

Simulation Studies of Carbon Nanotube Field-Effect Transistors

by

David Llewellyn John

B.A.Sc., The University of British Columbia, 2002

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

The Faculty of Graduate Studies

(Electrical and Computer Engineering)

The University Of British Columbia

July 2006

© David Llewellyn John 2006

Abstract

Simulation studies of carbon nanotube field-effect transistors (CNFETs) are presented using models of increasing rigour and versatility that have been systematically developed. Firstly, it is demonstrated how one may compute the standard tight-binding band structure. From this foundation, a self-consistent solution for computing the equilibrium energy band diagram of devices with Schottky-barrier source and drain contacts is developed. While this does provide insight into the likely behaviour of CNFETs, a non-equilibrium model is required in order to predict the current-voltage relation. To this end, the effective-mass approximation is utilized, where a parabolic fit to the band structure is used in order to develop a Schrödinger-Poisson solver. This model is employed to predict both DC behaviour and switching times for CNFETs, and was one of the first models that captured quantum effects, such as tunneling and resonance, in these devices. In addition, this model has been used in order to validate compact models that incorporated tunneling via the WKB approximation. A modified WKB derivation is provided in order to account for the non-zero reflection of carriers above a potential energy step. In order to allow for greater flexibility in the CNFET geometries, and to lift the effective-mass approximation, a non-equilibrium Green's function method is finally developed, which uses an atomistic tight-binding Hamiltonian to model doped-contact, as opposed to Schottky-barrier-contact, devices. This approach benefits by being able to account for both inter- and intra-band tunneling, and by utilizing a quadratic matrix equation in order to improve the computation time for the required self-energy matrices. Within this technique, an expression for the local inter-atomic current is derived in order to provide more detailed information than the usual compact expression for the terminal current. With this final model, an investigation is presented into the effects of geometrical variations, contact thicknesses, and azimuthal variation in the surface potential of the nanotube.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Figures	vi
Acknowledgements	viii
Dedication	ix
1 Introduction	1
1.1 Background	1
1.2 Simulation Studies	4
2 Band Structure	6
2.1 Graphene	6
2.2 Zone-Folding	9
3 Equilibrium	13
3.1 Electrostatics	14
3.1.1 Closed Geometry	14
3.1.2 Open Geometry	15
3.2 Solution Procedure	16
3.3 Results and Discussion	17
4 The Effective-Mass Method	21
4.1 Finite Difference Electrostatics	22

4.2	Treating Quantum Phenomena	23
4.2.1	Asymptotic Treatment of Tunneling	23
4.2.2	Solving Schrödinger's Equation	26
4.2.3	Quantum Capacitance	30
4.3	Planar Devices	34
4.3.1	Finite Element Electrostatics	35
4.4	Solution Procedure	37
4.5	Results and Discussion	39
4.5.1	DC	39
4.5.2	Switching Speed	43
5	The Tight-Binding Method	51
5.1	Relevance of the Green's Function	51
5.2	The Hamiltonian	55
5.3	Computing the Green's Function	56
5.3.1	Green's Function for a Semi-Infinite Lead	59
5.3.2	Test Cases	61
5.4	Current Operators	64
5.4.1	Effective-Mass	65
5.4.2	Tight-Binding	66
5.4.3	Current Operator Summary	70
5.5	Alternate Current Expression	72
5.6	Solution Procedure	74
5.7	Results and Discussion	78
6	Conclusions	87
6.1	Current Progress	87
6.2	Future Work	90
	Bibliography	92

A	Solution to Laplace's Equation	107
A.1	Potential Applied to the Curved Surface	108
A.2	Potential Applied to One End	110
A.3	Total Solution	112
B	Green's Function for a Cylindrical Box	113
C	Conformal Mapping of the Cylindrical Laplace Equation	117
C.1	Transformation of the Cartesian Laplace Equation	118
C.2	Transformation of the Cylindrical Laplace Equation	120
D	The WKB Approximation	122
D.1	Approximate Solution	123
D.2	Error	123
D.3	Turning Points and Matching	124
E	Solving for the Newton Update	129
E.1	Computing the Raw Newton Update	129
E.2	Using Exact Line Searches	130

List of Figures

1.1	Schematic graphene lattice	2
2.1	Dispersion relation and density of states for a (16,0) carbon nanotube	12
3.1	Schematics of the closed and open cylindrical CNFET geometries	13
3.2	Equilibrium energy bands near the source for a 5.6 nm gate-oxide thickness	19
3.3	Equilibrium energy bands near the source for a 30.9 nm gate-oxide thickness	19
3.4	Equilibrium energy bands with a permittivity five times higher than in Fig. 3.2	20
4.1	Model energy band diagram	24
4.2	A parabolic fit to the (16,0) CN band structure	27
4.3	1D quantum capacitance at equilibrium	32
4.4	1D quantum capacitance for applied bias voltages	35
4.5	Planar geometry CNFET	36
4.6	1D piecewise linear basis function	38
4.7	Conduction band edges for $V_{GS} = 0.5$ V, and $V_{DS} = 0$ and 0.4 V	40
4.8	Conduction band edges and transmission probabilities	41
4.9	Carrier concentration as a function of energy	42
4.10	\mathcal{I}_D - V_{GS} for a closed cylindrical device	42
4.11	Schematic showing the ambipolar conduction mechanism	43
4.12	Circuit for computing the delay time	44
4.13	Computation of the gate capacitance	45
4.14	Cross-section of the planar geometry CNFET	46
4.15	\mathcal{I}_D - V_{GS} for three planar devices	47

4.16	Intrinsic switching time, and ON/OFF ratio	48
4.17	CN charge as a function of the gate voltage	49
4.18	Valence bands for two oxide thicknesses	50
4.19	Extrinsic switching time, and ON/OFF ratio	50
5.1	Tight-binding density of states using the self-energy method	61
5.2	Local density of states for longitudinal and azimuthal potential variation	63
5.3	Sample one- and two-dimensional lattices	67
5.4	Hexagonal lattice with a primitive unit cell	69
5.5	Atomic structure of a (2, 0) carbon nanotube	72
5.6	Schematics of the double-gate and semi-cylindrical-gate CNFET geometries	78
5.7	Local CN potential for the double-gate and the semi-cylindrical-gate devices	79
5.8	Local CN potential as a function of azimuthal angle	80
5.9	Local density of states and energy spectrum of the electron density	81
5.10	Energy spectrum of the current	82
5.11	\mathcal{I}_D - V_{GS} for double-gate and semi-cylindrical-gate devices	83
5.12	Energy band diagrams for the semi-cylindrical-gate geometry	84
5.13	Inter-atomic current for a double-gate device	85
5.14	Comparison of NEGF results to those using the Flietner dispersion relation	86

Acknowledgements

I have been very fortunate to work with an outstanding group over the course of my studies. In particular, I am grateful for early discussions with Paul Pereira and Jason Clifford, and more recently, with Dylan McGuire. I have also benefited from many detailed discussions with Leonardo Castro, as we simultaneously pursued a deeper understanding of carbon nanotube devices. Finally, it has been a privilege to work with my mentor, Professor David Pulfrey, whose guidance, expertise, and insight have been vital to this research.

Dedication

This work would not have been possible without the love and support of my family. My children, Alicia, Emlyn, and Arden, are constant sources of wonder and inspiration. In addition, I owe everything to my wife, Kerrene, who has contributed to this work in more ways than I can list here. She is an inspiration in all of the aspects of my life, and this thesis is as much a result of her hard work and dedication as it is mine. I dedicate this work, and much more, to them.

Chapter 1

Introduction

The aggressive scaling of the metal-oxide-semiconductor field-effect transistor (MOSFET) to ever-smaller physical dimensions has required the introduction of a wide variety of technologies in order to fabricate these devices [1]. With each generation, however, the manufacturing challenges are becoming increasingly difficult, and according to the International Technology Roadmap for Semiconductors [1], intensive research is needed in order to continue this process, and indeed, to develop novel devices and methods that will move the technology improvements in other directions.

A candidate transistor that may allow for both the shrinking process to continue, and for the development of novel architectures, is the carbon nanotube field-effect transistor (CNFET). Essentially, this device is a conventional MOSFET, except that its channel is made up of a carbon nanotube (CN), and that the source and drain interfaces may be Schottky contacts. In the following section, a brief background to CNs is provided, with reference to some of their interesting properties.

1.1 Background

In 1976, the growth of cylindrical tubes of carbon was reported by Oberlin *et al.* [2], but they were not recognized as having the CN structure, and were not studied in detail. Intensive interest in CNs did not begin until 1991 when Iijima noticed their growth on an arc discharge electrode [3], and used electron microscopy to probe their atomic structure. Since that time, a great deal of theoretical and experimental work has gone into understanding carbon nanotubes.

To begin, consider the CN atomic structure. Conceptually, it may be thought of as a graphene sheet that has been rolled into a seamless cylinder. Shown in Fig. 1.1 is a schematic of the hexagonal lattice of carbon atoms that form the atomic monolayer of graphene. The two primitive translation vectors, \mathbf{a}_1 and \mathbf{a}_2 , point to equivalent sites in the lattice, *i.e.*, points that differ by integer multiples of these two vectors are identical in an infinite graphene sheet. Considering two nearest-neighbour

atoms, primitive translations may be performed in order to fill the entire graphene sheet with no overlaps or gaps. As a result, graphene may be defined by simply specifying the primitive translation vectors and the two-atom basis.

In order to form a seamless cylinder, a patch is taken out of the infinite sheet, and rolled such that the circumference is defined by the chiral vector $\mathbf{L} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2$, where n_1 and n_2 are integers. This maps an atom in the lattice to coincide with an equivalent point specified by \mathbf{L} .

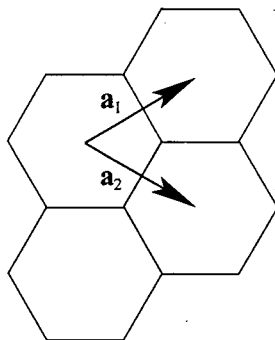


Figure 1.1: Schematic graphene lattice showing the primitive translation vectors.

Carbon nanotubes are uniquely defined, then, by simply specifying the (n_1, n_2) indices. Note that, in Cartesian coordinates,

$$\mathbf{a}_1 = \frac{3a}{2}\hat{\mathbf{x}} + \frac{a\sqrt{3}}{2}\hat{\mathbf{y}}, \quad (1.1)$$

$$\mathbf{a}_2 = \frac{3a}{2}\hat{\mathbf{x}} - \frac{a\sqrt{3}}{2}\hat{\mathbf{y}}, \quad (1.2)$$

where a is the nearest-neighbour carbon-carbon spacing of about 1.4 \AA [4], and $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are unit vectors in the x - and y -directions, respectively. In terms of some basic terminology, CNs are named “armchair” if $n_1 = n_2$, “zigzag” if $n_2 = 0$, and “chiral” otherwise. The armchair and zigzag names relate to the carbon-carbon bond pattern that may be observed when looking along the circumference of the tube.

An interesting property of CNs is that they may be metallic or semiconducting depending on the choice of (n_1, n_2) . Specifically, the tube is metallic if $n_1 - n_2$ is a multiple of 3, and is semiconducting otherwise with a bandgap, E_G , given approximately by [4–6]

$$E_G \simeq \frac{|\gamma|a}{R_C}, \quad (1.3)$$

where γ is an overlap integral which will be defined in Chapter 2, and R_C is the CN radius. Since semiconducting tubes are possible, this raises the prospect of using CNs in order to make semiconductor devices, such as transistors. Indeed, this was performed in 1998 by Tans *et al.* [7], and this early device stimulated a great deal of interest in producing, and simulating, reliable CNFETs.

Perhaps one of the most appealing features of CNs, for nanoelectronics, is their near-ballistic transport due to a limited carrier-phonon interaction [8]. Mean-free-paths for acoustic phonons have been reported in the range of 300 nm to greater than $2.9 \mu\text{m}$ [9–12], and for optical phonons the range is typically in the 10–100 nm range [12, 13]. In Ref. [14], it is asserted that transport should be ballistic in CNs up to 60 nm in length, while Ref. [10] gives a value of 20 nm for a CN in a low field. Mobilities of 10^4 – $10^5 \text{ cm}^2/\text{Vs}$ have been reported [9, 13], and a recent theoretical prediction also falls within this range at low field [15]. When compared to the bulk silicon values of around 10^2 and $10^3 \text{ cm}^2/\text{Vs}$ for holes and electrons [16], respectively, the CNFET appears to hold some promise. For silicon nanowire transistors, the comparison favours CNs even more heavily since the average mobility is in the range 30 – $560 \text{ cm}^2/\text{Vs}$, with a maximum value of $1350 \text{ cm}^2/\text{Vs}$ [17]. Couple this with the large current densities that CNs can withstand [18, 19], roughly three orders of magnitude greater than that reported for silicon nanowires [20], and the carbon nanotube looks even more promising.

In addition, with their cylindrical shape, CNs may lend themselves to the realization of a coaxial field-effect transistor, which is predicted to be the ultimate geometry in terms of good short-channel behaviour [21–23]. This structure is being pursued experimentally [24], and has been successfully made using vertical nanowires [25], but not yet with carbon nanotubes. High permittivity dielectrics are also a possibility towards reducing short-channel effects due to the increased capacitive coupling between the channel and the gate, and success has been achieved using ZrO_2 [26] and SrTiO_3 [27], with relative permittivities of 25 and 175, respectively. High permittivities have also been achieved using electrolytes [28, 29], where the relative permittivity of the solution could be around 80 [30], although liquid electrolytes are likely to have limited utility in highly integrated circuits.

Furthermore, there may be biological applications of CN devices [31], for example, in ultrafast DNA sequencing or as a DNA-incorporated switch [32]. Structures are also being investigated for use as chemical sensors [33, 34]. Additionally, with the unique electrical properties of CNs, one could envision all-carbon electronics with metallic nanotube leads connected to semiconducting nanotube devices [35, 36].

Given these properties, it is clear why there is interest in studying carbon nanotubes and carbon nanotube electronics. There are, however, some unresolved problems in CNFET fabrication. Early devices were fabricated by dispersing CNs from CN ropes, and then positioning them using an atomic force microscope tip [37]. While this method has been largely replaced by the more reliable chemical vapour deposition (CVD) technique [38,39], the major obstacle of chirality control remains an outstanding problem. Chirality control is important as this determines whether the CN is metallic or semiconducting, and if semiconducting, it also determines the energy bandgap.

Since chirality control is difficult, another option is to devise methods to separate CNs based on chirality or conductivity. This may be facilitated by the presence of single-stranded DNA [31]. Sorting of CNs by diameter [40], by metals and semiconductors [41], and by chirality [42,43], have been reported, and recent work has shown that semiconducting tubes may be preferentially grown in the plasma enhanced CVD method [44], with a yield of around 90% semiconducting CNs. With this latter technique, perhaps it would be sufficient to control only the diameter, and hence the bandgap. This may prove to be a more attainable goal.

1.2 Simulation Studies

In order to examine the expected performance, and performance limits, of CNFETs, detailed simulations of these devices are presented. In Chapter 2, the band structure and density of states are computed. This serves as the foundation for this work. Chapters 3–5 present a range of CNFET models and results: beginning with a simple case, invoking approximations that reduce the problem complexity, and then successively stripping away some of those simplifying assumptions such that each subsequent model captures more of the device physics.

In particular, Chapter 3 considers the equilibrium case of devices with Schottky-barrier contacts at the source and drain [45–50]. This is the simplest case to consider under the assumption that potential energy variations are sufficiently gradual to allow for the usual Fermi-Dirac statistics to apply. As will be seen in Chapter 4, rapid potential variations may lead to resonance effects which serve to alter the distribution function. These equilibrium simulations resulted in the development of a compact model for CNFETs that was an early attempt at including the effect of tunneling through the Schottky barriers [51].

For steep potential energy variations, and for the non-equilibrium case, Chapter 4 presents a self-

consistent Schrödinger-Poisson solver, which uses the effective-mass approximation to the CN band structure. At the time of this work, CNFETs were being modeled using electrostatics appropriate for an infinite CN, and under the assumption that transport was limited by thermionic emission over the top of a potential energy barrier [26,52]. Here, a modified WKB expression is contributed that is appropriate for use in compact models [53], and that accounts for the reflection of carriers above a potential energy step, such as in a metal-semiconductor contact where the Fermi wavevector changes abruptly upon crossing the interface [54]. It transpired that the new modified WKB expression was very successful in reconciling the compact model results with those of the, more detailed, full effective-mass model [53]. While a cylindrical structure is required in order for this method to be strictly valid, it may also be reasonable for less symmetric structures if the azimuthal potential energy variation is small enough [55].

In order to examine such variation, which had apparently been neglected in order to speed up similar calculations [56], and to lift the effective-mass approximation, Chapter 5 details a non-equilibrium Green's function method for computing the local density of states from Schrödinger's equation. This is, again, coupled with Poisson's equation in a self-consistent fashion, but here an atomistic, nearest-neighbour tight-binding Hamiltonian, as opposed to the previously employed effective-mass variety, is used. This method was developed independently of that due to Fiori *et al.*, despite the similarities in technique [57,58]. One key difference is in the computation of the self-energy matrices, where a different algorithm is presented herein. In addition, expressions for the inter-atomic current are derived, and these may prove useful when studying CNs with defects. The focus in this chapter is on doped-contact devices without Schottky-barriers at the end contacts, as may be appropriate for some devices [59–62]. The *p-i-n* case [63–65] is discussed as it may exhibit less than a 60 mV/dec subthreshold slope, which is the limit for devices governed by the thermionic emission of carriers over a barrier. A lower subthreshold slope would be a tremendous boost for highly integrated logic circuitry, which is presently curtailed by the static power dissipation from supposedly “OFF” gates.

Chapter 6 presents a summary of contributions, key conclusions to be drawn from this work, and outlines future work that may be built from the solid foundation presented herein. In addition, key derivations are included in the appendices in order to facilitate the reproduction, and extension, of this work by other investigators.

Chapter 2

Band Structure

In order to simulate a CNFET, the band structure of a carbon nanotube must first be determined. To this end, the nearest-neighbour, tight-binding approximation is applied. From the band structure, the density of states (DOS) may be calculated, and this is required in order to compute the amount of charge on the tube. This fundamental calculation is a basic tool, and is an important first step to take in understanding the behaviour of CNFETs.

2.1 Graphene

Consider, first, the band structure of graphene. The overall periodicity of the lattice is defined by a two-point basis which may be shifted by integer multiples of the primitive translation vectors in order to tile, with no overlaps or gaps, an infinite graphene crystal. The electron wavefunctions are governed by the time-independent Schrödinger equation,

$$H\Psi(\mathbf{r}) = E\Psi(\mathbf{r}), \quad (2.1)$$

where H is the Hamiltonian, Ψ is the wavefunction, \mathbf{r} is position, and E is energy. Following the usual tight-binding procedure [66], the Hamiltonian may be decomposed into

$$H = H_{\text{at}} + \Delta U(\mathbf{r}),$$

where H_{at} is the Hamiltonian for an isolated atom, and $\Delta U(\mathbf{r})$ represents the modification to this Hamiltonian due to the presence of all of the other atoms. The atomic Hamiltonian has a complete set of eigenfunctions, $\Psi_{\text{at},i}$, with their associated eigenenergies, E_i , *i.e.*, $H_{\text{at}}\Psi_{\text{at},i} = E_i\Psi_{\text{at},i}$. Further, due to the underlying periodicity of the graphene lattice, Bloch's theorem [66] may be used to write the total wavefunction as

$$\Psi(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} [B_1\Xi(\mathbf{r} - \mathbf{R}) + B_2\Xi(\mathbf{r} - \mathbf{d} - \mathbf{R})],$$

where the summation is over all reciprocal lattice vectors \mathbf{R} , \mathbf{k} is the wavevector, \mathbf{d} is a vector from one atom in the two-point basis to the other, B_1 and B_2 are constants, and Ξ is known as a Wannier function [66].

Since the $\Psi_{\text{at},i}$ are a complete set, Ξ may be written as

$$\Xi(\mathbf{r}) = \sum_i C_i \Psi_{\text{at},i}(\mathbf{r}),$$

where C_i are constants. Up until this point, the expressions have been exact. The first approximation is that the Wannier functions are well-described by a single choice of eigenfunction. Namely, let $C_i = 1$ for the p_z orbitals, defined by $i = 1$, and $C_i = 0$ otherwise.

Multiplying through by $\Psi_{\text{at},j}^\dagger(\mathbf{r})$ and integrating allows Eq. (2.1) to be written as

$$\int \Psi_{\text{at},j}^\dagger(\mathbf{r}) H_{\text{at}} \Psi(\mathbf{r}) d^3r + \int \Psi_{\text{at},j}^\dagger(\mathbf{r}) \Delta U(\mathbf{r}) \Psi(\mathbf{r}) d^3r = \int \Psi_{\text{at},j}^\dagger(\mathbf{r}) E \Psi(\mathbf{r}) d^3r, \quad (2.2)$$

where the integral is over all space, $\Psi_{\text{at},j}$ is an arbitrary eigenfunction, and \dagger denotes the Hermitian conjugate. Note, here, that E is a scalar, so it may simply come outside the integral. In addition, since the atomic Hamiltonian operator is Hermitian, $\Psi_{\text{at},j}^\dagger(\mathbf{r}) H_{\text{at}} = (H_{\text{at}} \Psi_{\text{at},j}(\mathbf{r}))^\dagger$. Substituting in the above definitions yields

$$\begin{aligned} & (E - E_j) \int \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} \Psi_{\text{at},j}^\dagger(\mathbf{r}) [B_1 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{R}) + B_2 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{d} - \mathbf{R})] d^3r \\ &= \int \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} \Psi_{\text{at},j}^\dagger(\mathbf{r}) \Delta U(\mathbf{r}) [B_1 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{R}) + B_2 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{d} - \mathbf{R})] d^3r. \end{aligned}$$

Considering the case where $j = 1$, and rearranging, gives

$$\begin{aligned} & (E - E_1) \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) [B_1 \Psi_{\text{at},1}(\mathbf{r}) + B_2 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{d})] d^3r \\ & - \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) \Delta U(\mathbf{r}) [B_1 \Psi_{\text{at},1}(\mathbf{r}) + B_2 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{d})] d^3r \\ & + \sum_{\mathbf{R} \neq 0} e^{i\mathbf{k} \cdot \mathbf{R}} \left\{ (E - E_1) \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) [B_1 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{R}) + B_2 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{d} - \mathbf{R})] d^3r \right. \\ & \quad \left. - \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) \Delta U(\mathbf{r}) [B_1 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{R}) + B_2 \Psi_{\text{at},1}(\mathbf{r} - \mathbf{d} - \mathbf{R})] d^3r \right\} = 0. \end{aligned}$$

At this point, a second approximation is invoked. Namely, assume that the wavefunction decays sufficiently fast that any term involving a product of wavefunctions, centred at positions that differ by more than the nearest-neighbour spacing, vanishes. Making a change of variables such that $\mathbf{c} = -\mathbf{d} - \mathbf{R}$, and utilizing the nearest-neighbour approximation, yields

$$\{E - E_1 + \beta\} B_1 + \left\{ \sum_{\mathbf{c} \in \text{nn}} e^{-i\mathbf{k} \cdot (\mathbf{c} + \mathbf{d})} [(E - E_1) \alpha(\mathbf{c}) - \gamma(\mathbf{c})] \right\} B_2 = 0,$$

where $\mathbf{c} \in \text{nn}$ implies that the summation is over all of the nearest neighbours,

$$\begin{aligned}\beta &= - \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) \Delta U(\mathbf{r}) \Psi_{\text{at},1}(\mathbf{r}) d^3r, \\ \alpha(\mathbf{c}) &= \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) \Psi_{\text{at},1}(\mathbf{r} + \mathbf{c}) d^3r, \\ \gamma(\mathbf{c}) &= \int \Psi_{\text{at},1}^\dagger(\mathbf{r}) \Delta U(\mathbf{r}) \Psi_{\text{at},1}(\mathbf{r} + \mathbf{c}) d^3r,\end{aligned}$$

and it is assumed that the $\Psi_{\text{at},i}$ are orthonormal.

If the same procedure is repeated, replacing Eq. (2.2) with

$$\int \Psi_{\text{at},j}^\dagger(\mathbf{r} - \mathbf{d}) H_{\text{at}} \Psi(\mathbf{r}) d^3r + \int \Psi_{\text{at},j}^\dagger(\mathbf{r} - \mathbf{d}) \Delta U(\mathbf{r}) \Psi(\mathbf{r}) d^3r = \int \Psi_{\text{at},j}^\dagger(\mathbf{r} - \mathbf{d}) E \Psi(\mathbf{r}) d^3r,$$

a second equation in terms of B_1 and B_2 emerges. In matrix form, where the arguments from α and γ have been dropped since their values do not depend on which neighbour we are considering, the equations may be summarized as

$$\begin{bmatrix} E - E_1 + \beta & \sum_{\mathbf{c} \in \text{nn}} e^{-i\mathbf{k} \cdot (\mathbf{c} + \mathbf{d})} [(E - E_1)\alpha - \gamma] \\ \sum_{\mathbf{c} \in \text{nn}} e^{i\mathbf{k} \cdot (\mathbf{c} + \mathbf{d})} [(E - E_1)\alpha - \gamma] & E - E_1 + \beta \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In order for this equation to have a non-trivial solution, the determinant of the 2×2 matrix must be set to zero. Solving yields

$$E = E_1 + \left(\frac{-\beta \pm \gamma \left| \sum_{\mathbf{c} \in \text{nn}} e^{i\mathbf{k} \cdot \mathbf{c}} \right|}{1 \pm \alpha \left| \sum_{\mathbf{c} \in \text{nn}} e^{i\mathbf{k} \cdot \mathbf{c}} \right|} \right),$$

where the plus and minus signs go together to give the bonding and antibonding π bands, respectively. For graphene, $\gamma = -3.033 \text{ eV}$ and $\alpha = 0.129$ [4]. β is of lesser importance since it does not multiply any function of \mathbf{k} . In addition, it is convenient to reference energies to $E_1 - \beta$, and neglect the α correction to the β term since both β and α are small. In fact, it is very common for α to be completely neglected. This is referred to as the Slater-Koster scheme [4], and this approximation will be used throughout the remainder of this work. Note that γ and α are not, in general, exact, and considerable variation may be found in these numbers. For example, a γ value of -2.5 eV has been reported in Ref. [5]; for carbon nanotubes, Ref. [67] suggests that the appropriate value is $-2.7 \pm 0.1 \text{ eV}$ in order to get good agreement between experimental bandgap measurements and the theoretical predictions. In the end, γ is essentially a fitting parameter to the models.

2.2 Zone-Folding

For a carbon nanotube, the “zone-folding” scheme is utilized to compute the dispersion, or E - k , relation from that of graphene [6]. In the Slater-Koster scheme, and referencing energies as discussed above, the tight-binding Hamiltonian may be written as

$$H = \begin{bmatrix} 0 & \gamma \sum_{\mathbf{c} \in \text{nn}} e^{i\mathbf{k} \cdot \mathbf{c}} \\ \gamma \sum_{\mathbf{c} \in \text{nn}} e^{-i\mathbf{k} \cdot \mathbf{c}} & 0 \end{bmatrix}, \quad (2.3)$$

where the eigenvalues of this matrix, as functions of \mathbf{k} , are the eigenenergies, E . It is convenient to work in a coordinate system where one axis is aligned with the CN axis, and the other is in the circumferential direction. If the standard xy -plane is oriented such that the primitive translation vectors, mapping to equivalent sites in the graphene lattice, are given by Eqs. (1.1) and (1.2), then the vector \mathbf{L} , wrapping around the circumference, is given by

$$\mathbf{L} = \frac{3a}{2} (n_1 + n_2) \hat{\mathbf{x}} + \frac{a\sqrt{3}}{2} (n_1 - n_2) \hat{\mathbf{y}} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2.$$

Similarly, the vector \mathbf{T} , which points to an equivalent site along the axis of the CN, including the helical angle, is given by [6]

$$\mathbf{T} = \frac{3a}{2} \left(\frac{n_2 - n_1}{d_R} \right) \hat{\mathbf{x}} + \frac{3a\sqrt{3}}{2} \left(\frac{n_1 + n_2}{d_R} \right) \hat{\mathbf{y}} = \frac{2n_2 + n_1}{d_R} \mathbf{a}_1 - \frac{2n_1 + n_2}{d_R} \mathbf{a}_2,$$

where d_R is the greatest common divisor of $2n_1 + n_2$ and $2n_2 + n_1$. It is useful to note that rotations from the xy -plane to the LT -plane may be performed via

$$\begin{bmatrix} \frac{\sqrt{3}}{2\eta} (n_1 + n_2) & \frac{n_1 - n_2}{2\eta} \\ \frac{n_2 - n_1}{2\eta} & \frac{\sqrt{3}}{2\eta} (n_1 + n_2) \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} v_L \\ v_T \end{bmatrix},$$

where v_i denotes the i -component of some arbitrary vector \mathbf{v} , and $\eta = \sqrt{n_1^2 + n_1 n_2 + n_2^2}$.

Note that \mathbf{L} and \mathbf{T} span the real-space primitive unit cell of the carbon nanotube, *i.e.*, an entire CN may be tiled, with no gaps or overlaps, by translating this cell through integer multiples of these two vectors. Since this cell is bigger than that of graphene, the tight-binding derivation detailed above would involve a larger Hamiltonian matrix than in Eq. (2.3), with one row for each atom in this enlarged cell. Here, the zone-folding scheme is employed in order to calculate the CN band

structure directly from the two-dimensional (2D) result for graphene. This scheme provides the identical result to that generated from treating each atom in the CN primitive unit cell, although the full tight-binding approach lends itself to performing the simulations discussed in Chapter 5.

Following the description in Ref. [4], the number of primitive graphene unit cells that fit inside the enlarged CN primitive unit cell is

$$N = \frac{|\mathbf{L} \times \mathbf{T}|}{|\mathbf{a}_1 \times \mathbf{a}_2|},$$

which means that there are $2N$ carbon atoms in the CN primitive unit cell.

The reciprocal lattice vectors of graphene, \mathbf{b}_1 and \mathbf{b}_2 , are computed by utilizing the defining relation

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

where δ_{ij} is the Kronecker delta function. The result is that

$$\begin{aligned}\mathbf{b}_1 &= \frac{2\pi}{a\sqrt{3}} \left(\frac{1}{\sqrt{3}}\hat{x} + \hat{y} \right), \\ \mathbf{b}_2 &= \frac{2\pi}{a\sqrt{3}} \left(\frac{1}{\sqrt{3}}\hat{x} - \hat{y} \right).\end{aligned}$$

Performing the analogous procedure for the reciprocal lattice vectors of the CN, \mathbf{K}_T and \mathbf{K}_L , yields

$$\begin{aligned}\mathbf{K}_T &= \frac{1}{N} (n_2\mathbf{b}_1 - n_1\mathbf{b}_2), \\ \mathbf{K}_L &= \frac{1}{N} \left(\frac{2n_1 + n_2}{d_R}\mathbf{b}_1 + \frac{2n_2 + n_1}{d_R}\mathbf{b}_2 \right).\end{aligned}$$

As in Ref. [4], note that $N\mathbf{K}_L$ is a reciprocal lattice vector of graphene, as may be seen based on the definition of d_R . For the CN, $j\mathbf{K}_L$ are all equivalent, where j is an integer. In addition, for $j \in \{0, 1, \dots, N-1\}$, $j\mathbf{K}_L$ are all reciprocal lattice vectors of the CN, but not of graphene. As a result, these discrete $j\mathbf{K}_L$ in the graphene E - k relation may simply be used for the CN dispersion relation. Hence, one-dimensional (1D) projections of the graphene relation are “folded” into the first Brillouin zone of the CN. For further details, see Ref. [4].

Specifically, if $\mathbf{k} = k_T\hat{\mathbf{T}} + k_L\hat{\mathbf{L}}$ in Eq. (2.3), then

$$E = \pm\gamma \left| \sum_{\mathbf{c} \in \text{nn}} e^{i\mathbf{k} \cdot \mathbf{c}} \right|,$$

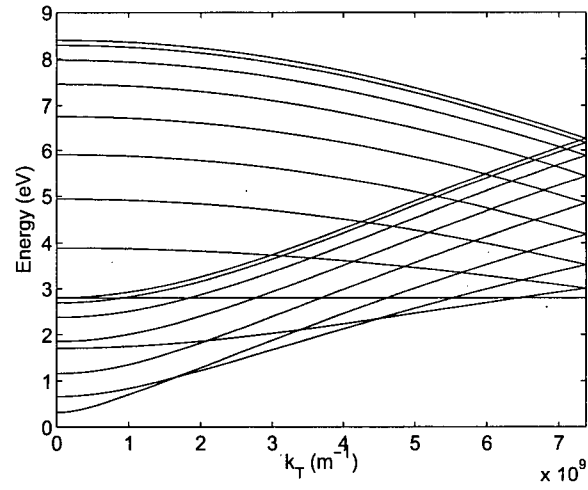
where $k_L\hat{\mathbf{L}} = j\mathbf{K}_L$ and

$$k_T \in \left[-\frac{\pi}{|\mathbf{T}|}, \frac{\pi}{|\mathbf{T}|} \right].$$

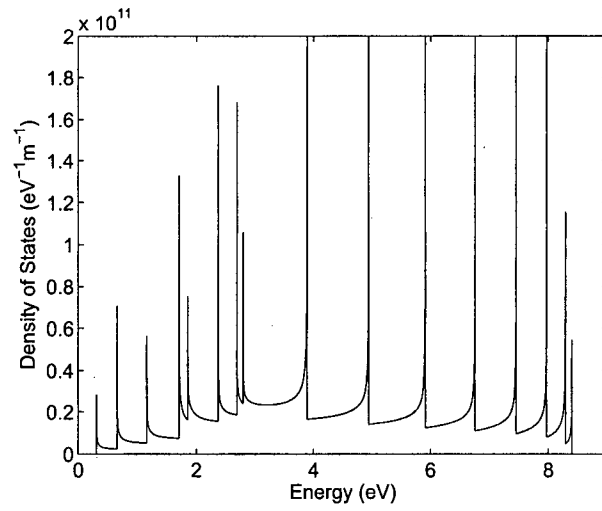
The DOS, $D(E)$, may then be computed numerically as

$$D(E) = \frac{2}{\pi} \sum_{j=0}^{N-1} \left| \frac{\partial E^{(j)}}{\partial k_T} \right|^{-1},$$

where spin degeneracy is included, and $E^{(j)}$ is the energy associated with $k_L \hat{\mathbf{L}} = j\mathbf{K}_L$. The E - k relation and DOS for a (16,0) CN are shown in Fig. 2.1. Each band in Fig. 2.1(a) corresponds to a different value of j , and some of these bands are multiply degenerate, *i.e.*, different values of j produce the same E - k_T relationship. In Fig. 2.1(b), note, too, the singularities at the band edges, as expected for a 1D DOS.



(a)



(b)

Figure 2.1: (a) Dispersion relation and (b) density of states for a (16,0) carbon nanotube.

Chapter 3

Equilibrium

The simplest case to consider for CNFETs is that of equilibrium. The goal is to compute the charge on the CN as a function of the applied potentials on the source, drain, and gate electrodes. From this relationship, the energy band diagram may be computed in order to gain some insight into CNFET operation. A Schottky-barrier device is considered, with metallic source and drain electrodes that contact an intrinsic, semiconducting nanotube. This work is the foundation for many subsequent efforts into non-equilibrium models [51, 53, 55, 68–76], and was initially used to study the closed [77] and open [71] cylindrical structures depicted in Fig. 3.1. The gate is translucent in order to view the inner structure of the device, and the source and drain cap either end of the inner CN cylinder. These coaxial structures are considered due to their enhanced Schottky-barrier modulation [22, 78], and in order to investigate the likely performance limits of these devices, particularly with respect to their short-channel behaviour [21, 23, 52, 70, 79].

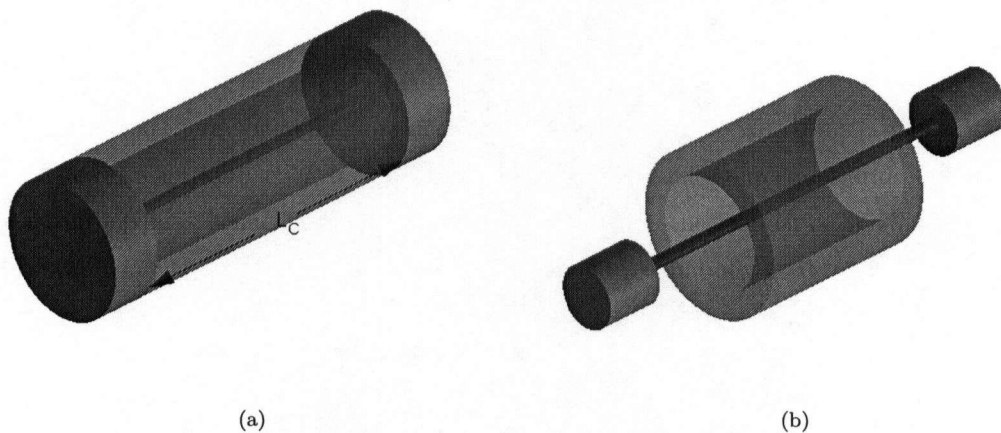


Figure 3.1: Schematics of the (a) closed and (b) open cylindrical CNFET geometries.

3.1 Electrostatics

Consider, first, the electrostatics problem. Poisson's equation is

$$\nabla^2 V = -\frac{q}{\epsilon} (p - n + N_f),$$

where V is the local electrostatic potential, q is the magnitude of the electronic charge, p is the hole concentration, n is the electron concentration, N_f is the concentration of fixed charge, and ϵ is the permittivity of the material. The boundary conditions, referencing applied potentials to the source, are

$$\begin{aligned} V_S &= -\frac{\Phi_S}{q}, \\ V_D &= V_{DS} - \frac{\Phi_D}{q}, \\ V_G &= V_{GS} - \frac{\Phi_G}{q}, \end{aligned}$$

where V_S , V_D and V_G are the boundary conditions on the source, drain, and gate electrodes, respectively, V_{DS} and V_{GS} are the applied drain-source and gate-source potential differences, respectively, and Φ_S , Φ_D and Φ_G are the workfunctions of the source, drain and gate metallizations, respectively. The boundary conditions on the source and drain are appropriate in the absence of Fermi-level pinning [80]. Note that $V_{DS} = 0$ in equilibrium.

The fixed charge is an input to the model, and will be set to zero unless otherwise stated. The electron and hole concentrations are computed according to the model being employed, and this calculation will be detailed when each model is discussed. At this point, it is sufficient to state that the CN charge is assumed to be confined to the surface of the CN in the form,

$$p - n = \delta(\rho - R_C) (p_{2D} - n_{2D}), \quad (3.1)$$

where $\delta(\cdot)$ is the Dirac delta function, ρ is the distance from the CN axis, and p_{2D} and n_{2D} are the two-dimensional carrier densities on the surface of the CN.

3.1.1 Closed Geometry

The simplest geometry to consider for the equilibrium electrostatics of a cylindrically-gated CNFET [77] is one where the source, drain and gate form a closed cylindrical box, as in Fig. 3.1(a). This case

is mathematically convenient since there are only Dirichlet boundary conditions on a finite domain. In this case, it is natural to work in cylindrical coordinates with

$$V(\rho, \phi, z) = V(\rho, \phi + 2\pi, z),$$

and finite $V(0, \phi, z)$, where ϕ is the azimuthal angle, and z is the distance along the CN axis. If p_{2D} and n_{2D} are not functions of ϕ , then V will also be independent of ϕ .

If the permittivity is constant throughout the device, it is convenient to solve Poisson's equation using superposition. The first problem is that of Laplace subject to the boundary conditions described above. This solution is analytic, and is derived in Appendix A. The solution to Poisson's equation subject to homogeneous boundary conditions may be added to the Laplace result in order to obtain the full solution. The Poisson problem is solved via the Green's function, which is derived in Appendix B, and the electrostatics problem is reduced to one of numerically integrating the Green's function multiplied by the charge distribution on the surface of the CN.

If the permittivity is not constant throughout the device, numerical methods are required. Typically, it is assumed that a material of constant permittivity surrounds the CN, and that the interior of the CN itself has unity permittivity [81]. Due to the permittivity discontinuity, there exists a matching condition at the CN surface [82]:

$$\epsilon_{\text{ox}} \nabla V|_{R_C^+} \cdot \hat{\mathbf{n}} - \epsilon_C \nabla V|_{R_C^-} \cdot \hat{\mathbf{n}} = -q \frac{p_{2D} - n_{2D}}{\epsilon_0}, \quad (3.2)$$

where ϵ_{ox} and ϵ_C are the relative permittivities outside and inside the CN, respectively, ϵ_0 is the permittivity of free space, R_C^+ and R_C^- are infinitesimally outside and inside the CN, respectively, and $\hat{\mathbf{n}}$ is the unit outward normal.

In cylindrical coordinates, and assuming no ϕ -dependence in p_{2D} and n_{2D} ,

$$\frac{\partial^2 V}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial V}{\partial \rho} + \frac{\partial^2 V}{\partial z^2} = 0 \quad (3.3)$$

is solved in two regions, $\rho > R_C$ and $\rho < R_C$, subject to Eq. (3.2) across $\rho = R_C$.

3.1.2 Open Geometry

For the physical case, where gaps will be present between the end contacts and the gate, as shown in Fig. 3.1(b), an analytical solution cannot be derived. One option to deal with this case is to simply

apply homogeneous Neumann boundary conditions in the gaps sufficiently far away [83] from the surfaces of interest, *i.e.*, away from the CN surface. If cylindrical symmetry is present, however, an alternate solution may be employed [71]. The 2D electrostatics may be solved by applying a conformal mapping to the far-field region [84]. This yields two domains which must be solved simultaneously. The near-field region is a central disk of radius R_B , which encloses the contacts and the CN, and the far-field region is the entire plane outside of this disk.

Since there are no electrodes in the far-field region, the appropriate equation to solve is Eq. (3.3), where the $\rho = R_B$ boundary introduces a matching condition between the two domains, where it is required that V is continuous and smooth. If $\mathcal{U} = z + i\rho$ and $\mathcal{T} = \zeta + i\sigma$, the conformal mapping

$$\mathcal{T} = \frac{R_B^2}{\mathcal{U}}$$

may be used in the far-field region in order to obtain a finite domain of the same size and shape as the near-field domain. Here, \mathcal{U} and \mathcal{T} represent the coordinates in the untransformed and transformed domains, respectively. Applying this transformation to Eq. (3.3) yields

$$(\zeta^2 + \sigma^2) \left(\frac{\partial^2 \bar{V}}{\partial \zeta^2} + \frac{\partial^2 \bar{V}}{\partial \sigma^2} \right) - 2\zeta \frac{\partial \bar{V}}{\partial \zeta} + \left(\frac{\zeta^2 - \sigma^2}{\sigma} \right) \frac{\partial \bar{V}}{\partial \sigma} = 0,$$

where the overbar indicates evaluation in the transformed coordinate system, and the derivation of this expression may be found in Appendix C.

Continuity is easily satisfied on both boundaries since the transformation does not modify the value of V . The smoothness condition is given by [71]

$$\nabla_{\mathcal{U}} V \cdot \hat{\mathbf{n}}_{\mathcal{U}} = -\nabla_{\mathcal{T}} \bar{V} \cdot \hat{\mathbf{n}}_{\mathcal{T}},$$

where $\nabla_{\mathcal{U}}$ and $\nabla_{\mathcal{T}}$ are the gradient operators in the untransformed and transformed domains, respectively, and $\hat{\mathbf{n}}_{\mathcal{U}}$ and $\hat{\mathbf{n}}_{\mathcal{T}}$ are the unit outward normals to the corresponding domains.

By posing the open boundary electrostatics problem as two coupled closed boundary problems, Poisson's equation has been cast in a form that lends itself to the finite element technique for solution [71].

3.2 Solution Procedure

Whichever type of cylindrically-symmetric geometry is considered, p_{2D} and n_{2D} must be defined. In equilibrium, this is a straightforward matter if it is assumed that V rigidly shifts the zero-bias band

structure as derived in Chapter 2. Due to the symmetry of the conduction and valence bands in the Slater-Koster scheme,

$$p_{2D} - n_{2D} = \frac{1}{2\pi R_C} \int_0^\infty D(E) [f(E + qV + \Phi_C) - f(E - qV - \Phi_C)] dE, \quad (3.4)$$

where Φ_C is the workfunction of the CN, and

$$f(E) = \frac{1}{\exp\left(\frac{E - E_F}{k_B T_K}\right) + 1}$$

is the Fermi-Dirac distribution function, where E_F is the Fermi level, k_B is Boltzmann's constant, T_K is temperature, and $E = 0$ corresponds to the lowest band edge in the density of states.

Initially, assume that there is no permittivity discontinuity, and consider the closed geometry. In this case, the electrostatics problem may be solved using the results derived in Appendices A and B. The solution in this case may be summarized as follows:

1. Solve Laplace's equation using Eq. (A.9) from Appendix A to get V_L .
2. Compute the charge via Eq. (3.4), where $V = V^{(i)}$ with the superscript indicating the iteration number, and $V^{(0)} = V_L$.
3. Compute the potential due to the charge $V_C^{(i)}$, via Eqs. (3.1) and (3.4), using a numerical integration scheme, such as Simpson's rule, to evaluate Eq. (B.2) in Appendix B.
4. Check that $V^{(i)} = V_L + V_C^{(i)}$ is consistent with the charge to within some specified tolerance. If not, increment i , and return to Step 2. If it is, convergence has been obtained.

If there is a permittivity discontinuity, or open boundaries, numerical methods are utilized. In Refs. [68, 71, 77], the finite element method was employed, through the use of FEMLAB [69, 85], to solve the nonlinear Poisson equation that results from casting p and n in Eq. (3.1) as functions of V via Eq. (3.4).

3.3 Results and Discussion

Results are presented, from Ref. [77], for the CNFET depicted schematically in Fig. 3.1(a) with a (16, 0) CN with a radius of 0.63 nm, a length of 100 nm, and a gate workfunction of 4.5 eV. The

electron affinity for the carbon nanotube is taken to be 4.2 eV, based on a workfunction of 4.5 eV [7], and an intrinsic-tube bandgap of around 0.6 eV. These values correspond to so-called “positive barrier” devices for electrons since the workfunctions of the source and drain are greater than the electron affinity of the CN. Similarly, “negative barrier” and “zero barrier” devices could be defined. For these positive barrier results, ϵ_C is taken to be the same as that of the gate dielectric. This permits the utilization of the analytic electrostatics solution. The temperature is taken to be 300 K.

At equilibrium, if Φ_S and Φ_D are equal, the energy bands along the tube will be symmetrical; thus, only profiles near one contact need to be shown. Fig. 3.2 shows these diagrams near the source for $R_G/R_C = 10$, where R_G is the gate radius, and data is presented for $\epsilon_{ox} = 3.9$, $\Phi_S = \Phi_D$, $V_{DS} = 0$, and $V_{GS} = 0.2$ and 0.5 V. In Fig. 3.2(a), $\Phi_S = 4.5$ eV, which corresponds to the case of equal workfunctions for the metal and the nanotube, whereas $\Phi_S = 4.33$ eV, as in Fig. 3.2(b), and $\Phi_S = 4.63$ eV, as in Fig. 3.2(c), refer to low- and high-metal workfunctions, respectively [80]. The potential in the body of the tube, away from either of the end contacts, depends directly on V_{GS} , leading to potential spikes near the source and drain of height determined by both $\Phi_{S,D}$ and V_{GS} . In the low- Φ_S case at low V_{GS} , thermionic emission will likely be the most significant contribution to the source current. In all other cases shown in Fig. 3.2, tunneling of electrons through the energy spike may also be important.

The band diagrams for the same workfunction cases as used in Fig. 3.2, but for $R_G/R_C = 50$, are shown in Fig. 3.3. The reduced band bending in the tube at the contacts, due to poorer coupling between the gate and the nanotube, is very evident, and will lead to a dramatic decrease in the tunneling current. Note that the present state of the art as regards gate-dielectric thinness is 2 nm [50].

Regarding the permittivity of the dielectric, Ref. [86] has reported the use of zirconia, for which ϵ_{ox} is around 5 times higher than that used in obtaining the above figures. The effect of such a change in ϵ_{ox} can be seen by comparing Figs. 3.2 and 3.4. At $V_{GS} = 0.5$ V, the increased capacitive coupling between the gate and the tube drives the mid-tube potential energy to lower values, yet does not change significantly the width of the source barrier at its base; thus, obviously, an increased current for a given bias will result from using a higher ϵ_{ox} . At lower V_{GS} , *e.g.*, 0.2 V, the increased ϵ_{ox} makes essentially no difference to the band diagram because, at least for $R_G/R_C = 10$, there is virtually no charge induced on the tube. In this case, the electrostatics are governed by the Laplace solution alone.

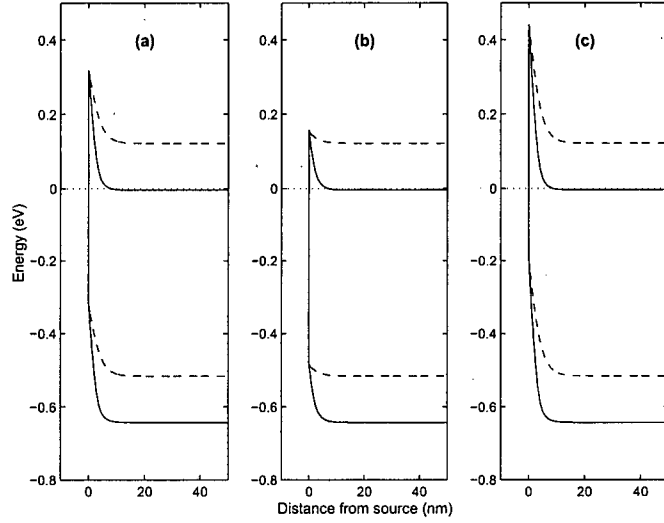


Figure 3.2: Equilibrium energy bands near the source for a 5.6 nm gate-oxide thickness, with $\Phi_S = \Phi_D$ set to: (a) 4.5 eV, (b) 4.33 eV, and (c) 4.63 eV. Shown are $V_{GS} = 0.2$ (dashed) and 0.5 V (solid). Energies are with respect to the Fermi level (dotted) [77].

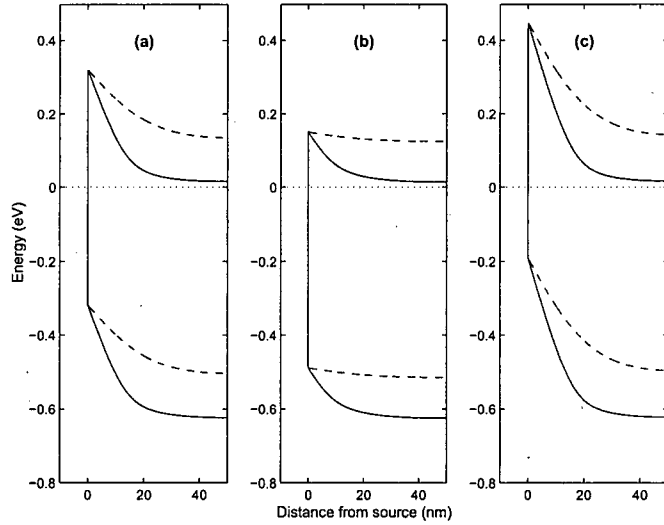


Figure 3.3: Equilibrium energy bands near the source for a 30.9 nm gate-oxide thickness, with $\Phi_S = \Phi_D$ set to: (a) 4.5 eV, (b) 4.33 eV, and (c) 4.63 eV. Shown are $V_{GS} = 0.2$ (dashed) and 0.5 V (solid). Energies are with respect to the Fermi level (dotted) [77].

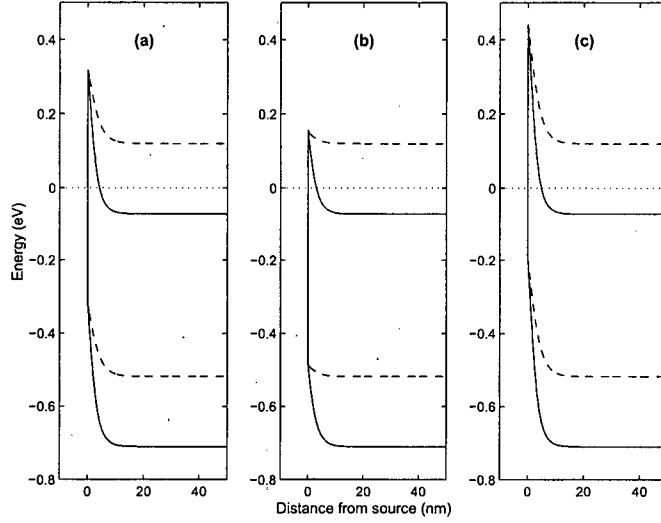


Figure 3.4: Equilibrium energy bands near the source for a 5.6 nm gate-oxide thickness with a permittivity five times higher than in Fig. 3.2, with $\Phi_S = \Phi_D$ set to: (a) 4.5 eV, (b) 4.33 eV, and (c) 4.63 eV. Shown are $V_{GS} = 0.2$ (dashed) and 0.5 V (solid). Energies are with respect to the Fermi level (dotted) [77].

From Figs. 3.2, 3.3 and 3.4, it appears that the width of the potential barrier at its base depends strongly on R_G for the contact geometry considered here, as has been remarked upon elsewhere [87], and here it is indicated that it also has a very weak dependence on V_{GS} . This has been used in the development of a compact model for the non-equilibrium behaviour, where an approximate expression describes the band bending near the contacts [51, 53].

Note that the effect of changing the gate workfunction from the value of 4.5 eV used here can be readily appreciated from the foregoing figures as an increase in Φ_G of 0.1 eV, for example, has the same effect as a corresponding decrease in V_{GS} .

Chapter 4

The Effective-Mass Method

Out of equilibrium, the modeling is complicated by the fact that Fermi distribution functions cannot be used to describe the energy spectrum of carriers in the CN. One solution is to neglect the effect of the CN charge on the electrostatics, and simply use a Laplace solution to define the non-equilibrium band structure [50]. Alternatively, one could employ infinite CN electrostatics, and consider only thermionically emitted carriers [52]. This approach is facilitated by the use of the so-called “quantum capacitance” [88] which describes the accumulation of CN charge due to changes in the local electrostatic potential [89]. Typically, this capacitance is treated as a constant [11, 23, 28, 30, 90, 91], although this is only strictly true when the charge is accumulating in a linear region of the E - k relation [89]. While this infinite-CN approach may be reasonable for certain devices, it does not allow for the treatment of carrier injection via tunneling. Tunneling has been included, via the WKB approximation, in compact models [51, 53], and in a quasi-Fermi level approach for CNFETs [68, 69] which follows a method similar to that used with success in HBT modeling [92]. The latter method, however, appears to predict earlier breakdown in the device saturation than the more rigorous effective-mass Schrödinger-Poisson solution presented here [72]. In this chapter, a modified WKB result is derived in order to account for the non-zero reflection of carriers above a potential energy step, and an algorithm for solving the coupled Schrödinger and Poisson equations is presented.

While FEMLAB [85] had been used for some modeling of CNFETs [68, 69, 71, 77], the program is somewhat unwieldy when making geometrical changes to the device structure, or when performing batch simulations. For these reasons, this chapter begins with a description of a finite difference solution for the Poisson equation.

4.1 Finite Difference Electrostatics

The finite difference method for solving Poisson's equation is well-established in Cartesian coordinates. For a closed cylindrical geometry, however, some modifications are required. Recall that Eq. (3.3) is being solved in two regions, subject to the matching condition given by Eq. (3.2) across the CN surface.

The finite difference approximation to Eq. (3.3) is

$$\frac{V_{i+1,j} - 2V_{i,j} + V_{i-1,j}}{\Delta\rho^2} + \frac{V_{i+1,j} - V_{i-1,j}}{2\rho_i\Delta\rho} + \frac{V_{i,j+1} - 2V_{i,j} + V_{i,j-1}}{\Delta z^2} = 0, \quad (4.1)$$

where $\Delta\rho$ and Δz are the uniform mesh spacings in the ρ - and z -directions, respectively, and the subscripts i, j give the node numbers in the radial and longitudinal directions, respectively. Note that, since the mesh spacing is assumed uniform, ρ_i may be written as $i\Delta\rho$, and that $0 \leq \rho \leq R_G$. This equation can be solved for $V_{i,j}$ for nodes where $\rho \neq 0$ and where $\rho \neq R_C$.

For $i\Delta\rho = R_C$, the finite difference approximation to Eq. (3.2) is

$$\epsilon_{\text{ox}} \left(\frac{V_{i+1,j} - V_{i,j}}{\Delta\rho} \right) - \epsilon_C \left(\frac{V_{i,j} - V_{i-1,j}}{\Delta\rho} \right) = -q \frac{p_{2D,j} - n_{2D,j}}{\epsilon_0}, \quad (4.2)$$

where $p_{2D,j}$ and $n_{2D,j}$ are the 2D hole and electron concentrations at the j^{th} node in the z -direction on the CN surface.

For $i = 0$, *i.e.*, on the CN axis, symmetry implies that

$$\left. \frac{\partial V}{\partial \rho} \right|_{\rho=0} = 0.$$

In this case, l'Hôpital's rule is applied to the singular term of Eq. (3.3). This yields

$$\lim_{\rho \rightarrow 0} \left(\frac{1}{\rho} \frac{\partial V}{\partial \rho} \right) = \left. \frac{\partial^2 V}{\partial \rho^2} \right|_{\rho=0}.$$

For $i = 0$, then, the appropriate equation is

$$4 \frac{V_{1,j} - V_{0,j}}{\Delta\rho^2} + \frac{V_{0,j+1} - 2V_{0,j} + V_{0,j-1}}{\Delta z^2} = 0, \quad (4.3)$$

where $V_{-1,j} \equiv V_{1,j}$ by symmetry.

Eqs. (4.1), (4.2), and (4.3), may be solved simultaneously by formulating the system of equations in matrix form. Note that it may be helpful to normalize these equations such that the coefficients of the matrix do not vary too greatly in magnitude. This is to aid the numerical matrix inversion.

4.2 Treating Quantum Phenomena

With the Poisson equation solved, consider, now, the computation of the CN charge. In equilibrium, this was facilitated by assuming that Fermi-Dirac statistics applied throughout the tube. Out of equilibrium, this is not the case. If the charge were either known or negligible, one could solve the electrostatics problem, and then simply compute the current corresponding to that band diagram. If ballistic transport is assumed, the current \mathcal{I} is given by the Landauer equation [93,94]

$$\mathcal{I} = \frac{4q}{h} \int_{-\infty}^{\infty} T(E) [f(E) - f(E - qV_{DS})] dE, \quad (4.4)$$

where $T(E)$ is the transmission probability for the scattering region, *i.e.*, the channel, h is Planck's constant, and spin- and CN band-degeneracy have been included. If the band diagram is known, then, it is sufficient to compute $T(E)$ in order to determine the terminal current.

4.2.1 Asymptotic Treatment of Tunneling

For the tunneling problem, the WKB approximation is perhaps the most widely used approach. The method is derived in Appendix D, with the standard result that

$$T(E) \propto \exp \left(-2 \int_{z_0}^{z_1} \sqrt{\frac{2m_{\text{eff}} [U(z) - E]}{\hbar^2}} dz \right), \quad (4.5)$$

where $U(z)$ is the band edge, \hbar is Dirac's constant, m_{eff} is the carrier effective mass, and the turning points, z_0 and z_1 , are such that $U(z_0) = U(z_1) = E$ with $U(z) > E$ for $z_0 < z < z_1$. If there are no points where this occurs, and $E > U(z)$, then $T(E) = 1$. If there is only one point where this occurs, or $E < U(z)$ everywhere, then $T(E) = 0$. Note that this expression assumes that $U(z)$ is slowly varying. If, as a further approximation, Eq. (4.5) is taken to be an equality, a simple $T(E)$ is obtained that may be used to compute the terminal current [51,68,77].

While this approximation may be useful, it can be shown, at least for Schottky barrier devices where there may be an abrupt change in $U(z)$ at the contacts, that this may result in a significant overestimate of the current [53]. This is due to the possibility of reflection from the potential energy step even for carriers above the barrier. These carriers would be assigned a unity $T(E)$ in the above scheme. A better expression for this thermionic current component may be derived while still using the WKB ansatz [53].

For phase-incoherent transport, the transmission through the regions close to the source and drain contacts may be computed separately. The total transmission probability may then be computed from

$$T^*(E) = \frac{T_S T_D}{T_S + T_D - T_S T_D},$$

where T_S and T_D are the source and drain transmission probabilities, respectively, and the star superscript indicates that phase-incoherent transport is being considered.

For specificity, the WKB transmission probability is derived for a source electron with energy E greater than any point in a “barrier” region of width w , where $U(z)$ has a significant discontinuity at some point. In particular, consider the case of the source barrier, as in Ref. [70], which has a band diagram that is similar in form to that shown in Fig. 4.1. Note that this figure is intended to be generic, and that the exact same calculation will hold for a barrier that has the opposite concavity, such as in a positive-barrier device, or for a barrier profile that is less smooth, as long as it is smooth enough for WKB to hold in the barrier region itself.

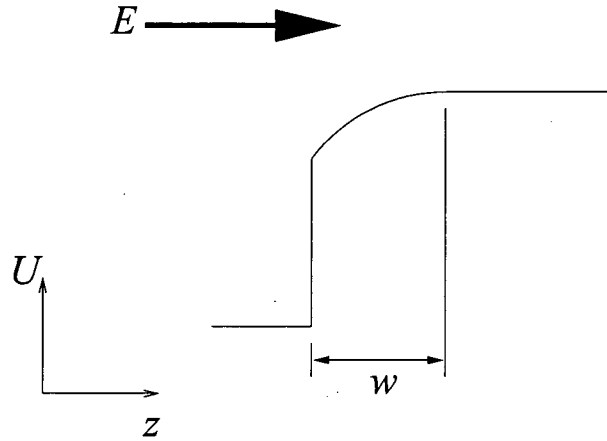


Figure 4.1: Model energy band diagram.

There are several important features of this energy band diagram: a discontinuity occurs at $z = 0$, the potential energy is constant for $z < 0$ and $z > w$, the potential energy and its derivative are continuous across $z = w$, carriers are only incident from the left, and one effective mass describes the carriers in all three regions.

The usual WKB form for the solution can be written as [95]

$$\Psi(z) = \begin{cases} C_1 \exp(ik_S z) + C_2 \exp(-ik_S z) & , \quad z < 0, \\ \frac{1}{\sqrt{k_B(z)}} \left[C_3 \exp\left(i \int_0^z k_B(\hat{z}) d\hat{z}\right) + C_4 \exp\left(-i \int_0^z k_B(\hat{z}) d\hat{z}\right) \right] & , \quad 0 < z < w, \\ C_5 \exp[ik_M(z-w)] & , \quad z > w, \end{cases}$$

where k_S , k_B , and k_M are the wavevectors when $z < 0$, $0 < z < w$, and $z > w$, respectively, and C_1 through C_5 are constants. Note that $k_B(w) = k_M$ and $k'_B(w) = 0$ from the key assumptions of this model problem, where the prime denotes a total derivative. In this case, only the transmission probabilities are of interest, so C_5 may be arbitrarily assigned a value of unity.

The usual matching condition [96] is that Ψ and its derivative must be continuous. The derivatives of the wavefunction are

$$\Psi'(z) = \begin{cases} ik_S [C_1 \exp(ik_S z) - C_2 \exp(-ik_S z)] & , \quad z < 0, \\ C_3 e^{i\nu} \left(\frac{2ik_B^2 - k'_B}{2k_B^{\frac{3}{2}}} \right) - C_4 e^{-i\nu} \left(\frac{2ik_B^2 + k'_B}{2k_B^{\frac{3}{2}}} \right) & , \quad 0 < z < w, \\ ik_M C_5 \exp[ik_M(z-w)] & , \quad z > w, \end{cases}$$

where

$$\nu = \int_0^z k_B(\hat{z}) d\hat{z},$$

and the explicit z -dependencies of k_B , k'_B , and ν have been dropped for notational convenience. Note that \hat{z} is a dummy integration variable used to avoid confusion with z in the integral limit.

These expressions may be further simplified. Since the barrier region joins smoothly onto the right-most region, it is expected that there will, in fact, be no reflected wave anywhere in the barrier region; the initial discontinuity at $z = 0$ is the only source of reflection. This essentially stems from the form of the solution assumed for the output wave, and from the fundamental WKB assumption. Of course, the same result will be arrived at even if C_4 is not set to zero at this point; there will just be a more difficult set of equations to solve.

The simplified system of matching equations is

$$C_1 + C_2 = \frac{C_3}{\sqrt{k_B(0)}},$$

$$\frac{C_3 e^{i\nu(w)}}{\sqrt{k_M}} = C_5,$$

$$ik_S (C_1 - C_2) = C_3 \left(\frac{2ik_B^2(0) - k'_B(0)}{2k_B^{\frac{3}{2}}(0)} \right),$$

with the solution, upon setting $C_5 = 1$,

$$C_1 = \frac{\sqrt{k_M} e^{-i\nu(w)}}{4k_S k_B^{\frac{3}{2}}(0)} [2k_S k_B(0) + 2k_B^2(0) + ik'_B(0)],$$

$$C_2 = \frac{\sqrt{k_M} e^{-i\nu(w)}}{4k_S k_B^{\frac{3}{2}}(0)} [2k_S k_B(0) - 2k_B^2(0) - ik'_B(0)],$$

$$C_3 = \sqrt{k_M} e^{-i\nu(w)}.$$

The source transmission probability is given by

$$T_S = \frac{k_M}{k_S} |C_1|^{-2}.$$

In this case,

$$|C_1|^2 = \frac{k_M}{16k_S^2 k_B^3(0)} \left\{ [k'_B(0)]^2 + 4 [k_B^2(0) + k_S k_B(0)]^2 \right\},$$

therefore

$$T_S = \frac{16k_S k_B^3(0)}{[k'_B(0)]^2 + 4 [k_B^2(0) + k_S k_B(0)]^2}, \quad (4.6)$$

which reduces to unity if $k'_B(0)$ is sufficiently small, and $k_B(0) = k_S$, *i.e.*, when $U(z)$ has no discontinuity, and is slowly varying. An analogous expression holds for T_D .

Incorporation of reflections from the abrupt potential energy step, while still using WKB, results in much better agreement with a full Schrödinger solution [53], and is much less computationally expensive than the full solution.

4.2.2 Solving Schrödinger's Equation

In order to perform a full solution to Schrödinger's equation [72], the scattering matrix method [97] is employed, where input wavefunctions are assumed at either end of the scattering region. The transmission matrix method [98] has also been attempted, where an input wavefunction is assumed at one end, and an output wavefunction at the other. This latter approach, however, while conceptually and implementationally simpler than the former, becomes numerically unstable for low transmission probability wavefunctions. As a result, the scattering matrix method, which has been sufficiently described in Ref. [97], is preferable.

In this system, the charge density is given by

$$p - n = \frac{(p_{1D} - n_{1D})\delta(\rho - R_C)}{2\pi\rho},$$

where $n_{1D}(z)$ and $p_{1D}(z)$ are computed via

$$\frac{\partial^2 \Psi}{\partial z^2} = -\frac{2m_{\text{eff}}}{\hbar^2} (E - U) \Psi, \quad (4.7)$$

where the nanotube effective mass is obtained by fitting a parabola to the tight-binding approximation of the band structure, and is the same for both electrons and holes due to symmetry [4]. The fit for the lowest conduction band of a (16,0) CN is depicted by the dashed line in Fig. 4.2. This approximation is reasonable for energies near the band minimum. For higher energies, the

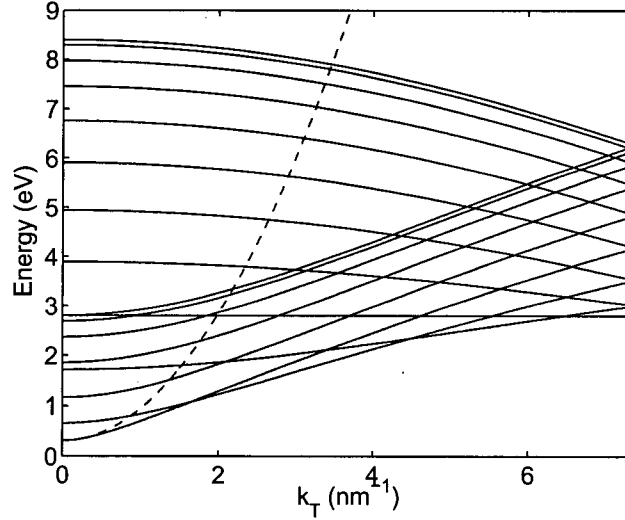


Figure 4.2: A parabolic fit (dashed) to the (16,0) CN band structure (solid).

non-parabolicity of the bands is clear from the figure. Only the first, doubly-degenerate band is included in the calculations presented herein. The potential energy, or band edge U , for each carrier type is specified by

$$U_e(z) = E_{\text{vac}}(z) - \chi_C,$$

$$U_h(z) = -U_e(z) + E_G,$$

where E_G and χ_C are, respectively, the nanotube bandgap and electron affinity, E_{vac} is the vacuum level which is given by $-qV$, and the e and h subscripts refer to electrons and holes, respectively. Note that the sign on U_h is chosen such that a subsequent integration will be over positive energies in order to get the hole concentration in the valence band.

The scattering matrix method provides a numerical solution for Eq. (4.7) by cascading 2×2 matrices [97]. In the required discretization of the potential energy, it is found that the use of a piecewise constant approximation, with plane-wave solutions, is preferable to a piecewise linear approximation, with Airy function solutions, due to the considerable reduction of simulation time without an appreciable increase in the error. Matching of the wavefunction and its derivative on the boundary between intervals i and $i + 1$, assuming a constant effective mass, is performed via the usual continuity and smoothness conditions.

In order to completely specify the wavefunction, two boundary conditions are required. In the contacts, the wavefunction at a given energy is of the form

$$\Psi = \begin{cases} C_1 e^{ik_S z} + C_2 e^{-ik_S z} & , \quad z < 0, \\ C_3 e^{ik_D z} + C_4 e^{-ik_D z} & , \quad z > L_C, \end{cases}$$

where k_S and k_D are the wavevectors in the source and drain contacts, respectively, L_C is the CN length, as depicted in Fig. 3.1(a), and C_1 through C_4 are constants. As an example, noting that an analogous calculation may be performed for the drain by exchange of variables, the case of source injection is now illustrated. For this case, $C_4 = 0$ for all energies. This is one boundary condition. In addition, it is expected that the Landauer equation [93] will hold for the flux, and must be equal to the probability current. For the transmitted wave, this yields [72]

$$\frac{2q}{\pi\hbar} f_S T = \frac{q\hbar}{m_{\text{eff}}} k_D |C_3|^2,$$

where the pre-factor of 2 accounts for the aforementioned band degeneracy, f_S is the Fermi-Dirac carrier distribution in the source, and T is specified by

$$T = \frac{k_D |C_3|^2}{k_S |C_1|^2}.$$

Simple manipulation yields the normalization condition

$$|C_1|^2 = \frac{2m_{\text{eff}}}{\pi\hbar^2} \frac{f_S}{k_S}. \quad (4.8)$$

At each energy, multiplication of the unnormalized wavefunction by a constant satisfies Eq. (4.8).

Including source and drain injection components, the normalized wavefunctions yield the total carrier densities in the system,

$$n_{1D}(z) = \int_{\mathcal{E}_e}^{\infty} (|\Psi_{e,S}|^2 + |\Psi_{e,D}|^2) dE,$$

$$p_{1D}(z) = \int_{\mathcal{E}_h}^{\infty} (|\Psi_{h,S}|^2 + |\Psi_{h,D}|^2) dE,$$

where the subscripts e and h refer to conduction and valence bands, respectively, S and D refer to source and drain injection, respectively, and $\mathcal{E}_{e,h}$ is taken to be the bottom of the band in the appropriate metallic contact. In this way, penetration of the carrier wavefunctions from the metal into the bandgap of the CN can be included [99]. In practice, the integrals are performed using adaptive Romberg integration, where repeated Richardson extrapolations are performed until a predefined tolerance is reached [100]. An adaptive integration method is a necessity in order to properly capture Ψ , which is typically highly peaked in energy for propagating modes. Without adequate sampling of the wavefunction, convergence is unlikely. Alternatively, one could employ a very fine discretization in energy; however, the Romberg method allows for the mesh size to change based on the requirements of the integrand, and results in a much improved simulation time.

As a final note to this section, a check is performed to confirm that the normalization results in the expected form for the electron and hole concentrations when there is no band bending anywhere. In this case, $f_S = f_D \equiv f(E)$, where f_D is the Fermi-Dirac distribution function in the drain, $k_S = k_D \equiv k(E)$, and $T = 1$. Considering electrons,

$$n_{1D}(z) = \int_{\mathcal{E}_e}^{\infty} \frac{4m_{\text{eff}}}{\pi\hbar^2} \frac{f(E)}{k(E)} dE.$$

Substituting in the definition for $k(E)$, *i.e.*,

$$k(E) = \sqrt{\frac{2m(E - U_e)}{\hbar^2}},$$

yields

$$n_{1D}(z) = \int_{\mathcal{E}_e}^{\infty} 2D(E)f(E) dE,$$

where $D(E)$ is the standard 1D, single band, effective-mass DOS, and the factor of two accounts for the double degeneracy of the lowest conduction band. This is indeed the expected expression.

4.2.3 Quantum Capacitance

Since a self-consistent charge-voltage relationship is being computed, it is instructive to consider the “capacitance” of a carbon nanotube [89], *i.e.*, the sensitivity of the charge accumulation to the local electrostatic potential. The term “quantum capacitance” has been used in the CNFET literature [30,90] in order to encapsulate this behaviour, and was most notably used by Luryi [88] in order to develop an equivalent circuit model for devices that incorporate a highly conducting two-dimensional electron gas. In the CNFET case, the results are somewhat different [89]. Here, this potentially important concept is discussed in the context of CNFETs, including a derivation of the 2D case, which has been previously presented in Ref. [88], in order to help illustrate some key differences.

In order to derive analytical expressions, it is assumed that the CNFET is in quasi-equilibrium, and that the carrier distribution functions, as usual, are rigidly shifted by the local electrostatic potential. The quasi-equilibrium assumption permits the use of Fermi-Dirac statistics for the carrier distributions in energy, and is reasonable for the case when tunneling is not significant. Since the DOS is symmetric with respect to E_F , the net carrier density, due to electrons and holes in the semiconductor, may be written as

$$p_d - n_d = \int_0^\infty D(E) \left[f\left(E + \frac{E_G}{2} + qV_{CS}\right) - f\left(E + \frac{E_G}{2} - qV_{CS}\right) \right] dE, \quad (4.9)$$

where the subscript d represents the dimensionality of a quantity, V_{CS} is the local electrostatic potential, E_G is the bandgap, and E_F is taken to be mid-gap when $V_{CS} = 0$. The quantum capacitance, C_Q , is defined as

$$C_Q = q \frac{\partial(p_d - n_d)}{\partial V_{CS}}, \quad (4.10)$$

and has units of F/m² and F/m in the 2D and 1D cases, respectively.

In the 2D case, if the effective-mass approximation with parabolic bands is employed, the DOS is given by

$$D(E) = \frac{m_{\text{eff}}}{\pi \hbar^2} M(E),$$

where $M(E)$ is the number of contributing bands at a given energy, and the m_{eff} for the conduction and valence bands are assumed to be equal as in 1D CNFETs. If this is combined with Eqs. (4.9)

and (4.10), and the order of differentiation and integration is exchanged, the result is that

$$C_Q = \frac{m_{\text{eff}} q^2}{4\pi\hbar^2 k_B T_K} \int_0^\infty M(E) \left[\text{sech}^2 \left(\frac{E + \frac{E_G}{2} - qV_{CS}}{2k_B T_K} \right) + \text{sech}^2 \left(\frac{E + \frac{E_G}{2} + qV_{CS}}{2k_B T_K} \right) \right] dE.$$

If M is a constant, the integration yields

$$C_Q = \frac{M m_{\text{eff}} q^2}{2\pi\hbar^2} \left[2 - \frac{\sinh \left(\frac{E_G}{2k_B T_K} \right)}{\cosh \left(\frac{\frac{E_G}{2} - qV_{CS}}{2k_B T_K} \right) \cosh \left(\frac{\frac{E_G}{2} + qV_{CS}}{2k_B T_K} \right)} \right],$$

which reduces to

$$C_Q = \frac{M m_{\text{eff}} q^2}{\pi\hbar^2}, \quad (4.11)$$

when $E_G = 0$, in agreement with Ref. [88], where metallic properties were assumed. Note that this function is piecewise constant in the metallic case, *i.e.*, it is quantized. In the semiconducting case, the function has continuous variation in it for a given M . For E_G greater than about $15k_B T_K$, however, the function makes a rapid transition from a small value to that given by Eq. (4.11) when V_{CS} crosses $E_G/2$, and is thus effectively quantized.

In the 1D, effective-mass case,

$$D(E) = \frac{M(E)}{\pi\hbar} \sqrt{\frac{2m_{\text{eff}}}{E}}.$$

The explicit energy dependence of this DOS complicates the evaluation of the integral for C_Q . The approach suggested in Ref. [30], using the fact that the derivative of $f(E)$ is peaked about E_F in order to approximate this integral using a Sommerfeld expansion [66], cannot be followed in general, due to the presence of singularities in the 1D DOS. This approximation may be useful if attention is restricted to situations where the DOS singularities are sufficiently far away from E_F .

The capacitance is given by

$$C_Q = \frac{q^2}{2k_B T_K \hbar} \sqrt{\frac{m_{\text{eff}}}{2}} \int_0^\infty \frac{M(E)}{\sqrt{E}} \left[\text{sech}^2 \left(\frac{E + \frac{E_G}{2} - qV_{CS}}{2k_B T_K} \right) + \text{sech}^2 \left(\frac{E + \frac{E_G}{2} + qV_{CS}}{2k_B T_K} \right) \right] dE. \quad (4.12)$$

For sufficiently large $|V_{CS}|$, one of the $\text{sech}^2(\cdot)$ terms may be completely neglected. As a simple example, if $V_{CS} = 0.1\text{V}$ for a material with $E_G \simeq 1\text{eV}$, the contribution to the integral from

the first term is roughly four orders of magnitude greater than the second. This approximation is equivalent to neglecting hole charge for positive V_{CS} , and electron charge for negative V_{CS} . The solid line in Fig. 4.3 shows the equilibrium C_Q as a function of V_{CS} for an arbitrary 1D semiconductor with two valence and conduction bands: at 0.2 and 0.6 eV away from the Fermi level. An effective mass of $0.06m_0$ is assumed, where m_0 is the free-electron mass. The van Hove singularities, at each band edge, result in corresponding peaks in C_Q .

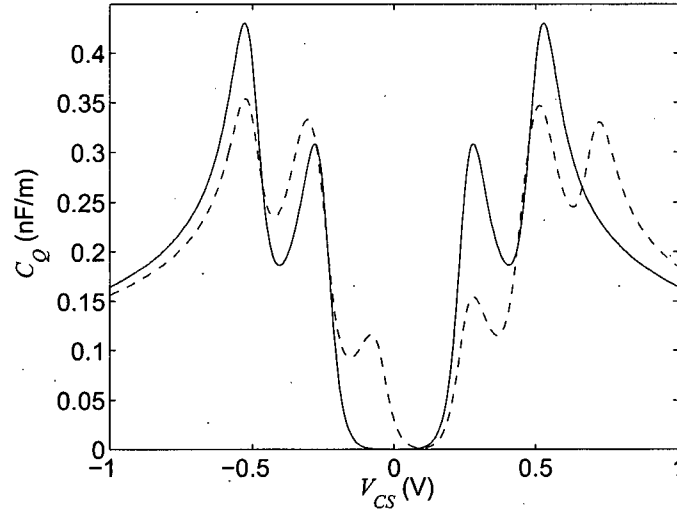


Figure 4.3: 1D quantum capacitance as a function of the local electrostatic potential at equilibrium (solid), and for the mid-length region of an end-contacted semiconductor with a bias voltage of 0.2 V (dashed) between the end contacts. The effective mass is taken to be $0.06m_0$, and energy bands are situated at 0.2 and 0.6 eV on either side of the Fermi level [89].

For a linear energy-wavevector relation, such as that near the Fermi level in graphene or a metallic CN, the DOS is constant. This is the case considered in Ref. [90] and is valid when V_{CS} is such that $f(E)$ is approximately zero before the first van Hove singularity is encountered in the integral. Since the higher energy bands are not relevant to the integration under such a condition, M is constant,

and the DOS is given by

$$D(E) = \frac{2M}{\hbar v_F},$$

where v_F is the Fermi velocity. The result is

$$C_Q = \frac{2Mq^2}{\hbar v_F},$$

which agrees with the expression quoted in Ref. [90].

Note that in Eq. (4.12), C_Q does not manifest itself as a multiple of some discrete amount, so “quantum capacitance” is not an appropriate description, unlike in the metallic 2D and 1D metallic-CN cases, where the capacitance is truly quantized.

The discussion can now be extended to include the non-equilibrium behaviour for a general, 1D, intrinsic semiconductor. All of the numerical results are based on the methods described in Refs. [70] and [72], which consider the cases when transport in the 1D semiconductor is either incoherent or coherent, respectively. While these methods were developed in order to describe CNFETs, their use of the effective-mass approximation allows them to be used for any device, and bias, where the semiconductor is described well by this approximation.

For phase-incoherent transport, a flux-balancing approach [51,70] is used to describe the charge in a CNFET. Considering only the electrons that are far away from the contacts, *i.e.*, in the mid-length region, Eq. (4.9) becomes [51]

$$p_d - n_d = -\frac{1}{2} \int_0^\infty D(E) T^* \left[f \left(E + \frac{E_G}{2} - qV_{CS} \right) \left(\frac{2}{T_D} - 1 \right) + f \left(E + \frac{E_G}{2} + q(V_{DS} - V_{CS}) \right) \left(\frac{2}{T_S} - 1 \right) \right] dE, \quad (4.13)$$

where V_{CS} is evaluated in the mid-length region. A similar expression holds for holes.

The first term in Eq. (4.13) resembles the equilibrium case, so a similar form for that contribution to C_Q is expected. The peak for each contributing band will occur at the same V_{CS} , but the overall magnitude will be smaller due to the multiplication by the transmission function. The second term is also similar except that these peaks will now be shifted by V_{DS} . This is depicted by the dashed curve in Fig. 4.3, where the case illustrated by the solid curve has been driven from equilibrium by $V_{DS} = 0.2$ V. Note the splitting of each large peak into two smaller peaks: one at the same point, and the other shifted by V_{DS} . Of course, the numerical value of the non-equilibrium capacitance depends on the exact geometry considered, as it will influence both V_{CS} and the transmission probabilities in Eq. (4.13), but the trends shown here are general and geometry independent.

For the coherent, non-equilibrium case, it is instructive to consider a metal-contacted device, in which the band discontinuities at the metal-semiconductor interfaces are sufficient to allow significant quantum-mechanical reflection of carriers even above the barrier. Further, this discussion is restricted to short devices since the importance of coherence effects is diminished as the device length is increased. Due to the phase coherence, then, the structure is very much like a quantum well, even for devices where tunneling through the contact barriers is not important. For this device, quasi-bound states are expected to emerge at the approximate energies

$$E_i \simeq \frac{i^2 \pi^2 \hbar^2}{2m_{\text{eff}} L_C^2}.$$

For $m_{\text{eff}} \simeq 0.06m_0$, such as in a (16, 0) CN, $E_i \simeq 6.3(i/L'_C)^2$ eV, where L'_C is in nanometres. This may be compared with the result for metallic CNs, where the linear energy-wavevector relationship yields a $1/L_C$ dependence [90, 101]. Fig. 4.4(a) displays C_Q as a function of position and V_{CS} for this choice of m_{eff} . The maxima, indicated as brighter patches, show a dependence on V_{CS} that reveals the population of quasi-bound states. Moreover, the maxima in position clearly show the characteristic modes expected from the simple square-well analogy. Note that the peak splitting occurs for coherent transport as well, as shown in Fig. 4.4(b), where the peaks have been split by $V_{DS} = 0.1$ V.

The main difference, between the coherent and incoherent cases, is the presence of the quasi-bound states. These serve to increase the number of C_Q peaks, since each quasi-bound state behaves like an energy band, and they also give rise to a strong spatial dependence. While Fig. 4.4 shows only a single-band, coherent result, inclusion of multiple bands would cause C_Q to exhibit peaks corresponding to each band, and to each quasi-bound state.

4.3 Planar Devices

Up to this point, attention has been paid only to cylindrically-gated devices. Experimentally, of course, planar devices, such as the one depicted in Fig. 4.5, are much more feasible. Without cylindrical symmetry, a three-dimensional Poisson equation must be solved for the local electrostatic potential.

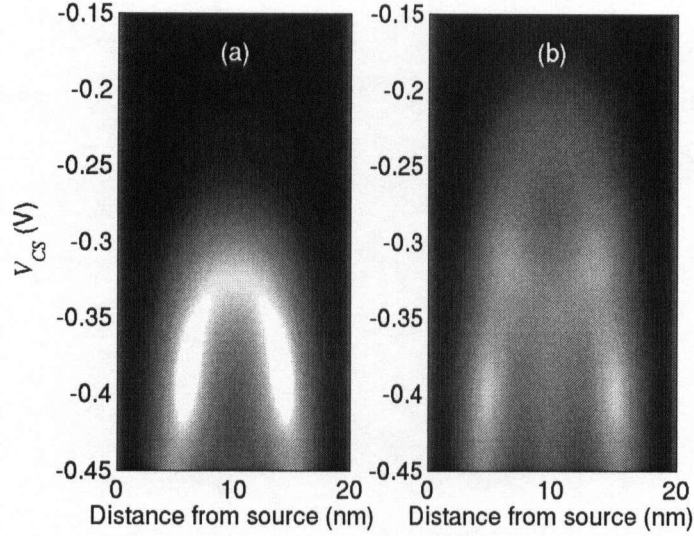


Figure 4.4: 1D quantum capacitance, in arbitrary units, for a short-channel, phase-coherent semiconductor as a function of position and the local electrostatic potential for applied bias voltages of (a) 0 and (b) 0.1 V between the end contacts. The bright areas indicate higher capacitance.

4.3.1 Finite Element Electrostatics

In order to simulate non-coaxial device structures, a numerical method for solving Poisson's equation, that can handle a variety of domain shapes, is desired. For a rectangular domain, the finite difference method described previously is extremely simple to implement. However, if the domain has a slightly more complicated structure, it quickly becomes unwieldy. In this case, the finite element method is chosen.

Since the charge is confined to the surface of the CN, by assumption, it is sufficient to solve Laplace's equation in two regions subject to a matching condition, as described previously. If all lengths are normalized to \mathcal{N} , the problem to solve is

$$\begin{aligned} \frac{1}{\mathcal{N}^2} \nabla \cdot (\epsilon_{\text{ox}} \nabla V) &= 0 \text{ outside the CN,} \\ \frac{1}{\mathcal{N}^2} \nabla \cdot (\epsilon_C \nabla V) &= 0 \text{ inside the CN,} \end{aligned}$$

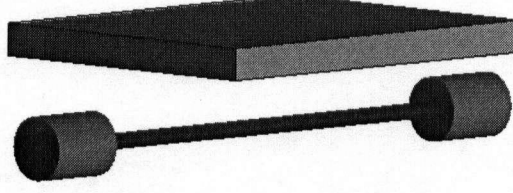


Figure 4.5: Planar geometry CNFET with cylindrical source and drain electrodes, and planar gate electrode. The volume outside of the CN and electrodes is filled with a homogeneous dielectric.

subject to the matching condition given by Eq. (3.2), except that ∇ is divided by \mathcal{N} to account for the length normalization. This can be written in the “weak form” by multiplying through by a test function v , and integrating over the volume of each region. The restriction on v is that it must vanish on any boundaries where a Dirichlet boundary condition applies. The equations to solve become

$$\begin{aligned} \frac{1}{\mathcal{N}^2} \iiint_{\Omega^+} v \nabla \cdot (\epsilon_{\text{ox}} \nabla V) d^3r &= 0, \\ \frac{1}{\mathcal{N}^2} \iiint_{\Omega^-} v \nabla \cdot (\epsilon_C \nabla V) d^3r &= 0, \end{aligned}$$

where Ω^+ and Ω^- are the volumes outside and inside the CN, respectively. Using a standard vector identity and the divergence theorem yields

$$\begin{aligned} -\frac{1}{\mathcal{N}^2} \left(\iiint_{\Omega^+} \epsilon_{\text{ox}} \nabla V \cdot \nabla v d^3r + \iint_{\partial\Omega_C} \epsilon_{\text{ox}} v \nabla V \cdot \hat{\mathbf{n}}|_{R_C^+} d^2r \right) &= 0, \\ -\frac{1}{\mathcal{N}^2} \left(\iiint_{\Omega^-} \epsilon_C \nabla V \cdot \nabla v d^3r - \iint_{\partial\Omega_C} \epsilon_C v \nabla V \cdot \hat{\mathbf{n}}|_{R_C^-} d\sigma \right) &= 0, \end{aligned}$$

where $\partial\Omega_C$ is the CN surface, d^2r is a differential area element, and the other surface integrals have vanished due to the restriction on the function v . Adding these together, and rearranging, gives

$$\iiint_{\Omega^+} \epsilon_{\text{ox}} \nabla V \cdot \nabla v d^3r + \iiint_{\Omega^-} \epsilon_C \nabla V \cdot \nabla v d^3r = - \iint_{\partial\Omega_C} v \left(\epsilon_{\text{ox}} \nabla V|_{R_C^+} - \epsilon_C \nabla V|_{R_C^-} \right) \cdot \hat{\mathbf{n}} d^2r.$$

Applying Eq. (3.2) yields

$$\iiint_{\Omega^+} \epsilon_{\text{ox}} \nabla V \cdot \nabla v d^3r + \iiint_{\Omega^-} \epsilon_C \nabla V \cdot \nabla v d^3r = q\mathcal{N} \iint_{\partial\Omega_C} v \left(\frac{p_{2D} - n_{2D}}{\epsilon_0} \right) d^2r. \quad (4.14)$$

This is known as the “weak form” of the electrostatics problem.

In the finite element method, the weak formulation is solved, and in the Galerkin technique, v and V are expanded in terms of basis functions, *i.e.*,

$$V = \sum_i V_i \omega_i,$$

$$v = \sum_i B_i \omega_i,$$

where B_i and V_i are constant coefficients to the basis function ω_i . These expansions can be substituted into Eq. (4.14), and since the expansion must be true for any choice of v that vanishes on the Dirichlet boundaries, the sum over the B_i coefficients may be treated term-by-term. This allows those coefficients to cancel out of the final expression, which is

$$\sum_j V_j \left(\iiint_{\Omega^+} \epsilon_{\text{ox}} \nabla \omega_j \cdot \nabla \omega_i \, d^3r + \iiint_{\Omega^-} \epsilon_C \nabla \omega_j \cdot \nabla \omega_i \, d^3r \right) = \frac{q\mathcal{N}}{\epsilon_0} \iint_{\partial\Omega_C} \omega_i (p_{2D} - n_{2D}) \, d^2r.$$

This may be written compactly as

$$\sum_j K_{ij} V_j = C_j,$$

or, in matrix form,

$$KV = C,$$

where K is known as the stiffness matrix.

It is convenient that K only needs to be computed once per geometry. The forcing term, C , needs to be computed at each iteration. In order to perform these calculations, the choice of basis function must be specified. In this work, piecewise linear basis functions are selected, where ω_i is unity at the i^{th} node, decreases linearly to vanish at the neighbouring nodes, and remains zero throughout the remainder of the domain. This is most easily seen in 1D, as shown in Fig. 4.6. Note that V_i will equal the potential at the i^{th} node. The basis function for a 3D domain is a straightforward generalization. In this work, the freeware program TetGen [102] is used to mesh the 3D domains.

4.4 Solution Procedure

The solution procedure is similar to that described in Chapter 3 except that the charge is now computed via a Schrödinger solution. Whether the finite difference or the finite element method is

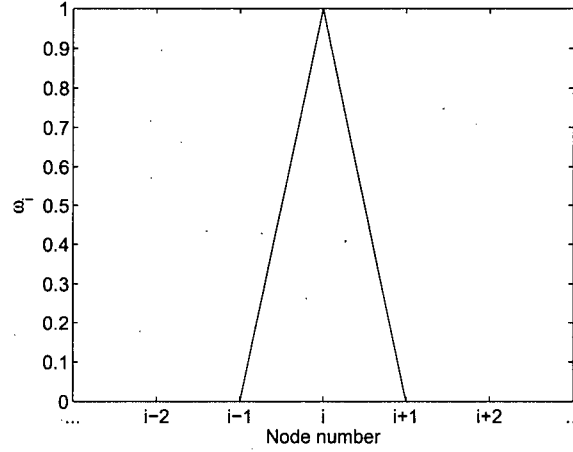


Figure 4.6: 1D piecewise linear basis function for use in the finite element method.

used to achieve the Poisson solution, the simulation procedure begins by precalculating the matrix corresponding to the discrete equivalent of the differential operation, *e.g.*, the stiffness matrix in the finite element case. The solution procedure proceeds as follows:

1. Solve Laplace's equation to get V_L .
2. Solve Schrödinger's equation over a sufficiently large energy range, and integrate to obtain the charge density. Here, $E_{\text{vac}} = -qV^{(i)}$, where $V^{(0)} = V_L$. In practice, this method treats one conduction band considering electrons, and one valence band considering holes. For electrons, the energy range is from the bottom of the band in the metal lead to around $15k_B T_K$ above the Fermi level of the injecting contact. Source and drain injection are treated independently. As mentioned previously, the integration over energy is done using adaptive Romberg integration [100].
3. If using the finite element method, the charge density will need to be multiplied by the basis functions and integrated in order to obtain the \mathcal{C} vector. In the finite difference case, this integration is not required, although the charge vector still requires generation when solving the system using matrices.
4. Compute the residual r , which, in the finite element notation, is given by $r^{(i)} = KV^{(i)} - \mathcal{C}^{(i)}$.

5. Update the potential according to $V^{(i+1)} = V^{(i)} - \vartheta K^{-1}r^{(i)}$, where ϑ is a damping factor that is required in order to get convergence. A value of 0.4 was typically used for this parameter. In practice, K^{-1} is not explicitly calculated, but rather, the backslash operator in Matlab is used in order to compute $K^{-1}r^{(i)}$ using Gaussian elimination.
6. Check that $\|K^{-1}r^{(i)}\|_{\infty}$ is below some tolerance, where $\|\cdot\|_{\infty}$ is the infinity norm, *i.e.*, the maximum value of $|K^{-1}r^{(i)}|$. If not, increment i , and return to Step 2. If it is, convergence has been obtained, and the current may be computed via Eq. (4.4), where $T(E)$ is stored from the computation in Step 2.

Note that this iteration scheme is similar to that used in Chapter 3 when $\vartheta = 1$. When $\vartheta \neq 1$, this is referred to as Picard iteration.

4.5 Results and Discussion

This method has been used to predict the performance of cylindrically-symmetric devices, and also to suggest the possible performance of other, asymmetric, devices. Here, results are presented for the DC behaviour of CNFETs, and a metric for the switching time, τ_d , is examined. τ_d is an important figure-of-merit for digital applications, and may be evaluated using DC quantities. Small-signal results have also been generated from these DC simulations [73–76, 103], and the analyses of these can be found in the appropriate references.

4.5.1 DC

For the DC characteristics, as in Chapter 3, the closed, cylindrically-symmetric device shown in Fig. 3.1(a) continues to serve as the model geometry, but now ϵ_C is allowed to differ from ϵ_{ox} . In particular, $\epsilon_C = 1$ [81] for all devices, and ϵ_{ox} is chosen as appropriate for the choice of dielectric. Due to the permittivity discontinuity, the finite difference Poisson solution is utilized.

The model device includes a 20 nm-long, (16, 0) CN, and a 2.5 nm-thick oxide. Additionally, all of the workfunctions are 4.5 eV, $\epsilon_{ox} = 25$, and $\mathcal{E}_{e,h}$ is taken to be 5.5 eV away from the contact Fermi level unless otherwise specified. This is equivalent to choosing a Fermi wavevector in the metallic contacts of around 0.3 \AA^{-1} when the effective mass is equal to that of the (16, 0) CN. This choice is similar to the smallest value used in Ref. [54].

Beginning with band diagrams, Fig. 4.7 shows the results for $V_{DS} = 0$ and 0.4 V with $V_{GS} = 0.5$ V. In equilibrium, the carrier concentrations, particularly away from the source and drain contacts, are

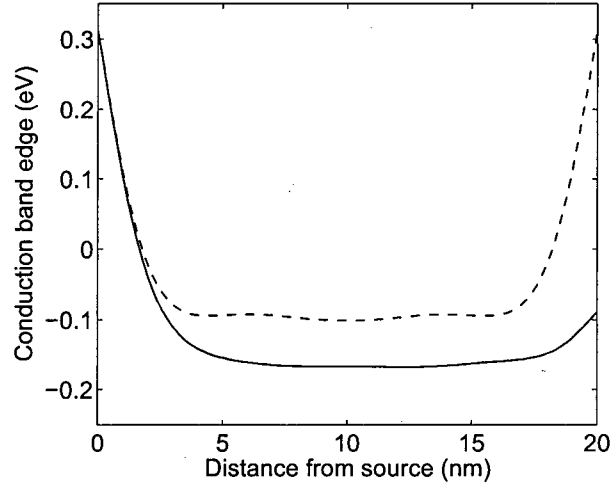


Figure 4.7: Conduction band edges for $V_{GS} = 0.5$ V, and $V_{DS} = 0$ (dashed) and 0.4 V (solid) [72].

in reasonable agreement with the results from using Fermi-Dirac statistics. It is interesting to note that the agreement is not exact, and this is attributed to the rapid potential energy variation at the end contacts, which leads to resonance in the carrier wavefunctions. Strictly speaking, then, a Schrödinger-Poisson solution is required even in equilibrium. Application of a drain voltage serves to decrease the number of electrons in the channel, and hence, the conduction band in the mid-tube region drops to a lower value.

In both cases, due to the short CN length and large band discontinuities at the contacts, interference effects and contact reflections modify the carrier distribution functions in the CN. Note that this is significant even for carriers above the contact barriers. To see this, consider Fig. 4.8, where the conduction band edges and transmission probabilities for electrons with $V_{GS} = 0.5$ V and $V_{DS} = 0.4$ V are shown. In (a) is a “negative barrier” device, where all of the carriers in the mid-tube region are thermionically emitted. Clearly, approximating the transmission probability as unity for these carriers would be grossly inaccurate. Use of Eq. (4.6) results in a much more accurate value despite the fact that it neglects interference effects, and only includes contact reflections [53]. The same result, but for a “positive barrier” device, is shown in (b). Fig. 4.9 shows the effect of

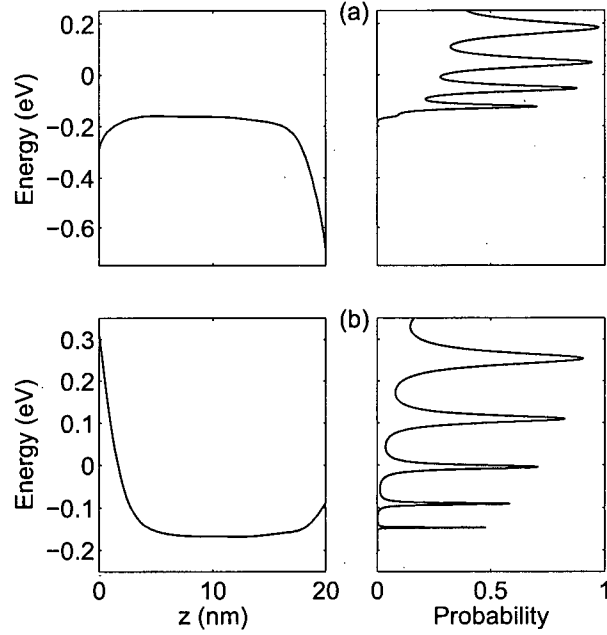


Figure 4.8: Conduction band edges and transmission probabilities for $V_{GS} = 0.5\text{ V}$, $V_{DS} = 0.4\text{ V}$, with (a) $\Phi_S = \Phi_D = 3.9\text{ eV}$, and (b) $\Phi_S = \Phi_D = 4.5\text{ eV}$ [72].

the resonant peaks in transmission on the carrier distribution in energy for the same device as in Fig. 4.8(a). Note that this function bears little similarity to a Fermi-Dirac distribution function multiplied by the density of states. As a result, quasi-Fermi level techniques [68,69], as used in HBT modeling [92], are difficult to rigorously justify.

Consider, now, the terminal characteristics. Fig. 4.10 shows the \mathcal{I}_D - V_{GS} characteristic for $V_{DS} = 0.4\text{ V}$ with $\Phi_S = \Phi_D$ set to 3.9, 4.3, and 4.5 eV. Note that Φ_G is left fixed at 4.5 eV since changes in this parameter, as noted previously, are equivalent to changes in V_{GS} . An important thing to note is that these CNFETs are ambipolar [104–107], *i.e.*, the current is either electron- or hole-dominated depending on V_{GS} . This is depicted schematically in Fig. 4.11, where low V_{GS} yields a hole-dominated current, and high V_{GS} yields an electron-dominated current. For the case shown in Fig. 4.11, where all workfunctions are equal, it is expected that the current minimum will occur at $V_{GS} = V_{DS}/2$. This is confirmed by Fig. 4.10. When the contact workfunctions are adjusted, this minimum moves, and may also change somewhat in magnitude.

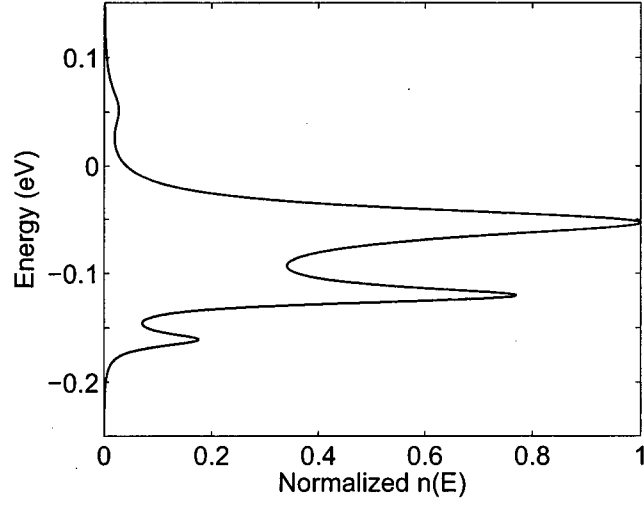


Figure 4.9: Carrier concentration as a function of energy, normalized to its maximum value, for $V_{GS} = 0.5 \text{ V}$, $V_{DS} = 0.4 \text{ V}$, with $\Phi_S = \Phi_D = 3.9 \text{ eV}$ [72].

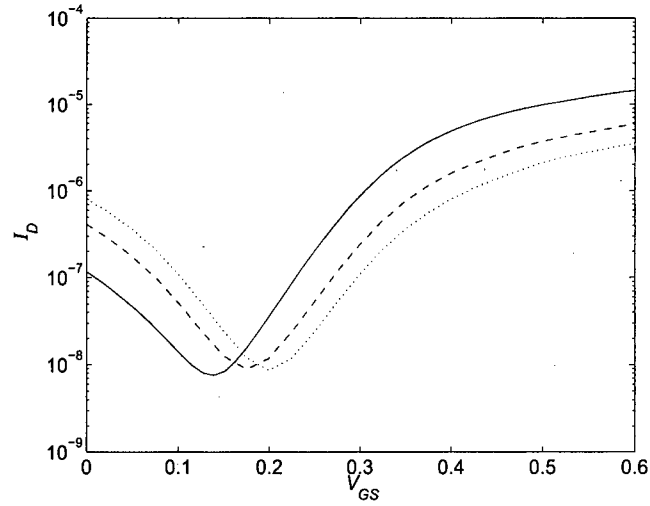


Figure 4.10: I_D - V_{GS} for a closed cylindrical device with $V_{DS} = 0.4 \text{ V}$, and $\Phi_S = \Phi_D$ set to 3.9 (solid), 4.3 (dashed), and 4.5 eV (dotted).

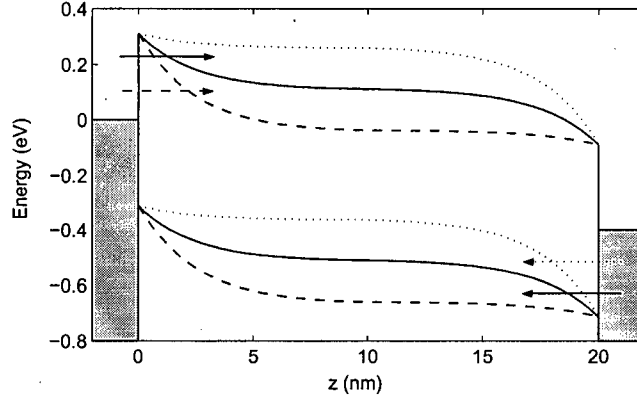


Figure 4.11: Schematic showing the ambipolar conduction mechanism. Shown are the hole (dotted) and electron (dashed) dominated cases along with the crossover (solid) between these two cases [70].

4.5.2 Switching Speed

In order to examine the suitability of CNFETs for digital applications, consider τ_d as analyzed in Ref. [55]. A method for computing this quantity from DC data, for devices with non-optimized threshold voltage [108], has been proposed in Ref. [30]. Here, this metric is examined, and results generated from the effective-mass Schrödinger-Poisson solver are presented.

In order to compute τ_d , consider the circuit shown in Fig. 4.12, where a load (transistor B) is discharged by a driver (transistor A) at a rate given by the current \mathcal{I} , defined as

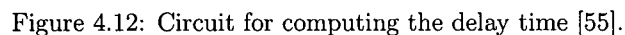
$$\mathcal{I}(V_{GS}^A, V_{DS}^A) = \frac{dQ_G^B(V_{GS}^B, V_{DS}^B)}{dt},$$

where $V_{GS}^A = V_{DS}^B = -V_{DD}$, where V_{DD} is the positive supply voltage given by the difference between the OFF and ON gate voltages on the load, namely:

$$V_{DD} = V_{GS,OFF}^B - V_{GS,ON}^B.$$

The discharge current $\mathcal{I} = -\mathcal{I}_D^A$ will vary as the load discharges. As an approximation, the practice is to evaluate \mathcal{I} at its highest value, namely:

$$\mathcal{I} = -\mathcal{I}_D^A(-V_{DD}, V_{GS,ON}^B) \equiv -\mathcal{I}_{ON}^A.$$


$$\frac{\partial Q_C^B}{\partial V_{GS}^B} = -C_G^B(V_{GS}^B, -V_{DD}),$$
$$C_G^B = C_G^B(V_{GS,ON}^B, -V_{DD}) \equiv C_{G,ON}^B.$$
$$\tau_d = \frac{C_{G,ON}^B V_{DD}}{|I_{ON}^A|}.$$

44

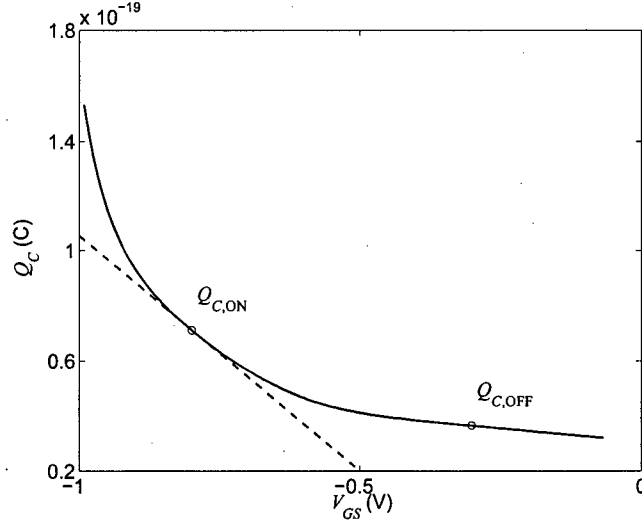


Figure 4.13: Computation of the gate capacitance in the ON-state via the slope of the dashed line [55].

$\mathcal{I}_{\text{ON}}/\mathcal{I}_{\text{OFF}}$, where the latter quantity is evaluated by sliding a window of width V_{DD} along a plot of \mathcal{I}_D vs. V_{GS} , as described previously in Ref. [109].

For these results, the planar geometry depicted in Fig. 4.5 is chosen, and the finite element method is used to solve Poisson's equation. Since the overall solution method uses a 1D Schrödinger equation, the axial potential of the CN is used when computing E_{vac} . Three devices are considered: a (22,0) CN with an 8 nm oxide, a (10,0) CN with an 8 nm oxide, and a (10,0) CN with a 2 nm oxide. Parameters that are common between all three devices are: $\epsilon_{\text{ox}} = 16$, as appropriate for hafnia, the end contacts are Pd, the underlap of the gate and end contacts, as depicted in Fig. 4.14, is $L_{SG} = L_{DG} = 4$ nm, the gate length, thickness and breadth is 20, 5, and 10 nm, respectively, and the end contacts have a radius $R_{\text{cont}} = 6$ nm and length $L_{\text{cont}} = 1$ nm. The (22,0), 8 nm device was chosen due to its similarity to a high-DC-performance experimental device from Ref. [110], and the other two devices are to show the effect of varying the tube diameter, and hence the bandgap, and the oxide thickness. In all cases, the workfunction of the nanotube is 4.7 eV [111], and the workfunction of the gate is adjusted to give reasonable \mathcal{I}_{ON} in the simulation range of $-1 < V_{GS}^B < -0.5$ V. The barrier height for the Pd/nanotube end contacts is taken from Ref. [112], and has values of -0.04 and

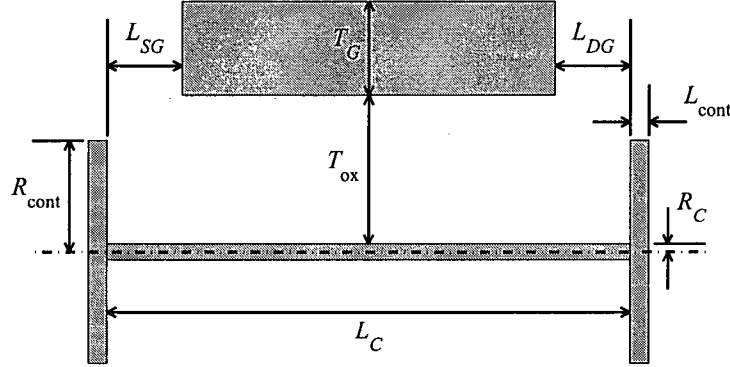


Figure 4.14: Cross-section of the planar geometry CNFET [55].

0.3 eV for holes in the (22, 0) and (10, 0) cases, respectively. Unless otherwise stated, simulations are performed at $V_{DS} = -0.5$ V, and I_{ON}/I_{OFF} is evaluated at $V_{DD} = 0.5$ V. The choice of this value is arbitrary [108], and is used here as it may be realistic for far-future, low-voltage logic applications.

Fig. 4.15 shows the I_D - V_{GS} characteristics for the three model devices. While the (22, 0) device has the best ON-current, due to its smaller bandgap, it also has the worst OFF-current. This is due to ambipolar conduction [105] setting in before a low OFF-current can be attained [108]. In order to suppress this effect, a larger bandgap CN is required [109] such as in the (10, 0), 8 nm device, where there is roughly twice the bandgap of the (22, 0) CN. Due to the higher barrier height for this device [112], a significant reduction in the ON-current occurs. This may be compensated for by thinning the oxide down to 2 nm, where reasonable ON and OFF currents are observed.

Fig. 4.16 displays the intrinsic switching time against the ON/OFF ratio. Due to the high ON-current in the (22, 0) device, there is an extremely short switching time. Unfortunately, this device has the worst ON/OFF ratio of the three devices. By choosing the smaller CN, the ON-current is sacrificed, but a gain in ON/OFF ratio is obtained. Note that these values are in rough agreement with the sub-picosecond switching times predicted for CNFETs out to $I_{ON}/I_{OFF} \approx 10^4$ in Ref. [109] and $\approx 10^2$ in Ref. [108]. In all three cases, although it is not apparent in the (10, 0), thick-insulator case until longer values of τ_d than are shown in Fig. 4.16, I_{ON}/I_{OFF} starts to decrease after reaching its maximum value. This corresponds to $V_{GS,OFF}$ passing through, in a positive direction for the p -type devices considered here, the point at which $|I_D|$ is a minimum [108]. Additionally, the (10, 0), thin-insulator case shows an increase in τ_d at the low end of its I_{ON}/I_{OFF} range, which corresponds

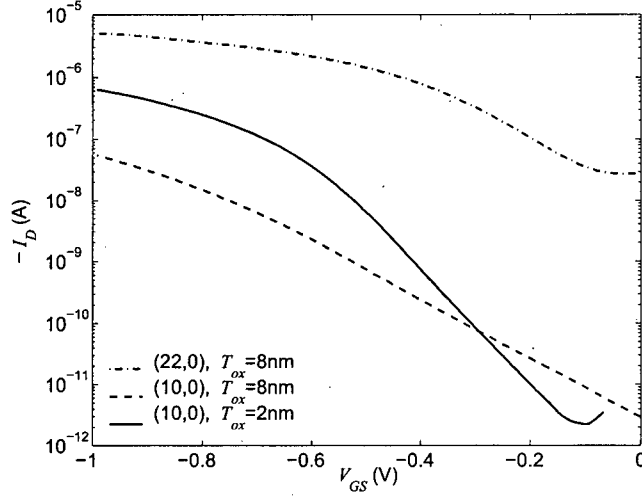


Figure 4.15: I_D - V_{GS} for three planar devices [55].

to operation at high, negative values of V_{GS} . To examine this phenomenon, consider Fig. 4.17, which shows the V_{GS} dependence of the charge in the nanotube, from which the intrinsic capacitance is computed. It is clear that there is a pronounced increase in C_G as the negative gate bias reaches high values, and this is responsible for the increase in τ_d . The rise in charge comes about because the thinning of the potential barrier at the drain, due to the use of a thinner insulator, allows holes to tunnel into the nanotube at high, negative gate bias. The valence-band profiles for the two (10,0) cases, displayed in Fig. 4.18, show that this does not happen with a thicker insulator. This restriction of the operating range for the aggressively scaled device was not apparent in earlier work [109], where it was assumed that a zero-height Schottky barrier could be obtained with a small-diameter nanotube.

While the intrinsic switching time may be a useful metric in the early stages of device design [108], it is well understood that parasitic capacitances become more important as devices are aggressively scaled. In CNFETs, the contacts are inevitably close together, and so it may also be useful to consider an extrinsic switching time which accounts for the inter-electrode capacitances when computing C_G . This is done by computing the charge on the gate itself, rather than using the charge on the CN. This is straightforwardly done from the finite element Poisson solution [55], and the resultant switching times and ON/OFF ratios are shown in Fig. 4.19. Even though modest dimensions have been

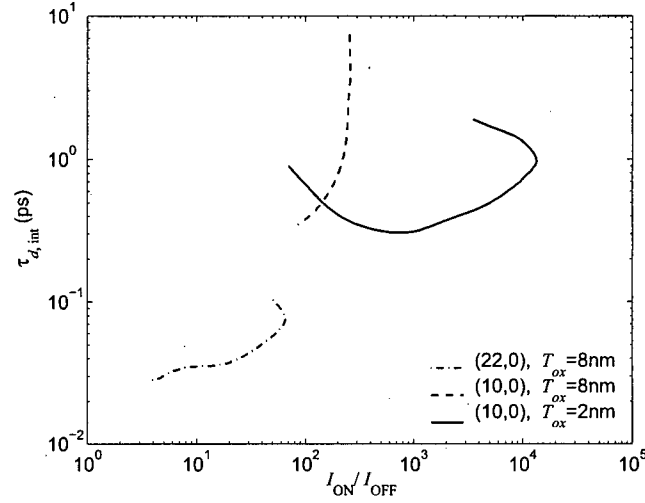


Figure 4.16: Intrinsic switching time, and ON/OFF ratio for three model CNFETs [55].

employed for the source and drain metallization, which would probably result in excessive series resistance, the contributions of the gate-source and gate-drain capacitances to C_G are huge, and lead to much larger switching times than predicted for the intrinsic cases. Numerically, average values of 0.2 and 6 aF are obtained for the intrinsic and extrinsic gate capacitances, respectively.

These results are placed in the context of modern devices by comparing them with calculations of the extrinsic switching time of present-day silicon MOSFETs. SPICE simulations¹ were performed for a minimum-size NMOSFET using parameters for a commercially available² 90 nm technology. An NMOS silicon device was chosen as its higher current drive makes it superior to PMOS devices [108]. In this way, a fair comparison can be made with the CNFETs studied in the present work, which employ Pd end contacts, which are known to give low barrier heights and, consequently, higher ON currents than are presently possible with n -type CNFETs. Simulations were performed at $V_{DS} = 1$ V, for which this technology is optimized, over a V_{GS} range of -1 to $+2$ V, from which the I_{ON}/I_{OFF} ratio was obtained for $V_{DD} = 1$ V. The total gate capacitance, as returned by SPICE as the capacitance C_{GG} at the chosen operating point, was used to estimate the extrinsic switching time. The result is shown by the dotted line in Fig. 4.19. It is clear that this device out-performs the small-

¹SPICE simulations were performed by D. Ho of the University of British Columbia.

²STMicroelectronics, <http://www.st.com/stonline/>.

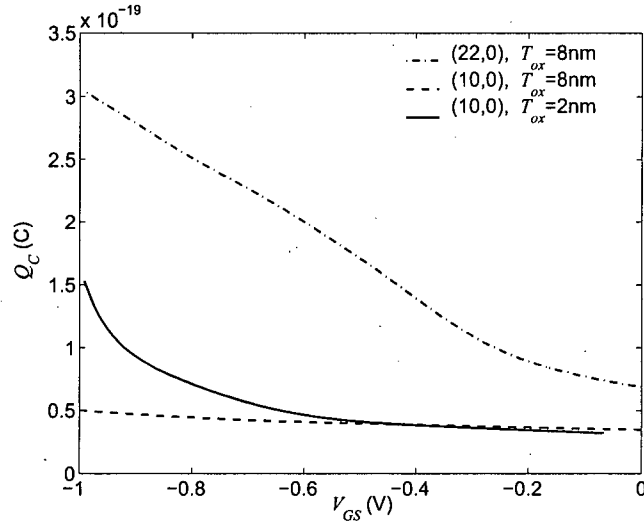


Figure 4.17: CN charge as a function of the gate voltage for three model CNFETs [55].

diameter, thin-oxide CNFET that appears to be the most practically relevant of the nanotube devices considered here. Typically, the total gate capacitance of the Si FET is around 200 aF, but superior switching performance is achieved by virtue of the higher drive current and the absence of ambipolar conduction. This suggests that a more extensive study of the switching characteristics of CNFETs, involving different geometries and doped-contact, rather than Schottky-contact, devices needs to be performed before any claims can be made as to the superiority of CNFETs over silicon MOSFETs. This result is in contrast to that of Guo *et al.*, where the sub-picosecond intrinsic switching times of CNFETs were compared to a silicon device [109] using an effective oxide thickness measurement to calculate the gate capacitance [113]. Their conclusion was that CNFETs would actually out-perform the silicon device [109].

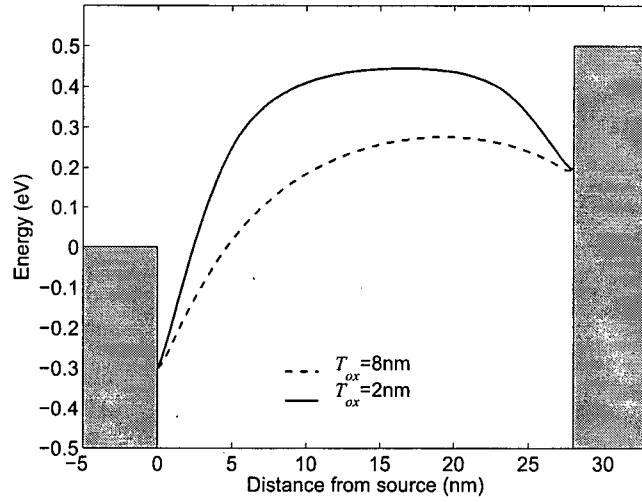


Figure 4.18: Valence bands for two oxide thicknesses showing the thinner tunneling barrier for holes in the thin-oxide case [55].

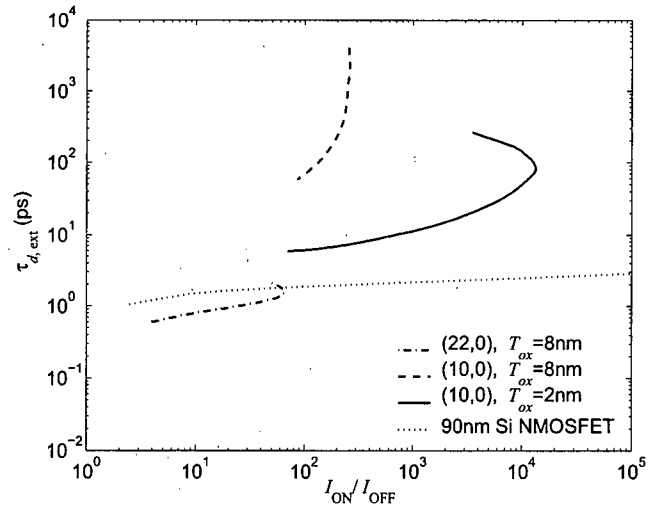


Figure 4.19: Extrinsic switching time, and ON/OFF ratio for three model CNFETs, and a present-day silicon MOSFET [55].

Chapter 5

The Tight-Binding Method

A limitation of the effective-mass approximation is its reliance on sufficiently low bias, such that the parabolic approximation to the dispersion relation holds. In order to improve on this model, consider a tight-binding Hamiltonian, and the non-equilibrium Green's function (NEGF) formalism [94, 114]. This method allows for non-parabolicity in the bands, as shown in Fig. 4.2, and continues to make the independent-electron approximation. In addition, all of the energy bands may now be considered simultaneously. This will allow for the inclusion of inter-band tunneling without resorting to an energy-dependent effective mass [63, 76], such as in the Flietner dispersion relation [115] with a complex band structure. For quantum wells, introducing an energy dependence into m_{eff} may lead to erroneous results [116], so that method should be used with caution.

While the NEGF approach has previously been employed to model CNFETs [56–58, 65], here the azimuthal potential variations are explicitly included, in contrast to Ref. [56], and a quadratic matrix equation is solved, in order to obtain the required self-energy matrices [117], instead of the usual recursive algorithm [118] which has been used in Refs. [56–58]. In addition, an extension is made to previous works by deriving an expression for the inter-atomic current, which may have use, for example, when considering structures with defects.

The key to this method is the computation of the Green's function for a tight-binding Hamiltonian, so first, a discussion regarding the relevance of this function in terms of computing the charge on the CN surface is presented.

5.1 Relevance of the Green's Function

In this section, the derivation presented in Ref. [119] is followed in order to derive the local density of states (LDOS), *i.e.*, the DOS as a function of position, for some atomic system via the Green's function.

Suppose that a solution to the time-independent Schrödinger equation,

$$[\mathcal{Z} - H(\mathbf{r})]\Psi(\mathbf{r}) = 0, \quad (5.1)$$

is required, where \mathcal{Z} is a complex number. For some particular choice of Hamiltonian and boundary conditions, this can, in principle, be solved for the wavefunction.

One technique for solving differential equations is through the use of a Green's function, G [120]. The appropriate equation to solve is

$$[\mathcal{Z} - H(\mathbf{r})]G(\mathbf{r}, \mathbf{r}'; \mathcal{Z}) = \delta(\mathbf{r} - \mathbf{r}'). \quad (5.2)$$

The Green's function, then, gives the response at some point \mathbf{r} due to an excitation at the point \mathbf{r}' , and this fact can be used to solve for $\Psi(\mathbf{r})$ in the entire domain, subject to the boundary conditions and any internal forcing functions. $G(\mathbf{r}, \mathbf{r}'; \mathcal{Z})$, however, also contains information that, when used with the NEGF formalism, obviates the need for an explicit calculation of $\Psi(\mathbf{r})$.

Let $\Psi_{c,i}(\mathbf{r})$ be the i^{th} eigenfunction of $H(\mathbf{r})$. Using completeness, G may be written as

$$G = \sum_i' C_i \Psi_{c,i}(\mathbf{r}),$$

where C_i are constants. Here, \sum' implies a sum over the eigenfunctions in the discrete spectrum plus an integral over the eigenfunctions in the continuous spectrum [119]. Eq. (5.2) becomes

$$\sum_i' C_i [\mathcal{Z} - H(\mathbf{r})] \Psi_{c,i}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}'). \quad (5.3)$$

Now, multiply both sides of Eq. (5.3) by $\Psi_{c,j}^\dagger$, and integrate over position. If E_i is the eigenvalue corresponding to $\Psi_{c,i}$, the result is, after using the orthonormality of the eigenfunctions,

$$\Psi_{c,i}^\dagger(\mathbf{r}') = C_i (\mathcal{Z} - E_i).$$

Therefore, the Green's function may be written as

$$G(\mathbf{r}, \mathbf{r}'; \mathcal{Z}) = \sum_i' \frac{\Psi_{c,i}^\dagger(\mathbf{r}') \Psi_{c,i}(\mathbf{r})}{\mathcal{Z} - E_i}. \quad (5.4)$$

Now, define

$$G^\pm(\mathbf{r}, \mathbf{r}'; E) = \lim_{\xi \rightarrow 0} G(\mathbf{r}, \mathbf{r}'; E \pm i\xi), \quad (5.5)$$

which is useful since there are poles in Eq. (5.4) at $Z = E_i$, where E_i is real, and it is convenient to shift them off of the real axis [119, 121]. Further, let

$$G_0(\mathbf{r}, \mathbf{r}'; E) = G^+(\mathbf{r}, \mathbf{r}'; E) - G^-(\mathbf{r}, \mathbf{r}'; E). \quad (5.6)$$

According to Ref. [121], on page 426,

$$\lim_{\xi \rightarrow 0} \frac{1}{v \pm i\xi} = \text{p.v.} \frac{1}{v} \mp i\pi\delta(v), \quad (5.7)$$

where p.v. denotes the principal value. Strictly speaking, this “identity” is an abbreviation for an integral identity resulting from complex analysis, and involves an explicit test function. Here, the full derivation is neglected, and Eq. (5.7) is treated as a true identity.

Using Eqs. (5.4)–(5.7) yields

$$G_0(\mathbf{r}, \mathbf{r}'; E) = -2\pi i \sum_i \delta(E - E_i) \Psi_{c,i}(\mathbf{r}) \Psi_{c,i}^\dagger(\mathbf{r}'),$$

where the delta function is either of the Kronecker or the Dirac variety, depending on whether the discrete or continuous spectrum is being considered, respectively. Defining

$$D(\mathbf{r}; E) = -\frac{1}{2\pi i} G_0(\mathbf{r}, \mathbf{r}; E), \quad (5.8)$$

it may be observed that this will represent the density of states at \mathbf{r} , and if an integration of $D(\mathbf{r}; E)$ over all \mathbf{r} is performed, the total number of states at a given energy emerges. Note that, if spin degeneracy is included, D should be multiplied by two. The density of states allows for the computation of the carrier distributions in the NEGF formalism, and hence, the wavefunctions do not require explicit calculation.

Note that, from Eqs. (5.4) and (5.5),

$$G^\dagger(\mathbf{r}, \mathbf{r}'; Z) = G(\mathbf{r}', \mathbf{r}; Z^\dagger),$$

$$G^-(\mathbf{r}, \mathbf{r}'; E) = [G^+(\mathbf{r}', \mathbf{r}; E)]^\dagger.$$

As a result, Eq. (5.8) can also be written as

$$D(\mathbf{r}; E) = \mp \frac{1}{\pi} \text{Im} \{ G^\pm(\mathbf{r}, \mathbf{r}; E) \}.$$

A generalized LDOS is defined as

$$D(\mathbf{r}, \mathbf{r}'; E) = \frac{i [G^+ - (G^+)^\dagger]}{2\pi} \equiv \frac{A}{2\pi}, \quad (5.9)$$

where A is known as the spectral function. If this function is discretized into matrix form, the equilibrium electron concentration, including spin degeneracy, may be written as

$$n_{2D} = \frac{1}{\pi} \text{diag} \left[\int_{-\infty}^{\infty} A f(E - E_F) dE \right], \quad (5.10)$$

where the $\text{diag}(\cdot)$ operator extracts the diagonal elements, *i.e.*, the elements corresponding to $\mathbf{r} = \mathbf{r}'$, of its matrix argument.

Out of equilibrium, states that are coupled to a particular lead are filled according to the Fermi function in that lead [114]. Here, a “lead” is defined as a region with constant potential energy in the transport direction, as opposed to the “scattering region” where arbitrary band bending is permitted. As shown in Refs. [94] and [114], a scattering region connected to two semi-infinite leads has a Green’s function given by

$$G_S^+ = \frac{1}{\Lambda} \lim_{\xi \rightarrow 0} [(E + i\xi) I - H_S - \Sigma_L - \Sigma_R]^{-1},$$

where I is the identity matrix, the S subscript denotes that the Green’s function provides information about the scattering region only, H_S is the Hamiltonian for the scattering region, Σ_L and Σ_R are called the self-energy matrices for the left and right leads, respectively, and Λ is the atomic density given by

$$\Lambda = \frac{3a^2\sqrt{3}}{4}.$$

This Λ factor comes about by assuming that the discrete wavefunction centred on an atom is constant over the area Λ , and is used in order to give G_S^+ the correct units. Unless otherwise specified, a set of units is employed such that $\Lambda = 1$.

As a result,

$$[G_S^+]^{-1} - [(G_S^+)^{-1}]^\dagger = \Sigma_L^\dagger - \Sigma_L + \Sigma_R^\dagger - \Sigma_R, \quad (5.11)$$

where ξ has gone to zero when we took the inverse. Note that $H_S = H_S^\dagger$, and the fact that the Hermitian conjugate of an inverse is equal to the inverse of a Hermitian conjugate, have been used. Eq. (5.11) is often written

$$[G_S^+]^{-1} - [(G_S^+)^{-1}]^\dagger = i(\Gamma_L + \Gamma_R),$$

where $\Gamma_i = i(\Sigma_i - \Sigma_i^\dagger)$ is known as a broadening function [114]. Performing a simple matrix multiplication allows Eq. (5.9) to be written as

$$\frac{A}{2\pi} = \frac{1}{2\pi} [G_S^+ \Gamma_L (G_S^+)^{\dagger} + G_S^+ \Gamma_R (G_S^+)^{\dagger}] \equiv \frac{1}{2\pi} [A_L + A_R],$$

where $A_i = G_S^+ \Gamma_i (G_S^+)^{\dagger}$ is the spectral function resulting from the coupling to the lead i . Out of equilibrium, Eq. (5.10) becomes

$$n_{2D} = \frac{1}{\pi} \text{diag} \left[\int_{-\infty}^{\infty} A_L f(E - E_{FL}) dE + \int_{-\infty}^{\infty} A_R f(E - E_{FR}) dE \right], \quad (5.12)$$

where E_{FL} and E_{FR} are the Fermi levels in the left and right leads, respectively. It now remains to calculate the Green's function and the self-energy matrices.

5.2 The Hamiltonian

Of course, if the Green's function is to be calculated, a suitable Hamiltonian must be defined. For infinite, periodic structures, it is convenient to employ Bloch's theorem and the tight-binding approximation [66], as discussed in Chapter 2. For a finite or semi-infinite crystal, where periodicity breaks near the ends, the similar Hückel approach [122] may be used. This could also apply to periodic crystals with aperiodic potentials applied to them.

In such aperiodic structures, each atomic site should be considered without appealing to Bloch's theorem. Of course, in the semi-infinite case, this results in a matrix of infinite size. However, as observed earlier, if the Green's function can be computed for this system, then the LDOS has also been obtained. Since a finite matrix may be inverted directly, this case will not be considered.

In a discretized form, the equation for the Green's function, without appealing to self-energies, is given by [94]

$$G^{\pm} = \lim_{\xi \rightarrow 0} [(E \pm i\xi) I - H]^{-1}.$$

The Hückel Hamiltonian looks very much like the tight-binding Hamiltonian derived earlier [122]. Using a suitable energy reference, considering a (2,0) CN for specificity, the Hamiltonian may be

written as

$$H = \begin{bmatrix} \ddots & & & & & & & & \\ \gamma & & U_8 & & \gamma & \gamma & & & \\ & \gamma & & U_7 & \gamma & \gamma & & & \\ & & \gamma & \gamma & U_6 & & \gamma & & \\ & & \gamma & \gamma & & U_5 & & \gamma & \\ & & & \gamma & & U_4 & & \gamma & \gamma \\ & & & & \gamma & & U_3 & \gamma & \gamma \\ & & & & & \gamma & \gamma & U_2 & \\ & & & & & \gamma & \gamma & & U_1 \end{bmatrix}, \quad (5.13)$$

such that the end of the tube is represented by elements in the bottom-right corner of the matrix. Note that the diagonal elements, U_i , are given by the local potential energy. Of course, the pattern that has been shown here is replicated along the diagonal out to infinity. The exact form of the matrix depends on the ordering of points, but this is the result for one particular ordering.

Since the Hückel derivation is similar to the tight-binding case, it is simply noted that by explicitly accounting for each atom, a matrix of a similar form as in tight-binding, but with no wavevector dependence, is obtained. The price is that there is now an infinite matrix to consider. Details may be found starting on page 627 in Ref. [122].

5.3 Computing the Green's Function

In this section, a technique for computing the Green's function [117] is presented. For an infinitely-long CN, a general version of the nearest-neighbour, tight-binding Hamiltonian of Eq. (5.13) may be written in the block form

$$H = \begin{bmatrix} \ddots & & & & & \\ \tau^\dagger & \Upsilon_L & \tau & & & \\ & \tau^\dagger & \Upsilon_S & \tau & & \\ & & \tau^\dagger & \Upsilon_R & \tau & \\ & & & \ddots & & \end{bmatrix},$$

where it may be recalled that a nanotube has periodicity in its atomic positions. Indeed, the following method could, in principle, be used for any crystalline material. The periodicity is defined by the

CN primitive unit cell, as discussed in Chapter 2, containing $2N$ atoms. Here, τ represents the coupling between these primitive unit cells, Υ represents both the local potential energy and the coupling between atoms within a primitive unit cell, and the subscripts L , S , and R represent the left lead, the scattering region, and the right lead, respectively. Note that all of these quantities are matrices, and that the left and right leads extend off to infinity. The magnitude of the atom-atom coupling is given by $|\gamma|$, and in this case, a nominal value of 2.8 eV is chosen as appropriate for the p_z orbitals [67]. In addition, note that rows corresponding to each unit cell of the leads repeat along the diagonal of H out to infinity due to the form of the potential energy there. Only potential energy variation normal to the transport direction is permitted in the leads, and any arbitrary potential variation is permitted in the scattering region.

The Green's function, which contains information about the LDOS [94, 114], is given by

$$G = (EI - H)^{-1},$$

where the notational change $G = G^+$ has been used, and the appropriate limit has been implicitly taken. Since the scattering region is the only relevant part of the calculation, the infinite matrix may be truncated using the method described in Ref. [94]. This allows the Green's function for the scattering region to be computed via

$$G_S = (EI - \Upsilon_S - \Sigma_L - \Sigma_R)^{-1},$$

where

$$\Sigma_L = \tau^\dagger G_L \tau, \tag{5.14}$$

$$\Sigma_R = \tau G_R \tau^\dagger \tag{5.15}$$

are the self-energies, and G_L and G_R are the Green's functions for the left and right leads, respectively. The self-energy matrices are finite, and as a result, the inverse may be computed directly, or by using a recursive algorithm [118]. For the self-energy calculations, two coupled, linear matrix equations may be solved simultaneously, as in Ref. [56]; however, a more efficient method is to solve a single, non-linear matrix equation.

Due to the tight-binding form of the Hamiltonian, the self-energies may be written in block form

as

$$\Sigma_L = \begin{bmatrix} \varsigma_L & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\Sigma_R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \varsigma_R \end{bmatrix},$$

where the diagonal elements, from upper-left to lower-right, have sizes $2N \times 2N$, $(n_S - 4N) \times (n_S - 4N)$, and $2N \times 2N$, respectively, where n_S is the total number of atoms in the scattering region, and it has been assumed, without loss of generality, that n_S is an integer multiple of $2N$. The LDOS is given by the diagonal elements [94, 114] of

$$G_S (\Sigma_L - \Sigma_L^\dagger + \Sigma_R - \Sigma_R^\dagger) G_S^\dagger,$$

and if G_S is written in block form as

$$G_S = \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix},$$

it may be observed that the second block column of G_S is irrelevant. As a result, it is sufficient to solve

$$(EI - \Upsilon_S - \Sigma_L - \Sigma_R) \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I \end{bmatrix},$$

for G_{ij} . The LDOS may, then, be computed as

$$D(\mathbf{r}; E) = \text{diag} \left[G_{i1}(\varsigma_L + \varsigma_L^\dagger) G_{i1}^\dagger + G_{i3}(\varsigma_R + \varsigma_R^\dagger) G_{i3}^\dagger \right],$$

where $i = 1, 2$, and 3 for atoms in the unit cell adjacent to the left lead, away from either lead, and adjacent to the right lead, respectively. It now remains to describe how one may compute G_L and G_R in order to obtain the self-energies from Eqs. (5.14) and (5.15).

5.3.1 Green's Function for a Semi-Infinite Lead

Previously, Appelbaum *et al.* have presented a method for calculating the required elements of G_L and G_R for the case where τ is a scalar [123], and Venugopal *et al.* have presented that for the case where $\tau = \tau^\dagger$ [124]. In the former work, the recursive algorithm [118] is implicitly employed to arrive at the solution, and this recursion may also be used for the general case considered here, as in Ref. [56]; however, as mentioned previously, by framing the problem as a quadratic matrix equation, the computation may be performed more efficiently.

G_L is the inverse of

$$EI - H_L = \begin{bmatrix} \ddots & & & \\ \tau^\dagger & EI - \Upsilon_L & & \tau \\ & \tau^\dagger & EI - \Upsilon_L & \tau \\ & & \tau^\dagger & EI - \Upsilon_L \end{bmatrix}.$$

Due to the infinite size of this matrix, this equation can also be written compactly as

$$EI - H_L = \begin{bmatrix} EI - H_L & \tau \\ \tau^\dagger & EI - \Upsilon_L \end{bmatrix}.$$

G_L is defined as

$$G_L = \begin{bmatrix} g_l & g_{ls} \\ g_{sl} & g_s \end{bmatrix} = \begin{bmatrix} EI - H_L & \tau \\ \tau^\dagger & EI - \Upsilon_L \end{bmatrix}^{-1}, \quad (5.16)$$

where g_s are the surface elements of G_L .

Eq. (5.16) may also be written as

$$\begin{bmatrix} EI - H_L & \tau \\ \tau^\dagger & EI - \Upsilon_L \end{bmatrix} \begin{bmatrix} g_l & g_{ls} \\ g_{sl} & g_s \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

This matrix equation can be solved to get

$$g_s = [EI - \Upsilon_L - \tau^\dagger(EI - H_L)^{-1}\tau]^{-1}.$$

Of course, this still requires the inversion of an infinite matrix. However, the nearest-neighbour, tight-binding structure of τ may be exploited in order to get that $\tau^\dagger(EI - H_L)^{-1}\tau = \tau^\dagger g_s \tau$. As a result, the problem now is to solve the nonlinear equation

$$g_s = [EI - \Upsilon_L - \tau^\dagger g_s \tau]^{-1}, \quad (5.17)$$

which may be most simply done using a fixed-point iterative scheme, as in Ref. [123], such as

$$g_s^{(i+1)} = [EI - \Upsilon_L - \tau^\dagger g_s^{(i)} \tau]^{-1},$$

where the superscript indicates the iteration number. This is a form of the recursive algorithm for self-energy calculations [118].

While this method does appear to converge for all cases attempted thus far, the convergence of the fixed-point, or recursive, method can be quite poor. As a result, the use of Newton's method on Eq. (5.17) is considered. The residual, r , is defined as

$$r(g_s) = \tau^\dagger g_s \tau g_s - (EI - \Upsilon_L)g_s + I,$$

which is equal to zero when g_s is the exact solution. This is a quadratic matrix equation, and the error matrix, ε , that may be applied to some approximate solution, \tilde{g}_s , is sought, *i.e.*,

$$r(\tilde{g}_s + \varepsilon) = \tau^\dagger (\tilde{g}_s + \varepsilon) \tau (\tilde{g}_s + \varepsilon) - (EI - \Upsilon_L)(\tilde{g}_s + \varepsilon) + I$$

is solved for ε , where $r(\tilde{g}_s + \varepsilon) = 0$. This leads to

$$r(\tilde{g}_s + \varepsilon) = r(\tilde{g}_s) + F_{\tilde{g}_s}(\varepsilon) + \tau^\dagger \varepsilon \tau \varepsilon,$$

where

$$F_{\tilde{g}_s}(\varepsilon) = \tau^\dagger \varepsilon \tau \tilde{g}_s + [\tau^\dagger \tilde{g}_s \tau - (EI - \Upsilon_L)]\varepsilon \quad (5.18)$$

is related to the Fréchet derivative of $r(\tilde{g}_s)$ in the direction ε . The Newton method is defined by solving $r(\tilde{g}_s) + F_{\tilde{g}_s}(\varepsilon) = 0$ for ε . As noted in Ref. [125], this is a special case of the generalized Sylvester equation, and may be solved using the method described in Ref. [126]. Due to the sensitivity of Newton's method to the initial guess, iterations are damped using exact line searches [125], where, instead of performing the usual Newton update of $g_s^{(i+1)} = g_s^{(i)} + \varepsilon$, the modified form of $g_s^{(i+1)} = g_s^{(i)} + \vartheta \varepsilon$ is used, where ϑ is a parameter chosen such that the square of the Frobenius norm of $r(g_s^{(i)} + \vartheta \varepsilon)$ is minimized at each step. Details of this procedure are provided in Ref. [125], and in Appendix E. In general, exact line searches are required in order to obtain convergence with Newton's method.

Due to the increased computational effort per iteration for this method, the iterative solution begins, in practice, with fixed-point iterations for a while, and then switches to Newton with exact line searches. Typically, only a few Newton iterations are required in this hybrid approach.

A completely analogous procedure holds for the calculation of the surface elements of G_R , except that τ is exchanged with τ^\dagger , and vice versa, in all of the above derivation.

5.3.2 Test Cases

The simplest case with which to check this method is for the case where the local potential is some fixed constant. This case is a periodic structure extending out to infinity in both directions, and therefore, the Bloch condition may be applied in order to compute the DOS using a single, primitive CN unit cell. Equivalently, one may employ the zone-folding scheme in order to obtain the DOS from the dispersion relation for graphene (see Refs. [4, 6], for example). Both of these methods are described in detail in Chapter 2. In this case, $\Upsilon_L = \Upsilon_S = \Upsilon_R$, and the NEGF method is used to compute the LDOS from G_S [94, 114]. In order to numerically handle the van Hove singularities in the CN DOS, energies were perturbed by a small, positive imaginary number, $i\xi$. For the results shown here, the perturbation was $\xi = 2 \times 10^{-6}$, and convergence was declared when the infinity norm of $r(\tilde{g}_s)$ became less than 10^{-6} . As shown in Fig. 5.1, this method produces the correct form.

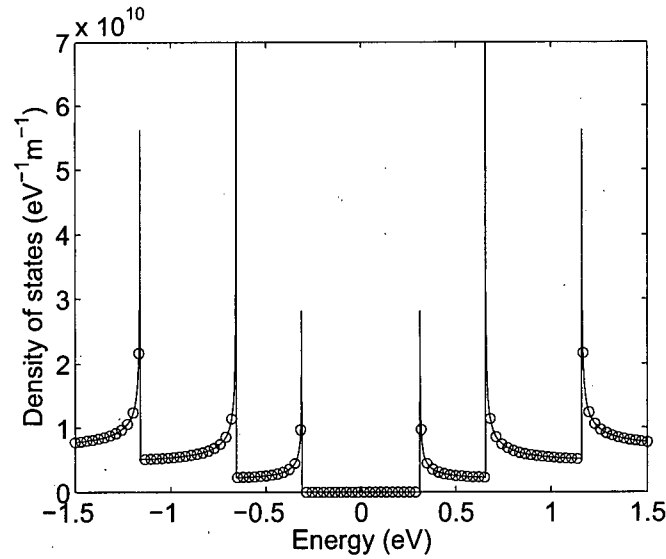
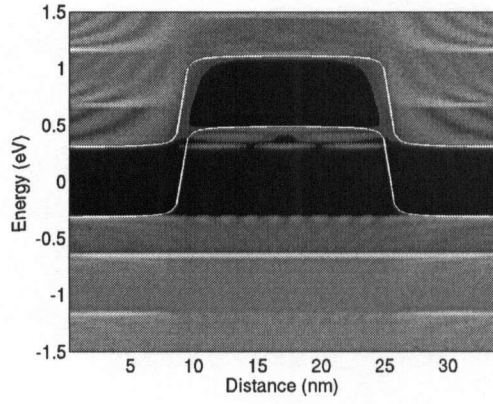


Figure 5.1: Comparison of the standard tight-binding density of states (solid line) with that computed using the self-energy method (\circ) for a (16,0) carbon nanotube [117].

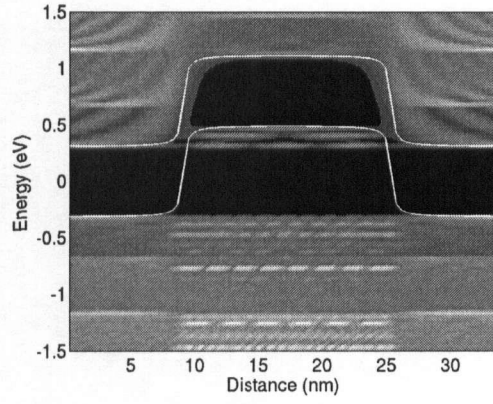
The second case to consider is that of a finite barrier. In this case, quantum mechanical reflection for carriers above the barrier, and the possibility of tunneling through the barrier, are expected. These phenomena should modify the form of the LDOS at each position. Fig. 5.2(a) shows the LDOS for this case, and the solid lines show the conduction and valence band edges shifted by the local potential energy. Regions of high electron concentration are shown as bright patches, and since the Green's function represents outgoing waves at infinity, a vanishing density is obtained in the scattering region for energies lined up with the bandgap in the leads. Also note that the phenomena of inter-band tunneling, which may have important technological applications [63, 127], is clearly shown in the figure. The jumps in density with energy are due to the contributions of higher bands, and intra-band tunneling is also evinced by the penetration of carriers into the bandgap region. It is noted, here, that in Ref. [114] a subsequent summation was required for each transverse mode, or band, and that these modes were assumed to be independent. In this case, a single calculation provides a complete multi-band result with allowed inter-band transitions.

Finally, the LDOS for a potential that varies both azimuthally, and along the length of the CN, is computed. The longitudinal potential variation is multiplied by $1 + 0.2 \sin(\phi)$. Fig. 5.2(b)–(d) shows the resultant LDOS at $\phi = 0, \pi/2$, and $3\pi/2$ radians, respectively. While the variation with angle is readily identifiable, particularly in the valence band, it is interesting to note that the local band structure appears to be more influenced by the average potential around the circumference than by the local potential. This may be seen by the superimposed conduction and valence band edges, which have simply been shifted by the local potential energy. The LDOS itself does not undergo such a shift directly; rather, its magnitude is modified, with the gap apparently taking the position appropriate for the average energy shift, *i.e.*, the shift at $\phi = 0$ in this case. This may bode well for methods that utilize an average potential on asymmetric structures [55, 56]. Careful examination reveals that the mid-length conduction band LDOS is highest for $\phi = 3\pi/2$, which corresponds to the angle with the lowest potential energy, followed by $\phi = 0$, and then $\phi = \pi/2$. This is the expected result; the unexpected result is that the LDOS is not rigidly shifted by the local potential energy.

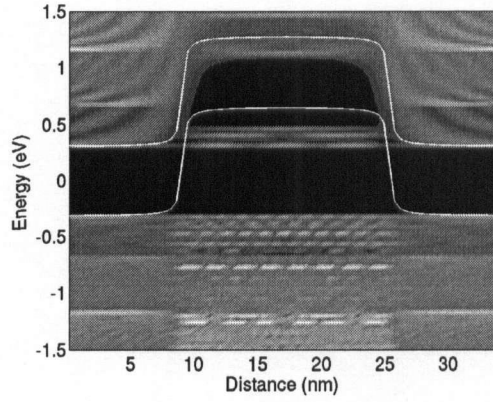
While this method yields excellent results, further analysis is required due to the possibility of a multiplicity of solutions that may satisfy the quadratic matrix equation. It is well-known that equations of this type may have any number of solutions, including no solution at all. At present, it can only be stated that there is at least one solution to the equation, and that it is a physically



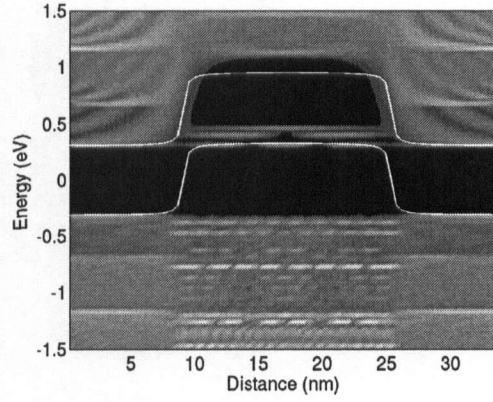
(a)



(b)



(c)



(d)

Figure 5.2: Local density of states for longitudinal and azimuthal potential variation. Shown are: (a) a barrier potential with a maximum shift of 0.8 eV longitudinally, and no azimuthal variation; the same barrier potential as in (a) multiplied by $1 + 0.2 \sin(\phi)$ and evaluated at angular positions of (b) 0, (c) $\pi/2$, and (d) $3\pi/2$ radians. Solid lines show the result of shifting the conduction and valence bands by the local potential energy.

reasonable one.

5.4 Current Operators

Ultimately, the current-voltage relationship is of interest. Consider, then, how the real-space current operator may be derived when considering quantum transport. Both effective-mass and tight-binding Hamiltonians are considered in order to better justify the tight-binding case by comparing it to a well-known result. One- and two-dimensional operators will be derived on a Cartesian grid, and the conclusion will involve the derivation of an expression for the current in a particular CN with a tight-binding Hamiltonian. The CN case is treated separately since the Hamiltonian allows for coupling between atomic sites that are not arranged on an orthogonal mesh. As a result, the CN matrix will differ from that of atoms on a rectangular or square crystal lattice.

The derivation begins with the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \quad (5.19)$$

where t is time. In addition, there is the conjugate equation,

$$i\hbar \frac{\partial \Psi^\dagger}{\partial t} = -(H\Psi)^\dagger. \quad (5.20)$$

Following the approach of Ref. [128], note that

$$\frac{\partial}{\partial t}(\Psi^\dagger \Psi) = \Psi^\dagger \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^\dagger}{\partial t} \Psi.$$

Using Eqs. (5.19) and (5.20),

$$\frac{\partial}{\partial t}(\Psi^\dagger \Psi) = \frac{1}{i\hbar} [\Psi^\dagger H\Psi - (H\Psi)^\dagger \Psi]. \quad (5.21)$$

An expression similar to this occurs in hydrodynamics if the associations

$$\begin{aligned} \tilde{\rho} &= \Psi^\dagger \Psi, \\ \nabla \cdot \tilde{\mathbf{J}} &= -\frac{1}{i\hbar} [\Psi^\dagger H\Psi - (H\Psi)^\dagger \Psi], \end{aligned} \quad (5.22)$$

are made, where $\tilde{\rho}$ is like a fluid density, and $\tilde{\mathbf{J}}$ is a probability current density. The tilde is a reminder that these quantities are functions of energy. In order to obtain the total probability current density \mathbf{J} , $\tilde{\mathbf{J}}$ must be integrated over E . To convert $\tilde{\mathbf{J}}$ to an electronic current density, simply multiply by the charge of a carrier q . It now remains to discover the appropriate expression for $\tilde{\mathbf{J}}$. The goal is to write the right-hand side of Eq. (5.22) as a divergence expression.

5.4.1 Effective-Mass

In the effective-mass approximation, the explicit form for the Hamiltonian is

$$H = -\frac{\hbar^2}{2m_{\text{eff}}} \nabla^2 + U(\mathbf{r}).$$

Since the potential energy is real, Eq. (5.22) can be written as

$$\nabla \cdot \tilde{\mathbf{J}} = \frac{\hbar}{2im_{\text{eff}}} [\Psi^\dagger \nabla^2 \Psi - (\nabla^2 \Psi^\dagger) \Psi].$$

If it is recalled that $\nabla^2 \Psi \equiv \nabla \cdot (\nabla \Psi)$, this becomes

$$\nabla \cdot \tilde{\mathbf{J}} = \frac{\hbar}{2im_{\text{eff}}} [\Psi^\dagger \nabla \cdot (\nabla \Psi) - \nabla \cdot (\nabla \Psi^\dagger) \Psi].$$

A standard vector identity is

$$\nabla \cdot (\Psi^\dagger \nabla \Psi) = \Psi^\dagger \nabla \cdot (\nabla \Psi) + \nabla \Psi \cdot \nabla \Psi^\dagger,$$

and similarly,

$$\nabla \cdot (\Psi \nabla \Psi^\dagger) = \Psi \nabla \cdot (\nabla \Psi^\dagger) + \nabla \Psi^\dagger \cdot \nabla \Psi.$$

Subtracting these identities implies that

$$\nabla \cdot \tilde{\mathbf{J}} = \frac{\hbar}{2im_{\text{eff}}} \nabla \cdot [\Psi^\dagger \nabla \Psi - \Psi \nabla \Psi^\dagger].$$

Therefore, the current density can be defined as

$$\tilde{\mathbf{J}} = \frac{\hbar}{2im_{\text{eff}}} [\Psi^\dagger \nabla \Psi - \Psi \nabla \Psi^\dagger] = \text{Re} \left[\frac{\hbar}{im_{\text{eff}}} \Psi^\dagger \nabla \Psi \right]. \quad (5.23)$$

Note that this expression is not unique. One could, for example, add any function to $\tilde{\mathbf{J}}$ that has a vanishing divergence, and still obtain Eq. (5.22). Here, an appeal is made to the correspondence principle, which states, from page 29 of Ref. [128], “Quantum Theory must approach Classical Theory asymptotically in the limit of large quantum numbers.” In order to satisfy this condition, “one establishes in principle that there exists a formal analogy between Quantum Theory and Classical Theory.” In this case, it is noted that the formal analogy of the classical momentum is the operator $(\hbar/i)\nabla$. As a result, for a plane-wave wavefunction, the form of Eq. (5.23) is analogous to the velocity multiplied by the density, as in the classical current. An additive term with a vanishing divergence is, therefore, not necessary.

5.4.2 Tight-Binding

In the tight-binding approach, the right-hand side of Eq. (5.22) will, again, be written as a divergence expression. As in the effective-mass case, it will be assumed that the current density is given by this expression without adding any function with a vanishing divergence to the result. Previously, an appeal was made to the correspondence principle. Here, the appeal is made to the discrete version of Eq. (5.23), and no attempt is made at trying to present a formal proof. It is assumed, throughout, that the atomic spacing, a , is uniform.

The first step is to discretize Eq. (5.21). If

$$|\Psi\rangle = \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \vdots \\ \Psi_N \end{bmatrix},$$

$$\langle\Psi| = [\Psi_1^\dagger \quad \Psi_2^\dagger \quad \Psi_3^\dagger \quad \dots \quad \Psi_N^\dagger],$$

where Ψ_i is the wavefunction at the i^{th} atom, and N is the total number of atoms in the system, then Eq. (5.19), and its dual equation, may be written as

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = H |\Psi\rangle,$$

$$i\hbar \frac{\partial}{\partial t} \langle\Psi| = -\langle\Psi| H^\dagger.$$

Proceeding as before,

$$\frac{\partial}{\partial t} (|\Psi\rangle \langle\Psi|) = \frac{1}{i\hbar} (H |\Psi\rangle \langle\Psi| - |\Psi\rangle \langle\Psi| H^\dagger), \quad (5.24)$$

where it is noted that the diagonal elements of this matrix equation are the discrete equivalent of evaluating Eq. (5.21) at the lattice sites.

In order to make the right-hand side of Eq. (5.24) look like a divergence expression, it is helpful to recall the formal definition of the divergence [129],

$$\nabla \cdot \tilde{\mathbf{J}} = \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} \int_s \tilde{\mathbf{J}} \cdot \hat{\mathbf{n}} d^m r, \quad (5.25)$$

where m is one less than the number of spatial dimensions of $\tilde{\mathbf{J}}$, Δ is the volume enclosed by the surface s which surrounds the point at which you are calculating the divergence, and $\hat{\mathbf{n}}$ is a unit

vector that points in the outward-normal direction to the infinitesimal area element $d^m r$. If working in dimensions different than three, the appropriate generalizations of the terms volume, area, and surface should be employed.

Cartesian Lattice

First, consider the case where the lattice of atoms forms a Cartesian grid, as in Fig. 5.3. The tight-binding Hamiltonian is assumed to include only the effects of the nearest neighbours.

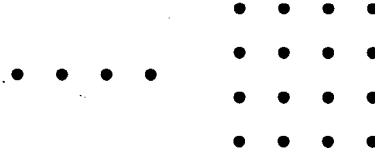


Figure 5.3: Sample one- and two-dimensional lattices on the left and right, respectively.

In one dimension, the Hamiltonian can be written as a tri-diagonal matrix of the form

$$H = \begin{bmatrix} 0 & \gamma & 0 & \cdots & 0 & 0 & 0 \\ \gamma & 0 & \gamma & & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & \gamma & 0 & \gamma \\ 0 & 0 & 0 & \cdots & 0 & \gamma & 0 \end{bmatrix},$$

where all the real, diagonal elements; such as the local potential energy, have been eliminated since they will cancel out in the diagonal elements of Eq. (5.24).

Substituting the 1D, tight-binding Hamiltonian into Eq. (5.24) yields diagonal elements, away from the ends, of the form

$$\frac{\partial}{\partial t} (\Psi_i \Psi_i^\dagger) = \frac{\gamma}{i\hbar} (\Psi_{i-1} \Psi_i^\dagger + \Psi_{i+1} \Psi_i^\dagger - \Psi_i \Psi_{i-1}^\dagger - \Psi_i \Psi_{i+1}^\dagger).$$

A 1D approximation to Eq. (5.25) is

$$\nabla \cdot \tilde{\mathbf{J}} \simeq \frac{1}{a} (\tilde{J}_{i+\frac{1}{2}} - \tilde{J}_{i-\frac{1}{2}}),$$

so

$$\tilde{J}_{i+\frac{1}{2}} = -\frac{\gamma a}{i\hbar} (\Psi_{i+1} \Psi_i^\dagger - \Psi_i \Psi_{i+1}^\dagger),$$

$$\tilde{J}_{i-\frac{1}{2}} = -\frac{\gamma a}{i\hbar} \left(\Psi_i \Psi_{i-1}^\dagger - \Psi_{i-1} \Psi_i^\dagger \right),$$

are defined, where $\tilde{J}_{i+\frac{1}{2}}$ is the magnitude of the current density between atoms i and $i+1$.

It now remains to compare these expressions with the effective-mass result. Using the finite-difference approximation, one can view the tight-binding Hamiltonian as an approximation to the effective-mass Hamiltonian with

$$\gamma = -\frac{\hbar^2}{2a^2 m_{\text{eff}}}. \quad (5.26)$$

Using finite differences on the effective-mass current expression leads to

$$\begin{aligned} \tilde{J}_{i+\frac{1}{2}} &\simeq \frac{\hbar}{2im_{\text{eff}}} \left[\left(\frac{\Psi_{i+1}^\dagger + \Psi_i^\dagger}{2} \right) \left(\frac{\Psi_{i+1} - \Psi_i}{a} \right) - \left(\frac{\Psi_{i+1} + \Psi_i}{2} \right) \left(\frac{\Psi_{i+1}^\dagger - \Psi_i^\dagger}{a} \right) \right] \\ &= \frac{\hbar}{2iam_{\text{eff}}} \left(\Psi_{i+1} \Psi_i^\dagger - \Psi_i \Psi_{i+1}^\dagger \right), \end{aligned}$$

where it has been assumed that the wavefunction between sites i and $i+1$ is equal to the average of the wavefunctions at the nearest atomic sites. Using Eq. (5.26), it becomes apparent that the tight-binding result is equivalent to the effective-mass result. Of course, this equivalence is expected since the tight-binding Hamiltonian may be viewed as an approximation to the effective-mass Hamiltonian when working with atoms on a uniform, Cartesian grid.

Using a 2D, tight-binding Hamiltonian, the result is that

$$\begin{aligned} \frac{\partial}{\partial t} \left(\Psi_{i,j} \Psi_{i,j}^\dagger \right) &= \frac{\gamma}{i\hbar} \left(\Psi_{i-1,j} \Psi_{i,j}^\dagger + \Psi_{i+1,j} \Psi_{i,j}^\dagger - \Psi_{i,j} \Psi_{i-1,j}^\dagger - \Psi_{i,j} \Psi_{i+1,j}^\dagger \right. \\ &\quad \left. + \Psi_{i,j-1} \Psi_{i,j}^\dagger + \Psi_{i,j+1} \Psi_{i,j}^\dagger - \Psi_{i,j} \Psi_{i,j-1}^\dagger - \Psi_{i,j} \Psi_{i,j+1}^\dagger \right), \end{aligned}$$

where a double subscript to the wavefunction, $\Psi_{i,j}$, is now required in order to denote the i^{th} site in the x -direction, and the j^{th} in the y -direction.

The 2D approximation to Eq. (5.25) is

$$\nabla \cdot \tilde{\mathbf{J}} \simeq \frac{1}{a} \left(\tilde{J}_{i+\frac{1}{2},j} - \tilde{J}_{i-\frac{1}{2},j} + \tilde{J}_{i,j+\frac{1}{2}} - \tilde{J}_{i,j-\frac{1}{2},j} \right),$$

where the surface, s , is defined as a square with area a^2 centred on the site at (i,j) . Note that it is assumed that the normal component of the current density is constant along each of the four sides of s , so the integral is just the normal component of the current density multiplied by a . This leads to similar definitions for \tilde{J} ,

$$\tilde{J}_{i+\frac{1}{2},j} = -\frac{\gamma a}{i\hbar} \left(\Psi_{i+1,j} \Psi_{i,j}^\dagger - \Psi_{i,j} \Psi_{i+1,j}^\dagger \right),$$

$$\begin{aligned}\tilde{J}_{i-\frac{1}{2},j} &= -\frac{\gamma a}{i\hbar} \left(\Psi_{i,j} \Psi_{i-1,j}^\dagger - \Psi_{i-1,j} \Psi_{i,j}^\dagger \right), \\ \tilde{J}_{i,j+\frac{1}{2}} &= -\frac{\gamma a}{i\hbar} \left(\Psi_{i,j+1} \Psi_{i,j}^\dagger - \Psi_{i,j} \Psi_{i,j+1}^\dagger \right), \\ \tilde{J}_{i,j-\frac{1}{2}} &= -\frac{\gamma a}{i\hbar} \left(\Psi_{i,j} \Psi_{i,j-1}^\dagger - \Psi_{i,j-1} \Psi_{i,j}^\dagger \right),\end{aligned}$$

where $\tilde{J}_{i+\frac{1}{2},j}$ is the magnitude of the current density component in the y -direction crossing the plane between rows i and $i+1$, and in the j^{th} column, of atoms. $\tilde{J}_{i,j+\frac{1}{2}}$ is defined similarly. By direct analogy to the 1D case, it is concluded that this expression is also equivalent to the effective-mass case.

Hexagonal Lattice

In a hexagonal lattice, such as in a CN, there is a two-point basis. As a result, two different equations emerge, depending on which atom in the basis is being considered. The primitive unit cell of the hexagonal lattice is a parallelogram [66]. If this unit cell is divided into two equilateral triangles, a convenient s over which to approximate Eq. (5.25) is found. The hexagonal lattice, primitive unit cell, and s are shown in Fig. 5.4.

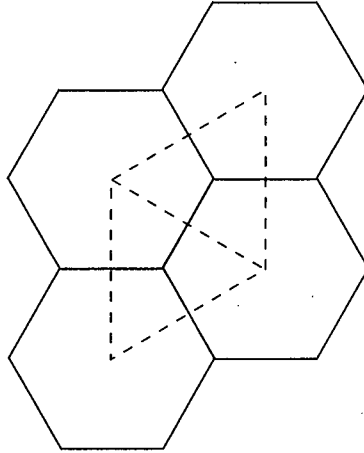


Figure 5.4: Hexagonal lattice (solid) with a primitive unit cell (dashed parallelogram) divided into two equilateral triangles s .

The appropriate approximation to Eq. (5.25), using the same assumptions as before, and con-

sidering the higher of the two atoms, denoted by a superscript h , in the primitive unit cell, is

$$\nabla \cdot \tilde{\mathbf{J}}^h \simeq \frac{4}{3a} \left(\tilde{J}_{\rightarrow}^h + \tilde{J}_{\nwarrow}^h + \tilde{J}_{\nearrow}^h \right),$$

where the subscript indicates the direction of the surface through which the current passes. Similarly, for the lower atom, denoted by a superscript l ,

$$\nabla \cdot \tilde{\mathbf{J}}^l \simeq \frac{4}{3a} \left(\tilde{J}_{\leftarrow}^l + \tilde{J}_{\nearrow}^l + \tilde{J}_{\nwarrow}^l \right).$$

Utilizing the tight-binding Hamiltonian,

$$\begin{aligned} \frac{\partial}{\partial t} [\Psi^h(\Psi^h)^\dagger] &= \frac{\gamma}{i\hbar} [\Psi_{\rightarrow}^h(\Psi^h)^\dagger + \Psi_{\nwarrow}^h(\Psi^h)^\dagger + \Psi_{\nearrow}^h(\Psi^h)^\dagger - \\ &\quad \Psi^h(\Psi_{\rightarrow}^h)^\dagger - \Psi^h(\Psi_{\nwarrow}^h)^\dagger - \Psi^h(\Psi_{\nearrow}^h)^\dagger], \\ \frac{\partial}{\partial t} [\Psi^l(\Psi^l)^\dagger] &= \frac{\gamma}{i\hbar} [\Psi_{\leftarrow}^l(\Psi^l)^\dagger + \Psi_{\nearrow}^l(\Psi^l)^\dagger + \Psi_{\nwarrow}^l(\Psi^l)^\dagger - \\ &\quad \Psi^l(\Psi_{\leftarrow}^l)^\dagger - \Psi^l(\Psi_{\nearrow}^l)^\dagger - \Psi^l(\Psi_{\nwarrow}^l)^\dagger]. \end{aligned}$$

Therefore,

$$\tilde{J}_{\rightarrow}^h = -\frac{3\gamma a}{4i\hbar} [\Psi_{\rightarrow}^h(\Psi^h)^\dagger - \Psi^h(\Psi_{\rightarrow}^h)^\dagger],$$

is defined, and similarly for the other current densities.

5.4.3 Current Operator Summary

After all of this work, a summary of the current operators, \mathcal{J} , is presented, where it is noted that it has been implicitly assumed that Ψ has been normalized to give the electron density at a given atomic site, *i.e.*, if $\psi^\dagger\psi$ gives the electron concentration per atom, then $\Psi^\dagger\Psi = \psi^\dagger\psi/\Lambda$. In this section, the Λ normalization is explicitly included.

1. The effective-mass current density is given by

$$\tilde{\mathbf{J}} = \frac{i\hbar}{2m} (\Psi \nabla \Psi^\dagger - \Psi^\dagger \nabla \Psi).$$

2. The 1D, tight-binding current density operator is given by

$$\mathcal{J}_{ij} = -\frac{\gamma a}{i\hbar} \begin{cases} 1 & , \quad j = i + 1, \\ 0 & , \quad j \neq i + 1, \end{cases} \quad (5.27)$$

where \mathcal{J}_{ij} is the element in the i^{th} row and j^{th} column of \mathcal{J} . The current density may be written as

$$\tilde{\mathbf{J}} = \text{diag} (\mathcal{J} |\Psi\rangle \langle \Psi| + |\Psi\rangle \langle \Psi| \mathcal{J}^\dagger), \quad (5.28)$$

where

$$\tilde{\mathbf{J}} = \begin{bmatrix} \tilde{J}_{\frac{3}{2}} \\ \tilde{J}_{\frac{5}{2}} \\ \tilde{J}_{\frac{7}{2}} \\ \vdots \end{bmatrix}.$$

3. In two dimensions, assume that the elements of $|\Psi\rangle$ are ordered by increasing x first, *i.e.*,

$$|\Psi\rangle = \begin{bmatrix} \Psi_{1,1} \\ \Psi_{2,1} \\ \vdots \\ \Psi_{N_x,1} \\ \Psi_{1,2} \\ \Psi_{2,2} \\ \vdots \end{bmatrix},$$

where N_x is the number of points in the x -direction. With this ordering, the current operator for the x -direction is also given by Eqs. (5.27) and (5.28), where $\tilde{\mathbf{J}}$ is ordered similarly to $|\Psi\rangle$. The only difference is that, for a finite number of atoms in x , the matrix will have only zeros in the rows corresponding to the last atom in that direction. A similar expression holds for the y -component of $\tilde{\mathbf{J}}$.

4. The current operator for a tight-binding Hamiltonian on a CN hexagonal lattice depends on the assumed chirality due to the periodicity in the circumferential direction, and it also depends on which atom is being considered in the two-point basis. Rather than trying to present all of the various permutations that are possible for a general CN, a demonstration of how one may compute the current through the cross-section of the CN, depicted in Fig. 5.5, is presented. From the current density expression, it may be seen that the current across the dashed line will be related to the wavefunctions at only four atomic sites. If a subscript denotes which

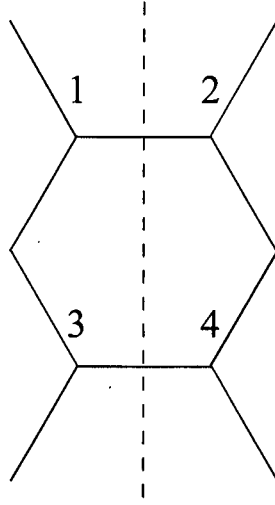


Figure 5.5: Atomic structure of a (2,0) carbon nanotube for which a calculation of the current across the dashed line is desired. Note that the atoms on the top of the figure are identical to those on the bottom due to the periodicity of the tube.

atomic site is being considered, the current, \tilde{I} , may be written as

$$\tilde{I} = -\frac{3q\gamma a^2\sqrt{3}}{4i\hbar} \left[\Psi_2\Psi_1^\dagger - \Psi_1\Psi_2^\dagger + \Psi_4\Psi_3^\dagger - \Psi_3\Psi_4^\dagger \right],$$

where spin degeneracy has been accounted for in the normalization of the wavefunction, and the tilde, again, implies that this quantity should be integrated over energy to give the total current. In order to get this result, it has been assumed that the current is constant over each face of the triangle depicted in Fig. 5.4, and an integration has been performed over the dashed line in Fig. 5.5.

5.5 Alternate Current Expression

In the previous section, current operators were derived that provide the current density spectrum as a function of position and energy. While this will be interesting in terms of being able to inspect the path of carrier flow in the device, the terminal characteristics are of primary importance. In this case, a compact expression may be derived for the terminal current following the approach detailed

in Ref. [130]. This compact expression is commonly used in mesoscopic physics.

The Schrödinger equation, in matrix form, is

$$\begin{bmatrix} (E + i\xi)I - H_L & -\tau & \\ -\tau^\dagger & (E + i\xi)I - H_S & -\tau \\ & -\tau^\dagger & (E + i\xi)I - H_R \end{bmatrix} \begin{bmatrix} \Psi_L \\ \Psi_S \\ \Psi_R \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

where $\Psi_{L,S,R} = \psi_{L,S,R}/\sqrt{\Lambda}$ is the wavefunction in each of the three regions under consideration, *i.e.*, the left lead, the scattering region, and the right lead, respectively; and it is recalled that the Green's function involves the limit as ξ approaches 0^+ . Using ψ_i from this point on, it is convenient to let $\psi_L = \varphi_L + \chi_L$ and $\psi_R = \varphi_R + \chi_R$, where φ_i satisfies

$$[(E + i\xi)I - H_i]\varphi_i = 0,$$

and represents the wavefunction in the lead before coupling to the scattering region. After coupling, the unperturbed wavefunction is modified by χ_i . Performing some algebra yields

$$\psi_S = [(E + i\xi)I - H_S - \Sigma_L - \Sigma_R]^{-1} [\tau^\dagger \varphi_L + \tau \varphi_R] \equiv G_S [\tau^\dagger \varphi_L + \tau \varphi_R], \quad (5.29)$$

$$\chi_L = G_L \tau \psi_S, \quad (5.30)$$

$$\chi_R = G_R \tau^\dagger \psi_S. \quad (5.31)$$

From Eq. (5.24),

$$\frac{\partial}{\partial t} \text{trace}(\psi_S \psi_S^\dagger) = \frac{1}{i\hbar} \text{trace}(\tau^\dagger \psi_L \psi_S^\dagger - \psi_S \psi_L^\dagger \tau + \tau \psi_R \psi_S^\dagger - \psi_S \psi_R^\dagger \tau^\dagger) = - \int_S \nabla \cdot \tilde{\mathbf{J}} d^2r,$$

since $H_S = H_S^\dagger$, and the trace, *i.e.*, the sum of the diagonal elements, is not dependent on the order of the matrix multiplications in its argument. Note that taking the trace over the scattering region elements is the discrete equivalent of integrating over the scattering region S up to a normalization factor of Λ . This factor cancels out when working with ψ_i instead of Ψ_i .

Utilizing the divergence theorem, and noting that current may only travel out through the leads, yields

$$-\tilde{\mathcal{I}}_L + \tilde{\mathcal{I}}_R = -\frac{q}{i\hbar} \text{trace}(\tau^\dagger \psi_L \psi_S^\dagger - \psi_S \psi_L^\dagger \tau + \tau \psi_R \psi_S^\dagger - \psi_S \psi_R^\dagger \tau^\dagger),$$

where $\tilde{\mathcal{I}}_L$ and $\tilde{\mathcal{I}}_R$ are the currents out of the scattering region at the left and right interfaces. Clearly, $\tilde{\mathcal{I}}_L = \tilde{\mathcal{I}}_R$ in steady-state, since the time derivative of $\psi^\dagger \psi$ vanishes. Of course, this implicitly assumes

that all of the charge in the scattering region is due to the leads, and hence there is no recombination or generation allowed in the scattering region.

Considering only the left lead,

$$\tilde{I}_L = \frac{q}{i\hbar} \text{trace} \left(\tau^\dagger \varphi_L \psi_S^\dagger + \tau^\dagger \chi_L \psi_S^\dagger - \psi_S \varphi_L^\dagger \tau - \psi_S \chi_L^\dagger \tau \right).$$

In order to write this in a more convenient form, substitute Eqs. (5.29)–(5.31) into this expression.

If it is noted that [130]

$$\varphi_i^\dagger \varphi_j = \frac{f_i A_i}{2\pi} \delta_{ij},$$

where f_i is the distribution function in the i^{th} lead, the result is, after some algebra,

$$\tilde{I}_L = \frac{q}{h} \text{trace} (f_L \Gamma_L A_R - f_R \Gamma_R A_L).$$

This may be simplified even further if it is noted that [130]

$$\text{trace} (\Gamma_L A_R) = \text{trace} (\Gamma_L G_S \Gamma_R G_S^\dagger) = \text{trace} (G_S \Gamma_L G_S^\dagger \Gamma_R) = \text{trace} (\Gamma_R A_L).$$

The final expression for the current is [114, 130]

$$\tilde{I}_L = \frac{q}{h} \text{trace} [\Gamma_L A_R (f_L - f_R)], \quad (5.32)$$

which should be multiplied by two if spin degeneracy is included. The same expression results from considering the right contact. In this form, Eq. (5.32) bears a striking resemblance to the Landauer current formula [93]. If the transmission probability for the scattering region is associated with $\text{trace} (\Gamma_L A_R)$ [114, 130], then the formulae are identical. Use of this compact expression is convenient since only diagonal elements of matrices that are already known from the charge calculation are required. In order to calculate the inter-atomic current density, further computation of off-diagonal elements would be needed.

5.6 Solution Procedure

In order to arrive at converged solutions, two algorithms have been implemented. The first is related to the standard Gummel technique as described, for example, in Ref. [131]. This method, when it converges, is useful due to its good convergence rate [56, 58]; however, it is found that a good solution

is actually less likely to be found with this method than with an alternate Newton iterative scheme with an approximate Jacobian [132]. Doped-contact CNFETs are considered, with a potential energy in the scattering region that joins smoothly to that in the leads. The scattering region, thus, includes some length of the doped contacts.

The Gummel iterative scheme proceeds as follows:

1. Ensure that an integral number of primitive unit cells are in the scattering region. This is really just a bookkeeping issue, and ensures that the leads have symmetric structures. This is useful when computing coupling matrices. Since the leads have no potential variation in the transport direction, the amount of doped contact included in the scattering region may be increased without changing the solution.
2. Tetrahedralize the domain for the finite element Poisson solver. In practice, this is done using the freeware program TetGen [102].
3. Compute the Hamiltonian, and atomic positions, for atoms in the scattering region. The Hamiltonian is a sparse matrix with non-zero entries only for nearest-neighbour atoms. The atomic positions are needed in order to interpolate between the atomic grid, used for the Green's function calculation, and the Poisson grid.
4. Compute the Hamiltonian for one CN unit cell of the leads. This is one block of the matrix computed in the previous step.
5. Compute the stiffness matrix for the finite element Poisson solution. Details of this calculation are provided in the Finite Element Electrostatics section of Chapter 4.
6. Guess initial quasi-Fermi levels for holes and electrons, E_{Fp} and E_{Fn} , respectively, for use in the nonlinear Poisson equation detailed in Step 7.
7. Solve a simplified nonlinear Poisson equation. This is done to improve the convergence properties of the overall iterative scheme [56, 58, 131]. The idea is to solve

$$\nabla^2 V = -\frac{q}{\epsilon} \delta(\rho - R_t) (p_{2D}(V) - n_{2D}(V) + N_f),$$

where

$$p_{2D} = \int_0^\infty \frac{D(E)}{\exp\left(\frac{E+qV+E_{Fp}}{k_B T}\right) + 1} dE, \quad (5.33)$$

$$n_{2D} = \int_0^\infty \frac{D(E)}{\exp\left(\frac{E - qV - E_{Fn}}{k_B T}\right) + 1} dE, \quad (5.34)$$

and $D(E)$ is the 2D density of states, based on symmetric electron and hole bands. In practice, the analytic $D(E)$ of Ref. [133] divided by the CN circumference is used. This Poisson equation may be straightforwardly solved using Newton's method. Note that, on subsequent iterations, E_{Fp} and E_{Fn} are calculated such that $p_{2D}(V)$ and $n_{2D}(V)$ are equal to the carrier concentrations computed via the Green's function. If the charge and potential are self-consistent, the solution to this nonlinear equation will converge in the first step.

8. Solve for the Green's function in the scattering region in order to obtain the charge there via Eq. (5.12). Note that the net charge density is given by $q(1 - n_{2D})$ where the first term represents the ionic charge left over if no p_z electrons are present. In order to facilitate calculation of the integrals in Eq. (5.12), the first term can be written as

$$\frac{1}{\pi} \text{diag} \left[\int_{-\infty}^{E_N} A_L dE - \int_{-\infty}^{E_N} A_L f(E_{FL} - E) dE + \int_{E_N}^{\infty} A_L f(E - E_{FL}) dE \right],$$

and similarly for the second term, where E_N is the "charge neutrality level" defined such that the integral of the total density of states up to this energy yields the atomic density Λ , where it is noted that the total LDOS is related to $A_L + A_R$, and that $\Lambda = 1$. Using this yields

$$1 - n_{2D} = -\frac{1}{\pi} \text{diag} \left\{ \int_{-\infty}^{\infty} \frac{\text{sgn}(E - E_N(z)) A_L}{\exp\left[\text{sgn}(E - E_N(z)) \left(\frac{E - E_{FL}}{k_B T_K}\right)\right] + 1} dE + \int_{-\infty}^{\infty} \frac{\text{sgn}(E - E_N(z)) A_R}{\exp\left[\text{sgn}(E - E_N(z)) \left(\frac{E - E_{FR}}{k_B T_K}\right)\right] + 1} dE \right\}, \quad (5.35)$$

where $E_N(z)$ is shifted by the local potential energy, and $\text{sgn}(\cdot)$ is the sign function. Using the charge neutrality level, and the sign function to obtain more compact and computationally tractable carrier integrals has been previously seen in Refs. [23, 56, 134], for example. It is sometimes convenient to associate $E - E_N(z) < 0$ with p_{2D} and $E - E_N(z) > 0$ with n_{2D} , particularly in the Gummel method where quasi-Fermi levels for both holes and electrons may require computation.

9. Compute the quasi-Fermi levels corresponding to the electron and hole concentrations from Step 8. This is done using Newton's method to compute E_{Fp} and E_{Fn} from Eqs. (5.33) and (5.34) given known p_{2D} and n_{2D} .

10. Repeat Steps 7–9 until the charge from Step 8 is consistent with the potential, *i.e.*, the simplified nonlinear Poisson equation converges in a single Newton step.

The Newton algorithm is identical in Steps 1–5. Following these precalculation steps, the algorithm proceeds as follows:

6. Guess an initial potential distribution. In practice, the initial guess is generated by performing Steps 6 and 7 from the Gummel algorithm above, with quasi-Fermi levels chosen such that

$$E_{Fp} = E_{Fn} = \begin{cases} 0 & , \text{ in the left lead and the left } \frac{1}{4} \text{ of the scattering region} \\ -qV_{DS} & , \text{ in the right lead and the right } \frac{1}{4} \text{ of the scattering region} \\ -\frac{qV_{DS}}{2} & , \text{ otherwise} \end{cases}$$

7. Solve for the Green's function in the scattering region in order to obtain the charge there. This is done via a charge neutrality level as in Step 8 from the previous algorithm. The approximate Jacobian of the Poisson equation,

$$KV = \mathcal{C}[V],$$

is also computed at this time, where it is noted, by the square brackets, that \mathcal{C} is a functional of V , *i.e.*, it depends on V everywhere on the CN surface, and not just on the value of V at a single point. The elements of the full Jacobian matrix are given by

$$K_{ij} = \frac{\partial \mathcal{C}_i}{\partial V_j},$$

where the first term is the finite element stiffness matrix, which contains mostly zero elements. As a result, sparse matrix methods may be used to store, and perform operations with, this term. The second term is also largely zero with nonzero elements only for i and j corresponding to points on the CN surface. Unfortunately, this term is still much too computationally intensive to evaluate fully, and is still too large to store if each CN atom is being treated. As a result, it is approximated by taking the derivative with respect to the contact Fermi levels [132]. This results in a matrix with nonzero elements only along the main diagonal.

8. Update the potential as usual in Newton's method with the approximate Jacobian.
9. Repeat Steps 7–8 until the potential converges, *i.e.*, until the charge density and potential are self-consistent.

Due to the time-intensive nature of these computations, it has been helpful to use parallel processing in order to compute the charge density. This was implemented on the energy integrals using the message passage interface [135]. The domain of integration was split into several subdomains, and these smaller integrals could be computed simultaneously. Upon completion, each small integral was sent to the master process for summation in order to complete the computation of the total integral.

5.7 Results and Discussion

This section begins by considering two device structures, depicted in Fig. 5.6: an asymmetric double-gate (DG) device, and a semi-cylindrical-gate (SC) device. In both cases, the source and drain are

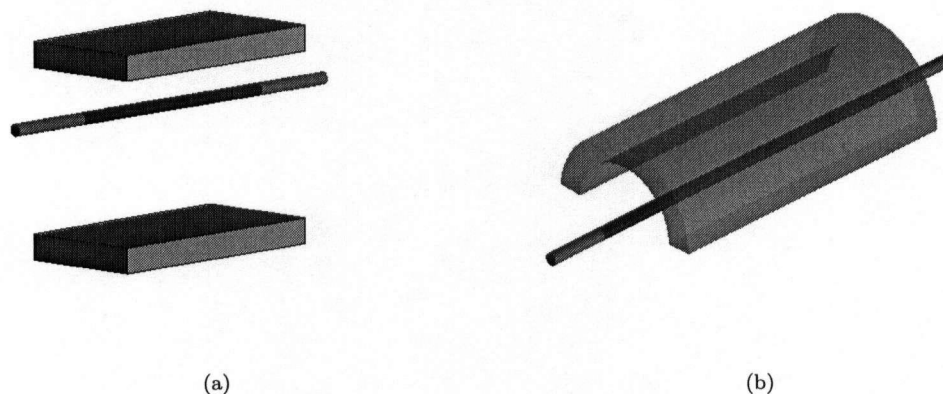


Figure 5.6: Schematics of the (a) double-gate and (b) semi-cylindrical-gate CNFET geometries.

doped nanotubes of the same chirality as the intrinsic channel, and are shown by the different shading in the figure. The DG case is close to structures presently being investigated experimentally, and the SC device may be a realizable tri-gate-like structure. It is assumed that the entire simulation space outside of the CN is filled with a homogeneous dielectric, taken to be SiO_2 unless otherwise specified. A Dirichlet boundary condition is applied on the gate, and homogeneous Neumann boundary conditions are used on the remaining boundaries of the simulation space, including on the outer ends of the source and drain, as in Ref. [57].

The reason for the Neumann conditions on the source and drain stems from the fact that the devices involve doped, semiconducting, ballistic contacts, and as a result, it is no longer appropriate to use Dirichlet boundary conditions, as was done in the metal-contacted, Schottky-barrier case [136]. Rather, recall that some length of the source and drain are included in the simulation space. Neumann conditions are applied at their ends [114, 136], such that the source and drain potentials effectively float at the edges of the domain, and the drain-source voltage is included only when evaluating Eq. (5.35), where $E_{FL} = 0$ is the Fermi level deep in the source, and $E_{FR} = -qV_{DS}$ is the Fermi level deep in the drain. In order for the solution to be valid, it is required that the local potential becomes constant in the longitudinal direction well before the end of the simulation space. To illustrate this, consider the potential in Fig. 5.7. Shown are the results for gate thicknesses of 10,

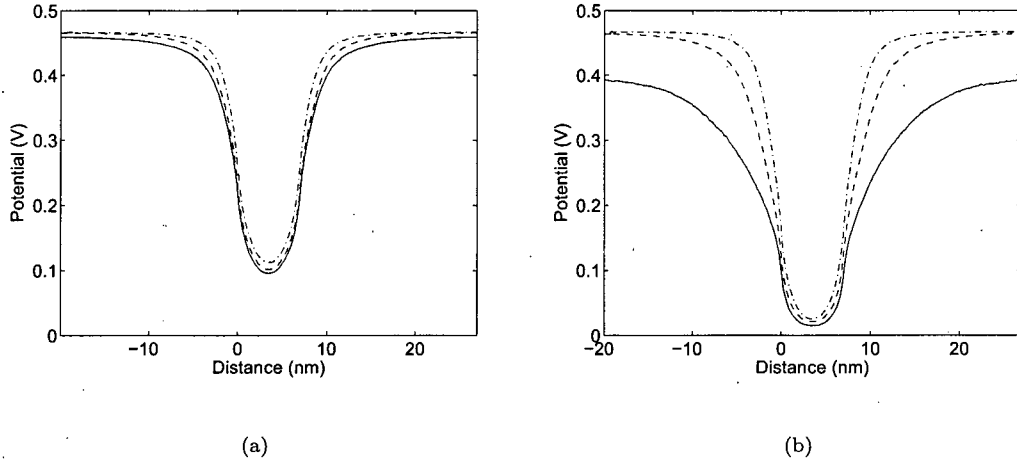


Figure 5.7: Local CN potential for (a) the double-gate and (b) the semi-cylindrical-gate devices with $V_{GS} = V_{DS} = 0$. The doped regions are 20 nm long at either end, and are doped n -type to a density of 0.38 nm^{-2} . The intrinsic region is 7 nm long. Shown are curves corresponding to gate thicknesses of 10 nm (solid), 2 nm (dashed), and 0.1 nm (dash-dotted). Data is taken along the top of the nanotube.

2, and 0.1 nm by the solid, dashed, and dash-dotted lines, respectively. In each case, the potential is forced to become flat by the imposition of the boundary conditions. In both the DG and SC cases, the potential becomes flat well before the boundary for the 2 and 0.1 nm thicknesses, and these

solutions may be considered to be valid. In the 10 nm case, particularly for the SC device, this is not the case, and the solution is not valid. For a valid solution, charge neutrality must be achieved within the simulation space. In order to correct this, longer doped regions must be simulated; however, due to the atomistic approach, this incurs a huge computational cost.

At equilibrium, insight may also be gained into the azimuthal potential variation that results from these non-cylindrical structures. This is depicted in Fig. 5.8 for the CN surface potential in the middle of the intrinsic region as a function of azimuthal angle, where $\phi = 0$ corresponds to the top of the CN, *i.e.*, the closest point to the top gate in the DG case. The DG structure exhibits

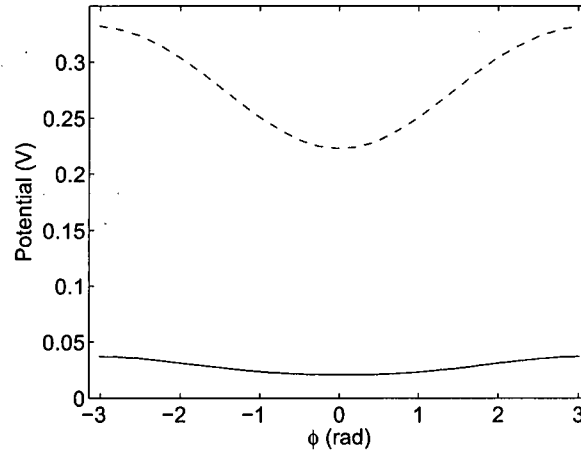


Figure 5.8: Local CN potential in the middle of the intrinsic region as a function of azimuthal angle for the double-gate (dashed) and semi-cylindrical-gate (solid) devices.

greater azimuthal variation as it is further from the completely symmetric, cylindrical geometry.

In order to investigate the impact of the azimuthal variation, and the geometrical differences, results are presented for five structures: a DG device with zero gate thickness, similar to that presented in Ref. [57]; a DG device with 5 nm gate thickness; a DG device with 5 nm gate thickness where the axial potential is used in the NEGF solution rather than the proper surface potential; an SC device with zero gate thickness; an SC device with 2 nm gate thickness. In all cases, a *p-i-n* device is considered, which is interesting due to its potential for an exceptional subthreshold slope [63,127]. The doping density is set to 0.1 nm^{-2} in the *p*- and *n*-regions.

These devices have also been studied using the Flietner dispersion relation [63, 127], and by Koswatta *et al.*, using a two-band NEGF simulation [65]. These results are an extension in that they use NEGF to consider all of the bands simultaneously, in addition to the consideration of geometrical variations. The simulation space includes 15 nm each of the source and drain doped regions, in addition to the 7 nm intrinsic channel, and a 2 nm oxide thickness is chosen on top of the CN, and 20 nm below in the DG case.

In order to better understand the operation of a *p-i-n* device, consider the local density of states, and the energy spectrum of the electron density for the SC device. These are shown in Fig. 5.9. The bright patches show regions of high density, the solid lines show the field-free conduction- and

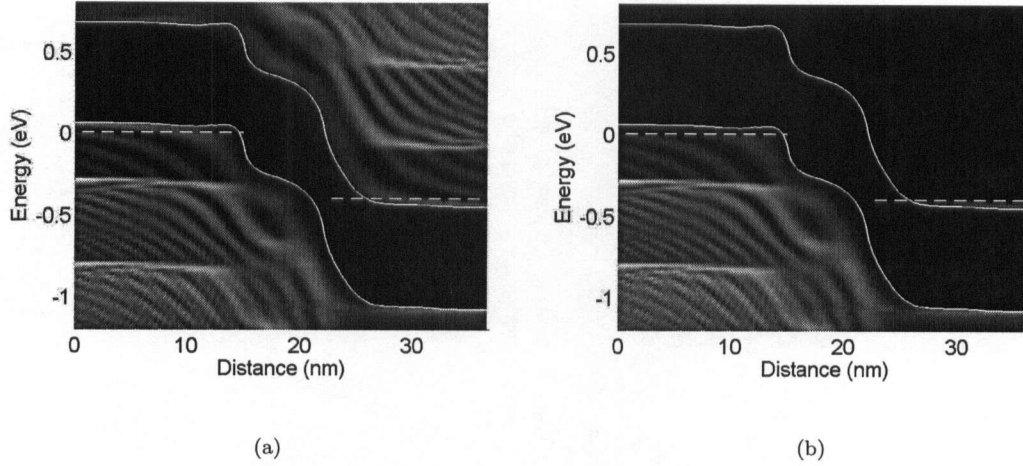


Figure 5.9: (a) Local density of states and (b) energy spectrum of the electron density for the semi-cylindrical-gate device. Solid lines correspond to the field-free conduction- and valence-band edges shifted by the local potential energy, and the dashed lines show the source and drain Fermi levels deep in the contacts. The dashed lines terminate at the ends of the doped regions. Data is taken for $V_{GS} = -0.057$ V and $V_{DS} = 0.4$ V.

valence-band edges shifted by the local potential energy, and the dashed lines indicate the positions of the source and drain Fermi levels deep in their respective contacts. The dashed lines extend throughout the doped regions. In Fig. 5.9(a), the electronic states that are available to be populated by source and drain injection are shown. As seen in Chapter 4, the jumps in density correspond

to the multiple bands present in the CN. Fig. 5.9(b) shows how these states are actually populated by source and drain injection. It appears that electrons may tunnel from the valence band to the conduction band near the doped region in the drain.

In order to confirm this, the energy spectrum of the current is plotted in Fig. 5.10, which clearly shows electrons tunneling from occupied states in the valence band to empty states in the conduction band. Note, too, that many of the electrons originate in the second valence band, and flow into

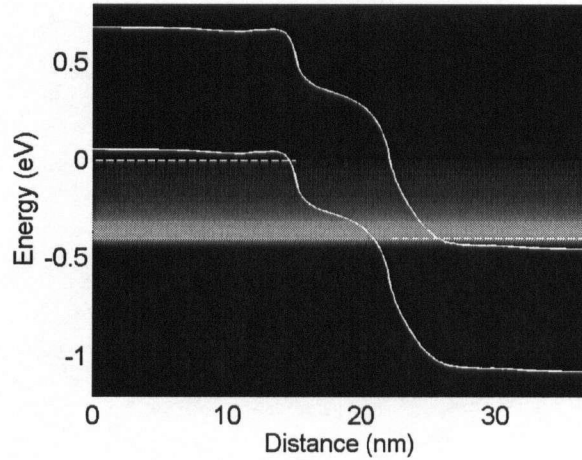


Figure 5.10: Energy spectrum of the current for the same device and conditions as in Fig. 5.9.

the first conduction band. This result highlights the importance of being able to perform a full multi-band simulation, and would be missed in Ref. [65], for example, where it appears that only the highest valence- and lowest conduction-bands were considered.

Turning to the terminal characteristics of the model devices, the \mathcal{I}_D - V_{GS} data is summarized in Fig. 5.11. For these devices, the subthreshold slope is significantly worse than the thermionic limit of 60 mV/dec, with a value around 300 and 700 mV/dec in the semi-cylindrical- and double-gate cases, respectively. In order to investigate this poor performance, the energy spectrum of the current may be utilized. Close inspection reveals that a significant tunneling current exists for all energies between the contact Fermi levels with a maximum for energies where the tunneling distance is shortest, as shown in Fig. 5.10 for example. For these model devices, the current is appreciable even in the OFF state, with a value on the order of $0.3 \mu\text{A}$. This OFF-current is the root cause of

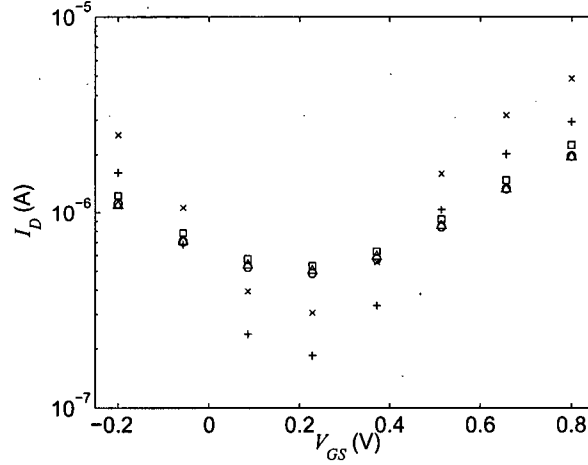


Figure 5.11: I_D - V_{GS} for double-gate and semi-cylindrical-gate devices. The DG devices are zero gate thickness (\square), 5 nm gate thickness (\circ), and 5 nm gate thickness where the axial potential is used (\triangle). The SC devices are zero gate thickness (\times), and 2 nm gate thickness ($+$).

the poor subthreshold slope, and may be reduced by increasing the length of the intrinsic region from the value of 7 nm used here. In previous works, both experimental [64] and theoretical [65], the intrinsic regions were at least 20 nm in length. This work indicates that there is a lower limit to the channel length in order to realize the benefit of an exceptional subthreshold slope, *i.e.*, using a *p-i-n* device does not necessarily imply that this figure-of-merit will be better than the thermionic limit.

From the figure, it is also clear that neglecting the azimuthal potential variation leads to a fairly small error in the DG terminal characteristic. In addition, note that the gate thickness has a similarly small effect for the thicknesses considered here. For the SC case, however, the gate thickness plays a larger role, with roughly 1.6 times as much current if a zero gate thickness is assumed. This discrepancy exists even for the aggressive value of 2 nm for the gate thickness simulated here, which would lead to excessive series resistance in a practical device. In order to understand this increase, consider the energy band diagrams for the semi-cylindrical geometries in Fig. 5.12. The solid line is for a gate thickness of 2 nm, and it may be seen that the inter-band tunneling distance is greater

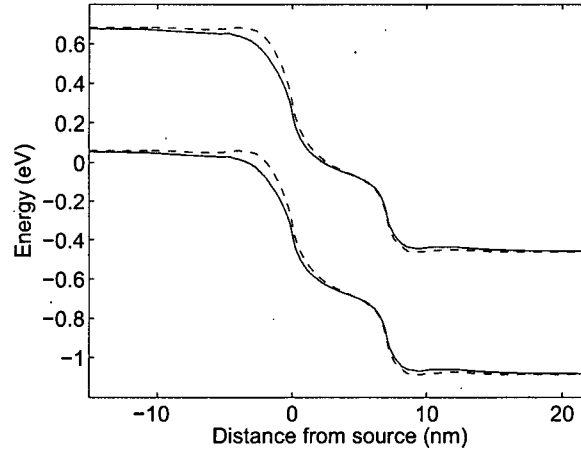


Figure 5.12: Energy band diagrams for the semi-cylindrical-gate geometry with a gate thickness of 2 nm (solid) and 0 nm (dashed).

in this case than for the zero-gate-thickness device, shown by the dashed line. This is due to the fringing field from the gate, which extends to greater distances in the thick-gate device. The steeper band-bending in the thin-gate case results in increased tunneling, and hence, more current.

Turning now to the atomic current, two of the DG devices are considered. For the case where the axial potential is used, the current shows no azimuthal variation, and is distributed evenly around the circumference of the CN. For the device with zero gate thickness, where the proper surface potential at each atomic site was utilized, considerable structure emerges, as shown in Fig. 5.13. While the ultimate utility of this calculation is not yet known, perhaps this approach could be used to investigate the effect of defects on CNFET behaviour, *i.e.*, the impact of defects on the electron flow pattern through the CN, or the flow patterns across junctions between CNs of differing chirality.

Finally, the NEGF model is used to compare to results generated³ from using the Flietner dispersion relation throughout the full energy range [76] on a *p-i-n* device. A cylindrical structure is used for this comparison, such that the 1D Schrödinger solution used in Ref. [76] applies. The device consists of a 7 nm intrinsic channel connected to doped CN leads, and a 2 nm SiO₂ layer between

³Doped-contact results using the Flietner dispersion relation were provided by L. Castro of the University of British Columbia.

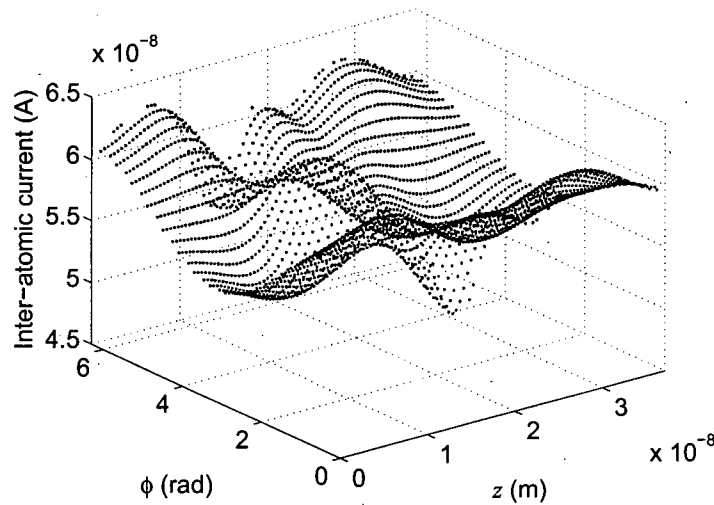


Figure 5.13: Inter-atomic current for a double-gate device with zero gate thickness.

the channel and the gate. The doping density is 0.1 nm^{-2} as for the previous NEGF results. As Fig. 5.14 shows, the Flietner relation provides very good agreement with the NEGF results in the low-current regime. At higher current, the discrepancy between the predictions grows to a maximum value of about 32% at $V_{GS} \simeq -0.06 \text{ V}$, where it is noted that the Flietner band structure results in an overestimate of the current. It remains to be seen how this comparison will turn out when considering non-cylindrical geometries.

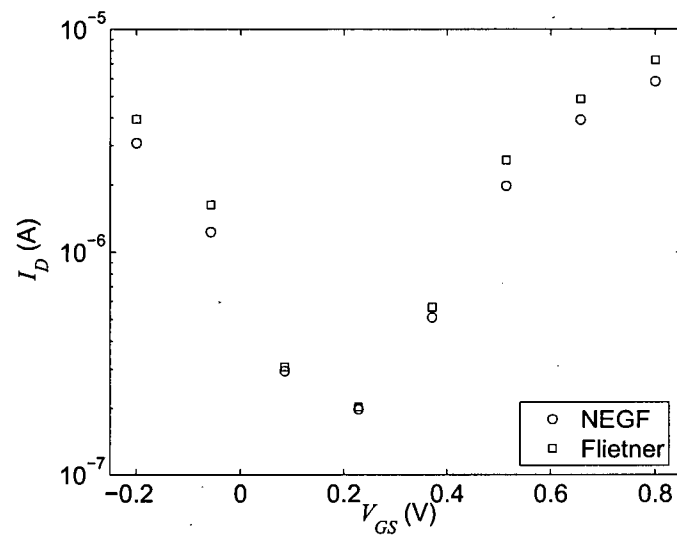


Figure 5.14: Comparison of NEGF results (○) to those using the Flietner dispersion relation (□) for a cylindrical *p-i-n* device.

Chapter 6

Conclusions

In this thesis, we have presented a variety of CNFET models valid in different operating regimes, and subject to successively less restrictive approximations. We have focused on including quantum phenomena that had been neglected at the time this work commenced. For example, the “top-of-the-barrier” [23, 52] model, while it may be useful for negative barrier devices, it does not deal with tunneling or quantum resonance. Also, the quasi-Fermi level model [68, 69, 71], while it was one of the earliest models to predict bipolar conduction at high V_{DS} , predicted a much earlier onset of this behaviour than the more rigorous effective-mass model due to an unphysical thinning of the Schottky barriers at high bias.

6.1 Current Progress

Our first contribution was to provide equilibrium results, using an analytical expression for the Poisson solution, from which Castro *et al.*, derived a compact model [51] that incorporated the WKB approximation to tunneling. In Ref. [77], we used this model to present some preliminary current-voltage data, and also showed the dependence of the equilibrium band structure on oxide permittivity and contact workfunction. A modified WKB result was derived to also account for above-barrier reflection, something that a classical treatment would also ignore, and this led to significant improvements in Castro’s compact model [53].

Following this, we proceeded to develop a coupled solution to the Poisson and effective-mass Schrödinger equations [72]. This work served as the early benchmark against which compact models could be validated, and also served as the method for performing small-signal analyses [73–75, 103] and switching-time calculations [55]. In addition, we derived, in detail, expressions for the so-called quantum capacitance [89], which had been previously used in qualitative discussions of CNFETs [23, 30, 90]. In so doing, we observed that this capacitance is not truly quantized due to the form

of the one-dimensional DOS. Our expression may be used to assist in the quantitative modeling of CNFETs.

In order to deal with the non-parabolicity of the CN bands, researchers started to investigate simulations using the NEGF formalism. In Ref. [56], a method was presented for simulating devices where the potential on the surface of the CN was not a function of azimuthal angle. In our work, we extended that technique by considering a real-space basis where the potential at each atom was computed via a three-dimensional finite element solution. While completing this method, we became aware of Refs. [57, 58] where essentially the same technique has been used. One difference in our implementations was in the computation of the self-energy matrices, where we derived a quadratic matrix equation [117] which may be solved with a generalization of Newton's method. In Refs. [56–58], a recursive algorithm was employed [118], which we have found to be less computationally efficient than our method. In addition, these works have used a Gummel technique in order to obtain convergence, whereas we have chosen an approximate Jacobian [132] in Newton's method, which seems to be more stable, albeit more time-consuming, in the cases that we have investigated thus far.

Despite the similarities in algorithm, we have further expanded on the work of Fiori *et al.*, where devices were simulated assuming that the contact thickness played no role in the device characteristics. This was implicitly assumed by the imposition of Neumann boundary conditions along surfaces that extended from the nearest face of the gate-contact to the CN channel [57, 58]. In this work, we investigated the effect of contact thickness, and also used our finite element Poisson solver in order to consider curved gate contacts. In addition, Fiori *et al.*, present results only for the (11, 0) CN, with a diameter of 0.86 nm. It is known that CNs with diameters below 1 nm undergo curvature-induced hybridization of the σ and π states [137] which significantly alter the CN band structure from the result of considering only the π states. In most of the literature, and in this thesis, we consider only the π states. As a result, we present results primarily for the (16, 0) CN, with a diameter of 1.25 nm, for which our calculations should be reasonable. Finally, while the ultimate utility of this aspect of the thesis is not yet clear, we have also derived the current operator for a carbon nanotube such that we may show the current between individual atoms of the CN.

From these contributions to the understanding of CNFET behaviour, we may conclude the following:

1. equilibrium calculations may be performed to gain insight into CNFET operation, and to aid in the development of compact models;
2. the tight-binding approach may be used to various levels of approximation in order to simulate CNFETs to varying degrees of physical rigour and computational effort;
3. effective-mass Schrödinger-Poisson solutions may be used for DC, small-signal AC, and switching time calculations;
4. an integral expression may be derived for the quantum capacitance;
5. significant optimization of CNFET geometries will be required in order to obtain switching times that rival existing Si CMOS, such as, for example, using parallel CNs in order to increase the drive current, or decreasing the source and drain contact thicknesses in order to reduce the extrinsic capacitances;
6. for large-signal behaviour, it is desirable to have a large bandgap, in order to obtain a good $\mathcal{I}_{\text{ON}}/\mathcal{I}_{\text{OFF}}$, and a thin oxide, in order to decrease the switching time by increasing \mathcal{I}_{ON} ;
7. the non-equilibrium Green's function formalism may be used with an atomistic, real-space, Hamiltonian in order to treat multi-band effects, and azimuthal asymmetries in CNFETs;
8. a quadratic matrix equation may be solved in order to improve the computation time of the self-energy matrices representing the semi-infinite leads connected to the CN channel;
9. contact thickness may play an important role in a device's behaviour for geometries where the gate-channel coupling is strong;
10. the current operator for a CN may be used to investigate electron flow patterns at the inter-atomic level;
11. while the azimuthal potential variation does not significantly affect the terminal characteristics, for the devices considered here, the inter-atomic current shows considerable structure, which would be missed if azimuthal effects are ignored;
12. in order to realize the exceptional subthreshold slope made possible by the use of a *p-i-n* structure, the channel length needs to be long enough to suppress the significant inter-band tunneling of carriers in the OFF-state;

13. use of the Flietner dispersion relation for a cylindrical device results in reasonable agreement with the NEGF results, particularly in the low-current regime, with greater error when the device is ON.

6.2 Future Work

Using these advances as the foundation for further study, we finally consider some possibilities for future work. Of course, there is initial work on geometry optimization that may be performed directly from this work in order to confirm, and extend upon, the results of Refs. [109,138]. Along with this, a detailed evaluation of existing models may be performed in order to validate less computationally intensive methods than the thorough NEGF algorithm. This work was begun in Chapter 5, where we compared to results generated using the Flietner dispersion relation, and may be continued to include a larger bias range, and non-cylindrical devices.

Most importantly, these models need to be rigorously validated against experimental results. This has been impeded by the many technological challenges involved with CNFET fabrication; however, our atomistic approach will better lend itself to the simulation of experimental structures than the earlier methods which relied on cylindrical symmetry. Models having the same general form as employed here have been compared, and fit, to individual device characteristics [56,109,112,139], and this is an encouraging sign of the validity and usefulness of our modeling approach. It also gives confidence in the quantitatively predictive quality of our model, but the ultimate validation of this will require a systematic experimental investigation which has not yet been achieved.

In terms of continuing the progression of these models towards more physically rigorous situations, the inclusion of electron-phonon interactions has started to be considered by some authors [12,65], and this may also be incorporated into our model through additional self-energy matrices [12,94]. In the DC case, this is not expected to impact on performance, but will have an effect on the AC behaviour [12].

Another extension, already alluded to in this work, would be to consider the role of defects in CNs, and to treat junctions between dissimilar tubes. Such defects have been shown to produce single-electron transistor characteristics [140] due to local band structure modifications [141], and have also been predicted to induce metal-semiconductor transitions in some tubes [141,142]. This may assist with the integration of devices if a single tube could be used to realize a metal-contacted

semiconductor channel [142]. Knowledge of the carrier flow patterns may be helpful in this investigation. Deformations of pristine nanotubes due to their physical interaction with a substrate could also be treated by modifying the overlap integral γ such that it varies with azimuthal angle. A density functional theory calculation could be used in order to obtain appropriate values.

Further, while DC simulations, such as those described herein, can be used to perform small-signal and switching-time analyses, it would be helpful to perform time-dependent simulations in order to validate previous results [73–76, 103]. For a switching calculation, for example, our DC solution could be used as an initial condition, with the evolution of the wavefunction given by the time-dependent Schrödinger equation.

Finally, it may be important to consider electron-electron interactions, as may lead to Luttinger liquid behaviour [143]; excitonic interactions, as may be important in electro-optical devices; or even just to include the effects of exchange and correlation, as in a density functional theory approach. Some investigations into the Luttinger liquid phenomena have been performed, *e.g.*, Refs. [144, 145], but not much in the context of CNFETs except perhaps for a qualitative treatment in Ref. [146]. The importance of this phenomena for CNFETs is not yet clear as it does not appear to be present at some energy scales [147], and electron-electron interactions are expected to have little effect on conductance when the contacts are highly transparent [148], such as would be desirable in a CNFET. However, many-body effects have also been shown to significantly modify the CN bandgap [149], and if this occurs in CNFETs, it would be critical to incorporate this effect.

The solid framework provided by this thesis will be an excellent starting point for these, and possibly other, further investigations. By continuing to successively strip away some of the more restrictive assumptions of our model, significant progress may continue towards a better understanding of these fascinating devices.

Bibliography

- [1] Semiconductor Industry Association. *International Technology Roadmap for Semiconductors—2004 Update: Overview and Summaries*, 2004. [Online.] Available: <http://www.itrs.net/Common/2004Update/2004Update.htm>.
- [2] A. Oberlin, M. Endo, and A. T. Koyama. Filamentous growth of carbon through benzene decomposition. *J. Cryst. Grow.*, 32(3):335–349, 1976.
- [3] Sumio Iijima. Helical microtubules of graphitic carbon. *Nature*, 354:56–58, 1991.
- [4] R. Saito, G. Dresselhaus, and M. S. Dresselhaus. *Physical Properties of Carbon Nanotubes*. Imperial College Press, London, 1st edition, 1998.
- [5] M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund. *Science of Fullerenes and Carbon Nanotubes*. Academic Press, Toronto, 1996.
- [6] Keivan Esfarjani, Amir A. Farajian, Yuichi Hashi, and Yoshiyuki Kawazoe. Electronic, transport and mechanical properties of carbon nanotubes. In Y. Kawazoe, T. Kondow, and K. Ohno, editors, *Clusters and Nanomaterials—Theory and Experiment*, Springer Series in Cluster Physics, pages 187–220. Springer-Verlag, Berlin, 2002.
- [7] Sander J. Tans, Alwin R. M. Verschueren, and Cees Dekker. Room-temperature transistor based on a single carbon nanotube. *Nature*, 393:49–52, 1998.
- [8] Paul L. McEuen, Michael S. Fuhrer, and Hongkun Park. Single-walled carbon nanotube electronics. *IEEE Trans. Nanotechnol.*, 1(1):78–85, 2002.
- [9] T. Dürkop, E. Cobas, and M. S. Fuhrer. High-mobility semiconducting nanotubes. In *Molecular Nanostructures*, pages 524–527, 2003.

- [10] Ali Javey, Jing Guo, Magnus Paulsson, Qian Wang, David Mann, Mark Lundstrom, and Hongjie Dai. High-field, quasi-ballistic transport in short carbon nanotubes. *Phys. Rev. Lett.*, 92(10):106804–1–106804–4, 2004.
- [11] Yu-Ming Lin, Joerg Appenzeller, Zhihong Chen, Zhi-Gang Chen, Hui-Ming Cheng, and Phaedon Avouris. High-performance dual-gate carbon nanotube FETs with 40-nm gate length. *IEEE Electron Device Lett.*, 26(11):823–825, 2005.
- [12] Jing Guo. A quantum-mechanical treatment of phonon scattering in carbon nanotube transistors. *J. Appl. Phys.*, 98:063519–1–063519–6, 2005.
- [13] Akin Akturk, Gary Pennington, and Neil Goldsman. Quantum modeling and proposed designs of CNT-embedded nanoscale MOSFETs. *IEEE Trans. Electron Devices*, 52(4):577–584, 2005.
- [14] Khairul Alam and Roger K. Lake. Leakage and performance of zero-Schottky-barrier carbon nanotube transistors. *J. Appl. Phys.*, 98:064307–1–064307–8, 2005.
- [15] Vasili Perebeinos, J. Tersoff, and Phaedon Avouris. Mobility in semiconducting carbon nanotubes at finite carrier density. *Nano Lett.*, 6(2):205–208, 2006.
- [16] David L. Pulfrey and N. Garry Tarr. *Introduction to Microelectronic Devices*. Prentice Hall Series in Solid State Physical Electronics. Prentice Hall, New Jersey, 1989.
- [17] Yi Cui, Zhaohui Zhong, Deli Wang, Wayne U. Wang, and Charles M. Lieber. High performance silicon nanowire field effect transistors. *Nano Lett.*, 3(2):149–152, 2003.
- [18] Cees Dekker. Carbon nanotubes as molecular quantum wires. *Phys. Today*, 52(5):22–28, 1999.
- [19] Zhen Yao, Charles L. Kane, and Cees Dekker. High-field electrical transport in single-wall carbon nanotubes. *Phys. Rev. Lett.*, 84(13):2941–2944, 2000.
- [20] A. Tilke, L. Pescini, A. Erbe, H. Lorenz, and R. H. Blick. Electron-phonon interaction in suspended highly doped silicon nanowires. *Nanotechnology*, 13:491–494, 2002.
- [21] Sang-Hyun Oh, Don Monroe, and J. M. Hergenrother. Analytic description of short-channel effects in fully-depleted double-gate and cylindrical, surrounding-gate MOSFETs. *IEEE Electron Device Lett.*, 21(9):445–447, 2000.

- [22] Brian Winstead and Umberto Ravaioli. Simulation of Schottky barrier MOSFET's with a coupled quantum injection/Monte Carlo technique. *IEEE Trans. Electron Devices*, 47(6):1241–1246, 2000.
- [23] Jing Guo, Sebastien Goasguen, Mark Lundstrom, and Supriyo Datta. Metal-insulator-semiconductor electrostatics of carbon nanotubes. *Appl. Phys. Lett.*, 81(8):1486–1488, 2002.
- [24] Wolfgang Hoenlein, Franz Kreupl, Georg Stefan Duesberg, Andrew Peter Graham, Maik Liebau, Robert Viktor Seidel, and Eugen Unger. Carbon nanotube applications in microelectronics. *IEEE Trans. Compon. Pack. T.*, 27(4):629–634, 2004.
- [25] Hou T. Ng, J. Han, Toshishige Yamada, P. Nguyen, Yi P. Chen, and M. Meyyappan. Single crystal nanowire vertical surround-gate field-effect transistor. *Nano Lett.*, 4(7):1247–1252, 2004.
- [26] Ali Javey, Hyounsub Kim, Markus Brink, Qian Wang, Ant Ural, Jing Guo, Paul McIntyre, Paul McEuen, Mark Lundstrom, and Hongjie Dai. High- κ dielectrics for advanced carbon-nanotube transistors and logic gates. *Nature Mater.*, 1:241–246, 2002.
- [27] B. M. Kim, T. Brintlinger, E. Cobas, M. S. Fuhrer, Haimei Zheng, Z. Yu, R. Droopad, J. Ramdani, and K. Eisenbeiser. High-performance carbon nanotube transistors on SrTiO₃/Si substrates. *Appl. Phys. Lett.*, 84(11):1946–1948, 2004.
- [28] Sami Rosenblatt, Yuval Yaish, Jiwoong Park, Jeff Gore, Vera Sazonova, and Paul L. McEuen. High performance electrolyte gated carbon nanotube transistors. *Nano Lett.*, 2(8):869–872, 2002.
- [29] S. B. Cronin, R. Barnett, M. Tinkham, S. G. Chou, O. Rabin, M. S. Dresselhaus, A. K. Swan, M. S. Ünlü, and B. B. Goldberg. Electrochemical gating of individual single-wall carbon nanotubes observed by electron transport measurements and resonant Raman spectroscopy. *Appl. Phys. Lett.*, 84(12):2052–2054, 2004.
- [30] Anisur Rahman, Jing Guo, Supriyo Datta, and Mark S. Lundstrom. Theory of ballistic nanotransistors. *IEEE Trans. Electron Devices*, 50(9):1853–1864, 2003.

- [31] Ming Zheng, Anand Jagota, Ellen D. Semke, Bruce A. Diner, Robert S. McLean, Steve R. Lustig, Raymond E. Richardson, and Nancy G. Tassi. DNA-assisted dispersion and separation of carbon nanotubes. *Nature Mater.*, 2:338–342, 2003.
- [32] Gang Lu, Paul Maragakis, and Efthimios Kaxiras. Carbon nanotube interaction with DNA. *Nano Lett.*, 5(5):897–900, 2005.
- [33] Jing Kong, Nathan R. Franklin, Chongwu Zhou, Michael G. Chapline, Shu Peng, Kyeongjae Cho, and Hongjie Dai. Nanotube molecular wires as chemical sensors. *Science*, 287:622–625, 2000.
- [34] Keith Bradley, Jean-Christophe P. Gabriel, Alexander Star, and George Grüner. Short-channel effects in contact-passivated nanotube chemical sensors. *Appl. Phys. Lett.*, 83(18):3821–3823, 2003.
- [35] Philip G. Collins and Phaedon Avouris. Nanotubes for electronics. *Sci. Am.*, pages 62–69, December 2000.
- [36] R. Gutierrez, G. Fagas, G. Cuniberti, F. Grossmann, R. Schmidt, and K. Richter. Theory of an all-carbon molecular switch. *Phys. Rev. B*, 65:113410–1–113410–4, 2002.
- [37] R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and Ph. Avouris. Single- and multi-wall carbon nanotube field-effect transistors. *Appl. Phys. Lett.*, 73(17):2447–2449, 1998.
- [38] Hongjie Dai. Nanotube growth and characterization. In Mildred S. Dresselhaus, Gene Dresselhaus, and Phaedon Avouris, editors, *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications*, volume 80 of *Topics in Applied Physics*, pages 29–53. Springer, Berlin, 2001.
- [39] Hongjie Dai. Carbon nanotubes: Synthesis, integration, and properties. *Acc. Chem. Res.*, 35(12):1035–1044, 2002.
- [40] Ralph Krupke, Frank Hennrich, Hilbert v. Löhneysen, and Manfred M. Kappes. Separation of metallic from semiconducting single-walled carbon nanotubes. *Science*, 301:344–347, 2003.
- [41] Michael S. Strano, Christopher A. Dyke, Monica L. Usrey, Paul W. Barone, Matthew J. Allen, Hongwei Shan, Carter Kittrell, Robert H. Hauge, James M. Tour, and Richard E. Smal-

- ley. Electronic structure control of single-walled carbon nanotube functionalization. *Science*, 301:1519–1522, 2003.
- [42] Yutaka Ohno, Shigeru Kishimoto, Takashi Mizutani, Toshiya Okazaki, and Hisanori Shinohara. Chirality assignment of individual single-walled carbon nanotubes in carbon nanotube field-effect transistors by micro-photocurrent spectroscopy. *Appl. Phys. Lett.*, 84(8):1368–1370, 2004.
- [43] Aleksey N. Kolmogorov, Vincent H. Crespi, Monica H. Schleier-Smith, and James C. Ellenbogen. Nanotube-substrate interactions: Distinguishing carbon nanotubes by the helical angle. *Phys. Rev. Lett.*, 92(8):085503–1–085503–4, 2004.
- [44] Yiming Li, David Mann, Marco Rolandi, Woong Kim, Ant Ural, Steven Hung, Ali Javey, Jien Cao, Dunwei Wang, Erhan Yenilmez, Qian Wang, James F. Gibbons, Yoshio Nishi, and Hongjie Dai. Preferential growth of semiconducting single-walled carbon nanotubes by a plasma enhanced CVD method. *Nano Lett.*, 4(2):317–321, 2004.
- [45] J. Appenzeller, J. Knoch, R. Martel, V. Derycke, S. Wind, and Ph. Avouris. Short-channel like effects in Schottky barrier carbon nanotube field-effect transistors. In *IEDM Tech. Digest*, pages 285–288, 2002.
- [46] Ph. Avouris, R. Martel, S. Heinze, M. Radosavljevic, S. Wind, V. Derycke, J. Appenzeller, and J. Terso. The role of Schottky barriers on the behavior of carbon nanotube field-effect transistors. In H. Kuzmany, J. Fink, M. Mehring, and S. Roth, editors, *Structural and Electronic Properties of Molecular Nanostructures*, XVI International Winterschool on Electronic Properties of Novel Materials, pages 508–512, 2002.
- [47] S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, and Ph. Avouris. Carbon nanotubes as Schottky barrier transistors. *Phys. Rev. Lett.*, 89(10):106801–1–106801–4, 2002.
- [48] J. Knoch and J. Appenzeller. Impact of the channel thickness on the performance of Schottky barrier metal-oxide-semiconductor field-effect transistors. *Appl. Phys. Lett.*, 81(16):3082–3084, 2002.

- [49] J. Appenzeller, J. Knoch, and Ph. Avouris. Carbon nanotube field-effect transistors—an example of an ultra-thin body Schottky barrier device. In *Proc. TMS 61st Annual Device Research Conference*, pages 167–170, Salt Lake City, U.S.A., June 2003.
- [50] S. Heinze, M. Radosavljević, J. Tersoff, and Ph. Avouris. Unexpected scaling of the performance of carbon nanotube Schottky-barrier transistors. *Phys. Rev. B*, 68:235418–1–235418–5, 2003.
- [51] L. C. Castro, D. L. John, and D. L. Pulfrey. Towards a compact model for Schottky-barrier nanotube FETs. In *Proc. IEEE Conf. on Optoelectronic and Microelectronic Materials and Devices*, pages 303–306, Sydney, Australia, December 2002.
- [52] Jing Guo, Mark Lundstrom, and Supriyo Datta. Performance projections for ballistic carbon nanotube field-effect transistors. *Appl. Phys. Lett.*, 80(17):3192–3194, 2002.
- [53] L. C. Castro, D. L. John, and D. L. Pulfrey. An improved evaluation of the DC performance of carbon nanotube field-effect transistors. *Smart Mater. Struct.*, 15(1):S9–S13, 2006.
- [54] M. P. Anantram, S. Datta, and Yongqiang Xue. Coupling of carbon nanotubes to metallic contacts. *Phys. Rev. B*, 61(20):14219–14225, 2000.
- [55] D. L. John and D. L. Pulfrey. Switching-speed calculations for Schottky-barrier carbon nanotube field-effect transistors. *J. Vac. Sci. Technol. A*, 24(3):708–712, 2006.
- [56] Jing Guo. *Carbon nanotube electronics: Modeling, physics, and applications*. PhD thesis, Purdue University, West Lafayette, U.S.A., 2004.
- [57] G. Fiori, G. Iannaccone, and G. Klimeck. Performance of carbon nanotube field effect transistors with doped source and drain extensions and arbitrary geometry. In *IEDM Tech. Digest*, pages 529–532, December 2005.
- [58] Gianluca Fiori, Giuseppe Iannaccone, Mark Lundstrom, and Gerhard Klimeck. Three-dimensional atomistic simulation of carbon nanotube FETs with realistic geometry. In *Proc. European Solid-State Device Research Conference*, Grenoble, France, September 2005.
- [59] M. Radosavljević, J. Appenzeller, Ph. Avouris, and J. Knoch. High performance of potassium *n*-doped carbon nanotube field-effect transistors. *Appl. Phys. Lett.*, 84(18):3693–3695, 2004.

- [60] Jia Chen, Christian Klinke, Ali Afzali, Kevin Chan, and Phaedon Avouris. Self-aligned carbon nanotube transistors with novel chemical doping. In *IEDM Tech. Digest*, pages 695–698, December 2004.
- [61] Yu-Ming Lin, Joerg Appenzeller, and Phaedon Avouris. Novel carbon nanotube FET design with tunable polarity. In *IEDM Tech. Digest*, pages 687–690, December 2004.
- [62] Ali Javey, Ryan Tu, Damon B. Farmer, Jing Guo, Roy G. Gordon, and Hongjie Dai. High performance n-type carbon nanotube field-effect transistors with chemically doped contacts. *Nano Lett.*, 5(2):345–348, 2005.
- [63] J. Appenzeller, Y.-M. Lin, J. Knoch, and Ph. Avouris. Band-to-band tunneling in carbon nanotube field-effect transistors. *Phys. Rev. Lett.*, 93(19):196805–1–196805–4, 2004.
- [64] Joerg Appenzeller, Yu-Ming Lin, Joachim Knoch, Zhihong Chen, and Phaedon Avouris. Comparing carbon nanotube transistors—the ideal choice: A novel tunneling device design. *IEEE Trans. Electron Devices*, 52(12):2568–2576, 2005.
- [65] Siyuranga O. Koswatta, Dmitri E. Nikonov, and Mark S. Lundstrom. Computational study of carbon nanotube p-i-n tunnel FETs. In *IEDM Tech. Digest*, pages 525–528, December 2005.
- [66] Neil W. Ashcroft and N. David Mermin. *Solid State Physics*. Harcourt College Publishers, New York, 1st edition, 1976.
- [67] Jeroen W. G. Wildöer, Liesbeth C. Venema, Andrew G. Rinzler, Richard E. Smalley, and Cees Dekker. Electronic structure of atomically resolved carbon nanotubes. *Nature*, 391:59–62, 1998.
- [68] Jason Clifford, D. L. John, and David L. Pulfrey. Bipolar conduction and drain-induced barrier thinning in carbon nanotube FETs. *IEEE Trans. Nanotechnol.*, 2(3):181–185, 2003.
- [69] Jason Paul Clifford. A self-consistent numerical model for bipolar transport in carbon nanotube field-effect transistors. Master’s thesis, University of British Columbia, Vancouver, Canada, December 2003.
- [70] L. C. Castro, D. L. John, and D. L. Pulfrey. Carbon nanotube transistors: An evaluation. In *Proc. SPIE Conf. Device and Process Technologies for MEMS, Microelectronics, and Photonics III*, volume 5276, pages 1–10, Perth, Australia, April 2004.

- [71] J. P. Clifford, D. L. John, L. C. Castro, and D. L. Pulfrey. Electrostatics of partially gated carbon nanotube FETs. *IEEE Trans. Nanotechnol.*, 3(2):281–286, 2004.
- [72] D. L. John, L. C. Castro, P. J. S. Pereira, and D. L. Pulfrey. A Schrödinger-Poisson solver for modeling carbon nanotube FETs. In *Tech. Proc. of the 2004 NSTI Nanotechnology Conf. and Trade Show*, volume 3, pages 65–68, Boston, U.S.A., March 2004.
- [73] L. C. Castro, D. L. John, D. L. Pulfrey, M. Pourfath, A. Gehring, and H. Kosina. Method for predicting f_T for carbon nanotube FETs. *IEEE Trans. Nanotechnol.*, 4(6):699–704, 2005.
- [74] L. C. Castro and D. L. Pulfrey. Extrapolated f_{\max} for carbon nanotube field-effect transistors. *Nanotechnology*, 17(1):300–304, 2006.
- [75] L. C. Castro, D. L. Pulfrey, and D. L. John. High-frequency capability of Schottky-barrier carbon nanotube FETs. *Solid-State Phenomena*, 2005. Accepted December 22, 2005.
- [76] Leonardo C. Castro. *Modeling of carbon nanotube field-effect transistors*. PhD thesis, University of British Columbia, Vancouver, Canada, 2006.
- [77] D. L. John, Leonardo C. Castro, Jason Clifford, and David L. Pulfrey. Electrostatics of coaxial Schottky-barrier nanotube field-effect transistors. *IEEE Trans. Nanotechnol.*, 2(3):175–180, 2003.
- [78] Christopher P. Auth and James D. Plummer. Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's. *IEEE Electron Device Lett.*, 18(2):74–76, 1997.
- [79] Jing Guo, Ali Javey, Hongjie Dai, Supriyo Datta, and Mark Lundstrom. Predicted performance advantages of carbon nanotube transistors with doped nanotubes as source/drain. [Online.] Available: <http://arxiv.org/pdf/cond-mat/0309039>, 2003.
- [80] François Léonard and J. Tersoff. Role of Fermi-level pinning in nanotube Schottky diodes. *Phys. Rev. Lett.*, 84(20):4693–4696, 2000.
- [81] François Léonard and J. Tersoff. Dielectric response of semiconducting carbon nanotubes. *Appl. Phys. Lett.*, 81(25):4835–4837, 2002.

- [82] John David Jackson. *Classical Electrodynamics*. John Wiley and Sons, Toronto, 3rd edition, 1999.
- [83] D. L. McGuire and D. L. Pulfrey. Error analysis of boundary condition approximations in the modeling of coaxially-gated carbon nanotube field-effect transistors. *Phys. Stat. Sol. (a)*, 2006. Accepted February 21, 2006.
- [84] E. M. Freeman and D. A. Lowther. A novel mapping technique for open boundary finite element solutions to Poisson's equation. *IEEE Trans. Magn.*, 24(6):2934–2936, 1988.
- [85] FEMLAB reference manual. <http://www.femlab.com>, 2002.
- [86] Ali Javey, Moonsub Shim, and Hongjie Dai. Electrical properties and devices of large-diameter single-walled carbon nanotubes. *Appl. Phys. Lett.*, 80(6):1064–1066, 2002.
- [87] Takeshi Nakanishi, Adrian Bachtold, and Cees Dekker. Transport through the interface between a semiconducting carbon nanotube and a metal electrode. *Phys. Rev. B*, 66:073307–1–073307–4, 2002.
- [88] Serge Luryi. Quantum capacitance devices. *Appl. Phys. Lett.*, 52(6):501–503, 1988.
- [89] D. L. John, L. C. Castro, and D. L. Pulfrey. Quantum capacitance in nanoscale device modeling. *J. Appl. Phys.*, 96(9):5180–5184, 2004.
- [90] P. J. Burke. An RF circuit model for carbon nanotubes. *IEEE Trans. Nanotechnol.*, 2(1):55–58, 2003.
- [91] Peter J. Burke. AC performance of nanoelectronics: Towards a ballistic THz nanotube transistor. *Solid-State Electron.*, 48:1981–1986, 2004.
- [92] D. L. Pulfrey. Modeling high-performance HBTs. In Patrick Roblin and Hans Rohdin, editors, *High-Speed Heterostructure Devices*, chapter 18. Cambridge University Press, 2002.
- [93] David K. Ferry and Stephen M. Goodnick. *Transport in Nanostructures*. Cambridge University Press, New York, 1997.

- [94] Supriyo Datta. *Electronic Transport in Mesoscopic Systems*, volume 3 of *Cambridge Studies in Semiconductor Physics and Microelectronic Engineering*. Cambridge University Press, New York, 1995.
- [95] M. H. Holmes. *Introduction to Perturbation Methods*, volume 20 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995.
- [96] David J. Griffiths. *Introduction to Quantum Mechanics*. Prentice Hall, New Jersey, 1994.
- [97] David Yuk Kei Ko and J. C. Inkson. Matrix method for tunneling in heterostructures: Resonant tunneling in multilayer systems. *Phys. Rev. B*, 38(14):9945–9951, 1988.
- [98] Wayne W. Lui and Masao Fukuma. Exact solution of the Schrodinger equation across an arbitrary one-dimensional piecewise-linear potential barrier. *J. Appl. Phys.*, 60(5):1555–1559, 1986.
- [99] J. Tersoff. Schottky barrier heights and the continuum of gap states. *Phys. Rev. Lett.*, 52(6):465–468, 1984.
- [100] Lee W. Johnson and R. Dean Riess. *Numerical Analysis*. Addison-Wesley, Don Mills, Ontario, 1977.
- [101] Zhen Yao, Cees Dekker, and Phaedon Avouris. Electrical transport through single-wall carbon nanotubes. In Mildred S. Dresselhaus, Gene Dresselhaus, and Phaedon Avouris, editors, *Carbon Nanotubes*, volume 80 of *Topics Appl. Phys.*, pages 147–171. Springer-Verlag, Berlin, 2001.
- [102] Hang Si. *TetGen: A quality tetrahedral mesh generator and three-dimensional Delaunay triangulator (Version 1.3)*, June 2005. [Online.] Available: <http://tetgen.berlios.de>.
- [103] D. L. Pulfrey, D. L. John, and L. C. Castro. Carbon nanotube field-effect transistors: AC performance capabilities. In *Proc. 13th Intl. Workshop Phys. Semiconductor Dev.*, pages 7–13, Delhi, India, December 2005. Invited.
- [104] Richard Martel, Hon-Sum Philip Wong, Kevin Chan, and Phaedon Avouris. Carbon nanotube field effect transistors for logic applications. In *IEDM Tech. Digest*, pages 159–162, 2001.

- [105] R. Martel, V. Derycke, C. Lavoie, J. Appenzeller, K. K. Chan, J. Tersoff, and Ph. Avouris. Ambipolar electrical transport in semiconducting single-wall carbon nanotubes. *Phys. Rev. Lett.*, 87(25):256805–1–256805–4, 2001.
- [106] R. Martel, V. Derycke, J. Appenzeller, S. Wind, and Ph. Avouris. Carbon nanotube field-effect transistors and logic circuits. In *Proc. Design Automation Conf.*, pages 94–98, June 2002.
- [107] M. Radosavljević, S. Heinze, J. Tersoff, and Ph. Avouris. Thermal carrier injection into ambipolar carbon nanotube field effect transistors. In *Molecular Nanostructures*, March 2003.
- [108] Robert Chau, Suman Datta, Mark Doczy, Brian Doyle, Ben Jin, Jack Kavalieros, Amlan Majumdar, Matthew Metz, and Marko Radosavljevic. Benchmarking nanotechnology for high-performance and low-power logic transistor applications. *IEEE Trans. Nanotechnol.*, 4(2):153–158, 2005.
- [109] Jing Guo, Ali Javey, Hongjie Dai, and Mark Lundstrom. Performance analysis and design optimization of near ballistic carbon nanotube field-effect transistors. In *IEDM Tech. Digest*, pages 703–706, December 2004.
- [110] Ali Javey, Jing Guo, Damon B. Farmer, Qian Wang, Erhan Yenilmez, Roy G. Gordon, Mark Lundstrom, and Hongjie Dai. Self-aligned ballistic molecular transistors and electrically parallel nanotube arrays. *Nano Lett.*, 4(7):1319–1322, 2004.
- [111] Jijun Zhao, Jie Han, and Jian Ping Lu. Work functions of pristine and alkali-metal intercalated carbon nanotubes and bundles. *Phys. Rev. B*, 65:193401–1–193401–4, 2002.
- [112] Zhihong Chen, Joerg Appenzeller, Joachim Knoch, Yu-Ming Lin, and Phaedon Avouris. The role of metal-nanotube contact in the performance of carbon nanotube field-effect transistors. *Nano Lett.*, 5(7):1497–1502, 2005.
- [113] Jing Guo. Personal communication.
- [114] Supriyo Datta. Nanoscale device modeling: The Green’s function method. *Superlattices Microstruct.*, 28(4):253–278, 2000.
- [115] H. Flietner. The $E(k)$ relation for a two-band scheme of semiconductors and the application to the metal-semiconductor contact. *Phys. Stat. Sol. (b)*, 54:201–208, 1972.

- [116] A. Persson and R. M. Cohen. Reformulated Hamiltonian for nonparabolic bands in semiconductor quantum wells. *Phys. Rev. B*, 38(8):5568–5575, 1988.
- [117] D. L. John and D. L. Pulfrey. Green’s function calculations for semi-infinite carbon nanotubes. *Phys. Stat. Sol. (b)*, 243(2):442–448, 2006.
- [118] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal. Two-dimensional quantum mechanical modeling of nanotransistors. *J. Appl. Phys.*, 91(4):2343–2354, 2002.
- [119] E. N. Economou. *Green’s Functions in Quantum Physics*, volume 7 of *Springer Series in Solid-State Sciences*. Springer-Verlag, New York, 2nd edition, 1983.
- [120] Erich Zauderer. *Partial Differential Equations of Applied Mathematics*. John Wiley and Sons, Toronto, 2nd edition, 1998.
- [121] E. B. Saff and A. D. Snider. *Fundamentals of Complex Analysis for Mathematics, Science, and Engineering*. Prentice Hall, New Jersey, 2nd edition, 1993.
- [122] Ira N. Levine. *Quantum Chemistry*. Prentice Hall, New Jersey, 5th edition, 2000.
- [123] Ian Appelbaum, Tairan Wang, J. D. Joannopoulos, and V. Narayanamurti. Ballistic hot-electron transport in nanoscale semiconductor heterostructures: Exact self-energy of a three-dimensional periodic tight-binding Hamiltonian. *Phys. Rev. B*, 69:165301–1–165301–6, 2004.
- [124] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic. Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches. *J. Appl. Phys.*, 92(7):3730–3739, 2002.
- [125] Nicholas J. Higham and Hyun-Min Kim. Solving a quadratic matrix equation by Newton’s method with exact line searches. *SIAM J. Matrix Anal. Appl.*, 23(2):303–316, 2001.
- [126] Judith D. Gardiner, Alan J. Laub, James J. Amato, and Cleve B. Moler. Solution of the Sylvester matrix equation $axb^t + cxd^t = e$. *ACM Trans. Math. Softw.*, 18(2):223–231, 1992.
- [127] J. Knoch and J. Appenzeller. A novel concept for field-effect transistors—the tunneling carbon nanotube FET. In *Proc. TMS 63rd Annual Device Research Conference*, pages 100–103, Santa Barbara, U.S.A., June 2005.

- [128] Albert Messiah. *Quantum Mechanics*. Dover, New York, 1999.
- [129] Frederick W. Byron Jr. and Robert W. Fuller. *Mathematics of Classical and Quantum Physics*. Dover, New York, 1992.
- [130] Supriyo Datta. *Quantum Transport: Atom to Transistor*. Cambridge University Press, New York, 2005.
- [131] Zhiben Ren. *Nanoscale MOSFETs: Physics, simulation and design*. PhD thesis, Purdue University, West Lafayette, U.S.A., October 2001.
- [132] R. Lake, G. Klimeck, R. C. Bowen, D. Jovanovic, D. Blanks, and M. Swaminathan. Quantum transport with band-structure and Schottky contacts. *Phys. Stat. Sol. (b)*, 204:354–357, 1997.
- [133] J. W. Mintmire and C. T. White. Universal density of states for carbon nanotubes. *Phys. Rev. Lett.*, 81(12):2506–2509, 1998.
- [134] Arkadi A. Odintsov. Schottky barriers in carbon nanotube heterojunctions. *Phys. Rev. Lett.*, 85(1):150–153, 2000.
- [135] MPI - The message passing interface standard. <http://www-unix.mcs.anl.gov/mpi>, 2000.
- [136] Ramesh Venugopal, Zhibin Ren, and Mark S. Lundstrom. Simulating quantum transport in nanoscale MOSFETs: Ballistic hole transport, subband engineering and boundary conditions. *IEEE Trans. Nanotechnol.*, 2(3):135–143, 2003.
- [137] Steven G. Louie. Electronic properties, junctions, and defects of carbon nanotubes. In Mildred S. Dresselhaus, Gene Dresselhaus, and Phaedon Avouris, editors, *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications*, volume 80 of *Topics in Applied Physics*, pages 113–143. Springer, Berlin, 2001.
- [138] Enzo Ungersboeck, Mahdi Pourfath, Hans Kosina, Andreas Gehring, Byoung-Ho Cheong, Wan-Jun Park, and Siegfried Selberherr. Optimization of single-gate carbon-nanotube field-effect transistors. *IEEE Trans. Nanotechnol.*, 4(5):533–538, 2005.
- [139] M. Pourfath, H. Kosina, B. H. Cheong, W. J. Park, and S. Selberherr. Improving DC and AC characteristics of ohmic contact carbon nanotube field effect transistors. In Gérard Ghibaudo,

- Thomas Skotnicki, Sorin Cristoloveanu, and Michel Brillouët, editors, *Proc. European Solid-State Device Research Conference*, pages 541–544, Grenoble, France, September 2005.
- [140] T. Kamimura, K. Sakamoto, M. Maeda, K. Kurachi, and K. Matsumoto. Fabrication and I-V characterization of carbon nanotube single electron transistor operated at room temperature. In *Proc. TMS 61st Annual Device Research Conference*, pages 177–178, Salt Lake City, U.S.A., June 2003.
 - [141] Jia-Lin Zhu, Hai-Feng Song, and Xiao Hu. Transverse-field and defect-azimuth effects in achiral carbon nanotubes. *J. Phys.: Condens. Matter*, 17:4629–4636, 2005.
 - [142] L. Chico and R. Pérez-Álvarez. Continuum model for long-wavelength optical phonons in cylinders: Application to carbon nanotubes. *Phys. Rev. B*, 69:035419–1–035419–6, 2004.
 - [143] Gerald D. Mahan. *Many-Particle Physics*. Kluwer Academic/Plenum Publishers, New York, 3rd edition, 2000.
 - [144] Mark Bockrath, David H. Cobden, Jia Lu, Andrew G. Rinzler, Richard E. Smalley, Leon Balents, and Paul L. McEuen. Luttinger-liquid behaviour in carbon nanotubes. *Nature*, 397:598–601, 1999.
 - [145] A. Kasumov, M. Kociak, M. Ferrier, R. Deblock, S. Guéron, B. Reulet, I. Khodos, O. Stéphan, and H. Bouchiat. Quantum transport through carbon nanotubes: Proximity-induced and intrinsic superconductivity. *Phys. Rev. B*, 68:214521–1–214521–16, 2003.
 - [146] P. J. Burke. Luttinger liquid theory as a model of the gigahertz electrical properties of carbon nanotubes. *IEEE Trans. Nanotechnol.*, 1(3):129–144, 2002.
 - [147] S. Bellucci, J. González, and P. Onorato. Doping- and size-dependent suppression of tunneling in carbon nanotubes. *Phys. Rev. B*, 69:085404–1–085404–6, 2004.
 - [148] Calin Buia, Alper Buldum, and Jian Ping Lu. Quantum interference effects in electronic transport through nanotube contacts. *Phys. Rev. B*, 67:113409–1–113409–4, 2003.
 - [149] Gordana Dukovic, Feng Wang, Dauhua Song, Matthew Y. Sfeir, Tony F. Heinz, and Louis E. Brus. Structural dependence of excitonic optical transitions and band-gap energies in carbon nanotubes. *Nano Lett.*, 5(11):2314–2318, 2005.

- [150] Murray R. Spiegel. *Mathematical Handbook of Formulas and Tables*. Schaum's Outline Series in Mathematics. McGraw Hill, Toronto, 1968.

Appendix A

Solution to Laplace's Equation

For a closed cylindrical geometry, one may calculate the analytical solution to Laplace's equation with constant potentials applied to the various surfaces of a cylindrical box. The results presented here are standard, see Ref. [82] for example, but we present the derivation here for completeness.

We solve

$$\nabla^2 V = 0 \tag{A.1}$$

in cylindrical coordinates subject to

$$V(0, \phi, z) \text{ is finite,}$$

$$V(R_G, \phi, z) = V_G,$$

$$V(\rho, \phi, 0) = 0,$$

$$V(\rho, \phi, L_C) = V_D,$$

$$V(\rho, \phi, z) = V(\rho, \phi + 2\pi, z),$$

where $\rho = R_G$ corresponds to the gate electrode, $z = 0$ is the source, $z = L_C$ is the drain, and we have chosen our potential reference here such that $V_S = 0$. Note that this solution computes the potential inside the cylindrical box before the nanotube is inserted, *i.e.*, there is no charge inside the box. Using superposition, we can combine this solution with that for the case of homogeneous boundary conditions with charge inside, as derived in Appendix B.

The Laplace problem will be solved as a linear superposition of two solutions. The first solution will consider the case where $V_D = 0$, and the second will be for $V_G = 0$. Since each solution independently satisfies Eq. (A.1), the total solution will, too. It may be seen, as well, that the superposition will satisfy the desired boundary conditions listed above.

A.1 Potential Applied to the Curved Surface

The problem to solve, in cylindrical coordinates, is

$$\frac{\partial^2 V}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial V}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 V}{\partial \phi^2} + \frac{\partial^2 V}{\partial z^2} = 0, \quad (\text{A.2})$$

subject to the boundary conditions mentioned previously. Using separation of variables, we assume a solution of the form

$$V(\rho, \phi, z) = R(\rho)\Theta(\phi)Z(z), \quad (\text{A.3})$$

where

$$R(0) \text{ is finite,}$$

$$R(R_G) = V_G,$$

$$Z(0) = Z(L_C) = 0, \quad (\text{A.4})$$

$$\Theta(\phi) = \Theta(\phi + 2\pi). \quad (\text{A.5})$$

Substitution of Eq. (A.3) into Eq. (A.2) yields

$$\frac{R''}{R} + \frac{R'}{\rho R} + \frac{\Theta''}{\rho^2 \Theta} = \frac{-Z''}{Z}, \quad (\text{A.6})$$

and since the left-hand side is independent of z , and the right-hand side is independent of ϕ and ρ , each side must be equal to a constant, κ^2 . This gives an equation for $Z(z)$ as

$$Z'' + \kappa^2 Z = 0,$$

subject to the boundary conditions given by Eq. (A.4). The general solution is

$$Z(z) = C_1 \sin(\kappa z) + C_2 \cos(\kappa z).$$

$Z(0) = 0$ implies that $C_2 = 0$, and $Z(L_C) = 0$ implies that $\kappa = m\pi/L_C$, $m \in [1, 2, \dots, \infty)$.

The left hand side of Eq. (A.6) can be separated further to give

$$\rho^2 \frac{R''}{R} + \rho \frac{R'}{R} - \kappa^2 \rho^2 = \frac{-\Theta''}{\Theta} = \mu^2, \quad (\text{A.7})$$

where μ is another constant of separation like κ was in the previous separation step. The equation for $\Theta(\phi)$ is similar to the one for $Z(z)$. The general solution is

$$\Theta(\phi) = C_3 \sin(\mu\phi) + C_4 \cos(\mu\phi), \quad (\text{A.8})$$

where Eq. (A.5) implies that $\mu = l$, $l \in [0, 1, \dots, \infty)$.

The equation to be solved for $R(\rho)$ is

$$\rho^2 R'' + \rho R' - (\kappa^2 \rho^2 + l^2) R = 0.$$

Making the substitution $\bar{\rho} = \kappa \rho$ gives

$$\bar{\rho}^2 R'' + \bar{\rho} R' - (\bar{\rho}^2 + l^2) R = 0,$$

which is Bessel's modified differential equation of order l [150]. The general solution is

$$R(\rho) = C_5 I_l(\kappa \rho) + C_6 K_l(\kappa \rho),$$

where I_l and K_l are the modified Bessel functions of the first and second kind, respectively, of order l . Since we require $R(0)$ to be finite, $C_6 = 0$.

Combining the results obtained thus far, and combining the constants, gives

$$V(\rho, \phi, z) = \sum_{l=0}^{\infty} \sum_{m=1}^{\infty} I_l \left(\frac{m\pi\rho}{L_C} \right) \sin \left(\frac{m\pi z}{L_C} \right) [B_{lm} \cos(l\phi) + C_{lm} \sin(l\phi)],$$

where B_{lm} and C_{lm} can be determined, using orthogonality, to be

$$B_{lm} = \frac{\int_0^{2\pi} \int_0^{L_C} V_G \sin \left(\frac{m\pi z}{L_C} \right) \cos(l\phi) dz d\phi}{\int_0^{2\pi} \int_0^{L_C} I_l \left(\frac{m\pi R_G}{L_C} \right) \sin^2 \left(\frac{m\pi z}{L_C} \right) \cos^2(l\phi) dz d\phi},$$

$$C_{lm} = \frac{\int_0^{2\pi} \int_0^{L_C} V_G \sin \left(\frac{m\pi z}{L_C} \right) \sin(l\phi) dz d\phi}{\int_0^{2\pi} \int_0^{L_C} I_l \left(\frac{m\pi R_G}{L_C} \right) \sin^2 \left(\frac{m\pi z}{L_C} \right) \sin^2(l\phi) dz d\phi}.$$

Note that, in our case, since V_G is not a function of ϕ , only the $l = 0$ term will contribute. Using this condition, and performing the integrations, noting that C_{lm} drops out, gives

$$V(\rho, \phi, z) = \sum_{m=1}^{\infty} B_m I_0 \left(\frac{m\pi\rho}{L_C} \right) \sin \left(\frac{m\pi z}{L_C} \right),$$

where

$$B_m = \frac{2V_G(1 - (-1)^m)}{m\pi I_0 \left(\frac{m\pi R_G}{L_C} \right)}.$$

A.2 Potential Applied to One End

Again, we wish to solve Eq. (A.2), but subject to our second set of boundary conditions where the potential is zero everywhere on the boundary except on one end. As before, we separate variables, but for simplicity, we slightly change the separation constant, so that we get the first separation step to look like

$$\frac{R''}{R} + \frac{R'}{\rho R} + \frac{\Theta''}{\rho^2 \Theta} = \frac{-Z''}{Z} = -\kappa^2.$$

In this case, the equation to solve for $Z(z)$ is

$$Z'' - \kappa^2 Z = 0.$$

The general solution is

$$Z(z) = C_1 \sinh(\kappa z) + C_2 \cosh(\kappa z),$$

where $C_2 = 0$ so that $Z(0) = 0$.

The second separation step looks like Eq. (A.7) except with a different sign on the $\kappa^2 \rho^2$ term.

$$\rho^2 \frac{R''}{R} + \rho \frac{R'}{R} + \kappa^2 \rho^2 = \frac{-\Theta''}{\Theta} = \mu^2.$$

The solution for $\Theta(\phi)$ is identical to Eq. (A.8).

The equation to be solved for $R(\rho)$ is now

$$\rho^2 R'' + \rho R' + (\kappa^2 \rho^2 + l^2) R = 0.$$

Making the same substitution as before, $\bar{\rho} = \kappa \rho$, gives

$$\bar{\rho}^2 R'' + \bar{\rho} R' + (\bar{\rho}^2 + l^2) R = 0,$$

which is Bessel's differential equation of order l [150]. The general solution is

$$R(\rho) = C_5 J_l(\kappa \rho) + C_6 Y_l(\kappa \rho),$$

where J_l and Y_l are the Bessel functions of the first and second kind, respectively, of order l . Since we require $R(0)$ to be finite, $C_6 = 0$. We also require that $R(R_G) = 0$ for this case. This implies that

$$\kappa_{lm} = \frac{x_{lm}}{R_G},$$

where x_{lm} is the m th solution to $J_l(x_{lm}) = 0$.

Combining these results gives

$$V(\rho, \phi, z) = \sum_{l=0}^{\infty} \sum_{m=1}^{\infty} J_l\left(\frac{x_{lm}\rho}{R_G}\right) \sinh\left(\frac{x_{lm}z}{R_G}\right) (B_{lm} \sin(l\phi) + C_{lm} \cos(l\phi)),$$

where orthogonality can once again be used to determine our new B_{lm} and C_{lm} according to

$$B_{lm} = \frac{\int_0^{2\pi} \int_0^{R_G} V_D J_l\left(\frac{x_{lm}\rho}{R_G}\right) \sin(l\phi) \rho \, d\rho \, d\phi}{\int_0^{2\pi} \int_0^{R_G} \sinh\left(\frac{x_{lm}L_C}{R_G}\right) J_l^2\left(\frac{x_{lm}\rho}{R_G}\right) \sin^2(l\phi) \rho \, d\rho \, d\phi},$$

$$C_{lm} = \frac{\int_0^{2\pi} \int_0^{R_G} V_D J_l\left(\frac{x_{lm}\rho}{R_G}\right) \cos(l\phi) \rho \, d\rho \, d\phi}{\int_0^{2\pi} \int_0^{R_G} \sinh\left(\frac{x_{lm}L_C}{R_G}\right) J_l^2\left(\frac{x_{lm}\rho}{R_G}\right) \cos^2(l\phi) \rho \, d\rho \, d\phi}.$$

If V_D has no ϕ dependence, only the $l = 0$ term contributes, thus B_{lm} drops out, and

$$C_m = \frac{\int_0^{R_G} V_D J_0\left(\frac{x_{0m}\rho}{R_G}\right) \rho \, d\rho}{\int_0^{R_G} \sinh\left(\frac{x_{0m}L_C}{R_G}\right) J_0^2\left(\frac{x_{0m}\rho}{R_G}\right) \rho \, d\rho},$$

where the integrals may be performed analytically for the case where V_D is a constant.

If we let

$$\hat{\rho} = \frac{x_{0m}\rho}{R_G},$$

then the numerator becomes

$$\frac{R_G^2 V_D}{x_{0m}^2} \int_0^{x_{0m}} \hat{\rho} J_0(\hat{\rho}) \, d\hat{\rho}.$$

This is a standard integral [150] with solution

$$\frac{R_G^2 V_D}{x_{0m}} J_1(x_{0m}).$$

If we now let

$$\hat{\rho} = \frac{\rho}{R_G},$$

then the denominator becomes

$$R_G^2 \sinh\left(\frac{x_{0m}L_C}{R_G}\right) \int_0^1 \hat{\rho} J_0^2(x_{0m}\hat{\rho}) \, d\hat{\rho}.$$

This is also a standard integral which simplifies, since $J_0(x_{0m}) = 0$, to

$$\frac{R_G^2}{2} \sinh\left(\frac{x_{0m}L_C}{R_G}\right) (J_0'(x_{0m}))^2,$$

or

$$\frac{R_G^2}{2} \sinh\left(\frac{x_{0m} L_C}{R_G}\right) J_1^2(x_{0m}).$$

The final solution is

$$V(\rho, \phi, z) = \sum_{m=1}^{\infty} C_m J_0\left(\frac{x_{0m} \rho}{R_G}\right) \sinh\left(\frac{x_{0m} z}{R_G}\right),$$

where

$$C_m = \frac{2V_D}{x_{0m} \sinh\left(\frac{x_{0m} L_C}{R_G}\right) J_1(x_{0m})}.$$

A.3 Total Solution

The total solution to Laplace's equation, subject to our initial boundary conditions, is

$$V(\rho, \phi, z) = \sum_{m=1}^{\infty} \left[B_m I_0\left(\frac{m\pi\rho}{L_C}\right) \sin\left(\frac{m\pi z}{L_C}\right) + C_m J_0\left(\frac{x_{0m}\rho}{R_G}\right) \sinh\left(\frac{x_{0m}z}{R_G}\right) \right], \quad (\text{A.9})$$

where

$$B_m = \frac{2V_G(1 - (-1)^m)}{m\pi I_0\left(\frac{m\pi R_G}{L_C}\right)},$$

$$C_m = \frac{2V_D}{x_{0m} \sinh\left(\frac{x_{0m} L_C}{R_G}\right) J_1(x_{0m})}.$$

Appendix B

Green's Function for a Cylindrical Box

A useful technique in solving Poisson's equation is to use a Green's function. Here, we derive this function for the case of a closed, cylindrical box with homogeneous boundary conditions. This solution may be added to that presented in Appendix A in order to solve Poisson's equation for the closed geometry cylindrical CNFET.

The basic idea is to solve

$$\nabla^2 V = -\frac{q(p-n)}{\epsilon}.$$

In cylindrical coordinates, this is

$$\frac{\partial^2 V}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial V}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 V}{\partial \phi^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{q[p(\rho, \phi, z) - n(\rho, \phi, z)]}{\epsilon}. \quad (\text{B.1})$$

Using the Green's function formalism, we can solve a different equation and use it to solve Eq. (B.1).

We are seeking a function such that

$$V(\rho, \phi, z) = -\frac{q}{\epsilon} \int [p(\rho', \phi', z') - n(\rho', \phi', z')] G_P(\rho, \phi, z; \rho', \phi', z') \rho' d\rho' d\phi' dz', \quad (\text{B.2})$$

where the integral is over the whole space, and G_P is the Green's function. The appropriate equation for G_P [82] is

$$\frac{\partial^2 G_P}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial G_P}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 G_P}{\partial \phi^2} + \frac{\partial^2 G_P}{\partial z^2} = \frac{1}{\rho} \delta(\rho - \rho') \delta(\phi - \phi') \delta(z - z'). \quad (\text{B.3})$$

G_P , then, represents the impulse response of the system to a point charge.

The Green's function for the potential inside a grounded cylindrical box, $0 \leq \rho \leq R_G$ and $0 \leq z \leq L_C$, may be found by solving Eq. (B.3) under the following boundary conditions

$$G_P(0, \phi, z) \text{ is finite,}$$

$$G_P(R_G, \phi, z) = 0,$$

$$G_P(\rho, \phi, 0) = 0,$$

$$G_P(\rho, \phi, L_C) = 0,$$

$$G_P(\rho, \phi, z) = G_P(\rho, \phi + 2\pi, z).$$

This section will demonstrate the technique of eigenfunction expansions [120] for this problem.

We look for the eigenfunctions and eigenvalues of the differential operator, and then express the Green's function as a linear superposition of these functions. If we let \mathcal{L} represent our operator, then Eq. (B.3) may be written as

$$\mathcal{L}G_P = \frac{1}{\rho}\delta(\rho - \rho')\delta(\phi - \phi')\delta(z - z'). \quad (\text{B.4})$$

Let the eigenfunctions of \mathcal{L} be \mathcal{M}_m with their associated eigenvalues, λ_m . Since the \mathcal{M}_m functions are a complete set,

$$G_P = \sum_m N_m \mathcal{M}_m(\rho, \phi, z), \quad (\text{B.5})$$

where N_m are constants determined by the boundary conditions. Eq. (B.4) then becomes

$$\sum_m N_m \mathcal{L}\mathcal{M}_m = \frac{1}{\rho}\delta(\rho - \rho')\delta(\phi - \phi')\delta(z - z'). \quad (\text{B.6})$$

Since \mathcal{M}_m are eigenfunctions of \mathcal{L} ,

$$\mathcal{L}\mathcal{M}_m = \lambda_m \mathcal{M}_m.$$

Since they are also orthogonal, we can multiply both sides of Eq. (B.6) by the complex conjugate of one of the eigenfunctions, \mathcal{M}_m^* , and integrate. If the functions are chosen to be orthonormal, we get

$$\mathcal{M}_m^*(\rho', \phi', z') = N_m \lambda_m,$$

which can be rewritten and substituted into Eq. (B.5) to give

$$G_P(\rho, \phi, z; \rho', \phi', z') = \sum_{m=1}^{\infty} \frac{\mathcal{M}_m^*(\rho', \phi', z') \mathcal{M}_m(\rho, \phi, z)}{\lambda_m}, \quad (\text{B.7})$$

where we assume that none of the eigenvalues vanish.

The eigenfunctions of the differential operator considered here are found by solving

$$\frac{\partial^2 \mathcal{M}}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial \mathcal{M}}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 \mathcal{M}}{\partial \phi^2} + \frac{\partial^2 \mathcal{M}}{\partial z^2} - \lambda \mathcal{M} = 0, \quad (\text{B.8})$$

subject to the same boundary conditions as for G_P above. Using separation of variables, we assume a solution of the form

$$\mathcal{M}(\rho, \phi, z) = R(\rho)\Theta(\phi)Z(z).$$

Substitution of this form into Eq. (B.8) leads to the following set of equations in the usual way:

$$Z'' + \kappa^2 Z = 0, \quad (\text{B.9})$$

$$\Theta'' + \mu^2 \Theta = 0, \quad (\text{B.10})$$

$$\rho^2 R'' + \rho R' + [\rho^2(-\kappa^2 - \lambda) - \mu^2] R = 0, \quad (\text{B.11})$$

where κ and μ are separation constants.

The solution to Eq. (B.9) is

$$Z(z) = C_1 \sin(\kappa z) + C_2 \cos(\kappa z),$$

subject to $Z(0) = Z(L_C) = 0$. This implies that $C_2 = 0$ and $\kappa = j\pi/L_C$, $j \in [1, 2, \dots, \infty)$. Eq. (B.10) has a similar solution in form, but is solved under the constraint that $\Theta(\phi) = \Theta(\phi + 2\pi)$.

This is satisfied for

$$\Theta(\phi) = C_3 e^{i\mu\phi},$$

where the separation constant, μ , must be an integer, l .

For Eq. (B.11), it is useful to make the substitution $\bar{\rho} = \rho\sqrt{-\kappa^2 - \lambda}$ in order to change the form of the equation to

$$\bar{\rho}^2 R'' + \bar{\rho} R' + (\bar{\rho}^2 - \mu^2) R = 0,$$

which we recognize from Appendix A as Bessel's differential equation of order μ . The general solution, noting that $\mu = l$ from the Θ solution, is

$$R(\rho) = C_5 J_l(\bar{\rho}) + C_6 Y_l(\bar{\rho}).$$

Since $R(0)$ must be finite, $C_6 = 0$. The other boundary condition is $R(R_G) = 0$ which implies that

$$R_G \sqrt{-\kappa^2 - \lambda} = x_{lm},$$

where x_{lm} has been defined in Appendix A.

With some algebra, we get the eigenvalues for \mathcal{M} as

$$\lambda_{lmn} = -\left(\frac{x_{lm}}{R_G}\right)^2 - \left(\frac{j\pi}{L_C}\right)^2.$$

Combining the above solutions, we get the eigenfunctions as

$$\mathcal{M}_m(\rho, \phi, z) = \sum_{j=1}^{\infty} \sum_{l=-\infty}^{\infty} C_{jl} e^{il\phi} J_l\left(\frac{x_{lm}\rho}{R_G}\right) \sin\left(\frac{j\pi z}{L_C}\right),$$

where C_{jl} are constants that may be chosen using the orthogonality of eigenfunctions. In this case, we wanted to have orthonormal eigenfunctions, so these constants are chosen such that

$$\int_0^{2\pi} \int_0^{L_C} \int_0^{R_G} C_{jl}^2 J_l^2\left(\frac{x_{lm}\rho}{R_G}\right) \sin^2\left(\frac{j\pi z}{L_C}\right) \rho d\rho dz d\phi = 1.$$

The z and ϕ integrals are trivial, and the ρ integral can be rearranged to become one of those listed in Ref. [150]. The ρ integral is

$$\frac{R_G^2}{2} \left\{ [J'_l(x_{lm})]^2 + \left(1 - \frac{l^2}{x_{lm}^2}\right) J_l^2(x_{lm}) \right\},$$

where the second term vanishes by the definition of x_{lm} . Using Ref. [150] again, this integral becomes

$$\frac{R_G^2}{2} J_{l+1}^2(x_{lm}).$$

Substituting all of this work into Eq. (B.7) produces

$$G_P = -\frac{2}{\pi L_C R_G^2} \sum_{j=1}^{\infty} \sum_{l=-\infty}^{\infty} \sum_{m=1}^{\infty} \frac{e^{il(\phi-\phi')} J_l\left(\frac{x_{lm}\rho}{R_G}\right) J_l\left(\frac{x_{lm}\rho'}{R_G}\right) \sin\left(\frac{j\pi z}{L_C}\right) \sin\left(\frac{j\pi z'}{L_C}\right)}{J_{l+1}^2(x_{lm}) \left[\left(\frac{x_{lm}}{R_G}\right)^2 + \left(\frac{j\pi}{L_C}\right)^2 \right]},$$

in agreement with problem 3.23 in Ref. [82]. Note that only the $l = 0$ term will survive if there is no azimuthal variation.

Appendix C

Conformal Mapping of the Cylindrical Laplace Equation

We show how Laplace's equation changes under a conformal mapping using the methods of complex analysis [121]. This derivation will then be employed to transform the azimuthally-invariant cylindrical Laplace equation into an inverse space [71, 84] as is useful when modeling open boundary CNFETs. First, we will start with some definitions. If we define a function $g(\mathcal{U}) = u(x, y) + iv(x, y)$ where $\mathcal{U} = x + iy$, then we discover that, for g to be differentiable at some point, $\mathcal{U}_0 = x_0 + iy_0$, we require that the Cauchy-Riemann equations hold at that point. These are

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad (C.1)$$

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (C.2)$$

To see this, we can take the definition of the derivative

$$g'(\mathcal{U}_0) = \lim_{\xi \rightarrow 0} \frac{g(\mathcal{U}_0 + \xi) - g(\mathcal{U}_0)}{\xi},$$

and let ξ , a complex number, approach from any direction in the complex plane. If, for example, we approach along the horizontal direction, it is easy to show that

$$g'(\mathcal{U}_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0).$$

If we approach along the vertical direction, we get

$$g'(\mathcal{U}_0) = -i \frac{\partial u}{\partial y}(x_0, y_0) + \frac{\partial v}{\partial y}(x_0, y_0).$$

These two expressions must be equal for the derivative to exist, therefore Eqs. (C.1) and (C.2) are necessary conditions for the function to be differentiable. It turns out that if the first partial

derivatives of u and v are continuous and satisfy the Cauchy-Riemann equations at all points in the domain, then g is analytic in the domain, *i.e.*, it is differentiable everywhere inside the domain. See page 59 of Ref. [121] for a proof of this.

C.1 Transformation of the Cartesian Laplace Equation

Suppose now that we have Laplace's equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

where V is a real-valued function. The source of V may be represented as the real and imaginary parts of an analytic function. Since the real and imaginary parts of an analytic function have continuous partial derivatives to any order,

$$\frac{\partial}{\partial y} \frac{\partial u}{\partial x} = \frac{\partial}{\partial x} \frac{\partial u}{\partial y} \quad (\text{C.3})$$

implies that

$$\frac{\partial^2 v}{\partial y^2} = -\frac{\partial^2 v}{\partial x^2}, \quad (\text{C.4})$$

where the Cauchy-Riemann equations have been used to get Eq. (C.4) from the elementary calculus result of Eq. (C.3). A similar result can be derived for u . Since the real and imaginary parts of an analytic function satisfy Laplace's equation, we can define a complex potential, and solve in the complex plane. This allows us to use conformal mapping to solve our problem.

We now show how Laplace's equation transforms under a conformal mapping. The first important point is that the mapping must be one-to-one. This means that each point in the original \mathcal{U} -plane must map to a unique point in the mapped space, the \mathcal{T} -plane. If $\mathcal{T} = \mathcal{T}(\mathcal{U})$ represents the mapping from one space to the other, this condition is $\mathcal{T}(\mathcal{U}_1) = \mathcal{T}(\mathcal{U}_2)$ if and only if $\mathcal{U}_1 = \mathcal{U}_2$. It will be shown that, if $\mathcal{T}(\mathcal{U})$ is an analytic function itself, then Laplace's equation transforms in a very useful way. Namely, we recover Laplace's equation exactly in either domain.

We will consider the specific case of the

$$\mathcal{T}(\mathcal{U}) = \frac{1}{\mathcal{U}}$$

transformation as this is of current interest, but the results and steps may be easily used with any other analytic mapping. First, we will confirm that this is an analytic function by checking that it

satisfies the Cauchy-Riemann equations, and that it is continuous everywhere. Note that we will consider a domain in the \mathcal{U} -plane where the origin is not included, so this technicality will not be considered. Essentially, we consider the transformation of the space outside some disk into the interior of a different disk.

$$\frac{1}{\mathcal{U}} = \frac{1}{x + iy} = \frac{x - iy}{x^2 + y^2} \equiv u + iv$$

implies that

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} = -\frac{x^2 - y^2}{(x^2 + y^2)^2}, \\ \frac{\partial v}{\partial y} &= \frac{-(x^2 + y^2) + 2y^2}{(x^2 + y^2)^2} = -\frac{x^2 - y^2}{(x^2 + y^2)^2}, \\ \frac{\partial u}{\partial y} &= -\frac{2xy}{(x^2 + y^2)^2} = -\frac{\partial v}{\partial x},\end{aligned}$$

and hence the first partial derivatives are continuous and satisfy the Cauchy-Riemann equations, so $\mathcal{T}(\mathcal{U})$ is analytic.

Following the steps outlined in problem 2 of Section 7.1 in Ref. [121], we let

$$\bar{V}(u, v) = V(x(u, v), y(u, v)) \tag{C.5}$$

where u and v are the coordinates in the \mathcal{T} -plane. The gradient of V is given by

$$\nabla_{\mathcal{U}} V = \left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y} \right) = \frac{\partial V}{\partial x} + i \frac{\partial V}{\partial y},$$

where the subscript indicates that the gradient is performed in the \mathcal{U} -plane. Similarly,

$$\nabla_{\mathcal{T}} \bar{V} = \left(\frac{\partial \bar{V}}{\partial u}, \frac{\partial \bar{V}}{\partial v} \right) = \frac{\partial \bar{V}}{\partial u} + i \frac{\partial \bar{V}}{\partial v}.$$

Using the chain rule on Eq. (C.5) yields

$$\begin{aligned}\frac{\partial \bar{V}}{\partial u} &= \frac{\partial V}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial u}, \\ \frac{\partial \bar{V}}{\partial v} &= \frac{\partial V}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial v}.\end{aligned}$$

This gives a relationship between $\nabla_{\mathcal{T}} \bar{V}$ and $\nabla_{\mathcal{U}} V$. Using the Cauchy-Riemann equations,

$$\nabla_{\mathcal{T}} \bar{V} = \frac{\partial V}{\partial x} \frac{\partial y}{\partial v} - \frac{\partial V}{\partial y} \frac{\partial x}{\partial v} + i \left(\frac{\partial V}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial v} \right).$$

Rearranging and factoring yields

$$\nabla_T \bar{V} = \nabla_U V \left(\frac{\partial y}{\partial v} + i \frac{\partial x}{\partial v} \right)$$

where we can, again, use the Cauchy-Riemann equations to get

$$\nabla_T \bar{V} = \nabla_U V \left(\frac{\partial x}{\partial u} - i \frac{\partial y}{\partial u} \right) = \nabla_U V \left(\frac{dU}{dT} \right)^*$$

where the star denotes a complex conjugate. This motivates us to define the gradient operator in the T -plane as

$$\nabla_T = \left(\frac{dU}{dT} \right)^* \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right).$$

The Laplacian of a function, g , is defined by $\nabla^2 g = \nabla \cdot \nabla g$, therefore we can write ∇_T^2 as

$$\nabla_T^2 = \left| \frac{dU}{dT} \right|^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right),$$

where we recall that a dot product of a vector with itself in the complex plane involves taking the complex conjugate of one of the arguments and multiplying. Finally, we have

$$\frac{\partial^2 \bar{V}}{\partial u^2} + \frac{\partial^2 \bar{V}}{\partial v^2} = \left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} \right) \left| \frac{dU}{dT} \right|^2. \quad (\text{C.6})$$

If

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = N_f(x, y),$$

then

$$\left(\frac{\partial^2 \bar{V}}{\partial u^2} + \frac{\partial^2 \bar{V}}{\partial v^2} \right) \left| \frac{dT}{dU} \right|^2 = N_f(x(u, v), y(u, v))$$

may be solved in the transformed space. If $N_f(x, y) = 0$, then we can solve

$$\frac{\partial^2 \bar{V}}{\partial u^2} + \frac{\partial^2 \bar{V}}{\partial v^2} = 0.$$

In other words, Laplace's equation is invariant under the transformation.

C.2 Transformation of the Cylindrical Laplace Equation

This section transforms a three-dimensional Laplace equation, with azimuthal symmetry, into the space defined by

$$\mathcal{T} = \frac{1}{\mathcal{U}}.$$

The equation we wish to transform is

$$\frac{\partial^2 V}{\partial x^2} + \frac{1}{y} \frac{\partial V}{\partial y} + \frac{\partial^2 V}{\partial y^2} = 0,$$

where we recall the second derivatives may be transformed using Eq. (C.6). It now remains to transform the second term in this equation. First, we recall that

$$x = \frac{u}{u^2 + v^2},$$

$$y = -\frac{v}{u^2 + v^2},$$

for a $T = 1/\mathcal{U}$ transformation. For convenience, we reproduce our starting four equations:

$$\frac{\partial \bar{V}}{\partial u} = \frac{\partial V}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial u}, \quad (\text{C.7})$$

$$\frac{\partial \bar{V}}{\partial v} = \frac{\partial V}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial v}, \quad (\text{C.8})$$

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y},$$

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

We can multiply Eqs. (C.7) and (C.8) in order to help us eliminate the $\partial V/\partial x$ terms:

$$\frac{\partial x}{\partial v} \frac{\partial \bar{V}}{\partial u} = \frac{\partial V}{\partial x} \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial u} \frac{\partial x}{\partial v},$$

$$\frac{\partial x}{\partial u} \frac{\partial \bar{V}}{\partial v} = \frac{\partial V}{\partial x} \frac{\partial x}{\partial u} \frac{\partial x}{\partial v} + \frac{\partial V}{\partial y} \frac{\partial y}{\partial v} \frac{\partial x}{\partial u}.$$

Subtracting these gives

$$\frac{\partial x}{\partial v} \frac{\partial \bar{V}}{\partial u} - \frac{\partial x}{\partial u} \frac{\partial \bar{V}}{\partial v} = \frac{\partial V}{\partial y} \left(\frac{\partial y}{\partial u} \frac{\partial x}{\partial v} - \frac{\partial y}{\partial v} \frac{\partial x}{\partial u} \right),$$

which can be rewritten using the Cauchy-Riemann equations as

$$\frac{\partial x}{\partial v} \frac{\partial \bar{V}}{\partial u} - \frac{\partial x}{\partial u} \frac{\partial \bar{V}}{\partial v} = -\frac{\partial V}{\partial y} \left[\left(\frac{\partial x}{\partial u} \right)^2 + \left(\frac{\partial y}{\partial u} \right)^2 \right].$$

This implies that

$$\frac{\partial V}{\partial y} = \left[\left(\frac{\partial x}{\partial u} \right)^2 + \left(\frac{\partial y}{\partial u} \right)^2 \right]^{-1} \left(\frac{\partial y}{\partial u} \frac{\partial \bar{V}}{\partial u} + \frac{\partial x}{\partial u} \frac{\partial \bar{V}}{\partial v} \right),$$

where Cauchy-Riemann has, again, been used. Putting this all together, and dividing out common factors, gives the transformed equation as

$$(u^2 + v^2) \left(\frac{\partial^2 \bar{V}}{\partial u^2} + \frac{\partial^2 \bar{V}}{\partial v^2} \right) - 2u \frac{\partial \bar{V}}{\partial u} + \left(\frac{u^2 - v^2}{v} \right) \frac{\partial \bar{V}}{\partial v} = 0.$$

Appendix D

The WKB Approximation

Here, we provide details on using the WKB approximation for solving the one-dimensional, effective-mass Schrödinger equation following the derivation in Ref. [95].

We wish to solve

$$\frac{\hbar^2}{2m_{\text{eff}}} \frac{\partial^2 \psi}{\partial z^2} = (U(z) - E) \psi, \quad (\text{D.1})$$

for the case where we have an incident electron, of energy E and wavefunction ψ , on a potential energy barrier $U(z)$. Note that m_{eff} is the effective mass which is assumed constant throughout the solution space. It will be assumed that $E < U(z)$ for $z_0 < z < z_1$, $E > U(z)$ when $z < z_0$ or $z > z_1$, and $E = U(z)$ when $z = z_0$ or $z = z_1$.

If we define

$$\bar{z} = \frac{z}{w},$$

$$\xi = \frac{\hbar^2}{2m_{\text{eff}} w^2},$$

where $w = z_1 - z_0$, then we can rewrite Eq. (D.1) as

$$\xi \frac{\partial^2 \psi}{\partial \bar{z}^2} = (U(\bar{z}) - E) \psi. \quad (\text{D.2})$$

The key to the WKB method is assuming a solution of the form

$$\psi(\bar{z}) = \mathcal{K}(\bar{z}) \exp \left(i \frac{\theta(\bar{z})}{\xi^\iota} \right),$$

where $\mathcal{K}(\bar{z})$ and $\theta(\bar{z})$ are complex functions, $\xi \ll 1$, and ι is a constant. We further assume that $\mathcal{K}(\bar{z})$ can be expressed in a series in powers of ξ , *i.e.*,

$$\mathcal{K}(\bar{z}) = \mathcal{K}_0(\bar{z}) + \xi^\iota \mathcal{K}_1(\bar{z}) + \xi^{2\iota} \mathcal{K}_2(\bar{z}) + \dots$$

Substituting these into Eq. (D.2) and rearranging yields

$$\xi \mathcal{K}_{\bar{z}\bar{z}} - \xi^{1-2\iota} \mathcal{K}(\theta_{\bar{z}})^2 + 2i\xi^{1-\iota} \mathcal{K}_{\bar{z}} \theta_{\bar{z}} + i\xi^{1-\iota} \mathcal{K} \theta_{\bar{z}\bar{z}} + (E - U) \mathcal{K} = 0,$$

where a subscripted coordinate denotes a partial derivative with respect to that variable. Note that the exponential term cancels out of this expression. We now have an equation that can be solved for asymptotically.

D.1 Approximate Solution

At this point, we note that by setting $\iota = 1/2$, and substituting our power series for \mathcal{K} , we get a sequence of equations in powers of ξ . Comparing terms that do not depend on ξ gives

$$-\mathcal{K}_0 (\theta_{\bar{z}})^2 + (E - U) \mathcal{K}_0 = 0,$$

which is known as the eikonal equation if we cancel the \mathcal{K}_0 out of this expression. Solving yields

$$\theta(\bar{z}) = \pm \int^{\bar{z}} \sqrt{E - U(\hat{z})} d\hat{z},$$

where the lower limit of integration may be chosen to suit the problem.

Comparing terms that involve ξ^1 yields

$$-\mathcal{K}_1 (\theta_{\bar{z}})^2 + 2i\mathcal{K}_{0\bar{z}}\theta_{\bar{z}} + i\mathcal{K}_0\theta_{\bar{z}\bar{z}} + (E - U) \mathcal{K}_1 = 0.$$

This is referred to as the transport equation. Since θ is a solution to the eikonal equation, this expression simplifies to

$$\frac{\mathcal{K}_{0\bar{z}}}{\mathcal{K}_0} = -\frac{\theta_{\bar{z}\bar{z}}}{2\theta_{\bar{z}}},$$

with the solution

$$\mathcal{K}_0 = C \left(\frac{1}{|E - U|} \right)^{\frac{1}{4}},$$

where C is a constant.

D.2 Error

In order to get a measure of the error, we can look at the next term in the expansion. Comparing terms involving powers of $\xi^{2\iota}$ gives

$$\mathcal{K}_{0\bar{z}\bar{z}} - \mathcal{K}_2 (\theta_{\bar{z}})^2 + 2i\mathcal{K}_{1\bar{z}}\theta_{\bar{z}} + i\mathcal{K}_1\theta_{\bar{z}\bar{z}} + (E - U) \mathcal{K}_2 = 0,$$

which again simplifies, due to the eikonal equation, giving

$$\mathcal{K}_{0\bar{z}\bar{z}} + 2i\mathcal{K}_{1\bar{z}}\theta_{\bar{z}} + i\mathcal{K}_1\theta_{\bar{z}\bar{z}} = 0.$$

If we let $\mathcal{K}_1(\bar{z}) = i\mathcal{K}_0(\bar{z})W(\bar{z})$, and substitute for \mathcal{K}_0 in terms of θ , we get

$$W_{\bar{z}} = \frac{1}{4} \frac{\theta_{\bar{z}\bar{z}\bar{z}}}{(\theta_{\bar{z}})^2} - \frac{3}{8} \frac{(\theta_{\bar{z}\bar{z}})^2}{(\theta_{\bar{z}})^3},$$

which implies that

$$W = C + \frac{1}{4} \frac{\theta_{\bar{z}\bar{z}}}{(\theta_{\bar{z}})^2} - \frac{1}{8} \int^{\bar{z}} \frac{(\theta_{\bar{z}\bar{z}}(\hat{z}))^2}{(\theta_{\bar{z}}(\hat{z}))^3} d\hat{z}.$$

We have a good expansion if $|\xi\mathcal{K}_1| \ll |\mathcal{K}_0|$, i.e., , if $|\xi W| \ll 1$. If we write this condition in terms of $E - U$, we observe that the expansion breaks down near the turning points. We need to perform a more careful analysis about these points.

D.3 Turning Points and Matching

In regions where the WKB expansion is valid, we can write down the general form of the wavefunction subject to the conditions specified above. This is

$$\psi(\bar{z}) = \begin{cases} \left(\frac{1}{E-U}\right)^{\frac{1}{4}} \left(C_1 e^{i \int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{E-U(\hat{z})}{\xi}} d\hat{z}} + C_2 e^{-i \int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{E-U(\hat{z})}{\xi}} d\hat{z}} \right), & \bar{z} < \bar{z}_0 \\ \left(\frac{1}{U-E}\right)^{\frac{1}{4}} \left(C_3 e^{\int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z}} + C_4 e^{-\int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z}} \right), & \bar{z}_0 < \bar{z} < \bar{z}_1 \\ \left(\frac{1}{E-U}\right)^{\frac{1}{4}} C_5 e^{i \int_{\bar{z}_1}^{\bar{z}} \sqrt{\frac{E-U(\hat{z})}{\xi}} d\hat{z}}, & \bar{z} > \bar{z}_1 \end{cases}$$

where C_1 through C_5 are constants that need to be determined, $\bar{z}_0 = z_0/w$, and $\bar{z}_1 = z_1/w$. As we have already mentioned, these results are not valid close to \bar{z}_0 and \bar{z}_1 .

If we examine the region close to \bar{z}_0 , and expand $E - U(\bar{z})$ about this point, we get

$$E - U(\bar{z}) \simeq -U'(\bar{z}_0)(\bar{z} - \bar{z}_0),$$

where the primed notation denotes a total derivative, and we have assumed that U' exists. In order to examine this region closely, we define $\bar{z} - \bar{z}_0 = \xi^{\beta_0} \zeta$ where β_0 is a constant and ζ is an expanded coordinate representation. Substitution gives

$$\xi^{1-2\beta_0} \psi_{\zeta\zeta} - U'(\bar{z}_0) \xi^{\beta_0} \zeta \psi = 0,$$

where we note that $U'(\bar{z}_0) > 0$ from our problem definition, and we choose $\beta_0 = 1/3$ to simplify the equation. If we further define

$$\alpha_0 = [U'(\bar{z}_0)]^{\frac{1}{3}} \zeta,$$

we get

$$\psi_{\alpha_0 \alpha_0} = \alpha_0 \psi,$$

which is Airy's equation. The general solution is

$$\psi(\alpha_0) = C_6 \text{Ai}(\alpha_0) + C_7 \text{Bi}(\alpha_0),$$

where C_6 and C_7 are constants, and $\text{Ai}(\cdot)$ and $\text{Bi}(\cdot)$ are the Airy functions of the first and second kinds, respectively.

As $\alpha_0 \rightarrow \infty$, we can look at the asymptotic behaviour of these functions. Using the forms presented in Ref. [96], we get that

$$\psi(\alpha_0) \sim \frac{C_6}{2\sqrt{\pi}\alpha_0^{\frac{1}{4}}} \exp\left(-\frac{2}{3}\alpha_0^{\frac{3}{2}}\right) + \frac{C_7}{\sqrt{\pi}\alpha_0^{\frac{1}{4}}} \exp\left(\frac{2}{3}\alpha_0^{\frac{3}{2}}\right).$$

Similarly, as $\alpha_0 \rightarrow -\infty$, we have

$$\frac{\exp\left(i\frac{2}{3}(-\alpha_0)^{\frac{3}{2}}\right)}{2\sqrt{\pi}(-\alpha_0)^{\frac{1}{4}}} \left(C_6 e^{-\frac{i\pi}{4}} + C_7 e^{\frac{i\pi}{4}}\right) + \frac{\exp\left(-i\frac{2}{3}(-\alpha_0)^{\frac{3}{2}}\right)}{2\sqrt{\pi}(-\alpha_0)^{\frac{1}{4}}} \left(C_6 e^{\frac{i\pi}{4}} + C_7 e^{-\frac{i\pi}{4}}\right).$$

In order to have a good solution, the asymptotic solution must join smoothly to the solution that is valid away from the turning point. If we recall that $U(\bar{z}) - E \simeq U'(\bar{z}_0)(\bar{z} - \bar{z}_0)$, then

$$\frac{1}{U - E} \simeq \frac{1}{[U'(\bar{z}_0)]^{\frac{2}{3}} \xi^{\frac{1}{3}} \alpha_0}.$$

Similarly,

$$\int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{U(\hat{z}) - E}{\xi}} d\hat{z} \simeq \sqrt{\frac{U'(\bar{z}_0)}{\xi}} \int_{\bar{z}_0}^{\bar{z}} \sqrt{\hat{z} - \bar{z}_0} d\hat{z} = \frac{2}{3} \alpha_0^{\frac{3}{2}}.$$

Near \bar{z}_0 , for $\bar{z} > \bar{z}_0$, we can write our WKB solution as

$$\frac{1}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}} \alpha_0^{\frac{1}{4}}} \left[C_3 \exp\left(\frac{2}{3}\alpha_0^{\frac{3}{2}}\right) + C_4 \exp\left(-\frac{2}{3}\alpha_0^{\frac{3}{2}}\right) \right].$$

Matching coefficients yields

$$\frac{C_3}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}}} = \frac{C_7}{\sqrt{\pi}},$$

$$\frac{C_4}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}}} = \frac{C_6}{2\sqrt{\pi}}.$$

If we consider $\bar{z} < \bar{z}_0$ near \bar{z}_0 , we note that

$$\frac{1}{E - U} \simeq \frac{1}{[U'(\bar{z}_0)]^{\frac{2}{3}} \xi^{\frac{1}{3}} (-\alpha_0)},$$

and

$$\int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{E - U(\hat{z})}{\xi}} d\hat{z} \simeq \sqrt{\frac{U'(\bar{z}_0)}{\xi}} \int_{\bar{z}_0}^{\bar{z}} \sqrt{\bar{z}_0 - \hat{z}} d\hat{z} = -\frac{2}{3}(-\alpha_0)^{\frac{3}{2}},$$

which provides the approximate form for the wavefunction in this region as

$$\frac{1}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}} (-\alpha_0)^{\frac{1}{4}}} \left[C_1 \exp\left(-i\frac{2}{3}(-\alpha_0)^{\frac{3}{2}}\right) + C_2 \exp\left(i\frac{2}{3}(-\alpha_0)^{\frac{3}{2}}\right) \right].$$

If we equate coefficients once again, we get

$$\begin{aligned} \frac{C_1}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}}} &= \frac{1}{2\sqrt{\pi}} \left(C_6 e^{\frac{i\pi}{4}} + C_7 e^{-\frac{i\pi}{4}} \right), \\ \frac{C_2}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}}} &= \frac{1}{2\sqrt{\pi}} \left(C_6 e^{-\frac{i\pi}{4}} + C_7 e^{\frac{i\pi}{4}} \right). \end{aligned}$$

This concludes the matching that is performed near the first turning point.

If we examine the region close to \bar{z}_1 , we get

$$E - U(\bar{z}) \simeq |U'(\bar{z}_1)|(\bar{z} - \bar{z}_1).$$

Again, we perform a stretching transformation via $\bar{z} - \bar{z}_1 = \xi^{\beta_1} \zeta$ where we choose $\beta_1 = 1/3$ as before.

If we let

$$\alpha_1 = |U'(\bar{z}_1)|^{\frac{1}{3}} \zeta,$$

the equation to solve is

$$\psi_{\alpha_1 \alpha_1} = -\alpha_1 \psi,$$

with the general solution

$$C_8 \text{Ai}(-\alpha_1) + C_9 \text{Bi}(-\alpha_1),$$

where C_8 and C_9 are constants.

Now, as $-\alpha_1 \rightarrow \infty$, we have

$$\psi(\alpha_1) \sim \frac{C_8}{2\sqrt{\pi}(-\alpha_1)^{\frac{1}{4}}} \exp\left(-\frac{2}{3}(-\alpha_1)^{\frac{3}{2}}\right) + \frac{C_9}{\sqrt{\pi}(-\alpha_1)^{\frac{1}{4}}} \exp\left(\frac{2}{3}(-\alpha_1)^{\frac{3}{2}}\right).$$

Proceeding as before, we note that

$$\frac{1}{U-E} \simeq \frac{1}{|U'(\bar{z}_1)|^{\frac{2}{3}} \xi^{\frac{1}{3}} (-\alpha_1)}.$$

Note that

$$\int_{\bar{z}_0}^{\bar{z}} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z} = \int_{\bar{z}_0}^{\bar{z}_1} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z} + \int_{\bar{z}_1}^{\bar{z}} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z},$$

so we can define

$$\nu = \int_{\bar{z}_0}^{\bar{z}_1} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z},$$

which is independent of \bar{z} . Furthermore,

$$\int_{\bar{z}_1}^{\bar{z}} \sqrt{\frac{U(\hat{z})-E}{\xi}} d\hat{z} \simeq \sqrt{\frac{|U'(\bar{z}_1)|}{\xi}} \int_{\bar{z}_1}^{\bar{z}} \sqrt{\bar{z}_1 - \hat{z}} d\hat{z} = -\frac{2}{3}(-\alpha_1)^{\frac{3}{2}},$$

so our wavefunction in this region is

$$\frac{1}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}} (-\alpha_1)^{\frac{1}{4}}} \left[C_3 \exp\left(\nu - \frac{2}{3}(-\alpha_1)^{\frac{3}{2}}\right) + C_4 \exp\left(-\nu + \frac{2}{3}(-\alpha_1)^{\frac{3}{2}}\right) \right].$$

Performing the matching yields

$$\begin{aligned} \frac{C_3}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}}} e^{\nu} &= \frac{C_8}{2\sqrt{\pi}}, \\ \frac{C_4}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}}} e^{-\nu} &= \frac{C_9}{\sqrt{\pi}}. \end{aligned}$$

Finally, we must match as $-\alpha_1 \rightarrow -\infty$ where the asymptotic form gives us

$$\frac{\exp\left(i\frac{2}{3}\alpha_1^{\frac{3}{2}}\right)}{2\sqrt{\pi}\alpha_1^{\frac{1}{4}}} \left(C_8 e^{-\frac{i\pi}{4}} + C_9 e^{\frac{i\pi}{4}}\right) + \frac{\exp\left(-i\frac{2}{3}\alpha_1^{\frac{3}{2}}\right)}{2\sqrt{\pi}\alpha_1^{\frac{1}{4}}} \left(C_8 e^{\frac{i\pi}{4}} + C_9 e^{-\frac{i\pi}{4}}\right).$$

As before,

$$\begin{aligned} \frac{1}{E-U} &\simeq \frac{1}{|U'(\bar{z}_1)|^{\frac{2}{3}} \xi^{\frac{1}{3}} \alpha_1}, \\ \int_{\bar{z}_1}^{\bar{z}} \sqrt{\frac{E-U(\hat{z})}{\xi}} d\hat{z} &\simeq \sqrt{\frac{|U'(\bar{z}_1)|}{\xi}} \int_{\bar{z}_1}^{\bar{z}} \sqrt{\hat{z} - \bar{z}_1} d\hat{z} = \frac{2}{3}\alpha_1^{\frac{3}{2}}. \end{aligned}$$

The function to match to becomes

$$\frac{C_5}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}} \alpha_1^{\frac{1}{4}}} \exp\left(i\frac{2}{3}\alpha_1^{\frac{3}{2}}\right),$$

which implies that

$$\frac{C_5}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}}} = \frac{1}{2\sqrt{\pi}} \left(C_8 e^{-\frac{i\pi}{4}} + C_9 e^{\frac{i\pi}{4}} \right),$$

$$C_8 e^{\frac{i\pi}{4}} = -C_9 e^{-\frac{i\pi}{4}}.$$

The matching is now complete, and we can solve for all but one of the unknown constants. The result, in terms of C_1 , is

$$C_2 = -\frac{iC_1 (4e^{2\nu} - 1)}{4e^{2\nu} + 1},$$

$$C_3 = \frac{2C_1 \sqrt{2}}{(4e^{2\nu} + 1)(1 - i)},$$

$$C_4 = -\frac{4iC_1 \sqrt{2}e^{2\nu}}{(4e^{2\nu} + 1)(1 - i)},$$

$$C_5 = \frac{4C_1 e^\nu}{4e^{2\nu} + 1},$$

$$C_6 = -\frac{8iC_1 \sqrt{2\pi}e^{2\nu}}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}} (4e^{2\nu} + 1)(1 - i)},$$

$$C_7 = \frac{2C_1 \sqrt{2\pi}}{[U'(\bar{z}_0)]^{\frac{1}{6}} \xi^{\frac{1}{12}} (4e^{2\nu} + 1)(1 - i)},$$

$$C_8 = \frac{4C_1 \sqrt{2\pi}e^\nu}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}} (4e^{2\nu} + 1)(1 - i)},$$

$$C_9 = -\frac{4iC_1 \sqrt{2\pi}e^\nu}{|U'(\bar{z}_1)|^{\frac{1}{6}} \xi^{\frac{1}{12}} (4e^{2\nu} + 1)(1 - i)}.$$

Finally, we can compute the transmission probability T . If $U(\bar{z})$ is constant outside of the turning points, and if we define

$$U(\bar{z}) = \begin{cases} U_S, & \bar{z} < \bar{z}_0 \\ U_D, & \bar{z} > \bar{z}_1 \end{cases}$$

then the transmission probability is given by

$$\begin{aligned} T &= \sqrt{\frac{E - U_D}{E - U_S}} \frac{|C_5|^2}{|C_1|^2}, \\ &= \sqrt{\frac{E - U_D}{E - U_S}} \left(\frac{4e^\nu}{4e^{2\nu} + 1} \right)^2, \\ &\simeq \sqrt{\frac{E - U_D}{E - U_S}} e^{-2\nu}, \end{aligned}$$

where the last line is valid if $e^{2\nu} \gg 1$.

Appendix E

Solving for the Newton Update

In Chapter 5, we formulated a quadratic matrix equation for the self-energy matrix representing a semi-infinite lead. Here, we show how one may compute the Newton update at each step of the iterative procedure, and how one may employ exact line searches in order to improve the convergence rate.

E.1 Computing the Raw Newton Update

We consider the equation, from Chapter 5,

$$r(\tilde{g}_s) + \tau^\dagger \varepsilon \tau \tilde{g}_s + [\tau^\dagger \tilde{g}_s \tau - (EI - \Upsilon_L)] \varepsilon = 0, \quad (\text{E.1})$$

where we seek ε . We define [125]

$$\bar{\mathcal{P}} = Q_1 \tau^\dagger Z_1,$$

$$\bar{\mathcal{S}} = Q_1 [\tau^\dagger \tilde{g}_s \tau - (EI - \Upsilon_L)] Z_1,$$

$$\bar{\mathcal{T}} = Q_2 Z_2,$$

$$\bar{\mathcal{R}} = Q_2 \tilde{g}_s^\dagger \tau^\dagger Z_2,$$

where $\bar{\mathcal{P}}$, $\bar{\mathcal{S}}$, $\bar{\mathcal{T}}$, and $\bar{\mathcal{R}}$ are, in this case, upper triangular, and Q_i and Z_i are unitary. These matrices may all be found via QZ factorizations, *i.e.*, $\bar{\mathcal{P}}$, $\bar{\mathcal{S}}$, Q_1 , and Z_1 are found from a QZ factorization of τ^\dagger and $\tau^\dagger \tilde{g}_s \tau - (EI - \Upsilon_L)$, and $\bar{\mathcal{T}}$, $\bar{\mathcal{R}}$, Q_2 , and Z_2 are found from a QZ factorization of the identity matrix and $\tilde{g}_s^\dagger \tau^\dagger$.

Premultiplying by Q_1 and postmultiplying by Q_2^\dagger allows Eq. (E.1) to be written in the equivalent form [125]

$$Q_1 \tau^\dagger Z_1 Z_1^\dagger \varepsilon Z_2 Z_2^\dagger \tau \tilde{g}_s Q_2^\dagger + Q_1 [\tau^\dagger \tilde{g}_s \tau - (EI - \Upsilon_L)] Z_1 Z_1^\dagger \varepsilon Z_2 Z_2^\dagger Q_2^\dagger = -Q_1 r(\tilde{g}_s) Q_2^\dagger.$$

Making the definitions $\mathcal{Y} = Z_1^\dagger \varepsilon Z_2$, and $\mathcal{F} = -Q_1^T(\bar{g}_s)Q_2^\dagger$ yields

$$\bar{\mathcal{P}}\mathcal{Y}\bar{\mathcal{R}}^\dagger + \bar{\mathcal{S}}\mathcal{Y}\bar{\mathcal{T}}^\dagger = \mathcal{F},$$

where we now seek \mathcal{Y} . If subscripts to a matrix represent the row and column entry, we may write this in the equivalent form

$$\mathcal{F}_{i,j} = \sum_{m=1}^N \sum_{l=1}^N (\bar{\mathcal{P}}_{i,m} \mathcal{Y}_{m,l} \bar{\mathcal{R}}_{j,l}^* + \bar{\mathcal{S}}_{i,m} \mathcal{Y}_{m,l} \bar{\mathcal{T}}_{j,l}^*),$$

where the star denotes complex conjugation. Since $\bar{\mathcal{P}}$, $\bar{\mathcal{S}}$, $\bar{\mathcal{R}}$, and $\bar{\mathcal{T}}$ result from a QZ factorization, many of the terms in each sum vanish. Eliminating these terms yields

$$\mathcal{F}_{i,j} = \sum_{m=i}^N \sum_{l=j}^N (\bar{\mathcal{P}}_{i,m} \mathcal{Y}_{m,l} \bar{\mathcal{R}}_{j,l}^* + \bar{\mathcal{S}}_{i,m} \mathcal{Y}_{m,l} \bar{\mathcal{T}}_{j,l}^*).$$

If we use the notation of Ref. [126], where a single subscript to a matrix denotes a column of the matrix, we can also write this as

$$\mathcal{F}_m = \bar{\mathcal{P}} \sum_{j=m}^N \mathcal{Y}_j \bar{\mathcal{R}}_{m,j}^* + \bar{\mathcal{S}} \sum_{j=m}^N \mathcal{Y}_j \bar{\mathcal{T}}_{m,j}^*.$$

Note that this differs slightly from the expression in Ref. [126], where the QZ factorization resulted in upper quasi-triangular, as opposed to upper triangular, matrices. Rearranging yields

$$(\bar{\mathcal{R}}_{m,m}^* \bar{\mathcal{P}} + \bar{\mathcal{T}}_{m,m}^* \bar{\mathcal{S}}) \mathcal{Y}_m = \mathcal{F}_m - \sum_{j=m+1}^N (\bar{\mathcal{R}}_{m,j}^* \bar{\mathcal{P}} + \bar{\mathcal{T}}_{m,j}^* \bar{\mathcal{S}}) \mathcal{Y}_j,$$

which lends itself to a recursive computation [126].

E.2 Using Exact Line Searches

The idea behind exact line searches is that, instead of performing the usual Newton update of $\bar{g}_s^{(i+1)} = \bar{g}_s^{(i)} + \varepsilon$, where the superscript denotes the iteration number, we update according to $\bar{g}_s^{(i+1)} = \bar{g}_s^{(i)} + \vartheta \varepsilon$, where $0 < \vartheta \leq 2$ is a damping factor. The interval for ϑ is a well-known restriction, required in order to obtain convergence with a damped Newton method, and may be derived by considering an appropriate norm of the residual at a particular iteration. The choice of ϑ is determined by using additional information from the original quadratic matrix equation.

We have that

$$r(\tilde{g}_s + \vartheta \varepsilon) = r(\tilde{g}_s) + \vartheta \tau^\dagger \varepsilon \tau \tilde{g}_s + \vartheta [\tau^\dagger \tilde{g}_s \tau - (EI - \Upsilon_L)] \varepsilon + \vartheta^2 \tau^\dagger \varepsilon \tau \varepsilon.$$

We define $p(\vartheta) \equiv \|r(\tilde{g}_s + \vartheta \varepsilon)\|_F^2$, where $\|A\|_F = \sqrt{\text{trace}(A^\dagger A)}$ is the Frobenius norm, and choose ϑ such that $p(\vartheta)$, a fourth-order polynomial, is minimized at each step with $0 < \vartheta \leq 2$. It may be shown [125] that $p(\vartheta)$ either has an inflection point at $\vartheta = 2$ or a minima in $(0, 2]$, so such a ϑ may always be found.