boundary condition on the right-hand side was set to zero such that the only inflow was from the contact on the left-hand side. Also in this case one could observe a growth of oscillations in $n(x)$ on badly-resolved grids, similar to what is shown in Fig. 5.3 (c), with oscillations extending far from the region of the barriers to the left and to the right. Only that this situation is not physically correct. The point is that an oscillating pattern in the region on the right-hand side of the barriers is not allowed in this case due to the fact that in this region exists solely a beam of $f(p, x, t)$ moving to the right but none moving to the left. But viewed from the point of Schrödinger's formulation of quantum mechanics, a standing wave pattern can only be formed when a wave moving to the left and one to the right interfere. Therefore, the additional oscillations seen in Fig. 5.3 (c) for larger times can be viewed as spurious and unphysical.

In Fig. 5.4 a simulation for a doubled number of $N_p = 798$ grid points is illustrated. Here, the later times $t = 5$ k and $t = 10$ k are compared. One can observe a qualitative similar behaviour, only that the stripe of oscillations can be properly resolved for longer times and the turning point is now located at a position of about $x \approx 275$ a.u.. Again, once the oscillating pattern is not well-resolved anymore the oscillations expand in an erroneous manner and spread out in phase space regions with $p < 0$ a.u.. As a result, spurious oscillations are formed in $n(x)$ and very different values for the current are observed.

The results presented in Fig. 5.5 can be viewed as properly converged and should come close to the real physical situation. As one can see from the plot shown in (c), the particle density shows a strongly oscillating pattern with two minima at the positions of the barriers and a maximum in between. Therefore, a strong occupation of the well state can be deduced. Furthermore, the constant value of the particle density on the right-hand side is larger than the one on the left-hand side, as a result of the net-flow of particles from the left to the right due to the resonant tunneling together with fixed boundary conditions. More realistic simulations would require a charge neutrality condition at the contacts limiting the inflow of particles, see e.g. [3]. The result for $j(x)$ is a fairly even distribution such that the divergence of $j(x)$ practically vanishes

as desired for the case $\partial_t n(x) = 0$. It should be mentioned that the grid spacings chosen for the simulation of Fig. 5.5 are well-suited for the particular domain but might fail to produce convergent results on a larger x -domain.

After this qualitative discussion we finally turn to the point of assessing the benefit of the piecewise linear approximation compared to the piecewise constant one. Figures 5.1 and 5.2 display the time evolution of $I(t)$ and the time-averaged values of $I(t)$ for different numbers of grid points N_p and two different x -grid spacings. In either case for Δx it is evident that the temporal evolution of $I(t)$ for low values of N_p is improved when using the piecewise linear approximation. On the one hand, the

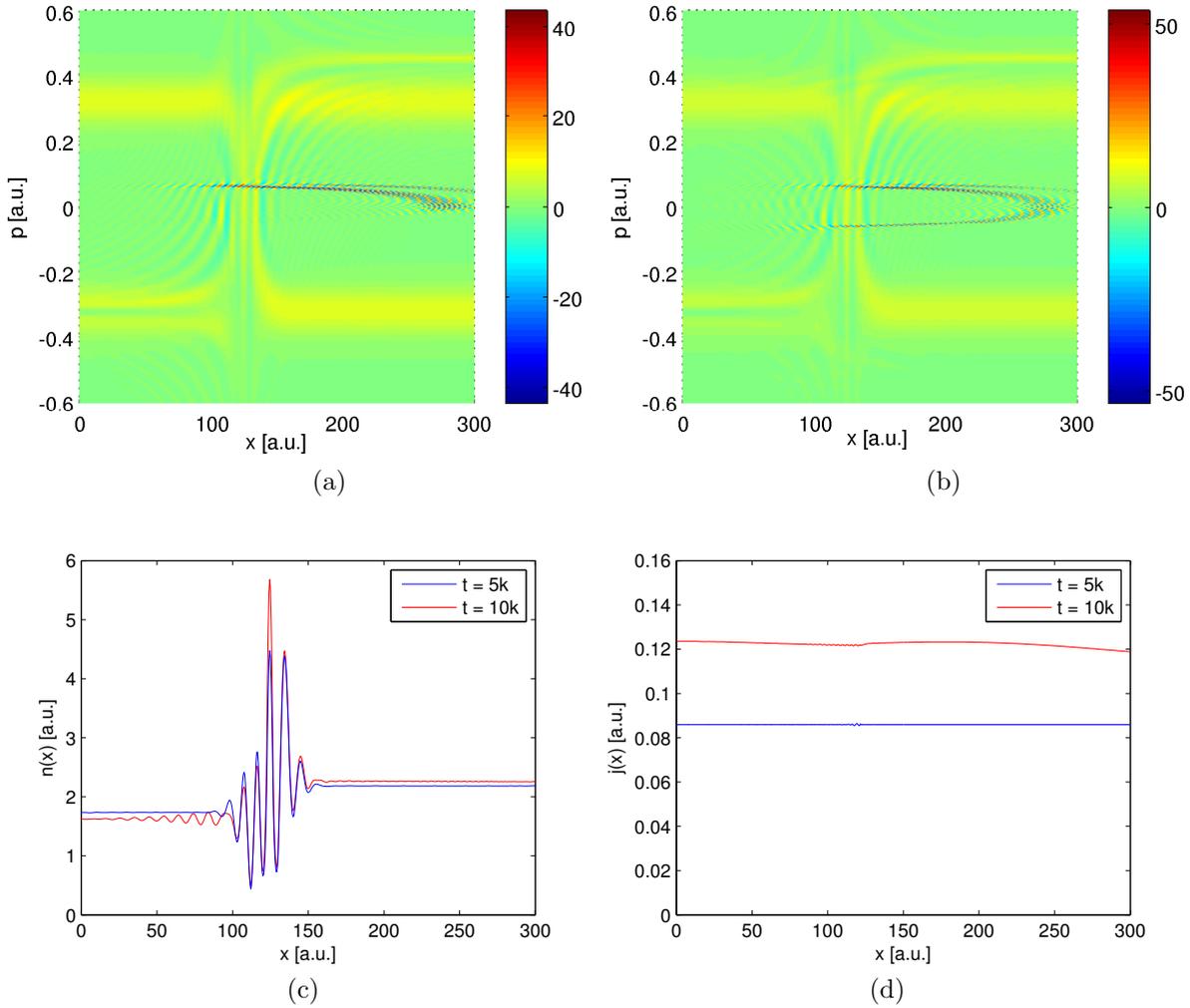


Figure 5.4: Illustrations of $f(p, x, t)$, $n(x)$ and $j(x)$ for the two times $t = 5k, 10k$, resulting from a simulation using the piecewise constant approximation as well as $N_x = 601$ and $N_p = 798$ grid points. A phase space plot of $f(p, x, t)$ is shown in (a) for $t = 2k$ and in (b) for $t = 5k$. In (c) $n(x)$ and in (d) $j(x)$ are displayed for the two times.

size of the jumps in $I(t)$ is mitigated and on the other hand, for simulations in the transient area between well-resolved and badly-resolved situations the application of the piecewise linear approximation may improve the temporal evolution enough to attribute a smooth convergent behaviour, see e.g. the cases $N_p = 798$ in (a),(b) and $N_p = 998$ in (c),(d). Therefore, one can conclude that the approach of using a piecewise polynomial approximation shows the desired tendency that a higher order method enables to produce accurate results on coarser grids.

In principle, one could analyze the degree of improvement of the piecewise linear approximation further by calculating the order of convergence for each of the two methods. This was tried to do, based on the data plotted in Fig. 5.2 but without success. The problem is that the asymptotic behaviour cannot be clearly determined in a logarithmic plot. One can only notice a rapid convergence for under-resolved grids but the results for the properly resolved grids, i.e. $N_p \geq 998$ for $N_x = 301$ and $N_p \geq 1198$ for $N_x = 601$ oscillate too strong to assign meaningful values. The problem of oscillations in the calculated quantities $j(x)$ and $n(x)$ seems to be ever present and their magnitude to depend greatly on the parameters of the (p, x) -grid. The dependency of the results on the other parameters like the length of the x domain, the extension of the p grid and the spacing Δx is analyzed in more detail in the next section. For this purpose a fine enough p grid is chosen and kept the same for most of the simulations. This strategy is justified by the results presented here since in first place it is important to resolve all oscillation patterns in $f(p, x, t)$ properly and a further increase of N_p has only minor effect on the obtained results, compare Fig. 5.2.

Another point worth noting is that a finer grid spacing with respect to x apparently requires a finer p -grid spacing as well, comp. Fig. 5.2. We believe that this is again a consequence of the shear movement of the stripe of oscillations as illustrated in Fig. 5.5 (b). A finer spacing with respect to x may reveal oscillation patterns on smaller scales which, upon advection along x , influence the variation of $f(p, x, t)$ with respect to p as well. It seems that this effect is only significant on coarse grids, as further test simulations indicated and as can also be seen to some degree from the simulations presented in the next section. As soon as Δx has been chosen fine enough to resolve all physical oscillation patterns a further decrease of Δx does not require to adjust the p -grid as well.

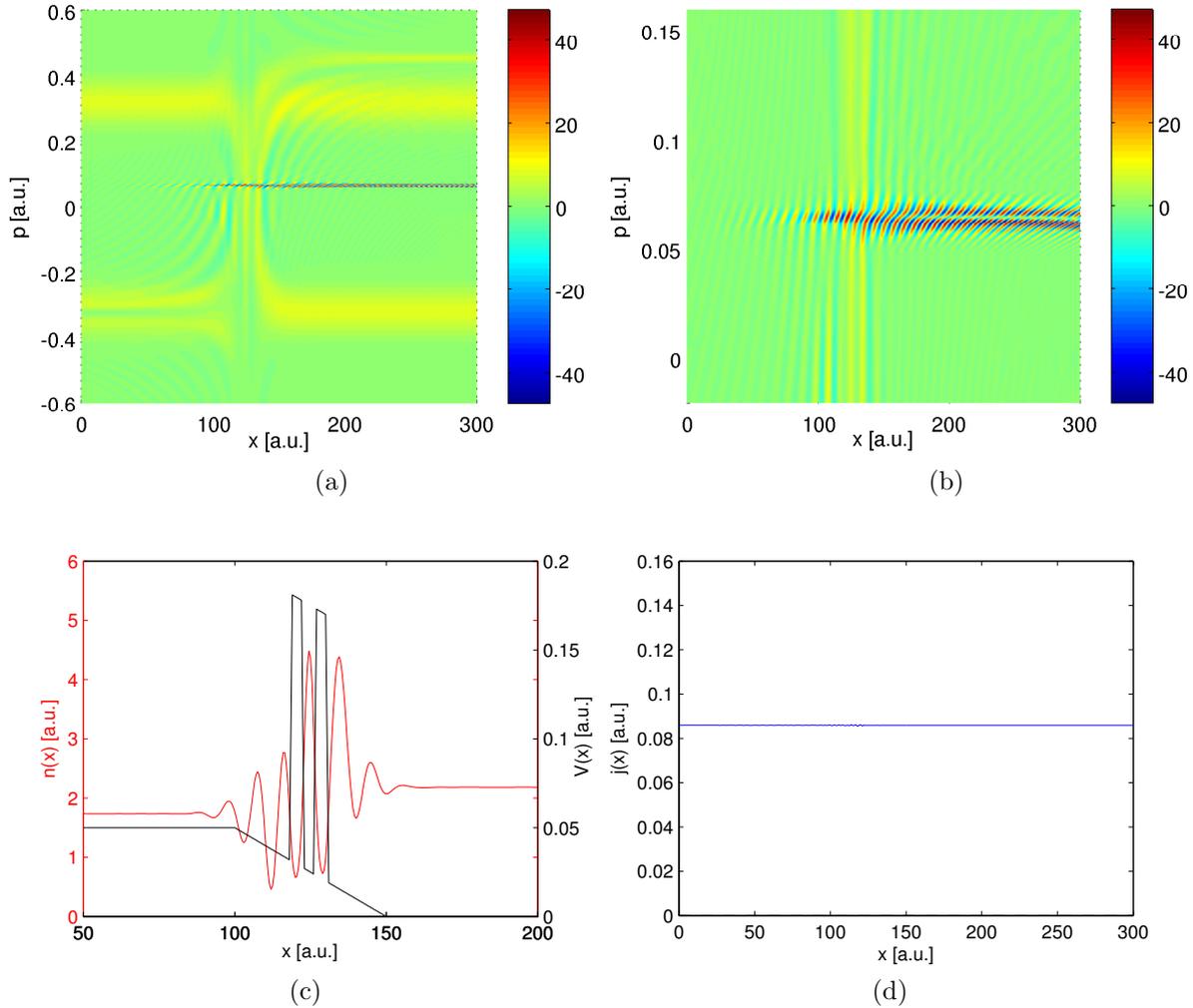


Figure 5.5: Illustrations of $f(p, x, t)$, $n(x)$ and $j(x)$ for the time $t = 20$ k, resulting from a simulation using the piecewise linear approximation as well as $N_x = 601$ and $N_p = 1596$ grid points. (a) and (b) display two different sections of the phase space for $f(p, x, t)$, (c) illustrates the particle density $n(x)$ together with the potential $V(x)$ and (d) shows a plot of $j(x)$.

5.2 Numerical study on the accuracy and convergence of the algorithm

As a particular application of the numerical method for the solution of the Wigner transport equation we consider in the following a resonant tunneling diode consisting of an $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}/\text{Al}_x\text{Ga}_{1-x}\text{As}$ heterostructure, as depicted schematically in Fig. 5.6 (a). The accuracy of the numerical method for the WTE is evaluated by comparing the steady state $I(V)$ curves obtained with the WTE calculations against a NEGF reference solution. In particular, the NEGF calculation with $\Delta x = a/64$ is chosen, for which we know from Sec. 4.2 to produce very accurate results.

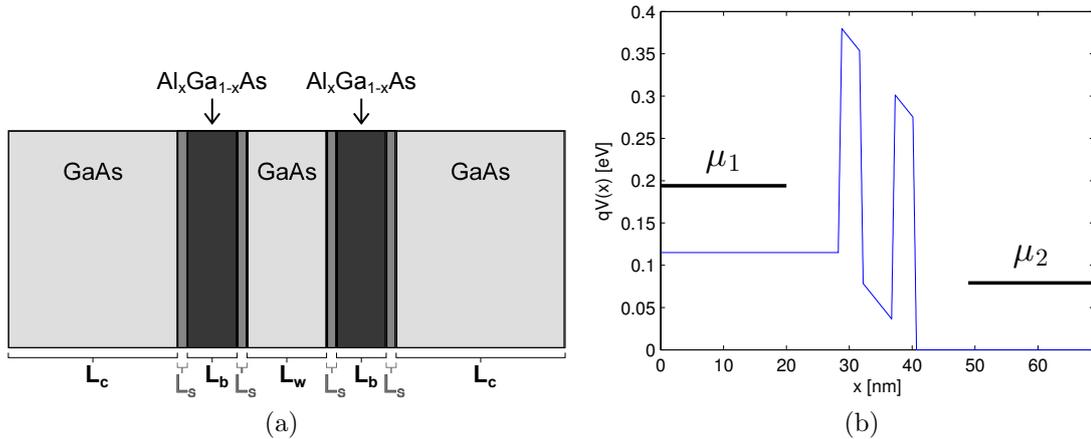


Figure 5.6: Illustration of the considered RTD consisting of a $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}/\text{Al}_x\text{Ga}_{1-x}\text{As}$ heterostructure in (a) and the corresponding potential $qV(x)$ in (b). The distances marked in (a) are the contact length L_c included in the x domain of the simulation, the barrier length $L_b = 5a$, the well width $L_w = 8a$ and the length of the slopes $L_s = 1a$, i.e. the length of the transition region at a $\text{AlGaAs}/\text{GaAs}$ interface, with the lattice constant $a = 0.565 \text{ nm}$ of GaAs . The plot of $qV(x)$ shown in (b) depicts the case of a rather short contact length of $L_c = 50a$ and a bias voltage of $V_{DS} = 0.115 \text{ V}$. In addition, the chemical potentials $\mu_1 = 0.19 \text{ eV}$ and $\mu_2 = 0.08 \text{ eV}$ of each contact are indicated.

The $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}/\text{Al}_x\text{Ga}_{1-x}\text{As}$ RTD is a standard problem to test numerical methods for quantum transport calculations in the field of device simulations and has been considered in various publications, as for instance in [4] [5] [6] [7]. We restrict ourselves to the simplest case of a homogenous doping profile, without aiming for a self-consistent solution by coupling the WTE to Poisson's equation, so that the heterostructure can be modeled by the static, effective potential $qV(x)$ depicted in Fig. 5.6 (b). The bias voltage V_{DS} is assumed to cause a linear potential drop in the region of the double barrier. Outside the heterostructure we assume a strictly constant potential to model the contact region. The length of the contacts included

in the simulated x domain for the Wigner function is labeled by L_c . In the case of the specific composition $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, the energy band offset has a magnitude of 0.27 eV [5]. Furthermore, expressed in terms of the lattice constant $a = 0.565$ nm of GaAs, we choose for the width of the barriers $L_b = 5a$, for the length of the well in between $L_w = 8a$ and include for the barriers a finite slope over a length of $L_s = 1a$, similar to the situation considered in [4] [6]. A spatial dependence of the effective mass is neglected and the value $m^* = 0.067 m_e$ for GaAs is used on the whole x domain [6].

The Fermi-Dirac distribution for the boundary conditions is characterized by a chemical potential μ_α for each contact, which in turn can be calculated from the knowledge of the donor density N_D and the temperature T by finding the root of the equation [4]

$$\frac{N_D}{N_C} = \mathcal{F}_{1/2}[\beta(\mu_\alpha - v^\alpha)] . \quad (5.1)$$

As stated already in Sec. (4.2), v^α labels the potential energy in each contact, $\beta = \frac{1}{k_B T}$, $\mathcal{F}_{1/2}$ is the Fermi-Dirac integral of order $\frac{1}{2}$ [4] [42] and N_C is given by [4]

$$N_C = 2 \left(\frac{m^*}{2\pi\hbar^2\beta} \right)^{\frac{3}{2}} . \quad (5.2)$$

For the simulations we choose a donor density of $N_D = 2 \times 10^{18} \text{ cm}^{-3}$ and a temperature of $T = 300$ K [6]. In Fig. 5.6 (b) the calculated chemical potentials $\mu_1 = 0.19$ eV and $\mu_2 = 0.08$ eV for the case $V_{DS} = 0.115$ V are indicated as well. From the knowledge of the chemical potentials for each contact, the cell averages of the ghost cells $F_{m,j}$ with $j = -2, -1, 0, N_x + 1, N_x + 2, N_x + 3$ are set according to Eqs. (3.255) and (3.256).

To calculate the drift matrix we make use of the general expression for the terms $I_{m+1/2,i+1/2,j}^F$ and $I_{m+1/2,i+/-j}^\sigma$, Eq. (3.215), together with Eqs. (3.196) and (3.197). Due to the simple piecewise linear potential shape as depicted in Fig. 5.6 (b), the number of potential nodes N_V is equal to eight. For all simulations presented in the following the piecewise linear approximation of $f(p, x, t)$ is used to calculate the drift matrix. For the time stepping a Strang splitting is applied in order to calculate in each sub-step the action of the drift term and that of the advection term separately. The time stepping for the drift term is performed by making use of Eq. (3.267), whereby the exponential of the drift matrix is calculated with the Matlab routine *expm*. In the other sub-step, for the advection term, the WENO5-SSP(3,3) scheme, as presented in Sec. (3.3), is applied with a Courant number of $\nu \approx 0.9$. Finally, the measurable quantities $n(x, t)$ and $j(x, t)$ are calculated via Eqs. (3.117) and (3.265), respectively. To obtain a value for the terminal current density $I(t)$, it is convenient to average $j(x, t)$ (because of possible oscillations) with respect to the position variable.

The grid for the x variable is chosen to be equidistant, with Δx expressed in terms of some fraction of the lattice constant a . The contact length L_c is varied in different simulations. In the case of the p grid it is necessary to make extensive use of the possibility to implement non-equidistant grids, in order to resolve all the oscillation

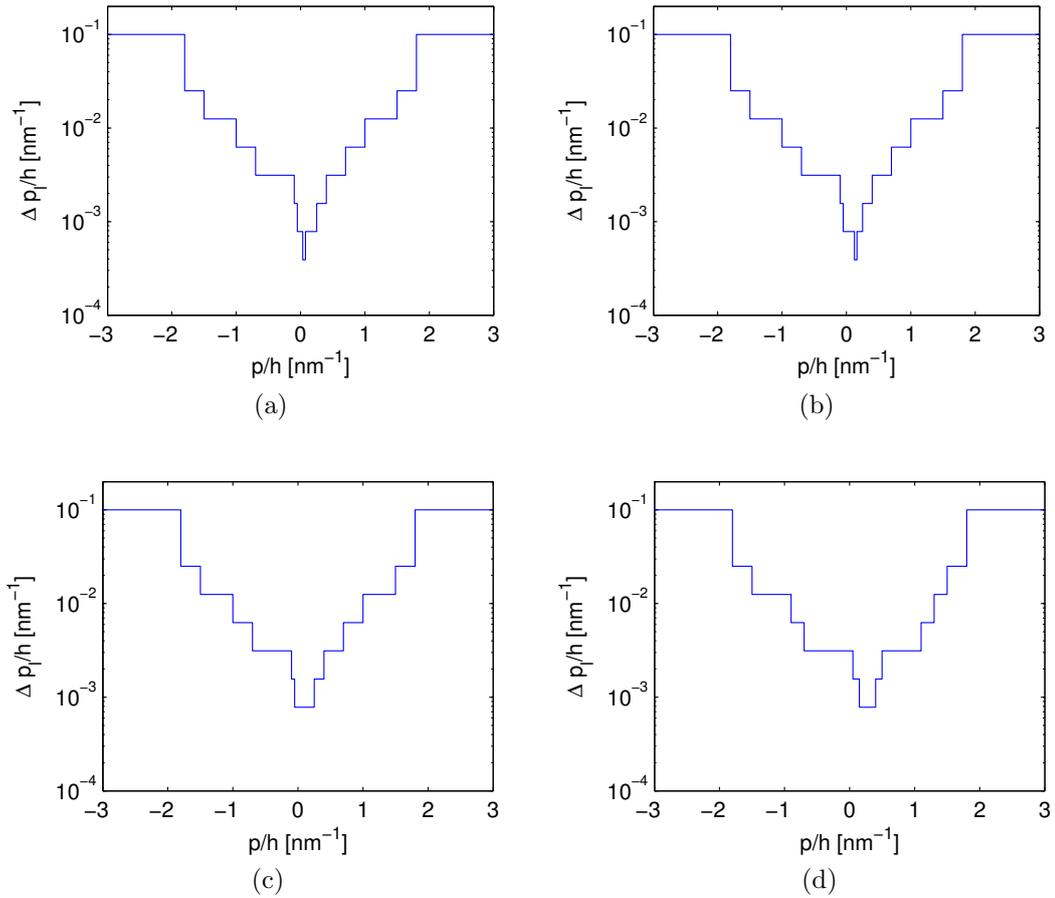


Figure 5.7: Illustration of the dependence of the grid spacing Δp_l on the particular value of p and the applied bias V_{DS} . The plots in (a) are for the case $V_{DS} = 0.05$ V, in (b) for $V_{DS} = 0.115$ V, in (c) for $V_{DS} = 0.185$ V and in (d) for $V_{DS} = 0.3$ V. The exact values can be obtained from Eq. (5.3) and Tab. 5.1 when setting $\Delta p_{max}/\hbar = 0.1$ nm $^{-1}$ and $p_{max}/\hbar = 3$ nm $^{-1}$.

patterns of $f(p, x, t)$ well enough and to thus achieve a smooth convergent behaviour. From the numerical point of view it is advisable to choose all grid spacings Δp_l as multiples of a minimal spacing Δp^0 . We specify a certain p_{max} which determines the considered p domain by $(-p_{max}, p_{max})$, as well as a maximum spacing Δp_{max} for the outermost region of the p domain. All the other Δp^i of the interior subdivisions of the p domain are then expressed as a fraction of Δp_{max} . In order not to introduce too large steps in the grid constants the spacings are divided by two in each refinement

step, except for the outermost interval:

$$\begin{aligned}
 \Delta p^{N_{\Delta p}-1} &= \Delta p_{max}, \\
 \Delta p^{N_{\Delta p}-2} &= \frac{\Delta p^{N_{\Delta p}-1}}{4}, \\
 \Delta p^i &= \frac{\Delta p^{i+1}}{2}, \quad i = 0, \dots, N_{\Delta p} - 3.
 \end{aligned} \tag{5.3}$$

In addition, an asymmetric and adaptive grid is used with different subdivisions depending on the applied bias voltage V_{DS} . The particular grid parameters are listed in Tab. 5.1 and an illustration of the chosen grid spacing can be found in Fig. 5.7.

Table 5.1: Subdivisions of the p grid with the corresponding grid spacings Δp^i depending on the considered bias voltage V_{DS} . The maximum value $p_{max}/\hbar > 1.8 \text{ nm}^{-1}$ is varied in different simulations. The particular spacings Δp^i are calculated by Eq. (5.3) for a specific value of Δp_{max} . For low voltages, $V_{DS} \leq 0.15 \text{ V}$, the grid is further refined in the region of the oscillations, for which the position is estimated by $p_{osc}^l = a^l + b^l V_{DS} + c^l V_{DS}^2$, with $a^l/\hbar = 0.006 \text{ nm}^{-1}$, $b^l/\hbar = 0.8 \text{ (Vnm)}^{-1}$ and $c^l/\hbar = 3.5 \text{ (V}^2\text{nm)}^{-1}$. An adaptive grid is used for larger voltages $V_{DS} \geq 0.2 \text{ V}$ as well, where the most oscillating region was estimated by $p_{osc}^h = a^h + b^h V_{DS}$ with $a^h/\hbar = 0.125 \text{ nm}^{-1}$ and $b^h/\hbar = 0.5 \text{ (Vnm)}^{-1}$. For the case that a value p_a or p_b does not coincide with a cell boundary $p_{m+1/2}$, the switching point p_a or p_b is rounded to the next possible cell boundary in favour of a higher resolution, i.e. the region with the smaller grid spacing is extended.

$p_a \leq p_{m+1/2} < p_b$						
	$V_{DS} \leq 0.15 \text{ V}$		$0.15 \text{ V} < V_{DS} < 0.2 \text{ V}$		$V_{DS} \geq 0.2 \text{ V}$	
Δp_l	$p_a/\hbar [\text{nm}^{-1}]$	$p_b/\hbar [\text{nm}^{-1}]$	$p_a/\hbar [\text{nm}^{-1}]$	$p_b/\hbar [\text{nm}^{-1}]$	$p_a/\hbar [\text{nm}^{-1}]$	$p_b/\hbar [\text{nm}^{-1}]$
Δp^7	$-p_{max}/\hbar$	-1.8				
Δp^6	-1.8	-1.5	$-p_{max}/\hbar$	-1.8	$-p_{max}/\hbar$	-1.8
Δp^5	-1.5	-1.0	-1.8	-1.5	-1.8	-1.5
Δp^4	-1.0	-0.7	-1.5	-1.0	-1.5	-0.9
Δp^3	-0.7	-0.1	-1.0	-0.7	-0.9	-0.7
Δp^2	-0.1	-0.05	-0.7	-0.1	-0.7	0.05
Δp^1	-0.05	$p_{osc}^l/\hbar - 0.02$	-0.1	-0.05	0.05	$p_{osc}^h/\hbar - 0.125$
Δp^0	$p_{osc}^l/\hbar - 0.02$	$p_{osc}^l/\hbar + 0.02$	-0.05	0.25	$p_{osc}^h/\hbar - 0.125$	$p_{osc}^h/\hbar + 0.125$
Δp^1	$p_{osc}^l/\hbar + 0.02$	0.25	0.25	0.4	$p_{osc}^h/\hbar + 0.125$	0.5
Δp^2	0.25	0.4	0.4	0.7	0.5	1.1
Δp^3	0.4	0.7	0.7	1.0	1.1	1.3
Δp^4	0.7	1.0	1.0	1.5	1.3	1.5
Δp^5	1.0	1.5	1.5	1.8	1.5	1.8
Δp^6	1.5	1.8	1.8	p_{max}/\hbar	1.8	p_{max}/\hbar
Δp^7	1.8	p_{max}/\hbar				

For the initial condition, $f(p, x, t=0)$, the Fermi-Dirac distribution in the contacts is extended into the region of constant $qV(x)$ inside the simulated x domain, labeled by L_c in Fig. 5.6 (a). In the region of the heterostructure and close to the barriers $f(p, x, t=0)$ is set to zero in order not to introduce too large momenta and long-lasting interference patterns. To test if a steady state is reached in the course of the time evolution, the terminal current is monitored and subsequent values are compared. Since an oscillating behaviour is ever-present, it was decided to calculate a new value of $I(t_n)$ after every ≈ 50 fs and furthermore, to average over ten subsequent values of $I(t_n)$ to arrive at one value for the convergence test. Therefore, only current values averaged over approximately 500 fs are compared. The averaging interval was found to be large enough to enable an accurate convergence test with an relative tolerance of $\varepsilon_{rel} = 10^{-3}$, but still small enough to detect the advent of strong oscillations in the case of a badly resolved solution $f(p, x, t)$. In addition, an initial time of 2000 fs is waited before the first convergence test to suppress the appearance of interference patterns from the initial distribution in the final solution $f(p, x, t)$.

For the first simulations and accuracy tests presented here, the spacing of the x grid is set to $\Delta x = a$, corresponding to the maximum possible spacing when choosing for the barriers a slope of length $L_s = 1 a$. For the p grid we choose $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$ and the other grid parameters, except for p_{max} , follow from Eq. (5.3) and Tab. 5.1. Examples for the different grid spacings depending on the bias voltage V_{DS} are illustrated in Fig. 5.7. Test simulations revealed that this combination of the x and p spacings is suited to properly resolve all of the oscillation patterns appearing in $f(p, x, t)$ and therefore, a relatively smooth convergence behaviour from the initial to the stationary solution is achieved. Grid parameters still left open to be specified, are the sizes of the x and p domain. A good starting point to choose the size of the p domain is to look at the boundary conditions and set the maximum p value in such a way that the major part of the Fermi-Dirac distributions of the contacts is covered and the values of $f(p, x, t)$ at the endpoints are small compared to the values of $f(p, x, t)$ in the interior region. From this one would choose a value of approximately $p_{max}/\hbar \approx 1 \text{ nm}^{-1}$ but test simulations revealed that it is advisable to use at least $p_{max}/\hbar = 2 \text{ nm}^{-1}$. Otherwise $f(\pm p_{max}, x, t)$ takes on quite large values, indicating that the p domain has been chosen too small and therefore, that a large error in the calculation of the action of the pseudo-differential operator on $f(p, x, t)$ is encountered. To analyze the impact of the size of the simulated x domain we consider simulations with different values of L_c . The inflow/outflow boundary conditions for the WTE do not stem from a quantum mechanical derivation, as in the case of the self-energies in the NEGF technique, but are in principle adapted from classical transport theory. The modeling of the inflow and outflow at the boundaries as presented in Sec. (3.3) implicitly assumes that the momentum distribution of $f(p, x, t)$ at the boundaries is maintained in the contacts as well. This assumption is only appropriate if, in the vicinity of the boundaries, the gradients of $f(p, x, t)$ with respect to x are negligible. This is only the case if the action of the pseudo-differential operator close to the boundaries is not significant. To find a proper position for the boundaries of the simulated x domain, results for different values of L_c are compared in the following.

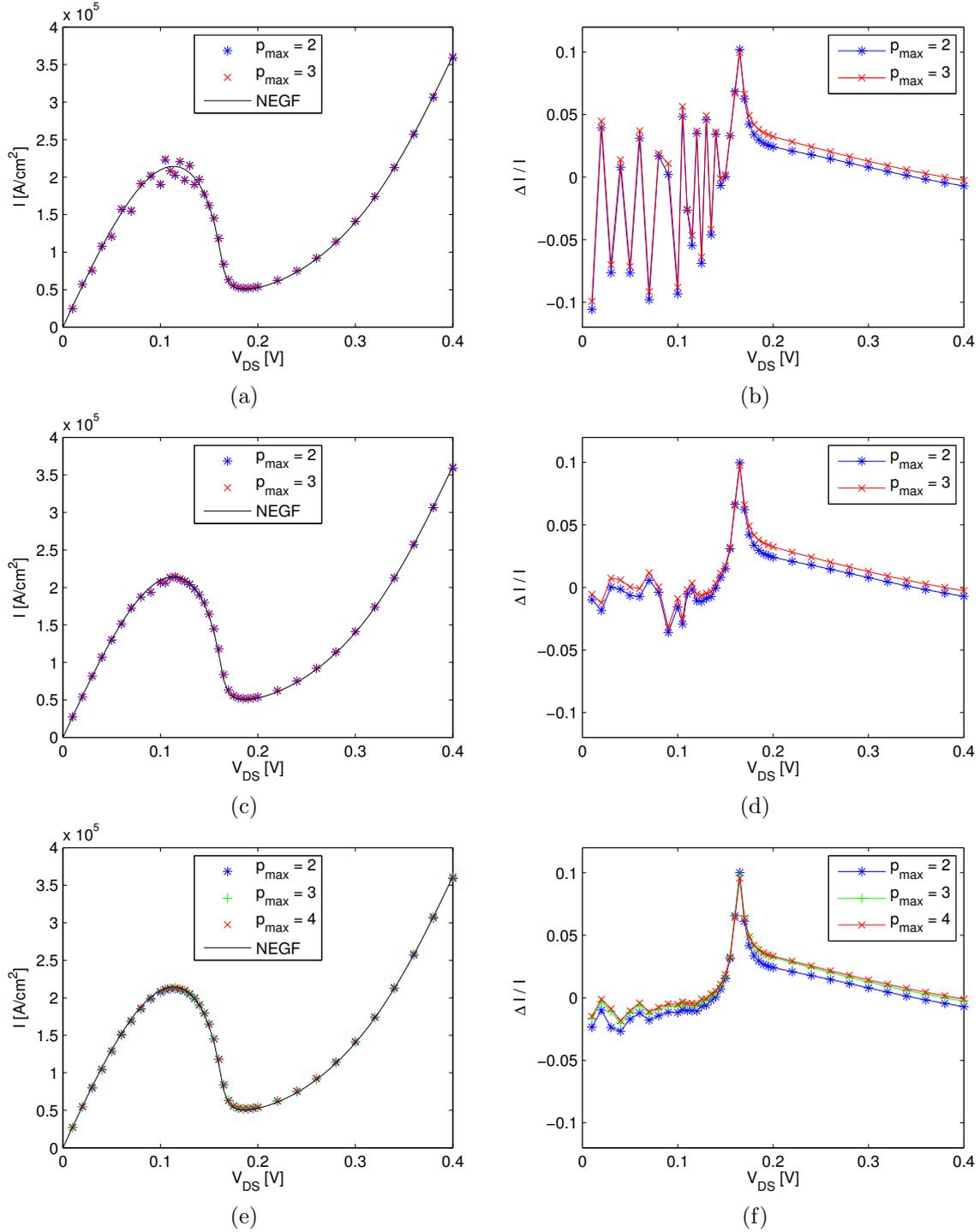


Figure 5.8: Comparison of the simulated $I(V)$ curves for different values of the contact length L_c and different extensions of the p domain, specified by $p_{max}/\hbar = 2 \text{ nm}^{-1}$, $p_{max}/\hbar = 3 \text{ nm}^{-1}$ and $p_{max}/\hbar = 4 \text{ nm}^{-1}$. For all of the simulations the grid spacings are chosen to $\Delta x = a$ and $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$, i.e. $N_x = 223, 423, 623$ and $N_p \approx 1000$. The results depicted in (a),(b) are for $L_c = 100 a$, those in (c),(d) are for $L_c = 200 a$ and those in (e),(f) are for $L_c = 300 a$. The relative differences of the currents with respect to the NEGF reference currents are presented in the plots on the right-hand side.

In Fig. 5.8 one can find the calculated $I(V)$ curves for simulations with different values of the contact length L_c , as well as different values of p_{max} . Clearly, the difference of the results for the cases $p_{max}/\hbar = 2 \text{ nm}^{-1}$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$ are comparatively small but the simulated contact length has great impact on the accuracy of the results. One has to bear in mind that the considered heterostructure exhibits in total a length of only $22a$, compare Fig. 5.6. Intuitively one could think that already a distance of $L_c = 100a$ between the double barrier structure and a boundary should suffice to cause only small errors. But, one can see from the results shown in Fig. 5.8 that the influence of the boundaries on the overall device behaviour is quite prominent for the case $L_c = 100a$ and still present for $L_c = 200a$. In the case of $L_c = 300a$ the errors from the boundaries have nearly disappeared and one can see a fairly good agreement between the Wigner and the NEGF calculations. For this case the three different values $p_{max}/\hbar = 2 \text{ nm}^{-1}$, $p_{max}/\hbar = 3 \text{ nm}^{-1}$ and $p_{max}/\hbar = 4 \text{ nm}^{-1}$ are considered. It is evident that the results for $p_{max}/\hbar = 3 \text{ nm}^{-1}$ and $p_{max}/\hbar = 4 \text{ nm}^{-1}$ show only minor differences, so that the extra cost in computation time for the larger value of p_{max} is not justified. Therefore, a size of the (p, x) domain determined by the values $L_c = 300a$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$ seems to be a good choice for the considered physical problem.

When examining the errors of the single simulations in Fig. 5.8 in detail, it is evident that the calculated current values in the upper region of the $I(V)$ curve, for values of approximately $V_{DS} \gtrsim 0.16 \text{ V}$ do not exhibit a significant dependence on the particular value of L_c . In contrast, for lower bias voltages of $V_{DS} \lesssim 0.16 \text{ V}$ one can see the advent of oscillations in the $I(V)$ curve when L_c is chosen too small. To understand this behaviour, Fig. 5.9 depicts two phase space plots of $f(p, x, t)$ for the bias voltages $V_{DS} = 0.05 \text{ V}$ and $V_{DS} = 0.3 \text{ V}$ as examples. In the case of $V_{DS} = 0.05 \text{ V}$ resonant tunneling processes are possible and the Wigner function $f(p, x, t)$ thus exhibits strong oscillation patterns. On the one hand, a sharp, short-scaled stripe of oscillations is present at low values of p and on the other hand, longer-scaled and more wave-like oscillations are formed over large regions of phase space. The oscillations extend to positions far away from the double barrier structure and the action of the pseudo-differential operator is thus non-negligible even in large distance to the barriers. In comparison, $f(p, x, t)$ exhibits for the case $V_{DS} = 0.3 \text{ V}$ only minor oscillation patterns. Therefore, applying the boundaries at short distances L_c and thus extending $f(p, x, t)$ with constant shape with respect to p into the contacts is in this case clearly a much better approximation than in the case $V_{DS} = 0.05 \text{ V}$. Interesting to note is that the two bias voltages result in very similar current values, as can be seen from Fig. 5.8. For $V_{DS} = 0.05 \text{ V}$ a rather large part of the electron distribution can pass through the double barrier by a resonant tunneling process but is accelerated only slightly. In contrast, for $V_{DS} = 0.3 \text{ V}$ comparatively few electrons tunnel through the barriers but are therefore accelerated much more, as can be seen from the beam at $p/\hbar \approx 0.8 \text{ nm}^{-1}$ on the right-hand side of the barriers in Fig. 5.9.

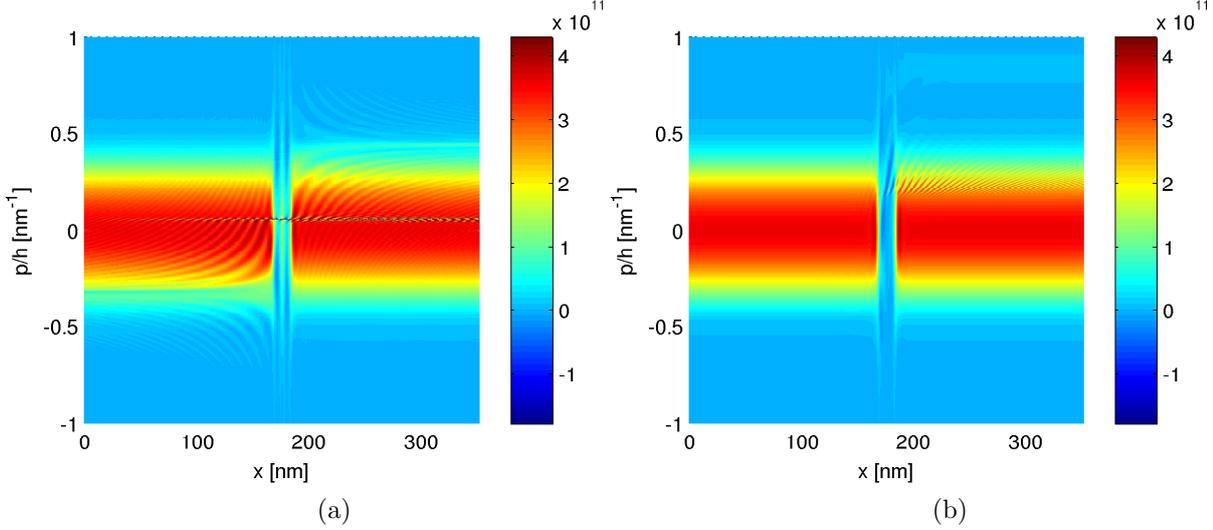


Figure 5.9: Illustrations of the steady state solutions of $f(p, x, t)$ obtained with simulations with $L_c = 300 a$, $p_{max}/\hbar = 3 \text{ nm}^{-1}$, $\Delta x = a$ and $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$ for two different bias voltages. Depicted in (a) is the Wigner function for $V_{DS} = 0.05 \text{ V}$ and in (b) the one for $V_{DS} = 0.3 \text{ V}$, where in each plot only a part of the calculated phase space up to values $p/\hbar = \pm 1 \text{ nm}^{-1}$ is shown. The number of grid points for the two cases are $N_x = 623$, $N_p = 1076$ and $N_x = 623$, $N_p = 1056$ for $V_{DS} = 0.05 \text{ V}$ and $V_{DS} = 0.3 \text{ V}$, respectively.

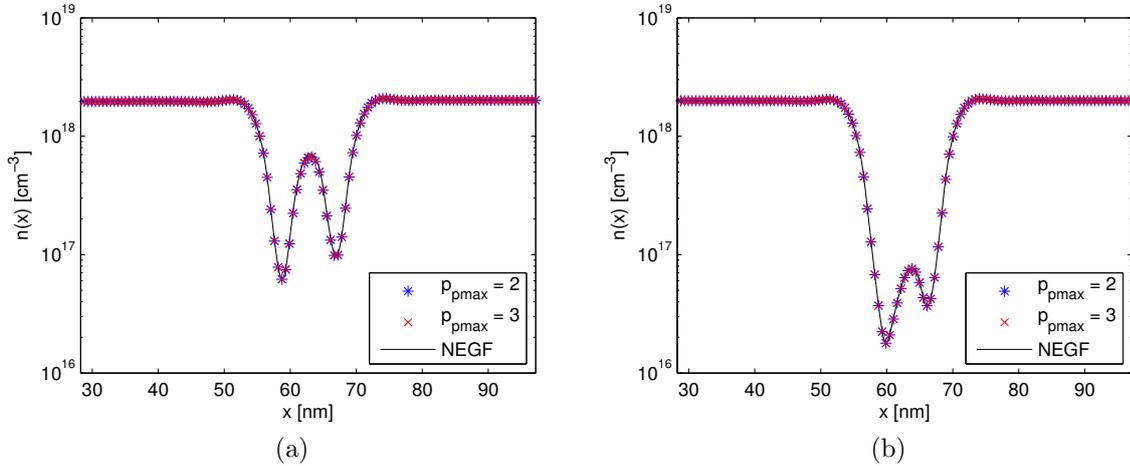


Figure 5.10: Calculated particle density obtained for $L_c = 100 a$ and compared with the NEGF reference particle density. The plot shown in (a) depicts the case $V_{DS} = 0.115 \text{ V}$ and the one in (b) $V_{DS} = 0.185 \text{ V}$. The other parameters are $\Delta x = a$, $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$ and thus $N_x = 223$ and $N_p \approx 1000$. The individual simulations correspond to $p_{max}/\hbar = 2 \text{ nm}^{-1}$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$.

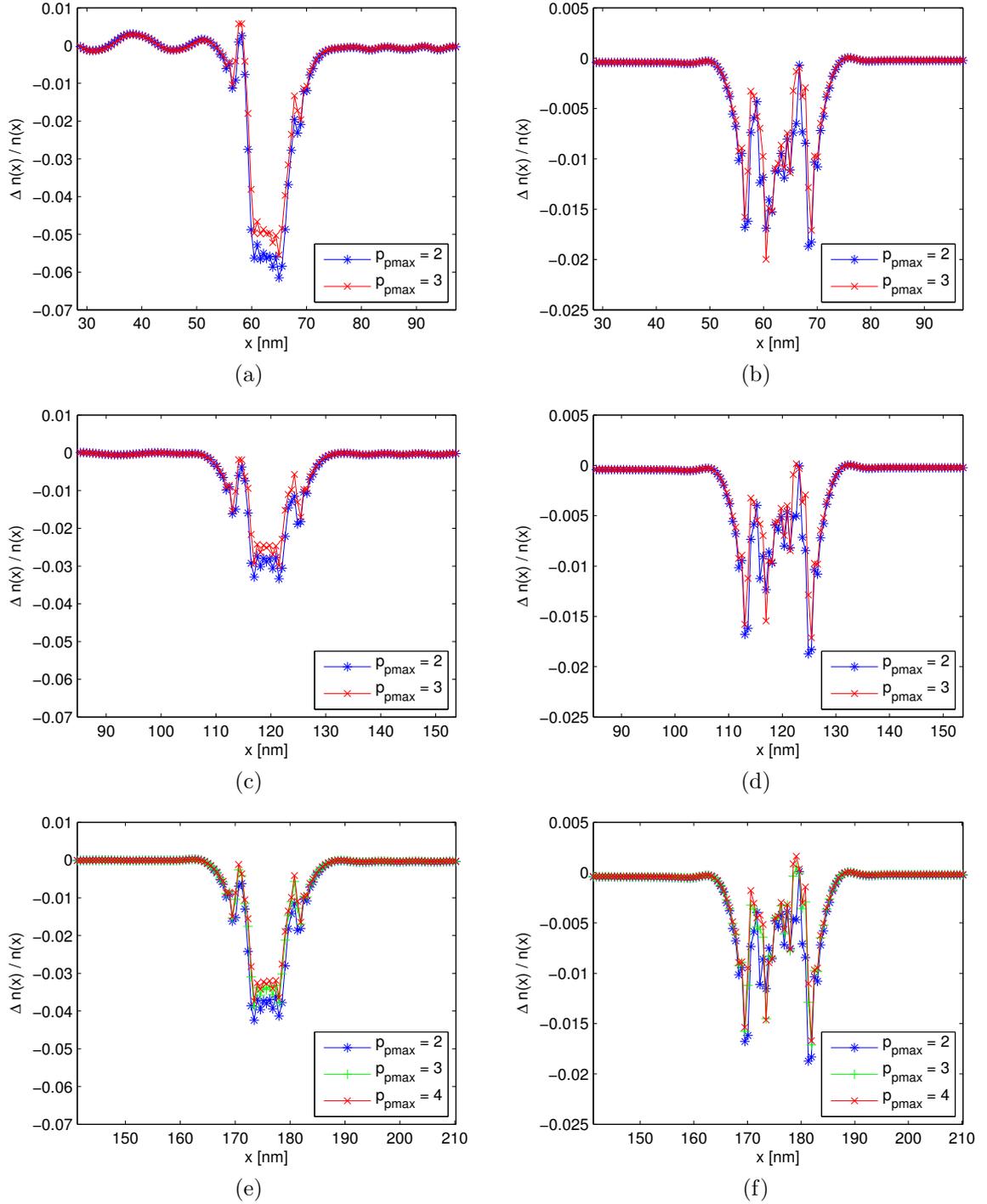


Figure 5.11: Relative differences of the calculated particle densities compared to the NEGF solution. The plots on the left-hand side show the results for $V_{DS} = 0.115 \text{ V}$ and the ones on the right-hand side the results for $V_{DS} = 0.185 \text{ V}$. (a),(b) correspond to $L_c = 100 a$, (c),(d) to $L_c = 200 a$ and (e),(f) to $L_c = 300 a$. The other parameters are $\Delta x = a$ and $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$. The individual $n(x)$ calculations correspond to $p_{max}/\hbar = 2 \text{ nm}^{-1}$, $p_{max}/\hbar = 3 \text{ nm}^{-1}$ and $p_{max}/\hbar = 4 \text{ nm}^{-1}$.

Besides the calculated current values $I(V)$ the particle density $n(x)$ is a quantity of interest. For the purpose of estimating the accuracy of the calculated $n(x)$ values, the results from the WTE calculations with $L_c = 100a$, $L_c = 200a$ and $L_c = 300a$ are compared with the NEGF solution in Figs. 5.10 and 5.11 for the two bias voltages $V_{DS} = 0.115\text{ V}$ and $V_{DS} = 0.185\text{ V}$ as examples. One can obviously see from Fig. 5.10 a very good agreement between the NEGF and the WTE calculations already for the case $L_c = 100a$. Fig. 5.11 thus depicts solely the relative differences between the results obtained with the WTE and those by means of the NEGF approach. Again, the results for the higher bias voltage of $V_{DS} = 0.185\text{ V}$ are fairly insensitive to the particular value of L_c and the results for the peak voltage $V_{DS} = 0.115\text{ V}$ show an increase in the degree of oscillations in $n(x)$ for smaller values of L_c . The peak voltage $V_{DS} = 0.115\text{ V}$ was chosen since in this case resonant tunneling processes are most prominent, causing far reaching oscillation patterns in $f(p, x, t)$.

The previous results indicate that a size of the simulated (p, x) domain characterized by the values $L_c = 300a$ and $p_{max}/\hbar = 3\text{ nm}^{-1}$ suffices to enable accurate simulations for the considered physical situation. We now focus on the question, how much the accuracy of the results can be improved by reducing the grid spacings. As a first step we reduce the x spacing to $\Delta x = \frac{a}{2}$ and perform simulations for three different values of Δp_{max} . The results can be found in Fig. 5.12 and for all three quantities $I(V)$, $n(x)$ for $V_{DS} = 0.115\text{ V}$ and $n(x)$ for $V_{DS} = 0.185\text{ V}$ an improvement of the results can be seen, at least if Δp_{max} is small enough. For the largest value $\Delta p_{max}/\hbar = 0.1\text{ nm}^{-1}$ an oscillating behaviour of the calculated quantities can be observed. One can thus suppose that a value of $\Delta x = a$ does not suffice to resolve all physical oscillations of $f(p, x, t)$ and therefore, when reducing the spacing to $\Delta x = \frac{a}{2}$ shorter-scaled oscillations are possible which ultimately also affect the requirements on the p grid. It is apparent to see from the calculated $I(V)$ curve that the problems occur especially for lower voltages of $V_{DS} \lesssim 0.1\text{ V}$ where resonant tunneling processes are prominent. The $I(V)$ curves as well as $n(x)$ for $V_{DS} = 0.185\text{ V}$ calculated with the parameters $\Delta p_{max}/\hbar = 0.075\text{ nm}^{-1}$ and $\Delta p_{max}/\hbar = 0.05\text{ nm}^{-1}$ show only small differences, solely $n(x)$ for the case $V_{DS} = 0.115\text{ V}$ requires the finest resolution of $\Delta p_{max}/\hbar = 0.05\text{ nm}^{-1}$ to suppress the creation of spurious oscillations. Still, a value of $\Delta p_{max}/\hbar = 0.075\text{ nm}^{-1}$ enables us to produce quite accurate results and is therefore used for another convergence study, comparing different x spacings.

In Fig. 5.13 the results for three different spacings $\Delta x = a$, $\Delta x = \frac{a}{2}$ and $\Delta x = \frac{a}{4}$ are compared. Only the region of negative differential resistance (NDR) of $I(V)$ is plotted since the largest error in the previously calculated curves occurred in this region and in addition, especially the NDR in the $I(V)$ curve is of interest for device applications. One can see a monotone decrease of the relative errors with decreasing values of Δx and a very good agreement between the NEGF reference and the WTE calculations for the case $\Delta x = \frac{a}{4}$, with a relative error $|\Delta I/I|$ below one percent. This results clearly demonstrate the ability of the developed algorithm to produce very accurate results and furthermore, the convergence of the results when refining the grid parameters. To achieve even more accurate results one would need to increase the values of p_{max} and L_c as well. From Fig. 5.13 (b) it is obvious to see an oscillating

behaviour of $\Delta I/I$ in the region of the peak voltage $V_{DS} = 0.115$ V, indicating again the influence of the domain boundaries and thus a too small value for L_c (comp. Fig. 5.8). To quantify the results, the 2-norm error and the order of convergence are calculated by using the analogous expressions to Eqs. (3.221) and (3.223) and the values for $\|\Delta I\|_2$ and $\|\Delta n(x)\|_2$ are listed in Tabs. 5.2 and 5.3, respectively. In the case of $\|\Delta I\|_2$, the order of convergence was calculated for two different bias voltage intervals, with and without including the region of the peak voltage. When examining the results listed in Tabs. 5.2 and 5.3, one can estimate the order of convergence of the algorithm roughly by $\mathcal{O}_{\Delta x} \approx 2$, when keeping in mind that the limited values of p_{max} and L_c are sources of error as well and in addition, that especially the region of the peak voltage is affected by too small values of L_c .

Table 5.2: Estimation of the order of convergence from the simulated $I(V)$ curves as shown in Fig. 5.13, for the region $0.11 \text{ V} \leq V_{DS} \leq 0.19 \text{ V}$ and for $0.14 \text{ V} \leq V_{DS} \leq 0.19 \text{ V}$. Compared is the 2-norm error as defined in Eq. (3.221) and the order of convergence $\mathcal{O}_{\Delta x}$ is determined by the logarithm to the base of 2 of the quotient of two subsequent 2-norm errors.

	$0.11 \text{ V} \leq V_{DS} \leq 0.19 \text{ V}$		$0.14 \text{ V} \leq V_{DS} \leq 0.19 \text{ V}$	
Δx	$\ \Delta I\ _2$	$\mathcal{O}_{\Delta x}$	$\ \Delta I\ _2$	$\mathcal{O}_{\Delta x}$
a	940	2.06	925	2.16
$a/2$	226	0.95	208	1.51
$a/4$	117		73	

Table 5.3: Estimation of the order of convergence from the calculation of the particle densities $n(x)$ from the simulations shown in Fig. 5.13. The 2-norm error and the order of convergence $\mathcal{O}_{\Delta x}$ are determined in the same manner as described in Tab. 5.2.

	$V_{DS} = 0.115 \text{ V}$		$V_{DS} = 0.185 \text{ V}$	
Δx	$\ \Delta n(x)\ _2$	$\mathcal{O}_{\Delta x}$	$\ \Delta n(x)\ _2$	$\mathcal{O}_{\Delta x}$
a	4.82E16	1.57	2.22E16	2.01
$a/2$	1.62E16	0.82	0.55E16	1.51
$a/4$	0.92E16		0.19E16	

Figure 5.14 depicts phase space plots of $f(p, x, t)$ for the two cases of the peak voltage $V_{DS} = 0.115$ V and the valley voltage $V_{DS} = 0.185$ V. Note that the colour mapping for the plots has been chosen in such a way that the long-scaled oscillation patterns in $f(p, x, t)$ are clear to see, with the disadvantage that the magnitude of the stripe of oscillations (dark red and dark blue, at $p/\hbar \approx 0.15 \text{ nm}^{-1}$) in a (p, x) -plot cannot be seen. Therefore, in Fig. 5.14 (c) a side view, a $(f(p, x, t), x)$ -plot, is shown. Apparent for the resonant case is the creation of strong oscillation patterns throughout the whole phase space and also a sharp stripe of oscillations at $p/\hbar \approx 0.15 \text{ nm}^{-1}$. Figure 5.14 (b) focuses on the shape of $f(p, x, t)$ in the central region, close to the

barriers. The detailed dynamics are involved but one can see in the reflected part of the distribution on the left-hand side that momenta of approximately $p/\hbar \approx -0.2 \text{ nm}^{-1}$ are extenuated, corresponding to the part for which a resonant tunneling process is accessible. On the right-hand side, one can see the outgoing, accelerated beam at $p/\hbar \approx 0.5 \text{ nm}^{-1}$. The resonant tunneling process itself is represented by the stripe of oscillations at $p/\hbar \approx 0.15 \text{ nm}^{-1}$, at least this is what we suppose after analyzing various simulations. The occupation of the well state can be seen from the increased, positive values of $f(p, x, t)$ at $x \approx 175 \text{ nm}$. The enhanced negative region at $x \approx 170 \text{ nm}$ and $-0.25 \text{ nm}^{-1} \lesssim p/\hbar \lesssim 0 \text{ nm}^{-1}$ could be interpreted as a manifestation of the destructive interference of the backscattered part of the beam. In the case of the valley voltage $V_{DS} = 0.185 \text{ V}$, the situation is very different and $f(p, x, t)$ is a rather smooth function with a much less pronounced oscillation pattern in the region $0.1 \text{ nm}^{-1} \lesssim p/\hbar \lesssim 0.4 \text{ nm}^{-1}$. Also in this case tunneling processes take place and the outgoing beam is vaguely perceptible at $p/\hbar \gtrsim 0.5 \text{ nm}^{-1}$ on the right-hand side of the barriers.

To conclude, Fig. 5.15 (a) depicts the calculated particle densities $n(x)$ for a set of bias voltages. It can be clearly seen that a high particle density between the barriers builds up only in the case of the lowest bias voltage, where resonant tunneling and thus the occupation of the quasi-bound well state is possible. The particle densities for $V_{DS} = 0.3 \text{ V}$ and $V_{DS} = 0.4 \text{ V}$ exhibit a pronounced minimum in the center of the barriers. From Fig. 5.15 (b) it is obvious to see that the conduction band edge of the left-sided contact exceeds the second barrier in the case of $V_{DS} = 0.4 \text{ V}$, such that tunneling occurs solely at the first barrier and the second one causes solely interference effects. The particular shape of $n(x)$ in the central region for the case $V_{DS} = 0.4 \text{ V}$ is thus a result of interferences.

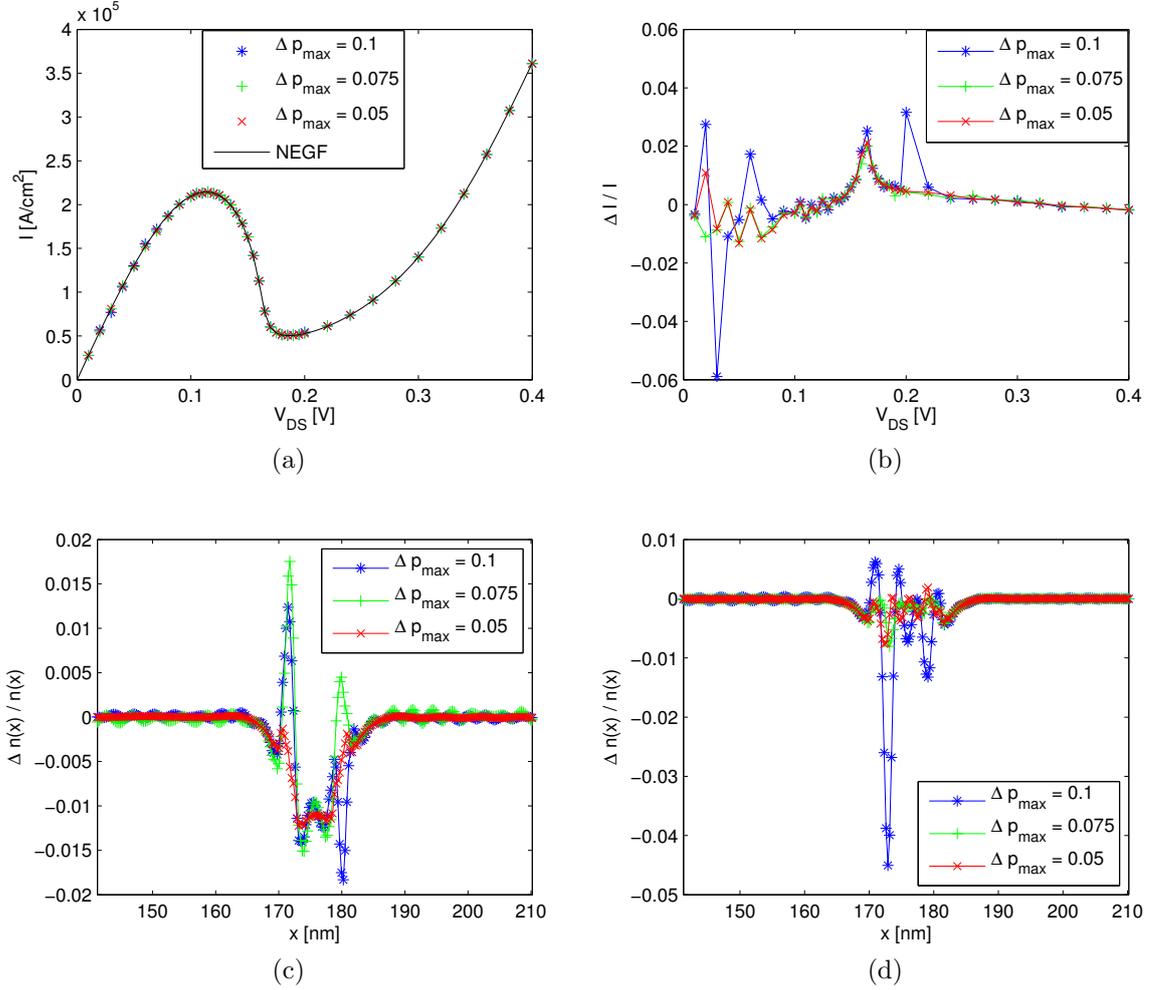


Figure 5.12: Plot (a) depicts the obtained $I(V)$ curve, (b) the relative difference of the $I(V)$ values to the NEGF solution and (c) and (d) the relative difference of the particle density $n(x)$ to the NEGF solution for $V_{DS} = 0.115$ V and $V_{DS} = 0.185$ V, respectively. The results are obtained by using a reduced x spacing of $\Delta x = \frac{a}{2}$ and three different values of Δp_{max} . The size of the (p, x) domain is the same in all three simulations and determined by the values $L_c = 300 a$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$. The number of grid points is $N_x = 1245$ and $N_p \approx 1000, 1400, 2000$.

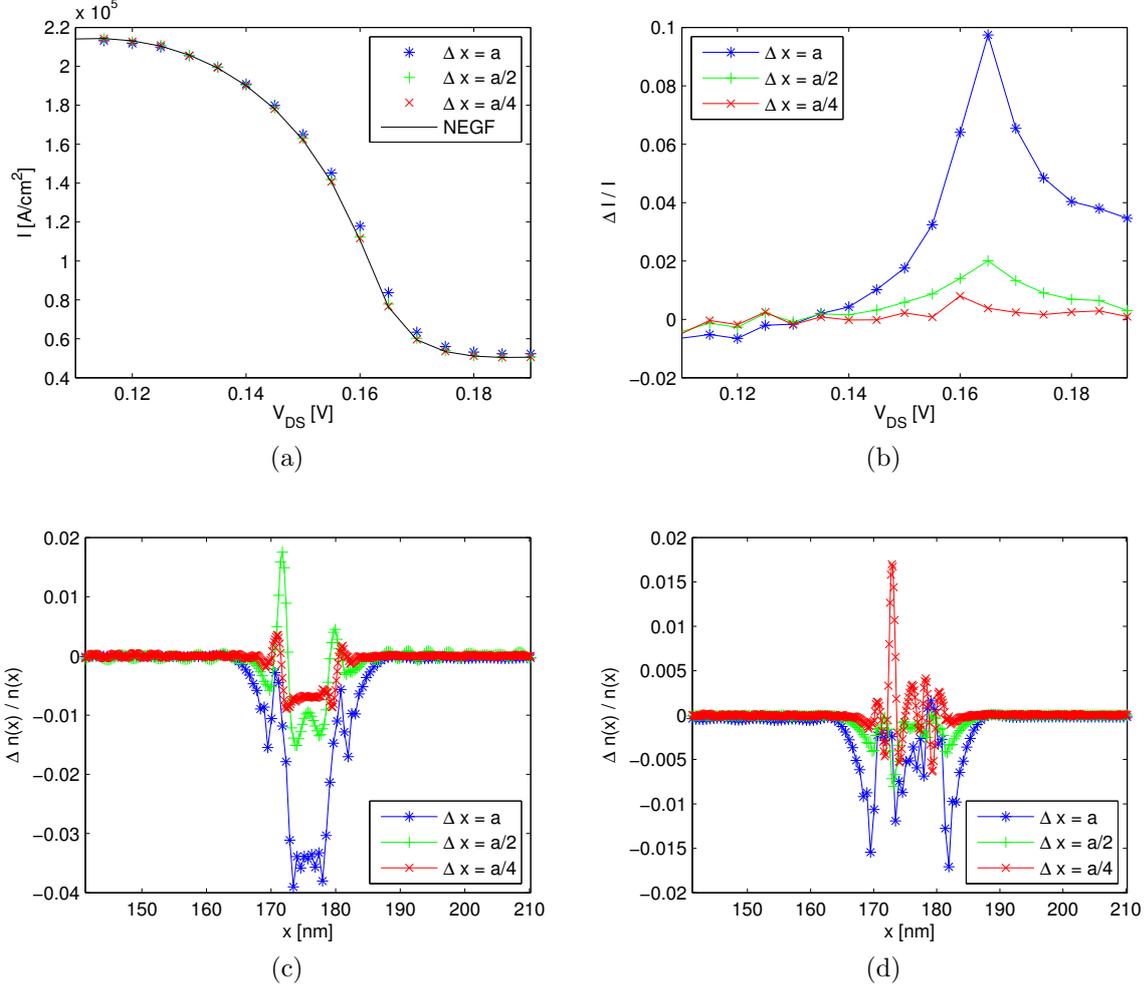


Figure 5.13: Plot (a) depicts the obtained $I(V)$ curve for the region of negative differential resistance (NDR), (b) the relative difference of the $I(V)$ values to the NEGF solution and (c) and (d) the particle density for $V_{DS} = 0.115$ V and $V_{DS} = 0.185$ V, respectively. The results are obtained by using $\Delta p_{max}/\hbar = 0.075$ nm⁻¹ and three different spacings $\Delta x = a, \frac{a}{2}, \frac{a}{4}$. The size of the (p, x) domain is the same in all three simulations and determined by $L_c = 300 a$ and $p_{max}/\hbar = 3$ nm⁻¹. The number of grid points is $N_x = 623, 1245, 2489$ and $N_p \approx 1400$.

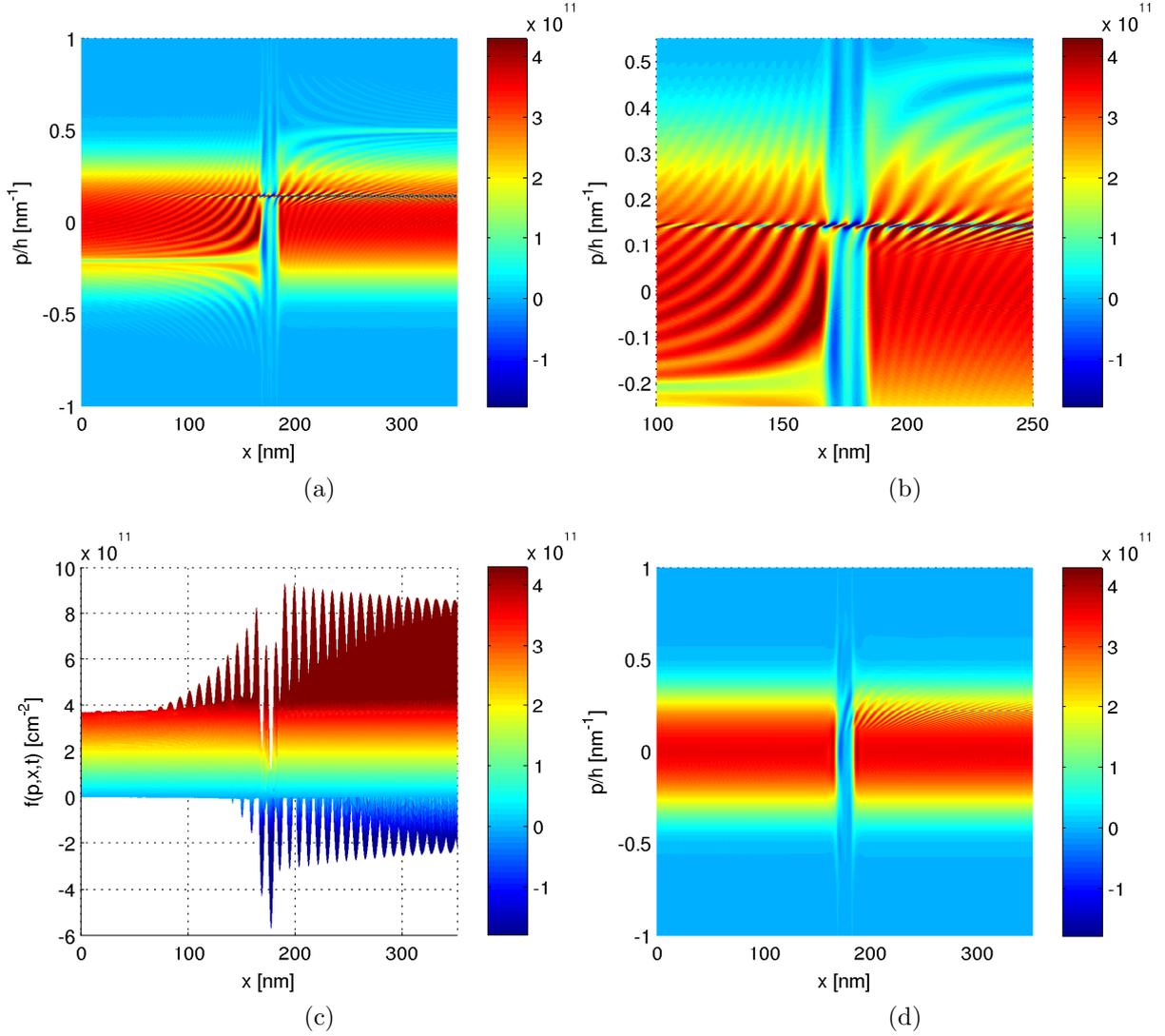


Figure 5.14: Illustrations of the steady state solutions of $f(p, x, t)$ obtained by using the parameters $\Delta x = \frac{a}{2}$, $\Delta p_{max}/\hbar = 0.05 \text{ nm}^{-1}$, $L_c = 300 a$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$ for two different bias voltages. Depicted in (a),(b),(c) are the solutions for $V_{DS} = 0.115 \text{ V}$ and in (d) the one for $V_{DS} = 0.185 \text{ V}$. The number of grid points for the two cases are $N_x = 1245$, $N_p = 2151$ and $N_x = 1245$, $N_p = 2048$ for $V_{DS} = 0.115 \text{ V}$ and $V_{DS} = 0.185 \text{ V}$, respectively.

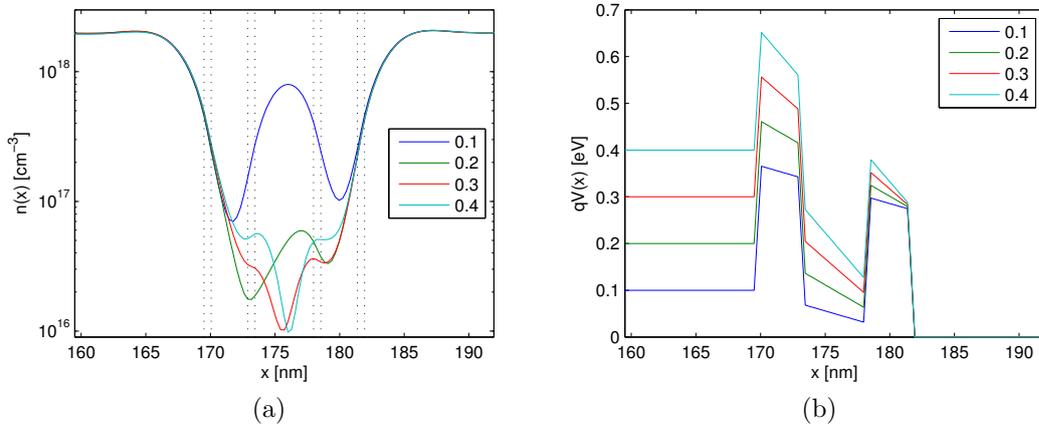


Figure 5.15: Logarithmic plot of the particle density $n(x)$ in the region of the double barrier in (a) and the corresponding shape of $qV(x)$ in (b), each for bias voltages of $V_{DS} = 0.1$ V, 0.2 V, 0.3 V, 0.4 V. In (a) the double barrier structure is indicated by dotted lines. The parameters of the simulations are: $\Delta x = \frac{a}{2}$, $\Delta p_{max}/\hbar = 0.05 \text{ nm}^{-1}$, $L_c = 300 a$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$.

5.3 Transient response simulation

The simulations presented in the last section revealed the ability of the developed WTE algorithm to produce very accurate results. The comparison with steady states obtained with the NEGF technique showed only minor differences when choosing the (p, x) domain large enough. Now, we present as an application the simulation of a time evolution. The NEGF technique has proven to be able to simulate accurately steady states, see e.g. [15] [7], but fully time-dependent calculations in the field of device simulations seem to be out of reach, at the moment at least. The great power of the WTE is the ability to simulate time-dependent problems since it is inherently an integro-differential equation describing the temporal evolution of $f(p, x, t)$.

A simple, time-dependent problem is to simulate the large-signal transient response of a RTD [4]. For this we consider the same heterostructure as in the previous section 5.2. To simulate a transient response, we start with the steady state obtained for the peak voltage, $V_{DS} = 0.115$ V, switch the bias voltage at $t = 0$ fs to the valley voltage $V_{DS} = 0.185$ V and simulate the time evolution. The same is done the other way around as well, with the steady state for $V_{DS} = 0.185$ V as initial condition and switching at $t = 0$ fs to $V_{DS} = 0.115$ V. In Fig. 5.16 the time evolution of the terminal current $I(t)$ is plotted for both cases. For the case of switching from the peak to the valley voltage, one can see an initial rise of the current before it starts to decrease towards the steady state value. For the case of switching from the valley to the peak voltage an analogous behaviour is observed, namely that the current decreases at first before starting to rise, only that it is less pronounced in this case. In addition, $I(t)$

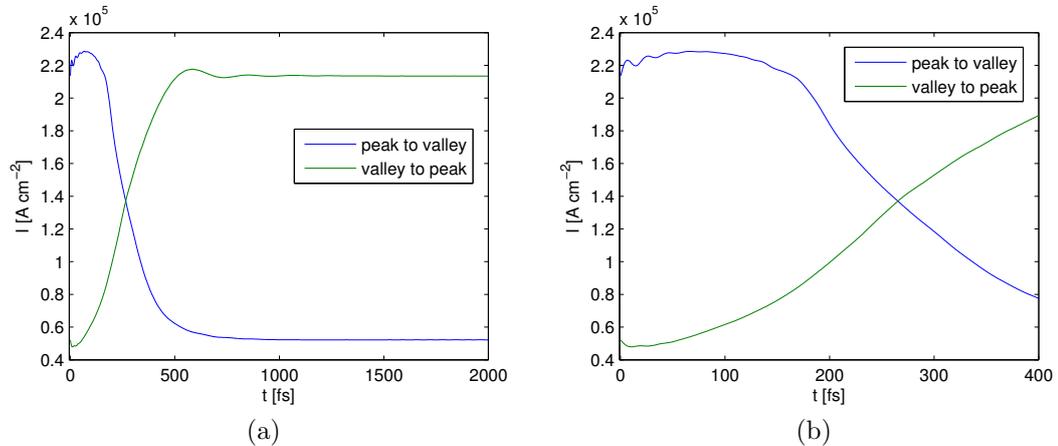


Figure 5.16: Time evolution of the terminal current density $I(t)$ when switching from the peak voltage $V_{DS} = 0.115$ V to valley voltage $V_{DS} = 0.185$ V and vice versa (a). Plot (b) displays the temporal evolution during the first 400 fs in higher resolution. The switching is done in each of the two cases at $t = 0$ fs. The parameters for the simulations are $\Delta x = a$, $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$, $L_c = 300 a$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$, i.e. $N_x = 623$ and $N_p = 1076$. In both cases the same p grid is used, namely the one optimized for the peak voltage.

exhibits oscillations when converging to the steady state current value of the peak voltage.

This particular behaviour can be understood when looking at the time evolution of the current density $j(x, t)$, depicted in Fig. 5.17. The plot in (a) shows the time evolution for the case of switching from peak to valley and one can see an increased current density starting from the region of the double barrier and propagating to the contact on the right-hand side. This represents the part of the electron distribution which has occupied the well state at $t = 0$ fs and is now accelerated to higher momenta by the increased bias voltage. The acceleration of this part of the Wigner distribution causes the initial net rise of the current $I(t)$. As time proceeds the well state is emptied and $j(x, t)$, on the right-hand side of the barriers, slowly decreases to the stationary value. The change of $j(x, t)$ on the left-hand side is much quicker. One can see that, after initial oscillations, $j(x, t)$ rapidly reduces to the stationary value at positions close to the double barrier. This is due to the fact that the well state suddenly becomes inaccessible to the electrons from the contact on the left-hand side and the reflection of the incoming distribution increases from one moment to the next.

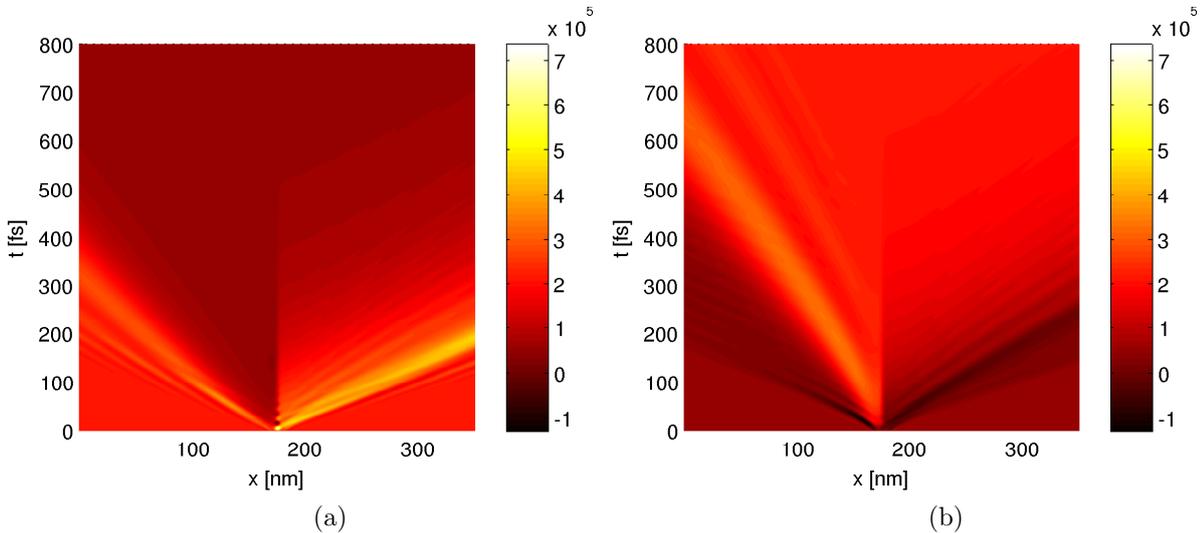


Figure 5.17: Time evolution of the current density $j(x, t)$ for the transient response simulations. The plot in (a) depicts the switching from the peak to the valley voltage and the one in (b) the switching from the valley to the peak voltage.

In Fig. 5.17 (b) the time evolution of $j(x, t)$ for the case of switching from the valley to the peak voltage is depicted. Similar to the previous case, a beam of $j(x, t)$ starting in the central region and propagating to the right can be seen, only that this region exhibits now lower current density values than the steady state at $t = 0$ fs. Again, the electrons tunneling through the device are suddenly accelerated in a different way, now less drastically due to the lower bias voltage. This effect is quickly covered by an increased current density on the left-hand side of the double barrier, resulting in a net rise of $I(t)$ after short times. This increased current density results from the

extenuated momenta in the reflected part of the distribution on the left-hand side since electrons with momenta exhibiting an energy alignment with the well state can now tunnel through the barriers.

Figure 5.18 depicts the time evolution of the particle density $n(x, t)$ in the region of the double barrier structure. As one can see, the filling of the well state by resonant tunneling, see the plot in (b), turns out to be a faster process than emptying the well state by ordinary tunneling through the second barrier, see (a).

The full dynamics can be seen when looking at the Wigner distribution itself and it is thus elucidating to examine the time evolution of $f(p, x, t)$ in phase space. Figures 5.19 and 5.20 illustrate the time evolution of $f(p, x, t)$ by depicting plots of the phase space distribution for different times. In Fig. 5.19 the case of switching from the peak to the valley voltage is shown. One can clearly see in the plots in (b),(c),(d) the accelerated electrons from the well state at $p/\hbar \approx 0.5 \text{ nm}^{-1}$, causing the initial rise of the current. On the left-hand side of the barriers, the previously extenuated part of $f(p, x, t)$ at momenta of $p/\hbar \approx -0.2 \text{ nm}^{-1}$ quickly rises again since the well state has become inaccessible. This corresponds to the quick decrease of $j(x, t)$ on the left-hand side of the barriers seen in Fig. 5.17 (a). Figure 5.20 depicts the case of switching from the valley to the peak voltage. It is obvious that electrons injected from the left quickly tunnel into the well state since on the one hand, the corresponding part in the reflected distribution is immediately extenuated, at $p/\hbar \approx -0.2 \text{ nm}^{-1}$, and on the other hand, since the values of $f(p, x, t)$ in the central region at $x \approx 175 \text{ nm}$ show a rapid increase and already at $t = 100 \text{ fs}$ a significant occupation of the well state can be seen. Characteristic for the resonant tunneling process is the formation of a sharp oscillation pattern. One can see from Fig. 5.20 that the oscillations at $p/\hbar \approx 0.15 \text{ nm}^{-1}$ start out with a rather long-scaled pattern and then become sharper together with spreading out over the phase space domain in the course of the time evolution.

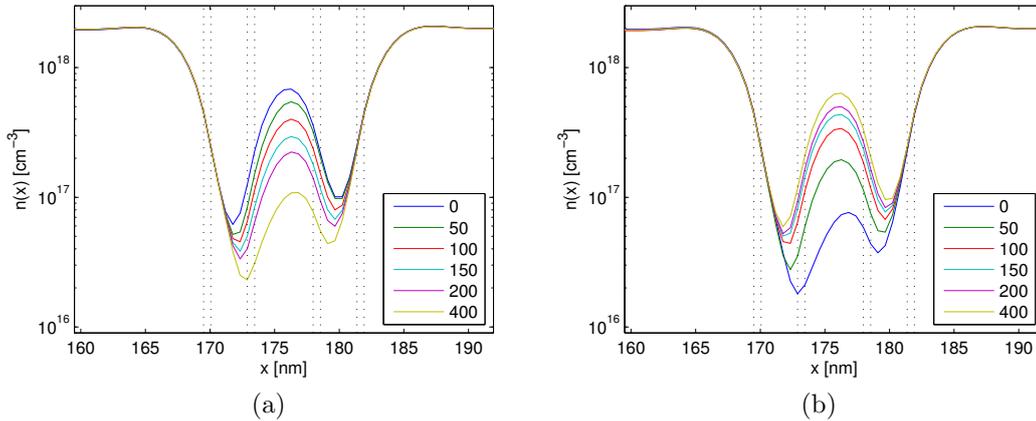


Figure 5.18: Time evolution of the particle density $n(x, t)$ for the transient response simulations. The plot in (a) depicts the switching from the peak to the valley voltage and the one in (b) the switching from the valley to the peak voltage, each one for times from $t = 0 \text{ fs}$ to $t = 400 \text{ fs}$. The positions of the barriers are indicated with dotted lines.

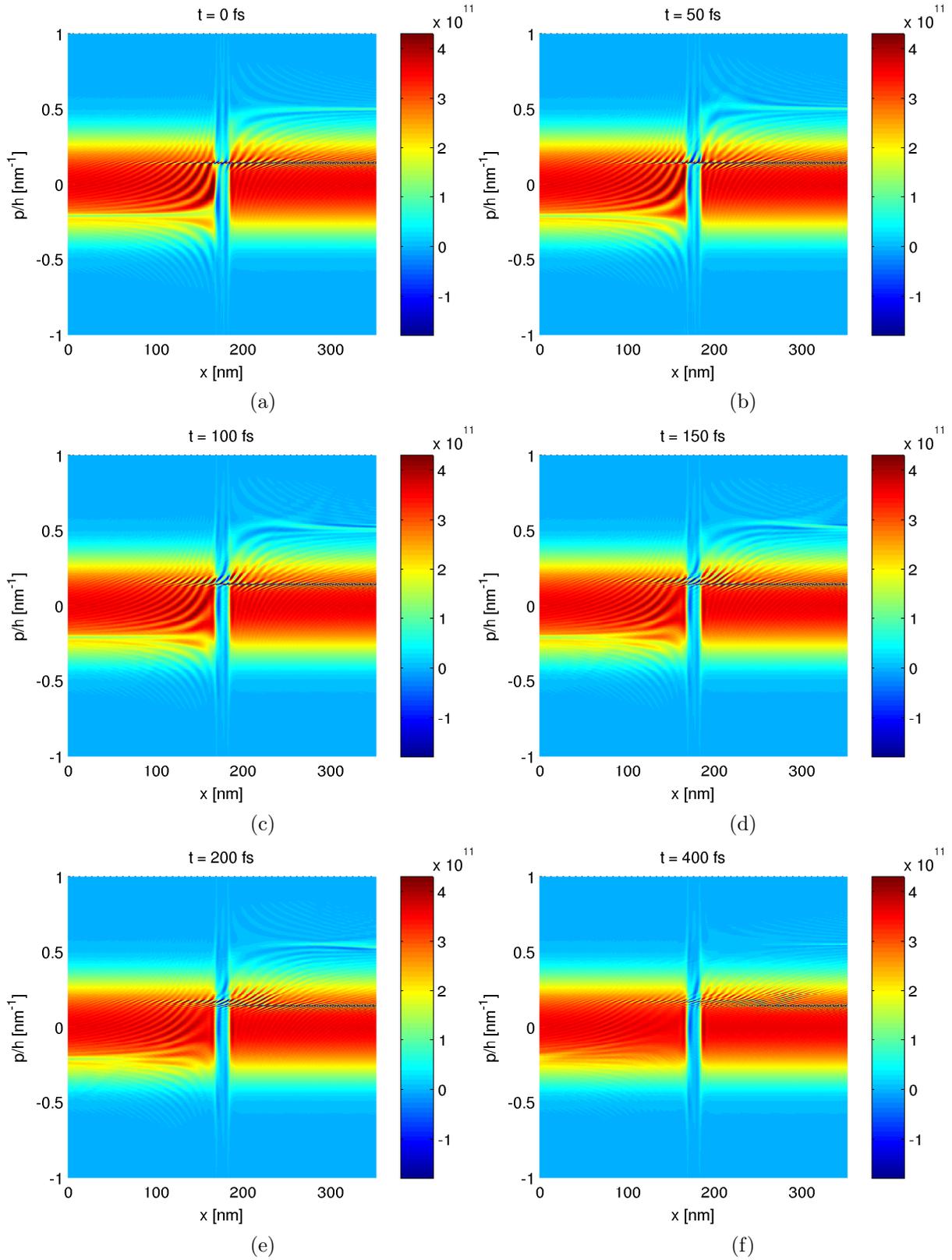


Figure 5.19: Time evolution of $f(p, x, t)$ in phase space when switching from the peak voltage $V_{DS} = 0.115$ V to the valley voltage $V_{DS} = 0.185$ V at $t = 0$ fs. Simulation for $\Delta x = a$, $\Delta p_{max}/\hbar = 0.1$ nm $^{-1}$, $L_c = 300 a$ and $p_{max}/\hbar = 3$ nm $^{-1}$.

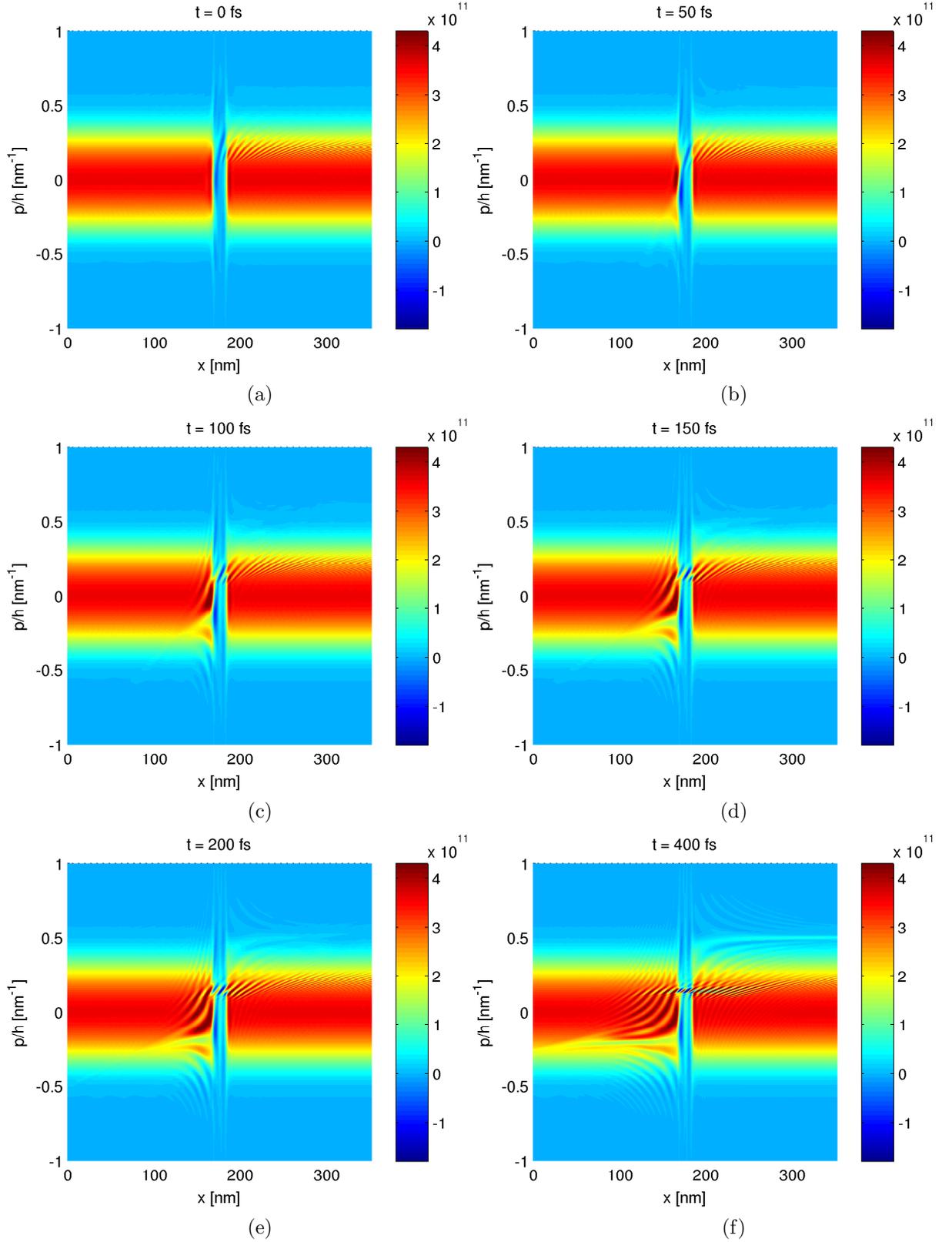


Figure 5.20: Time evolution of $f(p, x, t)$ in phase space when switching from the valley voltage $V_{DS} = 0.185$ V to the peak voltage $V_{DS} = 0.115$ V at $t = 0$ fs. Simulation for $\Delta x = a$, $\Delta p_{max}/\hbar = 0.1 \text{ nm}^{-1}$, $L_c = 300 a$ and $p_{max}/\hbar = 3 \text{ nm}^{-1}$.

Conclusion and outlook

In this work, a novel numerical method for the Wigner transport equation (WTE) was developed. This method relies on similar ideas as used in finite element (FE) methods or as in the case of the REA (reconstruct-evolve-average) algorithm of finite volume (FV) methods. The central point is to make use of a local basis to approximate the Wigner function. This enables us to apply a non-equidistant grid and thus to adapt the numerical domain to the physical situation without the constraints known from the common approach based on the application of a discrete Fourier transformation. Before the development of the final, successful method, the first attempt was to approximate the Wigner function by a continuous, piecewise linear approximation. But, the requirement of a continuous approximation was discarded thereafter due to the following two reasons: On the one hand, several numerical disadvantages resulted from demanding the continuity of the Wigner function and on the other hand, it was experienced that the effect of steps in the Wigner function due to a non-continuous approximation is negligible on fine enough grids. Therefore, the second and finally developed method uses a more general, non necessarily continuous, piecewise polynomial approximation of the Wigner function. To ensure the consistency with the continuity equation and thus the conservation of the particle density for arbitrary grids, properly chosen grid cells were defined and the WTE was averaged over this cells. In particular, a first-order approximation of the Wigner function with respect to the momentum variable was used to calculate the action of the pseudo-differential operator. Upon averaging the WTE over grid cells, the advection term was rewritten in terms of the difference of the fluxes at the cell boundaries. This fluxes were determined by a well approved scheme from FV methods, namely by a weighted essentially non oscillatory scheme with an accuracy up to fifth order (WENO5).

Before examining the overall accuracy of the numerical method for the WTE, the time evolution of an initial state towards the sought-for stationary Wigner function was analyzed in detail. This preliminary tests were necessary to comprehend the origin of the non-smooth convergence behaviour for certain grids. The algorithm revealed the peculiar behaviour that a sudden buildup of errors may occur in the course of the time evolution to a steady state. The various simulations performed enabled us to link this sudden buildup of errors to a bad resolution of the sharp oscillating features in phase space. As discovered in the course of this work, the stationary Wigner function for tunneling barriers may show a stripe of rapidly varying and large-valued oscillations. To resolve this sharp oscillations at all it seems to be inevitable to make use of a non-equidistant grid spacing with respect to the momentum variable. When examining the time evolution to a steady state this stripe of oscillations builds up comparatively slowly as time progresses. If the grid resolution is not fine enough, the

oscillations cannot be properly resolved anymore at a certain time and errors build up. By making extensive use of the ability to adapt the spacings of the momentum variable, it was possible to well resolve all of the oscillating features in the Wigner function for the test case of a resonant tunneling diode (RTD). This enabled us to achieve a smooth convergence behaviour and a stable stationary state.

Once it was understood how to suppress the creation of spurious oscillations and how to ensure a smooth convergence behaviour to a truly stationary state, the obtained steady state current voltage characteristic of the simulated RTD was compared with a non-equilibrium Green's function (NEGF) reference. The results revealed a very good agreement of the NEGF and the Wigner calculations for large enough grids. As we found out, it is not only important to consider a large enough momentum domain but especially to apply the boundary conditions, i.e. to define the contacts, at rather large distances from the region of the double barriers to not distort the device behaviour. For the case of a sufficiently large computational domain, the spatial grid spacing was reduced to examine the convergence of the results to the NEGF reference. The convergence could be demonstrated and it was possible to achieve a relative difference of the current values below one percent in the best case. This can be seen as a major success of the developed algorithm when comparing the results with those of other deterministic algorithms for the WTE presented so far. The high accuracy of the developed algorithm is attributable to the high-order approximation of the fluxes by the WENO5 method and particularly to the ability to adapt the grid to the severe oscillations of the stationary Wigner functions. The sharp oscillating features of the Wigner function discovered in the simulations are interesting from a physical point of view and seem to be closely related to the resonant tunneling process itself and to the coherence of the reflected and the transmitted parts in phase space.

In the last part of this work, a large-signal transient response simulation was performed as an example for a simple time-dependent situation. Especially for the case of RTDs, fully time-dependent potentials would be of interest, since possible applications of such devices include high frequency (THz) oscillators. For this purpose, one could investigate the frequency dependence of relevant quantities of a RTD, such as the impedance or capacity, by applying a modulated bias voltage. From a numerical point of view, a time-dependent simulation in this manner is definitely feasible for the present algorithm and comparatively simple to achieve. Other possible extensions are to couple the WTE to Poisson's equation in order to take charge variations into account, to include scattering effects via a collision operator or to include a spatially varying effective mass. In principle, these extensions have already been considered in literature. But despite of this, it could be of great interest to include them in the developed algorithm for the purpose of a more realistic device description, together with the ability to achieve very accurate simulations. Especially the adaptable grid seems to be a major advantage of the presented algorithm, which enables us to adjust the computational domain to the particular physical situation without additional constraints. As discovered in this work, the Wigner function is in general by no means as a smooth and slowly varying function as its classical analogue, the Boltzmann distribution, but may vary over length scales differing in orders of magnitude.

Acknowledgments

First of all, I want to thank my supervisor Prof. Dr. Ferdinand Schürerer for offering me the possibility to work on this interesting topic. Besides stating specific goals for this thesis he gave me great freedom to contribute own ideas, even if they turned out to be rather time consuming. What started more or less as a side project covers now most of the present thesis and I really want to thank him, not only for offering me this invaluable possibility but also for his support and encouragement throughout the progress of the thesis. A big thank you for a very pleasant supervision in general and, above all, for the large amount of time he invested in all of the meetings and especially in a careful proofreading and inspection of the calculations.

I also want to thank Dr. Omar Morandi for introducing me to his interesting work on the Wigner function for graphene and to the subject of modeling the charge transport with the Wigner function formalism in general. Unfortunately the time available did not suffice to do practical computations for graphene within this thesis, but it was definitely enriching to get a basic understanding of a multiband description of the carrier dynamics.

Furthermore, I want to thank all the other members of the research group. In particular I would like to thank Dr. Karin Zojer, Dr. Manfred Gruber, Dr. Stefan Possanner and DI Iris Hehn for a pleasant time together, fruitful discussions about physical problems and, of course, also for various not-so-physical coffee breaks.

Last but not least, I want to thank especially my parents for their great support, not only during my studies but whenever I needed them. A big thank you to my whole family and also to all of my friends for encouraging me and for cheering me up when things did not work out in the first place.

6 Appendix

6.1 Singularities in the Wigner function for the square potential barrier

The square potential barrier was used in the course of this work as a first test case to evaluate the performance and accuracy of the developed numerical method. The test problem was chosen since a simple, analytical solution for the wave function is known. The analytical solution served as a reference for the calculated steady states with the Wigner transport equation and a fairly good agreement could be found for rather short simulation times. But, no smooth convergent behaviour could be achieved and the creation of spurious oscillations posed a problem. The problems manifested itself in the time evolution of the current density and the particle density in a very similar manner as in the case of the results for the under-resolved grids presented in Sec. 5.1. In phase space one could observe the emergence of a very sharp stripe of oscillations around $p = 0$, which was not confined to the central region but spread out along x in the course of time. Exactly this behaviour increased the degree of oscillations of the Wigner function on the whole considered domain and lead to wrong results. Since it was not clear at this stage whether the oscillations resulted from an error in the algorithm or were truly physical, the analytical solution was examined in more detail. In principle one could calculate the whole Wigner function out of the knowledge of the wave function but for our purpose it is sufficient to focus on the part around $p = 0$.

An important aspect can already be seen from a much simpler example. For this we return to the calculations outlined in Sec. 1.1, in particular to the results for the two examples of a coherent and an incoherent superposition of two plane waves traveling in opposite directions. For the coherent case, with the density operator given by

$$\hat{\rho}_k = \frac{1}{2}(|k\rangle\langle k| + |-k\rangle\langle -k| + |-k\rangle\langle k| + |k\rangle\langle -k|) \quad (6.1)$$

together with

$$\langle x | \pm k \rangle = \exp(\pm ikx), \quad (6.2)$$

we found the following Wigner function

$$f_k(p, x) = \frac{1}{2}[\delta(p - \hbar k) + \delta(p + \hbar k)] + \cos(2kx)\delta(p), \quad (6.3)$$

compare Eq. (1.33). For simplicity we restrict ourselves now to the one-dimensional case $p, x, k \in \mathbb{R}$. The two delta functions $\delta(p \pm \hbar k)$ correspond to the two plane waves

with wave numbers $\pm k$, whereas the delta function at $p = 0$ refers to the interference among them. For an actual comparison with numerical results and to model the inflow of a continuum of states from the contacts, it is convenient to consider the following density operator

$$\hat{\rho} = \int_{-\infty}^{\infty} \hat{\rho}_k \mathcal{N}(k; k_0, \sigma) dk, \quad (6.4)$$

where $\mathcal{N}(k; k_0, \sigma)$ denotes in this context the normal distribution

$$\mathcal{N}(k; k_0, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(k - k_0)^2}{2\sigma^2} \right]. \quad (6.5)$$

A normal distribution was chosen for simplicity and any other distribution could be used equally well. In principle, to model the inflow from contacts one should limit the integration in Eq. (6.4) to the interval $(0, \infty)$ rather than $(-\infty, \infty)$, but for a sufficiently small σ in combination with a large value of k_0 the two expressions are approximately equal. Due to the linearity of the Wigner transformation the corresponding Wigner function is given by the analogous expression

$$\begin{aligned} f(p, x) &= \int_{-\infty}^{\infty} f_k(p, x) \mathcal{N}(k; k_0, \sigma) dk \\ &= \frac{1}{2\hbar} \left[\mathcal{N}\left(\frac{p}{\hbar}; k_0, \sigma\right) + \mathcal{N}\left(-\frac{p}{\hbar}; k_0, \sigma\right) \right] + \delta(p) \int_{-\infty}^{\infty} \cos(2kx) \mathcal{N}(k; k_0, \sigma) dk \\ &= \frac{1}{2\hbar} \left[\mathcal{N}\left(\frac{p}{\hbar}; k_0, \sigma\right) + \mathcal{N}\left(-\frac{p}{\hbar}; k_0, \sigma\right) \right] + \delta(p) \cos(2k_0x) \exp(-2x^2\sigma^2), \end{aligned} \quad (6.6)$$

where the first two terms are easily obtained and the last integral can be evaluated with the formulas for Gauss integrals, compare e.g. [43]. As one can see, despite of the continuum of plane wave states a singularity is still present at $p = 0$. In contrast to Eq. (6.3), the oscillation pattern decays exponentially with $|x| \rightarrow \infty$ but may still be prominent in some regions. In this case the position $x = 0$ has a special role, since no phase difference was included in Eq. (6.1) and as a consequence all of the single plane waves are in phase at $x = 0$.

This simple example illustrates the possibility of singularities in the Wigner function even though boundary conditions with a smooth momentum dependence are applied. It seems natural that the sharp stripe of oscillations at $p = 0$ observed in the simulations on the square barrier can be explained in a similar way. To see if these oscillations are truly physical, we make use of the analytical solution of the wave function for the square barrier and write down the expression for the corresponding Wigner function. Numerous integrals are involved and we will calculate solely those that can result in singular expressions. As just observed, especially singular expressions where the positions of the poles do not depend on the wave number k pose a problem.

The square potential barrier is chosen to exhibit a length L , a height U_0 and to be centered around $x = 0$. In particular the potential $U(x)$ is then given by

$$U(x) = \begin{cases} U_0 & \text{if } |x| \leq \frac{L}{2}, \\ 0 & \text{elsewhere.} \end{cases} \quad (6.7)$$

Since the potential is piecewise constant, plane waves can be used as ansatz for the wave function in the single regions, so that [44]

$$\phi_k(x) = \begin{cases} \exp(ikx) + A_k \exp(-ikx) & \text{if } x < -\frac{L}{2}, \\ B_k \exp(\kappa x) + B'_k \exp(-\kappa x) & \text{if } |x| \leq \frac{L}{2}, \\ C_k \exp(ikx) & \text{if } x > \frac{L}{2}, \end{cases} \quad (6.8)$$

where κ can be either real or complex, depending on the height U_0 relative to the energy $E = \frac{\hbar^2 k^2}{2m}$ and is given by [44]

$$\kappa = \sqrt{\frac{2m(U_0 - E)}{\hbar^2}}. \quad (6.9)$$

In the following we assume κ to be real, i.e. that $E < U_0$. The single coefficients A_k , B_k , B'_k and C_k can be determined by demanding the continuity of $\phi_k(x)$ and of its first derivative with respect to x at the boundaries $x = \pm \frac{L}{2}$. Details are not recapped here and can be found for instance in [44] or [12]. In order to calculate the corresponding Wigner function one could make direct use of the position space representation Eq. (6.8), as done for instance in [45]. Due to the required, tedious case differentiation to account for the different intervals in Eq. (6.8), it was deemed simpler to start from the momentum space representation of $|\phi_k\rangle$. This was found especially useful when one is solely interested in identifying singular terms in the Wigner function. An equivalent definition of the Wigner transformation is given by [2]

$$\begin{aligned} f(p, x) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \left\langle x + \frac{1}{2}\xi \left| \hat{\rho} \right| x - \frac{1}{2}\xi \right\rangle \exp\left(-i\frac{p}{\hbar}\xi\right) d\xi \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \left\langle p + \frac{1}{2}q \left| \hat{\rho} \right| p - \frac{1}{2}q \right\rangle \exp\left(i\frac{q}{\hbar}x\right) dq, \end{aligned} \quad (6.10)$$

which can be found when inserting two times the following expression for the identity $\mathbb{1} = \frac{1}{2} \int |p \pm \frac{q'}{2}\rangle \langle p \pm \frac{q'}{2}| dq'$ and by making use of the relation

$$\langle x | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(i\frac{p}{\hbar}x\right), \quad (6.11)$$

as well as the representation of the delta function given in Eq. (1.17). The momentum space representation of $|\phi_k\rangle$ is obtained by a Fourier transformation

$$\tilde{\phi}_k(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \phi_k(x) \exp\left(-i\frac{p}{\hbar}x\right) dx. \quad (6.12)$$

Now, unlike in the rest of this work, we use a slightly different definition of the Fourier transformation by distributing the factor $(2\pi\hbar)^{-1}$ in a symmetric way on the expressions for the ordinary and the inverse Fourier transformation, as can be seen from Eq. (6.12). This was found to be necessary to be consistent with Eqs. (6.11) and (6.10). To evaluate $\tilde{\phi}_k(p)$ for the wave function given in Eq. (6.8), we need to calculate the following integrals

$$\begin{aligned}\tilde{\phi}_k(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{-\frac{L}{2}} [\exp(ikx) + A_k \exp(-ikx)] \exp\left(-i\frac{p}{\hbar}x\right) dx \\ &\quad + \frac{1}{\sqrt{2\pi\hbar}} \int_{-\frac{L}{2}}^{\frac{L}{2}} [B_k \exp(\kappa x) + B'_k \exp(-\kappa x)] \exp\left(-i\frac{p}{\hbar}x\right) dx \\ &\quad + \frac{1}{\sqrt{2\pi\hbar}} \int_{\frac{L}{2}}^{\infty} C_k \exp(ikx) \exp\left(-i\frac{p}{\hbar}x\right) dx.\end{aligned}\quad (6.13)$$

To calculate the semi-infinite integrals we make once more use of the expression for the Fourier transform of the Heaviside step function, Eq. (3.121), in order to arrive at

$$\begin{aligned}\tilde{\phi}_k(p) &= \sqrt{2\pi\hbar} \left\{ \left[\frac{1}{2\pi} \frac{i}{p - \hbar k} + \frac{1}{2} \delta(p - \hbar k) \right] \exp\left[i \left(\frac{p}{\hbar} - k \right) \frac{L}{2} \right] \right. \\ &\quad \left. + A_k \left[\frac{1}{2\pi} \frac{i}{p + \hbar k} + \frac{1}{2} \delta(p + \hbar k) \right] \exp\left[i \left(\frac{p}{\hbar} + k \right) \frac{L}{2} \right] \right\} \\ &\quad + \frac{1}{\sqrt{2\pi\hbar}} \left\{ B_k \frac{1}{\kappa - i\frac{p}{\hbar}} \left\{ \exp\left[\left(\kappa - i\frac{p}{\hbar} \right) \frac{L}{2} \right] - \exp\left[- \left(\kappa - i\frac{p}{\hbar} \right) \frac{L}{2} \right] \right\} \right. \\ &\quad \left. + B'_k \frac{1}{\kappa + i\frac{p}{\hbar}} \left\{ \exp\left[\left(\kappa + i\frac{p}{\hbar} \right) \frac{L}{2} \right] - \exp\left[- \left(\kappa + i\frac{p}{\hbar} \right) \frac{L}{2} \right] \right\} \right\} \\ &\quad + C_k \sqrt{2\pi\hbar} \left[\frac{1}{2\pi} \frac{-i}{p - \hbar k} + \frac{1}{2} \delta(p - \hbar k) \right] \exp\left[-i \left(\frac{p}{\hbar} - k \right) \frac{L}{2} \right].\end{aligned}\quad (6.14)$$

One can see again the emergence of singular terms. In addition to the delta peaks known from the previous example, singular functions of the form $(p \pm \hbar k)^{-1}$ are present as well. The terms in the third and fourth line with B_k and B'_k as coefficients are considered to be smooth and well-behaved functions, since we assume that κ is real and nonzero. It is convenient to regroup the different terms by their functional

behaviour and we therefore introduce

$$\begin{aligned}
 \tilde{\phi}_k^\delta(p) &= \sqrt{\frac{\pi\hbar}{2}} [(1 + C_k)\delta(p - \hbar k) + A_k\delta(p + \hbar k)], \\
 \tilde{\phi}_k^{\frac{1}{z}}(p) &= \sqrt{\frac{\hbar}{2\pi}} \left\{ \frac{i}{p - \hbar k} \exp\left[i\left(\frac{p}{\hbar} - k\right)\frac{L}{2}\right] + A_k \frac{i}{p + \hbar k} \exp\left[i\left(\frac{p}{\hbar} + k\right)\frac{L}{2}\right] \right. \\
 &\quad \left. + C_k \frac{-i}{p - \hbar k} \exp\left[-i\left(\frac{p}{\hbar} - k\right)\frac{L}{2}\right] \right\}, \\
 \tilde{\phi}_k^{\text{sm.}}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \left\{ B_k \frac{1}{\kappa - i\frac{p}{\hbar}} \left\{ \exp\left[\left(\kappa - i\frac{p}{\hbar}\right)\frac{L}{2}\right] - \exp\left[-\left(\kappa - i\frac{p}{\hbar}\right)\frac{L}{2}\right] \right\} \right. \\
 &\quad \left. + B'_k \frac{1}{\kappa + i\frac{p}{\hbar}} \left\{ \exp\left[\left(\kappa + i\frac{p}{\hbar}\right)\frac{L}{2}\right] - \exp\left[-\left(\kappa + i\frac{p}{\hbar}\right)\frac{L}{2}\right] \right\} \right\}, \quad (6.15)
 \end{aligned}$$

so that the Fourier transformed wave function can be written as

$$\tilde{\phi}_k(p) = \tilde{\phi}_k^\delta(p) + \tilde{\phi}_k^{\frac{1}{z}}(p) + \tilde{\phi}_k^{\text{sm.}}(p). \quad (6.16)$$

According to Eq. (6.10), the corresponding Wigner function is then given by

$$f_k(p, x) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \tilde{\phi}_k\left(p + \frac{1}{2}q\right) \left[\tilde{\phi}_k\left(p - \frac{1}{2}q\right)\right]^* \exp\left(i\frac{q}{\hbar}x\right) dq. \quad (6.17)$$

To calculate the full Wigner function $f_k(p, x)$ one would need to solve several integrals. But, for our purposes it suffices to consider the contributions from the two singular functions $\tilde{\phi}_k^\delta(p)$ and $\tilde{\phi}_k^{\frac{1}{z}}(p)$. To analyze if a delta function peak is present at $p = 0$, we calculate as a first step

$$\begin{aligned}
 f_k^a(p, x) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \tilde{\phi}_k^\delta\left(p + \frac{1}{2}q\right) \left[\tilde{\phi}_k^\delta\left(p - \frac{1}{2}q\right)\right]^* \exp\left(i\frac{q}{\hbar}x\right) dq, \\
 &= \frac{1}{4} \left[\int_{-\infty}^{\infty} |1 + C_k|^2 \delta\left(p + \frac{q}{2} - \hbar k\right) \delta\left(p - \frac{q}{2} - \hbar k\right) \exp\left(i\frac{q}{\hbar}x\right) dq \right. \\
 &\quad + \int_{-\infty}^{\infty} |A_k|^2 \delta\left(p + \frac{q}{2} + \hbar k\right) \delta\left(p - \frac{q}{2} + \hbar k\right) \exp\left(i\frac{q}{\hbar}x\right) dq \\
 &\quad + \int_{-\infty}^{\infty} (1 + C_k) A_k^* \delta\left(p + \frac{q}{2} - \hbar k\right) \delta\left(p - \frac{q}{2} + \hbar k\right) \exp\left(i\frac{q}{\hbar}x\right) dq \\
 &\quad \left. + \int_{-\infty}^{\infty} (1 + C_k^*) A_k \delta\left(p + \frac{q}{2} + \hbar k\right) \delta\left(p - \frac{q}{2} - \hbar k\right) \exp\left(i\frac{q}{\hbar}x\right) dq \right]. \quad (6.18)
 \end{aligned}$$

The integrals are easily evaluated when making use of the relation

$$\delta(a + b)\delta(a - b) = \frac{1}{2}\delta(a)\delta(b), \quad (6.19)$$

which can be found when using the specific representation of the delta function given in Eq. (1.17) and performing a change of the integration variables. One thus finds

$$f_k^a(p, x) = \frac{1}{4} \left[|1 + C_k|^2 \delta(p - \hbar k) + |A_k|^2 \delta(p + \hbar k) \right] + \frac{1}{2} \Re \{ (1 + C_k) A_k^* \exp(i2kx) \} \delta(p). \quad (6.20)$$

The obtained expression is in close analogy to the Wigner function for the previous example, see Eq. (6.3). Again, a delta function is present at $p = 0$. The coefficient in front depends on the product of $(1 + C_k)$ and A_k^* , i.e. on the amplitudes of the plane waves moving to the right and to the left, respectively, compare Eq. (6.8).

Two other terms we look at in more detail are

$$\begin{aligned} f_k^{b'}(p, x) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \tilde{\phi}_k^\delta \left(p + \frac{1}{2}q \right) \left[\tilde{\phi}_k^{\frac{1}{z}} \left(p - \frac{1}{2}q \right) \right]^* \exp\left(i\frac{q}{\hbar}x\right) dq, \\ &= \frac{1}{\sqrt{2\pi\hbar}} (1 + C_k) \exp\left(-i2\frac{p - \hbar k}{\hbar}x\right) \left[\tilde{\phi}_k^{\frac{1}{z}}(2p - \hbar k) \right]^* \\ &\quad + \frac{1}{\sqrt{2\pi\hbar}} A_k \exp\left(-i2\frac{p + \hbar k}{\hbar}x\right) \left[\tilde{\phi}_k^{\frac{1}{z}}(2p + \hbar k) \right]^*, \end{aligned} \quad (6.21)$$

and

$$\begin{aligned} f_k^{b''}(p, x) &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \tilde{\phi}_k^{\frac{1}{z}} \left(p + \frac{1}{2}q \right) \left[\tilde{\phi}_k^\delta \left(p - \frac{1}{2}q \right) \right]^* \exp\left(i\frac{q}{\hbar}x\right) dq, \\ &= \frac{1}{\sqrt{2\pi\hbar}} (1 + C_k^*) \exp\left(i2\frac{p - \hbar k}{\hbar}x\right) \tilde{\phi}_k^{\frac{1}{z}}(2p - \hbar k) \\ &\quad + \frac{1}{\sqrt{2\pi\hbar}} A_k^* \exp\left(i2\frac{p + \hbar k}{\hbar}x\right) \tilde{\phi}_k^{\frac{1}{z}}(2p + \hbar k). \end{aligned} \quad (6.22)$$

Since $f_k^{b''}(p, x) = [f_k^{b'}(p, x)]^*$, the sum of the two terms is given by

$$\begin{aligned} f_k^b(p, x) &= f_k^{b'}(p, x) + f_k^{b''}(p, x) \\ &= 2\Re\{f_k^{b'}(p, x)\}. \end{aligned} \quad (6.23)$$

In particular, this expression can be evaluated to

$$\begin{aligned} f_k^b(p, x) &= \Re \left\{ \frac{1}{2\pi} (1 + C_k^*) \exp\left(i2\frac{p - \hbar k}{\hbar}x\right) \left[\frac{i}{p - \hbar k} \exp\left(i\frac{p - \hbar k}{\hbar}L\right) \right. \right. \\ &\quad \left. \left. + A_k \frac{i}{p} \exp\left(i\frac{p}{\hbar}L\right) - C_k \frac{i}{p - \hbar k} \exp\left(-i\frac{p - \hbar k}{\hbar}L\right) \right] \right. \\ &\quad \left. + \frac{1}{2\pi} A_k^* \exp\left(i2\frac{p + \hbar k}{\hbar}x\right) \left[\frac{i}{p} \exp\left(i\frac{p}{\hbar}L\right) \right. \right. \\ &\quad \left. \left. + A_k \frac{i}{p + \hbar k} \exp\left(i\frac{p + \hbar k}{\hbar}L\right) - C_k \frac{i}{p} \exp\left(-i\frac{p}{\hbar}L\right) \right] \right\}. \end{aligned} \quad (6.24)$$

After regrouping the terms with respect to the singular functions one arrives at

$$\begin{aligned}
 f_k^b(p, x) = & -\frac{1}{2\pi} \Im \left\{ \frac{1}{p - \hbar k} (1 + C_k^*) \left\{ \exp \left[i \left(\frac{p}{\hbar} - k \right) (2x + L) \right] \right. \right. \\
 & \left. \left. - C_k \exp \left[i \left(\frac{p}{\hbar} - k \right) (2x - L) \right] \right\} \right. \\
 & \left. + \frac{1}{p} \exp \left(i \frac{2p}{\hbar} x \right) \left\{ [(1 + C_k^*) A_k \exp(-i2kx) + A_k^* \exp(i2kx)] \exp \left(i \frac{p}{\hbar} L \right) \right. \right. \\
 & \left. \left. - C_k A_k^* \exp(i2kx) \exp \left(-i \frac{p}{\hbar} L \right) \right\} \right. \\
 & \left. + \frac{1}{p + \hbar k} |A_k|^2 \exp \left[i \left(\frac{p}{\hbar} + k \right) (2x + L) \right] \right\}. \tag{6.25}
 \end{aligned}$$

As a result, one can notice the occurrence of another singular term at $p = 0$. This particular term can also be written as

$$\begin{aligned}
 f_k^{b,0}(p, x) = & -\frac{1}{2\pi} \Im \left\{ \frac{1}{p} \exp \left(i \frac{2p}{\hbar} x \right) \left[2\Re \{ A_k \exp(-i2kx) \} \exp \left(i \frac{p}{\hbar} L \right) \right. \right. \\
 & \left. \left. - 2i \Im \{ A_k C_k^* \exp(-i2kx) \exp \left(i \frac{p}{\hbar} L \right) \} \right] \right\} \\
 = & -\frac{1}{\pi} \left\{ \frac{1}{p} \sin \left[\frac{p}{\hbar} (2x + L) \right] \Re \{ A_k \exp(-i2kx) \} \right\} \\
 & + \frac{1}{\pi} \left\{ \frac{1}{p} \cos \left(\frac{2p}{\hbar} x \right) \left[\Re \{ A_k C_k^* \exp(-i2kx) \} \sin \left(\frac{p}{\hbar} L \right) \right. \right. \\
 & \left. \left. + \Im \{ A_k C_k^* \exp(-i2kx) \} \cos \left(\frac{p}{\hbar} L \right) \right] \right\}. \tag{6.26}
 \end{aligned}$$

All expressions containing a sine function stay finite for $p \rightarrow 0$ and the only singularity appears in the very last term. We introduce the abbreviation

$$c_k^{\frac{1}{2}}(p, x) = \frac{1}{\pi} \Im \{ A_k C_k^* \exp(-i2kx) \} \cos \left(2 \frac{p}{\hbar} x \right) \cos \left(\frac{p}{\hbar} L \right), \tag{6.27}$$

as well as for the coefficient of the delta function $\delta(p)$ in Eq. (6.20)

$$c_k^\delta(x) = \frac{1}{2} \Re \{ A_k (1 + C_k^*) \exp(-i2kx) \}, \tag{6.28}$$

to write the final result as

$$f_k(p, x) = c_k^\delta(x) \delta(p) + c_k^{\frac{1}{2}}(p, x) \frac{1}{p} + f_k^{\text{rest}}(p, x). \tag{6.29}$$

The expression $f_k^{\text{rest}}(p, x)$ stands for all the other terms of the Wigner function, which are assumed to probably result in a smooth function as soon as one performs an integration over a continuum of k states, as done in Eq. (6.6). Of course, this is not guaranteed and to be certain that $f_k^{\text{rest}}(p, x)$ does not contain any other singular functions with poles at $p = 0$, one would need to examine further terms. Especially the part

$$f_k^c(p, x) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \tilde{\phi}_k^{\frac{1}{z}} \left(p + \frac{1}{2}q \right) \left[\tilde{\phi}_k^{\frac{1}{z}} \left(p - \frac{1}{2}q \right) \right]^* \exp \left(i \frac{q}{\hbar} x \right) dq \quad (6.30)$$

could be of interest. However, this is not done here because we assume that already the results of Eqs. (6.27) to (6.29) reveal the central, singular features of the Wigner function for the case of the square potential barrier. To briefly discuss the coefficients $c_k^\delta(x)$ and $c_k^{\frac{1}{z}}(p, x)$ we recap that the factor of 1 corresponds to the amplitude of the incoming wave, the factor A_k to the amplitude of the reflected one and C_k to the amplitude of the transmitted wave, compare Eq. (6.8). Consider the situation $|C_k| \approx 0$, which occurs for instance in the case of a very high barrier, $U_0 \gg E$, or a very wide one, $L \gg 0$. In this case the coefficient $c_k^{\frac{1}{z}}(p, x)$ of the $1/p$ singularity is approximately zero, but the coefficient $c_k^\delta(x)$ of the delta function can still be prominent. In particular it is then given by $c_k^\delta(x) \approx \frac{1}{2} \Re \{ A_k 1 \exp(-i2kx) \}$ and can be interpreted as a manifestation of the interference between the incoming and the reflected plane wave. For the case that the amplitude of the transmitted wave takes on significant values $|C_k| \neq 0$, i.e. that tunneling is enhanced, the coefficient of the delta function changes on the one hand, and more important, the $1/p$ singularity becomes prominent. This reveals the close relation between the sharp oscillation pattern produced by the function $c_k^{\frac{1}{z}}(p, x) \frac{1}{p}$ and the tunneling processes. Due to the factor $A_k C_k^*$ in the expression for $c_k^{\frac{1}{z}}(p, x)$, one can further relate the oscillation pattern to the coherence between the reflected and the transmitted wave.

The occurrence of the peculiar, singular terms in Eq. (6.29) may be interesting from a physical point of view, but causes severe problems for the numerical solution of the Wigner transport equation. To account for the possibility of singularities at $p = 0$, the second developed method with a discontinuous approximation of the Wigner function was implemented on a grid including the point $p^0 = 0$. The inclusion of a grid point at $p = 0$ should allow for a proper formation of the delta peak. It is interesting to note that for comparatively short time evolutions a fairly good (qualitative) agreement between the factor $c_k^\delta(x)$ and the numerical values $f(p^0, x, t) \Delta p^0$ could be found. Here, Δp^0 labels the width of the grid cell at $p^0 = 0$. But on the whole, no smooth convergent behaviour could be observed and problems with spurious oscillations remained. The appearance of singular terms in the Wigner function poses a problem for the present algorithm and one can expect the build-up of large errors in such situations. In addition to the problem of resolving a singular function properly, the Fourier transform of the term $\frac{1}{p}$ is problematic. In Fourier transformed space it corresponds to a step at $\eta = 0$, compare, e.g., Eq. (3.121). This step cannot be properly resolved in a

numerical simulation and one encounters overshooting and oscillations in $\tilde{f}(\eta, x, t)$ as known from Gibb's phenomenon [25]. To reduce the errors encountered one could, in principle, increase the size of the p domain to reduce the length scale of the oscillations in $\tilde{f}(\eta, x, t)$ and then apply an averaging to obtain a smooth function $\tilde{f}'(\eta, x, t)$ with a decreased deviation from the true step. However, this strategy seems to be unfeasible for an actual numerical simulation and the square barrier was therefore not considered any further.

To avoid the occurrence of singular terms it is important to include a bias voltage in the simulations. This breaks the symmetry of the potential shape and shifts the oscillation patterns away from the point $p = 0$. This effect can be seen for instance when redoing the simple example of a coherent superposition of two plane waves given in Eq. (6.1), but now for the asymmetric case of two plane waves with wave numbers k' and k'' with $|k'| \neq |k''|$. Instead of the delta function $\delta(p)$ in Eq. (6.3) one then obtains the delta function $\delta(p - \hbar \frac{k'+k''}{2})$, i.e. shifted by the mean value of the two wave numbers. In addition, practical simulations revealed that it seems to be important that the bias voltage drops over a sufficiently large region. A single potential step for instance could cause singularities in the Wigner function as well. Also due to this a finite length for the slopes of the double barriers was included in all of the simulations presented in Sec. 5.2. In general it seems to be a good advice to avoid situations with a step in $V(x)$. The results for the RTD presented in Sec. 5.2 reveal the occurrence of a very sharp stripe of oscillations in the Wigner function, as soon as resonant tunneling is prominent. Since it was possible to resolve all of the oscillations one can see that no singularities are present in this case, only heavily oscillating regions in phase space. The higher the bias voltage, the more the stripe of oscillations is shifted to larger values of p . For $V_{DS} \rightarrow 0$ also the position of the oscillations moves to $p = 0$ together with becoming even shorter scaled. In addition, when examining Fig. (5.14) (b) in more detail, one can see that the transmitted part of $f(p, x, t)$ has an approximate momentum of $p'/\hbar \approx 0.5 \text{ nm}^{-1}$. The corresponding reflected part on the left-hand side of the barriers can be identified as the region of the extenuated momenta, at approximately $p''/\hbar \approx -0.2 \text{ nm}^{-1}$. If one calculates the mean of p' and p'' one arrives at $\frac{p'+p''}{2}/\hbar \approx 0.15 \text{ nm}^{-1}$, which is just the position of the stripe of oscillations. This is in accordance to the example of the two plane waves with k' and k'' outlined at the beginning of this paragraph. Therefore, one can conclude that the observed stripe of oscillations in the Wigner function for the RTD can be regarded to be truly physical and furthermore, to be closely related to the coherence of the transmitted and the reflected part of the Wigner function.

Bibliography

- [1] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser, *Semiconductor Equations*. Wien New York: Springer-Verlag, 1990.
- [2] R. L. Liboff, *Kinetic Theory*. New York: Springer, 2003.
- [3] M. Galler, *Multigroup Equations for the Description of the Particle Transport in Semiconductors (Series on Advances in Mathematics for Applied Sciences, V. 70)*. Singapore: World Scientific Publishing Company, 2005.
- [4] W. R. Frensley, “Boundary conditions for open quantum systems driven far from equilibrium,” *Rev. Mod. Phys.*, vol. 62, pp. 745–791, 1990.
- [5] K.-Y. Kim and B. Lee, “On the high-order numerical calculation schemes for the Wigner transport equation,” *Solid-State Electronics*, vol. 43, pp. 2243–2245, 1999.
- [6] H. Jiang, W. Cai, and R. Tsu, “Accuracy of the Frensley inflow boundary condition for Wigner equations in simulating resonant tunneling diodes,” *J. Comput. Phys.*, vol. 230, pp. 2031–2044, 2011.
- [7] V. N. Do, P. Dollfus, and V. L. Nguyen, “Transport and noise in resonant tunneling diode using self-consistent Green’s function calculation,” *Journal of Applied Physics*, vol. 100, p. 093705, 2006.
- [8] P. Vogl and T. Kubis, “The non-equilibrium Green’s function method: an introduction,” *J. Comput. Electron.*, vol. 9, pp. 237–242, 2010.
- [9] T. L. C. Cosmas K. Zachos, David B. Fairlie, *Quantum Mechanics in Phase Space: An Overview with Selected Papers (World Scientific)*. Singapore: World Scientific Publishing Company, 2005.
- [10] W. P. Schleich, *Quantum Optics in Phase Space*. Berlin: Wiley-VCH, 2001.
- [11] R. J. LeVeque, *Finite Volume Methods for Hyperbolic Problems (Cambridge Texts in Applied Mathematics)*. Cambridge: Cambridge University Press, 2002.
- [12] R. Shankar, *Principles of Quantum Mechanics*. New York: Springer, 1994.
- [13] V. May and O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems*. Weinheim: Wiley-VCH, 2011.

- [14] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors (Springer Series in Solid-State Sciences)*. New York: Springer, 2007.
- [15] H. Jiang, S. Shao, W. Cai, and P. Zhang, “Boundary treatments in non-equilibrium Green’s function (NEGF) methods for quantum transport in nano-MOSFETs,” *J. Comput. Phys.*, vol. 227, pp. 6553–6573, 2008.
- [16] J. Watling, J. Barker, and S. Roy, “Quantum potential corrections for spatially dependent effective masses with application to charge confinement at heterostructure interfaces,” *J. Comput. Electron.*, vol. 1, pp. 279–282, 2002.
- [17] H. Tsuchiya, M. Ogawa, and T. Miyoshi, “Simulation of quantum transport in quantum devices with spatially varying effective mass,” *IEEE Transactions on Electron Devices*, vol. 38, pp. 1246–1252, 1991.
- [18] R. J. LeVeque, *Numerical Methods for Conservation Laws*. Basel: Birkhäuser, 2005.
- [19] Y. Yamada, H. Tsuchiya, and M. Ogawa, “Quantum transport simulation of silicon-nanowire transistors based on direct solution approach of the Wigner transport equation,” *IEEE Transactions on Electron Devices*, vol. 56, pp. 1396–1401, 2009.
- [20] S. Shao, T. Lu, and W. Cai, “Adaptive conservative cell average spectral element methods for transient Wigner equation in quantum transport,” *Commun. Comput. Phys.*, 2011.
- [21] O. Frosali, G. Morandi, “A quantum kinetic approach for modeling a two-band resonant tunneling diode,” *Transp. Theory Stat. Phys.*, vol. 36, pp. 159–177, 2007.
- [22] A. Arnold and C. Ringhofer, “An operator splitting method for the Wigner–Poisson problem,” *SIAM J. Numer. Anal.*, vol. 33, pp. 1622–1643, 1996.
- [23] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations (Springer Series in Computational Mathematics)*. New York: Springer, 2006.
- [24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge: Cambridge University Press, 2007.
- [25] M. A. Pinsky, *Introduction to Fourier Analysis and Wavelets (Graduate Studies in Mathematics)*. Pacific Grove: Brooks/Cole, 2002.
- [26] Wolfram Research, “Cosine integral.” <http://mathworld.wolfram.com/CosineIntegral.html>. Web. 29 Aug. 2012.

-
- [27] Wikipedia, “Trigonometric integral.” http://en.wikipedia.org/wiki/Trigonometric_integral. Web. 29 Aug. 2012.
- [28] Wolfram Research, “Sine integral.” <http://mathworld.wolfram.com/SineIntegral.html>. Web. 29 Aug. 2012.
- [29] K. Jänich, *Analysis für Physiker und Ingenieure: Funktionentheorie, Differentialgleichungen, Spezielle Funktionen*. Berlin: Springer, 2001.
- [30] W. Walter, *Einführung in die Theorie der Distributionen*. Mannheim: BI-Wissenschaftsverlag, 1994.
- [31] E. Bueler, “The Fourier transform of the Heaviside function: a tragedy.” <http://www.cs.uaf.edu/~bueler/M611heaviside.pdf>. Web. 29 Aug. 2012.
- [32] X.-D. Liu, S. Osher, and T. Chan, “Weighted essentially non-oscillatory schemes,” *J. Comput. Phys.*, vol. 115, pp. 200–212, 1994.
- [33] G.-S. Jiang and C.-W. Shu, “Efficient implementation of weighted ENO schemes,” *J. Comput. Phys.*, vol. 126, pp. 202 – 228, 1996.
- [34] J. A. Carrillo, A. Majorana, F. Vecil, “A semi-Lagrangian deterministic solver for the semiconductor Boltzmann–Poisson system,” *Commun. Comput. Phys.*, vol. 2, pp. 1027–1054, 2007.
- [35] R. Wang and R. J. Spiteri, “Linear instability of the fifth-order WENO method,” *SIAM J. Numer. Anal.*, vol. 45, pp. 1871–1901, 2007.
- [36] S. Datta, *Quantum Transport: Atom to Transistor*. Cambridge: Cambridge University Press, 2005.
- [37] N. J. Higham, “The scaling and squaring method for the matrix exponential revisited,” *SIAM J. Matrix Anal. Appl.*, vol. 26, pp. 1179–1193, 2005.
- [38] S. Datta, *Electronic Transport in Mesoscopic Systems (Cambridge Studies in Semiconductor Physics and Microelectronic Engineering)*. Cambridge: Cambridge University Press, 1997.
- [39] O. Baumgartner, M. Karner, S. Holzer, M. Pourfath, T. Grasser, and H. Kosina, “Adaptive energy integration of non-equilibrium Green’s function,” *NSTI-Nanotech*, vol. 3, pp. 145 – 148, 2007.
- [40] D. Ferry and S. M. Goodnick, *Transport in Nanostructures (Cambridge Studies in Semiconductor Physics and Microelectronic Engineering)*. Cambridge: Cambridge University Press, 1999.
- [41] T. Espelid, “Doubly adaptive quadrature routines based on Newton–Cotes rules,” *BIT Numerical Mathematics*, vol. 43, pp. 319–337, 2003.

- [42] Wikipedia, “Complete Fermi–Dirac integral.” http://en.wikipedia.org/wiki/Complete_Fermi-Dirac_integral. Web. 27 Sept. 2012.
- [43] Wikipedia, “Integral of a Gaussian function.” http://en.wikipedia.org/wiki/Integral_of_a_Gaussian_function. Web. 29 Okt. 2012.
- [44] H. G. Evertz, “Quantum mechanics.” http://itp.tugraz.at/LV/evertz/QM_Skript/qm.html. Lecture notes. Web. 29 Okt. 2012.
- [45] J. Muga, R. Sala, and S. Brouard, “Wigner function for the square barrier,” *Solid State Commun.*, vol. 94, pp. 877 – 882, 1995.
- [46] Wikipedia, “Sinc function.” http://en.wikipedia.org/wiki/Sinc_function. Web. 29 Aug. 2012.