Quantum Transport in Carbon-based Nanostructures

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)
der Naturwissenschaftlichen Fakultät II – Physik
der Universität Regensburg

vorgelegt von
Norbert Nemec
aus Regensburg

Juli 2007
Die Arbeit wurde angeleitet von:
  Prof. Gianaurelio Cuniberti

Promotionsgesuch eingereicht am:
  2. Juli 2007

Gutachter:
  Prof. Gianaurelio Cuniberti
  Prof. Milena Grifoni

Bewertung der Dissertation:
  summa cum laude

Tag der mündlichen Prüfung:
  27. Juli 2007

Prüfungsausschuss:
  Prof. Christoph Strunk
  Prof. Gianaurelio Cuniberti
  Prof. Milena Grifoni
  Prof. Andreas Schäfer

Bewertung des Kolloquiums:
  summa cum laude

Gesamtnote:
  summa cum laude
## Contents

4. **Electrical contacts to nanotubes and -ribbons**  
4.1. Conventional contact models ............................................. 74  
4.1.1. Carbon nanotube electrodes ........................................... 75  
4.1.2. Structureless electrodes ............................................... 76  
4.1.3. Atomically modeled electrodes ....................................... 76  
4.2. Extended contacts ............................................................ 76  
4.2.1. Analytic model ......................................................... 79  
4.2.2. Generalization to arbitrary injection energies ...................... 84  
4.2.3. Non-diagonal contacts .................................................. 86  
4.2.4. Realistic contacts to carbon nanotubes and graphene nanoribbons . 86  
4.2.5. Three-terminal setup and Fabry-Pérot physics ....................... 89  
4.2.6. Non-epitaxial contacts ................................................ 91  
4.2.7. Material related calculations ........................................ 91  
4.3. Ferromagnetic contacts and spin transport ............................. 96  
4.3.1. Modeling ferromagnetic leads ....................................... 97  
4.3.2. Magnetoresistance of an ordered nanotube ......................... 99  
4.3.3. Effects of disorder .................................................... 99  

5. **Disorder and defects** ..................................................... 103  
5.1. Anderson model for disorder ............................................. 105  
5.2. The elastic mean free path .............................................. 106  
5.3. Strong localization ....................................................... 109  
5.4. Vacancies and defects ................................................... 111  
5.5. Graphene nanoribbons ................................................... 114  

6. **Multilayer graphene and carbon nanotubes** ......................... 117  
6.1. Commensurability ....................................................... 118  
6.2. Modeling the interlayer coupling ..................................... 120  
6.3. Bilayer graphene ......................................................... 122  
6.4. Double-wall carbon nanotubes ....................................... 124  
6.4.1. Commensurate double-wall tubes ................................... 125  
6.4.2. Incommensurate double-wall tubes .................................. 127  
6.5. Telescopic carbon nanotubes .......................................... 129  
6.5.1. Armchair telescopic tubes .......................................... 129  
6.5.2. Zigzag telescopic tubes ............................................ 132  
6.5.3. Model of telescopic quantum wires .................................. 132  

7. **Magnetoelectronic structure and fractality** ......................... 137  
7.1. Peierls substitution .................................................... 139  
7.2. Hofstadter butterfly ................................................... 140  
7.3. Butterfly and anomalous Landau levels of graphene .................. 142  
7.4. Butterfly of single-wall nanotubes ................................... 144  
7.5. Graphene nanoribbons .................................................. 147  
7.6. Periodic gauge ............................................................ 147  
7.7. Bilayer graphene ........................................................ 149  
7.8. Butterfly of double-wall nanotubes ................................... 151  

Conclusions and perspectives .................................................. 155  
Appendices .............................................................................. 159  

6
Introduction

Carbon is the most versatile of all elements. It is the structural backbone of all organic molecules and thereby of life itself. Pure carbon is found in nature both as diamond, the hardest material known to mankind, crystal clear and insulating, or graphite, an excellent lubricant, steel black and conducting, both differing only by the arrangement of the carbon atoms in the crystal structure.

Besides these two natural forms, a whole new class of artificial pure carbon materials has been discovered over the past twenty years: small spherical molecules called fullerenes (1985), tubular fibers called carbon nanotubes (1991), monoatomic sheets called graphene (2005), and many more structures of various geometric shapes at a scale of nanometers. The physical properties of these materials are promising: fibers stronger and lighter than steel, Kevlar™ or spider silk, electrical conductors better than silver or copper, semiconductors with a tunable band gap. These and many more visions have been realized and demonstrated in the laboratory and are about to enter the market.

This development would have been impossible without the progress that was made in parallel in directly observing and manipulating nano-scale structures with atomic precision. Complex electronic devices can be built using single carbon nanotubes and individual molecules, measuring currents electron by electron.

The technological interest in miniaturization has continued unbroken for decades. Electronic circuits are becoming smaller, faster, cheaper and more energy efficient at constant rate. Currently, the semiconductor industry is preparing to start mass production in 45 nm technology. The roadmap for submicroelectronic device fabrication is already laid out until 2018, when the individual devices on a computer chip are expected to reach a size of 16 nm. This would correspond to a width of just about 60 silicon atoms.

It is unclear how long silicon will actually be the adequate material to keep up with this trend. The precision in production, the long-term stability, the need for chemical doping and finally the laws of quantum
mechanics itself are all factors that are likely to present a hard wall to the ongoing trend of miniaturization.

Carbon nanotubes can be produced reliably at a thickness of less than 1 nm, are thermally stable up to 2800°C, can be either metallic or semiconducting depending on their structure alone, without need for chemical modification and their quantum electronic properties are well defined.

However, the technological importance of carbon nanostructures goes beyond that of a new material for building traditional devices. Intrinsically new concepts have been proved: Electromechanical nanomachines and chemical sensors have been built from carbon nanotubes, and the extremely long spin coherence make nanotubes a prominent candidate for spintronic applications.

Graphene is a very recent addition to the family of available carbon nanomaterials, carrying perhaps even more technological potential than carbon nanotubes. As a flat sheet, graphene holds the promise of allowing the lithographic production of complex structures. Its spin coherence is equally excellent as in nanotubes. Its unique electronic structure is so fundamentally different from anything that was known before, that many of implications may not even have been considered yet.

This thesis presents our results of the investigation of several core issues of quantum transport properties and electronic structure of carbon nanostructures, namely, carbon nanotubes and graphene, based on the use of numerical and analytical tools. The understanding of electrical contacts is of immediate importance to any experimental effort in optimizing charge injection. Defects and disorder are crucial ingredients in any attempt to capture physical reality in a theoretical model. The interlayer coupling in multilayer structures has been a field of hot debate for several years among experimentalists as well as theoreticians and the magnetoelectronic structure, closely linked to the anomalous quantum Hall effect observed in graphene, is one of the best examples demonstrating the extraordinary electronic properties of these novel materials.

The outline of this thesis is organized as follows:

**Chap. 1: Carbon materials** A general introduction of basic carbon nanostructures is given, including a brief sketch of the history of their discoveries and the most important methods of production.

**Chap. 2: Electronic structure** The models of electronic structure are introduced and the most general electronic properties of \( sp^2 \)-hybridized carbon structures derived from them.
Chap. 3: Quantum transport  The theory of coherent quantum transport in mesoscopic systems is introduced. An overview of the various mesoscopic length scales and their competition in different regimes of transport is given. The Landauer theory of coherent quantum transport is derived and demonstrated on a number of characteristic model cases that display many of the core features also present in carbon nanostructures.

Chap. 4: Contacts  A model for realistic contacts to a carbon nanotube in a typical experimental transport measurement setup is studied in detail. A very general dependency is found between contact length, coupling strength and contact transparency. The result is demonstrated to be qualitatively robust against various kinds of disorder and independent of details of the modeling. The quantitative result is combined with the results from ab initio calculations to derive clear predictions for experiment.

Chap. 5: Disorder and defects  The regimes of diffusive transport and strong localization are studied for carbon nanotubes and graphene nanoribbons with homogeneous model disorder or randomly distributed point defects. The elastic mean free path and the localization length are derived perturbatively and compared with the numerical results.

Chap. 6: Multilayer structures  Commensurate and incommensurate double-wall nanotubes and graphene bilayers are investigated. The bipartite systems are demonstrated to be determined by approximate symmetries if the interlayer coupling is chosen smooth enough. Telescopic nanotubes are studied for their resonant transport properties. A minimal model is solved and used to explain the resonances.

Chap. 7: Magnetoelectronics  The electronic structure of carbon nanostructures in external magnetic fields is computed nonperturbatively and visualized in Hofstadter butterfly plots. A periodic gauge is developed that allows the handling of graphene bilayers at arbitrary relative positions. The general structure of the butterfly plots is discussed and the anomalous Landau level structure of graphene is analyzed for graphene bilayers and double-wall nanotubes.
Introduction

Part of the work presented in this thesis has already been published in the following international journal articles (see reprints in App. D):


Part of the work presented in Sec. 6.5 of this thesis has been included in the diploma thesis by D. Darau [5].
Chapter 1.
Carbon-based nanostructures

Carbon occupies a special position in the periodic table of chemical elements. Due to the ability of each carbon atom to form up to four strong covalent bonds to neighboring atoms, carbon can build up a huge variety of complex networks forming the structural backbone of the infinite variety of organic molecules which are the material building blocks of life itself. All of these organic molecules have in common that additional chemical ingredients (most commonly hydrogen and oxygen) are necessary to terminate the carbon network and to stabilize the molecule.

For a long time, it was believed that the list of pure carbon structures was restricted to the two crystalline modifications diamond and graphite. Only with the discovery of fullerenes in 1985 [149], it became clear that pure carbon itself also has the potential to form alternative stable structures and with the identification of carbon nanotubes in 1991 [123], the list of observed and proposed carbon nanostructures began to grow at high pace, soon including single- and double- and multiwall tubes, multiwall fullerenes, horns, foam and many others. One of the most recent additions to this list was given by the successful isolation of single sheets of graphene [194], completing the set of zero-, one-, two and three dimensional crystal structures of pure carbon.

Without doubt, the general rise of nanotechnology was mainly initiated by the development of imaging and manipulation techniques that allow us today to clearly observe and even directly act on structures at atomic scales. With optical microscopy, the wave length of light strictly limits the possible resolution to more than a hundred nanometers. Scanning or tunneling electron microscopy (SEM/TEM) in principle allows to break this barrier, but with charged particles, the possibilities for observation are rather limited. Only with the development of the scanning tunneling microscope (STM) by G. Binnew and H. Rohrer in 1981 [35] it became possible to non-destructively observe structures down to the level of single atoms. With the atomic force microscope (AFM) introduced 1986 by
Chapter 1. Carbon-based nanostructures

Figure 1.1.: The shape of the relevant orbitals of Carbon in their real, unhybridized form. Displayed are the isosurfaces at positive and negative sign. (Images taken from Ref. [69])

the same team [34] the direct manipulation of such structures also became possible. With these tools at hand, processes at a scale that had before been accessible only through indirect methods, could be controlled and studied with growing precision. This allowed the individual identification and exploration of a great variety of pure carbon structures, that would have been impossible to discern using traditional methods.

In this chapter, we will review the various structures formed from pure $sp^2$-hybridized carbon and the basic results about their electronic structure that will be used throughout this thesis.

1.1. Hybridization of carbon orbitals

Carbon (chemical symbol C) has the atomic number 6. Two of its electrons are in the 1s orbital as core electrons. The remaining four electrons reside in the 2s and 2p orbitals and are available to form chemical bonds. As the two 2s orbitals are energetically slightly lower than the 2p orbitals, the former are filled in the ground state by two electrons, while the other two electrons reside in 2p orbitals. The energy difference, however, is small enough, that both sets of orbitals can easily form hybrid bonds in various ways:

The electronic wave function of an atom is typically derived as eigenstates of the angular momentum operator in one selected orientation, say, the z direction. In spherical coordinates, these can be expressed as:

$$
\langle r \vartheta \phi | 1s \rangle = f_{1s}(r)
$$

$$
\langle r \vartheta \phi | 2s \rangle = f_{2s}(r)
$$

$$
\langle r \vartheta \phi | 2p^0 \rangle = f_{2p}(r) \sin \vartheta
$$

$$
\langle r \vartheta \phi | 2p^\pm \rangle = f_{2p}(r) \cos \vartheta e^{\pm i\phi}
$$
1.1. Hybridization of carbon orbitals

![Diagram showing sp, sp², and sp³ hybridization]

**Figure 1.2.** The three hybridized forms of carbon in chemical compounds: $sp$ hybridization occurs only in some typically rather reactive organic compounds, the alkynes. $sp^2$ hybridization leads to planar structures like graphite layers, $sp^3$-hybridized carbon atoms form sterical structures like the diamond crystal. (Images taken from Ref. [69])

where the radial functions $f$ are known analytically for the hydrogen atom, involving the Laguerre polynomials. For atoms with several electrons, they are qualitatively the same, but the actual shape is known only approximately.

As the $2p$ orbitals are energetically degenerate, they can be linearly combined in a different way, yielding purely real wave functions:

$$
|2p_x\rangle = \frac{1}{\sqrt{2}} (|2p^+\rangle + |2p^-\rangle) \\
|2p_y\rangle = \frac{i}{\sqrt{2}} (|2p^+\rangle - |2p^-\rangle) \\
|2p_z\rangle = |2p^0\rangle
$$

These three orbitals have the identical geometrical shape in three different orientations. See Fig. 1.1 for their commonly known visualization.
When forming chemical bonds, carbon is not limited to these atomic energy eigenstates but tends to merge the energetically close $2s$ and $2p$ orbitals, forming linear combinations in various ways:

- Hybridizing the $2s$ with one $2p$ orbital gives a set of two $sp$ orbitals in diametrically opposed directions:

$$|2sp_a\rangle = \frac{1}{\sqrt{2}} (|2s\rangle + |2p_y\rangle)$$

$$|2sp_b\rangle = \frac{1}{\sqrt{2}} (|2s\rangle - |2p_y\rangle)$$

Though relevant in organic molecules, this hybridization is not prevalent in pure carbon structures.

- Far more relevant is the hybridization of the $2s$ orbital with two different $2p$ orbitals resulting in three equivalent $sp^2$ orbitals arranged in plane at an angle of $120^\circ$:

$$|2sp^2_a\rangle = \frac{1}{\sqrt{6}} \left( \sqrt{2} |2s\rangle + 2 |2p_x\rangle \right)$$

$$|2sp^2_b\rangle = \frac{1}{\sqrt{6}} \left( \sqrt{2} |2s\rangle - |2p_x\rangle + \sqrt{3} |2p_y\rangle \right)$$

$$|2sp^2_c\rangle = \frac{1}{\sqrt{6}} \left( \sqrt{2} |2s\rangle - |2p_x\rangle - \sqrt{3} |2p_y\rangle \right)$$

These orbitals are capable of forming strong covalent bonds with other carbon atoms resulting in planar structures. The remaining, unhybridized $p_z$ orbital is perpendicular to the plane and may join with parallel $p_z$ orbitals of neighboring atoms, forming a strongly delocalized $\pi$ orbital that is responsible for the electronic properties of the resulting structures. By convention, this remaining atomic orbital in an $sp^2$-hybridized carbon structure is itself referred to as $\pi$ orbital.

- Finally, a hybridization of the $2s$ with all three $2p$ orbitals is also possible, yielding four equivalent $sp^3$ orbitals, oriented in a tetrahedral
1.1. Hybridization of carbon orbitals

arrangement at an angle of \( \sim 109^\circ \) (tetrahedral angle):

\[
\begin{align*}
|2sp_a^3\rangle &= \frac{1}{\sqrt{12}} \left( \sqrt{3} |2s\rangle + \sqrt{3} |2p_z\rangle \right) \\
|2sp_b^3\rangle &= \frac{1}{\sqrt{12}} \left( \sqrt{3} |2s\rangle - \sqrt{1} |2p_z\rangle + \sqrt{8} |2p_x\rangle \right) \\
|2sp_c^3\rangle &= \frac{1}{\sqrt{12}} \left( \sqrt{3} |2s\rangle - \sqrt{1} |2p_z\rangle - \sqrt{2} |2p_x\rangle + \sqrt{6} |2p_y\rangle \right) \\
|2sp_d^3\rangle &= \frac{1}{\sqrt{12}} \left( \sqrt{3} |2s\rangle - \sqrt{1} |2p_z\rangle - \sqrt{2} |2p_x\rangle - \sqrt{6} |2p_y\rangle \right)
\end{align*}
\]

These orbitals form strong covalent bonds with neighboring atoms but are electronically inactive due to their low energy.

Pure-carbon materials can be classified in two major groups given by the hybridization: diamond is a tetrahedral structure with each atom bound via four equivalent \( sp^3 \) orbitals forming \( \sigma \)-bonds with its neighbors. Generally, this kind of bond leads to an extremely stiff geometry and a large gap between valence and conduction electrons. Thus, diamond is hard, transparent and insulating.

Graphite, on the other hand, has a layer structure. Each layer is a monoatomic sheet of carbon atoms arranged in a honeycomb structure. Each atom has three in-plane neighbors to which it forms equivalent \( \sigma \)-bonds via \( sp^2 \)-hybridized orbitals. The remaining \( \pi \) orbital is free to combine with those of all other atoms into a completely delocalized \( \pi \) band similar to the delocalized orbital known from aromatic molecules. The layers are held together by weak van der Waals forces allowing them to slide against each other with minimal friction.

A single layer graphitic structure is called graphene. Beginning at its two-dimensional structure it is possible to derive the geometry of a whole family of materials. (See Fig. 1.3).

The in-plane \( \sigma \)-bonds of \( sp^2 \)-hybridized carbon are stiff against longitudinal forces but soft for angular deformations. In a single layer of graphene this is the cause for the observed rippling [179]. More generally, this flexibility opens the door for the large range of stable graphitic nanostructures. The electronic properties of \( sp^2 \)-hybridized carbon materials are generally determined by the delocalized \( \pi \) orbital. Unlike the strongly bonding \( \sigma \) orbitals, the \( \pi \) orbitals are generally near to the Fermi energy, leading to the conducting or semiconducting properties of graphitic materials and opening the path to a wealth of potential applications in future nanoelectronics.
Chapter 1. Carbon-based nanostructures

1.2. Graphite

Graphite is the chemically most stable allotrope of pure carbon. It is found in nature as a polycrystal with fairly small grains up to a few micrometers. It is mechanically very soft and electrically conducting. Single crystals have a highly anisotropic conductivity due to its layer structure: the in-plane conductivity is much larger than the conductivity perpendicular to the layers. Individual layers slide easily against each other, making graphene an excellent, technologically important lubricant. Though it is, in principle, the purest form of coal, it is hard to ignite and is therefore not usually used as an energy source.

Figure 1.3.: Theoretical construction of various carbon structures. Starting from planar graphene in two dimensions, it can be rolled up to quasi-zero-dimensional fullerenes or quasi-one-dimensional nanotubes. Stacks of graphene sheets are graphite. (Image taken from Ref. [93])
1.2. Graphite

Figure 1.4.: Natural graphite (image originates from “Minerals in Your World” project, a cooperative effort between the United States Geological Survey and the Mineral Information Institute)

Geometry

The first description of the hexagonal, layered crystal structure of graphene was given by A. W. Hull in 1917 [121], based on x-ray diffraction. Few years later, J. D. Bernal could identify the individual planar layers [30], completing the picture that we have today. Graphite consists of parallel graphene sheets spaced by \( d_{\text{interlayer}} = 3.34 \, \text{Å} \), stacked in an alternative series (ABAB, commonly called Bernal stacking, see Fig. 1.5).

Figure 1.5.: Structure of natural graphite (Bernal stacking): Two layers repeat in alternating positions A and B. One half of the atoms are aligned on top of each other, the other half is aligned with the plaquette centers of the neighboring layers.

A modification of this structure is rhombohedral graphite with the stacking order ABCABC. It has been shown that natural graphite often contains a certain amount in rhombohedral stacking, so, even though this form has never been isolated, it still has inspired a number of theoretical works [106,
Figure 1.6.: STM (left) and AFM (right) image of a graphite surface. While the AFM images each atom equally, the STM is sensitive to whether an atom is on top of another atom or a hole in the second layer. This way, the Bernal stacking of the top two layers is indirectly confirmed. (Image taken from Ref. [115])

47, 57], especially as an intermediate state in the transition from graphite to diamond [139, 84].

A further modification is turbostratic graphite where the individual layers are rotated by random angles against each other. Such structures arise from mechanical deformation of graphite. In this case, the crystalline nature of graphite is lost and quasi-crystalline properties are to be expected.

1.3. Graphene

Single layer graphene—the latest addition to the family of experimentally realized carbon nanomaterials [194]—is actually the most basic building block for the theoretical understanding of all other $sp^2$-hybridized carbon structures: graphite, as a stack of graphene sheets, or nanotubes and fullerenes as rolled-up graphene. Two-dimensional crystals were long believed not to exist in nature, based on theoretical reasoning saying that quantum mechanical length fluctuation of individual bonds would add up logarithmically over distance, destroying the long-range order that defines a crystal [211, 178]. In fact, however, freely suspended graphene layers do exist [179, 180], and it is now widely believed that the observed fluctuations in the third dimensions help in stabilizing the structure. An alternative conjecture would, of course, be that the theoretical veto simply may not apply for finite size patches.
1.3. Graphene

Figure 1.7.: SEM image of a graphene flake deposited on a SiO$_2$ substrate. The three shades of gray show sections of mono-, bi- and trilayer graphene. (Courtesy of U. Stöberl, Uni Regensburg)

1.3.1. Isolation by exfoliation

Considering the general desire to isolate single layers of graphene, it is surprising how long it took before the first success and even more surprising, considering the simple approach that finally led to success: using scotch tape to repeatedly peel off flakes from pyrolytic graphite that become thinner with each step until finally, there is a single layer left that can be placed on a clean surface for further handling [193]. Soon, the method of exfoliation from bulk graphite was further simplified to what could be described as searching a “pencil trace” for monolayer flakes [194]. The main difficulty with this search is that such thin structures are generally invisible by optical means, so—lacking any known electronic signature that would simplify the search—samples have to be screened tediously via AFM.

Figure 1.8.: AFM image of graphene flake. (Image taken from Ref. [194])

The alternative way of using wet chemistry to exfoliate graphene also has shown first promising results [246].
1.3.2. Synthesis by epitaxial growth

Graphite has been grown epitaxially using pyrolysis of methane on Ni crystals already 40 years ago [135]. This technique has been refined to produce monoatomic layers, i.e. graphene sheets on various crystal surfaces [200]. Even ribbons of well defined width (1.3 nm) have been produced and analyzed [252].

![An epitaxially grown graphene monolayer on a TaC(111) surface. (Image taken from Ref. [200])](image)

Alternatively, epitaxial growth of graphene can also been achieved by segregation of C atoms from inside a substrate to its surface, e.g. from doped Ni, Pt, Pd or Co [125, 108] or SiC [28, 29]. With these epitaxial growth methods, very high crystal qualities can be obtained, sometimes at a modified lattice constant, and other times also incommensurate to the surface with the lattice defined by the strong C-C bonds. Successful attempts to lift off epitaxially grown graphene from the surface are not known to us, but neither is any fundamental obstacle to prevent this in the future.

1.3.3. Geometry

Graphene can be understood as a two-dimensional crystal. Its honeycomb structure, displayed in Fig. 1.10, is spanned by a trigonal periodic lattice and a two-atomic basis. The distance between neighboring atoms is \( d_{CC} = 1.42 \, \text{Å} \), the lattice constant accordingly \( a = \sqrt{3}d_{CC} \approx 2.46 \, \text{Å} \). The reciprocal lattice is again trigonal, resulting in a hexagonal shape of the first Brillouin zone which appears rotated by 90° against the hexagons of the real-space structure.

As it was said, there is a long standing argument by R. E. PEIERLS, that in one or two dimensions crystals should not have any long range order, because the quantum mechanical fluctuations of each bond-length would add up with the square-root of the distance in one dimension or
1.3. Graphene

Figure 1.10.: The honeycomb structure of graphene along with its first Brillouin zone in reciprocal space. With the original lattice vectors $a_1$ and $a_2$ at $60^\circ$ against each other and each of length $a = \sqrt{3}d_{CC} \approx 2.46 \text{Å}$, the reciprocal lattice vectors, defined by $a_i \cdot \tilde{a}_j = 2\pi \delta_{ij}$ are at an angle of $120^\circ$ and of length $|a_i| = 4\pi/(\sqrt{3}d_{CC}) \approx 5.11 \text{Å}^{-1}$.

Figure 1.11.: Suspended graphene sheets. Left: TEM-image of few-layer graphene membrane near its edge. Below: Illustration of the rippled structure of suspended graphene. The out-of-plane fluctuations are presumed to be necessary for the stabilization of the 2D crystal structure. (Images taken from Ref. [179])

A remarkable detail about graphene crystals is that they preferentially break at crystal edges of two kinds: a zigzag edge runs parallel to a graphene lattice vector, an armchair edge runs parallel to the carbon bonds (see Fig. 1.12). While zigzag edges carry an electronic edge-state [186, 91], photonic edge-states are present at armchair edges [122]. These become especially relevant for finite width ribbons (see Sec. 1.4).
Chapter 1. Carbon-based nanostructures

Figure 1.12.: STM image of a large graphene monolayer obtained via exfoliation. The crystal edges are preferentially oriented at zigzag (red) or armchair (blue) edges. (Image taken from Ref. [93], originating from an unpublished work by T. J. Booth, K. S. Novoselov, P. Blake and A. K. Geim)

1.4. Graphene nanoribbons

Graphene nanoribbons (GNRs) were first considered theoretically as a convenient model to study the electronic edge state in graphene [186, 91] and their phononic counterpart [122] without much emphasis on how such structures could realistically be produced. In 2002, however, before the boom around graphene had even been initiated by successful exfoliation, T. Tanaka et al. indeed managed to grow well-defined, narrow GNRs on a TiC surface and measured the phononic edge modes [252] (see Fig. 1.13).

Figure 1.13.: STM image and schematic of GNRs grown on a TiC (755) surface in region (A). Regions (B) and (C) are (111) graphene-covered terraces that show a Moiré pattern due to a lattice mismatch. (Images taken from Ref. [252])

When the graphene boom began, interest in GNRs soon began to rise as well and ribbons were considered by theorists as a serious alternative to carbon nanotubes (CNTs) as quantum wires and devices [80, 244, 41, 203, 185, 234] and very recently, first experimental results have also been reported by Z. Chen et al. [51] on successful conductance measurements in GNRs of various widths (see Fig. 1.14).
1.4. Graphene nanoribbons

The structure of GNRs can be obtained directly from unrolling single-wall CNTs of various chirality. As the stability of edges has to be taken into account, only achiral GNRs are typically considered where the naming is based upon the edges. This leads to the somewhat awkward convention that unrolling an armchair CNT results in a zigzag GNR and a zigzag CNT unrolls into an armchair GNR.

The classification of individual GNRs follows a slightly different convention than that of CNTs. A \((N,N)\) armchair CNT unrolls into a zigzag GNR with a width of \(N_z = 2N\) zigzag carbon strands. A \((N,0)\) zigzag CNT unrolls into an armchair GNR with \(N_a = 2N\) lines of carbon dimers. Odd values of \(N_a\) and \(N_z\) refer to asymmetric GNR that cannot be constructed by unrolling CNTs (see Fig. 1.15).

**Figure 1.14.:** SEM image and room temperature resistivity measurement data of patterned GNRs from Ref. [51].

**Figure 1.15.:** Structure of the two classes of GNRs: Armchair GNRs are characterized by the number \(N_a\) of parallel lines of carbon dimers, zigzag GNRs by the number \(N_z\) or parallel zigzag strands. (Figure taken from Ref. [244])
Chapter 1. Carbon-based nanostructures

1.5. Carbon nanotubes

The history of the discovery of carbon nanotubes (CNTs) has been an issue of hot discussion [38, 182]. A Soviet team, L. V. Radushkevich and V. M. Lukyanovich, can claim to have published and correctly identified the first images of multiwall CNTs already in 1952 in the Russian Journal of Physical Chemistry [219], unnoticed by the international community (see Fig. 1.16). Similarly, the rediscovery by A. Obrin, M. Endo and T. Koyama, published in 1976 in the highly specialized Journal of Crystal Growth [196] (see Fig. 1.17), was not recognized for its significance until the real boom of carbon nanotube research had been initiated by the famous work of S. Iijima in 1991 [123] (see Fig. 1.18). The first observation of a single-wall CNT was reported soon afterwards in 1993 by two groups independently: S. Iijima and T. Ichihashi [124] as well as D. S. Bethune et al. [31].

Figure 1.16.: Historically the first published and correctly identified images of multiwall CNTs by L. V. Radushkevich and V. M. Lukyanovich in 1952 [219].

The structure of CNTs, which can be described as a cylindrically rolled-up graphene ribbon, is cause for an extreme mechanical strength in the longitudinal direction possibly even exceeding that of diamond. This theoretical property alone has inspired a multitude of potential applications ranging from ultra-strong textiles over compound materials all the way to the famous idea of the space-elevator which would demand a strength-to-weight ratio that has is reached exclusively by the theoretical predictions for CNTs.

It is generally believed that CNTs exist only as a synthetic material, though there are also indications that natural carbon soot contains certain amounts of these structures mixed in with all other forms of amorphous carbon. Recently, it has been found that nanotube synthesis may actually have been accessible in medieval times already, even though the produc-
1.5. Carbon nanotubes

**Figure 1.17.** Images of CVD-grown CNTs by A. Oberlin, M. Endo and T. Koyama, published in 1976 in the *Journal of Crystal Growth* [196].

**Figure 1.18.** High resolution electron micrograph images of CNTs grown by arc-discharge. Published 1991 by S. Iijima in the famous *Nature* article that initiated the boom of CNT research [123].
ers of the legendary Damascus sabers [222] certainly had no idea about the nanoscale physics of their production techniques. Recent reviews about CNTs in general and about their electronic and transport properties can be found in Ref. [16] and [46].

1.5.1. Synthesis methods

An excellent overview over the state-of-the-art synthesis methods can be found in Ref. [62]. Out of the countless ways of producing nanotubes, the three major methods will be outlined in the following.

Arc discharge

In 1991, S. Iijima tried to produce fullerenes using the conventional method of driving a 100 A current through graphite electrodes in an arc discharge and discovered carbon nanotubes in the soot [123]. Shortly afterwards, the efficiency of the method was improved to yield macroscopic quantities of nanotubes [75]. The method is fairly easy to set up but provides very limited control over the production parameters. The nanotubes are generally very short, have a wide distribution of diameters and are submerged with other forms of amorphous carbon. Arc discharge nanotubes typically have few defects and contain no catalyst residue. The production can be done in open air.
1.5. Carbon nanotubes

Figure 1.20.: Schematic view of a laser ablation apparatus, along with a TEM image of a resulting SWCNT bundle. (Image taken from Ref. [62])

Laser ablation

This method was pioneered by R. E. Smalley in 1995 [104]. Similar to the arc discharge method, pure graphite is thermally evaporized, though, instead of an electrical current, high-powered laser pulses are used. By fine-tuning the parameters, yields of high purity nanotubes can be achieved. The diameter can be controlled rather well. The main drawback of this method is the need for expensive equipment and high power.

Chemical vapor deposition

The most commonly used low-cost method for the production of carbon nanotubes is the growth via chemical vapor deposition (CVD). Indeed, this is the method that was used by A. Öberlin and M. Endo for their first observation of carbon nanotubes in 1976 [196, 78]. Generally, this method is based on cracking atomic carbon from a chemical compound and depositing it on a catalytic surface where nanotube can then grow in very controlled ways. The type and quality of the grown nanotubes depends delicately on the parameters of the growth. It is possible to selectively grow a narrow diameter range of single-wall tubes [163] or double-wall tubes [54], control the direction of growth [45] or grow highly aligned arrays of tubes [67]. A common drawback of CVD methods is the contamination by catalyst particles and the relatively high defect rate. On the other hand, the method is easiest to scale up to industrial production rates.
1.5.2. Single-wall carbon nanotubes

A single-wall CNT (SWCNT) can be understood as a ribbon of graphene with both edges joined to form a tube. Each of the various ways of forming such a tube can be uniquely identified by its chiral vector \((M, N)\) which specifies that lattice vector of the flat graphene sheet which corresponds exactly to one circumference of the rolled-up tube (see Fig. 1.22). Two special cases are CNTs with a chiral vector \((N, N)\), called armchair CNT, and those with a chiral vector \((N, 0)\), called zigzag CNT. Both cases are called achiral CNT and have in common that they are mirror symmetric and have a comparably small unit cell containing only \(4N\) atoms. Other tubes, called chiral CNT, have a non-mirror-symmetric helical structure and the number of atoms in the unit cell, given by

\[
N_{\text{atoms}} = 4 \left( M^2 + N^2 + MN \right) / \gcd (M + 2N, 2M + N)
\]

is usually much larger than in achiral tubes of similar diameter. A full derivation of the structure of a CNT, along with an algorithm for the atomic coordinates is given in App. C.3.

The radius obtained from this idealized rolling-up of a graphene ribbon can be obtained directly from the length of the chiral vector that specifies its circumference as

\[
r = \sqrt{3d_{\text{CC}}} \sqrt{M^2 + N^2 + MN} / 2\pi.
\]

Small diameter CNTs are known to deviate from this idealized structure due to the strong curvature of the graphene wall squeezing the \(\pi\) orbitals at the inner side of the wall and leading to a repulsion of neighboring carbon atoms [143]. The radius therefore tends to be larger than the idealized.
1.5. Carbon nanotubes

Figure 1.22: A carbon nanotube can be thought of as “rolled-up” graphene. The two kinds of achiral CNT shown here are rolled up along the two different graphene crystal edges. (Images taken from Ref. [69])

value, depending on both diameter and chirality. This leads to an inequivalence of the bond lengths, opening a gap in originally metallic non-armchair CNTs (see Sec. 2.3).

Though SWCNTs have been predicted to be stable down to a diameter of 0.4 nm [235] and even 0.3 nm [286], the smallest freestanding SWCNTs typically observed in experiment have a diameter of 0.7 nm, corresponding to a C$_{60}$ molecule [9]. However, smaller tubes down to 0.4 nm have been observed either as innermost shell of MWCNTs [218, 250] or embedded in porous crystals [272]. Even short segments of 0.33 nm wide SWCNTs have been observed and could be shown to form stable—though energetically unfavorable—structures [213]. This latest case corresponds to a chirality (4,0).

The largest observed SWCNT have a diameter of up to 7 nm [140, 160]. At this size, however, SWCNTs generally become deformed due to surface adhesion and have the tendency to collapse due to the van der Waals force between opposite walls [77, 253].
1.5.3. Multiwall carbon nanotubes

Multiple SWCNTs can enclose each other to form a double-wall CNTs (DWCNT) or multiwall CNTs (MWCNT). Experimentally, MWCNTs were discovered even before their single-wall counterparts [123, 9]. Due to their much larger diameter and the consequentially easier handling, far more experimental results are available on MWCNTs than there are on SWCNTs. From the theoretical side, much effort has been done in investigating the effects of the combination of several walls. Yet, due to their greater complexity, the theoretical understanding is less clear than that of the well-understood SWCNTs. While the spectrum of SWCNTs can be computed to high precision using comparatively simple models, for DWCNTs even fundamental qualitative properties like their metallicity are still subject to active research [290].

Typically, the interwall distance of such structures is similar to the interlayer distance of graphite \( d_{\text{interwall}} \approx 3.34 \, \text{Å} \). For a pair of armchair CNTs, this distance can be achieved by a combination as \((N, N) @ (N + 5, N + 5)\), leading to \( d_{\text{interwall}} = 3.39 \, \text{Å} \). The best match for a pair of zigzag CNTs is \((N, 0) @ (N + 9, 0)\) giving \( d_{\text{interwall}} = 3.52 \, \text{Å} \). Other combinations can be found by doing a algorithmic search. Experimentally, a wide distribution between 3.4 Å and 3.8 Å has been observed with no specific correlation of the chiralities [117].

In the case of pure armchair or pure zigzag MWCNT, the length of the unit cell is the same as for each individual shell. For most other combinations individual shells have unit cells of different length. In some cases, these length are commensurate, allowing the definition of a supercell of common periodicity. Most generally, however, the lengths are incommensurate, preventing the definition of any common supercell. For a detailed analysis of this issue see Sec. 6.1.

1.6. Fullerenes

Fullerenes were discovered in 1985 by the team of H. W. Kroto, J. R. Heath, S. C. O’Brien, R. F. Curl, and R. E. Smalley [149] and named after the geodesic domes by architect R. Buckminster Fuller. In general, Fullerenes consist of a varying number of carbon atoms, forming 12 pentagons and a varying number of hexagons in a sphere. Most prominent is the highly symmetric \( C_{60} \) molecule, nick-named bucky ball, forming a truncated icosahedron, more commonly known as the structure of a soccer ball (see Fig. 1.23).
1.6. Fullerenes

Figure 1.23: Structure of a $C_{60}$ fullerene, nicknamed *bucky ball*. The 12 pentagons and 20 hexagons form a structure with the highest nontrivial symmetry among all molecules.

Fullerenes were later found to occur naturally e.g. in regular candle soot. The arc-discharge method allows the easy production of grams of fullerenes, but the purification remains a challenge. Fullerenes form crystals called *fullerites* that occur naturally within shungite.
Chapter 1. Carbon-based nanostructures
Chapter 2.

Electronic structure

The theoretical modeling of carbon materials can generally be done at various levels of detail and precision that we can roughly group into three categories:

- *ab initio* methods allow the quantitative computation of material properties without experimental input, i.e. without free parameters that need to be adjusted (apart from atomic masses and fundamental physical constants). The method that is most commonly used within solid state physics is density functional theory (DFT) based on the theory by W. Kohn and L. J. Sham [144].

The general advantage of *ab initio* methods is, that they can produce precise quantitative results. The main disadvantages are the high computational cost and the difficulty in gaining a deeper understanding from the results.

- *atomistic models* generally take into account the full atomic structure of the system with each atom at a position either in a predefined geometry (like those given in the preceding sections), or computed within the model itself. For the description of mechanical properties, the most common semi-empirical descriptions are force-constant models, describing the mechanical forces between neighboring atoms in a harmonic approximation [170, 289]. For the description of the electronic structure, a large family of models is based on the concept of the linear combination of atomic orbitals (LCAO), also known as tight-binding method (TB-method). Originally introduced by Bloch for the description of simple periodic structures [36], the method was refined by J. C. Slater and G. F. Koster [241] and is today widely used for the efficient, flexible and fairly accurate computations both model systems and real physical structures.

Compared to *ab initio* methods, atomistic models can generally be handled with far less computational effort and allow an easy tuning
of internal parameters, helping in understanding the individual physical effects. However, the quantitative significance of the results always depends on the parameters that are needed as an input.

- **Effective models** comprise a collection of various models that do not take into account the detailed atomic structure, but instead approximate the electronic structure. The most common representatives of this class are continuum models to describe mechanical properties and effective mass models for the description of crystal electrons. In the special case of graphene, we will find that lattice electrons can be described by an effective model resembling massless, relativistic particles.

Effective models often allow an analytic treatment of the physical problems, which may allow a much deeper understanding than purely numerical results.

In the following, we will mostly concentrate on the tight-binding approximation which offers a reasonable trade-off between precision and computational complexity and which is well-suited for the handling of carbon nanostructures.

### 2.1. The tight-binding approximation

The first step of the tight binding (TB) approximation is the choice of an atomic basis with the individual basis states defined by their wave functions in real space as:

$$\chi_{n\ell m}(r) = R_{n\ell m}(|r - r_i|) Y_{\ell m}(\vartheta, \varphi)$$

where $Y_{\ell m}(\vartheta, \varphi)$ are the spherical harmonics. The quantum numbers are the index $i$ of the individual atom, and the three atomic quantum numbers $n$, $\ell$ and $m$ (principal, angular and magnetic). The radial function $R_{n\ell m}(r)$ can be chosen in various ways and is an important factor for the precision of the approximation.

The crucial step, that turns this overcomplete basis into a valuable approximation is the reduction to a very limited set of orbitals per atom. The electronic properties are usually determined by a small number of orbitals near the Fermi energy, so projecting the Hamiltonian to these orbitals produces a reasonably small error.

The Hamiltonian of this system is given by its matrix elements in this atomic basis

$$\mathcal{H}_{n'i'\ell'm'} = \langle \chi_{n'\ell'm'} | \hat{\mathcal{H}} | \chi_{n\ell m} \rangle,$$
2.1. The tight-binding approximation

generally called the hopping matrix. Furthermore, this basis is in general non-orthogonal, leading to non diagonal entries in the overlap matrix

\[ S_{\ell m, \ell' m'} = \langle \chi_{\ell m} | \chi_{\ell' m'} \rangle. \]

In a non-orthogonal basis, the Schrödinger equation has to be expressed as a generalized eigenvalue equation:

\[ \hat{H} \Psi = E \hat{S} \Psi, \]

and accordingly the expression for the Green functions (see, e.g. App. A) get generalized as:

\[ \hat{G}(E) = \left( E \hat{S} + i0^+ - \hat{H} \right)^{-1}. \]

In the special case of the orthogonal TB-approximation the matrix \( S \) is set to identity and does not need to be considered further.

2.1.1. Obtaining a tight-binding Hamiltonian

In general, a TB Hamiltonian can be obtained in various ways:

- In principle, the full many-particle Hamiltonian could be directly expressed in an atomic basis. That will, however, lead to an interacting Hamiltonian that can only be solved using sophisticated techniques involving further approximations.

- The Kohn-Sham Hamiltonian of a density-functional-theory calculation done in atomic orbitals (p.e., using the SIESTA code [242]) can be viewed directly as a TB-Hamiltonian. In the special case of the DFTB-method, the DFT calculation is done for each pair of atoms independently and the entries of the resulting small Hamiltonians are then combined to a TB-Hamiltonian of the complete system [90]. In either case one should be aware that the Kohn-Sham theorem assigns a physical meaning only to the total energy obtained from the effective single-particle Hamiltonian and the energy of the highest occupied band [13]. Any other quantities—like the individual band energies—have no clear physical meaning and should be treated with care.

- The entries of the Hamiltonian can be viewed as parameters that are adjusted to either experiment or certain results from ab initio computations. To achieve transferability of the obtained parameterization,
the values can be assumed to follow a simple functional form depending on the geometry, so once the parameters of these functions are determined, they can be used for alternative geometries. Most common is the approach by J. C. Slater and G. F. Koster [241], assuming a dependence of the hopping integral on the distance between the two involved atoms only. One excellent collection of Slater-Koster (SK) parameters for a wide range of elements was build up by M. J. Mehl and D. A. Papaconstantopoulos [177].

2.1.2. Single orbital tight-binding approximation of graphitic structures

For graphitic structures, i.e. $sp^2$-hybridized carbon, a good approximation of the electronic properties at low energy can be found already for a single orbital per atom: In flat graphene, the in-plane $sp^2$ orbitals are symmetric and the $\pi$ orbitals are antisymmetric with respect to the mirror-symmetry at the graphene plane. This suppresses any matrix elements coupling both kinds of orbitals. In the band-structure of the periodic system, these two groups of orbitals form two independent sets of bands: the $\sigma$ bands, formed by the $sp^2$-hybridized orbitals, are responsible for the structure and the mechanical properties within single graphene layers and single shells of nanotubes. These lie far below and above the Fermi energy, so they have negligible influence on the electronic properties at energies relevant for electronic transport. The $\pi$ bands, formed by the atomic $\pi$ orbitals, are half filled, right around the Fermi energy, and are responsible for transport and other electronic properties at low energy. We can, therefore, restrict the tight-binding basis to a single spin-degenerate $\pi$ orbital per carbon atom. It is known that the curvature of small-diameter carbon nanotubes causes a certain amount of mixture between the bands [231, 142]. For larger diameter tubes, however, these effects are negligible.

The intralayer Hamiltonian

In a single graphene sheet, the high symmetry of the system allows the definition of a tight-binding Hamiltonian with a single adjustable parameter, the hopping $\gamma_0$ between the $\pi$ orbitals of neighboring atoms:

$$\mathcal{H} = \varepsilon_i \sum_i |i\rangle\langle i| - \gamma_0 \sum_{\langle ij \rangle} |i\rangle\langle j|$$
2.1. The tight-binding approximation

where $|i\rangle$ is the $\pi$ orbital of the carbon atom indexed by $i$ and $\langle i, j \rangle$ specifies the sum over nearest neighbors. The choice of $\gamma_0 = 2.66$ eV used throughout this work results in a Fermi velocity of graphene of $v_F \approx 8.7 \times 10^5$ m/s which is close to experimental findings [230, 202]. The on-site energy $\epsilon_i$ of each atom can be used to implement a non-constant electric potential. As long as it is constant it gives a plain energy offset which we will neglect for all $i$: $\epsilon_i = \epsilon_0 = 0$

Refinements like a non-zero overlap between neighboring orbitals or contributions of second and third nearest neighbors have been shown to give considerable corrections to the band structure at higher energies [223] but do not have much effect near the Fermi energy.

The interlayer Hamiltonian

The modeling of the coupling between layers of graphite and walls of multiwall CNTs is less obvious and the literature offers a long list of parameterizations that have been used for theoretical studies. Often, the variations seem to be insignificant but result in qualitatively different results.

For graphite, most theoretical studies use parameterizations that are restricted to Bernal-stacking as it exists in nature, requiring only a small number of fixed values for the hopping integral of nearest neighbors. One widely used parameter set for this specific case was given R. C. Tatar and S. Rabii [256], showing agreement with a wide range of experimental and ab initio data. For multiwall CNTs, the curvature of the walls generally requires a more flexible parameterization that allows arbitrary relative
orientations of the layers. The parameterization that we chose for our work is taken from a work by S. Roche et al. [224], which uses a slight modification of the original version by P. Lambin et al. [151]. The matrix element for the intershell hopping between atoms $i$ and $j$ is:

$$\gamma_{i,j}^{\text{intershell}} = \beta \cos \delta_{ij} \exp \left( \frac{r_0 - |r_i - r_j|}{\delta} \right) \quad (2.1)$$

with $r_0 = 3.34$ Å set to the interlayer distance of graphite. The parameter $\beta = \gamma_0/8$ controls the strength of the interlayer coupling relative to the nearest neighbor intralayer coupling. The parameter $\delta = 0.45$ Å controls the exponential drop of the coupling between atoms that are not exactly aligned. The angle $\delta_{ij}$ between the orientations of the two $\pi$ orbitals is, of course, irrelevant for graphene multilayers. A length cutoff for the interlayer hopping parameter is chosen as $r_{\text{cutoff}} = r_0 + 5\delta$, long enough to ensure a sufficiently smooth decay and smooth dependence on continuous displacements. A detailed discussion of this parameterization is given in Chap. 6.

### 2.1.3. Slater-Koster parameterization of carbon structures

The $\pi$ orbital approximation is simple and efficient, but it is generally restricted to systems of low curvature and a well-defined interlayer distance. Far better transferability can be achieved by Slater-Koster (SK) parameterizations that include a more complete set of orbitals without presuming a certain hybridization. For carbon, the relevant orbitals are $2s$ and $2p$, carrying four electrons in $(1_{2s} + 3_{2p}) \times 2_{\text{spin}} = 8$ orbitals. If spin does not need to be considered, four orbitals per atom remain.

D. A. Papaconstantopoulos et al. developed a parameterization transferable both to diamond and graphite structure [205]. As we will show in Chap. 6, this parameterization is in very good agreement with the $\pi$ orbital parameterizations described above.

### 2.2. Band structure of graphene

The peculiar electronic structure of graphene is a corner stone in understanding the quantum transport properties of carbon nanostructures. As we will see in section Sec. 3.4.2, the conductance of an idealized carbon
nanotube can be deduced directly from its band structure and the anomalous quantum hall effect of graphene, that will be discussed in Chap. 7, is a direct consequence of the linear band dispersion in graphene at zero field. The band structure of graphene, as deduced first by P. R. WALLACE in 1947 [271] can be deduced analytically within the \( \pi \) orbital tight-binding approximation:

With the Bloch theorem applied to the two-dimensionally periodic lattice shown in Fig. 1.10, the Hamiltonian in reciprocal space is a 2x2 matrix dependent on the \( k \)-vector:

\[
\mathcal{H}(k) = \begin{pmatrix} 
\epsilon_0 & 1 + e^{ika_1} + e^{ika_2} \\
1 + e^{-ika_1} + e^{-ika_2} & 0 
\end{pmatrix},
\]

which has two eigenvalues symmetric around \( \epsilon_0 \):

\[
E(k) = \epsilon_0 \pm \gamma_0 \left| 1 + e^{ika_1} + e^{ika_2} \right|
\]

In cartesian coordinates the lattice vectors can be written as \( a_{1/2} = a(\sqrt{3}e_x \pm e_y)/2 \) with \( a = \sqrt{3}d_{CC} \) and the energies can be expressed as:

\[
E(k) = \epsilon_0 \pm \gamma_0 \sqrt{1 + 4 \cos(\sqrt{3}k_xa/2) \cos(k_ya/2) + 4 \cos^2(k_ya/2)}.
\]

The characteristic shape of the resulting band structure is displayed in Fig. 2.2.

**Figure 2.2.** Band structure of a graphene in sheet as given by Eq. (2.2). The Fermi “surface”—for two-dimensional systems usually a one-dimensional manifold—is reduced to Fermi points at the corners of the hexagonal Brillouin zone. Of the six corners, only two are inequivalent (\( K \) and \( K' \)).

The \( \pi \) bands of an undoped graphene sheet are half-filled, so the Fermi energy is exactly at \( E_F = \epsilon_0 \), resulting in an exact electron-hole symmetry between \( E_F + \delta E \) and \( E_F - \delta E \).

Doping or electrical gating may move the physical Fermi energy away from its neutral value. To consistently refer to the energy where the
graphene bands touch in the $K$ points, we will therefore use the term charge neutrality point (CNP). This point will generally be the reference point of our energy scale $\varepsilon_0 = 0$.

At the center of the Brillouin zone (the $\Gamma$ point), the band structure shows a simple parabolic band structure that allows a quadratic approximation

$$E(k) \approx \pm 3\gamma_0 \mp \frac{\hbar^2 k^2}{2m^*}$$

with the effective mass $m^* = \frac{2\hbar^2}{a^2\gamma_0}$. The region of much more physical relevance, however, is that around the CNP. Here, the valence and the conduction band touch at single points, forming rotationally invariant double-cone structures approximated by the linear dispersion relation:

$$E(K^{(i)} + \delta k) \approx \pm v_F |\delta k|$$

with $K^{(i)} \in \{K, K'\}$ and the Fermi velocity $v_F = \sqrt{3a\gamma_0/2\hbar}$ (see Fig. 2.3). As it turns out, this dispersion relation can be derived from the linearization of the Hamiltonian itself, which can be expressed in a form equivalent to a relativistic Dirac particle in two dimensions. In Sec. 7.3 and App. B.1, this will be used in the derivation of the anomalous Landau level spectrum of graphene.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{graphene_bands.png}
\caption{Figure 2.3.: Zoom into the band structure of a graphene sheet around the CNP: Around each of the two non-equivalent $K$-points, the bands can be approximated by rotationally symmetric double cones with a slope determined by the Fermi velocity $v_F = \sqrt{3a\gamma_0/2\hbar}$.}
\end{figure}

\subsection*{2.2.1. Graphene with full basis set}

The band structure developed in Eq. (2.2) is based on the highly idealized $\pi$ orbital tight-binding model with only nearest neighbor hopping terms.
To show the quality of this simplified model, we can compare it to the more refined parameterization from Sec. 2.1.3. A direct comparison of the band structure in both models is shown in Fig. 2.4. As can be seen, the $\pi$ bands are well described by the simple model and the $\sigma$ bands do not come close to the Fermi energy and therefore have negligible influence on most electronic properties. The degeneracy of the bands at the $K$ points and the double-cone shape near these points is exactly preserved due to the symmetry of the lattice.

Figure 2.4.: Comparing the bands resulting from the orthogonal $\pi$ orbital parameterization ($\gamma_0 = 2.66$ eV) with those of the full non-orthogonal $sp^3$ parameterization (Sec. 2.1.3). As can be seen, the additional $\sigma$ bands lie energetically well separate from the Fermi energy. The shape of the $\pi$ band is nearly identical in both cases.

Note that the parameters of the refined model were adjusted to match the total energy, so the close match of the $\pi$ bands actually is a significant result, indicating that the simple $\pi$ orbital model may be an even better approximation than it is often assumed.

2.3. Band structure of single-wall carbon nanotubes

In the zone-folding approximation, a carbon nanotube is be viewed as a graphene sheet with periodic boundary conditions, neglecting curvature effects [107, 231, 275]. For a $(M, N)$ CNT, vectors in real space are identified modulo the chiral vector $L = Ma_1 + Na_2$. Any electronic wave function must obey this condition as $\psi(r + L) = \psi(r)$. In reciprocal space, this maps to a selection criterion for allowed $k$ vectors based on the condition:

$$\frac{L \cdot k}{2\pi} \in \mathbb{Z}$$

The $K$ point at the corner of the Brillouin zone can be expressed as $K = (2a_1 + a_2) / 3$ leading to a simple rule that determines whether this point
Chapter 2. Electronic structure

Figure 2.5.: Zone-folding diagram for an armchair CNT. The selected lines in \( k \)-space always cut through the \( K \)-points, so the CNT always is truly metallic. Left: band structure of one-dimensional system, right: density of states (DOS) exhibiting van Hove singularities.

belongs to the set of allowed \( k \) vectors in the rolled up system:

\[
\frac{L \cdot K}{2\pi} = \frac{(M_1 + N_1) + (2\bar{a}_1 + \bar{a}_2)}{2\pi} = \frac{2M + N}{3} \in \mathbb{Z}, \quad (2.3)
\]

equivalent to the condition on \( M - N \) to be an integer multiple of 3.

Whenever this condition holds, the \( K \) point fulfills the periodic boundary conditions, so the Fermi energy is in the spectrum of the CNT, i.e. the CNT is metallic. In any other case, the closest lines of allowed \( k \) vectors miss the Fermi point by

\[
\delta k = \frac{2\pi}{3V} \cdot \frac{2\pi}{3V} = \frac{2}{3d}
\]

with the tube diameter \( d \). Within the linear approximation at the \( K \) points, this results in a gap of [275]:

\[
\Delta E_{\text{gap}} = 2v_F \hbar \delta k = \frac{4v_F \hbar}{3d}
\]

Generally, the zone-folding approximation works reasonably well for large diameter CNTs but breaks down in thin CNTs due to curvature effects. In armchair CNTs this does not have much qualitative effect, because the bands crossing at the CNP are strictly protected by the intrinsic supersymmetry of the system [161]. In “metallic” zigzag and chiral CNTs, however (i.e. those that would be metallic according to Eq. (2.3)), the calculation with more refined tight-binding parameterizations reveals a tiny gap opening at the Fermi energy [181], which has also been confirmed experimentally [201] (see also Fig. 2.8).
2.3. Band structure of single-wall carbon nanotubes

Figure 2.6.: Zone-folding diagram for two different zigzag CNTs. The (10,0) CNT does not meet the condition given in Eq. (2.3). The discretization lines miss the $K$-point so the CNT is semiconducting. The (9,0) CNT is metallic within this simple approximation but shows a small gap when curvature is taken into account.

Figure 2.7.: Zone-folding procedure for a chiral CNT. The periodicity of $(4N,N)$ CNTs is still fairly short and the Brillouin zone therefore large enough to visualize the band structure. The mirror symmetry is broken. As for the $(3N,0)$ zigzag CNT, the metallicity of this system will also be broken by curvature effects.
2.4. Band structure of graphene nanoribbons

The energy bands of GNRs can be constructed similarly to those of CNTs except for the boundary effects. Where zone folding of CNTs is done by imposing periodic boundary conditions in the lateral direction, the edges of GNRs demand for fixed boundaries, resulting in a different discretization condition. Besides this, however, the peculiar edge state of zigzag GNRs cannot be described by zone folding arguments at all but needs to be obtained from the finite-width Hamiltonian directly.

For armchair GNRs, a similar rule for the metallicity exists as it does for CNTs (see Fig. 2.9):

$$\frac{N_a + 1}{3} \in \mathbb{N}$$  \hspace{2cm} (2.4)

The situation for zigzag GNRs is slightly different: the discrete lines of allowed $k$ vectors never cross the $K$ point in the graphene Brillouin zone. However, the edge state always closes the gap that would otherwise remain (see Fig. 2.10). As we will see in Sec. 5.5 the localization caused by the zero velocity state very efficiently prohibits any charge transport even for very low disorder.

2.5. Beyond tight-binding: Density functional theory

One general disadvantage of the parameterized tight-binding approach is that it does not offer any means by itself to obtain the individual matrix elements. As shown in Sec. 2.1.1, these can be obtained by a fit to
2.5. Beyond tight-binding: Density functional theory

Figure 2.9.: Zone-folding diagram for two different armchair GNRs. For \( N_a = 8 \) the ribbon is semimetallic. The bands crossing at the CNP are easily split into a gap by perturbations. For \( N_a = 9 \), the ribbon has a finite width gap, as predicted by Eq. (2.4).

Figure 2.10.: Band structure and DOS of a zigzag GNR. The zone-folding diagram in this case fails to describe the edge state visible as the flat band and the peak in the DOS at the CNP. All other bands are described correctly by the diagram, similarly to an armchair CNT.
Chapter 2. Electronic structure

experimental results, but the functional form still has to be guessed and transferability is always questionably.

An alternative is given by ab initio methods that allow to computationally obtain quantitative results of quantum mechanical multiparticle systems based on fundamental physical constants only. The most common approach of this kind is density function theory (DFT), based on the theorems by P. Hohenberg and W. Kohn [120] and the equations by W. Kohn and L. J. Sham [144], which will be sketched out in the following, based on Ref. [131].

We start from the full Hamiltonian of \( N \) interacting electrons in a static external potential \( V_{\text{ext}}(\mathbf{r}) \):

\[
\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m_e} + \frac{e^2}{4\pi\varepsilon_0} \sum_{i \neq j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{N} V_{\text{ext}}(\mathbf{r}_i)
\]

The key quantity of DFT is the electron density:

\[
n(\mathbf{r}) = N \int d\mathbf{r}_2 \ldots d\mathbf{r}_N \Psi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)
\]

It is important to note that the electron density of the ground state completely determines the ground state wave function, so in principle all ground state observables could be obtained from it.

On which the Levy-functional is defined as:

\[
F[n] = \min_{\Psi \to n} \langle \Psi | \mathcal{T} + V_{\text{ee}} | \Psi \rangle
\]

where the minimum \( \min_{\Psi \to n} \) runs over all \( N \)-particle wave functions \( \Psi \) that lead to the electron density \( n \). Note that this is a very general functional independent of the system in question.

If we denote energy, wave function and electron density of the ground state as \( E_{\text{GS}}, \Psi_{\text{GS}} \) and \( n_{\text{GS}} \), the two fundamental theorems of DFT can now be expressed as:

\[
E[n] \equiv F[n] + \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \geq E_{\text{GS}}
\]

and

\[
E[n_{\text{GS}}] \equiv F[n_{\text{GS}}] + \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) n_{\text{GS}}(\mathbf{r}) = E_{\text{GS}}
\]
That means the ground state electron density \( n_{\text{GS}} \) can in principle be found by minimization of the functional and, if the ground state is nondegenerate, also the ground state wave function and all observables.

The important step from here is now the insight that the ground state electron density of this fully interaction multiparticle system can also be interpreted as the ground state electron density of a different system without electron interaction and with a different potential. The Hamiltonian of this system called *Kohn-Sham Hamiltonian* can be formally expressed as:

\[
\mathcal{H}_{\text{KS}} = \frac{\mathbf{p}^2}{2m_e} + V_{\text{ext}}(\mathbf{r}) + V_{\text{ee}}(\mathbf{r}, [n])
\]

which is a single-particle Hamiltonian that contains all the complexity of the many-particle system hidden in the density-dependent functional \( V_{\text{ee}} \).

As a next step, the Hartree potential \( V_{\text{H}} \), which effectively is a mean field approximation of the Coulomb interaction, can be split out from \( V_{\text{ee}} \):

\[
V_{\text{ee}}(\mathbf{r}, [n]) = \frac{e^2}{4\pi\varepsilon_0} \int \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} n(\mathbf{r}') + V_{\text{xc}}(\mathbf{r}, [n])
\]

Finally, the remaining term \( V_{\text{xc}} \) can be expressed formally as a functional derivative:

\[
V_{\text{xc}}(\mathbf{r}, [n]) = \frac{dE_{\text{xc}}[n]}{dn(\mathbf{r})}
\]

of the *exchange correlation functional* \( E_{\text{xc}} \).

Up to this point, the complicated many-particle problem has been rewritten in a clever way as effective single-particle problem without any approximations. The real problem, however, is not solved at all. Instead, all complications are well hidden within the exchange correlation functional. The real power of the whole DFT method comes now from the fact, that already a very simple approximation of \( E_{\text{xc}} \) gives surprisingly good quantitative results: The *local density approximation* (LDA) assumes a very simple form:

\[
E_{\text{xc}}[n] = \int d\mathbf{r} e_{\text{xc}}(n(\mathbf{r}))
\]

where the function \( e_{\text{xc}} \) is simply taken from the homogeneous electron gas.
With this, a self-consistency loop can be set up by choosing some initial $n(r)$, diagonalizing the Kohn-Sham Hamiltonian and summing up the occupied states to obtain a new density $n(r)$ until convergence is reached. In this loop, the integration of $V_H$ can optionally be avoided by solving the equivalent Poisson equation:

$$\nabla^2 V_H = -\frac{e^2}{\epsilon_0} n(r)$$

Alternatives to the LDA like the \textit{generalized gradient approximation} (GGA) have been developed, that improve the precision in certain cases but typically either degrade precision in other cases or severely reduce performance.

A large number of implementations of the DFT method has been developed over the years with different points of focus, offering different features and extensions. One general purpose implementation that offers a special linear scaling approach that allows the efficient handling of large systems, is the SIESTA code [242], which we used for our study of metal contacts to CNTs (see Chap. 4).
Chapter 3.

Theory of quantum transport

Transport processes are the physical core of technology, with very different types of transport—energy, matter, charge, light, heat and in the near future even spin—governed in principle by the same set of laws. The historical starting point in the discovery of these laws was provided by the phenomenological description of electricity that was discovered over 200 years ago by some of the pioneers of physics, namely A.-M. Ampère, A. Volta, G. S. Ohm, C. A. de Coulomb, M. Faraday and others. One century later, when atomistic models of matter were starting to be accepted, a microscopic theory of electron transport was presented by P. Drude [72, 73], giving an intuitive explanation for the observed Ohmic resistance in metals. With the advent of quantum theory, these originally classical models were soon improved by A. Sommerfeld [243], taking into account the correct Fermi-Dirac statistics of electrons and thereby correcting the quantitative deficiencies in the purely classical theory.

So far, however, quantum transport theory was mostly concerned with materials, assumed to be shaped in macroscopic dimensions. It took several decades before technology allowed the reliable production and handling of structures at a scale that would itself give rise to quantum effects. Theory developed in parallel until in the 1980s coherent electron transport in quantum systems was clearly formulated in the form that we will use in this work by R. Landauer [154, 155].

From the technological point of view, this theory is just becoming an issue more and more, as the size of individual devices in integrated circuits is decreasing below the limit where quantum effects begin to play a crucial role in transport. In the following, we will give a short review of the concepts of quantum transport, closely following the presentation flow in the work by S. Datta [64].
3.1. Mesoscopic length scales

Somewhere in between single molecules, that need to be described by a fully quantum mechanical microscopic theory, and macroscopic bulk systems, where geometry and material properties can be handled independently, lies the regime of mesoscopic physics. Quantum transport in this regime is governed by the competition of various characteristic length scales that will be discussed in the following.

The system size

The most obvious length scale is the system size, defined by the geometry of the structure that is studied. Typically, the size of a sample varies greatly between the different directions, like in two-dimensional electron gas (2DEG) structures with thickness of a few atomic layers and lateral extensions at the scale of microns or quantum wires that are strongly restricted in two directions. SWCNTs typically are a few nanometers in both lateral directions and several hundred nanometers in length. For MWCNT, the lateral size may be 10 to 20 nanometers at similar length. Graphene has a thickness of only a few Angstrom while it extends up to many micrometers in two dimensions. For more complex structures, the specification of the system size may need more parameters, like the width of restrictions or the independent sizes of distinct components.

The Fermi wavelength

The equilibrium of electrons in any system is given by the Fermi-Dirac statistics. The electronic states are filled from the bottom of the spectrum up to the Fermi energy $E_F$. At moderate temperatures, most states in the spectrum are either fully filled or empty. Only a narrow region around $E_F$ effectively contributes to the degrees of freedom of the system and thereby determines its electronic properties.
3.1. Mesoscopic length scales

In crystalline structures, the band structure relates the energy $E_F$ to a specific wave vector $k_F$ and the associated wavelength $\lambda_F = 2\pi/k_F$. In general, this defines the length scale at which quantum mechanical interferences are to be expected.

When the system size of a sample is of the same order of magnitude as the Fermi wavelength or below, the spectrum is quantized to a discrete set of states. This allows a selective tuning of the dimensionality of a system as can be shown with the example of the free electron gas in a rectangular box:

The energy of a free electron in three dimensions is

$$E(k) = \frac{\hbar^2 k^2}{2m}$$

Giving a density of states of:

$$g_{3D}(E) = \frac{V}{(2\pi)^3} \int d^3k \delta (E - E(k))$$

$$= \frac{V}{(2\pi)^3} \int d^3k \delta \left( E - \frac{\hbar^2 k^2}{2m} \right)$$

$$= \frac{V}{(2\pi)^3} \frac{d}{dE} \left( \frac{4}{3} \pi (k(E))^3 \right)$$

$$= \frac{\hbar^3 V}{\pi^2} \sqrt{2mE}$$

including the initial factor 2 originating from the spin degeneracy. The volume $V$ in this case is assumed to be large compared with the Fermi wavelength $\lambda_F = 2\pi/k_F = 2\pi\hbar/\sqrt{2mE}$ in all three directions. Restricting the height $H$ of the system, this is not true any more and the integral has to be replaced by the corresponding sum over allowed $k$ vectors $k_z = \pi n_z/H$ given by the boundary conditions of the box:

$$g_{2D}(E) = \frac{2WL}{(2\pi)^2} \int dk_x dk_y \sum_{n_z} \delta \left( E - \frac{\hbar^2}{2m} \left[ k_x^2 + k_y^2 + \left( \frac{\pi n_z}{H} \right)^2 \right] \right)$$

$$= \frac{WL}{2\pi^2} \sum_{n_z} \frac{d}{dE} \left( \pi \left[ k \left( E - E_{n_z} \right) \right]^2 \right)$$

$$= \frac{WLm}{\pi\hbar^2} \sum_{n_z} \Theta \left( E - E_{n_z} \right)$$

$$E_{n_z} = \frac{(\hbar\pi n_z)^2}{2mH^2}$$
Chapter 3. Theory of quantum transport

with the Heaviside step function \( \Theta(x) = \frac{(|x| + x)}{2x} \). Applying the same restriction to the width \( W \) yields the characteristic van Hove singularities [268]:

\[
g_{1D}(E) = \frac{2L}{2\pi} \int dk_x \sum_{n_y,n_z} \delta \left( E - \frac{\hbar^2}{2m} \left[ k_x^2 + \left( \frac{\pi n_y}{W} \right)^2 + \left( \frac{\pi n_z}{H} \right)^2 \right] \right)
\]

\[
= \frac{L}{\pi} \sum_{n_y,n_z} \frac{d}{dE} \left( 2k \left( E - E_{n_y} - E_{n_z} \right) \right)
\]

\[
= \frac{L\sqrt{2m}}{\pi\hbar} \sum_{n_y,n_z} \left( E - E_{n_y} - E_{n_z} \right)^{-1/2}
\]

And finally, restricting the length \( L \) as well:

\[
g_{0D}(E) = \frac{2}{\pi} \sum_{n_x,n_y,n_z} \delta \left( E - \frac{\hbar^2}{2m} \left[ \left( \frac{\pi n_x}{L} \right)^2 + \left( \frac{\pi n_y}{W} \right)^2 + \left( \frac{\pi n_z}{H} \right)^2 \right] \right)
\]

\[
= \frac{2}{\pi} \sum_{n_x,n_y,n_z} \delta \left( E - E_{n_x} - E_{n_y} - E_{n_z} \right)
\]

<table>
<thead>
<tr>
<th>condition</th>
<th>dimensionality</th>
<th>example material</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_F \ll H, W, L )</td>
<td>3D</td>
<td>bulk crystals</td>
</tr>
<tr>
<td>( H \sim \lambda_F \ll W, L )</td>
<td>quasi-2D</td>
<td>thin films</td>
</tr>
<tr>
<td>( H \ll \lambda_F \ll W, L )</td>
<td>2D</td>
<td>2DEG, graphene</td>
</tr>
<tr>
<td>( H \ll W \sim \lambda_F \ll W )</td>
<td>quasi-1D</td>
<td>quantum wires, MWCNTs</td>
</tr>
<tr>
<td>( H, W \ll \lambda_F \ll W )</td>
<td>1D</td>
<td>atomic chains, SWCNTs</td>
</tr>
<tr>
<td>( H, W, L &lt; \lambda_F )</td>
<td>0D</td>
<td>quantum dots, molecules</td>
</tr>
</tbody>
</table>

Table 3.1.: Approximate classification of the dimensionality of systems according to Fermi wavelength \( \lambda_F \) and system sizes. The conditions \( H, W, L < \lambda_F \) include the case that in certain directions the system is too small to define anything like a wavelength at all.

In each case, if the extensions are of a similar magnitude as \( \lambda_F \), the sum runs over a few energy levels (leading to the “quasi-nD” cases) while for extensions much smaller than \( \lambda_F \), only the lowest level is relevant, resulting in truly reduced dimensions (See Tab. 3.1 and Fig. 3.2).

For systems without periodicity (like single molecules or CNTs in their lateral direction) the Fermi wavelength is undefined. Instead, the exact atomic structure has to be considered to determine the number of discrete levels/bands.
3.1. Mesoscopic length scales

**Figure 3.2.** Sketch of the spectra in systems of different dimensionalities: 0D-quantum dots have discrete levels. In 1D, the $1/\sqrt{E}$-shaped van Hove singularities [268] are present at band edges. The same band edges in 2D produce steps. Finally, in 3D, the band edge results in a $\sqrt{E}$-shaped spectrum.

The phase coherence length

The phase coherence length $\ell_\varphi$ is the mean distance that electrons can travel before losing their quantum mechanical phase due to scattering with other electrons or other dynamic scatterers like phonons or magnetic impurities carrying a spin degree of freedom. Different paths of one electron can only interfere with each other if the phases are preserved coherently.

The length $\ell_\varphi$ is not influenced by static impurities as these do not have an internal degree of freedom and thus influence the phase of a scattering electron in a deterministic way. Taking an ensemble average over various samples of static impurities, however, may result in a similar effect as true dephasing.

The phase coherence length is connected directly with the phase coherence time $\tau_\varphi$ via the Fermi velocity as:

$$\ell_\varphi = \nu_F \tau_\varphi$$

Typically, the rate of phase decoherence is the same as the rate of inelastic scattering events, as dephasing is usually linked with energy transfer. The length $\ell_\varphi$ generally decreases with increasing temperature or increasing DOS at the Fermi energy, as both result in more electron-electron collisions.

The elastic mean free path

The elastic mean free path $\ell_{el}$ is the mean distance between elastic, phase-preserving scattering events. This length is generally given by the defect density and the amount of static disorder that perturbs the periodicity
of the crystal lattice. Again, the Fermi velocity gives a direct connection with the time between elastic scattering events:

\[ \ell_{\text{el}} = v_F \tau_{\text{el}} \]

For weak disorder that does not relax the electron momentum completely in each scattering process, the elastic mean free path is identified with the momentum relaxation length, which is the average distance over which the information about the original momentum of an electron is completely lost due to scattering events.

### The localization length

The localization length \( \ell_{\text{loc}} \) is given by the average spatial extent of quantum mechanical states. In a perfectly periodic system, the Bloch states are spread out infinitely (i.e. \( \ell_{\text{loc}} = \infty \)). In 1958, it was proposed by P. W. Anderson that defects and disorder may lead to a complete localization of all quantum mechanical states in a system, i.e. to a restriction of each state to a finite spatial extent, depending on dimensionality of the system and possibly the strength of the disorder [18] (see Chap. 5 for details). In this case, the localization length takes on a finite value. Formally, the localization length can be expressed via the spatial Green function as:

\[
\ell_{\text{loc}} = -\lim_{\ell \to \infty} \frac{\ell}{\ln |G(r, r + \ell)|^2}
\]

If \( \ell_{\text{loc}} \) is much smaller than the system size, the system is in the regime of strong localization. If both are of similar magnitude, it is in the regime of weak localization.

### The magnetic length

A further length scale is given by the magnitude of an applied external magnetic field \( B \): The magnetic length \( \ell_{\text{mag}} \) is typically defined as

\[
\ell_{\text{mag}} = \sqrt{\frac{\hbar}{eB}}
\]

and relates to the area that carries one magnetic flux quantum \( \Phi_0 = h/e \) in a given external magnetic field. The magnetic length determines the scale at which Aharonov-Bohm effect, negative magnetoresistance due to weak localization and many other mesoscopic effects involving magnetic fields take place. In Chap. 7, some of these effects will be discussed in more detail.
3.2. The transport regimes

Based on the competition of the various length scales introduced in the last section, we can distinguish various regimes of quantum transport depending on the system length $L$:

For $L \sim \lambda_F$, the system has a discrete spectrum. Typical for this case are effects like Coulomb blockade or resonant tunneling that probe the individual states of the system. For $L \gg \lambda_F$ with a continuous spectrum the other length scales become relevant.

At $L \ll \ell_{el}$, the system is in the regime of ballistic transport. Here, the resistance of a system is independent of $L$ and only depends on the contacts. For $L \gg \ell_{el}$, in the regime of diffusive transport, transport is dominated by multiple scatterings and resistance grows with $L$.

Typically, only a fraction of the scattering events destroy the phase, so $\ell_\phi$ is generally larger than $\ell_{el}$ and determines how the resistance grows with $L$: In the regime of coherent transport ($L \ll \ell_\phi$), resistance grows exponentially with the length. Only when coherence is destroyed at a scale much smaller than the length of the conductor ($L \gg \ell_\phi$), Ohm’s law is recovered with a linear dependence of resistance on length.

Finally, the localization length competes directly with the phase coherence length. In the delocalized regime ($L \ll \ell_{loc}$) the individual electronic states spread from one contact to the other and allow coherent transport of electrons. In this case, decoherence is detrimental for the transport and resistance will grow with increasing dephasing (e.g. by increasing temperature). In the regime of strong localization ($L \gg \ell_{loc}$), on the other hand, coherent transport within a single state is exponentially suppressed. Conductance can only originate from thermally activated hopping between the levels, so the resistance will go down along with increasing dephasing.

<table>
<thead>
<tr>
<th>conditions</th>
<th>regime</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_F \sim L \ll \ell_{el}, \ell_\phi, \ell_{loc}$</td>
<td>discrete spectrum (resonant tunneling)</td>
</tr>
<tr>
<td>$\lambda_F \ll L \ll \ell_{el}, \ell_\phi, \ell_{loc}$</td>
<td>ballistic transport</td>
</tr>
<tr>
<td>$\lambda_F, \ell_{el} \ll L \ll \ell_\phi, \ell_{loc}$</td>
<td>diffusive transport</td>
</tr>
<tr>
<td>$\lambda_F, \ell_{el}, \ell_\phi \ll L, \ell_{loc}$</td>
<td>Ohmic resistance</td>
</tr>
<tr>
<td>$\lambda_F, \ell_{el} \ll \ell_{loc} \sim L \ll \ell_\phi$</td>
<td>weak localization</td>
</tr>
<tr>
<td>$\lambda_F, \ell_{el}, \ell_{loc} \ll L \ll \ell_\phi$</td>
<td>strong localization (exponentially suppressed)</td>
</tr>
<tr>
<td>$\lambda_F, \ell_{el}, \ell_{loc}, \ell_\phi \ll L$</td>
<td>strong localization (thermally activated)</td>
</tr>
</tbody>
</table>

Table 3.2.: Overview of various transport regimes
Chapter 3. Theory of quantum transport

3.3. Landauer transport formalism

The quantum mechanical problem of coherent electron transport through a nanoscale device can be formulated within scattering theory: electrons from one lead enter the conductor, are either reflected or transmitted with a certain probability and move out into either one of the leads. The total current that can be measured in dependence of the voltage and other parameters is then given by the balance of the partial currents between the individual leads.

It is assumed that within each lead, the electrons are in thermal equilibrium at any time. Non-equilibrium distributions caused by electrons flowing in from other leads are supposed to be relaxed at much shorter time scales than the processes in the scattering region that acts as bottleneck for the electron flow.

Within each lead \( p \), the electrons therefore follow the Fermi-distribution according to its chemical potential \( \mu_p \) and the temperature \( T \):

\[
n_p(E) = f(E - \mu_p) = \left[ \exp \left( \frac{E - \mu_p}{k_B T} \right) + 1 \right]^{-1}
\]

The current through the system from lead \( p \) to lead \( q \) the system is now given via the Landauer-Büttiker formula:

\[
I_{p \rightarrow q} = -\frac{2e}{h} \int dE \left[ n_p(E) - n_q(E) \right] T_{p \rightarrow q}(E)
\]  

(3.1)

Where \( T_{p \rightarrow q}(E) \) is the total transmission function at the energy \( E \), which, as a sum of the individual channel transmission probabilities, may exceed unity. A derivation of this formula will be given in the next section, along with an expression for the transmission function.

In the special case of a symmetric two-terminal setup, the chemical potentials in both leads are given directly by the bias voltage as \( \mu_{L/R} = E_F \mp eV_b/2 \) and the transmission function symmetric in both directions. The differential conductance can in this case be derived as:

\[
\frac{dI}{dV_b} = \frac{2e^2}{h} \int dE \frac{-f'(E - E_F + \frac{eV_b}{2}) - f'(E - E_F - \frac{eV_b}{2})}{2} T(E, V_b)
\]

A further specialization is now the limit of zero temperature in which case the Fermi distribution turns into a step-function and its derivative into
3.4. The quantum mechanical transmission

A $\delta$-distribution $f'(E) = -\delta(E)$. Further assuming that the transmission function does not depend on the bias ($T(E, V_b) = T(E)$)

$$ \frac{dI}{dV_b} = \frac{2e^2}{h} T \left( \frac{E_F - eV_b}{2} \right) + T \left( \frac{E_F + eV_b}{2} \right) $$

(3.2)

Setting $V_b = 0$ now results in the Landauer conductance formula in its simplest form:

$$ G := \left. \frac{dI}{dV_b} \right|_{V_b=0} = \frac{2e^2}{h} T(E_F) $$

with the prefactor $G_0 = \frac{2e^2}{h} \approx (12.9 \text{ k}\Omega)^{-1}$ called conductance quantum, traditionally including the factor 2 originating from spin degeneracy.

3.4. The quantum mechanical transmission

Common derivations of an expression for the transmission function $T(E)$ are based on leads with internal translational invariance or periodicity that allows free electrons or Bloch waves with a well defined velocity. As it turns out, this condition is unnecessarily restrictive. In the following a straightforward derivation will be given that holds for arbitrary tripartite systems.

The most general two-terminal system is given by a tripartite Hamiltonian of the form:

$$ \mathcal{H} = \begin{pmatrix} H_L & H_{Lc} & H_{LR} \\ H_{cL} & H_c & H_{cR} \\ H_{RL} & H_{Rc} & H_R \end{pmatrix} $$

(3.3)

where $H_L$ and $H_R$ correspond to the left and right lead and $H_c$ to the conductor. These leads should be infinite or at least large enough to provide a spectrum that can be considered continuous. The conductor is often chosen finite but any infinite bath or additional leads of a multiterminal setup would also be considered part of $H_c$ for the following calculation.

The full Hamiltonian can be separated as:

$$ \mathcal{H}_0 = \begin{pmatrix} H_L & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & H_R \end{pmatrix} $$

$$ \mathcal{H} = \mathcal{H}_0 + \mathcal{V} $$
where $\mathcal{H}_0$ has two distinct sets of eigenstates in either lead: $\mathcal{H}_L \ket{l} = E_l \ket{l}$ and $\mathcal{H}_R \ket{r} = E_r \ket{r}$.

The commonly known expression for the transition rate from one state $\ket{l}$ to a state $\ket{r}$ due to a small perturbation $\mathcal{V}$ would be $\Gamma_{r\leftarrow l} = \frac{2\pi}{\hbar} \delta(E_l - E_r) \left| \langle r|\mathcal{V}|l \rangle \right|^2$. In this case, however $\mathcal{V}$ cannot be considered small, so we have instead to take into account the full expansion, leading to the nonperturbative expression:

$$\Gamma_{r\leftarrow l} = \frac{2\pi}{\hbar} \delta(E_l - E_r) \left| \langle r|\mathcal{V} + \mathcal{V}\mathcal{G}'(E_l)\mathcal{V}|l \rangle \right|^2$$

with the retarded Green function of the full system:

$$\mathcal{G}'(E) = (E + i0^+ - \mathcal{H})^{-1}$$

Both leads are filled according to their different chemical potentials, meaning that state $\ket{l}$ is filled with probability $n_L(E_l)$ and $\ket{r}$ with probability $n_R(E_r)$. This leads to an expression for the charge current from $L$ to $R$ based on classical probability flow:

$$I = -2e \sum_l \sum_r \left[ n_L(E_l) - n_R(E_r) \right] \Gamma_{r\leftarrow l}$$

$$= -2e \sum_l \sum_r \left[ n_L(E_l) - n_R(E_r) \right] \frac{2\pi}{\hbar} \delta(E_l - E_r) \left| \langle r|\mathcal{V} + \mathcal{V}\mathcal{G}'(E_l)\mathcal{V}|l \rangle \right|^2$$

$$= -\frac{2e}{\hbar} \int dE \left[ n_l(E) - n_R(E) \right] T(E),$$

which is exactly of the form given in (Eq. (3.1)), including the spin-degeneracy factor 2. The transmission function can be expressed as:

$$T(E) = \sum_l \sum_r 2\pi \delta(E_l - E_r) 2\pi \delta(E_l - E_r) \left| \langle r|\mathcal{V} + \mathcal{V}\mathcal{G}'(E_l)\mathcal{V}|l \rangle \right|^2$$

$$= \sum_l \sum_r 2\pi \delta(E_l - E_r) \left| \langle r|\mathcal{V} + \mathcal{V}\mathcal{G}'(E_l)\mathcal{V}|l \rangle \times \right.\left. \times 2\pi \delta(E_l - E_r) \langle l|\mathcal{V} + \mathcal{V}\mathcal{G}'(E_l)\mathcal{V}|r \rangle \right.$$
3.4. The quantum mechanical transmission

at which point the spectral function of the leads, defined via the Green function of the disconnected lead $G'_0 = (E + i0^+ - H_0)^{-1}$:

$$A_L(E) = i \left( G'_{0L}(E) - G'_{0R}(E) \right) = \sum_l |l \rangle 2\pi \delta (E - E_l) \langle l |$$

(and correspondingly for $A_R$) can be used to write:

$$T(E) = \text{Tr} \left[ A_R \left( V + V G' (E) V \right) A_L \left( V + V G' (E) V \right) \right] \quad (3.4)$$

Usually, the original system is chosen such that there is no direct link between the left and the right lead, i.e. $H_{LR} = H_{RL} = 0$. In this case, we can exploit the fact that the spectral functions have support only within the corresponding lead and write:

$$T(E) = \text{Tr} \left[ A_R \left( H_{RL} \right) H_{cL} \right] \left[ A_L \left( H_{cR} \right) \right] \quad (3.4)$$

where $G'_c$ is the Green function projected to the conductor region. Using the decimation technique (see App. A), this can conveniently be expressed as

$$G'_c = (E - \mathcal{H}_c - \Sigma'_L - \Sigma'_R)^{-1}$$

using the self energy from the leads defined as:

$$\Sigma'_L = \mathcal{H}_{cL} (E + i0^+ - \mathcal{H}_L)^{-1} \mathcal{H}_{cL}$$

Finally, one can define

$$\Gamma_L = H_{cL} A_L H_{cL}$$

$$\Gamma'_L = i \left( \Sigma'_L - \Sigma'_L \right)$$

to simplify the expression above to its final form:

$$T(E) = \text{Tr} \left( \Gamma'_R G'_c \Gamma'_L \right) \quad (3.5)$$

This equation is sometimes attributed to D. S. Fisher and P. A. Lee. In their work [87] they actually they actually gave an equivalent form that can be obtained by exploiting the fact that the spectral function is positive definite and Hermitian (so it has a well-defined square root $\chi$ with $\chi^2 = \mathcal{A}$) and writing:

$$T(E) = \text{Tr} \left[ \chi L H_{cL} G'_{c} H_{cR} \chi R H_{cR} G'_{c} H_{cL} \chi L \right]$$

$$= \text{Tr} \left[ t^L t \right] \quad (3.6)$$
defining \( t := \chi_R H_R \mathcal{G} \chi_L \). The individual transmission eigenchannels can be obtained from this matrix via singular value decomposition, yielding:

\[
\begin{align*}
 t &= \sum_{n=1}^{N_{\text{ch}}} |r_i\rangle t_i \langle l_i| \\
\end{align*}
\]

(3.7)

with the number of nonzero transmission channels \( N_{\text{ch}} \) and the individual channel transmission amplitudes \( 0 < t_i \leq 1 \). The concept of transmission eigenchannels was first mentioned by M. Büttiker \([44, 126, 208]\).

### 3.4.1. Transmission through single molecules

A first application of this formalism can be found in the description of conductance through single molecules as it is measured in various experiments \([221, 288, 99]\). In the simplest case, this system can be modeled by the Hamiltonian of the molecule with \( N \) energy levels \( E_n \) equally connected to two structureless metallic leads. The energy dependence in the leads is negligible compared to that of the molecule, so \( \Sigma_L \) and \( \Sigma_R \) can both be set to a constant. According to the Kramers-Kronig relations \([68, 145]\), a constant Green function must necessarily be purely imaginary, thus:

\[
\mathcal{G}_L = -i \mathcal{A}_L / 2 ; \quad \mathcal{G}_R = -i \mathcal{A}_R / 2
\]

Further, the leads are assumed to couple equally to all states of the molecule, so the self energies are diagonal with constant diagonal entries:

\[
\Sigma_L = -i (\Delta_L/2) \mathbb{1} ; \quad \Sigma_R = -i (\Delta_R/2) \mathbb{1}
\]

where the positive constant \( \Delta_L = \Delta_R = \Delta \), the Breit-Wigner formula for the transmission is readily obtained \([40, 249, 92]\):

\[
T(E) = \frac{\Delta^2}{(E - E_n)^2 + \Delta^2}
\]

(3.8)

which results in very characteristic resonant peaks in the transmission as demonstrated in Fig. 3.3. A review of the issue of molecular junctions from
the theoretical point of view can be found in Ref. [191]. Several aspects of this that are relevant to transport in CNTs will also be discussed in Chap. 4.

Figure 3.3.: Transmission through a single molecule between structureless leads as given by the Breit-Wigner formula [see Eq. (3.8)]. For $\Delta$ smaller than $\delta E$ (the distance between adjacent energy levels), the transmission consists of independent peaks of width $\Delta$ and height 1. For larger $\Delta$, close peaks are smeared out.

3.4.2. Transmission of periodic systems

A quasi-one-dimensional periodic system can generally be expressed by a Hamiltonian of the form:

$$
\mathcal{H} = \begin{pmatrix}
0 & H_{10} & H_{00} & H_{01} & 0 & \vdots \\
\vdots & 0 & H_{00} & H_{01} & 0 & \vdots \\
0 & \cdots & 0 & H_{00} & H_{01} & 0
\end{pmatrix}
$$

One possibility to handle such a system would be to virtually partition it according to (Eq. (3.3)). While this is a reasonable approach when using numerical tools, it is more instructive analytically to exploit the fact that such a Hamiltonian can be diagonalized using Bloch states:

Given the periodic length $a$ of the system, an effective Hamiltonian can be constructed for wave number $k$ as:

$$
H_{\text{eff}}^k = H_{00} + e^{ika} H_{01} + e^{-ika} H_{10}
$$
which is a finite matrix that can be diagonalized:

\[ H_k^b \Psi_{kb} = E_{kb} \Psi_{kb} \]

where \( b \) is the band index. A periodic system generally has as many bands as there are orbitals within one unit-cell.

A Bloch state is now easily constructed as:

\[ |\Psi_{kb}(n)\rangle = e^{ikn} |\psi_{kb}\rangle \]

where \( n \) is now the index of the unit cell and \( m \) the index of the orbital within a unit cell. This state is an eigenstate of the original periodic Hamiltonian \( \mathcal{H} \) with the energy \( E_{kb} \).

The quantum mechanical probability current density of a Bloch wave is given by:

\[
J_{kb}^{n \rightarrow n+1} = \frac{i}{\hbar} \left( \langle \Psi_{kb}(n) | H_{01} | \Psi_{kb}(n+1) \rangle - \langle \Psi_{kb}(n+1) | H_{10} | \Psi_{kb}(n) \rangle \right) \\
= \frac{i}{\hbar} \langle \psi_{kb} | \left( e^{ika} H_{01} - e^{-ika} H_{10} \right) | \psi_{kb} \rangle \\
= \frac{1}{\hbar a} \langle \psi_{kb} | \frac{dH_k^b}{dk} | \psi_{kb} \rangle \\
= \frac{1}{\hbar a} \frac{dE_{kb}}{dk}
\]

Which can be expressed directly in terms of the velocity of Bloch electrons:

\[ v_{kb} = \frac{1}{\hbar} \frac{dE_{kb}}{dk} \]

The charge current can now be constructed directly from the quantum mechanical current density and the filling factors at both ends of the system.

\[
I = -e \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} dk \sum_b \left[ n_L(E_{kb}) J_{kb} \Theta(J_{kb}) + n_R(E_{kb}) J_{kb} \Theta(-J_{kb}) \right] \\
= -e \int dE \left[ n_L(E) - n_R(E) \right] \int_{-\pi/a}^{\pi/a} dk \sum_b \delta(E - E_{kb}) \frac{dE_{kb}}{dk} \Theta \left( \frac{dE_{kb}}{dk} \right)
\]

The step functions \( \Theta(x) = (x + |x|)/2x \) in the first line of this equation distinguish between left- and right moving Bloch waves. Both contributions cancel out exactly for every energy, preventing a current without voltage.
From the final line, we see that the transmission of a periodic system is an integer number that can be obtained directly from counting the number of right-moving bands present at any given energy \( E \) which thereby present the conduction channels of the system:

\[
T(E) = N_{\text{ch}}(E) = \int_{-\pi/a}^{\pi/a} dk \sum_b \delta(E - E_{bk}) \frac{dE_{bk}}{dk} \Theta\left(\frac{dE_{bk}}{dk}\right)
\]

One example where this principle can nicely be observed is the linear ladder, displayed in Fig. 3.4.

**Figure 3.4.** The linear ladder as a demonstrative example for a periodic system. With two orbitals per unit cell, there are two cosine shaped bands of width \( 4\gamma \), split by \( 2\gamma \) against each other. In the common energy range there are two transmission channels open, above and below, only one channel each.

### 3.4.3. Systems with periodic leads

The concept of conduction channels of periodic systems developed in the previous section can now be generalized to systems where an arbitrarily formed scattering region is contacted by periodic leads. Having periodic leads means that the energy eigenstates of the complete system asymptotically deep within the leads have to be linear combinations of Bloch states. In addition, there may be states localized around the scattering center that extend into the leads with exponentially decaying tails. These bound states, however, do not contribute to the conductance of the system.

The total current in such a system is strictly limited by the current that can be carried by either of the leads, which is—as demonstrated in the
previous section—given by the number of channels. For the transmission that means:

\[ T(E) \leq N_{\text{ch}}(E) = \min(N_{\text{ch}}^{L}(E), N_{\text{ch}}^{R}(E)) \]

Another way to arrive at this relation is through the decomposition of the transmission matrix into channels (Eq. (3.7)): As the spectral function of one lead \( \mathcal{A}_{L/R}(E) \) has as many nonzero eigenvalues as there are channels within this lead. In the basis of these channels in each lead, the transmission matrix \( t \) therefore is a \( N_{\text{ch}}^{L}(E) \times N_{\text{ch}}^{R}(E) \) matrix, permitting \( N_{\text{ch}}(E) \) transmitting channels with a maximum transmission amplitude of one each.

**Figure 3.5.** Transmission through a linear ladder with both semi-infinite ends at different potential. As the potential step is smoothed out over five unit cells, scattering is negligible. However, transmission at every energy strictly limited by the lowest transmission of both leads. As the bands are shifted in opposite directions in both leads, the window of transmission is closed to the overlapping region.

### 3.4.4. Tunneling contacts

A tunneling contact is formed by a point contact in a quantum device that forms the single weakest link along the device and thereby dominates the total resistance. Tunneling contacts are formed for example by an STM tip on a surface, where the exponential decrease of the conductance with increasing tunneling distance is exploited to determine the distance between tip and surface with atomic precision without physically touching the surface. Another way to form tunneling contacts are mechanically controlled break junctions.

Within the formalism developed before, such a contact can be described by a Hamiltonian of the form given in (Eq. (3.3)) where the Hamiltonian

66
of the conductor $H_c$ is reduced to size zero and the entries of the hopping matrix $H_{LR}$ are small enough to be handled perturbatively. In this case, the $V^2$ contributions in (Eq. (3.4)) can be dropped and the transmission is approximated by:

$$T(E) = \text{Tr} \left[ A_R(E) H_{RL} A_L(E) H_{LR} \right].$$

The transmission is therefore determined by the spectral weight of the individual channels at the surface of both leads. The structure of the tunneling matrix $H_{LR}$ selects which channels of the leads are connected to conduction channels.

Note that this approximation fails to describe the correct behavior at van Hove singularities in the case of one-dimensional leads: even though the spectral weight has a singularity, the actual transmission should still be limited by the number of channels.

![Figure 3.6: A tunneling contact between two semi-infinite linear ladders. The tunneling transmission can be obtained perturbatively in $\gamma_c/\gamma_0$ as:]

$$T(E) \propto \sum_{n=1}^{N_{ch}} g_n(E)^2 \left( \frac{\gamma_c}{\gamma_0} \right)^2$$

where $g_n(E)$ is the local DOS (LDOS) at the last atoms of one semi-infinite chain of a single band $n$ (due to symmetry protection between the two bands).

### 3.4.5. Resonant tunneling and Fabry-Pérot physics

When placing two tunneling contacts into the same chain at a certain distance, the effects of quantum mechanical interference in coherent transport become most obvious. In Fig. 3.7, it can be observed how such system shows resonant tunneling comparably with a molecular junction.

An interesting parallel can be drawn between this behavior and the optical Fabry-Pérot interferometer [83, 184]. Shining a laser of wavelength
Chapter 3. Theory of quantum transport

Figure 3.7.: Transmission of a linear ladder with weak links in two points at distance $N = 10$. The vertical lines indicate the discrete energy levels of a finite chain of length $N = 10$. As can be seen, the transmission is similar to a single molecule between contacts (see Fig. 3.3). The width of the Breit-Wigner resonances varies with energy, caused by the non-constant DOS of the leads.

A beam through two consecutive parallel semi-transparent mirrors placed at a distance $L$ results in oscillatory interference patterns with peaks in transparency exactly when a standing wave can build up inside the cavity, i.e. when $L = n\lambda$ for $n$ integer.

In an electron transport setup, for electron wave lengths near the Fermi energy in a fixed resonator length $L$, the spacing between allowed states $\Delta \lambda = \frac{\lambda_F^2}{L}$ relates to an energy spacing between individual resonances of:

$$\Delta E = \frac{v_F h}{L}$$

Figure 3.8.: Experimental data showing the Fabry-Pérot resonances in electron transport through a finite carbon nanotube: The physical length $L \approx 200$ nm of the device and Fermi velocity $v_F \approx 8 \times 10^5$ m/s give $\Delta E \approx 0.016$ eV. The oscillations in the gate voltage with $\Delta V_g \approx 0.83$ V are related to $\Delta E$ by the gate capacitance $C_g$. (Figure taken [164])

Experimentally, the Fermi energy in a conductance measurement can be controlled by the gate voltage $V_b$, allowing the direct measurement of
3.4. The quantum mechanical transmission

Fabry-Pérot oscillations in the regime of linear conductance as displayed in Fig. 3.8. However, the proportionality constant between $V_g$ and $E_F$ is unknown as it is determined by the gate capacitance

$$C_g = \frac{L g_F e E_F}{V_g}$$

where $g_F = 8/\hbar v_F$ is the DOS per length at $E_F$ of the CNT.

To circumvent this unknown factor, it is possible to use the data from a finite bias measurement which—neglecting charging and other interaction effects—is described by Eq. (3.2) and plotted in Fig. 3.9.

Figure 3.9: Theoretical (left) and experimental (right) data showing Fabry-Pérot resonances in electron transport through a finite carbon nanotube at finite bias $V$. The vertical size of the diamonds $\Delta V_b \approx 15$ mV allows the direct extraction of the resonator length $L = v_F h/\Delta E = v_F e h/\Delta V_b \approx 220$ nm which corresponds to the physical dimension of the sample. (Figure taken from Ref. [164])

3.4.6. Structureless leads

In Sec. 3.4.1, the concept of wide-band leads was introduced for modeling structureless leads to a molecular junction. The same kind of leads can also be used for a finite piece of a periodic system at the center. As can be seen from Fig. 3.10, the behavior for small $\Delta$ is similar to that in the resonant tunneling case with periodic leads. The difference, though, becomes clear for strong $\Delta$, where the resonances correspond to the levels of a reduced system. Furthermore, one can observe that for no value of $\Delta$ the system is as transparent as the perfectly homogeneous system.
Figure 3.10.: Linear ladder between wide-band leads. For small $\Delta$, the transmission shows resonances at the energy levels of the finite $N = 10$ linear ladder, similar to the resonances for the single molecule. For large $\Delta$, the resonances coincide with the levels of a $N = 8$ ladder, indicating that the strong coupling to the leads effectively pulls the contact atoms into the leads. For intermediate $\Delta$, the transmission comes close to the transmission of the infinite periodic chain in Fig. 3.4.

3.5. Beyond coherent transport: interactions and decoherence

The Landauer transport formalism captures only coherent quantum transport without correlations between electrons or inelastic scattering processes. The Coulomb interaction between electrons can be captured in a mean field approximation and decoherence can to some degree be mimicked by non-Hermitean terms in the Hamiltonian. For a real description multiparticle effects, however, one has to use a different formalism.

The extreme case of decoherence can be found in the description of Coulomb blockade in a quantum dot. When the rate of decoherence is larger that at which electrons move in and out of the dot, the system can be described in the orthodox theory, basically considering the electrons to follow classical statistics without quantum interference between individual state transitions [14]. A carbon nanotube with very weak contacts shows such Coulomb blockade [233, 86]. The purely statistical description used for the description of this completely incoherent system, however, cannot be extended to include effects of quantum coherence in any straightforward way.

A theoretically sound way to include any kind of interactions between electrons or interactions between electrons and environment perturbatively is that of nonequilibrium Green functions (NEGF) based on the formali-
ism by L. V. Keldysh [138, 17, 66, 65, 207]. In the absence of interactions, this formalism is equivalent to the Landauer formalism that we presented before. This method is widely used and can be directly incorporated with ab initio methods [259].

Some of the most interesting physical effect caused by electron interaction, however, cannot be described by such perturbative methods, one prominent example being the Kondo effect that is typically observed in metals but has also been found in CNTs [25], another one being the physics of the Luttinger liquid [169] which was proposed in 1963 by J. M. Luttinger as an exactly solvable theoretical model and has now been discussed for some time as a model for the transport in SWCNTs [37, 76, 236, 217].

Most of these effects, however, have in common that the effects caused by interactions are a distinct but small deviation from the results obtained for noninteracting electrons. A thorough understanding of the physics of coherent transport is therefore essential before any attempt to incorporate interactions into a simulation.
Chapter 4.

Electrical contacts to nanotubes and -ribbons

The high conductivity of carbon nanotubes (CNTs) can be attained thanks to a unique combination of several features. The quasi-one-dimensional crystal structure together with a low density of defects allow to explore the theoretical limit of conductance of $4e^2/h$ at the charge neutrality point. The stiffness of carbon-carbon bonds reduces the effect of electron-phonon coupling at room temperature [95]. Also, restricting electron movement to a single dimension results in a very small phase space which strongly reduces the effectiveness of scattering. A further reduction of backscattering is caused by the low density of states (DOS) at the charge neutrality point (CNP) in combination with a high Fermi velocity (see Chap. 5). Considering all these factors, measured ballistic lengths of several microns [172] become understandable. Yet, to exploit the potential for carrying current densities of up to $10^9$ A/cm$^2$ [278], the contacts at nanometer scale become crucial.

Experimentally a crucial factor for obtaining good metallic contacts are the wetting properties of the material. Thus, it has been observed that Ti, Ni and Pd form continuous coatings on single-wall CNTs (SWCNTs) while Au, Al and Fe form isolated particles [283]. Furthermore, among several common contact metals, Ti was found to be the only one where true chemical bonds could be observed, while the others showed only weak van der Walls interactions [282]. Surprisingly enough, Pd—traditionally known as a rather poor conductor—was found to form better and more reliable Ohmic contacts to CNTs than Ti [172] and could be successfully applied to produce a CNT field effect transistor with Ohmic contacts [127]. It is generally believed that this superiority of Pd is due to its high work function ($\phi_{\text{Pd}} = 5.1$ eV) that matches well with that of CNTs (e.g. $\phi_{(7,0)\text{CNT}} = 5.1$ eV [239]) and thereby avoids a high Schottky barrier. Pt, which has an even higher work function, would therefore be expected
to perform even better as a contacting material, but as it turns out it does not form Ohmic contacts at all [128].

A further experimental puzzle is the question of the effective length of contacts formed by a coating metal layer: While some studies report that transport occurs only at the edge of the contact [172], others state that the contact resistance depends on the length of the contact [270].

Various theoretical studies have been conducted to investigate these issues: in ab initio studies comparing Au, Pd and Pt contacting a (8, 0)-CNT, Pd was found to have the lowest Schottky barrier [238]. Another ab initio study comparing the metal-graphene bonding of the same three metals indicated a very weak binding energy for Au. For Pd it is somewhat stronger while for Pt it is yet stronger [171]. The bad contacts formed by graphene are here blamed on a clustering effect of larger metal grains. Direct ab initio simulations of transport in a metal-contacted CNT compared Pd and Au [137], finding again that Pd forms superior contacts.

Apart from these practical issues in explaining and improving the quality of materials, the study of contact models is also of great theoretical relevance: In studying the physics of electronic devices at the nanometer scale, it is generally crucial to have detailed control over the contacts [116, 61]. Indeed, specifying properties of nanoelectronic devices generally is completely meaningless without clearly stating the way the system was contacted or—for theoretical studies—how the contact was modeled.

In the following chapter, we will therefore first take a look at the effect of various traditional theoretical models of nanotube contacts before demonstrating a more realistic model as we had proposed in Ref. [2]. A detailed analytical study of a minimal model will be followed by various generalizations and finally a direct application to quantitative results obtained from microscopic ab initio computations. Based on these results we will give an alternative interpretation of the seemingly contradicting experimental findings.

The last section of this chapter will then deal with spin transport in CNTs under the influence of weak disorder, which is, as it turns out, determined mostly by the contacts.

### 4.1. Conventional contact models

In theoretical studies of electron transport in CNTs, different approaches for modeling the contacts can be found. CNT electrodes are an elegant way to avoid the issue of contacts altogether by continuing the CNT
4.1. Conventional contact models

indefinitely in both directions. Structureless electrodes are a concept known from modeling transport in single molecules and appreciated for its simplicity. And finally a whole family of contact models is based on a very specific geometry, often placing each atom of the contact region at an exact position. In the following we will take a look at each approach individually.

4.1.1. Carbon nanotube electrodes

One commonly used and elegant way to model the leads in a CNT-transport setup is to continue the CNT infinitely in both directions, thereby avoiding any additional parameters and any potential scattering point in the system. Though this approach does not claim to be a realistic model of any experimental setup, it conveniently allows to theoretically separate out contact effects from intrinsic effects of the CNT itself that may be subjected to various perturbations in the central region.

Without such perturbations, the system corresponds exactly to the perfectly periodic system as it was presented for the example of the linear ladder in Sec. 3.4.3. The transmission is an integer-valued step function that corresponds exactly to counting the number of channels, i.e. of positive-velocity bands at a given energy as found in Fig. 4.1.

Figure 4.1.: Transmission of a (5,5)-CNT contacted by homogeneous CNT electrodes. Effectively, the tripartite system is equivalent to a perfectly periodic CNT, so the transmission equals the number of channels that can be counted in the band structure. The band edges, visible as van Hove singularities in the DOS, coincide with the steps in the transmission.

In Chap. 5 and in Sec. 6.5 we will make extensive use of this contact model for studying the effects of disorder or a modified structure within the central region. However, as we will find out in this chapter, the concept of CNT electrodes is not only a theoretical construct, but it is also
the well defined limiting case of the more realistic contact model that we will describe.

### 4.1.2. Structureless electrodes

In Sec. 3.4.1, we introduced so-called *wide-band leads* to model metallic contacts neglecting their internal structure. As this approach is widely used for modeling transport through single molecules, it is also commonly used to model the contacts to CNTs, attempting to avoid the idealized but highly unrealistic semi-infinite CNTs \[148\]. However, the shortcomings of this approach are visible already from the linear ladder model system in Sec. 3.4.6 and remain valid for CNTs: The observed physics is strongly dominated by contact effects and depends strongly on the choice of the free parameter $\Delta$. Instead of providing a robust model that might describe a wide range of realistic situations, the simplistic wide band leads describe the very specific and highly unrealistic situation of a finite nanotube of exactly defined length perfectly contacted by metal at each end in the last ring of atoms only. While such atomically exact contacts might be a reasonable model for molecular electronics with chemically defined contacts, it is far from reality for CNT devices.

Fig. 4.2 shows the transmission of a short CNT with such structureless electrodes. The observed physics is very similar to the Fabry-Pérot resonances described in Sec. 3.4.5 which is indeed observed CNT transport measurements. We will see later on in this chapter, that this contact model is indeed similar to an extreme case of our more realistic model.

### 4.1.3. Atomically modeled electrodes

In the literature, various approaches have been described to model CNT-metal contacts for transport calculations with atomistic detail, sometimes within a tight-binding formalism \[147, 70\], but more often using DFT to determine the electronic structure \[189, 204, 166\]. As the results in Fig. 4.3 and 4.4 show, the resonances are clearly visible in every case, but the result depends strongly on microscopic details of the system and details of the model.

### 4.2. Extended contacts

In Sec. 3.4.1 the transmission of a single molecule between two leads was derived, leading to the Breit-Wigner equation Eq. (3.8). The same equation
4.2. Extended contacts

**Figure 4.2.** Transmission of a 5 nm long (5,5)-CNT contacted by wide-band leads at both ends. For low $\Delta$, the transmission shows a resonance at each energy level of the finite CNT. For high $\Delta$, the resonances are aligned with the levels of a CNT of a length reduced by one unit cell: Those atoms that are in direct contact with the leads are effectively removed from the resonator. At intermediate $\Delta$, transmission is closest to the optimum (indicated by the gray steps). Yet, it still shows resonant oscillations.

**Figure 4.3.** Example for a transport calculation through a (9,0)-CNT with gold contacts modeled with atomistic detail. As the two different transmission curves (left: extended Hückel, right: tight-binding) show, the observed behavior depends strongly on details of the model that cannot be controlled with enough precision. (Figure from Ref. [70])
limited to a single energy level \( \varepsilon \) and generalized for asymmetric leads \( \Gamma_L \) and \( \Gamma_R \) has the form:

\[
T(E) = \frac{\Gamma_L \Gamma_R}{(E - \varepsilon)^2 + (\Gamma_L + \Gamma_R)^2/4},
\]

(4.1)

The bell-shaped peak in this expression as a function of the energy is well known. What is rarely noted in literature, however, is the fact that also for fixed energy \( E \) and one fixed contact \( \Gamma_L \), the transmission as a function of the other contact \( T(\Gamma_R) \) has a bell shape with an optimum at the balanced coupling \( \Gamma_R = \Gamma_L \). A new perspective to this old issue was provided in the recent experiments by L. Grüter \textit{et al.} [99]. Small couplings (\( \Gamma_R \ll \Gamma_L \)) result in a linear \( \Gamma_R \) dependence of the transmission typical of tunneling phenomena, while for large coupling (\( \Gamma_R \gg \Gamma_L \)) such better contact \( \Gamma_R \) results in an overall suppressed transmission. One way to understand this counterintuitive behavior is to consider the tunneling rate \( \Gamma \) as a measure for the chemical bond between the conducting orbitals of the molecule and the lead: A strong bond to one of the leads causes the molecule itself to virtually become part of that lead so that we observe the physics of a single point contact. Furthermore, the strong bonding results in a strong redistribution of the spectral weight of the energy level in the molecule, i.e., in a low local DOS (LDOS) at the energy \( \varepsilon \). The tunneling transmission, which directly probes this LDOS will therefore be suppressed by large \( \Gamma_R \).

Once the length \( N \) of a contact is increased for an extended molecule, the optimal value decreases monotonically with the number of contact points \( N \), as displayed in Fig. 4.5. In particular for large \( N \) it will be shown that such critical value scales like \( \Gamma_R = \Gamma_L \ln N/N \) [see Eq. (4.7)].

In Ref. [2], we had considered a finite CNT contacted at both ends, which can indeed be viewed as a typical nanodevice measured in experiments. In contrast to a regular molecule contacted in single atomic positions, this system has the length of the contact as an additional physical ingredient.
4.2. Extended contacts

Figure 4.5.: Breit-Wigner resonance in an extended molecule. The internal hopping is fixed as $2\gamma = \Gamma_L$ to allow an optimal match to the left contact. In the case $N = 1$, the system is identical to the molecular junction described by Eq. (4.1). The transmission shows the shift of the Breit-Wigner peak towards lower $\Gamma_R$ with growing $N$. The functional form of this shift can be approximated for large $N$ as $\Gamma_R = \Gamma_R \ln N/N$ [see Eq. (4.7)].

Our results will be first displayed for a one extended contact only and will be generalized to the realistic situation of a finite scattering region of length $L_0$ between two extended contacts of length $L_c$. In experiments with CNTs these lengths typically measure several hundred nanometers.

4.2.1. Analytic model

A minimal model which captures the essential physics of extended nanotube contacts is set up as follows. The tube itself is represented by a linear chain of atoms with the hopping integral $\gamma$ and the on-site energy $\epsilon = 0$ (fixing the energy offset). A two-probe setup is defined by selecting an arbitrary single atom as the “conductor” and the semi-infinite sections at both ends as “leads”. In this unmodified setup, the system is fully transparent, the transmission $T(E)$ is equal to the number of channels $N_{\text{ch}}$ at any given energy. The single, cosine-shaped band of the linear chain provides a single transmission channel

$$T_{\text{band}}(E) = \Theta(E + 2\gamma) \Theta(-E - 2\gamma)$$

which presents a theoretical upper transmission limit when scattering at the contacts could be neglected.

An “extended contact” to the nanotube is now modeled by replacing the semi-infinite lead by a finite $N$-atom chain contacted in each atom individually by a wide-band lead of strength $\Delta$. (For a sketch of the model, see Fig. 4.6)

A full solution of this model is obtained by calculating the conductance as the quantum mechanical transmission probability:

$$G(E) = \frac{2e^2}{\hbar} T(E).$$
Chapter 4. Electrical contacts to nanotubes and -ribbons

Figure 4.6.: Minimal model for extended contacts solvable analytically: a linear chain of identical atoms (one orbital per atom) with hopping integral $\gamma$ between nearest neighbors. To the left, the chain continues infinitely, at the right end, $N$ atoms are contacted, each by an independent wide-band lead of strength $\Delta$. For defining the transmission, the system is virtually split into three regions: the “conductor” $C$ and the “leads” $L$ and $R$.

This can be done within the Landauer approach to transport by means of the Green function formalism as shown in App. B.2. As a result one obtains (for $\epsilon = 0$):

$$T(E) = \frac{8 \sqrt{4 - E^2/\gamma^2} \text{Im} \left( f_N \left( E/2\gamma - i\Delta/4\gamma \right) \right)}{|E/\gamma - i \sqrt{4 - E^2/\gamma^2} - 2f_N \left( E/2\gamma - i\Delta/4\gamma \right)|^2},$$

with $f_N(x) = U_{N-1}(x)/U_N(x)$. $U_N(x)$ are the Chebyshev polynomials of the second kind as given in Eq. (B.4). To gain full understanding of the physics described by this expression, the transmission $T(E)$ is plotted as a function of the energy for different values of the two parameters $N$ and $\Delta$ (i.e. the length and the quality of the contact region) in Fig. 4.7. Two regimes can be identified: An $N$-resonant regime for low $\Delta$/small $N$ where the transmission shows about as many peaks as atoms present in the contact region and an $N$-independent regime for high $\Delta$/large $N$ where the transmission shows no resonances and depends only on $\Delta$.

The two different plots in Fig. 4.7 illustrate two aspects: For fixed $N$ with increasing $\Delta$, the transparency of the system improves, goes through an optimum point and degrades again, while for fixed $\Delta$, the transparency improves with growing $N$ and saturates at an $N$-independent optimum. In both cases, the transmission goes through the two different regimes.

Especially the last point can be seen more clearly by looking at the reflection $R = 1 - T$ in the energy range of the single channel of our system:

$$R = \left| \frac{E + i \sqrt{4\gamma^2 - E^2} - 2\gamma f_N \left( E/2\gamma - i\Delta/4\gamma \right)}{E - i \sqrt{4\gamma^2 - E^2} - 2\gamma f_N \left( E/2\gamma - i\Delta/4\gamma \right)} \right|^2 \quad (4.2)$$

In Fig. 4.8 this observable is plotted in logarithmic scale, illustrating that the average value of the transmission already saturates at $N = 100$ (specifically for $\Delta = 0.1\gamma$). For larger values of $N$, the overall transparency is
4.2. Extended contacts

Figure 4.7.: Transmission through the system displayed in Fig. 4.6, as given in Eq. (B.8). Left panel: for a fixed contact length $N$, starting from low $\Delta$, the transmission first improves, reaches an optimum and then degrades again at high $\Delta$. Right panel: for fixed contact strength $\Delta$: transmission improves with growing $N$ and saturates for large $N$. not improved any further, but the $N$-dependent resonances are smoothed out.

$\Delta = 1/\gamma$

$\Delta = 0.1/\gamma$

$\Delta = 10/\gamma$

$\Delta = 0.01/\gamma$

$N = 15$

$N = 100$

$N = 30$

$N = 15$

$N = 100$

$N = 30$

Figure 4.8.: Contact reflection in the system displayed in Fig. 4.6, as given in (Eq. (4.2)). At fixed contact strength $\Delta$, with growing $N$, the contact becomes more transparent and saturates at an $N$-independent value.

To better understand the origin of this saturation we fix the energy to the half filling case $E = \epsilon$ and study the transmission and the reflection for varying contact lengths $N$ (see Fig. 4.9):

$$R(E = \epsilon) = \frac{1 + \frac{1}{2} f_N \left( -i \Delta / 4 \gamma \right) \left( 1 + \frac{1}{2} f_N \left( -i \Delta / 4 \gamma \right) \right)}{1 - \frac{1}{2} f_N \left( -i \Delta / 4 \gamma \right) \left( 1 - \frac{1}{2} f_N \left( -i \Delta / 4 \gamma \right) \right)^2}$$

(4.3)

Similar data was obtained before numerically for CNTs [187]. New insight, however, can be gained from the reflection in log-scale: Ignoring the even-odd oscillations in $N$, one observes first an exponential decay of
Figure 4.9.: Transmission $T$ (left panel) and contact reflection $R$ (right panel) in the system displayed in Fig. 4.6 at fixed energy $E = \varepsilon$ for varying contact length $N$ and selected values of the contact strength $\Delta$.

$R$ with increasing $N$, followed by an abrupt crossover to a $N$-independent value. Both the rate of decay and the saturation value depend on $\Delta$ in such a way, that for lower values of $\Delta$ the transparency initially improves slower with the contact length, but ultimately $R$ saturates at a lower value, meaning higher contact transparency. This result, which we presented before based on numerical calculations on CNTs [2], will be studied in more detail in the following, using our analytical expressions.

An expression for the $N$-independent regime can easily be obtained as the limit $N \to \infty$ of Eq. (B.6) as

$$
R_{N \to \infty}^{E=\varepsilon} = \frac{1 + i f_\infty \left( -i \Delta/4\gamma \right)^2}{1 - i f_\infty \left( -i \Delta/4\gamma \right)}
$$

$$
= \left( \frac{\sqrt{\Delta^2/4 + 4\gamma^2} - \Delta/2 - 2\gamma}{\sqrt{\Delta^2/4 + 4\gamma^2} - \Delta/2 + 2\gamma} \right)^2
$$

(4.4)

which can be further simplified for $\Delta \ll \gamma$ to obtain

$$
R_{N \to \infty, \Delta \ll \gamma}^{E=\varepsilon} = \frac{\Delta^2}{64\gamma^2}
$$

The validity of this approximation is illustrated in Fig. 4.10.

The approach for finding the corresponding approximation for the $N$-resonant regime is less rigorous, since a simple limit is not sufficient to capture the behavior in this case. A far better approximation is found graphically: the straight section in a semi-log-scale plot (inset of Fig. 4.10)
4.2. Extended contacts

Figure 4.10.: \( \Delta \)-dependence of the contact reflection \( R \) for various even (upper panel) and odd (lower panel) contact lengths \( N \). Solid: the exact value as given in [Eq. (4.2)]. Dashed: the limit \( R_{N \to \infty} \) given in [Eq. (4.4)], along with its approximation \( \Delta^2/64 \gamma^2 \), valid for \( \Delta \ll 1 \). Dotted: the approximation \( \exp(-N\Delta/\gamma) \), valid for even \( N \) in the \( N \)-resonant regime. The inset shows the identical data in semi-log-scale, further illustrating the precision of the \( \exp(-N\Delta/\gamma) \) approximation in the \( N \)-resonant regime.

indicates a clean exponential law. The missing coefficients are easily found from a Taylor expansion in \( \Delta = 0 \), yielding

\[
\begin{align*}
R_{N \text{ resonant}}^{\text{even}} & = \exp(-N\Delta/\gamma) \\
R_{N \text{ resonant}}^{\text{odd}} & = \exp(-(N+1)\Delta/\gamma)
\end{align*}
\]

both of which can be seen to fit precisely over the whole \( N \)-resonant region.

Having found good approximations for both regimes, the last missing piece is the crossover. For even \( N \), the smooth shape of the crossover in Fig. 4.10 suggests a simple function of the form \( R = \sqrt[n]{A^n + B^n} \) and indeed, we find that for the case \( n = 1/2 \)

\[
R_{\text{crossover}}^{\text{even}} = \left( \sqrt{R_{N \to \infty}^{\text{even}}} + \sqrt{R_{N \text{ resonant}}^{\text{even}}} \right)^2 \tag{4.5}
\]

gives an extremely good match over the full range of \( \Delta \). Moreover, a very similar function is found to match the crossover for odd values of \( N \):

\[
R_{\text{crossover}}^{\text{odd}} = \left( \sqrt{R_{N \to \infty}^{\text{odd}}} - \sqrt{R_{N \text{ resonant}}^{\text{odd}}} \right)^2
\]
Both approximations show slight deviations from the exact value for small $N$ but match with high precision for larger $N$. Obviously, the two reflection probabilities behave like squares of quantum mechanical amplitudes interfering either constructively or destructively with each other.

The parameter values where $R_{\text{resonant}}$ and $R_{N\to\infty}$ coincide are of special interest. In the case that $N \gg 1$, where this coincidence happens for $\Delta \ll \gamma$, the condition for this is simply

$$\exp\left(-N\Delta/\gamma\right) = \Delta^2/64\gamma^2$$

leading to an expression for the $\Delta$-dependent effective contact length

$$N_{\text{eff}}(\Delta) = \frac{2\gamma}{\Delta} \ln\left(\frac{8\gamma}{\Delta}\right) \quad (4.6)$$

over which a longer contact does not further modify transport. This can be interpreted as the length that contributes to the electron transmission for a very long contact.

The inverse of Eq. (4.6) can expressed using the Lambert-$W$ function [56]:

$$\Delta_{\text{opt}}(N) = 2\gamma W(4N)/N$$

which can be approximated in the range of interest as:

$$\Delta_{\text{opt}}(N) \approx 2\gamma \ln N/N \quad (4.7)$$

A complete overview of the contact reflection and both parameters $N$ and $\Delta$ is shown in Fig. 4.11.

### 4.2.2. Generalization to arbitrary injection energies

Having found the transport relations at the fixed energy $E = \varepsilon$, we can now continue with generalizing the results for $E \neq \varepsilon$. Assuming the general functional form of the reflection

$$R_{\text{resonant}} = \exp(-2N\Delta/\alpha_1)$$

$$R_{N\to\infty} = \Delta^2/\alpha_2^2$$

we find numerically:

$$\alpha_1 = \sqrt{4\gamma^2 - E^2(1 + \zeta)}$$
Extended contacts

Figure 4.11.: The contact reflection for even \( N \) obtained from Eq. (4.5). Well visible are the two regimes separated by the minimal line \( \Delta_{\text{opt}}(N) \). The “exact” value for \( \Delta_{\text{opt}} \) is the true minimum for fixed \( N \). The “approximate” value comes from Eq. (4.7).

which holds for arbitrary fixed \( N \) with an approximate error estimate \( |\zeta| \lesssim 1/N \) capturing the resonant oscillations. Considering the characteristic form of the DOS of the linear chain:

\[
g(E) = \left( \pi \sqrt{4 \gamma^2 - E^2} \right)^{-1},
\]

we can rewrite the last expression as:

\[\alpha_1 \approx 1/\pi g(E),\]

reflecting the similarity to a weak point contact where the tunneling transmission is proportional to the DOS on either side. The last relation could be confirmed numerically to hold very generally, as will be discussed below when considering realistic contacts.

For the regime of \( N \to \infty \), we can similarly find an expression numerically:

\[\alpha_2 = \frac{2}{\gamma} \left( 4 \gamma^2 - E^2 \right)\]

which fits the exact formula Eq. (4.2) with arbitrary precision for fixed \( \Delta \ll \gamma \) small enough and \( E \) not too near to the band edges. Unlike the formula for \( \alpha_1 \), however, expressing \( \alpha_2 \) as a function of the DOS alone does not help to generalize the relation to other structures.

The crossover region, generally governed by interference effects, can be approximated by averaging over quantum mechanical phases, resulting in:

\[R_{\text{crossover}} = R_{N \to \infty} + R_{\text{resonant}}\]
which gives a good approximation for the full parameter space with $N \gg 1, \Delta \ll \gamma$ and $E$ away from band edges. Apart from the resonant oscillations, this now allows the full description of the reflection and we find a precise numerical confirmation of the previously obtained expression of the effective contact length:

$$N_{\text{eff}}(\Delta) = \frac{1}{\pi g(E) \Delta} \ln \frac{8\gamma^2 - 2E^2}{\gamma \Delta}$$

### 4.2.3. Non-diagonal contacts

To generalize our results beyond the diagonal contact approximation, we model the contacting metal not as single-parameter wide-band lead but as material with an internal structure, leading to off-diagonal terms in the contact matrix (see Fig. 4.12).

Fig. 4.13 illustrates that the off-diagonal terms in the self-energy do not bring any qualitative changes to the behavior described before. An exact quantitative mapping, would depend strongly on the details of the model.

### 4.2.4. Realistic contacts to carbon nanotubes and graphene nanoribbons

For the case of carbon nanotubes, the method of the Chebyshev polynomials cannot be used to obtain an analytical solution, due to the the
4.2. Extended contacts

**Figure 4.13.** Nondiagonal contacts: Reflection $R$ of the generalized model displayed in Fig. 4.12. Relating the parameter $\Delta = \gamma_2^2 / \gamma_\perp$ to the parameter $\Delta$ of the wide-band leads, the results are qualitatively similar to those of the original model (Figs. 4.9 and 4.10). One prominent difference is the enlarged reflection $R(\Delta)$ for $\Delta/\gamma$ between 1 and 10: While the diagonal self-energy was uniform for every atom along the contact, the nondiagonal self-energy now is sensitive to the edge of the contact. For large values of $\Delta$, where only the atoms near the edge contribute to the transport, this causes the visible deviation from the $R = \Delta^2 / \alpha_2^2$ law. The same reason is behind the visible irregularities in the resonant oscillations of $R(N)$.

**Figure 4.14.** Extended contacts to a (6,6) CNT (left) and the corresponding $N_z = 12$ zigzag GNR (right). Top panel: DOS with characteristic van Hove singularities. The zigzag-edge-state in the GNR causes a peak at the CNP. Center panel: The value $\alpha_1 = -2N\Delta / \ln R$, here computed for $N = 40$ and $\Delta = 10^{-5}$ eV, lies already very near to the limiting case $\alpha_1 = N_{\text{ch}} / \pi g(E)$. Bottom panel: The value $\alpha_2 = \Delta / \sqrt{R}$ is well converged for $N \to \infty$ and $\Delta = 10^{-2}$ eV. Note the suppression of both $\alpha_1$ and $\alpha_2$ in the ribbon at the CNP where the presence of the localized edge state suppresses the conductance in the contact region.
non-commutativity of the partial Hamiltonians of the periodic structure. In numerical studies, however, we find that the behavior is identical to that of the linear chain, except for a quantitative adjustment of the parameters $\alpha_1$ and $\alpha_2$ (see Fig. 4.14). The $N$-resonant regime can be described precisely by a simple generalization of the law found for the linear chain

$$R_{\text{resonant}} = \exp(-2N\Delta/\alpha_1) \quad \alpha_1 = N_{\text{ch}}/\pi g(E)$$

where $N_{\text{ch}}$ is the number of channels and $g(E)$ the total DOS per unit cell. Generally, both values are dependent on the energy and the chirality of the tube. For metallic CNTs near the CNP, however, one finds the general values of $N_{\text{ch}} = 2$, $g = 2N_{\text{ch}}/3\gamma d_{CC}$ and therefore $\alpha_1 = 3\gamma d_{CC}/2\pi \ell_{uc}$. ($\gamma = 2.66$ eV and $d_{CC} = 1.42$ Å). Introducing the physical length of the contact region $L = \ell_{uc}N$ with the length of the unit cell $\ell_{uc}$, the previous formula can be rewritten as:

$$R_{\text{resonant}} = \exp(-2L\Delta/\alpha_1 \ell_{uc}) \quad \alpha_1 \ell_{uc} = 1.80 \text{ eV } \text{Å}$$

For the $N$-independent regime, the general law of $R_{N \to \infty} = \Delta/\alpha_2$ still holds, but the functional form of the parameter $\alpha_2$ at arbitrary energies could not be determined. Generally, it turns out that $\alpha_2$ is suppressed at van Hove singularities in a similar to $\alpha_1$. Furthermore, metallic CNTs have a fairly constant value of $\alpha_2$ around $E_F$. At the CNP, we find $\alpha_2 = 4.24 \gamma$ for armchair CNTs and $\alpha_2 = 5.66 \gamma$ for metallic zigzag CNTs.

For graphene nanoribbons (GNRs), the situation is slightly more complex due to the presence of edge states in at zigzag edges [91, 186]. In metallic ribbons with armchair edges, i.e. unrolled zigzag-CNTs with a chiral vector of $(3N + 1, 0)$, the situation is similar to that of CNTs and we find at the CNP a value of $\alpha_2 = 8.0 \gamma$. The quantitative difference to the value of the corresponding CNTs can be explained by the presence of only one channel at the CNP. For GNRs with zigzag edges carrying localized edge states, however, we find that the constant $\alpha_2$ is completely suppressed at the CNP due to the peak in the DOS, caused by the edge state.

Physically, this suppressed value of both $\alpha_1$ and $\alpha_2$ would mean that the injection of electrons into the edge state via extended leads is generally inefficient. However, since the edge state itself has a very low dispersion and is therefore not a robust conduction channel, the physical relevance of electron injection is questionable in any case.
4.2.5. Three-terminal setup and Fabry-Pérot physics

A realistic setup for conduction measurements in CNTs and GNRs generally needs a second contact at the other end of the tube to close a circuit. Such a setup is well known to lead to Fabry-Pérot-like oscillations of the conductance along the energy range (see Sec. 3.4.5). For very bad contacts, Coulomb blockade has been observed, but we intentionally avoid this regime which would demand for the inclusion of charging effects.

One important aspect of Fabry-Pérot oscillations is their experimental use in measuring the length of the scattering region. In the zero-bias differential conductance, the spacing of the Fabry-Pérot resonances depends on the gate capacitance \( C_g \) as \( \delta V_g = e/C_g \). Only the diamond shapes in a plot of the finite bias differential conductance
\[
\frac{dI}{dV_b} = \frac{e^2}{h} \left[ T \left( \frac{V_g C_L}{e g} - \frac{e V_b}{2} \right) + T \left( \frac{V_g C_L}{e g} + \frac{e V_b}{2} \right) \right]
\]
give access to the spacing of the energy levels \( \delta E = \hbar v_F / L_0 \) and can thereby be used to measure the length \( L_0 \) of the resonator.

As visible in Fig. 4.15, however, the amplitude of these oscillations is strongly reduced as soon as the effective contact length exceeds the length of one unit cell. One could view this situation as smooth contacts that cause the Fabry-Pérot oscillations to be broadened and the resonator length \( L_0 \) to be ill-defined.

In some experiments using extended contacts on CNTs, the length of the scattering regions was measured to be just as long as the uncovered region of the tube [172], which could be explained based on our model by a strong contact \( \Delta \) and therefore a short effective contact length. For weaker contacts \( \Delta \) it is to be expected that Fabry-Pérot oscillations cannot be cleanly observed any more. A point defect inside the contacted region might, of course, act as a scattering point instead and give rise to oscillations that indicate a resonator longer than the uncovered region of the CNT.

In our previous study [2], we chose to average this oscillating conductance over \( E_F \pm 0.5 \text{ eV} \) in order to separate the finite-length Fabry-Pérot effects from the effects caused by the contacts themselves. Physically, this is similar to the thermal effects caused by high enough temperature. For the chosen conductor length of \( L_0 = 100 \text{ nm} \), this approach was very successful in canceling all Fabry-Pérot oscillations and reproducing the physics of a single extended contact. The resonance oscillations within the contact were, of course, also strongly suppressed by the averaging, leaving only a minimal signature that we correctly identified as such.
Chapter 4. Electrical contacts to nanotubes and -ribbons

Figure 4.15: Differential conductance through linear chain with symmetric, infinite length extended contacts. The central region consists of $N_0 = 200$ atoms. The Fabry-Pérot-oscillations in the gate voltage $V_g$ depend on the total gate capacitance $C_g$ as $\delta V_g = e/C_g$. The extent of the diamonds in the bias voltage $V_b$ depends directly on the level spacing $\delta E = \nu_F h/2L_0 = 2\gamma \pi / N_0$ as $\delta V_b = e \delta E$. Top: strong coupling $\Delta = 10 \gamma$, leading to a short effective contact length producing sharp resonances and a distinct diamond pattern. Bottom: moderate coupling $\Delta = 1 \gamma$ leading to an effective contact length of 4 unit cells and sinusoidal oscillations with small amplitude. For yet weaker coupling as it is to be expected for Pd or Ti, the amplitude is reduced even further. A strong defect within the contact region may act as point of scattering and recover sharp resonances.
4.2.6. Non-epitaxial contacts

Unlike the theoretical model contacts presented so far, realistic samples produced in experiment are never perfectly epitaxial but contain imperfections due to fabrication faults, lattice mismatch or metal faceting. To check whether the effects described so far are robust to such perturbations, we have investigated various kinds of disorder at the contact. Relatively weak disorder was implemented as random fluctuations of the contact parameter $\Delta$ on each atom $i$ as

$$\Delta^\text{fluct}_i = \Delta (1 + \xi^\text{fluct}_i W)$$

with an evenly distributed random variable $-1 \leq \xi^\text{fluct}_i \leq 1$ and the parameter $W$ specifying the relative strength of the fluctuations. As can be seen in Fig. 4.16, even for the strongest possible value $W = 1$, the effect of the disorder is moderate and purely quantitative.

Yet stronger disorder was realized by using a model of diluted contacts, where only a randomly selected fraction of the atoms in the contact region is contacted:

$$\Delta^\text{diluted}_i = \begin{cases} \Delta / P & \text{with probability } P \\ 0 & \text{with probability } 1 - P. \end{cases}$$

This kind of disorder gives considerable larger changes to the observed behavior, but even in the extreme case of a 1% dilution (i.e. $P = 0.01$), the general trend of the original model is well preserved (see Fig. 4.16).

4.2.7. Material related calculations

To link the model results obtained so far to the physical properties of real contact materials, we performed density functional theory (DFT) calculations of Ti and Pd monolayers interacting with a graphene layer as described before [2]. We described valence electrons by Troullier-Martins pseudopotentials and used the Perdew-Zunger form of the exchange-correlation functional in the local density approximation to DFT, as implemented in the SIESTA code [242]. With a double-zeta basis and a 100 Ry energy cutoff in the plane-wave expansions of the electron density and potential, we found the total energy to be converged to $\lesssim 1$ meV/atom. We performed a full structure optimization to determine the equilibrium adsorption geometry, the adsorption energy, and the local charge redistribution caused by the metal-graphene interaction. Since the interatomic distances in bulk Pd (2.7 Å) and Ti (2.95 Å) lie close to the honeycomb
Figure 4.16.: Effects of two different kinds of disorder on extended contacts. Left: relatively weak fluctuation disorder with varying contact strength $\Delta_i$ on each atom $i$. Right: stronger dilution disorder with only a randomly selected fraction of the atoms in the contact region attached to a lead. In each case, the parameter $\Delta$ refers to the average contact strength, i.e. for a dilution of $P = 0.1$, each individual contact has a strength of $10\Delta$ to keep the effects comparable.
4.2. Extended contacts

spacing in graphene (2.46 Å), we considered only epitaxial adsorption. For both Pd and Ti, we found a slight preference for the sixfold hollow site on graphite. For Pd, we found the equilibrium interlayer distance to be 3.2 Å, consistent with a relatively weak, mostly covalent bond energy of 0.3 eV per Pd atom. The bond between an epitaxial Ti monolayer and graphene was only insignificantly stronger with 0.4 eV per Ti atom at an interlayer distance of 3.0 Å.

Figure 4.17.: Charge density redistribution \( \Delta \rho(r) = \rho_{Me/C}(r) - \rho_{Me}(r) - \rho_C(r) \) in (a) Pd and (b) Ti monolayers interacting with a graphene layer, indicating regions of charge depletion and excess with respect to the superposition of isolated layers. (c) Schematic double-layer geometry in top view, with the cutting plane used in (a) and (b) indicated by the dash-dotted line.

The quality of nanotube-electrode contacts has been shown to depend sensitively on the Schottky barrier in semiconducting nanotubes [114] and band bending in metallic nanotubes, both reflecting the charge transfer within the junction. Our Mulliken population analysis indicates a net charge transfer of only 0.1 electrons from Pd and Ti to the graphene layer. More useful information is contained in the charge redistribution, depicted in Fig. 4.17. Results for Pd electrodes, shown in Fig. 4.17(a), suggest an accumulation of excess charge in the region between Pd and graphene layers. As seen in Fig. 4.17(b), the charge redistribution in Ti/C is very different, suggesting charge accumulation in the atomic layers, depopulation of the interlayer region, and thus an increase of the interlayer scattering potential. The lower scattering potential and the populated interlayer state at the Pd/C junction appear well suited for carrier injection into the nanotube, making the Pd/C contact superior to the Ti/C contact.

To study the electronic coupling between the two system, we study the band structures (see Fig. 4.18 and 4.19). Especially for Pd as a contacting
metal, the extraction of parameters for our model is greatly simplified by the fact that the band structures of the two sheets in isolation are very well preserved when the hybrid system is formed. One can see a rigid shift of the carbon bands by $E_C = 0.374$ eV while the palladium bands are shifted slightly in the opposite direction with $\Delta E_{Pd} = -0.020$ eV. On top of this rigid shift, one can observe slight hybridization effects in the band structure. For injecting conduction electrons into a graphene sheet, modeling the wall of a carbon nanotube, the most important area of the Brillouin-zone is the $K$-point, where graphene has its states near the Fermi energy. In Fig. 4.19, a small avoided crossing is visible near this region in the Pd/C band structure. To extract an estimate of tight-binding parameters from this data, we modeled a honeycomb lattice and a matching hexagonal lattice representing both sheets. As it turned out, a single orbital per atom is sufficient to obtain bands that can be fitted to the hybridizing bands in this region of interest with a single parameter each. Now, an additional coupling between the two sheets was introduced, linking each Pd-atom with its six neighboring C-atoms. This hopping parameter could then be tuned to reproduce a hybridization between the two tight-binding band structures, which is close to that in the hybrid band structure obtained from DFT, resulting in a coupling of $t_{Pd/C} \approx 0.15$ eV.

For the case of Ti, the distortions caused by the hybridization of the two layers are considerably stronger than for Pd. Still, a rigid shift can be determined as $\Delta E_C = -1.15$ eV. To determine the hopping parameter, the same procedure as for Pd could not be directly applied, because the relevant band of the Ti-monolayer cannot be reproduced with a single-orbital hexagonal lattice. Instead, a rough estimate was obtained by visually comparing the band structures themselves where the avoided crossing near the $K$-point is at least twice as large as for Pd, giving an estimate of $t_{Ti/C} \gtrsim 0.3$ eV.

To turn these parameters into values of $\Delta$ that can be directly placed into our model calculations, we need the surface DOS (SDOS) which is comparable for both materials at $N_{Me} \approx 1$ eV$^{-1}$. Finally, the connectivity at the interface is also important: Each C-atom contacted to three different metal atoms simply triples the value of $\Delta$. The internal connections inside the metal are already taken into account with the SDOS and do not have to be considered any further. With the relation $\Delta = t^2 N$, we obtain rough estimates of $\Delta_{Pd} \approx 0.06$ eV and $\Delta_{Ti} \gtrsim 0.3$ eV.

Unfortunately, this approach of computing a graphene layer and a layer of the contacting material within a common unit cell cannot necessarily
4.2. Extended contacts

Figure 4.18.: Scheme of the method applied for the analysis of the hybridization between a graphene sheet and a metal monolayer. The hybrid band structure matches well with an overlay of the band structures of the two individual systems rigidly shifted in energy and hybridization at some band crossings.

Figure 4.19.: Analysis of the hybridization between a graphene sheet and a metal monolayer by the method explained in Fig. 4.18. Highlighted are the regions of interest, i.e. those hybridizations that contribute most to the electron injection.
be transferred to other materials of interest in any straightforward way. There exist, however, ab initio calculations of various metals in contact with graphene or CNTs that show a clear trend [74, 171]: The highly conducting metals Au, Ag and Cu generally have a very weak binding energy, insufficient for wetting the carbon surface, so a clean contact is hard to achieve. Pd, Pt and Ti all have sufficient binding energies for wetting the surface. Pt and Ti both have higher binding energies than Pd.

The authors A. Maiti and A. Ricci [171] set out to explain the seeming contradiction of why Pt and Ti both form worse contacts than Pd, even though their chemical bonding is stronger. Finally they attribute this effect to the formation of metal clusters of different sizes. Our own results that were just presented offer a more fundamental explanation: It is exactly the weak bonding between Pd and graphene or CNTs—just large enough to wet the surface—that makes Pd such an excellent contact material.

4.3. Ferromagnetic contacts and spin transport

Due to their one-dimensional nature and their reported extremely long mean free paths over hundreds of nanometers, CNTs have been expected to play an important role in spintronics. It is, in fact, hoped that the spin-charge separation phenomenon [237] together with the almost negligible spin-orbit coupling for carbon would result in ballistic pure spin currents (no charge!) over submicrometer lengths. Still, despite the numerous theoretical indications, CNT spintronics could not show any surprising phenomena over the last decade. One of the reasons for that is believed to be found in the ill-defined electrical contacts between ferromagnets and CNTs over a very small area. For this reason it is important to realistically model ferromagnetic contacts to CNTs and to separate disorder at the contacts from possible disorder which could naturally be present along the CNT due to the fabrication process.

Experimental measurements of the magnetoresistance in CNTs contacted by ferromagnetic electrodes have revealed a clear signal of spin filtering and magnetoresistance of up to 9% [263, 199]. Theoretical works have proposed a mechanism based on Fabry-Pérot resonances that allow a device to be sensitive to the small energy differences that occur in spin polarized transport [209, 147, 146]. This theory is indeed confirmed by ex-

1Very recent experiments with graphene have, in fact, shown more encouraging results [261].
4.3. Ferromagnetic contacts and spin transport

Experiments that show the magnetoresistance of such a device to be tunable by electrical fields [228].

It was our interest to investigate the robustness of the spin transport in CNTs against disorder within the tube itself. As the question of spin transport in CNTs is mostly an issue of the contacts, we will present this study at this point, in anticipation of the more comprehensive study of the effects of disorder in general that will be the issue of Chap. 5.

4.3.1. Modeling ferromagnetic leads

The model used is similar to that described in Ref. [147]: The model by M. Julliere [132] defines a Stoner shift $\Delta E$ of the band structure for the different spin channels, depending on the magnetization of the leads. Those electrons with spin parallel to the surrounding spin polarization experience a lowered effective potential, those with spin antiparallel to the polarization are raised in energy.

There are two configurations of magnetization of the two leads:

*parallel* (p) or *anti-parallel* (ap)

In each case, there are two transport channels:

*spin-up* ($\uparrow$) and *spin-down* ($\downarrow$)

The total measured conductance is the sum of the two channels:

$$G^{\text{total}} = G^{\uparrow} + G^{\downarrow}.$$  \(4.8\)

The conductances of the two channels $G^{\uparrow}$ and $G^{\downarrow}$ are equal in the presence of paramagnetic leads. The conductance of each channel depends on the sign of the Stoner shifts in both leads. We can express this dependence by the symbols $G^{++}$, $G^{+-}$, $G^{-+}$ and $G^{--}$.

In the case of *parallel magnetization* of the two leads, the Stoner shifts of each channel are in the same direction for both leads L and R:

$$\Delta E^{\uparrow}_L = \Delta E^{\uparrow}_R = + |\Delta E| \quad \Rightarrow \quad G^{\uparrow}_p = G^{++}$$
$$\Delta E^{\downarrow}_L = \Delta E^{\downarrow}_R = - |\Delta E| \quad \Rightarrow \quad G^{\downarrow}_p = G^{--}$$

Which results in a total conductance of:

$$G^p = G^{\uparrow}_p + G^{\downarrow}_p = G^{++} + G^{--}$$
In the case of antiparallel magnetization, the Stoner shifts have opposite signs for both leads in each case:

\[ \Delta E^\uparrow_L = + |\Delta E| ; \Delta E^\uparrow_R = - |\Delta E| \Rightarrow G^\uparrow_{ap} = G^{+-} \]
\[ \Delta E^\downarrow_L = - |\Delta E| ; \Delta E^\downarrow_R = + |\Delta E| \Rightarrow G^\downarrow_{ap} = G^{++} \]

Which again results in a total conductance of:

\[ G_{ap} = G^\uparrow_{ap} + G^\downarrow_{ap} = G^{+-} + G^{++} \]

We consider a symmetric system, so both partial conductances are equal in the antiparallel case: \( G^{+-} = G^{--} \).

The naming of the various kinds of magnetoresistive effects reflects the historical development starting with the anisotropic magnetoresistance (AMR) later also known as ordinary magnetoresistance (OMR) that was known since the 19th century, continuing with the tunneling magnetoresistance (TMR) which was already a much larger effect, then the giant magnetoresistance (GMR) that was even larger finally the colossal magnetoresistance (CMR). Though the effect that is studied here is closely related to the ballistic magnetoresistance (BMR), we choose the symbol that is commonly used for the whole family of magnetoresistive effects in ferromagnetic materials (XMR).

The typical value that is measured as magnetoresistance is the normalized difference between the two magnetic configurations:

\[ \text{XMR} = \frac{G^p - G_{ap}}{G^p + G_{ap}} \]

This ratio quantifies the strength of the signal that can be measured in the conductance when the magnetization of the leads switches between parallel and antiparallel. Note that there are alternative definitions of this ratio commonly used, placing only one of the two values in the denominator: The optimistic definition \((G^p - G_{ap})/G_{ap}\) results in the highest nominal values possibly above 100%, while the pessimistic definition \((G^p - G_{ap})/G^p\) gives lower nominal values of maximal 100%. The definition used here, however, results in even lower nominal values, but it is symmetric for positive and negative sign and can be more immediately understood as signal strength, which is maximal, i.e. ±100%, if one of the two configurations completely blocks transport.

Unlike previous attempts to model the ferromagnetic leads to a CNT, we will assume the tube to be embedded in metallic leads similar to the model developed in the first part of this chapter. Further, we assume a
metal that forms a highly transparent contact, so the leads can actually be modeled as semi-infinite CNT electrodes, that experience a Stoner shift due to the surrounding magnetic polarization.

**Figure 4.20.** The various partial conductances and the XMR ratio of an ordered (4,4) CNT embedded at both ends in a ferromagnetic metal contact. The distance between the metal contacts is 100 nm. The individual partial conductances can be understood by the model calculation in Sec. 3.4.3. Near the CNP, the XMR ratio is very low.

### 4.3.2. Magnetoresistance of an ordered nanotube

The results in Fig. 4.20 demonstrate the extraction of the XMR-ratio from the partial conductances. The very homogeneous model results in near perfect conductance, except for the band mismatch due to the Stoner-shift. Near the band edges, this causes a considerable difference in the conductance depending of the magnetization of the leads.

### 4.3.3. Effects of disorder

To test the robustness of the XMR signal, we applied Anderson model disorder as it will be described in detail in Sec. 5.1. As can be seen in Fig. 4.21, the reduction of the signal is significant: Especially at the band edges, where the disorder-free signal was strongest, the individual conductances are affected most by the disorder.
Chapter 4. Electrical contacts to nanotubes and -ribbons

Figure 4.21.: The equivalent data as that of Fig. 4.20, this time under the influence of Anderson disorder $\sigma \approx 0.1 \text{ eV}$ in the region between the metal coatings. As can be seen, the disorder significantly reduces the XMR signal.

Figure 4.22.: Direct comparison of the data presented in Fig. 4.20 and Fig. 4.21: In the regions around $E = \pm 1.0 \text{ eV}$, the conductance in the antiparallel magnetization is enhanced by the disorder. Transmission between certain conductance channels in the leads, that was strictly prohibited by symmetry, is allowed when the symmetry is broken by disorder.
4.3. Ferromagnetic contacts and spin transport

The direct comparison of the data from the clean system with that from the disordered system in Fig. 4.22 reveals the reason for some of the loss of XMR signal: in the clean system, the conductance is strictly limited to channels that have the same angular momentum in both leads. In the antiparallel configuration, this limits the conductance below the possible value of each section of the system individually. With a small amount of disorder, conductance electrons can scatter between states of different symmetry, raising the conductance.
Chapter 4. Electrical contacts to nanotubes and ribbons
Chapter 5.

Disorder and defects

The theoretical models for transport calculations as described in Chap. 3 are highly idealized descriptions of physical reality which is never as clean and perfect as theorists like to see it. Producing experimental samples of nanoscale structures is an extremely delicate process and even though immense progress is being made in all steps necessary to create well-defined nanostructures, defects and disorder are and will always be an important aspect in nanoscale physics.

In the theory of transport in disordered systems, dimensionality plays a major role. Single-wall carbon nanotubes (SWCNTs) come near to an ideal one-dimensional system with only two conductance channels at the charge neutrality point (CNP) and a rigid crystal structure often free of defects over distances several orders of magnitude above their lateral size. Soon after SWCNTs became available in sufficient quality, it was found experimentally, that coherent transport is indeed possible over distances of several hundred nanometers [254]. Only a little while later, it was found that even multiwall CNTs (MWCNTs) have the capability of conducting ballistically without heat dissipation over several micrometers [89]. Recent experiments have even successfully demonstrated that the length dependence of the conductance in CNTs may be measured directly by scanning along the tube with an AFM tip, finding that the decrease of the conductance with increasing length was mostly caused by point defects [245].

From the theory side, similar efforts were made to understand the effect of disorder on the transport properties of CNTs. in 1998, C. T. White and T. N. Todorov analytically predicted exceptionally long elastic mean free paths in CNTs [276], in agreement with the results of numerical work done at the same time by M. P. Ananthram and T. R. Govindan [15]. More recently these studies were completed by the analysis of the elastic mean free path in the whole energy range by F. Trizion et al. [262] that allowed to understand the influence of a shift of the Fermi energy in doped
systems. Among the latest contributions to this area of research were those studying the effect of strong localization in CNTs by B. Biel et al. [33] and finally the success of linking the localized with the diffusive regime in CNTs by R. Avriller et al. [24], based on random-matrix theory [27].

The research on defects and disorder in GNRs is comparably young. Some of the knowledge from CNTs can be transferred directly, but the special situation at the edges offers a whole class of new questions that is just starting to be explored [162, 103, 168]. With the first experiments on transport in GNRs only just becoming public [51], this area of research is sure to gain much interest in the near future.

In the following chapter, we will review several theoretical predictions about transport in disordered CNTs and combine them with analytical results found in the literature. We will demonstrate that the numerical results for the elastic mean free path can indeed be described exactly by an analytical expression generalized to arbitrary energies. The localization length is also computed for arbitrary energies and is found to match well with a simple analytical expression except for the regions around the band edges.

In a study of point defects, considering single vacancies as the most basic example, we find that from the transmission of a single defect and the defect density, both the elastic mean free path and the localization length can be obtained with good accuracy.

Many theoretical results from CNTs can be directly applied to graphene nanoribbons (GNRs) in the same way. For point defects, however, the lateral position becomes relevant as will be demonstrated in several exemplary cases.

The model to be used for the numerical evaluation of the length scales is a finite, disordered CNT/GNR section of length $L$ contacted by semi-infinite electrodes as introduced in Sec. 4.1.1 (see Fig. 5.1). In the limiting case $L \rightarrow 0$, the transmission of this system obviously has to reach that of the perfectly periodic system: $T(E) \rightarrow N_{\text{ch}}$.

**Figure 5.1.** The model generally used for the study of disordered carbon nanotubes: A infinite ideal nanotube with a finite section of length $L$ modified by disorder or defects.
5.1. Anderson model for disorder

A commonly used model in the theoretical investigation of disorder effects is that of Anderson disorder: In a lattice model corresponding to the tight-binding models used in this work, the on-site energy of the individual atoms $\epsilon_i$ is modified randomly for each atom independently. The disorder can be expressed as a perturbative contribution to the total Hamiltonian $\mathcal{H} = \mathcal{H}_0 + W$ with the following properties:

- $P(W)$ is a probability measure: $\int dW P(W) = 1$
- $W$ is diagonal in an atomic basis: $\langle ij| W | j \rangle = \delta_{ij} W_i$
- each $W_i$ is centered around zero: $\int dW P(W) W_i = 0$
- the $W_i$ are uncorrelated: $\int dW P(W) W_i W_j = \delta_{ij} \sigma_i^2$

In his original work [18], Anderson chose each value $W_i$ to follow a uniform distribution of width $W$, leading to a standard deviation of $\sigma_\epsilon = W/\sqrt{12}$. Within the perturbative regime, however, the exact shape of the distribution does not have any influence on the physical results.

A slight variation of this model can be constructed by varying the nearest neighbor hopping $\gamma$ instead of the on-site energy. The physical effect of such disorder with the standard deviation $\sigma_\gamma$ can, however, be related to that of regular Anderson disorder via renormalization.

In Fig. 5.2 one can see the effect of Anderson disorder on the transmission of a CNT. Visible already from one sample, but much clearer from the sample average is the effect of the increased suppression of transmission at the band edges. The shape of the transmission curve can be understood to a certain degree perturbatively by the calculation of the elastic mean free path discussed in the following section.

The decay of the transmission with increasing length $L$ of the sample, which is already observable in Fig. 5.2 is displayed in detail in Fig. 5.3 for three different energies. For short samples, the system is in the diffusive regime, where the transmission follows the form:

$$T_{\text{diff}}(L) = N_\text{ch} \frac{\ell_{\text{el}}}{\ell_{\text{el}} + L}$$

with the elastic mean free path $\ell_{\text{el}}$. For longer samples, one can observe the transition to the localized regime, where the transmission decays exponentially:

$$T_{\text{loc}}(L) \propto \exp \left( -L/\ell_{\text{loc}} \right)$$
Chapter 5. Disorder and defects

Figure 5.2.: The transmission of a (5,5)-CNT with an Anderson-disordered section of length $L$ with $\sigma_\varepsilon = 0.5$ eV. Left: one specific, randomly generated sample. Right: average $\langle T \rangle$ over 1000 samples. The integer-step function is the transmission of the clean system for comparison. Note especially how the disorder is especially effective at band edges and especially ineffective around the CNP.

with the localization length $\ell_{\text{loc}}$. Both regimes and their interrelation will be subject of the following sections.

Clearly visible in Fig. 5.3 is also the effect of the different ways of sample averaging. As it turns out [19, 232, 50, 27], the transmission itself does not follow Gaussian statistics in the localized regime. The quantity $\langle T \rangle$ therefore does not converge towards any meaningful value. Instead, the proper value to average over is the logarithm of $T$. The statistical average $\langle \ln T \rangle$ converges very well and indeed gives the expected result. This fact will be used later on in extracting $\ell_{\text{loc}}$ from the numerical data.

5.2. The elastic mean free path

The elastic mean free path $\ell_{\text{el}}$ is the mean distance between scattering events of an electron in a disordered medium. In our case of a one-dimensional medium, where the electron has only two ways to go, only back-scattering events have an effect on the transmission, so we consider only these. Considering a short distance $L \ll \ell_{\text{el}}$, the probability of an electron to be back-scattered, i.e., reflected is $L/\ell_{\text{el}}$. The transmission is then simply given by:

$$T_{L \ll \ell_{\text{el}}} = N_{\text{ch}} \left( 1 - \frac{L}{\ell_{\text{el}}} \right)$$ (5.1)
5.2. The elastic mean free path

Figure 5.3.: Length dependence of the transmission in Anderson-disordered samples ($\sigma_\varepsilon = 0.5$ eV). The length $L$ is given in unit cells of $\ell_{uc} = 2.5$ Å.

Left column: transmission of a single sample. Center column: simple $\langle T \rangle$ averaging over 50 samples. Right column: logarithmic average $\langle \ln T \rangle$ over 50 samples.

Top row: diffusive regime $L < 500\ell_{uc}$ in linear scale. Center row: localized regime in logarithmic scale. Bottom row: the data converges against $\ell_{uc}$ with logarithmic sample averaging giving the lowest fluctuations.

As observed in Fig. 5.2, the disorder is most effective at the band edges and least effective at the CNP.
Chapter 5. Disorder and defects

For larger $L/\ell_{el}$, the possibility of multiple scattering has to be taken into account. As it is demonstrated by S. Datta [64] (section 2.2), the correct handling of multiple reflections, where each reflection is assumed to be statistically independent (i.e. neglecting quantum interferences between scattering events), directly leads to Ohm’s law:

$$T_{\text{Ohmic}} = N_{\text{ch}} \frac{\ell_{el}}{\ell_{el} + L}$$

(5.2)

better known in its inverse form for the resistance:

$$R_{\text{Ohmic}} = \frac{h}{2e^2N_{\text{ch}}} + \frac{h}{2e^2N_{\text{ch}}\ell_{el}}$$

(5.3)

This can be broken up into the contributions of the contact resistance $R_{\text{contact}}$ and the Ohmic resistance per length $\rho$. As we will see in the next chapter, quantum interference between scattering events will break this law at the scale of the localization length $\ell_{\text{loc}}$. However, for $L < \ell_{\text{loc}}$, i.e. in the diffusive regime, same expression holds also for quantum coherent transport:

$$T_{\text{diff}} = N_{\text{ch}} \frac{\ell_{el}}{\ell_{el} + L}$$

(5.4)

An analytic expression for the elastic mean free path in Anderson-disordered armchair CNTs at the CNP was derived by C. T. White and T. N. Todorov [276]. A more general derivation of this quantity for arbitrary energies (with pure Anderson disorder) can be found in App. B.3, resulting in Eq. (B.9):

$$\ell_{el} = \left( \frac{\ell_{uc}\sigma_{\varepsilon}^2}{16NN_{\text{ch}}} \left(2\pi g_L(E)\right)^2 \right)^{-1}$$

(5.5)

where $N$ is the chiral index of a $(N, N)$ CNT and $g_L(E)$ denotes the total DOS per length.

Numerically, the elastic mean free path is straightforward to obtain with high precision: For low $\sigma_{\varepsilon}$ and short $L$, the transmission can simply be averaged over a large number of disorder configurations. As can be seen, in Fig. 5.4, the numerical result converges nicely against Eq. (5.5).

Similar results for the elastic mean free path have been obtained from time-dependent diffusion calculations by F. Triozon et al. [262]. Another approach for studying diffusion coefficients is the evaluation of the fractal dimension of the energy spectrum as it was used on incommensurate MWCNTs [274].
5.3. Strong localization

The derivation of the Ohmic resistance in Eq. (5.3) is based upon statistically independent scattering events and breaks down when quantum interference has to be taken into account. Interference takes place in general, when one electron travels a closed path coherently. The average length of closed paths in a system is called localization length $\ell_{\text{loc}}$.

If the phase coherence length $\ell_{\phi}$ is shorter than $\ell_{\text{loc}}$, electrons lose their phase coherence before their paths can interfere. Without interference, scattering events are statistically independent and classical Ohmic behavior is observed. If however, $\ell_{\phi}$ is very large, electrons coherently move in closed paths, or in other words, electrons reside in bound states. In this case, $\ell_{\text{loc}}$ specifies the mean spatial extension of such states. For $L < \ell_{\text{loc}}$ this does not have any consequences: electrons will move all the way through the device without completing a circle, so interference is not relevant. For $L > \ell_{\text{loc}}$, however, electrons are prohibited from moving freely and instead stay trapped in bound states. This is called the regime of strong localization, where coherent transport is possible only through the exponentially decaying tails of localized states. Transmission in this regime falls off exponentially as:

$$T_{\text{loc}} \propto \exp \left( -\frac{L}{\ell_{\text{loc}}} \right).$$

(5.6)

In his original work in 1958, P. W. Anderson first described his theory about the localization of quantum mechanical states in disordered systems [18]: In three dimensions for weak disorder, an electron can move freely and is unlikely to ever again come back to its starting point. The system is said to be in the delocalized state. Above a certain critical strength of the disorder, however, all electrons are trapped in localized states and...
diffusion is completely suppressed. Today, this effect is called Anderson transition.

Only a short while later, N. F. Mott and W. D. Twose suggested [183], that in one-dimensional systems, this effect of localization should already appear at the presence of arbitrarily weak disorder: As electrons can only move forward or backward, any back-scattering probability will certainly bring the electron back to its origin.

In a strictly one-dimensional system, $\ell_{\text{loc}}$ is equal to $\ell_{\text{el}}$ because any electron is bound to return to its origin right after the first scattering event. In a quasi-one-dimensional system with $N_{\text{ch}}$ channels in each direction that the electron can reach with equal probability at each scattering event, an electron can travel back and forth several times before it ends up at the same place and in the same channel. With this argument, D. J. Thouless gave an estimate of $\ell_{\text{loc}} \sim N_{\text{ch}} \ell_{\text{el}}$ based on a free electron gas in a finite-width quantum wire [257, 258]. More recently, C. W. J. Beenakker refined this expression using random matrix theory and obtained [27, 24]:

$$\ell_{\text{loc}} = \beta \frac{(N_{\text{ch}} - 1) + 2}{2} \ell_{\text{el}}$$

(5.7)

where $\beta = 1$ for time-reversal-invariant systems (i.e. for our purpose, systems without magnetic field) and $\beta = 2$ otherwise.

To obtain the localization length from an ensemble of disorder configurations, it is necessary to understand the statistical behavior. As it turns out [232, 50], the only quantity that provides a meaningful thermodynamic limit is the mean value $\langle \ln T \rangle$. A proper definition of the localization length of a configuration ensemble is therefore:

$$\ell_{\text{loc}} = -\frac{L}{\langle \ln T \rangle}$$

(5.8)

Indeed this expression holds true not only for $L \to \infty$, but shows good convergence even for finite $L \gg \ell_{\text{loc}}$. In Fig. 5.5, one can see the computed value of the localization length $\ell_{\text{loc}}$ for one specific nanotube with Anderson disorder of fixed strength.

A common alternative way to obtain the localization length numerically is via iterative transfer matrix methods [212, 226, 190], determining the Lyapunov exponents. In our case of carbon nanotubes, however, where the hopping matrix between successive unit cells is not necessarily invertible, the decimation method described in App. A.3 gives better results. Furthermore, it turns out that the transmission through a finite scattering region between two homogeneous leads, as given by Eq. (3.5), actually gives better numerical stability than the raw Green function.
5.4. Vacancies and defects

The crossover between the diffusive regime described by Eq. (5.3) and the localized regime described by an exponentially suppressed transmission in Eq. (5.6) was first studied by A. V. Tartakovski [255] and later put into a closed expression for the resistance at arbitrary $L$, ranging from $L \ll \ell_{\text{loc}}$ all the way to $L \gg \ell_{\text{loc}}$, by T. N. Todorov [260]:

\[
R = \frac{h}{2e^2N_{\text{ch}}} \left( \frac{\ell_{\text{loc}}}{\ell_{\text{el}}} \sinh \frac{L}{\ell_{\text{loc}}} + \cosh \frac{L}{\ell_{\text{loc}}} \right)
\]

### Figure 5.5.

The localization length $\ell_{\text{loc}}$ computed for a (5,5)-CNT with Anderson disorder of strength $\sigma_{\varepsilon} = 0.5\ \text{eV}$ as average $\ell_{\text{loc}} = -L/\langle \ln T \rangle$ over 100 samples of length $L = 20N_{\text{ch}}\ell_{\text{el}}$. The dashed line follows the analytic value from Eq. (5.7). In the bottom panel, the ratio between both values is plotted, which deviates from 1.0 around the band edges, where the van Hove singularities in the DOS cause discontinuities in the analytical form of $\ell_{\text{el}}$.

5.4. Vacancies and defects

From the experimental point of view, point defects in CNTs are easier to study than homogeneously distributed disorder. While homogeneous disorder is in principle omnipresent to some degree due to environmental influences and intrinsic properties like vibrations and thermal fluctuations, it is fairly hard to control with precision. Individual point defects, on the other hand, can be defined with chemical precision and artificially produced with a controlled density. Thus, strong localization could be measured in dependence of the defect density caused by ion irradiation [96] and a single tunable point defect induced by an AFM tip could be used to control the transmission of a metallic CNT [206].

The theoretical handling of point defects demands somewhat more effort, as the result depends strongly on the detailed structure of a single defect. Point vacancies caused by removing a single atom from the tight binding network can easily be handled and were studied already back in 1996 by L. Chico et al. [52]. So-called Stone-Wales defects, caused by two
neighboring atoms that are rotated by 90° around their common center, are similarly straightforward to implement by changing the topology of the network \[173\]. However, these simple models are far from physical reality, as a defect will generally have a heavy influence on the electronic structure of its environment that needs to be handled with ab initio methods. This has been done with vacancies, Stone-Wales defects and substitutions by various doping atoms \[53, 136, 158, 33, 6\].

Picking one sample of possible point defects, we will here only consider single vacancies modeled by removing one atom from the tight-binding network. Mathematically, this “removing” is identical to setting its on-site energy \(\varepsilon_i\) to infinity, or numerically, a very large value, in our case \(10^5\) eV.

The transmission of such a system of a carbon nanotube with a single atom removed shows characteristic reduction to a single channel at the CNP.

![Figure 5.6](image)

**Figure 5.6:** Similar to the calculation first presented by L. Chico et al. \[52\], this figure shows the effect of a single point defect (i.e. a single missing atom) on the transmission of a (5,5)-CNT.

The characteristic shape of the single-point-defect transmission has a strong influence on the transmission of an extended region containing a random distribution of point defects (Fig. 5.7). The transmission of a single configuration of defect positions shows a nearly complete suppression of transmission at the CNP. A sample average over many configurations shows the emergence of a very different behavior than the one found for Anderson disorder: In contrast to a reduction of transmission near band edges, the point defects cause stronger scattering at energies with low DOS.

The elastic mean free path \(\ell_{el}\) can in this case be obtained from the transmission through a single defect \(T_{def}(E)\) as displayed in Fig. 5.6. The probability that an electron is reflected by a single defect is \(1 - T_{def}/T_0\), so if the defect density \(\rho_{def}\) is low enough that individual defects are unlikely
to occur in the same unit cell, $\ell_{el}$ can be expressed as:

$$\ell_{el} = \frac{\ell_{uc}}{(1 - T_{def}/T_0) \rho_{def} N_{atom}} \quad (5.9)$$

with the number of atoms per unit cell $N_{atom}$ and the length of a unit cell $\ell_{uc}$.

Fig. 5.7 demonstrates that the diffusive transmission in Eq. (5.4) based on this value for $\ell_{el}$ gives an excellent approximation of the numerical result.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.7.png}
\caption{The transmission through a finite length (5,5) CNT with a defect density of 1%: Left: one random sample (i.e. the positions of the defects are chosen once and fixed for the whole computation) Right: average $\exp\langle \ln T \rangle$ over 1000 samples. The dotted lines indicate the diffusive transmission based on the elastic mean free path obtained from Eq. (5.9) for comparison. The transmission suppression at the CNP was first demonstrated by M. P. Anantram [15].}
\end{figure}

Equally, the localization length in a system with randomly distributed vacancies follows fairly well the value given in Eq. (5.7) when the elastic mean free path is obtained from a single vacancy, as can be seen in Fig. 5.8. At the band edges, the transmission for a single vacancy touches the value of the clean sample, so the estimate for the elastic mean free path diverges and the localization length deviates strongly from the perturbative estimate.

A study of the localization length of nanotubes with realistically modeled point defects was presented by B. Biel et al. [33]. In contrast to the effects of our model defects, the conclusion of this study was that monovacancies give a very small contribution to resistance and only divacancies produce a significant localization. It is, however, not surprising that different models of point defects produce very different results.
Chapter 5. Disorder and defects

5.5. Graphene nanoribbons

For GNRs, the effect of homogeneous disorder is conceptually very similar as in CNTs: the elastic mean free path is given by the DOS and number of channels via Eq. (5.5), the localization length by Eq. (5.7). The transmission through short, weakly disordered samples is well described by Eq. (5.4), except for the band edges. The major difference between zigzag GNRs and armchair CNTs however, is the presence of zigzag edge states that produce a huge DOS right at the CNP. Due to this peak, the same principles that make the conductance of metallic CNTs very insensitive to disorder have exactly the opposite effect in metallic GNRs, as can be seen in Fig. 5.9.

The issue of point defects in GNRs however, is very different from that in CNTs. While in CNTs, due to their high symmetry, all atoms are basically equivalent, there is a distinction in GNRs not only between edge

---

**Figure 5.8.** The localization length $\ell_{\text{loc}}$ computed for a (5,5)-CNT with randomly distributed vacancy of density 1% as average over 100 samples of length $L = 20N_{\text{ch}}\ell_{\text{el}}$ with $\ell_{\text{el}}$ obtained from Eq. (5.9). The dashed line follows the analytic value from Eq. (5.7). In the bottom panel, the ratio between both values is plotted, which deviates from 1.0 around the band edges, where $\lambda_{\text{el}}$ diverges.

**Figure 5.9.** Transmission of a $N_z = 10$ zigzag GNR with Anderson disorder ($\sigma_\epsilon = 0.1$ eV) over a finite length. The huge DOS at the CNP, caused by the zigzag edge state, increases scattering and suppresses transmission. The dashed lines are the diffusive transmission obtained from the analytical value for $\ell_{\text{el}}$. 

The issue of point defects in GNRs however, is very different from that in CNTs. While in CNTs, due to their high symmetry, all atoms are basically equivalent, there is a distinction in GNRs not only between edge
 states, but also between the two sublattices. Fig. 5.10 shows two example cases for the transmission drop caused by a single defect either at the edge or at the center.

**Figure 5.10.**: Transmission of a $N_z = 10$ zigzag GNR with a single vacancy at two different locations. Removing an edge atom has little effect at the CNP but brings down the transmission at the band edges. Removing an atom from the center of the ribbon drastically reduces transmission at the CNP. (The peak at the CNP originates from the edge state and does not contribute to the conductance due to its low dispersion.)

However, when computing the transmission sample average through a finite length ribbon with fixed defect density, the situation turns out to be even more complicated. As displayed in Fig. 5.11, the computed transmission is much lower than the diffusive transmission in a broad region around the CNP. And what is even more: the transmission is very different depending on the density of defects for the same total number of defects. Obviously interference between defects is very strong for an average distance of five unit cells between individual vacancies. The effect is especially strong for edge vacancies: a single edge vacancy has little effect at the CNP but already a few vacancies taken together can bring down the transmission considerably. For defects at the center of the ribbon, the effect of interferences is less clear but still visible.

For a deeper understanding of these interference effects, the interplay of individual defects at specific relative positions in the ribbon will need to be investigated in more detail.
Figure 5.11.: Randomly distributed vacancies in a $N_z = 10$ zigzag GNR. The defect density $\rho_{\text{def}}$ corresponds to the total percentage of missing atoms, so for 20 atoms per unit cell and $\rho_{\text{def}} = 1\%$, there is on average one defect every five unit cells. The edge defects are aligned along one edge of the ribbon. The center defects are distributed over a range of $\pm 3$ atoms from the true center. The transmission average is taken as $\langle T \rangle$ over $> 1000$ samples. The dashed lines correspond to the diffusive transmission based on $\ell_{\text{el}}$ obtained from a single defect using Eq. (5.9).
Chapter 6.

Multilayer graphene and carbon nanotubes

Soon after the first success in isolating graphene via exfoliation from graphite, the focus of research included not only monolayers but also bi-, or more generally, few-layered systems. While the peculiar electronic structure of monolayer graphene already gives rise to many unique physical effects, the additional degrees of freedom of multilayers initiated yet another flurry of discoveries and theoretical predictions. Most prominent among them is the anomalous quantum Hall effect that differs fundamentally from either that in regular semiconductors or that in monolayer graphene [195] or an electrically tunable gap that holds the immediate promise of valuable applications [197, 174].

Likewise, double-wall carbon nanotubes (DWCNTs) show many features that go beyond the physics of independent SWCNTs and form an intermediate step between these—which show mostly one-dimensional physics—and multiwall carbon nanotubes (MWCNTs) with physical properties similar to bulk graphite.

For MWCNT, it has been a long standing debate whether electron transport would be ballistic or diffusive. Generally it is believed that MWCNTs are more likely to show regular disorder and defects than SWCNTs due to the larger surface and the various interactions between the layers. This would cause conventional diffusion as that described in Chap. 5 and, together with dephasing effects due to interactions between electrons, explain the Ohmic resistance found within and between the shells of MWCNTs [39]. However, ballistic transport seems to be possible as well [42, 157], so there is need for further investigations.

Besides the effect of explicit disorder and defects that may be reduced by refined production techniques, theorists have been curious about the possibility of intrinsic disorder caused by the very structure of DWCNTs themselves. S. Roche et al. have shown numerically that incommen-
surate MWCNTs may exhibit diffusive intra- and intershell transport in absence of disorder [224], even though momentum conservation is supposed to suppress intrashell tunneling [279, 273]. The analysis of the energy spectrum even points to an anomalous diffusion in the incommensurate CNTs [274].

In this chapter, we will study various aspects of commensurate and incommensurate bilayer carbon systems including DWCNTs and graphene bilayers. Special attention will be given to the modeling of the interlayer coupling within the tight-binding approach as this modeling turns out to be an important factor for the varying results on incommensurate DWCNTs. Both angular and longitudinal momentum are not strictly preserved in DWCNTs as the two underlying symmetries do not hold in the combined system. Still, as long as the interwall coupling does not give rise to abrupt fluctuations, the momenta are preserved to a high degree and scattering between states of different momentum is strongly suppressed. As it turns out, the presence of such fluctuations is not an intrinsic feature of the incommensurability, but rather an effect of details in the model for the interwall coupling.

In the remainder of the chapter, we will apply the results about the interlayer coupling in commensurate DWCNTs on telescopic CNTs, which are one of the most promising ideas for a future application of DWCNTs, allowing a precisely controlled coupling between electronic and mechanical degrees of freedom at the nanoscale.

6.1. Commensurability

The concept of commensurability goes back to the Euclid in 300 B.C. describing it in his book Elements [79] as the property of two lengths $a$ and $b$ to be measurable as integer multiples of the same rod $c$, i.e. $a = nc$ and $b = mc$ where $m, n \in \mathbb{N}$. In modern terms, this condition is equivalent with their ratio being a rational number: $a/b \in \mathbb{Q}$.

For MWCNTs, this concept generally refers to the periodic lengths $\ell_{uc}$ of the individual shells [given by Eq. (C.1)]. Two CNTs of chiralities $(M, N)$ and $(M', N')$ are commensurate if

$$\frac{\sqrt{M^2 + N^2 + MN}}{\sqrt{M'^2 + N'^2 + M'N'}} \in \mathbb{Q}$$

or incommensurate otherwise.

The simplest case are achiral CNTs: armchair CNTs have a periodic length $\ell_{uc} = \sqrt{3}d_{CC}$, for zigzag CNTs, it is $\ell_{uc} = 3d_{CC}$. A DWCNT made
6.1. Commensurability

up of two armchair CNTs only, like (5, 5) @ (10, 10) or of two zigzag CNTs, like (9, 0) @ (18, 0) is commensurate. A combination of both types, like (7, 0) @ (9, 9) or (5, 5) @ (17, 0) leads to a irrational ratio of \( \sqrt{3} \) between the two periodic lengths. In the first case, the combined system still is a quasi-1D crystal with the same periodic length as that of a single wall, in the latter case, the combined system does not have any periodicity, similar to quasi-crystals in two or three dimensions [71].

When chiral CNTs are involved in building a DWCNT, most combinations lead to incommensurate systems, but many nontrivial commensurate combinations can also be found, like (11, 2) @ (12, 12) or (8, 8) @ (22, 1). A detailed theoretical analysis of the possible combinations can be found in Ref. [63]. Experimentally, it has been found, that DWCNTs produced by arc discharge show no correlation in the chiralities of the walls [112, 117], so both commensurate and incommensurate combinations are equally possible.

In graphite, the issue of commensurability is closely related, in this case being determined by the angle between the layers. The energetically most favorable configuration of graphene bilayers is Bernal stacking [23]. Epitaxially grown bilayers, however, often feature a rotational angle between the layers [113] and even in the top layers of bulk graphite, Moiré patterns have been observed and were attributed to a rotational angle between the layers [227]. In general, such rotations result in incommensurate periodicities, yet, it is possible to find an arbitrary number of angles that result in periodic superstructures following a simple rule:

The choice of any graphene lattice vector \((M, N)\) defines an axis in the graphene plane crossing the center of a plaquette chosen as origin. Mirroring the sheet at this axis is equivalent to a rotation by an angle:

\[
\alpha = 2 \arctan \left( \frac{\sqrt{3}N}{2M + N} \right),
\]

with the guarantee that the lattice vector \((M, N)\) remains a periodic vector for the rotated sheet, along with its perpendicular counterpart \((M + 2N, -2M - N)\). With the set of rational numbers lying dense in the space of real numbers, this allows to approximate any rotational angle by one that produces a commensurate periodicity. Fig. 6.1 shows the structure obtained from mirroring at the (2,1) axis—or, equivalently, the (4,1) axis—resulting in a rotation by \( \alpha = 2 \arctan(\sqrt{3}/9) \approx 21.78^\circ \), which gives the smallest non-trivial supercell in a rotated bilayer system.

Mathematically, there is a very clear distinction between commensurate and incommensurate systems. The physical meaning of this distinction,
Figure 6.1.: Sketch of a graphene bilayer with one layer rotated by $2 \arctan(\sqrt{3}/9) \approx 21.78^\circ$ against the other. This is the simplest non-trivial case of a rotation that results in a finite super cell. In this case, the super cell contains a total of 28 atoms.

However, is less clear-cut. A theoretical periodicity that is larger than the system itself is meaningless. Already for a periodic length larger than the elastic mean free path or the coherence length, one expects to see similar physics as in a mathematically aperiodic system. For quasiperiodic systems, however, that consist of weakly linked parts of different periodicity, the underlying periodicities may still dominate the observed physics.

6.2. Modeling the interlayer coupling

In Sec. 2.1.2 and 2.1.3, we introduced two options for the tight-binding parameterization of the Hamiltonian matrix-elements between different layers. As it turns out, contradicting conclusions about the physics of multilayer systems in various theoretical works can often be traced back to differences in details of the modeling of this interlayer coupling. It is, therefore, crucial to take a closer look at these modelings that have been passed on from publication to publication.

Full ab initio calculations of the electronic properties of double-wall CNTs have been done [48, 198, 97, 277, 32, 290], but these are generally very limited in the number of atoms. For the efficient handling of systems with large diameter or large periodic length, and especially for the investigation of quasiperiodic systems, tight binding parameterizations are necessary.

For SWCNTs, the first-nearest-neighbor tight-binding approximation is a simple and obvious model that gives reasonable quantitative results and allows the qualitative investigation of various non-trivial physical
effects. For the interwall coupling of MWCNTs, such an obvious choice does not exist and a large set of different tight-binding parameterizations can be found in the literature. Even similar parameterizations often show variations in details, making it difficult to compare theoretical results with each other and even more difficult to make reliable predictions about experiments.

First theoretical calculations of DWCNTs were restricted on the highly symmetric (5,5)@(10,10) and (9,0)@(18,0) examples [229, 153], allowing the definition of a parameterization based on a small number of fixed parameters not transferable to other systems. A better transferability can be achieved by parameterizations that describe a functional dependence of the hopping integral on the relative geometry between the layers.

One common class of tight-binding parameterizations is based on a model originally developed for graphite by J.-C. Charlier et al. [49]. This parameterization was then modified step-by-step in various successive publications, often without even clearly indicating the differences. While early publications based on this parameterization considered different types of interlayer bonds depending on the atomic coordination number [151, 152], this detail is later often silently dropped [224, 262]. Furthermore, many works fail to document the length cutoff that was used, even though this is a crucial factor for the intershell conductance in disorder-free incommensurate DWCNTs as, we will see later on.

In our work, we use the intershell coupling as given in Eq. (2.1), following the orthogonal $\pi$ orbital model by S. Roche et al. [224], setting the undocumented cutoff length to $r_{\text{cutoff}} = a + 5\delta$. The details of this parameterization are described in Sec. 2.1.2.

For a comparison, we will also use the full non-orthogonal Slater-Koster parameterization including the 2s and 2p orbitals developed by D. A. Papaconstantopoulos and described in Sec. 2.1.3.

From ab initio calculations it is known that the work function of SWCNTs depends on their diameter [240, 239]. For thin MWCNTs, this is the cause for charge transfer between the shells that may close the gap of otherwise semiconducting tubes [198, 290]. In the tight-binding model used here, this charge transfer could be implemented by a shift of the on-site energy in each wall. For metallic CNTs, however, this effect is small, and for the issues that we are about to study, it can be neglected.
6.3. Bilayer graphene

The band structure of a graphene bilayer in Bernal-stacking (see Fig. 1.10) is displayed in Fig. 6.2. Comparing it to the bands of a graphene monolayer (see Fig. 2.4), one can identify several characteristic effects caused by the interlayer coupling: The $\sigma$ bands show splits in several regions, especially the antibonding bands at high energies. This, however, has little physical relevance as it is far away from the electronically active energy range around the Fermi energy.

![Band structure of a graphene bilayer (Bernal stacking)](image)

**Figure 6.2.** Band structure of a graphene bilayer (Bernal stacking), comparing the orthogonal $\pi$ orbital parameterization with the full non-orthogonal $sp^3$ Slater-Koster parameterization. The full $sp^3$ parameterization shows the additional $\sigma$ bands at low and high energies. The $\pi$ bands in both parameterizations show a distinct asymmetry in the energy range that is caused by the interlayer coupling.

More important are effects in the $\pi$ bands that are reproduced similarly for both parameterizations: The electron-hole symmetry (the mirror symmetry in the energy scale around $E_F$) which is exact for a single graphene layer the simplest parameterization (see Sec. 2.2) and still approximately preserved for the more detailed parameterization is broken by the interlayer coupling: For negative energies at the bottom of the $\pi$ band, the bands are split up maximally, while the $\pi$ band at positive energies shows only a negligible split. This asymmetry can be understood by considering the wave functions within a graphene monolayer (see Fig. 6.3).

At the large energy scale of the full band structure in Fig. 6.2, the exact stacking of the bilayer has no visible influence. Zooming in on the K-points, however, shows a distinctive difference between Bernal and aligned stacking (see Fig. 6.4): In Bernal stacking the cones are deformed into four parabolic bands, two of which touch a the Fermi energy (see also [197]). In the aligned graphene bilayer, on the other hand, the hybrid bands retain their original double-cone shape, now with two copies split by a rigid shift of $\pm 0.2$ eV.
6.3. Bilayer graphene

Figure 6.3.: Sketch of three characteristic electronic wave functions in a single layer of graphene:
At $E_{\text{min}}$, the phase is constant over all atoms. In a bilayer, these states can couple most efficiently between the layers.
At $E_{\text{max}}$, the wave function has opposite signs for neighboring atoms. The interlayer coupling sums over several neighboring atoms, canceling out very effectively.
At $E_{\text{F}}$, the system can be separated into two sublattices. Within one sublattice, three of the phases at the atoms around one plaquette cancel out exactly to zero. The effect of the coupling now depends on the exact stacking of the two layers (see Fig. 6.4).

Figure 6.4.: Zoom into the bands around the $K$-point in Bernal-stacked and aligned graphene bilayers. In both cases, the bands are approximately rotationally symmetric around the $K$-point.
For the rotated bilayer depicted in Fig. 6.1, the full band structure cannot easily be visualized due to large unit cell and the resulting folding of the Brillouin zone. Most relevant, however, is the spectrum near the Fermi energy, where the familiar cone structure is well-preserved. Fig. 6.5 displays a cut through one of these cones, featuring a split by only \(~\sim\) 1 meV, as compared to the \(~\sim\) 0.3 eV in the non-rotated bilayer (Fig. 6.4).

Figure 6.5.: The bands of the rotated graphene bilayer (displayed in Fig. 6.1) near Fermi energy: The cone of graphene monolayer bands is preserved with minimal hybridization, about two orders of magnitude smaller than in non-rotated graphene (Fig. 6.4). The rotational symmetry around the \(K\) point is preserved.

The reason for this difference by two orders of magnitude is momentum conservation. Even though the combined system has a reduced symmetry, the high symmetry of the original system is broken only by the interlayer coupling, which can be viewed as a small perturbation. Furthermore, the interlayer coupling is defined with a fairly long cutoff, resulting in a smooth distance dependence, conserving the momentum to a large degree. The Brillouin zones of the two layers, however, are rotated against each other, so the \(K\) points of the two unperturbed systems do not match. Even though the cones at the Fermi energy appear at the same point in the folded Brillouin zone for both layers (Fig. 6.5), they actually carry a different momentum in each of the two unperturbed systems. Thus, the smooth interlayer coupling cannot efficiently mix these states and the two cones of the two sheets are split only by a very small amount.

A slightly different approach to this issue was presented by S. Urqu [266] showing a suppression of the effective matrix elements in dependence of the rotational angle.

6.4. Double-wall carbon nanotubes

One long-standing issue in understanding charge transport in MWCNTs is the question whether the inner shells make any significant contribu-
6.4. Double-wall carbon nanotubes

It was shown that MWCNTs are capable of ballistic transport over long ranges in a single conductance channel \([89, 216]\). Furthermore, there have been clear experimental indications that only the outermost shell contributes to transport \([26, 247]\). Other measurements showed conductance proportional to the cross-section of the MWCNT \([118]\) or Ohmic behavior of MWCNTs that could only be explained by a model including significant conduction between the shells \([39]\).

Theoretical studies predict a considerable interwall conductance for pure armchair MWCNTs \([109]\). For other cases, it is generally prohibited by symmetry in other cases \([279, 273, 267, 277]\). However, while there is general agreement, that disorder will break those symmetries and therefore greatly enhance interwall conductance \([224, 266, 264]\), the discussion whether the incommensurability of nanotubes may itself be viewed as disorder is still ongoing \([8, 274]\).

In the following, we will take a closer look at several DWCNTs to understand the effect of incommensurability in the interwall coupling.

### 6.4.1. Commensurate double-wall tubes

The extreme case of commensurate DWCNTs are pairs of armchair and zigzag CNTs retaining their short periodicity. For such a system, the band structure can be computed in a straightforward way. When setting up the periodic Hamiltonian, special care has to be taken to make sure that no matrix elements are dropped by restricting the coupling to neighboring unit cells. Depending on the length cutoff of the interwall parameterization, more than nearest neighbor cells have to be considered in the periodic block matrix. A sample implementation of this can be found in App. C.5.

The band structure of an armchair DWCNT is displayed in Fig. 6.6. The breaking of the electron-hole symmetry corresponds directly to the case of bilayer graphene as shown in Fig. 6.2. In contrast to the rotationally symmetric cone at the \(K\) point in bilayer graphene, however, the bands at the Fermi energy show a strong asymmetry: while the bands with negative slope (near \(ka = +2\pi/3\)) split by an amount of 0.25 eV, the others remain nearly degenerate. An explanation for this difference can again be found by looking at the wave functions in two independent walls: While the former bands have constant phase for all atoms around the circumference of the tube, the latter bands have an alternating sign for each
Chapter 6. Multilayer graphene and carbon nanotubes

Figure 6.6.: Band structure, DOS and energy resolved number of channels of a (4,4)@(9,9) armchair DWCNT. As a reference, the bands of the two decoupled shells is displayed on the right-hand side. Most prominent feature caused by the intershell coupling is the breaking of the electron-hole symmetry: While the states at negative energies hybridize strongly, those at positive energies are mostly preserved.

pair of atoms, so the states of different diameter tubes cannot hybridize effectively.

Early theoretical studies of DWCNTs overlooked this asymmetry due to a over-simplified model of the interwall coupling [229]. For more refined models, however, the asymmetry is a very robust feature [153, 150, 266].

Figure 6.7.: Band structure, DOS and energy resolved number of channels of a (6,0)@(15,0) zigzag DWCNT. Unlike the case of the armchair DWCNT (Fig. 6.6), the hybridization at the Fermi level is minimal, due to mismatching angular momentum of the wave functions.

The situation for a zigzag DWCNT is very different (Fig. 6.7): Though the electron-hole asymmetry can equally be found between energy minimum and maximum, the band structure remains mostly unperturbed near the Fermi energy. The hybridization of the massless bands is one order of magnitude smaller than in the armchair case. The explanation can again be found by looking at the wave functions themselves: The SWCNT bands at the Fermi energy are associated via zone-folding (see Sec. 2.3) with graphene states on a line crossing a K point. The discrete angular momentum of these states depends on the tube circumference,
so the bands in the different shells generally have a different rotational symmetry and cannot hybridize.

Note that this protection by symmetry need not be mathematically exact: In general, the combined system does not share the rotational symmetry of the individual shells. Still, as long as the interwall coupling is a weak perturbation and is smooth enough, the original symmetry will be preserved and the hybridization will be strongly—though not completely—suppressed.

6.4.2. Incommensurate double-wall tubes

For incommensurate DWCNTs, and equally for DWCNTs that are mathematically commensurate but have a large common unit cell, the effect of “approximate symmetries” is relevant not only in angular but also in longitudinal direction. Explicit disorder potentials that break the symmetry generally allow the coupling of arbitrary states and thereby greatly increase intershell coupling [224, 266, 264]. Several independent theoretical studies, however, indicated, that even MWCNTs free of explicit disorder show signs like diffusive transport and intershell conductance that are interpreted as effects of implicit disorder induced by incommensurability itself [224, 225, 8, 274]. To elucidate this finding, we computed the localization length in incommensurate DWCNTs based on the limiting expression:

\[
\ell_{\text{loc}} = \lim_{L \to \infty} \ell_{\text{loc}}^{(L)} = -L / \ln \left( |G_{1,1}(E)|^2 \right)
\]

where \(G_{1,1}(E)\) is nonlocal Green function connecting both ends of a finite DWCNT of length \(L\). This common definition of the localization length is equivalent to Eq. (5.8) but avoids the need for defining contacts and leads. In Fig. 6.8 the convergence process \(\ell_{\text{loc}}^{(L)}\) is demonstrated for three nearly identical systems based on the same interlayer parameterization (see Sec. 2.1.2) differing only in the cutoff length \(r_{\text{cutoff}}\) of the interlayer coupling. As it turns out, even within the same simple interlayer parameterization, the presence of strong localization depends crucially on this parameter \(r_{\text{cutoff}}\), a detail that is not even documented in most publications.

The most influential general difference between the various models of the interlayer coupling for investigations of incommensurate systems turns out to be their “smoothness”. In a “smooth” interlayer parameterization, small changes in the atomic configuration produce only small changes in the tight-binding Hamiltonian.
For commensurate systems, this only becomes an issue when continuous displacements are directly compared against each other for studying aspects like the angle-dependence in DWCNTs or the transport in continuously telescoping CNTs. In incommensurate systems, however, the smoothness of the interlayer coupling has a direct effect on the intrinsic properties of the system in a fixed configuration: an incommensurate system is aperiodic in the sense that no two regions are exactly identical. However, local atomic configurations will repeat approximately in a quasiperiodic pattern. A “smooth” parameterization will therefore result in a quasiperiodic Hamiltonian that changes slowly with each approximate repetition of the atomic structure and causes little scattering. A parameterization that is less smooth—be it because of an abrupt cutoff or simply a steeper functional form—will result in a Hamiltonian that varies strongly even if the atomic configuration is similar and the geometric quasiperiodicity will turn into a disorder-like effect in the Hamiltonian.

The foremost question about any theoretical modeling must of course be, whether it is a correct description of physical reality. Considering the delocalized nature of the π orbitals suggests that the electronic coupling between the layers should be fairly smooth. Ab initio results showing
strong variations in the electronic structure of non-Bernal-stacked graphene multilayers \[23\] need not necessarily disagree, since we have shown here, that even a smooth pair-coupling between atoms may produce such a dependency due to the symmetries of the graphene lattice.

6.5. Telescopic carbon nanotubes

After telescopically sliding multiwall CNT structures had been proposed in 1993 soon after the discovery of CNTs \[48\], the first experimental confirmation was reported in 2000 by J. Cumings and A. Zettl \[58\] and independently by Min-Feng Fu et al. \[280\], demonstrating the possibility to slide the shells against each other with extremely low friction with a constant pulling force caused by the attractive van der Walls interaction. Soon the electronic transport properties of such structures were investigated theoretically. Though some studies predict in a mere monotonic increase in resistance with increasing extension \[109\] which would be in agreement with experimental measurements \[59, 60\], theoretical studies of larger scale armchair-armchair structures generally result in resonances depending on the length of the overlapping region \[141, 43, 97, 251\] which have recently been confirmed experimentally \[129\].

The following sections are a brief outline of the our results presented in the diploma thesis by D. Darau \[5\] about the conductance of commensurate telescopic CNTs of different chiralities. A minimal model is given that allows to explain the observed oscillation patterns.

A sketch of the system to be studied is given in Fig. 6.9. We used the interwall parameterization as described in Sec. 2.1.2 with a cutoff length chosen as \( r_{\text{cutoff}} = 0.97 \sqrt{(1.6a^2) + \Delta R^2} \), short enough to allow the efficient computational handling in cells of length \( a \) with only nearest neighbor interactions.

6.5.1. Armchair telescopic tubes

As we have seen in the previous chapter, armchair DWCNTs are the only ones where the states around the Fermi energy have the same symmetry in both walls and can hybridize efficiently. Compared to zigzag DWCNTs, where the hybridization at the Fermi energy is of the order of 10 meV, the bands in a DWCNT split by 0.2 eV, promising a far more efficient intershell transmission. We will therefore start our investigation with the case of armchair telescopic CNTs where, where we can expect the to find the best conduction properties.
Figure 6.9.: Sketch a telescopic CNT. The combinations of tubes are chosen the same way as it was done for DWCNT. For armchair systems, that gives pairings like \((N, N)@\(N + 5, N + 5\)). Each tube is continued as a semi-infinite lead to avoid contact issues to interfere with the features that we want to study. The overlap length \(L\) is defined such that at \(L = 0\) the first unit cell of the outer tube is aligned with the first missing unit cell of the inner tube. The relative angle \(\vartheta\) is defined between two arbitrarily chosen atoms. (Figure taken from Ref. [5]).

Figure 6.10.: Transmission of three different armchair telescopic CNTs in dependence of the overlap length \(L\). The behavior is nearly identical in all three cases: interference patterns resulting in a oscillation at a length of \(3\alpha/4\). For small overlaps, the minimum of the oscillations is near zero, the maximum grows linearly. For larger \(L\), the amplitude itself shows an oscillation of a length of about \(70\alpha\) to \(75\alpha\) in a beat-like pattern. The maximum transmission is 1, indicating that one conductance channel is suppressed. (Figure taken from Ref. [5]).
6.5. Telescopic carbon nanotubes

The transmission through the telescopic system in dependence of the overlap length is displayed in Fig. 6.10. The fine oscillations of length $3a/4$ can be understood directly by considering the Fermi wave length of a standing wave at the end of a semi-infinite tube: $\lambda_F = 3a/2$. Transmission in the telescopic system is suppressed when the waves are in both sections are orthogonal to each other, which happens twice per $\lambda_F$. The slow amplitude modulation over 70 to 75 unit cells is caused by resonances of the standing waves in the overlapping region and will be discussed in detail in Sec. 6.5.3.

Figure 6.11.: Energy dependence of the DOS and the transmission in an armchair telescopic CNTs at two different overlap lengths: on the left, at $L = 35.8a$, the transmission at $E_F$ is at a maximum, on the right, at $L = 70.8a$, at a minimum. The only significant difference is near $E_F$. In the DOS, the only energy spacing of the resonances reflects weak Fabry-Pérot oscillations in the finite-length overlap region. The strong asymmetry between energies below and above $E_F$ directly reflects the broken electron-hole symmetry observed already in bilayer graphene (see Sec. 6.3): at high energies, the layers cannot hybridize, so intershell conductance is strongly suppressed. (Figure taken from Ref. [5]).

Going beyond the physics at the Fermi energy, we find mostly the chirality independent effects known already from bilayer graphene (see Sec. 6.3). For low energies, the massive graphene band is well described by
Chapter 6. Multilayer graphene and carbon nanotubes

an isotropic continuum model, so any combinations of chiralities strongly hybridize with each other. For high energies, the two wave functions have opposite signs for the two graphene sublattices, resulting in a strong cancellation of the interlayer matrix elements and with it, a suppression of the transmission.

6.5.2. Zigzag telescopic tubes

In zigzag telescopic CNTs, the transmission is generally orders of magnitude lower than in armchair CNTs, which is in perfect agreement with the suppressed hybridization found in zigzag DWCNTs. As the Fermi wave number $k_F$ of zigzag CNTs is zero, there is no well-defined electron wave length that would lead to regular resonances like in armchair telescopic CNTs. As can be seen in Fig. 6.13 and 6.12, well-defined oscillations can only be found due to the nodes of the wave functions inside the unit cell. The wavelengths of the Bloch waves themselves are highly dependent on energy and give rise to irregular resonances.

The energy resolved transmission in zigzag telescopic CNTs (see Fig. 6.14) shows the same asymmetry that was observed for armchair telescopic CNTs as well, due to the suppressed hybridization in the positive energy range common for all bilayer structures.

6.5.3. Model of telescopic quantum wires

A minimal model that allows an explanation of the long-wavelength amplitude modulations in the armchair telescopic CNT can be constructed as follows. Each nanotube is modeled as a one-dimensional continuum quantum wire where a free electron can propagate with a fixed wave number $k_0$. The telescopic region corresponds to a region where two quantum wires run in parallel (see Fig. 6.15). The coupling between the two wires leads to an energy split of the bands into one parallel ($\uparrow\uparrow$) and one antiparallel ($\uparrow\downarrow$) band.

In first order, the energy split $E \rightarrow E \pm \delta E$ of the bands results in a symmetric split of the Fermi wave numbers $k_{\uparrow\uparrow} = k_0 + \delta k$ and $k_{\uparrow\downarrow} = k_0 - \delta k$ defined by the Fermi velocity $v_F = dE/dk\hbar$ as $\delta k = \delta E/\hbar v_F$.

Within the center region $0 < z < L$, the corresponding wave functions fulfill $\Psi_{\uparrow\uparrow}^{\text{left}}(z) = \Psi_{\uparrow\uparrow}^{\text{right}}(z)$ and $\Psi_{\uparrow\downarrow}^{\text{left}}(z) = -\Psi_{\uparrow\downarrow}^{\text{right}}(z)$, where left and right specify the two wires present in the center region by the direction in which they extend.
6.5. Telescopic carbon nanotubes

**Figure 6.12.** Transmission through a (6,0)@(15,0) zigzag telescopic CNT. Overall transmission is several orders of magnitude lower than in the armchair case. The short wavelength oscillations (right) coincide exactly with the periodic structure of the system and can be explained by the nodes of the wave functions inside each unit cell. At $E = 0$, the wave number $k_F$ is near zero and the system does not have a well-defined wavelength. At $E = 0.1$ eV the wavelength is better defined (see Fig. 6.7) and the system shows slightly more regular resonances. (Figure taken from Ref. [5])
Figure 6.13.: Transmission through a (3,0)@(12,0) zigzag telescopic CNT. The overall transmission is three orders higher than in the previous case (Fig. 6.12) due to the stronger hybridization in this extremely lower diameter case. At $E = 0.1 \text{eV}$, the wave vectors are large enough to produce clear resonances at long wavelengths, though they are still less regular than in the armchair case. (Figure taken from Ref. [5])
6.5. Telescopic carbon nanotubes

Figure 6.14.: Energy resolved transmission in a zigzag telescopic CNT. The peak at $E = 0$ at $L = 0.5a$ is caused by the zigzag edge state. (Figure taken from Ref. [5])

Figure 6.15.: Sketch of the model of telescopic quantum wires: two wires run in parallel over length $L$. Both extend infinitely in opposite directions. In the common region, a coupling causes an energy split of $\pm \delta E$ into one band of even and another band of odd symmetry. At the Fermi energy, this translates into a split of the $k$ vector by $\pm \delta k$. 

135
The transmission of the model can be obtained based on the method of free particle wave functions in all three regions that have to follow the correct boundary conditions at the two points 0 and $L$. A detailed solution of this problem is given in Ref. [5] and leads to an expression of the form:

$$T = \frac{256k_0^2}{|\det M|^2} \left( k_{\uparrow\uparrow} \sin (k_{\uparrow\uparrow}L) - k_{\uparrow\downarrow} \sin (k_{\uparrow\downarrow}L) \right)^2 \quad (6.1)$$

with

$$\det M = (2\delta k \cos (\delta k L/2) - 4k_0 i \sin (\delta k L/2))^2 - (4k_0)^2 \exp (-2ik_0 L)$$

The resonances arising from this analytic formula are displayed in Fig. 6.16 and are comparable to those in the armchair telescopic CNT in Sec. 6.5.1.

Figure 6.16.: Transmission through the telescopic quantum wires as described by Eq. (6.1) for three different magnitudes of the coupling between the quantum wires. For the weakest coupling (i.e. the smallest value of $\delta k$) the resonant behavior resemble qualitatively that of the armchair telescopic CNTs in Fig. 6.10. The slight additional modulation of those resonances missing in this model is caused by the weak interference of the two channels. (Figure taken from Ref. [5])
Chapter 7.

Magneto-electronic structure and fractality

The electronic spectrum of graphene under the influence of external magnetic fields was first studied by J. W. McClure in 1956 [176], soon after P. R. Wallace had given the first derivation of the graphene bands [271]. As this work preceded the discovery of the quantum Hall effect [269], the special distribution of the Landau levels, though clearly visible in the spectrum, was not given much attention. Furthermore, J. W. McClure based his work fully on the perturbative regime of the magnetic field, lacking numerical tools and unaware yet of the fractal spectrum of lattice electrons that would later be found by Hofstadter [119].

Half a century later, when monolayer graphene had shifted into focus of theory, an anomalous odd-integer quantum Hall effect was first predicted by Y. Zheng and T. Ando in 2002 [287] and was then one of the first effects observed when exfoliated monolayers became available [285, 192, 284]. Inspired by this experimental success, the quantum Hall effect in graphene has become the focus of numerous theoretical works [105, 134, 214, 100, 111]. For bilayers of graphene, an additional degeneracy of the Landau levels and a Berry phase of $2\pi$ were predicted to lead to an anomalous quantum Hall effect, different from either the regular massive electrons or the special Dirac-type electrons of single-layer graphene [175], which was confirmed in experiment shortly afterwards [195] and used for the characterization of bilayer samples [197].

In CNTs, the effects of external magnetic fields have been of interest from the very beginning. Already in 1993, H. Ajiki and T. Ando predicted an oscillation band gap in SWCNTs in a parallel magnetic field, based on an effective mass theory [11], which could later on be observed experimentally by optical measurements in fields of up to 45 T [281]. A few years later, T. Ando and T. Nakanishi predicted a positive magnetoresis-
tance in perpendicular magnetic fields [21] that could also be measured in experiments [159, 85].

Independent of these band structure effects, regular negative magnetoresistance due to weak localization has also been measured in MWCNTs [156, 165] and was found to be quenched near the band edges due to enhanced electron-electron scattering [247, 248].

Interestingly enough, the quantum Hall effect, though it has attracted much attention in graphene, has been the subject of only few theoretical studies in CNTs [20, 215]. Possibly this is due to the extreme magnetic field strength necessary to observe the formation of Landau levels in a narrow conductor like a CNT. Fig. 7.1 gives an overview of the scales of the magnetic fields in connection to the structure of graphene and CNTs.

Figure 7.1.: Sketch to illustrate the magnetic field scales relevant for the magnetoelectronic structure of graphitic nanostructures. An individual plaquette of the area $A_{\text{plaquette}} = (3/2) \sqrt{3}d_{CC}^2 \approx 5.26 \text{ Å}^2$ relates to a physical magnetic field of $B_\perp = \Phi_0/A_{\text{plaquette}} \approx 79 \text{ kT}$ with the flux quantum $\Phi_0 = h/e$, a value several orders of magnitude beyond experimental reach. The relevant scale for magnetic fields parallel to the axis of a CNT is determined on its cross section: $B_\parallel = \Phi_0/r^2 \pi$. While this is still an extremely large value for small SWCNT [e.g. $B_\parallel = 7.9 \text{ kT}$ for a (6,6)-CNT] it comes down to realistic magnitudes for shells of large MWCNTs [e.g. $B_\parallel = 28 \text{ T}$ for a (100,100) CNT].

In this chapter, we will present various results on the evolution of the electronic structure in graphene and CNTs under the influence of an external magnetic field. The Hofstadter butterfly, originally studied on a square lattice, will be demonstrated for the graphene honeycomb lattice and its equivalence for various the CNTs will be presented. A special
periodic gauge that we developed to extend the study of the Hofstadter butterfly to arbitrary atomic configurations, will allow us to handle bilayer graphene in continuously displaced stackings and give a perspective to the results on DWCNTs that we present afterwards.

7.1. Peierls substitution

In 1933, R. E. Peierls introduced a method of modeling an external field acting on lattice electrons [210], today known as Peierls substitution. The core idea is based on starting from the principle of least coupling, substituting the quantum mechanical momentum of a charged particle with the vector potential as:

$$\hat{p} \rightarrow \hat{p} - eA$$

with the vector potential $A$ chosen such that its curl gives the magnetic field: $B = \nabla \times A$. Considering now that the momentum $\hat{p} = -i\hbar\nabla$ is the operator of infinitesimal displacement and that the operator of finite displacement from $r_i$ to $r_j$ can be expressed as $\exp\left(i (r_j - r_i) \hat{p}/\hbar\right)$, the hopping matrix element between two atoms $i$ and $j$ in a tight binding Hamiltonian is

$$\gamma^B_{ij} = \gamma^0_{ij} \exp\left(\frac{2\pi i}{\Phi_0} \int_{r_i}^{r_j} d\mathbf{r} \cdot \mathbf{A}(\mathbf{r})\right) \quad (7.1)$$

with the flux quantum $\Phi_0 = h/e$. In principle, this expression is dependent on the integration path between $r_i$ and $r_j$. For weak magnetic fields, however, the phase is approximately constant for all paths close to the straight line. Furthermore, one has to be aware that the expression is based on the assumption that the individual orbitals are negligibly small, so it breaks down when the magnetic flux captured by single orbitals is not small compared to the flux quantum.

Visualized in Fig. 7.2 is the consequence of this model that a closed loop formed by tight-binding links captures a total magnetic phase equivalent to the magnetic flux captured by the enclosed area:

$$\varphi_{i \rightarrow j \rightarrow k \rightarrow i} = \frac{2\pi}{\Phi_0} \int_{j \rightarrow k \rightarrow i} d\mathbf{r} \cdot \mathbf{A}(\mathbf{r})$$

$$= \frac{2\pi}{\Phi_0} \int d^2r B$$

$$= \frac{2\pi}{\Phi_0} FB_\perp$$
where second line follows by Stokes’ theorem and the third holds for a constant magnetic field $B_\perp$ perpendicular to the enclose area $F$. It is only this phase on closed loops that is physically observable.

Figure 7.2.: Scheme illustrating the implementation of an external magnetic field in a tight-binding model via the Peierls substitution: a link between two atoms carries a phase defined by the vector potential along the connecting path. A closed loop carries a total phase defined by the field threading through the enclosed area. (see text)

### 7.2. Hofstadter butterfly

D. Hofstadter was the first to study the electronic spectrum of lattice electrons in nonperturbatively strong magnetic fields [119]. Starting from a square lattice periodic in two dimensions (see Fig. 7.3), he used Peierls substitution to implement the magnetic and arrived at Harper’s equation [110] which he then solved numerically on a Hewlett-Packard 8920A table-top calculator (see Fig. 7.6), visualizing the result in what he called the `gplot` and what is today known as the Hofstadter butterfly (see Fig. 7.5).

Figure 7.3.: A simple square lattice with on-site energy $\varepsilon_0$ and coupling $\gamma_0$ identical in both directions. The energy band is given by $E(k) = \varepsilon_0 - 2\gamma_0 (\cos ka + \cos ka')$.

To apply a homogeneous perpendicular magnetic field $B_z$ to a square lattice with lattice constant $a$, certain links need to be multiplied by a complex phase factor in such a way, that circling around any single plaquette, the total phase comes out as $\exp(2\pi i B_z a^2/\Phi_0)$ (see Fig. 7.4). For irrational values of $B_z a^2/\Phi_0$, this leads to a non-periodic system that does not allow any exact solution. For any rational value, however, a finite-volume super-cell can be defined for which the Hamiltonian is periodic and can be solved using Bloch theorem.
7.2. Hofstadter butterfly

Figure 7.4.: Square lattice (lattice constant $a$) in homogeneous perpendicular magnetic field $B_z$. The gauge is chosen as $A(r) = B_z x e_y$. Any link in $y$ direction gets multiplied by $\exp\left(2\pi i B_z a^2 k/\Phi_0\right)$. If $B_z a^2/\Phi_0$ is a rational fraction $n/m$ (e.g. $1/3$), the link with index $k = m$ carries phase unity and the system is strictly periodic with a super-cell of $m$ times the original size.

The original Hofstadter butterfly as displayed in Fig. 7.5 results from setting up the periodic Hamiltonian for various rational values of the flux per plaquette and diagonalizing the resulting Bloch Hamiltonian for various $k$-points in the folded Brillouin zone. Each spectral value is plotted as a dot. For rational flux values with a small denominator, the super-cell contains few atoms, resulting in few bands with fairly wide dispersion. As the denominator (in the fully reduced fraction) grows larger, the super-cell contains more atoms, giving more bands of lower dispersion. Physically, of course, the infinitely self-similar pattern is limited by finite precision of measurement and finite size effects.

Figure 7.5.: The Hofstadter butterfly of the square lattice. Note the fractal structure, the mirror symmetry in the energy (electron-hole symmetry of the square lattice) and the periodicity (resulting from the fact that all plaquettes have the same magnetic flux which is periodic with the flux quantum $\Phi_0$. The lines at the top and the bottom of the spectrum indicate the linear Landau levels originating from the effective mass at the energy minimum and maximum.

The full fractal structure of the Hofstadter butterfly can be explored experimentally in artificially generated super-lattices [12, 94]. For real, atomic lattices, only the regime of low flux per plaquette can be explored. In this regime, the linear Landau levels emerge from the continuum limit.
of the lattice for fixed effective electron mass (straight lines in Fig. 7.5). These levels follow the usual law:

\[ E = E_0 \pm \hbar \omega (B) \left( n + \frac{1}{2} \right) \]  

(7.2)

with the cyclotron frequency \( \omega (B) = eB/m^* \) determined by the effective mass of the square lattice band extrema \( m^* = \frac{\hbar^2}{2\gamma_0} \).

7.3. Butterfly and anomalous Landau levels of graphene

Subsequently to the work of Hofstadter on the square lattice, various alternative topologies have been studied. The first, obvious choice was the hexagonal lattice, which has no electron-hole symmetry, leading to an asymmetric butterfly [55, 102]. The honeycomb lattice was studied soon after [220], though without reference to the yet-unknown anomalous quantum Hall effect of graphene. Finally, special Lieb and Kagome lattice structures were also studied, featuring graphene-like massless bands in a square symmetry [22, 130], leading to similar anomalous Landau levels.

The Hofstadter butterfly of a honeycomb lattice is displayed in Fig. 7.7. At the top and the bottom of the energy spectrum, the structure closely resembles that of the square lattice. The linear Landau levels are caused by the massive bands at the \( \Gamma \) point and can again be described by Eq. (7.2), this time with an effective mass of \( m^* = 2\hbar^2/3\gamma_0d_{CC}^2 \approx 0.95m_e \). At the Fermi energy \( E_F \) however, a very different behavior can be observed based on the massless bands of graphene near the \( K \) points (see Sec. 2.2).

As it turns out, the linearized Hamiltonian of graphene near the Fermi energy can be expressed formally equivalent to the relativistic Dirac equation in two dimensions. This leads to a very special spectrum that can...
7.3. Butterfly and anomalous Landau levels of graphene

Figure 7.7.: Hofstadter butterfly of a graphene-like honeycomb lattice. At the minimum and the maximum of the spectrum, the same structure as in the square lattice is visible, featuring regular Landau levels that originate from the massive bands at the Γ point in graphene. At the Fermi energy, a very different structure is visible, consisting of one level at fixed energy, the supersymmetric Landau level (SuSyLL), surrounded by a set of square-root-shaped levels, the relativistic Landau levels.

be found by an analysis based on supersymmetric arguments as it was given by M. Ezawa [81, 82]. An excellent review of supersymmetry in the perspective of condensed matter physics is given by G. Junker [133].

The spectrum in a homogeneous external magnetic field $B$ consists of a set of relativistic Landau levels (ReLL):

$$E_{\pm n}^{\uparrow \downarrow} = \pm v_F \sqrt{2e\hbar \pm n}$$

with $n \in \{1, 2, \ldots\}$ and a special supersymmetric Landau level (SuSyLL)

$$E_0 = 0$$

The index $\uparrow$ or $\downarrow$ of the ReLL indicates that each is twice pseudospin degenerate. Note the square root scaling of the ReLL in contrast to the linear scaling of the conventional Landau levels of effective mass particles.

The SuSyLL is also twice degenerate, but with both pseudospins having the same sign, depending on the sign of the magnetic field $B_z$. In addition, all levels have a two-fold degeneracy due to the symmetry between the two cones in the Brillouin zone, giving in total a four-fold symmetry of each landau level. At charge neutrality, the SuSyLL is exactly half filled, making the graphene sheet conducting at arbitrarily strong magnetic fields. Most remarkable about the supersymmetric derivation is that is is robust also against arbitrary inhomogeneities of the magnetic field, as it was shown by Y. Aharonov and A. Casher [7].

A detailed derivation of this spectrum, based on the supersymmetric structure, is given in App. B.1.
7.4. Butterfly of single-wall nanotubes

A strong magnetic field imposed upon a CNT can be modeled in tight-binding in the same way as for the two dimensional lattices before, using Peierls substitution. A considerable simplification comes from the fact that the CNT is periodic only in one direction and finite in both others. Choosing the $z$ direction parallel to the CNT axis and the magnetic field as $\mathbf{B} = B_x \mathbf{e}_x + B_y \mathbf{e}_y + B_z \mathbf{e}_z$, a gauge can therefore be chosen straightforwardly as:

$$A(x) = B_z x \mathbf{e}_y + (B_x y - B_y x) \mathbf{e}_z$$

which is independent of the $z$ coordinate along the CNT. Knowing all the atomic positions in the unit cell, the complex phase for any tight-binding link can easily be computed, resulting in a periodic Hamiltonian that can be diagonalized using Bloch theorem. To visualize the continuous evolution of the electronic structure with changing magnetic field, the density of state (DOS) can be plotted as color-code against energy and field (see Fig. 7.8).

**Figure 7.8.** Visualization scheme for the continuous evolution of the band structure of a CNT in a homogeneous external magnetic field. The band structure, distorted differently by each magnetic field strength (right) can be projected to obtain the corresponding DOS which is then visualized as color in the $E$-$B$-plot plane.

The two special cases, the field exactly parallel or exactly perpendicular to the tube axis ($B_{\parallel}$ resp. $B_{\perp}$), show very different results: $B_{\parallel}$ produces periodic oscillations opening and closing a gap, based on the Aharonov-Bohm effect [10, 26]. $B_{\perp}$ on the other hand results in a much more complex
7.4. Butterfly of single-wall nanotubes

pattern of bands that shows strong similarity with the Hofstadter butterfly of planar graphene (Fig. 7.7). In Fig. 7.9, the Hofstadter butterfly of an armchair CNT is displayed. Remarkable is the fact that the perpendicular field does not open a gap at any strength of the field as it is protected by the special supersymmetry in armchair CNTs [161].

Figure 7.9.: Hofstadter butterfly of a (6,6) CNT both in perpendicular (left) and parallel (center) magnetic fields. The transmission through the perfectly periodic system (right) takes on only integer values, counting the number of channels at each energy and magnetic field.

In semiconducting CNTs (see Fig. 7.10 and Fig. 7.11), the electronic gap is opened and closed in a highly irregular pattern that can be understood by the large number of bands touching the Fermi energy at different fields.

Figure 7.10.: Hofstadter butterfly of a (11,0) CNT. The gap at the Fermi energy opens and closes in a highly irregular pattern.

The butterfly of the chiral CNT clearly shows the effect of the breaking of the helical symmetry by the magnetic fields. The large number of bands that are connected by the helical symmetry for zero magnetic field, are now broken up into separate bands. The statistical distribution of this larger number of bands causes the gap to be smaller than in the case of the zigzag CNT.
For extremely large diameter CNTs, the wall closely resembles a planar graphene sheet and the overall butterfly plot converges against an overlay of various graphene butterflies as to be seen in Fig. 7.12. Much of the physics in this system can be understood from a continuum model as it is described in our published work on this issue [3].

One detail that can be nicely observed in this example is the effect of the finite lateral size on the fractality of the original Hofstadter butterfly. The butterfly of a two-dimensionally infinite planar lattice is truly fractal in the sense, that the self-similarity repeats to arbitrary magnification of the pattern. A lattice of finite lateral size, on the other hand, results in a finite number of bands, causing the patterns to end in a continuum after a few levels of magnification, producing what we call a pseudofractal structure.

For experimentally accessible magnetic fields and low energy, the DOS of large CNTs, as displayed in Fig. 7.13, can be described well by the approximation of linear bands and follows the scaling law

\[ g_{(N,N)}(E,B) = \frac{N'}{N} g_{(N',N')} \left( \frac{N}{N'}E \frac{N^2}{N'^2} B \right) \]
7.5. Graphene nanoribbons

Clear similarities to the physics of CNTs can also be found in the magnetoelectronic structure of graphene nanoribbons (GNRs). While the whole butterfly has the same $\Phi_0$-symmetry as that of a two-dimensionally infinite graphene sheet, the full fractal structure of the same is not present, due to the finite width of the ribbon. Instead, a finite number of bands is found, which are continuously distorted with increasing magnetic field. As can be seen in Fig. 7.14, the magnetoelectronic structure at low fields closely resembles that in CNTs, featuring the same peak in the DOS arising at the Fermi energy.

7.6. Periodic gauge

The method described for setting up the Hamiltonian for the computation of the original Hofstadter butterfly of the square lattice in Sec. 7.2 worked very specifically for this structure. It can be generalized straightforwardly only to very specific simple structures like the hexagonal or the honeycomb lattice. To handle periodic structures with atoms in arbitrary positions, a gauge needs to be found that yields a vector potential which is defined everywhere and is itself periodic.
Figure 7.14.: Density of states of a finite width GNR. The width \( W = 20 \text{ nm} \) corresponds to an unrolled \((80,0)\) CNT. The gap of the originally semiconducting ribbon is narrowed by the magnetic field and a flat band appears at zero energy, resulting in a peak of the DOS. This flat band is the equivalent of the supersymmetric Landau level of two-dimensionally infinite graphene sheet. Also the square root shaped relativistic Landau levels are clearly visible through the shifting of the van Hove singularities that correspond to band edges.

For a strictly homogeneous nonzero magnetic field, such a gauge does not exist, as can be seen by a simple argument: Assume

\[
A(r) = A(r + a_x) = A(r + a_y)
\]

for two lattice vectors \( a_x \) and \( a_y \) in the \( x\)-\( y\)-plane. If follows that

\[
\int_{r_0}^{r_0 + a_x} \, dr \cdot A(r) = \int_{r_0}^{r_0 + a_y} \, dr \cdot A(r)
\]

and

\[
\int_{r_0}^{r_0 + a_y} \, dr \cdot A(r) = \int_{r_0}^{r_0 + a_x} \, dr \cdot A(r)
\]

so a circular integral along the four line segments will always cancel out exactly, resulting in a vanishing total magnetic flux over the area \( a_x \cdot a_y \).

This fundamental restriction can be circumvented by introducing singular flux vortices of the form:

\[
B(x, y, z) = \Phi_0 \delta(x - x_0) \delta(y - y_0) \, e_z
\]

(assuming the physical magnetic field in the \( z \) direction). Such a flux vortex of zero width has no physical effect, since magnetic fields can only be measured as the phase gathered by a circular path around an area. The phase change of exactly \( 2\pi \) in this case is physically undetectable. Distributing such vortices in a periodic array allows to obtain a vanishing
total flux per unit cell for a homogeneous nonzero physical magnetic field and thereby permits to find a periodic gauge.

For the lattice vectors $a_x$ and $a_y$ and their reciprocal counterparts $\tilde{a}_x$ and $\tilde{a}_y$ (defined by $a_i \cdot \tilde{a}_j = 2\pi \delta_{ij}$), an integer number $n$ of flux vortices per unit cell allows to compensate for a homogeneous magnetic field $B = n\Phi_0 (\tilde{a}_x \times \tilde{a}_y) / 4\pi^2$. The usual linear—but aperiodic—gauge would be $A_{\text{lin}}(r) = n\Phi_0 (r \cdot \tilde{a}_x) \tilde{a}_y$. A periodic gauge can now be defined as:

$$A(r) = n\Phi_0 \left( [r \cdot \tilde{a}_x] \tilde{a}_y - [r \cdot \tilde{a}_y] \delta (\{r \cdot \tilde{a}_x\}) \tilde{a}_x \right)$$

where $\{.\}$ denotes the fractional part of a real number ($\{x\} = x - \lfloor x \rfloor, 0 \leq \{x\} < 1 \forall x \in \mathbb{R}$). To make sure that the phase of every link between two atoms is well defined, the gauge field is displaced by an infinitesimal amount such that every atom is either left or right of the convergent line.

As with the simpler modeling for the square lattice that bypassed the explicit definition of a periodic vector potential, this method still is limited by the fundamental restriction to rational values of the flux per unit cell in order to be able to define a finite super-cell: the lattice of flux vortices has to be commensurate with the periodicity of the underlying physical structure. Within this unit cell, however, the atoms can be arbitrarily located and the phase of each link can be determined by integrating up the vector potential as prescribed in Eq. (7.1). A concise example implementation is given in App. C.8.

### 7.7. Bilayer graphene

The periodic gauge introduced in the previous section can be applied directly to graphene bilayers, removing any restriction to the relative position of the atoms within the unit cell or to the links in the tight binding network. Fig. 7.15 shows the two example configurations of highest symmetry which we will study first, before taking a look at the continuum of intermediate configurations.

The Hofstadter butterfly of bilayer graphene in these two configurations can be seen in Fig. 7.16. As to be expected, the relatively weak interlayer coupling preserves the general shape of the structure obtained from the monolayer in Fig. 7.7, but breaks some of the symmetries, namely the electron-hole symmetry (as observed already without magnetic field in Sec. 6.3) and the $\Phi_0$ periodicity.

The look at a finite width ribbon allows a detailed zoom into the region of experimentally accessible magnetic fields and low energy. The plots in
Chapter 7. Magnetoelectronic structure and fractality

Figure 7.15: The two highest symmetry configurations of bilayer graphene. Bernal stacking (top) is the energetically preferred configuration which is also found within natural graphite. The aligned stacking is energetically unstable, but it is bound to occur in regions of large DWCNTs or bilayers with a very small rotational angle and might also be realized in mechanically constricted samples.

Figure 7.16: The butterfly of bilayer graphene in the two different configurations shown in Fig. 7.15. Compared to Fig. 7.7, the two mirror-symmetries are broken: the electron-hole symmetry \( E \leftrightarrow -E \) is broken already for zero magnetic field, as described in Sec. 6.3. The breaking of the \( \Phi_0 \)-periodicity in the magnetic flux per plaquette can be explained for the Bernal-stacked bilayer by the presence of closed loops with the effective area of \( 1/3 \) of a plaquette. Beyond this, the additional long-distance-bonds, caused by the large cutoff length of the interlayer coupling also breaks the periodicity to a smaller degree, causing the slight asymmetry also for the aligned stacking. Also visible is a split of the SuSyLL in the aligned stacking configuration.
7.8. Butterfly of double-wall nanotubes

Fig. 7.17 reveal the split of the SuSyLL. It was stated before that this level should remain degenerate due to supersymmetry of the system also in bilayer graphene [195, 175, 82]. However, these analytic derivatives were generally based on Bernal-stacked graphene with an interlayer coupling that considered only hoppings between the nearest neighbors. In this numerical study that contains all hoppings up to the cutoff length, a tiny split of $\Delta E_{\text{SuSyLL}} \sim 0.01 \text{ eV}$ is caused in Bernal stacking by long-ranged interlayer coupling terms. The much larger split in aligned configuration originates from the shape of the wave functions in the individual monolayers that now have the same symmetry and can couple efficiently already by the short-distance terms in the interlayer coupling.

Figure 7.17.: Bilayer graphene ribbons in two different configurations at realistic magnetic fields up to 200 T at low energy. In Bernal stacking, the SuSyLL is split by a tiny amount of $\Delta E_{\text{SuSyLL}} \sim 0.01 \text{ eV}$. In aligned stacking configuration, the same split grows to $\Delta E_{\text{SuSyLL}} \sim 0.3 \text{ eV}$. The continuous transition between both extremes is displayed in Fig. 7.18.

Thanks to the periodic gauge introduced in the previous section, it is possible now to continuously slide the two layers against each other and observe the transition between the two high-symmetry configurations that we studied up to now. The two special configurations indeed turn out to give the extremal results for $\Delta E_{\text{SuSyLL}}$. In between, the transition is smooth, except for tiny discontinuities caused by the parameterization of the interlayer coupling.

7.8. Butterfly of double-wall nanotubes

The large diameter SWCNTs described at the end of Sec. 7.4 were of the size of the outer shell of a typical MWCNT. In a real MWCNT, however, the intershell coupling can be expected to have serious influence on the
magnetoelectronic structure, as we have seen from the effect in bilayer graphene in the last section. Indeed, as it can be observed in Fig. 7.19, the coupling causes a general split of the two copies of the electronic structure by an amount of $\Delta E \sim 0.1$ eV.

In parallel magnetic field, the two walls independently would have gaps opening in closing with similar periodicity. The split caused by the coupling closes these gaps nearly exactly, accompanied by several further indications of hybridization between the walls (avoided band crossings).

In perpendicular field, the same split acts on the SuSyLL, clearly producing two regions of increased DOS, which are further distorted in comparison to the very clean level at the Fermi energy in the decoupled case.

In contrast to the strong dependence of $\Delta E_{\text{SuSyLL}}$ on the exact configuration of bilayer graphene, the data in Fig. 7.20 shows a constant split of $\Delta E_{\text{SuSyLL}} \sim 0.1$ eV for any relative configuration. Though the details of the magnetoelectronic structure depend visibly on the relative rotation and shift of the walls in the DWCNT, the absolute value of the split is fixed. This is understandable when considering that the states near the Fermi energy are delocalized around the whole circumference and are therefore sensitive to an average of all the configurations that occur along the perimeter of the DWCNT.
7.8. Butterfly of double-wall nanotubes

**Figure 7.19.** Magnetoelectronic structure of a \((100,100)@(105,105)\) DWCNT without (left) and with (right) interwall coupling. The main effect of the coupling is a general split by \(\Delta E \sim 0.1\) eV. In parallel fields, this split closes the gaps that would open otherwise. In perpendicular fields, the SuSyLL is split by the same amount and then further distorted.

**Figure 7.20.** Magnetoelectronic structure of a \((95,95)@(100,100)\) DWCNT in various relative orientations. The schematics illustrate the atomic configuration in each case at the top and the bottom w.r.t. a vertically imposed magnetic field. Without magnetic field, the band structures of the systems are indistinguishable. The field breaks the rotational symmetry and causes a clear distinction between the configurations. Yet, the split of the SuSyLL remains fixed at 0.1 eV.
Conclusions and perspectives

In the course of this work, some light has been shed on several issues of immediate experimental relevance about quantum transport properties and electronic structure of carbon nanotubes and graphene. The theoretical methods have been described in detail and special attention has been given to experimental aspects, both in modeling and in the interpretation of the results.

In Chap. 4, we introduced a new model for electrical contacts to carbon nanotubes that captures the fact that the contacts in an experimental setup are typically extending over a length of several hundred nanometers, covering the nanotube with some contact metal. We have demonstrated the counterintuitive result that, given a metal coating of several nanometers length, the contact transparency is actually improved by using a metal that couples weaker to the surface of the nanotube. Using ab initio results of Ti and Pd as contact metals, we have demonstrated that Pd actually forms a weaker bond, giving a new explanation of the experimental finding that Pd forms good contacts. This finding suggests a possible route for future attempts in optimizing the charge injection in carbon nanotubes: find contacting materials that couple to the nanotube as softly as possible to exploit the available contact length. Using a related approach to model ferromagnetic contacts, we studied the magnetoconductance based on the concept transport in separate spin channels, demonstrating that disorder in the conducting region reduces the signal of magnetoresistance even in the absence of spin-flips, simply due to the fact that symmetries are broken and channels that would otherwise be prohibited, are now opened.

In a more detailed analysis of the effects of disorder and defects in Chap. 5, we have first demonstrated that the two regimes of diffusive transport and strong localization in carbon nanotubes can both be understood quantitatively from the band structure alone in the case of weak, homogeneous disorder. Even for point defects modeled as single atom vacancies, the diffusive and localized regime can be extrapolated to reasonable precision from the transmission of a single defect. In graphene
Conclusions and perspectives

nanoribbons, the situation for homogeneous disorder is different mainly due to the influence of the zigzag edge state at the Fermi energy. The same principles that make armchair nanotubes especially robust against disorder due to its low density of states at the Fermi energy lead to a severe reduction of the transmission in zigzag ribbons even for weak disorder. For vacancy defects in graphene nanoribbons, as an additional complication it has to be taken into account that the lateral position of the defect and even its sublattice strongly influence its effect on the transmission. For edge defects, we have also demonstrated that, even though a single defect has little influence on the transmission at the Fermi energy, multiple defects interfere strongly even over a distance of several unit cells, and reduce the total transmission more than it would be expected from individual defect. The experimental relevance of this last finding is fairly fundamental: It is clear that graphene ribbons are prone to defects especially at the edges. Knowing about the strong interference between multiple defects, one can expect that reducing the defect rate below a certain threshold will pay off disproportionately high.

In the study of the interlayer coupling in multiwall nanotubes and multilayer graphene in Chap. 6, we have presented some new insight about the interplay of incommensurability and the detailed modeling of the interlayer coupling. For graphene bilayers that are shifted but not rotated, the coupling of the states at the Fermi energy is generally strong, but the shape of the hybrid energy bands depends strongly on the relative position of the two layers. For rotated bilayers with a common supercell, the coupling of these bands is weak, even though the reduced symmetry would permit it. The reason is that a smooth enough interlayer coupling still approximately preserves the momenta in the individual layers which do not match for rotated layers. In double-wall carbon nanotubes, the same principle of approximate symmetries results in the difference in the coupling of the bands at the Fermi energy between either pure armchair or pure zigzag double-wall tubes. In the case of two armchair walls, the Fermi angular momentum is always zero, so the bands can couple efficiently. Otherwise, the Fermi angular momenta are different and prevent a coupling, even if the combined system shows now rotational symmetry at all. For double-wall nanotubes that are incommensurate in the longitudinal direction, we demonstrated that the smoothness of the interlayer coupling becomes really critical: The local atomic configuration never repeat identically, but one can always find quasiperiodic lengths at which the configurations repeat approximately. A smooth interlayer coupling guarantees that the Hamiltonian of such a system also is approximately
periodic and gives little rise to backscattering. We have demonstrated that within a widely used interlayer parameterization, the cutoff length has to be chosen very long to guarantee the necessary smoothness and prevent artificial disorder caused by the incommensurability. This insight does not solve the experimental issue, whether multiwall tubes should exhibit diffusive or ballistic transport. It does, however, indicate that whatever disorder it might be that causes diffusion, it is likely to be caused by external effects rather than intrinsic to the incommensurate system.

Finally, in Chap. 7, we have demonstrated in detail what effect external magnetic fields have on the band structure of carbon nanotubes and graphene. Using a visualization method similar to the Hofstadter butterfly, we showed that finite width systems like nanotubes of nanoribbons exhibit a pseudo-fractal structure that differs from real fractality by a finite depth of self-similarity. The aperiodic oscillations of the band gap of semiconducting nanotubes in a perpendicular field result have been displayed in the context of the full evolution of the band structure. We have developed a periodic gauge that permits the computation of Hofstadter butterfly plots for two-dimensionally periodic systems with atoms located at arbitrary positions within the unit cell. This gauge was then used to obtain the Hofstadter butterfly of bilayer graphene at arbitrary relative positions. The symmetries of the original butterfly of graphene have been demonstrated to be broken and a split of the special zero-energy Landau level was demonstrated to depend on the relative position of the two layers. In double-wall nanotubes, the same split was demonstrated to exist, but with a magnitude independent of the relative shift or rotation of the two tubes. Experimentally, the full fractal structure of the Hofstadter butterfly in graphene-like structures is far out of reach, due to the extreme magnetic fields that would be needed. However, anomalous Landau levels at accessible fields have been demonstrated in experiment already and certain fractal structures that were found in measurements of the quantum Hall effect in graphene can hopefully be explained by superlattice structures due to a substrate or intrinsic to rotated graphene bilayers.

There are several issues that demand for further investigation: The question of extended contacts still neglects several important aspects like that of band bending at the contact edges due to shifts caused by the observed charge transfer. Also the microscopic structure of the metal-carbon interface would need more detailed ab initio studies to offer better quantitative predictions. The investigation of further alternative contact-
Conclusions and perspectives

ing metals is also of high relevance and could be based upon the existing studies of similar interfaces.

The whole issue of spin transport in ferromagnetically contacted carbon nanotubes was, of course, only scratched on the surface by our study. Any in-depth work on this subject will certainly have to do a far more detailed modeling of the contacts. We believe, however, that our concept of extended contacts will prove valuable also when considering ferromagnetic metals.

For disorder in nanotubes and ribbons, the obvious next step will be a detailed investigation of the deviations from the perturbative regime. Some of the effects may already be explained by taking into account the modified density of states in a disordered system. Other effects may be captured by higher order perturbation theory. Alternatively, it will certainly be worthwhile to take a close look at the random-matrix theoretical approach, to see in what aspects our straightforward may yet be improved.

Point defects leave even more room for future studies. Individual point defects may have very different effects on the transmission, especially in ribbons, where the exact position also has to be taken into account. Even more complexity, however, is to be expected in the study of the interference between multiple defects.

In the study of the multilayer coupling, the most immediate next step would be the search for an improved tight binding parameterization that take into account the special demands of incommensurate systems. Fitting ab initio results of continuously displaced commensurate systems would be a possible step to proceed in this direction. At the time being, such attempts are yet unknown.

Finally, for studies of the magnetoelectronic structure, it would be of great interest to observe the effect of superlattice structures on the Landau levels of graphene. At an appropriate scale of such structures, the fractality of the Hofstadter butterfly should be directly experimentally observable and could give an interesting interplay with the anomalous Landau levels of the underlying system.
Appendix A.

Decimation techniques

Decimation is a powerful technique for the efficient handling of extended systems. This tutorial explains the basics as well as two of the more advanced applications. Once the reader is familiar with the basic principle behind these algorithms, one should quickly be able to model new use cases of different geometries of symmetries.

All of the following was written with a quasi-one-dimensional single-electron tight-binding model in mind (more specifically a carbon nanotube in π orbital approximation). The technique, however should be just as useful for many other areas that involve a localized basis and a local Hamiltonian in this basis that allows the exploitation of the locality of each decimation step.

A.1. The fundamental equation of decimation

Be \( \mathcal{A} \) an invertible matrix of the form:

\[
\mathcal{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}
\]

where all for \( A_{ij} \) are themselves matrices, \( A_{11} \) and \( A_{12} \) invertible.

The inverse \( \mathcal{B} = \mathcal{A}^{-1} \) can be divided up in the same way:

\[
\mathcal{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}
\]

Expanding the equation \( \mathcal{A}\mathcal{B} = 1 \) into these block matrices, we get for equations:

\[
\begin{align*}
A_{11}B_{11} + A_{12}B_{21} &= 1 \\
A_{21}B_{11} + A_{22}B_{21} &= 0 \\
A_{11}B_{12} + A_{12}B_{22} &= 0 \\
A_{21}B_{12} + A_{22}B_{22} &= 1
\end{align*}
\]
Rewriting Eq. (A.2) as
\[ B_{21} = -A^{-1}_{22}A_{21}B_{11} \]  
(A.3)
and substituting this into Eq. (A.1), we get
\[ (A_{11} - A_{12}A^{-1}_{22}A_{21}) B_{11} = 1 \]
so we have the fundamental equation of decimation:
\[ B_{11} = (A_{11} - A_{12}A^{-1}_{22}A_{21})^{-1} \]  
(A.4)

A.1.1. Application to Green functions: Bipartite systems

Be \( \mathcal{H} \) the Hamiltonian of a bipartite system in matrix form:
\[ \mathcal{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \]
where both parts may be finite or infinite. From \( \mathcal{H} = \mathcal{H}^\dagger \), it follows that \( H_{ij} = H_{ji}^\dagger \).

System 1 will later represent the system of interest, while system 2 will be the environment, which we don’t want to study per se, but only consider its effect on the system of interest.

The Green function of the complete system is defined as:
\[ \mathcal{G}(E) = \left( (E + i\eta) \mathbb{1} - \mathcal{H} \right)^{-1} = \begin{pmatrix} G_{11}(E) & G_{12}(E) \\ G_{21}(E) & G_{22}(E) \end{pmatrix} \]
where \( i\eta \) is an imaginary quantity needed to move the poles away from the real axis.

Mathematically it is generally considered infinitesimal and one takes the limit \( \lim_{\eta \to 0^+} \). Numerically, one usually picks a small but finite value \( \eta \), chosen small enough that its effects on the results are acceptable but large enough to avoid badly conditioned matrices for the numerical inversions. Physically, one can interpret such a finite imaginary part to the energy as effective dissipation to the environment.

In the following only retarded Green functions will be considered and the \( r \) will be dropped from the notation. For advanced Green functions all derivations can be done analogous.
From Eq. (A.4) it follows that:

\[ G_{11}(E) = \left( \left( (E + i\eta) 1 - H_{11} \right) - (-H_{12}) \left( (E + i\eta) 1 - H_{22} \right)^{-1} (-H_{21}) \right)^{-1} \]

In this, we can define the Green function of the isolated system 2:

\[ G_{22}^0(E) = \left( \left( (E + i\eta) 1 - H_{22} \right)^{-1} \right) \]

and the effective Hamiltonian of system 1:

\[ H_{11}^{\text{eff}}(E) = H_{11} + H_{12} G_{22}^0(E) H_{21} \quad (A.5) \]

where \( \Sigma_2(E) := H_{12} G_{22}^0(E) H_{21} \) is called the self-energy encapsulating the effect of the environment on the system of interest.

Beware the difference in the sign between Eq. (A.4) and Eq. (A.5). It arises from the negative sign in the definition of the Greens function and is a common cause for coding errors in numerical implementations.

A further helpful relation can be derived from Eq. (A.3) to access the off-diagonal blocks of an inverted 2 × 2 block matrix:

\[ G_{12}(E) = \left( (E - H_{11}) - H_{12} G_{22}^0 H_{21} \right)^{-1} H_{12} G_{22}^0 \]
\[ = G_{11}^0 H_{12} \left( (E - H_{22}) - H_{21} G_{11}^0 H_{12} \right)^{-1} \]

\section*{A.2. Tripartite systems}

Going one step further, we can start with the Hamiltonian of a tripartite system:

\[ \mathcal{H} = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \]

The Green function is

\[ \mathcal{G}(E) = \left( (E + i\eta) 1 - \mathcal{H} \right)^{-1} \]
\[ = \begin{pmatrix} G_{11}(E) & G_{12}(E) & G_{13}(E) \\ G_{21}(E) & G_{22}(E) & G_{23}(E) \\ G_{31}(E) & G_{32}(E) & G_{33}(E) \end{pmatrix} \]

Imagine now, we want to study only the boundaries 1 and 3 of this system, but not the interior 2. I.e., we want to know only \( G_{11}, G_{13}, G_{31} \) and \( G_{33} \). We
Appendix A. Decimation techniques

start by reordering the Hamiltonian (a simple unitary transformation):

$$\mathcal{H}' = \begin{pmatrix} H_{11} & H_{13} & H_{12} \\ H_{31} & H_{33} & H_{32} \\ H_{21} & H_{23} & H_{22} \end{pmatrix} = \begin{pmatrix} H_{bb} & H_{b2} \\ H_{2b} & H_{22} \end{pmatrix}$$

with $H_{bb} = \begin{pmatrix} H_{11} & H_{13} \\ H_{31} & H_{33} \end{pmatrix}$, $H_{b2} = \begin{pmatrix} H_{12} \\ H_{32} \end{pmatrix}$, and $H_{2b} = \begin{pmatrix} H_{21} \\ H_{23} \end{pmatrix}$.

Now, we have a bipartite system and can use Eq. (A.5):

$$\begin{pmatrix} H_{ee}^{bb} & H_{ee}^{b2} \end{pmatrix} = \begin{pmatrix} H_{bb} + H_{b2} G_{22}^{0} (E) H_{2b} \\ H_{31} + H_{32} G_{22}^{0} (E) H_{23} \end{pmatrix} = \begin{pmatrix} H_{11} + H_{12} G_{22}^{0} (E) H_{21} & H_{13} + H_{12} G_{22}^{0} (E) H_{23} \\ H_{31} + H_{32} G_{22}^{0} (E) H_{23} & H_{33} + H_{32} G_{22}^{0} (E) H_{23} \end{pmatrix}$$

The four interesting parts of the Green function can now be obtained by inverting this smaller matrix:

$$\begin{pmatrix} G_{11} (E) & G_{13} (E) \\ G_{31} (E) & G_{33} (E) \end{pmatrix} = \left( (E + i\eta) I - \begin{pmatrix} H_{11} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & H_{N-1,N} \end{pmatrix} \right)^{-1}$$

A.3. Finite block tridiagonal systems

Consider now an extended quasi one-dimension system with a Hamiltonian written in some localized basis (e.g. the tight-binding Hamiltonian of a carbon nanotube). The full system can be divided up into a chain of $N$ cells, large enough that there is no direct coupling beyond directly neighboring cells. The Hamiltonian of this system is a block tridiagonal matrix:

$$\mathcal{H} = \begin{pmatrix} H_{11} & H_{12} & 0 & \cdots & 0 & 0 \\ H_{21} & H_{22} & \ddots & \ddots & \vdots & \vdots \\ 0 & H_{32} & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & H_{N-2,N-1} & 0 \\ \vdots & \vdots & \ddots & \ddots & H_{N-1,N-1} & H_{N-1,N} \\ 0 & 0 & \cdots & 0 & H_{N,N-1} & H_{N,N} \end{pmatrix}$$
A.3. Finite block tridiagonal systems

The important property of this matrix is that $H_{ij} = 0$ for any $|i - j| > 1$.

The full Green function is defined as before:

$$\mathcal{G}(E) = \left((E + i\eta) \mathbb{1} - \mathcal{H}\right)^{-1}$$

$$= \begin{pmatrix}
G_{11}(E) & \cdots & G_{1N}(E) \\
\vdots & \ddots & \vdots \\
G_{N1}(E) & \cdots & G_{NN}(E)
\end{pmatrix}$$

Like in the case of the tripartite system, we are now only interested in the Green functions of the two surfaces and their connections: $G_{11}(E)$, $G_{1N}(E)$, $G_{N1}(E)$ and $G_{NN}(E)$. As before, we can decimate out the central region from 2 to $N - 1$, but this time not in one step but one block after the other. With each decimation of one block, we get an intermediate effective Hamiltonian denoted by $H_{ij}^{(n)}$ which is then again decimated to get $H_{ij}^{(n+1)}$. In the first step, we decimate out cell 2:

$$\begin{pmatrix}
H_{11}^{(1)} & H_{13}^{(1)} & H_{14}^{(1)} & \cdots & H_{1N}^{(1)} \\
H_{31}^{(1)} & H_{33}^{(1)} & H_{34}^{(1)} & \cdots & H_{3N}^{(1)} \\
H_{41}^{(1)} & H_{43}^{(1)} & H_{44}^{(1)} & \cdots & H_{4N}^{(1)} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
H_{N,1}^{(1)} & H_{N,3}^{(1)} & \cdots & H_{N,N}^{(1)}
\end{pmatrix} = \begin{pmatrix}
H_{11} + H_{12}G_{22}^{(0)}H_{21} & H_{12}G_{22}^{(0)}H_{23} & 0 & \cdots & 0 \\
H_{32}G_{22}^{(0)}H_{21} & H_{33} + H_{32}G_{22}^{(0)}H_{23} & H_{34} & 0 & \vdots \\
0 & H_{43} & H_{44} & \ddots & \vdots \\
\vdots & 0 & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & H_{N,N}
\end{pmatrix}$$

with $G_{22}^{(0)} = (E + i\eta - H_{22})^{-1}$. Next, we decimate out cell 3 and obtain:

$$\begin{pmatrix}
H_{11}^{(2)} & H_{13}^{(2)} & H_{14}^{(2)} & \cdots & H_{1N}^{(2)} \\
H_{31}^{(2)} & H_{33}^{(2)} & H_{34}^{(2)} & \cdots & H_{3N}^{(2)} \\
H_{41}^{(2)} & H_{43}^{(2)} & H_{44}^{(2)} & \cdots & H_{4N}^{(2)} \\
H_{51}^{(2)} & H_{53}^{(2)} & H_{54}^{(2)} & \cdots & H_{5N}^{(2)} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
H_{N,1}^{(2)} & H_{N,3}^{(2)} & \cdots & H_{N,N}^{(2)}
\end{pmatrix} = \begin{pmatrix}
H_{11}^{(1)} + H_{13}G_{33}^{(1)}H_{31}^{(1)} & H_{13}G_{33}^{(1)}H_{34} & 0 & \cdots & 0 \\
H_{43}G_{33}^{(1)}H_{31}^{(1)} & H_{44} + H_{43}G_{33}^{(1)}H_{34} & H_{45} & 0 & \vdots \\
0 & H_{45} & H_{55} & \ddots & \vdots \\
\vdots & 0 & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & H_{N,N}
\end{pmatrix}$$
Appendix A. Decimation techniques

with \( G^{(1)}_{33} = (E + i \eta - H^{(1)}_{33})^{-1} \), and so on. Notice that with each decimation step, only the cells that were in direct contact with the decimated cell are affected.

To simplify the notation, we introduce:

\[
H^{(n)}_{LL} := H^{(n)}_{11}, \\
H^{(n)}_{LR} := H^{(n)}_{1,n+2}, \\
H^{(n)}_{RL} := H^{(n)}_{n+1,1}, \\
H^{(n)}_{RR} := H^{(n)}_{n+2,n+2}.
\]

Now, with \( H^{(0)}_{ij} = H_{ij} \), we can run an iteration over \( n = 0,1,\ldots,N-2 \) following the steps:

\[
G^{(n)}_{XX} := \left( E + i \eta - H^{(n)}_{RR} \right)^{-1}, \\
H^{(n+1)}_{LL} = H^{(n)}_{LL} + H^{(n)}_{LR} G^{(n)}_{XX} H^{(n)}_{RL}, \\
H^{(n+1)}_{LR} = H^{(n)}_{LR} G^{(n)}_{XX} H^{(0)}_{n+1,n+2}, \\
H^{(n+1)}_{RL} = H^{(0)}_{n+2,n+1} G^{(n)}_{XX} H^{(n)}_{RL}, \\
H^{(n+1)}_{RR} = H^{(0)}_{n+2,n+2} + H^{(0)}_{n+2,n+1} G^{(n)}_{XX} H^{(0)}_{n+1,n+2},
\]

and arrive at the desired quantities:

\[
\begin{pmatrix}
H^{(N-2)}_{LL} & H^{(N-2)}_{LR} \\
H^{(N-2)}_{RL} & H^{(N-2)}_{RR}
\end{pmatrix}
= 
\begin{pmatrix}
H_{LL}^{(N)} & H_{LR}^{(N)} \\
H_{RL}^{(N)} & H_{RR}^{(N)}
\end{pmatrix}
\]

A simple implementation of this algorithm is given in App. C.6.

A.3.1. Thoughts about efficiency

The straightforward way to obtain the quantities \( G_{11}(E) \), \( G_{1,N}(E) \), \( G_{N,1}(E) \) and \( G_{N,N}(E) \) would be to invert the matrix \( E + i \eta - \mathcal{H} \) completely and read out the data from the corners. The general matrix inversion itself, however, is a order \( N^3 \) operation in time and \( N^2 \) in memory. The algorithm above, on the other hand is order \( N \) in both time and memory. For any system that can be split into more than a few cells, this will give a gain in speed and memory consumption that easily outweighs the slightly increased complexity.

It has been tried before to take the decimation method one step further and split the system into individual atoms, decimating one atom after
the other. This, however, will only bring additional gain if the $H_{ij}^{(0)}$ block matrices themselves are sparse, i.e., if there is low connectivity inside each cell. For dense block matrices, it is generally more efficient to construct the block matrices once and let some optimized library do the inversion.

A.4. Periodic systems

A periodic quasi-1D system with a localized basis has an infinite Hamiltonian of the form:

$$
\mathcal{H} = \begin{pmatrix}
\ddots & \ddots & \ddots & \ddots & \ddots & \cdots & 0 \\
0 & H_{10} & H_{00} & H_{01} & 0 & \vdots \\
\vdots & 0 & H_{10} & H_{00} & H_{01} & 0 \\
0 & \cdots & \ddots & \ddots & \ddots & \cdots & \ddots
\end{pmatrix}
$$

cutting this system in halves, we get two semi-infinite surface systems:

$$
\mathcal{H}_L = \begin{pmatrix}
\ddots & \ddots & \ddots & \ddots & \ddots & \cdots & 0 \\
0 & H_{10} & H_{00} & H_{01} & 0 & \vdots \\
\vdots & 0 & H_{10} & H_{00} & H_{01} & 0 \\
0 & \cdots & \ddots & \ddots & \ddots & \cdots & \ddots
\end{pmatrix}
$$

$$
\mathcal{H}_R = \begin{pmatrix}
H_{00} & H_{01} & 0 & \cdots & 0 \\
H_{10} & H_{00} & H_{01} & 0 & \vdots \\
0 & H_{10} & H_{00} & H_{01} & 0 \\
0 & \cdots & \ddots & \ddots & \ddots
\end{pmatrix}
$$

As always, the Green functions have the same dimensionalities as the corresponding Hamiltonian, but now, the entries in the bulk and at the
surfaces are different:

\[
\mathcal{G}(E) = (E + i\eta - \mathcal{H})^{-1} = \\
\begin{pmatrix}
\cdots & G^\text{bulk}_{20} & G^\text{bulk}_{10} & G^\text{bulk}_{00} & G^\text{bulk}_{01} & G^\text{bulk}_{02} & \cdots \\
\cdots & G^\text{bulk}_{10} & G^\text{bulk}_{00} & G^\text{bulk}_{01} & G^\text{bulk}_{02} & \cdots \\
\cdots & G^\text{bulk}_{00} & G^\text{bulk}_{01} & G^\text{bulk}_{02} & \cdots \\
\end{pmatrix}
\]

\[
\mathcal{G}_L(E) = (E + i\eta - \mathcal{H}_L)^{-1} = \\
\begin{pmatrix}
\cdots & G^L_{-1,-2} & G^L_{-1,-1} & G^L_{-1,0} & \cdots \\
\cdots & G^L_{0,-2} & G^L_{0,-1} & G^L_{0,0} & \cdots \\
\end{pmatrix}
\]

\[
\mathcal{G}_R(E) = (E + i\eta - \mathcal{H}_R)^{-1} = \\
\begin{pmatrix}
G^R_{00} & G^R_{01} & G^R_{02} & \cdots \\
G^R_{10} & G^R_{11} & G^R_{12} & \cdots \\
\end{pmatrix}
\]

The most basic quantities of interest here are the bulk Green function \(G^\text{bulk}(E) := G^\text{bulk}_{00}(E)\) and the surface Green functions \(G^s_L(E) := G^s_{00}(E)\) and \(G^s_R(E) := G^s_{00}(E)\).

### A.4.1. Simple iterative scheme

To calculate the surface Green functions in a straightforward way, one can start by writing the left surface Hamiltonian as:

\[
\mathcal{H}_L = \\
\begin{pmatrix}
\mathcal{H} & 0 \\
0 & H_{01} \\
\cdots & 0 & H_{10} \\
\end{pmatrix}
\]

this is a bipartite system as used before. Thus, we can use \(G^s_L(E) = (E + i\eta - H^{\text{eff}}_{00}(E))^{-1}\) with

\[
H^{\text{eff}}_{00}(E) = H_{00} + \left( \begin{array}{ccc} \cdots & 0 & H_{10} \\ \cdots & 0 & H_{10} \end{array} \right) (E + i\eta - \mathcal{H}_L)^{-1} \begin{pmatrix}
\cdots \\
0 \\
H_{01} \\
\end{pmatrix}
\]

\[
= H_{00} + H_{10} G^L_{00}(E) H_{01}
\]

This gives us recursive definition of the Green functions:

\[
G^L_{s}(E) = (E + i\eta - (H_{00} + H_{10} G^L_{s}(E) H_{01}))^{-1}
\]

\[
G^R_{s}(E) = (E + i\eta - (H_{00} + H_{01} G^R_{s}(E) H_{10}))^{-1}
\]
A.4. Periodic systems

which can be directly implemented numerically. This algorithm is easy
to implement and fairly robust, but it can take several thousand steps for
iteration.

A.4.2. Renormalization-decimation algorithm

A slightly more complex but far more efficient algorithm for calculating
not only the surface but also the bulk Green functions was introduced
by M. P. López Sancho, J. M. López Sancho and J. Rubio [167] and will
be described here in our notation. Equivalent algorithms were also pro-
posed by F. Guinea, C. Tejedor, F. Flores and E. Louis [101] as well as
M. B. Nardelli [188], who also first applied it to carbon nanotubes.

Consider an infinite block tridiagonal Hamiltonian:

\[
\mathcal{H} = \begin{pmatrix}
\cdots & \cdots & \cdots & 0 \\
\ldots & H_{0,-1} & H_{00} & H_{01} \\
& H_{10} & H_{11} & H_{12} \\
0 & H_{21} & H_{22} & H_{23} \\
\end{pmatrix}
\]

Decimating out each second cell leaves us with

\[
\mathcal{H}^{(1)} = \begin{pmatrix}
\cdots & \cdots & \cdots & 0 \\
\ldots & H_{0,-2}^{(1)} & H_{00}^{(1)} & H_{01}^{(1)} \\
& H_{20}^{(1)} & H_{22}^{(1)} & H_{24}^{(1)} \\
0 & \cdots & \cdots & \cdots \\
\end{pmatrix}
\]

with:

\[
H_{i,i}^{(1)} = H_{i,i}^{(0)} + H_{i-1,i-1}^{(0)} C_{i-1,i-1}^{(0)} H_{i-1,i}^{(0)} + H_{i,i+1}^{(0)} C_{i+1,i+1}^{(0)} H_{i+1,i}^{(0)}
\]

\[
H_{i,i+2}^{(1)} = H_{i,i+1}^{(0)} C_{i+1,i+1}^{(0)} H_{i+1,i+2}^{(0)}
\]

repeating this procedure gives:

\[
\mathcal{H}^{(2)} = \begin{pmatrix}
\cdots & \cdots & \cdots & 0 \\
\ldots & H_{0,-4}^{(2)} & H_{00}^{(2)} & H_{04}^{(2)} \\
& H_{40}^{(2)} & H_{44}^{(2)} & H_{48}^{(2)} \\
0 & \cdots & \cdots & \cdots \\
\end{pmatrix}
\]
Appendix A. Decimation techniques

with:

\[ H_{i,i}^{(2)} = H_{i,i}^{(1)} + H_{i,i-2}^{(1)} G_{i-2,i-2}^{(1)} H_{i-2,i}^{(1)} + H_{i,i+2}^{(1)} G_{i+2,i+2}^{(1)} H_{i+2,i}^{(1)} \]

\[ H_{i,i+4}^{(2)} = H_{i,i+2}^{(1)} G_{i+2,i+2}^{(1)} H_{i+2,i+4}^{(1)} \]

To write this more compactly, we define:

\[ \epsilon^{(n)} := E + i\eta - H_{00}^{(n)} \]

\[ \alpha^{(n)} := -H_{02}^{(n)} \]

\[ \beta^{(n)} := -H_{20}^{(n)} \]

and remember that we wanted to calculate a periodic system, so

\[ H_{2e,2e}^{(n)} = H_{00}^{(n)} \]

\[ H_{2e,i+12e}^{(n)} = H_{02e}^{(n)} \]

\[ H_{i+12e,2e}^{(n)} = H_{2e,0}^{(n)} \]

for any \( i \in \mathbb{Z} \).

The iteration rules above now become:

\[ \gamma^{(n)} := (\epsilon^{(n)})^{-1} \]

\[ \epsilon^{(n+1)} = \epsilon^{(n)} - \beta^{(n)} \gamma^{(n)} \alpha^{(n)} - \alpha^{(n)} \gamma^{(n)} \beta^{(n)} \]

\[ \alpha^{(n+1)} = -\alpha^{(n)} \gamma^{(n)} \alpha^{(n)} \]

\[ \beta^{(n+1)} = -\beta^{(n)} \gamma^{(n)} \beta^{(n)} \]

For a semi-infinite system, we can exploit the fact that after any iteration in the whole effective Hamiltonian only the surface element itself feels the presence of the surface:

\[ \mathcal{H}_R^{(1)} = \begin{pmatrix} H_{00}^{R(1)} & H_{02}^{R(1)} & H_{20}^{R(1)} & H_{22}^{R(1)} & \cdots \\ H_{20}^{R(1)} & H_{22}^{R(1)} & H_{24}^{R(1)} & \cdots & \cdots \end{pmatrix} \]

so we can define

\[ \epsilon_L^{(n)} = E + i\eta - H_{00}^{L(n)} \]

\[ \epsilon_R^{(n)} = E + i\eta - H_{00}^{R(n)} \]

and find two additional iteration rules:

\[ \epsilon_L^{(n+1)} = \epsilon_L^{(n)} - \beta^{(n)} \gamma^{(n)} \alpha^{(n)} \]

\[ \epsilon_R^{(n+1)} = \epsilon_R^{(n)} - \alpha^{(n)} \gamma^{(n)} \beta^{(n)} \]
Now, it turns out that $\alpha^{(n)}$ and $\beta^{(n)}$ converge against zero for almost all initial conditions. If this is the case, $\varepsilon^{(n)}$, $\varepsilon^{(n)}_L$ and $\varepsilon^{(n)}_R$ also converge towards values $\varepsilon$, $\varepsilon_L$, and $\varepsilon_R$, which we can use to retrieve the desired quantities:

\[
G_{\text{bulk}}(E) = \varepsilon^{-1}
\]
\[
G_{s}^L(E) = \varepsilon^{-1}_L
\]
\[
G_{s}^R(E) = \varepsilon^{-1}_R
\]

There are certain conditions when this convergence fails, usually near eigenenergies of $H_{00}$ for $\eta$ chosen too small: if $\varepsilon^{(0)}$ is badly conditioned, the first inversion creates a large error which is carried through with every iteration. Within physically relevant precision, however, the algorithm gives efficient and reliable results. A sample implementation of this algorithm is given in App. C.7.

**A.4.3. Umerski algorithm**

It should be mentioned for completeness, that there is a closed mathematical expression that does, in principle, allow the computation of the surface greens functions in one single step [265]. A. Umerski, the author, claims this method to be superior in precision and efficiency to any known iterative scheme.

The main drawback of this scheme is that it depends on the invertibility of $H_{01}$, which is not given for many of the systems we use. One can always work around this limitation by first decimating out parts of the unit cell, but to do so in full generality turns out to lead to severe numerical instabilities. It is, however, a powerful method that may be worth further investigation depending on the problem at hand.
Appendix B.

Analytic derivations

B.1. Supersymmetric spectrum of graphene

The linearized Hamiltonian of graphene near the Fermi energy can be expressed formally equivalent to a relativistic Dirac particle in two dimensions. This special form leads to a very special spectrum that can be found by an analysis based on supersymmetry as it was given by M. Ezawa [81, 82]. An excellent review of supersymmetry in the perspective of condensed matter physics is given by G. Junker [133]. In the following, a brief deduction of the arguments will be given:

The Dirac Hamiltonian for the two cones, located in the points $K$ and $K'$ of the Brillouin zone, has the form

$$H_D = v_F (\alpha_x P_x + \alpha_y P_y)$$

$$P_i = -i\hbar \partial_i - eA_i$$

with

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} ; \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad \sigma_y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

where $A$ is the vector potential of a magnetic field $B = \nabla \times A$ perpendicular to the graphene plane. The Pauli matrices $\sigma_i$ refer to the pseudospin that originates from the two cones. The real electron spin is not considered here, as it couples to the magnetic field at a completely different scale.

In matrix form, the same Hamiltonian can be expressed as

$$H_D = \begin{pmatrix} 0 & Q \\ Q & 0 \end{pmatrix}$$

$$Q = v_F (\sigma_x P_x + \sigma_y P_y)$$

A unitary transform allows the separation of $H_D$ into two symmetric subspaces

$$UH_DU^* = \begin{pmatrix} Q & 0 \\ 0 & -Q \end{pmatrix} ; \quad U = \sqrt{\frac{1}{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$
Appendix B. Analytic derivations

So we see that the each energy level has a partner of opposite sign. With this knowledge, it is sufficient to obtain the spectrum of the square of the Dirac Hamiltonian to deduce the spectrum of the original system:

\[ \mathcal{H}_D^2 = \begin{pmatrix} Q^2 & 0 \\ 0 & Q^2 \end{pmatrix} \]

We therefore consider the quantity:

\[ \mathcal{H}_P = 2Q^2 = 2v_F^2 \left( p_x^2 + p_y^2 \right) + 2v_F e \hbar B_z \sigma_z \]

which has exactly the form of a Pauli Hamiltonian with a “mass” of \( m = 1/4v_F \) in two dimensions with a perpendicular magnetic field \( B_z \). As described in Ref. [133], this is supersymmetric with two different supercharges:

\[ Q^+ = v_F \left( \sigma_x p_x + \sigma_y p_y \right) ; \quad Q^- = v_F \left( \sigma_x p_y - \sigma_y p_x \right) \]

following the superalgebra:

\[ \{Q^+, Q^+\} = \mathcal{H}_P ; \quad \{Q^+, Q^-\} = 0 ; \quad [\mathcal{H}_P, Q^\pm] = 0 \]

Such a system generally has a Witten parity operator

\[ \mathcal{W} = -i \mathcal{H}_P^{-1} [Q^+, Q^-] \]

that is defined at least on the nonzero subspace of \( \mathcal{H}_P \) and obeys:

\[ [\mathcal{W}, \mathcal{H}_P] = 0 ; \quad [\mathcal{W}, Q^\pm] = 0 ; \quad \mathcal{W}^2 = 1 \]

The Witten parity is a symmetry of the Hamiltonian, so both operators have a common set of eigenvectors:

\[ \mathcal{H}_P |\Psi_\pm\rangle = E |\Psi_\pm\rangle ; \quad \mathcal{W} |\Psi_\pm\rangle = \pm |\Psi_\pm\rangle \]

And indeed, for any state of nonzero energy and positive parity, a negative state of the same energy can be constructed and vice versa:

\[ |\Psi_\pm E\rangle = \sqrt{\frac{E}{2}} Q^+ |\Psi_\pm E\rangle ; \quad |\Psi_\mp E\rangle = \sqrt{\frac{E}{2}} Q^- |\Psi_\mp E\rangle \]

So there is a one-to-one correspondence that guarantees a two-fold degeneracy of any non-zero energy eigenvalue.

Note that the two subspaces of positive and negative sign are sometimes referred to as bosonic and fermionic sector. These rather misleading
B.1. Supersymmetric spectrum of graphene

terms originate from supersymmetric quantum field theories but do not have any clear correspondence in this single-particle quantum mechanical model, where the two sectors are simply orthogonal subspaces of the same Hilbert space.

To this point, nothing is known about the existence and degeneracy of \( E = 0 \) eigenstates of either parity which does not follow from the supersymmetric algebra on its own. In fact, the question of existence of zero energy eigenvalues is equivalent with the question whether the supersymmetry is a good symmetry of the system.

For the special case of the two-dimensional Pauli Hamiltonian, however, several additional points can be shown:

- One finds the Witten parity operator to be \( \mathcal{W} = \sigma_3 \) which is defined on the whole Hilbert space. This leads to the physical interpretation that all non-zero energy levels are pseudospin-degenerate and only for energy zero, the degeneracy is yet unknown.

- Y. Aharonov and A. Casher have shown that the supersymmetry is always a good symmetry of this special system if the total magnetic flux \( \Phi \) through the \( x\)-\( y \) plane is larger than one flux quantum \( \Phi_0 \) [7]. More precisely, for a finite total flux, the degeneracy of the zero-energy level is given as

\[
d = \left\lfloor \frac{\Phi}{\Phi_0} \right\rfloor
\]

where \( \lfloor \cdot \rfloor \) denotes the integer floor. Depending on the sign of the total flux, the zero-energy states are all either pseudospin-up or pseudospin-down (i.e. positive or negative Witten parity).

- For a homogeneous, positive magnetic field \( B_z \), the spectrum can be derived as follows. The supercharges can be expressed as:

\[
Q^+ = \begin{pmatrix} 0 & \mathcal{A} \\ \mathcal{A}^\dagger & 0 \end{pmatrix} \quad ; \quad Q^- = i \begin{pmatrix} 0 & -\mathcal{A} \\ \mathcal{A}^\dagger & 0 \end{pmatrix}
\]

with \( \mathcal{A} = v_F (P_x - i P_y) \). The Pauli Hamiltonian is then

\[
\mathcal{H}_0 = 2 Q^+ Q^+ = 2 Q^- Q^- = \begin{pmatrix} 2 \mathcal{A}^\dagger \mathcal{A} & 0 \\ 0 & 2 \mathcal{A} \mathcal{A}^\dagger \end{pmatrix} = \begin{pmatrix} \mathcal{H}_0^\dagger & 0 \\ 0 & \mathcal{H}_0 \end{pmatrix}
\]
Appendix B. Analytic derivations

with the commutator \([\mathcal{A}, \mathcal{A}^\dagger] = 2v_F^2 e \hbar B_z = (\hbar \omega_c)^2\). Defining \(a = \mathcal{A}/(\hbar \omega_c)\), we get:

\[
\mathcal{H}^\dagger = 2(\hbar \omega_c)^2 a^\dagger a; \quad \mathcal{H}^\dagger = 2(\hbar \omega_c)^2 aa^\dagger
\]

with \([a, a^\dagger] = 1\), so we can see the spectrum contains the twice degenerate values \(E_n^{\uparrow\downarrow} = 2(\hbar \omega_c)^2 n\) for \(n = 1, 2, \ldots\) and the additional nondegenerate value \(E_0^\dagger = 0\). In the case that \(B_z\) is negative, the signs have to be reversed and the nondegenerate zero eigenstate has negative parity (i.e. pseudospin) \(E_0^\dagger = 0\).

Knowing the spectrum of \(\mathcal{H}_F = 2Q^2\), we can now directly deduce that of \(\mathcal{H}_D = \text{diag}(Q, -Q)\): There are the relativistic Landau levels (ReLL)

\[
E_{n\uparrow\downarrow}^{\dagger} = \pm v_F \sqrt{2e \hbar B_z n}
\]

with \(n \in \{1, 2, \ldots\}\). The index \(\uparrow\) or \(\downarrow\) indicates that each is twice pseudospin degenerate. The name ReLL distinguishes them from the conventional Landau levels of the form \(E_n = \hbar e B_z (2n + 1)/2m^*\) for particles with effective mass \(m^*\). In addition there is a special supersymmetric Landau level (SuSyLL) \(E_0 = 0\) which is also twice degenerate, but with both pseudospins having the same sign, depending on the sign of the magnetic field \(B_z\).

Taking into account the regular electron spin, each Landau level in the graphene monolayer becomes four-fold degenerate.

B.2. The linear-chain model of extended contacts

B.2.1. Transmission calculations

The Hamiltonian of a two-probe system for transport calculations is given by:

\[
\mathcal{H} = \begin{pmatrix}
H_L & H_{Lc} & 0 \\
H_{cL} & H_c & H_{cR} \\
0 & H_{Rc} & H_R
\end{pmatrix}
\] (B.1)

where \(H_c\) describes the finite-size conductor region and \(H_{L/R}\) describes the leads, which are connected to independent reservoirs and have no direct contact with each other. From \(\mathcal{H} = \mathcal{H}^\dagger\) it follows that \(H_{Lc} = H_{cL}^\dagger\) and \(H_{Rc} = H_{cR}^\dagger\).
To simplify the notation, we first define the complex-energy Green function:
\[
\mathcal{G}(\mathcal{E}) = (\mathcal{E} - \mathcal{H})^{-1}
\]
and derive expressions for the retarded and advanced Green functions from it \((\mathcal{E} = E \pm i\eta)\):
\[
\mathcal{G}^r(E) = \lim_{\eta \to 0^+} \mathcal{G}(E + i\eta)
\]
\[
\mathcal{G}^a(E) = \lim_{\eta \to 0^+} \mathcal{G}(E - i\eta)
\]
The transmission through this system is given by [87, 64]
\[
T = \text{Tr} \left\{ \Gamma_L \mathcal{G}_c^r \Gamma_R \mathcal{G}_c^a \right\}
\tag{B.2}
\]
with
\[
\mathcal{G}_c = (\mathcal{E} - H_c - \Sigma_L - \Sigma_R)^{-1}
\]
\[
\Sigma_\alpha = H_{c,\alpha} \mathcal{G}_c H_{\alpha,c} \quad \alpha = L, R
\]
\[
\Gamma_\alpha = i \left( \Sigma_\alpha^r - \Sigma_\alpha^a \right)
\]
\[
\mathcal{G}_\alpha = (\mathcal{E} - H_\alpha)^{-1}
\]
The Hamiltonian of the model at hand, depicted in Fig. 4.6, can be split up according to Eq. (B.1): the conductor consists of just one atom, so its Hamiltonian is a 1 \times 1-matrix \(H_c = (\varepsilon)\). The left lead is a semi-infinite chain, contacted only at the last atom:
\[
H_{L} = \begin{pmatrix}
\vdots & \vdots & \vdots & \ddots & 0 \\
\cdots & 0 & -\gamma & \varepsilon & -\gamma \\
\cdots & 0 & -\gamma & \varepsilon & \ddots \\
\end{pmatrix}_{\infty \times \infty}
\]
\[
H_{cl} = \begin{pmatrix}
\varepsilon & \cdots & 0 & 0 & -\gamma \\
\cdots & \varepsilon & \cdots & \ddots & \ddots \\
\cdots & \cdots & \varepsilon & \cdots & \ddots \\
\end{pmatrix}_{1 \times \infty}
\]
The right lead consists of a chain of \(N\) atoms, each attached to a wide-band lead. This can be captured by defining an effective Hamiltonian of the form
\[
H_{R}^{\text{eff}} = \begin{pmatrix}
\varepsilon - \frac{i\Delta}{2} & -\gamma & \cdots & 0 \\
-\gamma & \varepsilon - \frac{i\Delta}{2} & \cdots & \vdots \\
-\gamma & \cdots & -\gamma & \ddots \\
\vdots & \cdots & \varepsilon - \frac{i\Delta}{2} & -\gamma \\
0 & \cdots & -\gamma & \varepsilon - \frac{i\Delta}{2}
\end{pmatrix}_{N \times N}
\tag{B.3}
\]

175
Appendix B. Analytic derivations

together with a contact point in the first atom only:

$$H_{cR} = \begin{pmatrix} -\gamma & 0 & 0 & \cdots & 0 \end{pmatrix}_{1 \times N}$$

Note, that $H_{R}^{\text{eff}}$ is the effective Hamiltonian containing the retarded self energy, so $G_{R}^{0} = (E + i0^{+} - H_{R}^{\text{eff}})^{-1}$ and $G_{R} = (E - i0^{+} - (H_{R}^{\text{eff}})^{\dagger})^{-1}$.

In the following, we will simplify the notation by setting $\gamma = 1$ and $\varepsilon = 0$. Both constants can be reintroduced in the final result Eq. (B.8) by substituting $E \rightarrow (E - \varepsilon) / \gamma$.

B.2.2. Inverse based on Chebyshev polynomials

As a starting point for a full analytical solution, we look at a finite linear chain of length $N$, which has the Hamiltonian:

$$H^{N} = \begin{pmatrix} 0 & -1 & \cdots & 0 \\ -1 & 0 & \ddots & \vdots \\ \vdots & \ddots & -1 & 0 \\ 0 & \cdots & -1 & 0 \end{pmatrix}_{N \times N}$$

The quantity of interest of this system is the 1,1-matrix-element of the Green function $G^{N}(E) = (E - H^{N})^{-1}$. The solution is based on the Chebyshev polynomials of the second kind [98] which can be defined via the determinant identity:

$$U_{n}(x) = \det \begin{pmatrix} 2x & 1 & 0 \\ 1 & 2x & \ddots \\ \vdots & \ddots & 1 \\ 0 & 1 & 2x \end{pmatrix}_{n \times n}$$

or, equivalently by the recursive definition:

$$U_{0}(x) = 1$$
$$U_{1}(x) = 2x$$
$$U_{n+1}(x) = 2xU_{n}(x) - U_{n-1}(x)$$  \hspace{1cm} (B.4)
B.2. The linear-chain model of extended contacts

We can now use the well-known identity for the matrix inverse:

\[
(A^{-1})_{ij} = \frac{1}{\det(A)} \det \begin{pmatrix}
A_{1,1} & \cdots & A_{1,j-1} & 0 & A_{1,j+1} & \cdots & A_{1,N} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
A_{i-1,1} & \cdots & A_{i-1,j-1} & 0 & A_{i-1,j+1} & \cdots & A_{i-1,N} \\
0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
A_{i+1,1} & \cdots & A_{i+1,j-1} & 0 & A_{i+1,j+1} & \cdots & A_{i+1,N} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
A_{N,1} & \cdots & A_{N,j-1} & 0 & A_{N,j+1} & \cdots & A_{N,N}
\end{pmatrix}
\]

to find

\[
(G^N(E))_{i,j} = \left( (E - H_N)^{-1} \right)_{i,j} \\
= (G^N(E))_{ji} \\
= \frac{i < j}{(-1)^{i-j} U_{i-1} (E/2) U_{N-j} (E/2)} \\
= \frac{U_{N-1} (E/2)}{U_N (E/2)} \\
= f_N (E/2).
\]

and specifically

\[
(G^N(E))_{1,1} = \frac{U_{N-1} (E/2)}{U_N (E/2)} =: f_N (E/2).
\]  

(B.5)

B.2.3. Surface of semi-infinite linear chain

The surface Green function of a semi-infinite linear chain can be defined as

\[
G_s (E) = \lim_{N \to \infty} (G^N(E))_{1,1} \\
= \lim_{N \to \infty} f_N (E/2) \\
=: f_\infty (E/2).
\]

To find an expression for \( f_\infty (x) \), we can use the recursive definition of the Chebyshev polynomials Eq. (B.4) and obtain

\[
f_N (x) = (2x - f_{N-1} (x))^{-1}.
\]

For \( N \to \infty \), this becomes

\[
f_\infty (x) = (2x - f_\infty (x))^{-1}.
\]
which has two solutions $f_{\infty}(x) = x \left(1 \pm \sqrt{1 - 1/x^2}\right)$. On the real axis, it follows from Eq. (B.4) by induction that $|f_N(x)| < 1$ when $|x| > 1$, so that we can select the correct solution

$$f_{\infty}(x) = x \left(1 - \sqrt{1 - 1/x^2}\right), \quad (B.6)$$

which can be continued analytically to $x \in \mathbb{C} \setminus (-1, 1)$ by reading the square root of a complex number as the principal square root, uniquely defined everywhere except on the negative real axis by the condition $\text{Re} \left(\sqrt{x}\right) \geq 0 \quad \forall x \in \mathbb{C}$.

The retarded surface Green function follows as

$$\mathcal{G}_r^s(E) = \lim_{\eta \to 0^+} \mathcal{G}_s(E + i\eta)$$

$$= \begin{cases} 
  E/2 - \sqrt{E^2/2 - 1} & \text{for } |E| \geq 2 \\
  E/2 - i \sqrt{1 - E^2/2} & \text{for } |E| \leq 2 
\end{cases} \quad (B.7)$$

### B.2.4. Transmission of the model system

With these results, we can now obtain the quantum mechanical transmission of our model system. The left lead is a semi-infinite chain giving a self energy of

$$\Sigma_L = H_{cL} (E - H_L)^{-1} H_{LC} = \mathcal{G}_s(E)$$

For $|E| \geq 2$, $\mathcal{G}_s(E)$ is real [see Eq. (B.7)], so $\Gamma_L$ and with it by Eq. (B.2) the whole transmission $T$ are strictly zero. In the following, we therefore assume $|E| < 2$ and select the second case in Eq. (B.7):

$$\Sigma_L^t = \frac{E}{2} \left(1 - i \sqrt{4/E^2 - 1}\right)$$

To find the self-energy of the right lead, we use the definition of $H_{R}^{\text{eff}}$ from Eq. (B.3) and find:

$$\Sigma_R^t = H_{cR} (E - H^N + i\Delta/2)^{-1} H_{Re}$$

$$= \left(\mathcal{G}_N(E + i\Delta/2)\right)_{1,1}$$

$$= f_N(E/2 + i\Delta/4)$$

178
Now we can put together all parts to calculate the transmission:

\[ G_r^c = (E - \Sigma_L^r - \Sigma_R^r)^{-1} \]
\[ = 2 \left( E + iE \sqrt{4/E^2 - 1} - 2f_N(E/2 + i\Delta/4) \right)^{-1} \]
\[ G_a^c = 2 \left( E - iE \sqrt{4/E^2 - 1} - 2f_N(E/2 - i\Delta/4) \right)^{-1} \]
\[ \Gamma_L = i \left( \Sigma_L^a - \Sigma_R^a \right) \]
\[ = E \sqrt{4/E^2 - 1} \]
\[ \Gamma_R = i \left( f_N(E/2 + i\Delta/4) - f_N(E/2 - i\Delta/4) \right) \]
\[ = 2 \text{Im} \left( f_N(E/2 - i\Delta/4) \right) \]
\[ T = \frac{8 \sqrt{4 - E^2} \text{Im} \left( f_N(E/2 - i\Delta/4) \right)}{\left| E - i \sqrt{4 - E^2} - 2f_N(E/2 - i\Delta/4) \right|^2} \]

(B.8)

### B.3. The elastic mean free path in carbon nanotubes

Starting point for the derivation of the elastic mean free path \( \ell_{\text{el}} \) is the Fermi golden rule, stating the rate of scattering from one state to another by a perturbation of the Hamiltonian:

\[ \tau^{-1}_{E,s \rightarrow d} = \frac{2\pi}{\hbar} g(E,d) \left| \langle E,d|\mathcal{H}'|E,s \rangle \right|^2 \]

where \( g(E,d) \) is the partial DOS of a certain band \( d \) present at energy \( E \).

As Anderson disorder defines a statistical ensemble for the actual value of \( \mathcal{H}' \), the scattering rate is given as the ensemble average:

\[ \tau^{-1}_{E,s \rightarrow d} = \frac{2\pi}{\hbar} g(E,d) \int d^4W \rho(W) \left| \langle E,d|W|E,s \rangle \right|^2 \]

In a local basis \( |i\rangle \), Anderson disorder is diagonal:

\[ \tau^{-1}_{E,s \rightarrow d} = \frac{2\pi}{\hbar} g(E,d) \int d^4W \rho(W) \left| \sum_{i=1}^{N_{\text{atoms}}} \langle E,d|i\rangle W_i \langle i|E,s \rangle \right|^2 \]
Appendix B. Analytic derivations

This integral of the form \( \int dW |\sum a_i W_i|^2 \) can be evaluated using the identities for the distribution \( P(W) \):

\[
\int dW P(W) \left| \sum a_i W_i \right|^2 = \int dW P(W) \left( \sum a_i W_i \right) \left( \sum a_j^* W_j \right) = \sigma^2 \sum_i a_i \sum_j a_j^* \delta_{ij} = \sigma^2 \sum_i |a_i|^2
\]

Due to the symmetry of a nanotube, a Bloch wave has an equal probability weight on each atom \( |\langle i|E,s \rangle|^2 = 1/N_{\text{atoms}} \), so the above expression can be fully evaluated:

\[
\tau_{E,s-d}^{-1} = \frac{2\pi}{\hbar} g(E,d) \sigma^2 \sum_{i=1}^{N_{\text{atoms}}} |\langle E,d|i \rangle \langle i|E,s \rangle|^2 = \frac{2\pi}{\hbar} g(E,d) \frac{\sigma^2}{N_{\text{atoms}}}
\]

The states \( s \) and \( d \) correspond to bands that are present at energy \( E \). For the transmission of a disordered system, only the backscattering events are important. The rate of backscattering events from a right-moving band \( s \) is then the sum of all rates of left-moving states \( d \):

\[
(\tau_{E,s}^{\text{back}})^{-1} = \frac{2\pi}{\hbar} \frac{\sigma^2}{N_{\text{atoms}}} \sum_{d\text{leftmv}} g(E,d)
\]

The mirror symmetry of a \((N,N)\)-nanotube guarantees that each left-moving state corresponds to one right-moving state of equal energy and velocity. The sum over the partial DOS of right-moving bands therefore is exactly half of the total DOS. The partial DOS of a single band \( d \) can be expressed as \( g(E,d) = L g_L(E,d) \) with the DOS per length of the conductor \( g_L(E,d) = \frac{1}{2\pi} \frac{dL}{dE} \).

In a \((N,N)\) carbon nanotube, each unit cell of length \( \ell_{\text{uc}} \) contains \( 4N \) atoms. Over a length \( L \), the total number of atoms is therefore given as \( N_{\text{atoms}} = 4NL/\ell_{\text{uc}} \):

\[
(\tau_{E,s}^{\text{back}})^{-1} = \frac{2\pi}{\hbar} \frac{\sigma^2}{4NL/\ell_{\text{uc}}} \frac{L g_L(E)}{2} = \frac{2\pi}{\hbar} \frac{\ell_{\text{uc}} \sigma^2}{8N} g_L(E)
\]
B.3. The elastic mean free path in carbon nanotubes

The mean free path of an electron in a certain band \(s\) is given by the mean time between backscattering events \(\tau_{E,s}^{\text{back}}\) and the band velocity:

\[
v_s = \frac{1}{\hbar} \frac{dE_s}{dk} = \left( \hbar g_L(E,d) \right)^{-1}
\]

leading to:

\[
\ell_{el} = v_s / \left( \frac{2\pi}{\hbar} \frac{\ell_{uc} \sigma_s^2}{8N} g_L(E) \right)
= \left( g_L(E,s) \left( 2\pi \right)^2 \frac{\ell_{uc} \sigma_s^2}{8N} g_L(E) \right)^{-1}
\]

To obtain an expression for the total elastic mean free path, we can go via its definition Eq. (5.1):

\[
T(E)\big|_{L<\ell_{el}} = N_{ch} \left( 1 - \frac{L}{\ell_{el}} \right)
= \sum_s N_{ch} \left( 1 - \frac{L}{\ell_{el}^s} \right)
= N_{ch} - L \sum_s \left( \ell_{el}^s \right)^{-1}
\]

\[
\ell_{el} = N_{ch} \left( \sum_s \left( \ell_{el}^s \right)^{-1} \right)^{-1}
= N_{ch} \left( \frac{1}{2} g_L(E) \left( 2\pi \right)^2 \frac{\ell_{uc} \sigma_s^2}{8N} g_L(E) \right)^{-1}
= \left( \frac{\ell_{uc} \sigma_s^2}{16NN_{ch}} \left( 2\pi g_L(E) \right)^2 \right)^{-1}
\]
Appendix B. Analytic derivations
Appendix C.

Numerical implementations

C.1. Programming language, libraries and tools

The numerical work presented in this thesis was done using Open Source tools exclusively:

- As a programming language, Python (http://www.python.org) by G. Rossum et al. was chosen due to its flexibility. Being an interpreted language, small scripts can be written with minimal overhead, allowing many ideas to be tried out. Yet, unlike other scripting languages, Python encourages the author to write clear and well-structured code that can be well maintained and reused.

- The library NumPy (http://www.numpy.org) by T. Oliphant et al. allows the very efficient and elegant handling of arrays of numerical data within Python. For many numerical problems, this approach completely eliminates the performance penalty of the interpreted language. Handled correctly, the bulk of the computations runs at the full speed of C or Fortran, with the user only writing Python code or, in special situations, small snippets of compiled code.

- The library SciPy (http://www.scipy.org) by E. Jones et al. contains a collection of numerical algorithms based on NumPy.

- The library PyTables (http://www.pytables.org) by F. Altet et al. was used for the storage of numerical data. This library gives highly efficient and elegant access to HDF5 data files, storing numerical data in full binary precision in flexible hierarchical structures. The library is fully integrated with NumPy.
Appendix C. Numerical implementations

- Plotting the data was mostly done using the Matplotlib library (http://matplotlib.sourceforge.net/). Using Python as scripting language, computation, postprocessing and plotting can be done without problems of interfacing.

Besides these libraries for the numerical work, the main tools used in the process of writing this thesis were Inkscape (http://www.inkscape.org/) for preparing the figures, Texmacs (www.texmacs.org) for authoring, Tellico (http://periapsis.org/tellico/) for the management of bibliographical data as well as \LaTeX for the final formatting.

C.2. Handling of physical units

The handling of physical units seems to be a trivial issue. Yet, without the proper care, it is a common cause of errors. A few simple rules have proven useful to avoid confusion:

- work consistently within SI
- use an explicit constant for each unit
- define a set of fundamental units with the numerical value 1 (e.g. the SI standard units any alternative set)
- define any derived units from these fundamental units
- do an explicit multiplication with the corresponding unit at any definition of constants or parameters in the source code
- for any input values, explicitly state the unit and multiply the value accordingly
- for any output values explicitly state the unit and divide the numerical value accordingly

the following code defines a number of important units and constants used later on.

```python
#!/usr/bin/env python
from math import pi

# numerical fundamental values
angstrom = 1.0
```

184
C.3. Constructing chiral carbon nanotubes

Though the conceptual and mathematical construction of chiral CNTs is fairly straightforward, an algorithmic determination of the atoms within one unit cell is a surprisingly complex task. A proper documentation of any algorithm could not be found in the literature. All implementations that could be found unnecessarily depend on floating point arithmetic and therefore bear the risk of rounding errors. In the following, I will describe my own, pure-integer algorithm to solve the problem in a clean way.

Assume a CNT with the chiral vector $M\mathbf{a}_1 + N\mathbf{a}_2$. Given the angle of $60^\circ$ between the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, a perpendicular lattice vector can always be constructed as $(M + 2N)\mathbf{a}_1 - (2M + N)\mathbf{a}_2$. This vector can still be divided by the greatest common divisor:

$$Q = \gcd(M + 2N, 2M + N),$$
leading to a new lattice vector:

\[
(M_\perp, N_\perp) = \left( \frac{M + 2N}{Q}, \frac{-2M + N}{Q} \right),
\]

that corresponds to the minimal periodic vector of the CNT. We can now determine the circumference and the length of the unit cell as:

\[
\ell_{\text{circ}} = a \sqrt{M^2 + N^2 + MN} \]
\[
\ell_{\text{uc}} = a \sqrt{M^2 \perp + N^2 \perp + M \perp N \perp}
= \sqrt{3} a \sqrt{M^2 + N^2 + MN/Q}. \tag{C.1}
\]

With each hexagonal plaquette having an area of \( A_{\text{plaquette}} = a^2 \sqrt{3}/4 \) and two atoms per plaquette, we can therefore find the number of atoms per unit cell by dividing the area of the tube surface per unit cell:

\[
N_{\text{atoms}} = \frac{2 \ell_{\text{circ}} \ell_{\text{uc}}}{A_{\text{plaquette}}}
= \frac{4 (M^2 + N^2 + MN)}{Q}.
\]

Relative to a plaquette that is chosen as origin, every plaquette of the graphene lattice can be identified by a pair of indices \((m, n)\). The chiral vector \((M, N)\) identifies plaquettes as \((m, n) \equiv (m + M, n + N)\), so it is sufficient to consider indices with \(0 \leq m + n < M + N\) to uniquely identify each plaquette along the periodic tube. New coordinates are introduced as:

\[
[i, j] = [m + n, m],
\]

with \(0 \leq i < M + N\). A plaquette \((m, n)\) is now considered part of the first unit cell, if

\[
0 \leq mN - nM < M_\perp N - N_\perp M
\]
\[
0 \leq jN - (i - j) M < M_\perp N - N_\perp M
\]
\[
iM \leq j (M + N) < iM + (M_\perp N - N_\perp M)
\]

For every \(i\), this can be turned into a pair of conditions for \(j\):

\[
\frac{iM}{M + N} \leq j < \frac{iM + M_\perp N - N_\perp M}{M + N}
\]

which can be computed exactly with pure integer operations. We can now count deterministically over the indices \([i, j]\) of all plaquettes in the
C.3. Constructing chiral carbon nanotubes

unit cell and create the plane coordinates of all atoms, two per plaquette as

\[ r_{[i,j]}^\pm = j a_1 + (i - j) a_2 \pm (a_1 + a_2)/6 \]

Finally, we can express the lattice vectors in cylindrical coordinates:

\[
\begin{align*}
Ma_1 + Na_2 &= (2\pi, 0) \\
M_\perp a_1 + N_\perp a_2 &= (0, \ell_{uc}) \\
a_1 &= (2\pi M_\perp - M_\ell_{uc}) / (M_\perp N - MN_\perp) \\
a_2 &= (2\pi N_\perp - M_\ell_{uc}) / (M_\perp N - MN_\perp)
\end{align*}
\]

leading to expressions for \((\phi, z)^\pm_{[i,j]}\) that can be used together with the tube radius \(\rho = \ell_{circ}/2\pi\) to express the coordinates of each atom in three dimensions.

```python
#!/usr/bin/env python

from numpy import *  # need at least NumPy 1.0, see http://www.numpy.org
from units import *

d_CC = 1.42 * angstrom

def gcd(a, b):
    if a > b:
        a, b = b, a
    while a != 0:
        a, b = b % a, a
    return b

def chiral(M, N):
    assert M >= 0 and N >= 0
    a = d_CC * 3**0.5  # length of a graphene lattice vector
    l_circ = a * (N*2 + M*2 + M*N)**0.5  # circumference of the tube
    l_uc = a * (3.0 * (M*2 + N*2 + M*N)**0.5 / M)  # length of one unit cell
    N_atoms = 4 * (N*2 + M*2 + M*N) / Q  # number of atoms per unit cell
    M_perp = (M*N) / Q  # periodic vector in lattice coordinates
    N_perp = - (M*M) / Q  # periodic vector in lattice coordinates

    # graphene lattice vectors in cylinder coordinates (phi,z):
    a_1 = array([2*pi*M_perp, l_uc*N]) / (M*N_perp - M_perp*M)
    a_2 = array([2*pi*N_perp, l_uc*M]) / (M*N_perp - N_perp*M)

    coords = []

    for i in range(0, N_atoms):
        # integer division always rounds to lower value (i.e.: (a/b)*b < a)
        # we need to round up, so we do -((-a)/b)
        j_min = -((-i*M)/M)
        j_max = -((-i*M + M_perp*N - M*N_perp)/(M*N))
        assert (j_min - 1) * (M*N) < i*M <= j_min * (M*N)
        assert (j_max - 1) * (M*N) < (i*M + M_perp*N - M*N_perp) <= j_max * (M*N)

        for j in range(j_min, j_max):
            for offset in [+ (a_1 + a_2)/6, - (a_1 + a_2)/6]:
                phi, z = j*a_1 + (i-j)*a_2 + offset
                x, y = rho*cos(phi), rho*sin(phi)
                coords += [(x, y, z)]
```

187
Appendix C. Numerical implementations

```python
assert len(coords) == N_atoms
return array(coords), l_uc

if __name__ == "__main__":
    coords, l_uc = chiral(7,5)
    # produce a sample file in cartesian xyz format
    xyFile = file("cnt-7-5.xyz", 'w')
    print >>xyFile, "%i AN" % (coords.shape[0])
    print >>xyFile, "xyz-file of (7,5)-CNT"
    for n in range(coords.shape[0]):
        print >>xyFile, "C %f %f %f" % (tuple(coords[n,:]))
    xyFile.close()
```

C.4. Construction of a periodic Hamiltonian

The periodic first-nearest-neighbor tight-binding Hamiltonian of the form

\[
H = \begin{pmatrix}
  \ddots & \ddots & \ddots & \cdots & 0 \\
  \vdots & H_1^T & H_0 & H_1 & \\
  \vdots & H_1^T & H_0 & H_1 & \\
  0 & \cdots & \ddots & \ddots & \ddots
\end{pmatrix}
\]

is constructed from the coordinates of a SWCNT obtained from the routine in the previous section by simply checking the distance for any pair of atoms and setting the corresponding matrix element to \(-\gamma\).
C.5. Tight-binding interlayer Hamiltonian

The following code snippet demonstrates the construction of the tight-binding Hamiltonian of a DWCNT based on the parameterization described in Eq. (2.1) as introduced by S. Roche et al. [224]. Due to the long cutoff chosen as \( r_{\text{cutoff}} = r_0 + 5\delta \), the Bloch Hamiltonian may contain nonzero-entries beyond neighboring unit cells:

\[
H = \begin{pmatrix}
  \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
  \cdots & 0 & H_N^1 & \cdots & H_1 & \cdots & H_0 & \cdots & 0 & \cdots \\
  \cdots & 0 & H_N^1 & \cdots & H_1 & \cdots & H_0 & \cdots & 0 & \cdots \\
  \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\end{pmatrix}
\]

which is respected by defining the \( k \)-dependent effective Hamiltonian as:

\[
H_{\text{eff}}(k) = H_0 + \sum_{n=1}^{N} e^{ik\ell_{\text{uc}}n} H_n^1 + e^{-ik\ell_{\text{uc}}n} H_n^\dagger
\]

The band structure obtained from this code can be found in Fig. 6.6, along with a discussion in Sec. 6.4.1.
Appendix C. Numerical implementations

def hopping(pos_a, pos_b):
    if abs(norm(pos_a[:2]) - norm(pos_b[:2])) > r_O:
        return 0.0
    elif abs(norm(pos_a[:2]) - norm(pos_b[:2])) < r_O + 0.001:
        if norm(pos_a - pos_b) < d_CC + 1.1:
            return -gamma
    else:
        r = norm(pos_b - pos_a)
        if r < cutoff:
            cos_theta = dot(pos_a[:2], pos_b[:2])/(norm(pos_a[:2])*norm(pos_b[:2]))
            return -beta * cos_theta * exp((-r_0 - r)/delta)
        return 0.0

def tight_binding_MWCNT(coords0, coords1=None):
    Natoms0 = len(coords0)
    if coords1 is None:
        H = matrix(zeros((Natoms0, Natoms0), 'D'))
        for i in range(Natoms0):
            for j in range(i+1, Natoms0):
                H[i,j] = hopping(coords0[i,:], coords0[j,:])
                H[j,i] = conj(H[i,j])
        else:
            Natoms1 = len(coords1)
            H = matrix(zeros((Natoms0, Natoms1), 'D'))
            for i in range(Natoms0):
                for j in range(Natoms1):
                    H[i,j] = hopping(coords0[i,:], coords1[j,:])
        return H

if __name__ == '__main__':
    from chiral import chiral
    import pylab

    # set up coordinates
    inner_coords.period = chiral(4, 4)
    outer_coords.period = chiral(9, 9)
    assert outer_coords.period == period
    coords = concatenate((inner_coords, outer_coords))

    # set up periodic Hamiltonian H[n]
    H = [ tight_binding_MWCNT(coords) ]
    for n in range(1, int(z_cutoff**2)):
        if abs(H[n] - H[n-1]) < 0.001:
            break
        H = [ H[n] for i in range(period)*2 ]

    # calculate band structure E[k, b]
    band_energies = zeros(len(ka), coords.shape[0])
    for k in range(len(ka)):
        H_eff = H[0] + exp(i*j*ka[k]) * H[1] + exp(-j*ka[k]) + H[1] * H[0]
        for n in range(2, len(H)):
            H_eff += exp(i*n*ka[k]) * H[n] + exp(-i*n*ka[k]) * H[n] + H[n] * H[0]
        band_energies[k, :n] = eigvalsh(H_eff)

    for e in range(band_energies.shape[1]):
        pylab.plot(ka, band_energies[:, e], 'b', alpha=0.2)
        pylab.xticks((-pi, pi, 2*pi), (r'$-\pi$','$\pi$','$2\pi$'))
        pylab.yticks((-N, 0, N), (r'$-N$','$0$','$N$'))
        pylab.ylim(-N, N)
        pylab.xlabel(r'${\vec k}$ (in units of $2\pi$)')
        pylab.ylabel(r'$E-E_F$ (in eV)')
        pylab.show()
C.6. Decimation of a finite block tridiagonal system

A finite block tridiagonal $N \times N$ Hamiltonian can be stored in two lists of block matrices, one of length $N$ for the diagonal blocks, the other of length $N - 1$ for the blocks on the superdiagonal. The inner blocks can be decimated using the algorithm described in App. A.3.

```python
#!/usr/bin/env python
from numpy import *

class block_tridiagonal:
    """a minimalistic implementation encapsulating
    a read-only hermitean block-tridiagonal matrix""
    def __init__(self, diag, superdiag):
        """diag: list of diagonal block matrices
        superdiag: list of block matrices on the first superdiag
        (first subdiag is deduced by hermiticity)"
        self.N = len(diag)
        assert len(superdiag) == self.N - 1
        self.diag = diag
        self.superdiag = superdiag
    def __getitem__(self, ij):
        i, j = ij
        if i == j:
            return asmatrix(self.diag[i])
        if i+1 == j:
            return asmatrix(self.superdiag[i])
        if i == j+1:
            return asmatrix(self.superdiag[j]).H
        raise KeyError()

def decimate_block_tridiagonal(E, H):
    assert imag(E) != 0
    assert isinstance(H, block_tridiagonal)
    H_00 = H[0,0] ; H_0n = H[0,1]
    H_n0 = H[1,0] ; H_nn = H[1,1]
    for n in range(2, H.N):
        G_XX = matrix(eye(len(H_nn)) * E - H_nn).I
        H_00 = H_00 + H_0n * G_XX * H_n0 ; H_0n = H_0n * G_XX + H[n-1,n]
        H_n0 = H[n,n-1] * G_XX + H_n0 ; H_nn = H[n,n] + H[n,n-1] * G_XX + H[n-1,n]
    H_eff = [[H_00, H_0n],[H_n0, H_nn]]
    return H_eff

if __name__ == '__main__':
    # simple test system: linear square ladder
    L, W = 100, 2
    H0 = eye(W,W) + eye(W,W-1)
    H1 = eye(W,W)
    H = block_tridiagonal([H0]*L, [H1]*(L-1))
    E = linspace(-4,4,101)
    LDOS = zeros(E.shape)
    for i in range(len(E)):
        H_eff = decimate_block_tridiagonal(E[i]+1j*eta, H)
        H_eff = bmat([H_eff, H_eff]) # fuse four block matrices together
        G = matrix(eye(len(H_eff))) / (E[i]+1j*eta) - H_eff.I
        LDOS[i] = -1/1j * 1/len(G)
    # plot local density of states at the ends
    import pylab
    pylab.plot(E,LDOS)
    pylab.show()
```

191
Appendix C. Numerical implementations

C.7. Renormalization decimation algorithm

The following is a compact implementation of the highly convergent algorithm for the computation of surface and bulk Green functions of periodic systems as described in App. A.4.2.

```python
#!/usr/bin/env python
from numpy import *
from numpy.linalg import norm
RD_MAXITER = 100
RD_CUTOFF = 1e-8

def renormalization_decimation(E, H_00, H_01):
    assert imag(E) != 0
    epsilon = asmatrix(eye(len(H_00)) * E - H_00)
    epsilon_L = epsilon
    epsilon_R = epsilon
    alpha = - asmatrix(H_01)
    beta = alpha.T
    countiter = 0
    while (norm(alpha) + norm(beta) > RD_CUTOFF):
        gamma = epsilon.T
        bga = beta * gamma * alpha
        agb = alpha * gamma * beta
        epsilon = epsilon - bga - agb
        epsilon_L = epsilon_L - bga
        epsilon_R = epsilon_R - agb
        alpha = - alpha * gamma * alpha
        beta = - beta * gamma * beta
        countiter += 1
        if countiter >= RD_MAXITER:
            raise Exception("Renormalization-decimation algorithm did not converge.")
    G_bulk = epsilon.T
    G_sL = epsilon_L.T
    G_sR = epsilon_R.T
    return (G_bulk, G_sL, G_sR)

if __name__ == "__main__":
    # simple test system: linear square ladder
    W = 2
    H0 = eye(W, W) + eye(W, W, -1)
    H1 = eye(W, W)
    eta = 1e-5 # chosen small enough to allow convergence
    E = linspace(-4, 4, 101)
    SDOS_bulk = zeros(E.shape)
    SDOS_surf = zeros(E.shape)
    for i in range(len(E)):
        G_bulk, G_sL, G_sR = renormalization_decimation(E[i] + 1j * eta, H0, H1)
        SDOS_bulk[i] = - 1/pi * imag(trace(G_bulk)) / W
        SDOS_surf[i] = - 1/pi * imag(trace(G_sL)) / W
    # plot local density of bulk and at surface (end) of system
    import pylab
    pylab.subplot(2,1,1); pylab.plot(E, SDOS_bulk)
    pylab.subplot(2,1,2); pylab.plot(E, SDOS_surf)
    pylab.show()
```

192
C.8. Periodic gauge

Consider the analytic expression of the periodic gauge as derived in Sec. 7.6:

\[ A(r) = \ell \Phi_0 \left( [r \cdot \tilde{a}_x] \tilde{a}_y - [r \cdot \tilde{a}_y] \delta([r \cdot \tilde{a}_x]) \tilde{a}_x \right) \]  
(C.2)

for which we need to evaluate the integral along a straight line between two arbitrary locations \(s\) and \(d\):  

\[ \varphi = \frac{1}{\Phi_0} \int_s^d dr \cdot A(r) \]

To do so, we split Eq. (C.2) into a sum of two parts:

\[ A_y(r) = \ell \Phi_0 [r_x] \tilde{a}_y \]
\[ A_x(r) = -\ell \Phi_0 [r_y] \delta([r_x]) \tilde{a}_x \]

with \(r_x = r \cdot \tilde{a}_x\) and \(r_y = r \cdot \tilde{a}_y\). The integral over \(A_y\) can be rewritten as:

\[ \varphi' = \ell \int_s^d dr \cdot \tilde{a}_y [r_x] \]
\[ = \ell \left( d_y - s_y \right) \left[ r_x \right] \]
\[ = \int_s^{d_x} dr_x \frac{[r_x]}{d_x - s_x} \]
\[ = \frac{[d_x] - [s_x] + [d_x]^2 - [s_x]^2}{2(d_x - s_x)} \]

where \([\cdot]\) denotes the integer floor (\(\lfloor x \rfloor \in \mathbb{Z}, x - 1 < \lfloor x \rfloor \leq x \forall x \in \mathbb{R}\)).

The integral over \(A_x\), expressed as a sum, gives the necessary correction:

\[ \varphi = \varphi' - \ell \int_s^d dr \cdot \tilde{a}_x [r_y] \delta([r_x]) \]
\[ = \begin{cases} 
\varphi' - \ell \sum_{x=[s_x]+1}^{[d_x]} \frac{x-s_x}{d_x-s_x} (d_y - s_y) & \text{for } s_x \leq d_x \\
\varphi' + \ell \sum_{x=[d_x]+1}^{[s_x]+1} \frac{x-s_x}{d_x-s_x} (d_y - s_y) & \text{for } s_x > d_x
\end{cases} \]
dx, dy = d

# periodic gauge:
if dx == sx:
    phase = (dy - sy) * (dx % 1)
else:
    si = int(sx // 1)
    sz = sx % 1
    di = int(dx // 1)
    dr = dx % 1

# first the part "forced" to be periodic
Aavg = 0.5 * (di - si + dx**2 - sz**2) / (dx - sx)
phase = (dy - sy) * Aavg

# then the correction
if di > si:
    for x in range(si+1, di+1):
        phase -= x * (dy-sy)/(dx-sx) + (dx*sy-sx*dy)/(dx-sx)
elif di < si:
    for x in range(di+1, si+1):
        phase += x * (dy-sy)/(dx-sx) + (dx*sy-sx*dy)/(dx-sx)

return 2*pi*phase%

def phase(rs, rd, ax, ay, l):
    sx = dot(rs, ax)
    sy = dot(rs, ay)
    dx = dot(rd, ax)
    dy = dot(rd, ay)

    return l*phase_in_lattice_coordinates((sx, sy), (dx, dy))
Appendix D.
Reprints of publications

Part of the work presented in this thesis has been published in the following international journal articles:


In the following, reprints of these articles are included for reference.
Spin transport in disordered single-wall carbon nanotubes contacted to ferromagnetic leads

S. Krompiewski*, 1, N. Nemec2, and G. Cuniberti2

1 Institute of Molecular Physics, Polish Academy of Sciences, M. Smoluchowskiego 17, 60-179 Poznan, Poland
2 Institute for Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany

Received 14 June 2005, revised 15 July 2005, accepted 25 July 2005
Published online 29 November 2005

PACS 72.25.Mk, 73.63.Fg, 75.47.De, 81.07.De, 85.35.Kt

Recent conductance measurements on multi-wall carbon nanotubes (CNTs) reveal an effective behavior similar to disordered single-wall CNTs. This is due to the fact that electric current flows essentially through the outermost shell and is strongly influenced by inhomogeneous electrostatic potential coming from the inner tubes. Here, we present theoretical studies of spin-dependent transport through disorder-free double-wall CNTs as well as single-wall CNTs with Anderson-type disorder. The CNTs are end-contacted to ferromagnetic electrodes modelled as fcc (111) surfaces. Our results shed additional light on the giant magnetoresistance effect in CNTs. Some reported results concern realistically long CNTs, up to several hundred nanometers.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1 Introduction

Over the last two decades, the magneto-electronics, based on all-metal multilayers, has proven to be very successful indeed [1]. The most important effect which should be invoked in this context is giant magnetoresistance (GMR) discovered in 1988 [2]. This effect makes it possible to control electric current flowing through magnetic materials by means of a magnetic field. In other words, the essence of the GMR effect lies in taking advantage of not just the electronic charge alone but also of its spin counterpart. Quite naturally researchers involved so far in physics of semiconductors, as well as those studying molecular systems have intensified their efforts in search for possible analogous effects in all-semiconducting [3] and/or hybrid systems (combinations among metals, semiconductors and molecules) [4, 5]. Consequently, a new field of science and technology has been triggered, under the name of spintronics [6, 7]. Here we report our results on the GMR effect in perfect and disordered carbon nanotubes sandwiched between ferromagnetic electrodes. There is no doubt nowadays that miniaturization requirements imposed on the emerging spintronics will be met by applying the so-called bottom-up approach as far as designing of new electronic devices is concerned. From this point of view carbon nanotubes are surely excellent candidates.

2 Double-wall CNTs

We start our studies with carbon nanotubes (CNTs) end-contacted to metal electrodes. Our present approach is essentially that described in detail in [8] with an improved simulation method of CNT/metal-electrode interface developed in [9]. Spin-polarization of the ferromagnetic leads is defined...
as \( P = 100(n_\sigma - n_{\bar{\sigma}})/(n_\sigma + n_{\bar{\sigma}}) \), where \( n_\sigma \) stands for a number of \( \sigma \)-spin electrons per lattice site. It is noteworthy that the structures in question are relaxed under the Lennard–Jones potential in order to find energetically favorable relative positions of CNTs' and electrodes' interface atoms. During the relaxation process the external electrodes are allowed to rotate and shift independently of each other, similarly the inner tube is also free to rotate. As regards the inter-tube hopping integrals, they are taken in the form as proposed in [10], i.e. set to \( t_{tt} = -(8/\cos \theta \exp[(d - b)/\delta]) \), where \( \theta \) is the angle between the \( \pi \) orbitals, \( d \) is a relative distance, \( t \) stands for the nearest neighbor hopping integral (chosen as energy unit), \( \delta = 0.45 \) Å and \( b = 3.34 \) Å. The GMR coefficient is defined in terms of the conductances, \( G \), as \( \text{GMR} = 1 - \frac{G_{\uparrow\uparrow}}{G_{\uparrow\downarrow}}, \) with \( \uparrow\downarrow \) denoting aligned (antialigned) magnetization orientation of the electrodes.

Most of the hitherto existing experiments on electronic transport suggest that current flowing through MWCNTs goes mostly through the outermost shell (see e.g. [11]). A precise role of the inner shells is still hardly known. Here we show the results on the GMR effect in two double-wall (DW) CNTs which have the same outer shell but differ – though non-conductive in each case – inner shells. Specifically the DWCNTs in question are: (i) the zigzag armchair (45-(5,0)@39-(8,8)) and (ii) the armchair armchair (38-(3,3)@39-(8,8)), using a short-hand notation \( L-(n, m) \), for the length (in carbon rings) and the chiral vector, respectively. In the former case the corresponding lengths are roughly the same (ca. 5 nm each) so both the inner tube and the outer one are well contacted to the magnetic electrodes. In the latter case, in turn, the inner shell is artificially shortened and forced thereby to be out of contact to the drain electrode. Figure 1 presents giant magnetoresistance for the two DWCNTs.

Despite the fact that both the systems are formally similar (at least in the presented “energy window” \((|E|/E_D| < 0.2)\), which falls into the zigzag-tube gap), the GMR curves are clearly different. We attribute these differences to intertube-quantum interfere nces which are present owing to the non-vanishing \( t_{\text{int}} \).

To mimic a possible effect of some additional disorder we present also GMR curves (thin lines) calcu-lated from \( \Delta E \)-averaged conductances (thin curves of the same style), where \( \Delta E \) is the inter-level spacing of the outer shell.

### 3 Single-wall CNTs with Anderson disorder

The simplest way to include the effect of disorder in a system described in terms of the tight-binding model is to allow all on-site (atomic) potentials to take random values within a given energy interval. Such an approach has been already applied to CNTs [12], but to our knowledge, only for tubes with non-magnetic leads. In the case of disordered systems, it is necessary to perform statistical (ensemble) averaging of the results corresponding to particular sets of on-site potential distributions (to be referred to as samples hereafter). This is a purely technical problem easy to overcome at the expense of the computa-
tion time. Another more serious problem is to develop a recursive procedure, which would make it possible to deal with big systems approaching macroscopic sizes of the order of several hundred nanometers. Recursive algorithms based on Dyson-type equations for the Green’s function are well-known [13, 14]. Here however we modify those methods in order to make them work in the case of highly non-homogenous systems composed of disordered carbon nanotubes and two adjacent atomic layers from each electrode (to be referred to as the extended molecule). While conductance computations are usually rather fast, the computations of electronic charge at all atoms of a big system are extremely computer time consuming and expensive. In order to surmount this difficulty we impose a global charge neutrality condition only on a disorder-free “parent” system and self-consistently determine detailed values of all its on-site potentials (ca. 4000 and 40000 atoms for the SWCNT(8,8), 30 and 300 nm long). On introducing disorder, these on-site potentials are modified by random corrections fluctuating around zero within an interval \([-W/2, W/2]\). So, on the average the global charge of the Anderson-disordered extended molecule might be regarded as roughly close to that of the neutral parent system. Our computations proceed according to the following protocol: First the surface Green functions are found (see [9] for details). Second, the set of on-site potentials which ensure the charge neutrality of the parent system is found. Third, conductance calculations along with the corresponding GMR coefficients are performed for 100 different samples with random on-site corrections. Finally the results are ensemble-averaged. The main results of this study is presented in Fig. 2, for the SWCNT (8,8) consisting of 240 carbon rings (120 unit cells ~30 nm). It is seen that although disorder suppresses the GMR, it happens to be of about the right value as compared to recent experiments on MWCNTs with transparent ferromagnetic contacts made from \(\text{Pd}(10\%\text{Ni})\) (device resistances are then as low as 5.6 k\(\Omega\) at 300 K) [15]. Other noteworthy points are: (i) on the average the GMR remains positive, and (ii) there exist some extra features in the GMR spectrum at energies close to \(\pm 0.4\) and \(\pm 0.7\) corresponding to higher sub-band onsets in the pristine (ideal) SWCNT.

For smaller \(W\), GMR increases and eventually oscillates with the amplitude of roughly \(\pm 0.2\) in the disorder-free (parent) case, as shown in Fig. 3 (l.h.s). Additionally the right-hand side of this figure highlights the length-effect on the period of oscillations. For the sake of simplicity this is shown for the paramagnetic leads. It is clearly seen that in the absence of disorder the observed periodicity reflects length-dependent interferences, as expected for a (quasi) ballistic transport. In the magnetic case the peaks are split due to the lifting of spin-degeneracy. The quasi-periodic behavior does always occur when there is no disorder, regardless of whether the electrodes are ferromagnetic or not (c.f. the inset in Fig. 3 with the thick solid curve on the r.h.s). From the present results one sees that the process of averaging of conductance and GMR spectra leads to a subtle interplay between the length and the amount of disorder in the CNTs.

**Fig. 2** Left hand side: GMR for individual SWCNTs (8,8), ca. 30 nm in length (points), along with the GMR (white curve) averaged over 100 samples with disorder-induced corrections to the on-site potentials (within \([-W/2, W/2]\) for \(W = 0.2\)). On the right hand side the GMR computed from the disorder-averaged conductances together with the standard-deviation error bars are shown in the vicinity of the charge neutrality point.
Fig. 3 (online colour at: www.pss-b.com) Left hand side: GMR for disorder-free SWCNT (8,8), $P = 50\%$, $W = 0$, $L = 240$ carbon rings ($\sim 30$ nm). Right hand side: Visualization of the length-dependent periodicity of the conductance for the case of paramagnetic leads, $P = 0$, and $L = 5$, $30$ and $300$ nm. Compare the inset with the thick solid curve to see that the quasi-period of oscillations is roughly maintained, but in the magnetic case the peaks are spin-split.

4 Summary

In this work it has been shown theoretically that the GMR effect in ferromagnetically contacted carbon nanotubes is quite considerable and may reach a few tens percent. Ideally, the GMR coefficient oscillates as a function of energy (gate-voltage) with a quasi-period close to the inter-level spacing of the CNT, which scales inversely proportional to the nanotube length. Yet, such a picture is to some extent too detailed if the system at hand is imperfect, e.g. due to some impurities, dopants or a presence of incommensurate inner shells in a MWCNT. The disorder-averaged GMR ranges from $6\%$ down to $2\%$ in the vicinity of the charge neutrality point, in conformity with recent experiments on MWCNTs with transparent ferromagnetic contacts. Furthermore, the aforementioned periodicity gets nearly completely suppressed, and there is no more tendency for the GMR to become negative.

Acknowledgements S. K. thanks the KBN project (PBZ-KBN-044/P03-2001), the Centre of Excellence (contract No. GSMA-CT-2002-04049) and the Poznan Supercomputing and Networking Center for the computing time. This work was partially funded by the Volkswagen Foundation and the Vielberth Foundation.

References

Contact Dependence of Carrier Injection in Carbon Nanotubes: An Ab Initio Study

Norbert Nemec,1 David Tománek,2 and Gianaurelio Cuniberti1

1Institute for Theoretical Physics, University of Regensburg, D-93040 Regensburg, Germany
2Physics and Astronomy Department, Michigan State University, East Lansing, Michigan 48824-2320, USA

(Received 25 August 2005; published 23 February 2006)

We combine ab initio density functional theory with transport calculations to provide a microscopic basis for distinguishing between good and poor metal contacts to nanotubes. Comparing Ti and Pd as examples of different contact metals, we trace back the observed superiority of Pd to the nature of the metal-nanotube hybridization. Based on large scale Landauer transport calculations, we suggest that the optimum metal-nanotube contact combines a weak hybridization with a large contact length between the metal and the nanotube.

DOI: 10.1103/PhysRevLett.96.076802 PACS numbers: 73.23.Ad, 73.40.Cg, 73.63.Fg, 73.63.Rt

A major challenge linked to the use of carbon nanotubes [1] in future electronic devices is to understand the profound effect of the nanotube-metal contact on transport. Weak nanotube-metal coupling, found in nanotubes deposited on metal electrodes, has been shown to cause Coulomb blockade behavior [2]. In spite of significant progress in maximizing the contact area by depositing metal on top of nanotubes [3], the transparency of such contacts exhibits strong sample-to-sample variations and depends strongly on the contact metal. Reports of low contact resistance between nanotubes and Au or Au/Cr [4,5] are in stark contrast to the high resistance observed in nanotube contacts with Au/Ti [6]. The transparency of Pd-based contacts has been reported as superior in comparison to using Ti, Pt, and Al as contact metals [7–9]. Additional modulation of the Pd-nanotube contact transparency has been reportedly achieved by modulating the gate voltage [10]. Reports suggesting that carrier injection occurs only at the edge of the contact region [8] appear to contradict the observed dependence of the contact resistance on the length of the contact [6].

Published theoretical results include studies of the electronic structure at a nanotube-Au interface and transport properties of a nanotube-Al junction [11]. Ab initio calculations furthermore suggest that Ti contacts may be superior to those with Al or Au [12], and that the Schottky barrier between semiconducting tubes and Pd is lower than with Au or Pt [13]. Because of the limitation to specific examples of different contact metals, we trace back the observed superiority of Pd to the nature of the metal-nanotube hybridization. Based on large scale Landauer transport calculations, we suggest that the optimum metal-nanotube contact combines a weak hybridization with a large contact length between the metal and the nanotube.

FIG. 1 (color online). Charge density redistribution \( \Delta \rho (r) = \rho_{\text{metal}}(r) - \rho_{\text{metal}}(r) - \rho_{\text{metal}}(r) \) in (a) Pd and (b) Ti monolayers interacting with a graphene layer, indicating regions of charge depletion and excess with respect to the superposition of isolated layers. (c) Schematic double-layer geometry in top view, with the cutting plane used in (a) and (b) indicated by the dash-dotted line.

0031-9007/06/96(7)/076802(4)$23.00 © 2006 The American Physical Society
tion. For both Pd and Ti, we found a slight preference for the sixfold hollow site on graphite, depicted schematically in Fig. 1(c). For Pd, we found the equilibrium interlayer distance to be 3.2 Å, consistent with a relatively weak, mostly covalent bond energy of 0.3 eV per Pt atom. The interaction between an epitaxial Ti monolayer and graphene was only insignificantly stronger with 0.4 eV per Ti atom at an interlayer distance of 3.0 Å.

The quality of nanotube-electrode contacts has been shown to depend sensitively on the Schottky barrier in semiconducting nanotubes [15] and band bending in metallic nanotubes, both reflecting the charge transfer within the junction. Our Mulliken population analysis indicates a net charge transfer of only ≈0.1 electrons from Pd and Ti to the graphene layer. More useful information is contained in the charge redistribution, depicted in Fig. 1. Results for Pd electrodes, shown in Fig. 1(a), suggest an accumulation of excess charge in the region between Pd and graphene layers. As seen in Fig. 1(b), the charge redistribution in Ti/graphene is very different, suggesting charge accumulation in the atomic layers, depopulation of the interlayer region, and thus an increase of the interlayer scattering potential. The lower scattering potential and the populated interlayer states in the Pd/graphene system, the rigid band shift at the junction, is much stronger, suggesting charge accumulation in the region between Pd and graphene, which is a prerequisite for a good contact. The electronic density of states can be extracted from our calculations, which is described by an energy-independent, purely imaginary self-energy $\Sigma(E) = -i\Delta$, where $\Delta = r_{M/C}N_M(E_F)$ can be extracted from our $ab initio$ results. Using the calculated density of states $N_M(E_F)$ at Pd and Ti surfaces, which is of the order $1 \text{ eV}^{-1}$, we obtain $\Delta_{\text{Pd}} = 0.06 \text{ eV}$ and $\Delta_{\text{Ti}} = 0.3 \text{ eV}$, assuming that each carbon atom is in direct contact with three metal atoms.

Transmission $T$ through a molecular conductor is generally limited by the number of channels $N_{\text{ch}}$, which depends on the band structure of a perfectly periodic system at $T \equiv T_{\text{bands}}$. The reason for the effective transmission through a conductor with contacts $T_{\text{con}}$, $a$ being lower than through the ideal infinite system $T_{\text{bands}}$, lies in

Especially interesting for the transparency of the contact is the nature of Pd-C rehybridization, which is best visible in Fig. 2(b) close to the Fermi level. Particularly clear is the hybridization between $\text{Pd}d_z^2$ and $Cp_z$ orbitals, which causes a $\approx 0.15 \text{ eV}$ band splitting about $0.5 \text{ eV}$ below $E_F$, in the vicinity of the $K$ point. Since $K$ denotes also the Fermi momentum of graphene, this occurrence of Pd-C hybridization near this $k$ point suggests an efficient way to inject carriers into graphene near the Fermi energy without involving phonons to conserve momentum.

To obtain quantitative information about the effect of the junction geometry and hybridization on the transparency of the contact, we performed large scale quantum transport calculations of nanotubes in contact with metal electrodes within the Landauer approach [17]. Our calculations for nanotube segments exceeding $10^5$ nm were facilitated using an efficient $O(N)$ decimation algorithm [18]. Our transport calculations were based on a simplified tight-binding Hamiltonian, describing only the interaction between $\text{Pd}d_z^2$ and $Cp_z$ orbitals. We found that the electronic band structure of the Pd/graphene system near $E_F$, depicted in Fig. 2(b), can be reproduced by considering the hybridization between the $pp\pi$ band of graphene, associated with $\gamma_0 = 2.66 \text{ eV}$, and a Pd-based band, using $t_{\text{Pd-C}} = 0.15 \text{ eV}$ for the hopping integral between Pd and each of the six C neighbors. Such a simple mapping turned out insufficient to describe the hybridization between Ti and graphene. Based on typical band repulsion observed in that system, the Ti-C hopping integral $t_{\text{Ti-C}} = 0.3 \text{ eV}$ should be about twice that found for Pd.

Results of our transport calculations for metal-nanotube junctions are summarized in Fig. 3. In our schematic view of an extended contact, depicted in Fig. 3(a), we distinguish three regions within a finite tube. The central region of length $L_{\text{c}}$, describing an unperturbed nanotube, is connected at both ends to contact regions of varying length $L_{\text{c}}$. In both extended contact regions, each atom is coupled to the coating metal electrode in a similar way, as previously considered in Ref. [19].

In the model examined in the following, we chose diagonal wide-band leads, which contacted each atom of the nanotube independently. In this case, the coupling is described by an energy-independent, purely imaginary self-energy $\Sigma(E) = -i\Delta$, where $\Delta = r_{M/C}N_M(E_F)$ can be extracted from our $ab initio$ results. Using the calculated density of states $N_M(E_F)$ at Pd and Ti surfaces, which is of the order $1 \text{ eV}^{-1}$, we obtain $\Delta_{\text{Pd}} = 0.06 \text{ eV}$ and $\Delta_{\text{Ti}} = 0.3 \text{ eV}$, assuming that each carbon atom is in direct contact with three metal atoms.

Transmission $T$ through a molecular conductor is generally limited by the number of channels $N_{\text{ch}}$, which depends on the band structure of a perfectly periodic system at $T \equiv T_{\text{bands}}$. The reason for the effective transmission through a conductor with contacts $T_{\text{con}}$, $a$ being lower than through the ideal infinite system $T_{\text{bands}}$, lies in
the reflection at the contacts. To quantify the quality of a nanotube contact, we define the contact reflection coefficient by
\[
\rho(L_c, \Delta) = \frac{1}{\tau(L_c, \Delta)} - \frac{1}{\tau_{\text{hand}}}
\]
(1)
The average is taken around the Fermi level, in a region between the first van Hove singularities, similar to a transmission convolution capturing hot electron effects [20]. Our results in Fig. 3, based on ensemble averaging in Eq. (1), agree quantitatively with those obtained by averaging the transmission using the Fermi distribution, for a wide range of temperatures. Physically, the total resistance \( R \) of such an idealized system could be separated as \( R = R_{\text{hand}} + R_c \), where \( R_{\text{hand}} = 1/(2G_0) \) is given by the quantum limit of two open channels with conductance \( G_0 = 2e^2/h \) each. Assuming zero temperature and neglecting disorder effects, \( R_c \approx \rho/G_0 \) originates only from suboptimal contacts and may theoretically be arbitrarily small.

We generally expect zero transmission in the limiting cases of vanishing contact length, \( L_c = 0 \), and vanishing coupling, \( \Delta = 0 \). For finite values of \( L_c \) and \( \Delta \), however, it is not obvious if a combination of strong coupling and short contact is superior to a combination of weak coupling and a long contact. To obtain a quantitative answer, we calculated \( \rho \) for a contact to a (6,6) armchair nanotube with \( L_0 = 100 \text{ nm} \) as a function of \( L_c \) and \( \Delta \). We found \( \rho \) to be independent of the tube diameter, as long as the energy range used in the averaging avoids subband-related van Hove singularities. We also found no dependence on \( L_0 \), as long as \( L_0 \) was much larger than a unit cell size. Consequently, \( \rho \) should be only a function of \( L_c \) and \( \Delta \).

Our results for \( \rho(L_c, \Delta) \) are depicted in Figs. 3(b) and 3(c). We find the contour plot of \( \rho \) in Fig. 3(b) separated into two regions by a line of “minimum contact reflection” \( L_{\text{eff}}^0(\Delta) \). In short contacts, the transparency is restricted by \( L_c \) and \( \rho \) increases with decreasing \( L_c \) due to a generalized Breit-Wigner broadening of the resonances. The pronounced ripples found at small values of \( \Delta \) and \( L_c \) are not numerical artifacts, but rather reflect the interplay between resonances in a finite nanotube segment and conduction electrons propagating with the Fermi momentum. For very weak coupling, we find \( \rho \sim \rho_0 \Delta^2 \). For larger values of \( L_c \), however, dissipation along the contact region modifies this simple behavior, yielding \( \rho \sim \exp(-\alpha L_c \Delta) \).

In sufficiently long contacts, defined by \( L_c > L_{\text{eff}}^0(\Delta) \), \( \rho \) becomes independent of \( L_c \) and is given by \( \rho = \rho_0 \Delta^2 \). We found that our results can be reproduced well by using \( \rho_0 = 0.016 \text{ eV}^{-2} \). The independence of \( \rho \) from \( L_c \) in long contacts is seen clearly in the plots of \( \rho(L_c) \), depicted in Fig. 3(c) for selected values of \( \Delta \). Particularly intriguing appears our result that reflection is minimized in case of a weaker specific coupling, provided the contact is sufficiently long.

This physical origin of this unexpected behavior is schematically illustrated in the insets in Fig. 3(c). Saturation transparency is reached for relatively short contacts in the case of strong coupling. In this case, however, the abrupt change in the electronic structure between the uncoated and the coated nanotube segment causes extra reflection. This change is less abrupt in the case of weaker coupling and a long effective contact region \( L_{\text{eff}}^0(\Delta) \), reducing the overall reflection.

To estimate the effective contact length \( L_{\text{eff}}^0(\Delta) \), we make use of the above described expressions for \( \rho \) in the adjacent regions in the \( (L_c, \Delta) \) plane. The line, where these two functions intersect, corresponds to the line of minimum \( \rho \), and is given by the analytical expression
\[
L_{\text{eff}}^0(\Delta) = \ell_{\text{uc}} \alpha_1 \frac{\alpha_1}{\Delta} \ln(\alpha_2/\Delta)
\]
(2)
Here, \( \ell_{\text{uc}} = 2.46 \text{ Å} \) is the unit cell length. The parameters \( \alpha_1 = 1.34 \text{ eV} \) and \( \alpha_2 = 9.14 \text{ eV} \) were obtained by fitting our numerical data. In the specific case of Pd and Ti, we found \( L_{\text{eff}}^0(\Delta_{\text{Pd}(C)}) \approx 30 \text{ nm} \) and \( L_{\text{eff}}^0(\Delta_{\text{Ti}(C)}) \approx 4 \text{ nm} \).

Since realistic metal-nanotube contacts are rarely epitaxial, we have considered various forms of disorder in the

076802-3
We modeled weak to moderate disorder by randomly perturbing the metal-nanotube coupling with respect to $\Delta_{\text{M/C}}$. We did not find noticeable change in $\rho$ even for perturbations of $\Delta$ as large as its reference value. As another extreme case, we modeled strongly diluted contacts by randomly suppressing the interaction between individual metal and nanotube atoms in the contact region, down to 1% of contacts with respect to the epitaxial case. Finally, we considered various forms of nondiagonal contributions to the self-energy, modeling the cross coupling between neighboring metal atoms. In all the studies, which addressed deviations from epitaxy at the interface and our description of the leads, we found the same behavior as depicted in Fig. 3 and analytically described by Eq. (2), with possibly modified numerical values of $\rho_{t}$, $\alpha_{t}$, and $\alpha_{z}$.

Our main conclusion, which proved to be robust with respect to variations in the details, is that each contact can be characterized by an effective contact length $L_{c}^{\text{eff}}$, which depends only on the local metal-nanotube coupling, not on the diameter of the tube. Assuming that the effective contact length between the nanotube and the electrode exceeds $L_{c}^{\text{eff}}$, which is likely the case in most experimental setups, then a higher contact transparency is expected when the metal-nanotube coupling is weak [19]. Especially in very long contacts, the sensitive $\rho = \rho_{0}\Delta^{2} \propto n_{\text{M/C}}^{4}$ dependence of $\rho$ on the coupling strength $n_{\text{M/C}}$ may be taken as an important guideline, suggesting to minimize coupling and maximize contact length to achieve a high contact transparency. In the specific case of Pd and Ti contacts, the weaker nanotube-metal coupling in the case of Pd is a good explanation for the superiority of Pd-based nanotube contacts.

In conclusion, we combined \textit{ab initio} density functional theory with transport calculations to distinguish microscopically between “good” and “poor” metal contacts to nanotubes. Comparing Pd and Ti as examples of different contact metals, we traced back the observed superiority of Pd to the nature of the metal-nanotube hybridization. Based on large scale Landauer transport calculations, we suggest that the “optimum” metal-nanotube contact generally combines a weak hybridization with a large contact length of typically few hundred nanometers between the metal and the nanotube.

Of immediate interest is, of course, the general validity of our results. We plan additional studies addressing the effect of nonepitaxial contacts and the nature of charge carriers [18]. Particularly interesting in this respect should be studying spin injection from ferromagnetic contacts and Andreev reflection at the contact to superconducting electrodes.

We acknowledge fruitful discussions with M. Grifoni, A. Nitzan, J. Fabian, and C. Strunk. This work has been funded by the Volkswagen Stiftung under Grant No. I/78 340. D. T. was partially supported by the Vieltherr Foundation, the DFG-GRK 638, NSF-NIRT Grants No. DMR-0103587 and No. ECS-0506309, and NSF NSEC Grant No. 425826.

[20] An alternative definition $\rho = \left(1/T - 1/T_{\text{ex}}\right)$ would differ from Eq. (1) only in the high-$\rho$ regime, which is not of primary interest here.
Hofstadter butterflies of carbon nanotubes: Pseudofractality of the magnetoelectronic spectrum
Norbert Nemec and Gianaurelio Cuniberti
Institute for Theoretical Physics, University of Regensburg, Germany
(Received 28 June 2006; revised manuscript received 29 August 2006; published 23 October 2006)

The electronic spectrum of a two-dimensional square lattice in a perpendicular magnetic field has become known as the Hofstadter butterfly [Hofstadter, Phys. Rev. B 14, 2239 (1976)]. We have calculated quasi-one-dimensional analogs of the Hofstadter butterfly for carbon nanotubes (CNTs). For the case of single-wall CNTs, it is straightforward to implement magnetic fields parallel to the tube axis by means of zone folding in the graphene reciprocal lattice. We have also studied perpendicular magnetic fields which, in contrast to the parallel case, lead to a much richer, pseudofractal spectrum. Moreover, we have investigated magnetic fields piercing double-wall CNTs and found strong signatures of intervalley interaction in the resulting Hofstadter butterfly spectrum, which can be understood with the help of a minimal model. Ubiquitous to all perpendicular magnetic field spectra is the presence of cusp catastrophes at specific values of energy and magnetic field. Resolving the density of states along the tube circumference allows recognition of the snake states already predicted for nonuniform magnetic fields in the two-dimensional electron gas. An analytic model of the magnetic spectrum of electrons on a cylindrical surface is used to explain some of the results.

I. INTRODUCTION

The availability of new materials for nano- electronic research allows for a detailed test of the emergence of the quantum physical nature of electrons, via transport or optical measurements. Carbon nanotubes2–7 (CNTs) are an example of a very peculiar electronic material, due to the extreme confinement of electrons on their π-conjugated “walls”. In these systems, many mesoscopic phenomena such as single-electron charging,8 and conductance quantization,9 as well as effects typical for semiconductor physics like s-like excitons,10 can be observed already at room temperature.

Since the prediction of band structure effects of carbon nanotubes in parallel external fields by Ajiki and Ando in 1993,11 it took only a few years until clear hallmarks of a 2D magnetic spectrum of electrons on a cylindrical surface is used to explain some of the results. The study of the local distribution of the spectral density will shed some light on the relation between the spectrum of a planar sheet of graphene and that of a CNT, strongly affected by curvature and finite size. A closer look at the spectrum will reveal the presence of cusp catastrophes, which are closely related to the quenching of the Bloch state velocity, induced by a magnetic field.

For magnetic fields parallel to the tube axis, the natural unit is that of one flux quantum per tube cross section r²πΘ. The density of states of a CNT in dependence on an external magnetic field parallel (left) or perpendicular (right) to the tube axis. For every value of the magnetic field, the DOS is unity normalized over energy. The units B||=Φ0/2π and B⊥=Φ0/A_lateral (see text) are scaled such that the physical field scale is the same for both segments of the plot.
FIG. 2. (Color online) The structure of a CNT [here, a (3,3)-CNT]: the hexagonal lattice of a graphene sheet is rolled up in such a way that the chiral vector \( (n,m) \) becomes the circumference of the resulting cylinder. Magnetic fields parallel to the tube axis pierce the tube cross section \( \pi r^2 \), while perpendicular magnetic fields pierce the wall made up from hexagonal plaquettes.

For a general \( (n,m) \) CNT the tube radius can be obtained with simple geometrical arguments:

\[
B_0 = \frac{\Phi_0}{\pi r^2} = \frac{2\sqrt{3} \pi}{m^2 + n^2 + mn} B_0^2.
\]

For a typical SWCNT with \( \sim 1 \text{ nm diameter} \), this gives a value of \( B_0^2 \approx 5 \times 10^3 \text{ T} \). It is thus understandable that multi-wall CNTs (MWCNTs) present a more interesting object for magnetic field experiments: For a typical MWCNT with a diameter of \( 20 \text{ nm} \), as a matter of fact, one can already observe the first Aharonov-Bohm oscillations accessible at around \( 12 \text{ T parallel fields} \).

As shown in this work, however, even for perpendicular fields low-field signatures could be visible within experimentally accessible field ranges if one takes into account the external shell of a MWCNT.

This paper is organized as follows. We first give definitions, introduce the method of computation and visualization, and point toward general features observable in quasi-1D systems. In Sec. III, we then do a systematic study of SWCNTs, including an analytic model and a detailed view of the range of experimentally accessible fields. In Sec. IV, we proceed with an analysis of the effects of the interwall interaction in DWCNTs on the magnetic spectrum and introduce a minimal model, closing with a discussion of the results in the last section.

II. DEFINITIONS, METHODS, AND OBSERVABLES

Lattice electrons in arbitrary external magnetic fields. All numerical calculations in this work are based on a tight-binding Hamiltonian of the form

\[
\hat{H}(B) = \sum_i e_i c_i^\dagger c_i - \sum_{i,j} \gamma_{ij}(B) c_i^\dagger c_j,
\]

where the indices denote the atomic orbitals. For the single-orbital approximation used hereafter, these coincide with the label of the atom so that \( \hat{H} \) can be represented by means of the matrix elements \( \gamma_{ij} \) between \( \pi \) orbitals centered on the atom \( i \) and \( j \) at the position \( r_i (r_j) \) of the CNT molecular network.

An external magnetic field is implemented using the Peierls substitution. Based on the principle of minimal coupling \( p \rightarrow p - eA \), the effect of a magnetic field \( B = \text{rot} A \) is absorbed in the translation operator \( \hat{T}(R) = \exp\left[\frac{e}{\hbar c} (p - eA) \cdot R\right] \). In the tight-binding Hamiltonian, this is reflected by representing the hopping matrix elements \( \gamma_{ij} = \langle \Psi_j | \hat{H} | \Psi_i \rangle \) between two \( \pi \) orbitals \( \Psi_i \) and \( \Psi_j \) localized at sites \( r_i \) and \( r_j \) as

\[
\gamma_{ij}(B) = \gamma_{ij}^0 \exp\left\{\frac{2\pi}{\Phi_0} \int_{r_i}^{r_j} \mathbf{r} \cdot \mathbf{A}(r) \right\}.
\]

The bare hopping at zero magnetic field \( \gamma_{ij}^0 \) acquires a complex phase expressed as an integral along the bond direction \( d_{ij} = r_j - r_i \).

With the CNTs oriented parallel to the \( z \) axis, it is advantageous to choose a gauge in such a way that \( \mathbf{A}_g \) is independent of \( r_z \). This is provided, e.g., by

\[
\mathbf{A}_g(r) = (0, r_y B_z, r_x B_y),
\]

giving a magnetic field \( \mathbf{B} = (B_z, 0, B_y) \) with known components perpendicular and parallel to the tube axis. Throughout this work fields will be considered either perpendicular \( (B_z = 0) \) or parallel \( (B_z = 0) \) to the tube axis. Arbitrary angles are of course possible as well, showing the expected crossover of both regimes.

Having chosen a linear gauge further simplifies the integration in Eq. (2) to a product:

\[
\gamma_{ij} = \gamma_{ij}^0 \exp\left\{\frac{2\pi}{\Phi_0} \mathbf{r}_j \cdot \mathbf{A}_g\left(\frac{r_i + r_j}{2}\right)\right\}.
\]

II. DEFINITIONS, METHODS, AND OBSERVABLES

Lattice electrons in arbitrary external magnetic fields. All numerical calculations in this work are based on a tight-binding Hamiltonian of the form

\[
\hat{H}(B) = \sum_i e_i c_i^\dagger c_i - \sum_{i,j} \gamma_{ij}(B) c_i^\dagger c_j,
\]

where the indices denote the atomic orbitals. For the single-orbital approximation used hereafter, these coincide with the label of the atom so that \( \hat{H} \) can be represented by means of the matrix elements \( \gamma_{ij} \) between \( \pi \) orbitals centered on the atom \( i \) and \( j \) at the position \( r_i (r_j) \) of the CNT molecular network.

An external magnetic field is implemented using the Peierls substitution. Based on the principle of minimal coupling \( p \rightarrow p - eA \), the effect of a magnetic field \( B = \text{rot} A \) is absorbed in the translation operator \( \hat{T}(R) = \exp\left[\frac{e}{\hbar c} (p - eA) \cdot R\right] \). In the tight-binding Hamiltonian, this is reflected by representing the hopping matrix elements \( \gamma_{ij} = \langle \Psi_j | \hat{H} | \Psi_i \rangle \) between two \( \pi \) orbitals \( \Psi_i \) and \( \Psi_j \) localized at sites \( r_i \) and \( r_j \) as

\[
\gamma_{ij}(B) = \gamma_{ij}^0 \exp\left\{\frac{2\pi}{\Phi_0} \int_{r_i}^{r_j} \mathbf{r} \cdot \mathbf{A}(r) \right\}.
\]

The bare hopping at zero magnetic field \( \gamma_{ij}^0 \) acquires a complex phase expressed as an integral along the bond direction \( d_{ij} = r_j - r_i \).

With the CNTs oriented parallel to the \( z \) axis, it is advantageous to choose a gauge in such a way that \( \mathbf{A}_g \) is independent of \( r_z \). This is provided, e.g., by

\[
\mathbf{A}_g(r) = (0, r_y B_z, r_x B_y),
\]

giving a magnetic field \( \mathbf{B} = (B_z, 0, B_y) \) with known components perpendicular and parallel to the tube axis. Throughout this work fields will be considered either perpendicular \( (B_z = 0) \) or parallel \( (B_z = 0) \) to the tube axis. Arbitrary angles are of course possible as well, showing the expected crossover of both regimes.

Having chosen a linear gauge further simplifies the integration in Eq. (2) to a product:

\[
\gamma_{ij} = \gamma_{ij}^0 \exp\left\{\frac{2\pi}{\Phi_0} \mathbf{r}_j \cdot \mathbf{A}_g\left(\frac{r_i + r_j}{2}\right)\right\}.
\]

In the presence of a perpendicular magnetic field, it is thus necessary to consider the exact coordinates of the molecular structure at hand rather than—as sufficient for parallel or vanishing magnetic fields—their simple topological connectivity.

Density of states. For such an \( r_z \)-independent gauge field, the Hamiltonian of any quasi-1D periodic structure like a CNT stays periodic in the presence of a magnetic field. This allows the use of the Bloch theorem to derive the corresponding band structure. As can be seen in Fig. 3, the band structure is in general strongly distorted by an applied magnetic field. The density of states (DOS) can be determined from the magnetic band structure \( E_z(k, B) \) via

\[
\gamma_{ij}(B) = \gamma_{ij}^0 \exp\left\{\frac{2\pi}{\Phi_0} \mathbf{r}_j \cdot \mathbf{A}_g\left(\frac{r_i + r_j}{2}\right)\right\}.
\]
\[ p_{\text{DOS}}(E, B) = \frac{\alpha}{2m N_b} \sum_{n,m} \int_{-\mu}^{\mu} dk \Im \delta(E - E_n(k, B)) \]  

where \( b \) is the band index and \( \alpha = 3d_{\text{CC}}^2 \mu^2 + \pi^2 + m^2 \) \( \text{gcd}(3n, n-m) \) the length of the unit cell of an \((n, m)\) CNT. \( E_n(k, B) \) is obtained by direct diagonalization of the CNT Hamiltonian via the Bloch ansatz (see Appendix A). Since we work in a basis of one orbital per atom, the number of bands \( N_b \) equals the number of atoms in the unit cell \( N = 4(\pi^2 + m^2) \text{gcd}(3n, n-m) \), growing with the diameter and dependent on the helicity angle (deviation from the armchair or zigzag configuration). The plot of the DOS directly reflects the distortions of the band structure caused by the magnetic field.

Alternatively, Green-function-based approaches allow one to resolve the DOS within different atoms in the same unit cell by introducing the local density of states

\[ p_{\text{DOS}}(E, B) = -\frac{1}{\pi} \text{Im} G_{r,r'}(E, B), \]  

where \( G_{r,r'} \) is the space-diagonal component of the lattice Green function matrix (see Appendix B)

\[ \delta(k, E) = (E - \mathcal{H}(B) + i\delta)^{-1}. \]  

Of course by tracing the LDOS within the different atoms of the same unit cell, one can restore the full DOS as \( p_{\text{DOS}}(E, B) = \sum_{r} p_{\text{DOS}}(E, B) \).

**Butterfly plots.** To capture the continuous evolution of the band structure with growing magnetic fields, it is very convenient to visualize the DOS in butterfly plots, as illustrated in Fig. 3. The resemblance to the well-known Hofstadter butterfly of 2D lattice electrons becomes very clear for CNTs of large diameter (see Fig. 7 below). A common feature to butterfly plots of all quasi-1D systems are the pronounced band edges, caused by van Hove singularities in the DOS.\(^{13}\)

In Fig. 1, a (6,6) CNT Hofstadter butterfly is plotted as a reference for further comparisons. For the parallel field, the behavior is perfectly periodic for integer multiples of the flux quantum \( \Phi_0 = b/e \) penetrating the tube cross section \( \pi^2 \). Starting as a metallic CNT at \( B = 0 \), the gap opens and closes periodically.\(^{11}\)

For perpendicular fields with their natural scale of one flux quantum per graphene plaquette, the overall behavior is not periodic. This can be understood due to the presence of plaquettes at various angles toward the field, capturing different, in general incommensurate, fractions of the flux quantum. However, a number of features from the underlying graphene structure are still visible at the diameter-independent scale of \( B_G \).

Important to note is the difference in the behavior for small fields: while the parallel field causes a linear Zeeman split of the states with opposite angular momentum, small perpendicular fields generally cause quadratic energy shifts.

All plots are of course symmetric in the magnetic field sign, which is why only half butterfly plots are shown. Different is the case of the \( E \rightarrow -E \) symmetry which is related to the particle-hole symmetry. The latter is present in the \( \pi \)-orbital description of SWCNTs but is broken by the interwall interaction in DWCNTs.

**Transport observables in quasi-1D systems.** As can also be seen in Fig. 3, it is straightforward to apply the same scheme not only to the density of states, but just as well to other properties like the quantum mechanical transmission \( T \) of a quasi-1D system. The latter is the dimensionless zero-temperature conductance after the Landauer theory of phase-coherent transport\(^{36}\) \( G = G_K T \), where \( G_K = 2e^2/h \) is the conductance quantum and inverse of the von Klitzing resistance. The calculation of the transmission, which involves a renormalization procedure for the semi-infinite carbon nanotube leads\(^{43}\) by means of the energy-dependent injection rates \( \Gamma_{\pm} \) and the Green function \( \check{G} \) projected on a finite nanotube partition, can be cast into the Fisher and Lee formula\(^{36}\)

\[ G = \frac{2e}{h} \text{Tr}(\Gamma_{\pm} \check{G} \Gamma_{\pm} \check{G}). \]
Still, for a periodic structure—as is the case for the systems at hand—the quantum mechanical transmission is simply a band-counting algorithm, and as such contains less information than the band structure itself or the DOS. This is very different from magnetotransport through finite CNTs: Scattering at the contacts leads to resonant tunneling, resulting in spectroscopy of the electronic states of the finite tube. This spectrum may show strong dependence on magnetic fields, even in regions of flat bands, resulting in quantum-dot-like physics.

Relation to 2D periodic structures. It is important to note some similarities, but also some fundamental differences between the butterfly plots of quasi-1D structures and those in the original work by Hofstadter and later generalizations which handled 2D periodic structures. Starting out from an analogous Hamiltonian and also using the Peierls substitution to implement the magnetic field, the most striking difference is that, for a 2D periodic structure, it is not possible in general to choose a gauge in such a way that the resulting Hamiltonian has the same translational symmetry as the underlying system. For rational values of the magnetic flux per unit cell, one can still find a larger effective unit cell but, for irrational values, this is not possible at all, which ultimately leads to the fractal structure of the energy spectrum found by Hofstadter, similar to that displayed in the lower panel of Fig. 7. In contrast, the quasi-1D structure of CNTs results in a fixed number of bands, leading to a pseudofractal spectrum, with the recursion of self-similarity limited by the transverse length cutoff of the system.

Graphene ribbons. Since the recent experimental success in isolating single sheets of graphene, the exotic Dirac-like electronic structure has become the focus of several studies. Epitaxially grown graphene has been used to laterally confine electrons and determine coherence lengths studying weak-localization effects in magnetotransport measurements. For understanding the relation between the butterfly of a 2D graphene sheet and these quasi-1D carbon nanotubes, it is instructive to take a look at graphene ribbons as an intermediate step. An \((n,m)\) graphene ribbon is simply a planar "unrolled" CNT, periodic in one dimension and finite in the other. As in the original Hofstadter butterfly, the ribbon butterfly plots are periodic as a function of the perpendicular magnetic field due to the equal flux piercing any hexagonal plaquette forming the honeycomb lattice (see Fig. 4). As the ribbon width increases the butterfly plots tend to the Hofstadter butterfly of a graphene layer as visible in the bottom panel of Fig. 7.

Cusp catastrophes. One striking detail ubiquitous in butterfly plots are the cusp catastrophes appearing at specific positions of energy and magnetic fields (see Fig. 5). These are points where, with changing magnetic field, some band is continuously deformed from a strictly monotonic curve into a band with two adjacent zero-group-velocity points. At the exact point where this mathematical catastrophe happens, both the first and the second derivatives of \(E(k)\) are zero. A wave packet of this energy and momentum will have both its velocity and its spreading suppressed leading to a special kind of localization not unlike that of Landau levels.

III. SINGLE-WALL CARBON NANOTUBES

For SWCNTs we consider only the radial \(p\) orbitals—forming the most electronically relevant \(\pi\) bands—and only
interactions between nearest neighbors, setting $\gamma_{ij}^0 = \gamma_0 = 2.66$ eV. This has been shown to be an excellent approximation in explaining electronic structural and transport properties of SWCNTs. The on-site energy $e_i = e_0$ is constant for all atoms and defines the Fermi energy $E_F = e_i$ of a neutral CNT. Ignoring an offset in the energy, we can simply superimpose the electronic bands of a SWCNT via the procedure described above, one could calculate the spectrum of graphene and then apply periodic boundary conditions in the angular direction of the CNT (zone folding). For magnetic fields parallel to the tube axis, this method is still applicable: the phase gathered by an electron moving on a closed loop around the tube circumference can simply be included in the boundary conditions. This results in a shift of the allowed discretized quasimomenta in the reciprocal space. For perpendicular magnetic fields, however, this method breaks down and one has to consider the full geometry of the CNT.

A. Structural properties

Chirality dependence. Several features can be found when comparing the magnetic spectra of tubes with different chiralities though similar diameter (see Figs. 1 and 6). (i) The behavior of the gap around the charge neutrality point $E = E_F$ is very helicity dependent: a parallel magnetic field always opens and closes the gap periodically as a consequence of the integer number of fluxes per nanotube cross section. This phenomenon is independent of whether the tube is metallic or semiconducting at $B=0$. In contrast, for perpendicular fields there are distinctions. Armchair CNTs stay strictly metallic for any perpendicular field, as can be understood from supersymmetry arguments. On the other hand, the gaps of the two semiconducting CNTs in Fig. 6 do open and close in an aperiodic, though oscillatory, pattern. The gap closes to zero in single points of specific values of $B_z$ and opens again. Close observations of a larger set of CNTs reveal that this also happens for semimetallic tubes like the $(6,5)$ zigzag CNTs. The precise opening and closing pattern carries an intrinsic complexity; its statistical behavior, however, seems to depend mostly on the number of atoms in the unit cell. (ii) Another general effect of the large unit cell in the $(6,5)$ CNT is that in this chiral tube with its large number of plaquettes at different angles towards the field, the high symmetry of the original system is broken down very efficiently by the magnetic field, resulting in a larger number of bands of very low dispersion. The magnetic field effectively localizes the electrons in nonpropagating Landau-like states.

Diameter dependence. Figure 7 shows the evolution toward the graphene Hofstadter butterfly of the magnetic spectrum of armchair CNTs as a function of their diameter. The (200,200) SWCNT has a diameter of 27 nm, comparable to...
the external shell of a typical MWCNT, and is thus of great interest. The overlaid visual effect—resembling watercolors “flowing” toward the right hand side—can also be understood by a simple picture. For large enough diameters, the CNT consists of regions of nearly flat graphene, each at a different angle toward the magnetic field, thereby experiencing a different normal component of the magnetic field, as visible in the inset of the middle panel of Fig. 7. Since the DOS is an average over the LDOSs at the different unit cell atoms, one ends up with a sum of different graphene Hofstadter butterflies, stretched to different effective fields, the stretch being minimal where the magnetic field is normal to the tube wall and maximal where it is tangential. Overlaying these differently stretched graphene butterflies results in the “flowing” appearance of the butterfly of large diameter tubes.

At the lower and upper energy edges, one can clearly see the emergence of linear Landau levels and the characteristic fractal structure of the graphene butterfly is unmistakably visible at the same scale of the magnetic field. In fact near the top and bottom of the graphene $\pi$ energy band (of width $2W = 6/3 \approx 16$ eV), electrons have an effective mass of $m^* = 2\hbar^2/3y_C^2 = 0.95m_e$, leading to a cyclotron frequency $\omega(B) = eB/\hbar/m^*$, so one could write

$$E = \pm \frac{W}{2} n + \frac{1}{2}$$

with $n = 0, 1, 2, \ldots$, which fits nicely with the numerical results (as indicated by the straight lines in the bottom panel of Fig. 7).

Around the Fermi energy, the Dirac-like dispersion of graphene leads to the so-called relativistic Landau levels, following

$$E = \pm \sqrt{\omega(B)n + \frac{1}{2}}$$

with $n = 0, 1, 2, \ldots$, and the Fermi velocity $v_F = 3y_C^2/2\hbar$. These levels, which can be clearly observed in the Hofstadter butterfly of graphene (left edge in the bottom panel of Fig. 7), are also responsible for the recently observed anomalous quantum Hall effect of graphene.42,43

**Snake states.** The view of the total DOS of a large SWCNT as the sum of different contributions from the regions at various angles around the tube circumference can be confirmed by taking a look at the LDOS at individual atom positions. As can be seen in Fig. 8, the LDOS at $\theta = 0$, where the magnetic field pierces the wall perpendicularly, resembles very much the butterfly of the planar graphene sheet. The electrons here show very low dispersion, similar to Landau levels. At $\theta = \pi/2$, on the other hand, the magnetic field is tangential to the CNT wall and therefore has far less effect on the electron dispersion. An understanding of the electronic states in these regions can be gained by considering classical electrons confined to the surface of a cylinder. As the effective magnetic field (the projection of the field onto the tube normal) changes sign at $\theta = \pi/2$, the curvature of an electron trajectory will also switch orientation each time the electron crosses this “equator” line, leading to a snakelike movement of the electron.43,48

---

**FIG. 7.** (Color online) Comparing different diameters: the large (200,200) CNT bears strong resemblance to the Hofstadter butterfly of graphene (Ref. 25) combined with the curvature effects (details in text). The straight lines at the lower left corner of the graphene butterfly (bottom panel) indicate the Landau states obtained from an effective mass continuum theory [see Eq. (8)]. The parabolic lines near $E_F$ in the same plot indicate the relativistic Landau levels obtained from the Dirac-like dispersion of graphene [see Eq. (9)].
B. Analytical model

In order to shed some light of intuition on our results, we may consider the physics of a structureless hollow cylinder, a tubule, in a perpendicular magnetic field. (Similar systems in parallel magnetic fields have been studied before.\textsuperscript{48,49} This system bears some similarity to a “Hall bar”, with the crucial difference that it does not have borders that could carry edge states. Instead, it has two flanks where the magnetic field is tangential to the tube and therefore the radial component of the magnetic field—which is the effective field experienced by electrons confined to the cylinder surface—vanishes. To understand where charges do accumulate, we consider this system in cylindrical coordinates $(\theta, z)$ at fixed radius $r$. By Eq. (3), a perpendicular magnetic field leads to a gauge field

$$A(\theta) = B \sin \theta,$$

in cylindrical coordinates. With this, the Hamiltonian of an electron restricted to the tube surface becomes

$$\mathcal{H} = \frac{1}{2mr^2}p_r^2 + \frac{1}{2m}(p_\theta - eB \sin \theta)^2,$$

which can be viewed as that of an electron in 2D with periodic boundaries in a nonuniform magnetic field.\textsuperscript{51,52} A similar system—a 2D strip ranging over $[-L/2, L/2]$ in the $y$ direction and infinite in the $z$ direction, placed in a linearly varying magnetic field $B = B_0 \sin \theta$—was first studied in 1992 by Müller,\textsuperscript{53} who identified two new classes of states: one at finite magnetic field propagating perpendicularly to the field gradient direction with looping trajectory and low velocity, the other around the line $B = 0$, propagating in the opposite direction at higher velocity with a snakelike trajectory. To solve our system, we can exploit the $z$ invariance and do an ansatz for the wave function: $\Psi(\theta, z) = \psi(\theta) e^{ikz}$. Our problem reduces to that of a particle in one dimension with a $k_y$-dependent potential:

$$\mathcal{H}_k = \frac{1}{2mr^2}p_x^2 + V_k(\theta),$$

$$V_k(\theta) = \frac{1}{2m} \left( \frac{eB}{2} \right)^2 \sin \theta^2.$$  \hspace{1cm} (10)

For $|\hbar k| < |eB/r|$, this potential has two minima at $\theta_{\text{min}} = \pi/2 \pm \arccos(k_e/eB \rho)$. A harmonic approximation at either of these minima yields the approximate Hamiltonian...
Placing this into Eq. (11), we can retrieve an expression for the velocity of electrons moving at certain angles:

\[ v_r(\theta) = \frac{\hbar}{mr} (n + 1/2) \tan \theta. \]

Now, the divergency at \( \theta = \pm \pi/2 \) originates from the fact that the harmonic approximation breaks down when the two minima of \( V_\perp(\theta) \) meet at this angle. Apart from this, however, one can see clearly the angular separation of electrons moving in both directions and the localization in Landau-like states at \( \theta = 0 \) and \( \theta = \pi \), where the magnetic field pierces the tube wall normally (see Fig. 9).

Such a continuum model can only be expected to hold for the CNTs at low magnetic fields with \( B_A \ll \Phi_0 \). As it turns out, some of the effects visible at higher fields can be understood qualitatively by studying a model of intermediate complexity: a square lattice cylindrical tube of lattice constant \( a \). Coming from the continuum model and following Ref. 53, we can replace the continuous coordinates by integer indices: \((am,an)=(r\theta,z)\). Using a tight-binding model with on-site energy \( e_0 \) and hopping parameter \( \gamma_0 \), the Hamiltonian acts on a wave function in the following way:

\[
H \Psi_{mn} = e_0 \Psi_{(m+1)n} - \gamma_0 (\Psi_{mn+1} + \Psi_{mn-1}) - \gamma_0 (e^{-i\pi n/2} \Psi_{mn+1} + e^{i\pi n/2} \Psi_{mn-1})
\]

where the phase factor \( q(m)=\frac{2m\alpha_0}{\Phi_0} \sin \frac{\pi m}{a} \) originates from the Peierls substitution Eq. (2). As in the continuum, the invariance in the z direction can be exploited, now using a Bloch ansatz due to the discreteness of the system:

\[ \Psi_{mn} = e^{ikz} \psi_n, \]

This leads to a finite Hamiltonian for any fixed \( k_z \) \in \([-\pi/a, \pi/a] \):

\[
H_{k_z} \psi_n = -\gamma_0 (\psi_{n+1} + \psi_{n-1}) + V_{k_z}(m) \psi_n,
\]

\[
V_{k_z}(m) = e_0 - 2\gamma_0 \cos (k_z a - \pi/m a) \frac{B_A}{\Phi_0} \sin \frac{m a}{\Phi_0}.
\]

The most significant difference to the effective potential of the continuum model Eq. (10) is the replacement of the square law by a cosine one. This has the effect that the potential does not grow indefinitely for large magnetic fields, but instead oscillates, forming several minima at various angles \( \theta \), as seen in Fig. 9. In combination with the discretization of the angle, this potential leads to the formation of a complex pattern in the angular dependence of the density of states, as it can be observed is the LDOS at high magnetic fields also displayed in Fig. 8.

To capture more details in a model, an appropriate step would be the implementation of the correct dispersion at the Fermi energy: The characteristic cones at the Fermi points of graphene can be approximated by a Dirac-like Hamiltonian. For a detailed study of the magnetic spectrum of Dirac electrons on a cylindrical surface, see Refs. 45 and 54.
C. Experimentally accessible perpendicular magnetic fields

In recent experiments, optical transitions in CNTs were studied in magnetic fields up to 75 T.\textsuperscript{55} New experiments are in preparation to go up to 200 T and even 2–3 kT.\textsuperscript{56}

In perpendicular fields of this magnitude, as displayed in Fig. 10, the first onset of the band structure distortions can be seen clearly in large CNTs, comparable with the outer shell of typical MWCNTs, measuring up to tens of nm in diameter.

In particular the plots show very clearly the rapidly changing van Hove singularities, resembling those of Fig. 1 for a \( /H_{20849}^{6,6}/H_{20850} \) CNT, but at much lower magnetic field scale.

Moreover, the \( /H_{20849}^{E_F} = /H_{20849}^{E_F} /H_{20850}^{B} \) graphene state, which is due to the peculiar distortion of the Dirac-like linear dispersion into a strongly nonlinear one,\textsuperscript{18} emerges at lower fields with increasing diameters.

Most notable is the scaling law that can be found in the butterfly plot of large tubes at low fields near the Fermi energy: For two different armchair CNTs with the chiral vectors \( /H_{20849}^{m, m} /H_{20850} \) and \( /H_{20849}^{m', m'} /H_{20850} \) it can be expressed as

\[
\rho_{\text{DOS}, \text{arm}}(E, B) = \frac{m'}{m} \rho_{\text{DOS}, \text{arm}'}(m', m', B).
\]

This scaling is followed approximately already for small CNTs and becomes very precise for large diameters, converging toward a DOS that is reproducible from a model of Dirac electrons on a continuum cylinder.\textsuperscript{45} The peak at the Fermi energy also follows this scaling law. Within the region of scaling, the maximum of the peak at \( /H_{20849}^{E_F} = /H_{20849}^{E_F} /H_{20850}^{B} \) grows exponentially with the magnetic field while its integral grows linearly.

It is important to note that the scaling is not an effect of the curvature, but of the discretization of the transversal momentum, since it can be observed in graphene ribbons as well.

IV. DOUBLE-WALL CARBON NANOTUBES

While SWCNTs and MWCNTs have been studied intensely over the past 15 years, it has only recently become possible to produce DWCNTs of high purity and quality,\textsuperscript{57,58} fueling the interest in details about the interwall interaction.

Previous studies have shown an interesting interplay between magnetic fields parallel to the DWCNT axis and the interwall interaction near the Fermi energy.\textsuperscript{59} A minimal Hamiltonian of a DWCNT can be set up as

\[
\end{equation}

FIG. 10. (Color online) Zoom into the butterfly of two different armchair SWCNTs. The scales, including the color scale, are chosen according to the scaling law given in the text to produce comparable representations of the data. The tubes correspond to diameters of 2.7 nm (top) and 27 nm (bottom). The inset in the bottom panel illustrates the shape of the peak at the Fermi level. The white lines crossing the plot in the upper panel are caused by small avoided crossings in the band structure.

FIG. 11. (Color online) Butterfly plot of a \( /H_{20849}^{6,6}/H_{20850} @ /H_{20849}^{11,11}/H_{20850} \) double-wall CNT. In the upper panel, the interwall interaction is switched off, resulting in an overlay of the butterflies of two independent SWCNTs. In the lower panel, the interwall interaction gives rise to a number of new features (see text for details).

HOFSTADTER BUTTERFLIES OF CARBON NANOTUBES… PHYSICAL REVIEW B 74, 165411 (2006)
two different shells. For the angle and the absolute distance between the two shells, the rotational periodic cell, leading to two interlaced modulations. Furthermore, the system has two atoms in the rotational periodic cell, leading to two interlaced modulations.

\[ \mathcal{H} = \sum_{\langle i,j \rangle} \gamma_i(B)c_i^\dagger c_j + \sum_{\langle i,j \rangle} \gamma_j(B)c_i^\dagger c_j \]

by defining the intrawall interaction as described for SWCNTs. For the interwall interaction, we can fix the hopping coefficients as

\[ \gamma_j(B) = \beta \cos \phi \exp \left( \frac{d_{ij} - a}{\beta} \right) \times \exp \left[ -\frac{2\pi}{q_0} \left( \frac{r_j + r_i}{2} \right) \right], \]

where \( \beta = \gamma_0/8, a = 3.34 \text{ Å}, \delta = 0.45 \text{ Å}, \) and \( \phi \) and \( d_{ij} \) stand for the angle and the absolute distance between the two \( \pi \) orbitals \((i,j)\) centered at positions \( r_i \) and \( r_j \) belonging to two different shells. As a representative example, the butterfly of a (6,6) carbon nanotube (DWCNT) is displayed in Fig. 11. For the parallel magnetic field, the interwall interaction gives rise to features at two levels: looking at individual bands, one can observe Van Hove singularities crossing and sometimes avoid a crossing. The complex rules whether a crossing is avoided are not obvious from studying the butterfly plot only, but can be deduced by looking at the band structure and taking into account the various symmetries of the system. At larger scale in the butterfly plot, one finds a modulation of the pattern crossing from \( E = E_B - 8 \text{ eV} \) at \( B_t = 0 \) to \( E = E_B + 8 \text{ eV} \) at \( B_t = 45B_0 \).

To understand this phenomenon, we have studied a single unit cell in a magnetic field perpendicular to the plane of the resulting concentric ring (Fig. 12). The spectrum shows a periodic behavior of the modulation with a period \( \Phi_{\text{interwall}} = -\Phi_i/(d_{\text{CC}} \Phi_{\text{interwall}}/4) \) relating to the area of the minimal circular path between both walls. This period and the shape of the modulation are independent of the diameter of the DWCNT. A physical explanation for this modulation is as follows. Like the interaction in a two-atomic molecule, splitting two atomic orbitals into a bonding and an antibonding molecular orbital, the interwall interaction may also hybridize SWCNT states of the same energy into bonding and antibonding DWCNT states. However, the system has an approximate rotational symmetry, so the interwall interaction may only hybridize states of the same angular momentum.

At zero magnetic field, the angular momentum of the states at the bottom of the spectrum is zero in both shells. This allows hybridization, causing a split in the hybrid spectrum. At the upper end of the spectrum, the angular momentum does not match, prohibiting a hybridization. This is the cause for the strong electron-hole asymmetry visible in DWCNT butterfly plots. By switching on a parallel magnetic field, the effective angular momentum is shifted by the Aharonov-Bohm phase gathered on a circular path around the tube. This shift depends on the cross section of the path, so it is different for the two shells. Therefore, the energy at which both angular momenta match depends on the magnetic field, causing the splitting region to travel over the energy range, which leads to the visible modulation in the parallel field butterfly plots.

For a simplified model—a double-wall square lattice tube—the modulation does follow a single cosine-shaped curve, as shown in the upper panel of Fig. 12. In comparison, the DWCNT shows an additional complexity: the underlying honeycomb lattice of graphene has a unit cell containing two atoms, resulting in two intertwined cosine curves, the second just becoming visible at the edge of Fig. 11. With growing field, this effect disappears, and the plot shows no remarkable global patterns.

V. CONCLUSIONS

The magnetic spectrum of two-dimensional infinite lattice electrons gives rise to the well-known Hofstadter butterfly. In this paper, we have shown that quasi-one-dimensional lattice electrons exhibit a spectrum which does resemble the fractal structure of the Hofstadter butterfly but with a finite cutoff due to the transversal confinement. We have calculated such pseudofractals for carbon nanotubes, a material at the focus of many nanoelectronic studies also in relation to the presence of external magnetic fields. We have calculated the density of states (butterfly plots) of several single wall carbon nanotubes and we could show (i) the strong dependence of the magnetic spectrum on the underlying chiral indices; (ii) the emergence of the graphene Hofstadter butterfly at increasing nanotube diameter. In particular, perpendicular fields induce an aperiodic and pseudofractal magnetic spec-
trum. Periodic structures have been obtained for graphene ribbons, demonstrating that the aperiodicity of the perpendicular field butterfly plots is due to the incommensurability of the magnetic flux captured by elementary (hexagon) plaquettes of a CNT oriented at different angles towards the external field.

By studying the angle-resolved electronic structure of a SWCNT one can observe the emergence of snake states already predicted for nonuniform magnetic fields in a Hall bar. In our case, we have been able to devise an analytical model for the states at the top and the bottom of the energy spectrum by means of an effective mass approximation. In this latter case a continuum theory can capture the striping of the wave function along the region of the tube with zero normal field. Inversely, near the Fermi level, one cannot bypass the Dirac neutrino nature of the electronic states. We have interpreted the wave function striping by writing a Harper equation for square lattice electrons with a cylindrical geometry.

While the effects of parallel fields are of comparable simplicity in SWCNTs (being an expression of the Aharonov-Bohm oscillations due to a rigid shift of the graphene band structure sampled via the zone-folding method), this is not the case for DWCNTs. The electron-hole symmetry of π bands in SWCNTs is broken once two shells are put in interaction. The resulting hybridization of inner and outer states could be clearly understood by means of two interacting Aharonov-Bohm rings.

Experimentally relevant effects have been calculated for SWCNTs of diameter of typical external shells in MWCNTs. There, underlying multifractal structure like that of Hofstadter can be observed already at a few tens of tesla, and an outstanding scaling law for the DOS at low magnetic fields near the Fermi energy has been given. The latter applies also to graphene ribbons and is intrinsically related to the massless dispersion at the charge neutrality point.

This study, though systematic, could not include very interesting issues which also deserve careful investigation, such as the effects of disorder on the butterfly plots of SWCNTs. Disordered SWCNTs can be thought in fact a model for the external shell of MWCNTs. More atomistically one could study the influence of the interwall interaction of the structure of large diameter DWCNTs (also reasonable models for MWCNTs).63

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with J. Kono, S. Krompiewski, E. Heller, U. Rößler, and C. Strunk. Ulrich Rößler made us aware of the interesting history of science involving Douglas Hofstadter, Gregory Wannier, and Gustav Obermeir which took place in Regensburg during the mid-1970s and eventually led to the discovery of the Hofstadter butterfly thirty years ago. This work was funded by the Volkswagen Foundation under Grant No. 178340, by the European Union grant CARDEQ under Contract No. IST-021258-2 and by the Deutsche Forschungsgemeinschaft within the Collaborative Research Center SFB 689. Support from the Vielberth Foundation is also gratefully acknowledged.

APPENDIX A: HISTOGRAM METHOD

A histogram method is the simplest method to get the magnetic spectrum of a quasi-1D system. It is also very efficient if the complete energy range has to be calculated. Starting from a periodic Hamiltonian of the form:

$$H = \begin{pmatrix} H_0 & H_1 & H_{10} & H_{01} & \cdots \end{pmatrix},$$  \hspace{1cm} (A1)

one can use Bloch theorem to get an effective Hamiltonian:

$$H_{\text{eff}}(k) = H_0 + e^{iak}H_1 + e^{-iak}H_1^*$$

where $a$ is the length of the unit cell. Numerically scanning the 1D Brillouin zone $-\pi < k \ll \pi$ with a uniform distribution, one can now diagonalize the finite matrix $H_{\text{eff}}(k)$ for each value $k$. The resulting eigenvalues from this diagonalization are counted in a linear histogram over the full energy range and normalized to the total number of states. Depending on the resolution of the $k$ sampling, this histogram will become an arbitrarily good approximation to the density of states. Figures 1, 6, and 7 (CNT panels), 4 and 11 were calculated using this method.

The calculation of the data in Fig. 10 was heavily optimized by using an adaptive $k$ sampling in combination with a linear interpolation to reduce the number of diagonalizations in regions of smooth band structure and increase the precision at band edges.

APPENDIX B: GREEN FUNCTION METHOD

Another, more flexible method is that using Green functions: The bulk Green function $G(E)$ of the infinite CNT can be calculated very efficiently by the following method.64

The periodic Hamiltonian in Eq. (A1) is used as the starting point of a recursive decimation scheme:

$$H_0^{(0)}(E) = H_0,$$

$$H_{01}^{(0)}(E) = H_1,$$

$$H_{0}^{(0)}(E) = H_{10}.$$

With each recursion, the length of the effective unit cell is now doubled by decimating out every second cell:

$$H_0^{(n+1)}(E) = H_0^{(n)}(E) + H_{01}^{(n)}(E) + H_{0}^{(n)}(E) + H_{0}^{(n)}(E),$$

$$H_{01}^{(n+1)}(E) = H_{01}^{(n)}(E) + H_{01}^{(n)}(E) + H_{01}^{(n)}(E),$$

$$H_{0}^{(n+1)}(E) = H_{0}^{(n)}(E) + H_{0}^{(n)}(E) + H_{0}^{(n)}(E),$$

where $\gamma = (E + i\alpha - H_0)/\gamma$ and $\alpha$ is a small positive numerical value, chosen smaller than the desired energy resolution but large enough to provide fast convergence and numerical stability.

We acknowledge fruitful discussions with J. Kono, S. Krompiewski, E. Heller, U. Rößler, and C. Strunk. Ulrich Rößler made us aware of the interesting history of science involving Douglas Hofstadter, Gregory Wannier, and Gustav Obermeir which took place in Regensburg during the mid-1970s and eventually led to the discovery of the Hofstadter butterfly thirty years ago. This work was funded by the Volkswagen Foundation under Grant No. 178340, by the European Union grant CARDEQ under Contract No. IST-021258-2 and by the Deutsche Forschungsgemeinschaft within the Collaborative Research Center SFB 689. Support from the Vielberth Foundation is also gratefully acknowledged.

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with J. Kono, S. Krompiewski, E. Heller, U. Rößler, and C. Strunk. Ulrich Rößler made us aware of the interesting history of science involving Douglas Hofstadter, Gregory Wannier, and Gustav Obermeir which took place in Regensburg during the mid-1970s and eventually led to the discovery of the Hofstadter butterfly thirty years ago. This work was funded by the Volkswagen Foundation under Grant No. 178340, by the European Union grant CARDEQ under Contract No. IST-021258-2 and by the Deutsche Forschungsgemeinschaft within the Collaborative Research Center SFB 689. Support from the Vielberth Foundation is also gratefully acknowledged.

We acknowledge fruitful discussions with J. Kono, S. Krompiewski, E. Heller, U. Rößler, and C. Strunk. Ulrich Rößler made us aware of the interesting history of science involving Douglas Hofstadter, Gregory Wannier, and Gustav Obermeir which took place in Regensburg during the mid-1970s and eventually led to the discovery of the Hofstadter butterfly thirty years ago. This work was funded by the Volkswagen Foundation under Grant No. 178340, by the European Union grant CARDEQ under Contract No. IST-021258-2 and by the Deutsche Forschungsgemeinschaft within the Collaborative Research Center SFB 689. Support from the Vielberth Foundation is also gratefully acknowledged.
Convergence is reached for $n > n'$ if $\|H_{01}^{n' n}(E) + H_{01}^{n' n}(E)\| < \varepsilon$ for some matrix norm $\| \|$ and some small cutoff $\varepsilon$. We can then retrieve the bulk Green function from the converged $H_{01}^{n}$ as

$$G_{bulk}(E) = (E + i \eta - H_{01}^{n})^{-1}.$$  

With the original Hamiltonian (A1) expressed in a $\pi$-orbital tight-binding basis, the resulting Green function $G_{bulk}$ is a matrix in the same atomic basis of one unit cell. Therefore, the local density of states in each atom is directly given by

$$\rho_{LDOS}(E) = -\frac{1}{\pi} \text{Im}[G(E)|_{i}]$$

summing up to the $\rho_{LDOS}(E) = \sum \rho_{LDOS}(E)$. In the same run, the surface Green functions $G_{n}^{\text{surf}}(E)$ can be used to calculate the transmission through the system using the Fisher-Lee relation\(^{38}\) with a single unit cell selected as conductor, as shown in Fig. 3.

The Green function method and the histogram method both give numerical approximations to the same mathematical quantity, but their numerical errors are very different: while the former method tends to give fluctuations that show up as graining structure in flat areas of the butterfly plot, the latter suffers from sampling problems around van Hove singularities. Both errors have to be countered with very high resolution scanning and down sampling of the data. The data presented in the figures of this article typically took several hours to weeks of computation time on standard PCs [Intel(R) Pentium(R) 4, 3 GHz].

The work presented here was done using the following Open Source\(^{65}\) (R) software: Python as programming language,\(^{66}\) NumPy (Refs. 67 and 68) and SciPy (Ref. 69) for numerical computations, PyTables for data storage and handling,\(^{71}\) matplotlib for data visualization,\(^{71}\) inkscape for figure preparation,\(^{72}\) and TeXmacs for authoring.\(^{73}\)

---

33. F. Peierls, Z. Phys. 80, 763 (1933).
Appendix D. Reprints of publications

HOFSTADTER BUTTERFLIES OF CARBON NANOTUBES …

PHYSICAL REVIEW B 74, 165411 (2006)

155423 (2004).
56 J. Kono (private communication).
57 J. L. Hutchisona, N. A. Kiselev, E. P. Krimchmayac, A. V. Kres-
Hofstadter butterflies of bilayer graphene

Norbert Nemec and Gianaurelio Cuniberti
Institute for Theoretical Physics, University of Regensburg, D-93040 Regensburg, Germany
(Received 14 February 2007; published 7 May 2007)

We calculate the electronic spectrum of bilayer graphene in perpendicular magnetic fields nonperturbatively. To accommodate arbitrary displacements between the two layers, we apply a periodic gauge based on singular flux vortices of phase $2\pi$. The resulting Hofstadter-like butterfly plots show a reduced symmetry, depending on the relative position of the two layers against each other. The split of the zero-energy relativistic Landau level differs by one order of magnitude between Bernal and non-Bernal stacking.

DOI: 10.1103/PhysRevB.75.201404 PACS number(s): 73.22.-f, 71.15.Dx, 71.70.Di, 81.05.Uw

After the theoretical prediction of the peculiar electronic properties of graphene in 1947 by Wallace and the subsequent studies of its magnetic spectrum, it took half a century until single layers of graphene could be isolated in experiment and the novel mesoscopic properties of these two-dimensional (2D) Dirac-like electronic systems, e.g., their anomalous quantum Hall effect, could be measured. Inspired by this experimental success, graphene has become the focus of numerous theoretical works. For bilayers of graphene, an additional degeneracy of the Landau levels and a Berry phase of $2\pi$ were predicted to lead to an anomalous quantum Hall effect, different from either the regular massive electrons or the special Dirac-type electrons of single-layer graphene, which was confirmed in experiment shortly afterward and used for the characterization of bilayer samples.

The low-energy electronic structure of a single layer of graphene is well described by a linearization near the corner points of the hexagonal Brillouin zone (K points), resulting in an effective Hamiltonian formally equivalent to that of massless Dirac particles in two dimensions. A related Hamiltonian can be constructed featuring a supersymmetric structure which can be exploited to derive the electronic spectrum in the presence of an external magnetic field. The level at zero energy, characteristic for any supersymmetric system, maps directly to a half-filled Landau level fixed at the Fermi energy $E_F$, henceforth called the supersymmetric Landau level (SUSYLL).

In this Rapid Communication, we use the nonperturbative method pioneered in 1933 by Peierls for the implementation of a magnetic field in a model, which led Hofstadter, in 1976, to the discovery of the fractal spectrum of 2D lattice electrons in a magnetic field. Since its discovery, various aspects of the so-called Hofstadter butterfly have been studied, particularly in relation to graphene-like honeycomb structures. Featuring a large variety of topologies, all these systems have in common that the atoms inside the unit cell are located at discrete coordinates. All closed loops have commensurate areas, and the atomic network is regular enough that the magnetic phases of all links can be determined individually without the need of a continuously defined gauge field. For bilayer graphene, such a direct scheme for implementing a magnetic field is possible only for highly symmetric configurations like Bernal stacking. To handle more general configurations, such as continuous displacements between the layers, it is in general unavoidable to choose a continuously defined gauge that fixes the phase for arbitrarily placed atoms. The difficulty that arises can be seen immediately: For any gauge field that is periodic in two dimensions, the magnetic phase of a closed loop around a single unit cell must cancel out exactly, corresponding to a vanishing total magnetic flux. Conversely, this means that any gauge field that results in a nonzero homogeneous magnetic field will invariably break the periodicity of the underlying system.

A possible way to bypass this problem is based on defining a magnetic flux vortex, here oriented in the $z$ direction and located in $(x_0, y_0)$, and...
In absence of a magnetic field, the Hamiltonian—based on a tight-binding parametrization originally used for multilayered carbon nanotubes\textsuperscript{20,23}—consists of a contribution for nearest neighbors within a layer (\(i, j\)) and one for pairs of atoms located on different sheets (\((i, j))\):

\[
\mathcal{H} = \sum_{(i, j)} \gamma_{ij}^\text{intra} \, \hat{c}_i^\dagger \hat{c}_j + \sum_{(i, j)} \gamma_{ij}^\text{inter} \, \hat{c}_i^\dagger \hat{c}_j.
\]

In absence of a magnetic field, the intralayer hopping is fixed to \(\gamma_{ij}^\text{intra} = 2.66\ eV\), while the interlayer hopping depends on the distance only,

\[
\gamma_{ij}^\text{inter} = \beta \exp(-|\mathbf{r}_i - \mathbf{r}_j| / \delta).
\]

with \(\beta = \gamma_0 / 8\), \(\alpha = 3.34\ \text{Å}\), and \(\delta = 0.45\ \text{Å}\). A cutoff is chosen as \(r_{\text{max}} = \alpha + 5\delta\). Following the Peierls substitution,\textsuperscript{18} the magnetic field \(B\) is now implemented by multiplying a magnetic phase factor to each link between two atoms \(i\) and \(j\):

\[
\gamma_{ij}(B) = \gamma_{ij}(B = 0) \exp\left(\frac{2\pi}{\Phi_0} \int_{\mathbf{r}_i}^{\mathbf{r}_j} A_B(r) \cdot dr\right),
\]

where the integral is computed on a straight line between the atomic positions \(\mathbf{r}_i\) and \(\mathbf{r}_j\).

For the bilayer graphene, we arrive thus at a periodic Hamiltonian with a two-dimensional unit cell containing four plaquettes:

\[
\mathcal{H}_{\text{plaquette}}(B) = (3/2) \Phi_0 |B|, \quad \text{where} \ |B| = |B_{\text{CC}}| = 1.42\ \text{Å} \text{ is the intralayer distance between neighboring carbon atoms. The }\]

FIG. 2. (Color online) Hofstadter butterfly of a bilayer graphene in two differently shifted configurations. Top panel: \(AA\) stacking (two layers exactly aligned). The band structure for this highly symmetric stacking (same rotational symmetry as for Bernal stacking in Fig. 1) shows the single-layer cone simply split up in energy. Bottom panel: Intermediate position between Bernal and \(AA\) stacking. The rotational symmetry is broken and the bands split into two cones at different offsets from the \(K\) point and different energies. The straight lines overlaid at the energy minimum and maximum are the regular Landau levels of the massive bands. Near \(E_F\), one can make out the parabolic traces of the relativistic Landau levels and the horizontal lines of the SUSYLLs (see text). Insets at the lower right of each panel: DOS of a finite-width ribbon shows the corresponding behavior in each case.

atoms and spanning the area of one hexagonal graphene plaquette: \(A_{\text{plaquette}} = (3/2) d_{\text{CC}}^2\), where \(d_{\text{CC}} = 1.42\ \text{Å}\) is the intralayer distance between neighboring carbon atoms. The
HOFSTADTER BUTTERFLIES OF BILAYER GRAPHENE …

The effect of a perpendicular magnetic field, measured in flux per plaquette $\Phi = A_{\text{plaquette}} / H$, can be calculated for commensurate values $\Phi = (p/q) \Phi_0$ ($p,q$ integers) by constructing a supercell of $q$ unit cells. The corresponding Bloch Hamiltonian $H(k)$ is a $4q \times 4q$ matrix that can be diagonalized for arbitrary values of $k$ in the two-dimensional Brillouin zone of area $4\pi^2 / q^2_{\text{plaquette}}$.

To obtain the butterfly plots as displayed in Figs. 1 and 2, we chose $0 < p < q = 512$, reducing the fraction $p/q$ for efficiency. For each value of $\Phi$ the density of states was calculated from a histogram over the spectral values for a random sampling of $k$ over the Brillouin zone. The number of sampling points was chosen individually for different values of $p$ to achieve convergence. In Figs. 1 and 2, the Hofstadter spectra of three differently aligned graphene bilayers are presented. The Bernal stacking (Fig. 1) stands out, as it is the configuration of layers in natural graphite. Alternative configurations like AA stacking were found in ab initio calculations to be energetically unfavorable; they can, however, be thought of as either mechanically shifted samples or sections of curved bilayers (e.g., sections of two shells in a large multivall carbon nanotube) where the alignment unavoidably varies over distance. Compared to the Hofstadter butterfly of a single sheet of graphene,22 two asymmetries are visible in all three plots: The electron-hole symmetry ($E = -E$) is broken down by the interlayer coupling already at zero magnetic field: while the lowest-energy states of a single graphene layer have constant phase over all atoms and can couple efficiently into symmetric and antisymmetric hybrid states of the bilayer system, the states at high energies have alternating phases for neighboring atoms, so interlayer hybridization is prohibited by the second-nearest-neighbor interlayer coupling. For low magnetic fields, two sets of Landau levels can therefore be observed at the bottom of the spectrum, indicating a split of the massive band of graphene at the $\Gamma$ point ($E_{\text{min}} = -3.29$, $m_\text{r} = 2h^2 / (3 \ell^2 m_0)$) into two bands at different energies and with different effective masses [$E_{\text{min}} = E_{\text{max}} = 1.1$ eV, $m_\text{r} = 2m_0 / (1 + 21/2\ell / \gamma_c)$], independent of the relative shift of the two layers; see the straight lines overlaid in the bottom panel of Fig. 2]. At the top of the spectrum, where the split is prohibited, only one degenerate set of Landau levels appears, as in single-layer graphene. The original periodic symmetry along the $B$-field axis at one flux quantum per graphene plaquette is broken down due to the smaller areas formed by interlayer loops. The breaking of this symmetry is comparably small in the AA-stacking configuration (Fig. 2, top) where loops of the full plaquette area are dominant. In the other two configurations smaller loops are more dominant, so the periodicity is perturbed more severely. In the intermediate configuration (Fig. 2, bottom), the fractal patterns appear slightly smeared out for high magnetic fields, due to the reduced symmetry of the system.

The right insets of Figs. 1 and 2 display the spectra of (200.0) bilayer graphene nanoribbons,25 each in a corresponding configuration, obtained by a method described before26 that allows handling of continuous magnetic fields. For low magnetic fields, these spectra are strongly influenced by finite-size effects.21 Only for magnetic fields larger than $B = 4\Phi_0 / d^2$, which for a ribbon of width $d = 50$ nm relates to $\sim 7$ T, do the spectra of two-dimensional bilayer graphene begin to emerge. Prominent in all three insets are the dark, horizontal pairs of lines at the center, the supersymmetric Landau levels. While these represent discrete levels in two-dimensional graphene sheets, they are broadened by the finite width of the ribbon to a peak of the same shape as in carbon nanotubes.23,24 The mesoscopic character of these split SUSYLLs in dependence on the width $W$ of the ribbon is captured by the functional form of the density of states:

$$\rho(E,W) = f((E - E_0) W, B W^2)$$

where $E_0$ is the position of the maximum. Single-layer graphene is known to feature an anomalous supersymmetric Landau level at the Fermi energy.2,4,17 Neglecting Zeeman splitting, this level is fourfold degenerate (two spin, twice valley) and half filled. For bilayer graphene in Bernal stacking (Fig. 1) the SUSYLLs of the two layers have been shown to be protected by symmetry and to remain degenerate, giving in total an eightfold degeneracy.13 In Fig. 2, this degeneracy can be observed to be lifted for displaced bilayers, leading to a split of the SUSYLLs.
SYLL into a bonding and an antibonding hybrid state in the two layers, each fourfold degenerate. The continuous evolution of the split for varying displacement of the two layers against each other is displayed in Fig. 3. The split reaches its maximum of \( \Delta E \sim 0.3 \) eV for the AA-stacking configuration and is minimal for Bernal stacking. For simpler tight-binding parametrizations that take into account only first- and second-nearest-neighbor interlayer hoppings, the degeneracy in the Bernal configuration is known to be exact.\(^{13}\) Here, in contrast, this degeneracy is split by \( \Delta E \sim 0.01 \) eV due to interlayer hoppings of a longer range, similar to the effect caused by second-nearest-neighbor interactions within one layer.\(^{33}\)

In conclusion, we have developed a method that allows the nonperturbative implementation of a magnetic field in periodic systems with arbitrarily positioned atoms. A \( \pi \) orbital parametrization for graphic interlayer interactions with arbitrary displacements was then used to calculate the Hofstadter spectrum of bilayer graphene in various configurations, revealing common features like electron-hole symmetry breaking, and differences, especially in the breaking of the magnetic-field periodicity. A close look at the supersymmetric Landau level at low fields near the Fermi energy revealed a breaking of the previously found symmetry, resulting in a split of the level, depending on the lateral displacement of the two graphene layers against each other.

We acknowledge fruitful discussions with I. Adagideli, C. Berger, V. Fal’ko, F. Guinea, and H. Schomerus. This work was funded by the Volkswagen Foundation under Grant No. I/78 340 and by the European Union grant CARDEQ under Contract No. IST-021285-2. Support from the Vielberth Foundation is also gratefully acknowledged.
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography

Materials: Carbon nanotubes in an ancient Damascus sabre. 


Bibliography


Bibliography


Acknowledgments

Although a few words do not do justice to their contribution, I would like to thank the following people for making this work possible:

- **Prof. G. Cuniberti**, my supervisor, for believing in me all the time, for all his patience with me and for teaching me infinitely more than just the tricks of the trade

- **Prof. D. Tománek** for the welcome in his group at Michigan State University over one month and for the inspiring cooperation on our work about contacts

- **Prof. K. Richter** for being such a great host to our whole group in Regensburg; him as well as **T. Breu** and **V. Agarwal** for their cooperation and fruitful discussion in the work on disorder and mesoscopic length scales

- **Prof. S. Krompiewski** and **N. Ranjan** for their cooperation in the work on spin transport

- **D. Darau** and **G. Niebler** for the great work during their diploma projects and the cooperation on telescopic carbon nanotubes respectively defects in graphene nanoribbons, as well as for their help in proof-reading

- **Prof. M. Grifoni**, **Dr. M. Marganska** and **Dr. S. Wang** for the fruitful discussions about incommensurability and about magnetic fields

- **Dr. I. Adagideli** for his help in the understanding of the supersymmetry in graphene

- **Prof. R. Römer** for his help on the issue of strong localization

- **Prof. C. Strunk** and his group for keeping us all in touch with experimental reality
Acknowledgments

- Dr. M. del Valle, Dr. R. Gutiérrez and F. Pump for being such great and dependable friends and colleagues and for their help in proof-reading

- all the members of our Molecular Computing Group for making this such a great place for working and for all the fun that we had over the past years; special thanks to Dr. M. Gheorge for shaping our whole group in her incomparable, cordial way

- Dr. C. Eltschka for his help in proof-reading and in solving crossword puzzles

- A. Reisser for being the best secretary one could imagine and for providing coffee, crossword puzzles and soul for our daily coffee-table

- all the members of our daily coffee-table for all the sociability and the fun of solving crossword puzzles

- all the great Open Source software authors for providing better tools than money could buy

- my dear friends M. Boyer, T. Burtsche, S. Fischer and D. Häseker for providing such great opportunities to spend physics-free weekends

- my parents for me being around at all and for their never ending love and support

- my wife Lydia, my son-to-be Paul and our friend S. Gonsor for all the love, the great support and the patience especially during the hot phase of the past weeks

This work was funded by the Volkswagen Foundation under Grant No. I/78 340, by the European Union grant CARDEQ under Contract No. IST-021285-2 and by the Deutsche Forschungsgemeinschaft within the Collaborative Research Center SFB 689. Support from the Vielberth Foundation is also gratefully acknowledged.