SPIN STATES AND SPIN-ORBIT COUPLING IN NANOSTRUCTURES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ferdinand Kuemmeth

January 2008

© 2008 Ferdinand Kuemmeth

ALL RIGHTS RESERVED

SPIN STATES AND SPIN-ORBIT COUPLING IN NANOSTRUCTURES

Ferdinand Kuemmeth, Ph.D.

Cornell University 2008

This dissertation describes electronic transport measurements which we performed on two different nanoscale systems: quantum dots based on gold nanoparticles approximately 10 nm in diameter, and quantum dots based on partially suspended carbon nanotubes a few hundred nanometers long. Both experiments employ tunneling spectroscopy at milliKelvin temperatures to study the effects of magnetic fields and spin-orbit interaction on the discrete spin-states of each system.

Using break junctions on top of gate electrodes made from oxidized aluminum, we were able to study the discrete spectra of individual gold nanoparticles as a function of their charge states. When adding electrons to the quantum dot, doubly-degenerate levels are filled as in a non-interacting model. This result is consistent with theoretical calculations which found weak electron-electron interactions for noble metals. We also fabricated single-electron transistors based on chemically synthesized gold nanoparticles, and observed effects of spin-orbit interaction and mesoscopic fluctuations on their discrete level spectra.

In the second experiment we suspended small-bandgap nanotubes in between two independent gate electrodes fabricated from silicon-on-insulator substrates. The spectra of such quantum dots depended on the electric field produced by the gate electrodes as well as a magnetic field applied parallel to the nanotubes.

We were able to accumulate electrons as well as holes on the quantum dot, and measured accurately the magnetic moment associated with their ground states as well as excited states. At zero magnetic field we observed a shell splitting which we attributed to spin-orbit interaction.

BIOGRAPHICAL SKETCH

Ferdinand Kümmeth was born as a rather unexpected surprise to his parents on March 9th, 1979 in Landsberg am Lech, Germany — entering the world some 10 minutes after his mother had already given birth to his two triplet siblings. He also has an older sister and an older brother with whom he spent a wonderful childhood growing up in a spacious farmhouse in the village of Ludenhausen. After attending the Ignaz-Kögler Gymnasium in Landsberg, collecting vintage tractors and playing a lot of classical music on his concert harp, he backpacked through Australia for a couple of months before he started studying "Diplom-Physik" at the Universität Augsburg in 1998. In the summer of 2001 he transferred to Cornell University on a Fulbright Scholarship and soon thereafter decided to obtain a Ph.D. under the guidance of Professor Dan Ralph. Ferdinand's next stop in life will be in Boston, doing postdoctoral work with Professor Charlie Marcus.

Für meine Großmutter, Eltern und Geschwister.

ACKNOWLEDGEMENTS

The last few years have been a great experience. I am grateful for having been exposed to many exciting experiments and for having met numerous wonderful people. First and foremost, I would like to thank my advisor Dan Ralph for making this possible. He taught me solid state physics as well as many tricks to survive in the lab, and always helped out if things went wrong. Thank you for your support and expertise whenever and wherever it was needed. Working in your group was a tremendous privilege which I will not forget. It was a lot of fun to work in lab, and I would like to thank everyone in the Ralph group. Alex was a thoughtful guide during my first year and a great source of advice and fun thereafter. Kirill has never failed to surprise me with his creative style and wide interests, and working together with him was always rewarding. Thanks to Abhay and Jason for teaching me nanofabrication, and to Sergey, Jack, Jacob, Kiran, Janice and Josh for their daily help and entertainment. Over the last year it has been fun to see Sufei, Yongtao, Eugenia and Chen become experts in nanofabrication, and I wish them all the best in their experiments. Thank you all for making this a fun group!

Thanks to Paul for letting me work in his lab as well. Playing with carbon nanotubes was fun and I learned many new things. Thank you Shahal for your stimulating ideas, your help in lab and your many insights and discussions.

I would like to thank my committee members for their help and advice whenever it was needed: Thanks to Paul McEuen, Piet Brouwer, Vinay Ambegaokar and David Muller. I would also like to thank Eric Smith, Bob Snedecker and Mick Thomas for making the basement of Clark Hall such a great place for experiments, and Rob Ilic and all the CNF staff for their help in making our devices small.

Life outside lab has never been boring thanks to Jim, Suniti, Dan, Jay, Faisal, Saikat, Jahan, Amar, Vera, Markus, Arend, Sarah, and many others. I would especially like to thank my roommates for their true friendship, encouragement and fun additions over the years. Thank you Louis for sailing, caving and Canadian adventures, thank you Edoardo for your love and passion for Italian cuisine and fine drinks, thank you Ben for fixing my model airplanes and teaching me telemarking, and thank you Marie for keeping me afloat during windsurfing and thesis writing.

I would like to thank Xak Bjerken for helping me store my harp in the music department, and Josh Parks for joining me on his violin on many occasions. It has been great playing with you over the last few years. Thank you for tolerating my mistakes and straightening my chords when I went out of tune.

My move to Ithaca would have not been possible without the help and encouragement from family and friends. Thank you Tobi, Natalie, Eva, and Tobias for advice and keeping my German fluent. Last but not least, I would like to acknowledge the loving support from my family in Germany. Thank you Mom and Dad, for your continuous encouragement and understanding over the years, and thank you Stephanie, Franz, Lisa and Max for always keeping in touch.

TABLE OF CONTENTS

Biographical Sketch **iii**

Dedication **iv**

Acknowledgements **v**

List of Figures **ix**

1 Gold, Carbon and Spin-orbit coupling **1**
- 1.1 Motivation . . . 1
- 1.2 Gold nanoparticles . . . 4
- 1.3 Spin-orbit interaction in gold nanoparticles . . . 7
- 1.4 Carbon-nanotube quantum dots . . . 10
- 1.5 Bandstructure of graphene . . . 11
- 1.6 Spin-orbit interaction in graphene . . . 15
- 1.7 Spin-orbit coupling in nanotubes . . . 20
- 1.8 Probing quantum states with a single-electron transistor . . . 26

References **35**

2 Fabrication of metal nanoparticle SETs with efficient gating **38**
- 2.1 Motivation . . . 38
- 2.2 Fabrication of the electrodes . . . 39
- 2.3 Evaporation of nanoparticles . . . 52
- 2.4 Absence of electron-electron interactions . . . 65
- 2.5 Remarks on this fabrication technique . . . 72

References **77**

3 Spin-orbit interaction in chemically synthesized gold nanoparticles **79**
- 3.1 Summary . . . 79
- 3.2 Fabrication . . . 80
- 3.3 Spin-orbit interaction in chemically synthesized gold nanoparticles . 94
- 3.4 Outlook . . . 108

References **110**

4 Fabrication of partially suspended, split-gate carbon nanotube devices **112**
- 4.1 Motivation and summary of fabrication . . . 112
- 4.2 Device schematic and design criteria . . . 113
- 4.3 Fabrication I: Whole wafer processing . . . 120

4.4 Fabrication II: Cleaving and final processing 131
4.5 Testing devices using a cryogenic probe station 135
4.6 Measurements in a dilution refrigerator 139

References **144**

5 Few-electron carbon nanotube quantum dots **145**
5.1 Significance of this work 145
5.2 Electric polarizability 146
5.3 Magnetic polarizability 151
5.4 Configuration changes induced by magnetic and electric fields . . . 157
5.5 Shell splitting in zero magnetic field 170
5.6 Outlook 176

References **179**

LIST OF FIGURES

1.1 Two very different quantum dots . . . 2
1.2 Bandstructure and spin-orbit scattering in gold . . . 6
1.3 Band structure of carbon nanotubes . . . 12
1.4 Spin-orbit interaction in graphene . . . 17
1.5 Spin-orbit interaction in nanotubes . . . 21
1.6 SET circuit schematic . . . 29
1.7 SET as a spectrometer . . . 30
1.8 Probing occupied and unoccupied states . . . 33

2.1 Fabrication . . . 40
2.2 Patterning of gate electrodes . . . 44
2.3 Electromigration . . . 49
2.4 Electromigration and grain structure . . . 50
2.5 Aluminum break junctions . . . 53
2.6 Evaporated Au, Cu, Pd and Pt nanoparticles . . . 54
2.7 Diagnosis via SEM . . . 56
2.8 4.2 K Coulomb diamonds of evaporated nanoparticle devices . . . 59
2.9 Coulomb stairs . . . 61
2.10 Discrete states while gating . . . 64
2.11 Spectra from different charge states . . . 68
2.12 Even-odd level filling . . . 69
2.13 Copper nanoparticle devices . . . 73
2.14 Large g-factors . . . 74

3.1 Various chemically synthesized nanoparticles . . . 81
3.2 TEM micro graphs of synthesized gold nanoparticles . . . 82
3.3 Deposition of NPs from solution . . . 84
3.4 From the break junction to the finished device . . . 85
3.5 Coulomb diamonds and discrete states . . . 87
3.6 Fluctuations of tunnel couplings . . . 89
3.7 More than 40 resolved resonances . . . 90
3.8 Correlation between N and N+1 excitations spectra . . . 91
3.9 G-factors from a symmetric device . . . 93
3.10 G-factors and level statistics . . . 101
3.11 Avoided level crossings and mesoscopic fluctuations . . . 105
3.12 Fluctuating g-factors . . . 107
3.13 Au(Fe) and Pd particles . . . 109

4.1 Device Schematic . . . 114
4.2 Variation of device layer thickness . . . 118
4.3 Etching the device layer . . . 125
4.4 Oxide removal and brushcleaning . . . 127
4.5 E-beam patterning of source and drain . . . 129

4.6 Challenges of tungsten and thin PMMA . . . 130
4.7 Crack guides . . . 132
4.8 Growth of nanotubes . . . 133
4.9 Testing devices with a parameter analyzer . . . 136
4.10 Conductance at 77 and 17 K . . . 137
4.11 Signatures of a double dot . . . 138

5.1 Polarizability of quantum states . . . 148
5.2 Even-odd magnetic field dependence . . . 152
5.3 Inconsistency between shell model and data . . . 153
5.4 The 3rd Coulomb peak . . . 156
5.5 More Coulomb oscillations . . . 158
5.6 Bright states and dark states . . . 159
5.7 Configuration changes . . . 162
5.8 Electron and hole Coulomb oscillations . . . 164
5.9 Detuning the 2nd Coulomb peak . . . 166
5.10 Exchange versus spin-orbit interaction . . . 168
5.11 One-electron excitations . . . 171
5.12 Shell splitting of the first electron . . . 172
5.13 Shell splitting of the first hole . . . 175
5.14 Outlook . . . 177

Chapter 1

Gold, Carbon and Spin-orbit coupling

1.1 Motivation

The invention of the ^{3}He-^{4}He dilution refrigerator in the 1960s and the advancement of nanofabrication techniques during the past 20 years have enabled the powerful tool of spectroscopy to be applied to nanoscale solid state systems. In this thesis, we use the single-electron transistor, operated at sufficiently low temperature, to probe individual quantum states of two very different materials: quantum dots based on gold nanoparticles and quantum dots based on carbon nanotubes (Figure 1.1). These quantum dots are very different in terms of their host, size, number of electrons, and symmetries. Carbon nanotubes inherit shell structure from the structural symmetries of graphene, whereas metal nanoparticles are characterized by chaotic single-electron dynamics and mesoscopic fluctuations. Despite these differences, among which is the vast difference in atomic number for gold and carbon, we find that in both types of quantum dots their electronic spectra are modified by the same fundamental interaction, namely the relativistic coupling between the electron's spin and orbital motion.

In the remainder of this chapter, we introduce the host materials and explain why they are interesting. We discuss basic theoretical concepts of spin-orbit interaction in gold nanoparticles and carbon nanotubes, and conclude with an introduction to single-electron transistors (SETs) and tunneling spectroscopy.

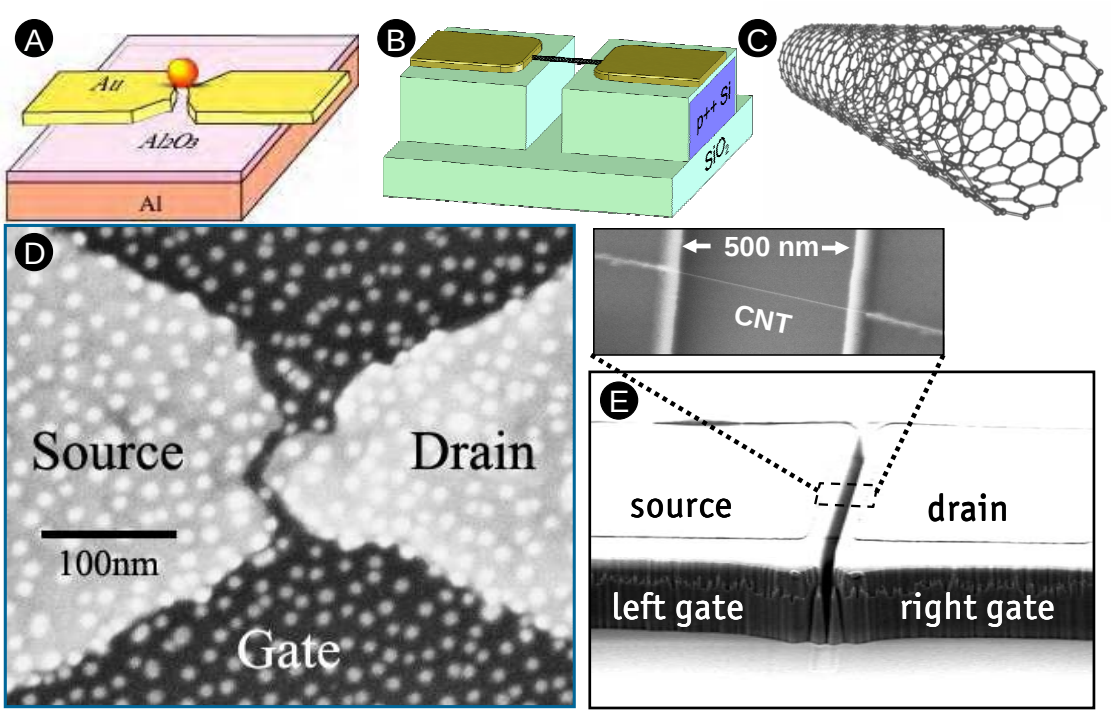


Figure 1.1: Schematic of single-electron transistors based on metallic nanoparticles (A) and partially suspended carbon nanotubes (B), and images of actual devices (D,E) obtained with a scanning-electron microscope. C) Schematic of a carbon nanotube (from [51]).

Table 1.1: The two types of quantum dots studied in this thesis (Fig. 1.1) are very different due to their different sizes and different host materials. Despite these myriad differences, we observe in this thesis that the discrete level spectrum of both quantum dots are modified by the same fundamental interaction between an electron's spin and its orbit ("spin-orbit interaction").

	Carbon ^{6}C	Gold ^{79}Au
Electron configuration	$1s^2 2s^2 2p^2$	$[Xe]4f^{14}5d^{10}6s^1$
Atomic spin-orbit splitting	∼10 meV	∼1.5 eV
Crystal structure	hexagonal $d = 2.46$ Å (graphene)	FCC $d = 4.08$ Å
Fermi wavelength	$>> d$	5.2 Å
Quantum dot size	0.5 μm	10 nm
Charging energy	∼10 meV	tens of meV
Level spacing	∼3 meV	∼0.2 meV
Degeneracy	4 (graphene)	2

1.2 Gold nanoparticles

Bulk gold is a dense, non-magnetic metal with fcc crystal structure and one conduction electron per unit cell. It is a very heavy element, in which relativistic effects are readily observed[1] even with the naked eye: the distinct color of gold (with a sudden onset of absorption at 520 nm) is due to a relativistic lowering of the $6s$ and raising of the $5d$ bands.

In this thesis relativity (in the form of spin-orbit interaction) meets quantum mechanics, by confining electrons to a very small grain of gold. Because of the high electron density, the Fermi wavelength is very small ($\sim$0.5 nm) and the nanoparticle needs to small ($\sim$10 nm) in order to show discrete quantum states in a dilution refrigerator (100 mK). Unlike in the carbon nanotube quantum dots, the number of conduction electrons is still in the many thousands.

We can estimate the spin-degenerate level spacing δ expected for a metallic nanoparticle from the free-electron expression for the density of states at the Fermi level [57]:

$$\delta = \frac{2\pi^2\hbar^2}{mk_F\mathrm{Vol}} = \frac{1.50\ \mathrm{eV\ nm}^2}{k_F\ \mathrm{Vol}} \tag{1.1}$$

which for a spherical gold grain of diameter 10 nm gives 0.24 meV (using $k_F = 12.1$ nm^{-1} for gold). The challenge to control the nanoparticle's size is taken up in Chapter 2.

Because the conduction band of gold is a mixture of $6s$ and $5d$-orbitals (Figure 1.2) the electronic states of the quantum dot are affected by spin-orbit interaction. However, this happens in a random fashion due to the random shape of a

[1] For example, the relativistic contraction of the innermost s-electron leads to the contraction of all outer shells (orthogonality requirement), to stronger bonds compared to silver, and eventually to a lattice constant slightly smaller than silver [63, 64].

nanoparticle. This also removes any orbital degeneracies which would result from a spherical shape (Section 3.3). Therefore, at $B = 0$ T each quantized state is only two-fold degenerate (due to Kramer's theorem). In the presence of spin-orbit coupling, spin is not a good quantum number, and each quantum state Zeeman splits with its own distinct g-factor [31].

Metals have large charging energies (10s of meV), hence it is difficult to change the total number of electrons via a gate electrode. Unlike in carbon nanotube quantum dots, the gate capacitance ($\leq$1 aF) is smaller than the capacitances associated with the tunnel barriers. This challenge of efficient gating in metal nanoparticle SETs is taken up in Chapter 2.

Metal nanoparticles are interesting because they allow the study of various fundamental interactions on the nanoscale. Previous experiments have addressed superconductivity [15, 61], ferromagnetism [60], spin-orbit interaction [2, 15] and electron-electron interactions[2] [62]. Gold in particular is interesting because it has been studied extensively in bulk and in thin films. Furthermore, its electronic properties can be modified by magnetic impurities, turning it into a dilute ferromagnet [25], a spin glass [33], or a Kondo alloy [31, 32]. Each of these interactions are interesting to study on the nanoscale, when the level spacing becomes comparable to these interactions. We believe that chemically synthesized metal nanoparticles (Chapter 3) open the door to many interesting experiments on the nanoscale.

[2]For relativistic calculations of interaction effects on the electronic structure of noble metals see [34].

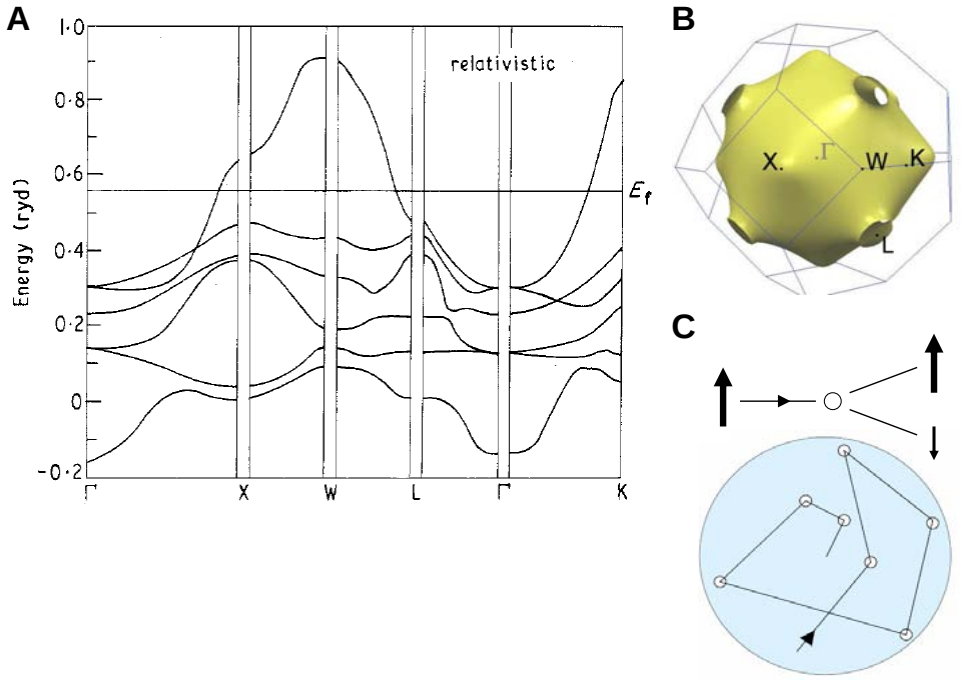


Figure 1.2: A) The Fermi surface of FCC gold in the 1st Brillouin zone includes 1 electron per atom, and deviates from the nearly-free electron model near the hexagonal face (neck) and near the square face (belly) [49]. B) Band energies of gold are affected by relativistic effects, raising the 5d bands and lowering the 6s bands [50]. The conduction band marked by E_f is a mixture of 6s and 5d orbitals, but the effects of spin-orbit coupling ($\sim$ 1.5 eV for the atomic 5d-shell) is quenched due to the crystal field. C) The dominant mechanism of spin-orbit scattering in a metal nanoparticles is through momentum scattering at boundaries and defects: Each time a momentum scattering event occurs (indicated by $\circ$) the spin of the electron flips with a small probability. Therefore, spin is not a good quantum number in the presence of spin-orbit scattering [14, 31, 32].

1.3 Spin-orbit interaction in gold nanoparticles

Spin-orbit interaction and the tight-binding model

The fundamental Hamiltonian describing spin-orbit interaction follows from the non-relativistic limit of the Dirac-equation [29] and is given by

$$H_{SO} = \frac{\hbar^2}{4m^2c^2}(\nabla V \times \boldsymbol{p}) \cdot \boldsymbol{s} \tag{1.2}$$

where $\boldsymbol{p}$ and $\boldsymbol{s}$ are the momentum and spin operators of an electron in an electric potential V. In the central field approximation ($V(\boldsymbol{r}) = V(r)$) this reduces to the well known form

$$H_{SO} = \xi(\boldsymbol{r})\boldsymbol{L} \cdot \boldsymbol{s} = \xi(\boldsymbol{r}) \left[\frac{L_+ s_- + L_- s_+}{2} + L_z s_z \right] \tag{1.3}$$

where $\xi(\boldsymbol{r}) = \frac{1}{2m^2c^2r}\frac{dV}{dr}$ and $\boldsymbol{L}$ and $\boldsymbol{s}$ are the usual total atomic angular momentum operator and total electronic spin operator respectively. The Hamiltonian of a free atom then reads

$$H_{atom} = \frac{p^2}{2m} + V(\boldsymbol{r}) + \xi(\boldsymbol{r})\boldsymbol{L}\boldsymbol{s} \tag{1.4}$$

and the tight-binding Hamiltonian becomes [48]

$$H_{TB} = \frac{\boldsymbol{p}^2}{2m} + \sum_j V(\boldsymbol{r} - \boldsymbol{R}_j) + \sum_j \xi(\boldsymbol{r} - \boldsymbol{R}_j)\boldsymbol{L}_j\boldsymbol{s}_j \tag{1.5}$$

where the sum is over all lattice sites $\boldsymbol{R}_j$. This Hamiltonian does not commute with the total spin and hence the resulting Bloch waves are not pure spin states. However, each state is still two-fold degenerate due to time reversal symmetry (Kramer's degeneracy) [31, 48].

g-factors in nanoparticles

Similar to the Bloch waves of a bulk crystal, eigenstates of a metal nanoparticle are not pure spin states, but linear superpositions of spin-up ($|\uparrow\rangle$) and spin-down ($|\downarrow\rangle$) states. Due to Kramer's degeneracy each orbital state (labeled by n) is two-fold degenerate (labeled by $''\uparrow''$ and $''\downarrow''$) and we may write

$$|n''\uparrow''\rangle = \alpha_n|\uparrow\rangle + \beta_n|\downarrow\rangle \tag{1.6}$$

and its time reversed partner

$$|n''\downarrow''\rangle = -\beta_n^*|\uparrow\rangle + \alpha_n^*|\downarrow\rangle \tag{1.7}$$

The Kramer's degeneracy is lifted by the interaction $\mu_B \boldsymbol{B}\cdot(\boldsymbol{L}+2\boldsymbol{s})$ with a magnetic field $\boldsymbol{B}$, and the g-factor g_n of a state may be defined as the matrix element of $2(\boldsymbol{L}+2\boldsymbol{s})$ in the direction of the magnetic field. In the absence of spin-orbit interaction the orbital angular momentum $\boldsymbol{L}$ is quenched [12] and Zeeman splitting due to the spin $\boldsymbol{s}$ is the only contribution. This leads to a constant g-factor of $g_n = 2$ for all quantum levels of the nanoparticle. However, in the presence of weak spin-orbit coupling each state is a mixture of spin-up and spin-down, and hence the Zeeman splitting is reduced[3]. Moreover, the g-factor is different for each quantum level. This can be understood from Elliott-Yafet scattering [31, 46] which constitutes the main spin-orbit scattering mechanism in metals: Each time a momentum scattering event occurs, the spin of an electron also flips with some small probability [14], giving rise to some mean spin scattering time τ_{SO}. This is schematically indicated in Figure 1.2C. Because scattering sites are distributed randomly in a metallic nanoparticle (for a ballistic nanoparticle, most scattering

[3]The effect of spin-orbit interaction on the spin-susceptibility of small metal particles was studied by J. Sone [43].

occurs on the boundary) the effective spin-orbit matrix element is different for each quantum level [9].

Avoided level crossing

In addition to reduced g-factors which depend on the microscopic details of a wavefunction, spin-orbit interaction also manifests itself as avoided level crossings between neighboring levels, as a function of magnetic field [9]. The magnitude of the avoided level crossing undergoes mesoscopic fluctuations as well. Brouwer *et al.* [8,9,54] and Matveev *et al.* [12] worked out sophisticated theories based on random matrix theory to describe the effects of spin-orbit interaction and mesoscopic fluctuations in metallic nanoparticles. The physical picture behind the random properties of the spin contribution lies in its coupling to the highly oscillatory, essentially random orbital wavefunctions of the nanoparticle. A sufficient coupling is established already by weak spin-orbit interaction.

Strong spin-orbit coupling

For the case of strong spin-orbit interaction (as appropriate for gold nanoparticles) it should be noted that the spin becomes totally randomized and hence the spin-contribution to the mean g-factor becomes small. For a ballistic nanoparticle, the dominating contribution then stems from the orbital magnetic moment. This is because in the limit of strong spin-orbit interaction ($\hbar/\tau_{SO} >> \delta$), wave functions are complex and do not have well-defined spin[4]. However, they do carry a current, and Matveev *et al.* estimate the corresponding orbital magnetic moment as $e|A|\delta/\hbar$ where A is the directed area covered by a trajectory corresponding to

[4]One may say that during the electron motion along a closed trajectory the spin flips many times [12].

level n. To find A they notice that during the time $\hbar/E_T$ (here $E_T = D/L^2$ is the Thouless energy and D the diffusion constant for electrons in the grain) the electron travels across the nanoparticle and covers the area $\sim L^2$. During the period of motion $\hbar/\delta$ the electron bounces off the boundaries $\sim E_T/\delta$ times. Since the direction of motion after each bounce is random, they estimate the total directed area as $|A| \sim L^2\sqrt{E_T/\delta}$. They conclude that the orbital magnetic moment is of order $2\mu_B \frac{m}{m^*}\sqrt{l/L}$ where l is the transport mean free path of electrons and and m^* is of the order of the effective mass derived from the density of states of the bulk material. More rigorously, and including both spin and orbital contributions, they obtain for the mean g-factor (averaged over different levels in a single particle)

$$<< g^2 >>= \frac{3g_0^2}{2\pi\hbar}\delta\tau_{S0} + \alpha\frac{l}{L} \tag{1.8}$$

where the dimensionless constant α is of order 1 (it depends on the geometry) and $g_0 = 2$ is the g-factor of the free electron. This important result indicates that for a ballistic nanoparticle (i.e. $l \sim L$) with strong spin-orbit coupling (i.e. $\delta\tau_{SO} << \hbar$), the orbital contribution to the g-factor is of order 1, whereas the spin contribution is very small. We will confirm this result experimentally in Chapter 3.

1.4 Carbon-nanotube quantum dots

Carbon is light compared to gold and shows an atomic spin-orbit splitting of ∼10 meV in its $2p$ shell [14]. It has a filled $1s$ shell and four $2p$-electrons. Presumably because of its soft core (lacking inner shells compared to Si and Ge [67]) it does not follow the trend of period IV toward the diamond structure as the lowest allotropic form. When arranged in a hexagonal lattice carbon atoms bond in a very neat way which gives graphene and nanotubes their unique electronic structure: an

electron-hole symmetric, linear dispersion near the K and K' points. The Dirac physics associated with this condensed-matter version of quantum electrodynamics has been the topic of many recent experimental and theoretical studies [65, 66]. In addition, because of the lack of nuclear spins in predominantly ^{12}C (weak hyperfine interaction) and a spin-orbit interaction suppressed by graphene's symmetries, spin-qubits based on carbon have been proposed [44, 45]. However, experimental data on spin-orbit coupling in graphene with broken symmetries (such as electric fields or curvature) is missing. An observation of spin-orbit induced shell splitting in nanotubes is presented in Section 5.5.

In our carbon nanotube devices (see Figure 1.1 B,E) the charge can be positive or negative ("electrons" or "holes") and is tunable between zero to many tens. The Fermi wavelength is comparable to the size of the quantum dot, and size quantization becomes apparent even if the nanotube is micrometers long. In contrast to our metal nanoparticle SETs, we find the gate capacitance ($\sim$ 3 aF) is larger than the tunnel capacitances, and hence the number of electrons can easily be changed by more than 10. The electronic states of carbon nanotubes [15, 16] follow from the electronic properties of graphene, which are discussed in the next section.

1.5 Bandstructure of graphene

The bandstructure of flat graphene is well described by a tight-binding model in which p_z-orbitals couple only to their nearest neighbors. Due to their symmetry they do not couple to the σ-bands, and hence it is sufficient to work with a four-fold basis ($A\uparrow$, $A\downarrow$, $B\uparrow$, $B\downarrow$) corresponding to spin-up/down p_z-orbitals located on the "A atom" and "B atom" of the unit cell (see Fig. 1.3A). Because of the intercalated structure of two inequivalent triangular lattices in real space, one

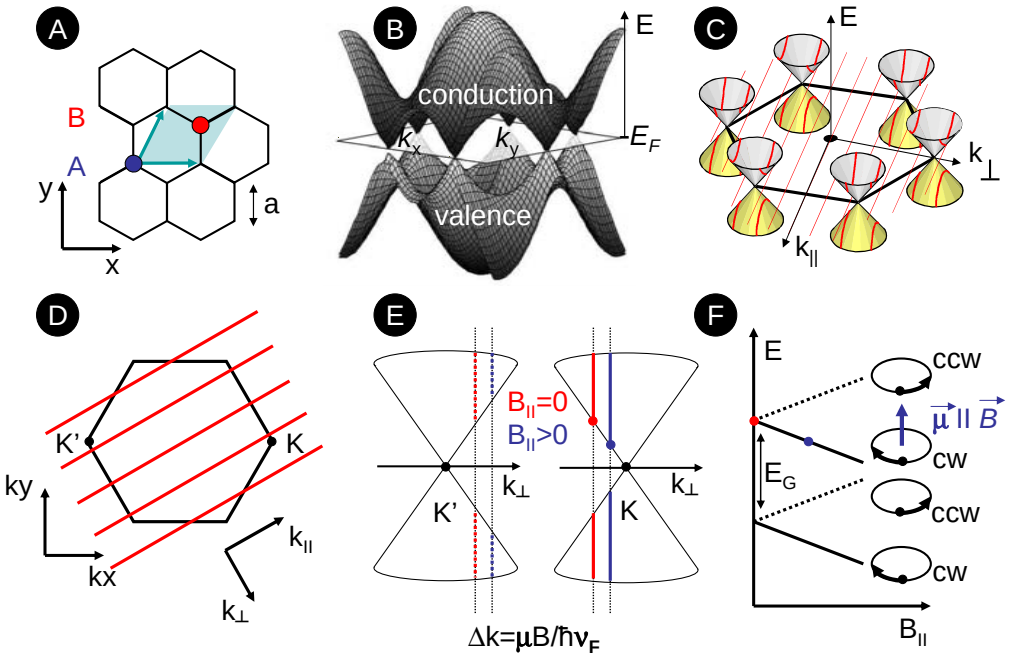


Figure 1.3: A) The unit cell of graphene contains two carbon atoms (called A and B) separated by $a \approx 1.42$ Å. B) Tight-binding band structure if only p_z-orbitals are considered (justified only if spin-orbit coupling is zero). The conduction and valence band meet at the K and K' points, where they can be approximated by a linear dispersion (panel C). In a carbon nanotube (Fig. 1.1C) only a discrete set of states (red lines) are allowed due to the boundary conditions imposed by wrapping the graphene sheet. D) In a small-bandgap nanotube, the quantization lines almost pass through a K point. The offset determines the bandgap E_G at $B = 0$ T. E) If a magnetic field is applied parallel to the nanotube, an Aharonov-Bohm phase modifies the quantization conditions, as indicated by the shift Δk. This decreases the bandgap associated with K electrons, and increases the bandgap for K' electrons. F) Ignoring spin contributions, each electronic state shifts in energy according to its orbital magnetic moment. For example, the level marked with a blue dot corresponds to a clockwise moving electron, i.e. its orbital magnetic moment is aligned parallel to the magnetic field.

obtains in reciprocal space two independent $k-$points, called K and K', where the conduction band and the valence band meet (panel B). In the neighborhood of the K and K' points the states can be described by two sets of two-dimensional chiral spinors, the spin and the pseudospin ($\boldsymbol{k} \cdot \boldsymbol{p}$ approximation).

Pseudospin and isospin

The pseudospin can be defined in such away that it points along the direction of propagation for "electrons" near the K point, and antiparallel near the K' point [28]. With "electrons" we mean states above the Fermi level. For states below the Fermi level ("holes") the assignment is reversed. Physically the pseudospin contains information about the character of the underlying molecular wavefunction within the unit cell [28] and should not be confused with the "isospin". The latter is sometimes used to distinguish between states near the K-points and states near the K'-point [16]. We can use the spin ("up" and "down") and isospin ("clockwise" and "counterclockwise") to label the four-fold degenerate states expected for carbon nanotubes. Neglecting spin one obtains the Hamiltonian

$$H = \hbar \nu_F \boldsymbol{\sigma} \boldsymbol{k} = -i\hbar \nu_F [\sigma_x \partial_x + \tau_z \sigma_y \partial_y] \tag{1.9}$$

where $\nu_F \approx 8 \times 10^5$ m/s is the Fermi velocity, $\boldsymbol{k}$ is the wavevector measured relative to the K (K') point, and $\boldsymbol{\sigma}$ are the Pauli matrices acting on the pseudospin. This is the Dirac-Weyl equation for graphene. $\tau_z = \pm 1$ denotes the K (K') point.

Bandstructure of carbon nanotubes

In carbon nanotubes (see Fig. 1.1C) not all states in the Brillouin zone of graphene are allowed, but only those which satisfy the quantization condition

along the circumference of the nanotube (we call this direction $k_\perp$)[5] as well as the longitudinal confinement. The latter leads to a level spacing of

$$\delta = \hbar\nu_F \frac{\pi}{\mathrm{L}} \approx \frac{1.67\ meV \cdot \mu m}{\mathrm{L}} \tag{1.10}$$

for metallic nanotubes of length L. For semiconducting nanotubes the Dirac particles acquire a non-zero effective mass and the level spacing is smaller [23]. In the absence of symmetry breaking each quantized state is four-fold degenerate (spin up/down, orbit clockwise/counterclockwise).

The states which are allowed by the transverse quantization condition (see Eq. 1.11 below) are indicated by red lines in panel C and D. They are known as the bands of a (infinitely long) nanotube. If a quantization line nearly passes through a K point, one obtains a nanotube with a small bandgap E_G. This is schematically indicated in panel E. Ignoring spin degeneracy, we note that each state near the K point (solid red line) has a time-reversed partner near the K' point (dashed red line) which is degenerate for $B = 0$ T. Application of a magnetic field parallel to the nanotube breaks this degeneracy through the Aharonov-Bohm phase [13] which has to be accounted for in the quantization condition along $k \perp$ (blue lines):

$$\pi D k_\perp + 2\pi \frac{\Phi}{\Phi_0} = 2\pi j \tag{1.11}$$

where j is an integer, D is the diameter of the nanotube, $\Phi_0 = h/e$ is the flux quantum and $\Phi = B_{||}\pi D^2/4$ is the magnetic flux inside the nanotube [22].

From $\Delta E = \hbar\nu_F \Delta k$ (Eq. 1.9) one obtains the shift in energy associated with the magnetic field

$$\Delta E = \frac{e\nu_F D}{4} B_{||}. \tag{1.12}$$

[5]The electronic structure of carbon nanotubes is more rigorously discussed for example in [15, 16].

The quantity $\mu_{orb} = e\nu_F D/4 \sim D[\text{nm}]\times 0.2$ meV/T is the orbital magnetic moment associated with an electron encircling a nanotube of diameter D at velocity ν_F [22]. Our convention is to call the states which lower their energy with increasing magnetic field "clockwise" and the states which increase in energy "counterclockwise" (panel F). Note that the orbital magnetic moment ($\mu_{orb} \approx\ D$ [nm] $\times$ 0.2 meV/T) in nanotubes is much larger than the Bohr magneton ($\mu_B \approx 0.058$ meV/T).

1.6 Spin-orbit interaction in graphene

Spin-orbit interaction has been discussed for graphite [10, 11], for diamond [9], for graphene [2, 2, 6–8] and for nanotubes [2–5]. Huertas-Hernando *et al.* [2] use a tight-binding model and obtain numerical estimates (see below), without assuming a specific chirality of the nanotube. Chico *et al.* [4] consider chiral nanotubes and find a spin splitting at the Fermi level in the absence of a magnetic field. For non-chiral tubes spin-degeneracy is preserved as expected from time-reversal and inversion symmetry. Table 1.2 summarizes theoretical results from different groups.

Spin-orbit interaction for the free carbon atom

The spin-orbit splitting of the free carbon atom is an important parameter for calculating spin-orbit effects in graphene and carbon nanotubes. We can evaluate Eq. 1.3 for hydrogenic $2p$-orbitals ($n = 2$, $\ell = 1$) to obtain the intra-atomic spin-orbit coupling Δ in carbon. Substituting

$$V(\boldsymbol{r}) = \frac{Ze^2}{4\pi\epsilon_0 r} \tag{1.13}$$

into Equation 1.3 and using [38]

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{\ell(\ell+1/2)(\ell+1)n^3 a_B^3} \tag{1.14}$$

where $a_B = \hbar/mc\alpha$ is the Bohr radius and $\alpha = e^2/\hbar c 4\pi\epsilon_0 \approx 1/137$ is the fine structure constant, we obtain

$$\Delta = Z^4 mc^2 \frac{\alpha^4}{4n^3} \frac{[j(j+1) - l(l+1) - s(s+1)]}{\ell(\ell+1/2)(\ell+1)} \tag{1.15}$$

where $j = 3/2, 1/2$ are the eigenvalues of the total angular momentum operator $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{s}$. Substituting $n = 2$, $\ell = 1$ and $Z \approx 6 - 2$ (effective nuclear charge seen by the p-electrons) we obtain a splitting of $\Delta \sim 12$ meV between the $J = 3/2$ and $J = 1/2$ eigenstates of carbon. This is consistent with more rigorous calculations [14, 30]. It is much smaller than the $5d$-splitting calculated for the gold atom (~ 1.5 eV [14]) due to low atomic number of carbon (note the strong dependence of the atomic spin-orbit coupling on the atomic number in Eq. 1.15).

Intrinsic spin-orbit interaction in graphene

The role of the electronic spin has been neglected in all panels of Figure 1.3. To include the effects of spin-orbit coupling on the electronic structure of graphene it is not sufficient to consider p_z-orbitals only, because the intra-atomic spin-orbit Hamiltonian

$$H_{atom} = \Delta \left[\frac{L_+ s_- + L_- s_+}{2} + L_z s_z \right] = \frac{\Delta}{2} \left(\begin{array}{ccc|ccc} 0 & 1 & i & & & \\ 1 & 0 & -i & & 0 & \\ -i & i & 0 & & & \\ \hline & & & 0 & -1 & i \\ & 0 & & -1 & 0 & -i \\ & & & -i & i & 0 \end{array} \right) \tag{1.16}$$

couples $p_z\uparrow$, $p_x\downarrow$ and $p_y\downarrow$ on the same atom (as well as $p_z\downarrow$, $p_x\uparrow$ and $p_y\uparrow$). In the right hand side of Eq. 1.16 the basis states are $|p_z\uparrow\rangle$, $|p_x\downarrow\rangle$ $|p_y\downarrow\rangle$, $|p_z\downarrow\rangle$, $|p_x\uparrow\rangle$ and $|p_y\uparrow\rangle$.

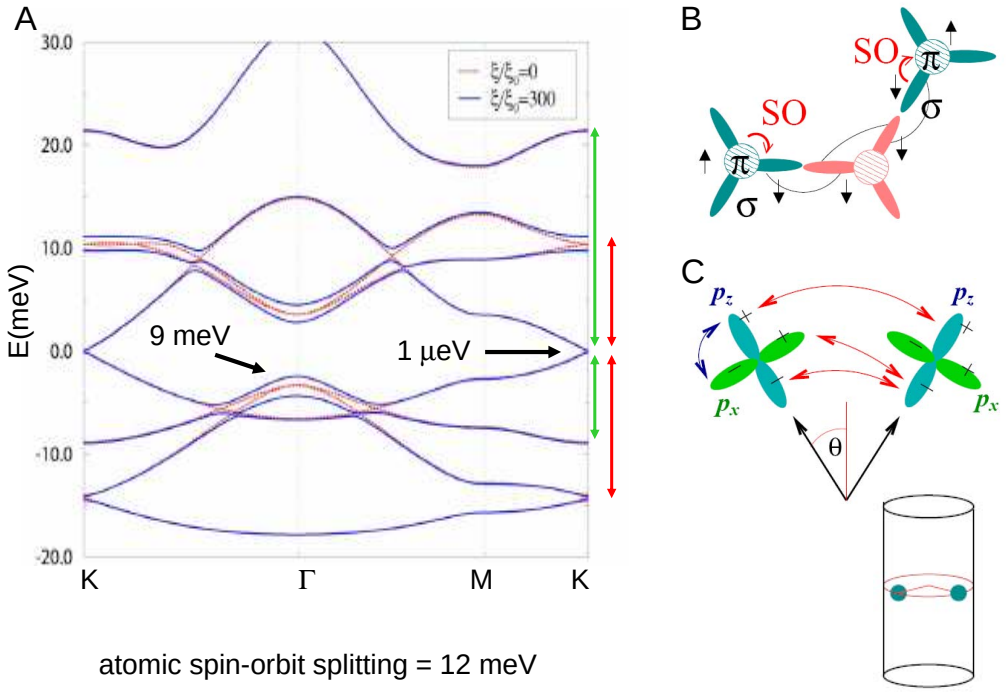


Figure 1.4: A) Tight-binding bandstructure of flat graphene without spin-orbit coupling (red) and with 300-fold enhanced spin-orbit interaction (blue). Note that several splittings occur comparable to the atomic spin-orbit coupling ($\Delta = 12$ meV). The intrinsic splitting at the K points however is greatly reduced ($\sim$1 μeV). From [8]. This suppression is due to the mirror symmetry about the graphene sheet, and second-nearest neighbor processes as shown in B need to be considered. B) Intrinsic spin-orbit coupling between two A atoms (cyan): A spin-up electron hops from the p_z-orbital onto a spin-down p_x-orbital (spin-orbit), then hops to the second nearest neighbor via the σ-band) and back onto a spin-up state via spin-orbit coupling. This process is second order in Δ and corresponds to the transitions marked with green arrows in panel A. The red arrows in A correspond to the same process for a spin-down electron, hence mixing with *all* bands occurs. C) Additional hoppings allowed by curvature are important in carbon nanotubes. Here, a spin-up electron on the left p_z-orbital hops onto the left p_x-orbital (now spin-down) and directly to the right p_z-orbital (spin-down). This process is first order in Δ. From [2].

The term $L_+s_- + L_-s_+$ of the intra-atomic spin-orbit Hamiltonian allows transitions between the states of the π-band and states from the σ-band. Using a tight-binding model and second-order perturbation theory Min *et al.* [8] calculated the bandstructure shown in Figure 1.4A. Note that due to spin-orbit coupling gaps of the same order as the intra-atomic coupling Δ opens at various points [7] (for better visibility a 300-fold enhanced Δ has been used to generate the blue curves). However, the intrinsic splitting near the K points is very small ($\sim$1 μeV). This is due to the mirror symmetry of the graphene sheet in respect to the x-y-plane [2], which leads to a cancellation of nearest-neighbor hopping terms [2] for the K and K' points. Hence the intrinsic spin-orbit coupling Δ_{int} near the K points is second order in Δ, and Min *et al.* find:

$$\Delta_{int} = \frac{|s|}{18(sp\sigma)^2}\Delta^2 \sim 0.5\ \mu\text{eV} \tag{1.17}$$

where $|s| = 8.868$ eV is the energy splitting between atomic s- and p-orbitals and $sp\sigma = 5.58$ eV is a band parameter in their model. They assume $\Delta = 6$ meV for the atomic spin orbit coupling and find an intrinsic spin-orbit gap of order 1 μeV at the K points. The second-order processes responsible for the gap are illustrated in Fig. 1.4B: An $p_z \uparrow$ electron on the left A atom (cyan) can hop onto a $p_x \downarrow$-orbit through the intra-atomic spin-orbit coupling. Then it can hop via the σ-band to the next B atom and further to the sp_2-orbits of the next A atom. Another spin-flip process induced by the intra-atomic coupling brings it back to an $p_z \uparrow$ state.

Spin-orbit coupling due to curvature

A more efficient way to couple the π-band to the σ-band is by application of a perpendicular electric field (Rashba) or by curvature. Figure 1.4C illustrates how the coupling of a tilted p_z-orbital can couple directly to a neighboring p_x-

orbital (this process is important for spin-orbit coupling) as well as a neighboring p_z-orbital (this process leads to a curvature induced bandgap even if there is no spin-orbit coupling [16]). Huertas-Hernando *et al.* [2] find for the curvature induced spin-orbit coupling Δ_{curv}

$$\Delta_{curv} = \frac{V_{pp\sigma} - V_{pp\pi}}{V_1}(\frac{a}{R1} + \frac{a}{R2})\frac{V_1^2}{V_2^2}\Delta \sim 1.60 \text{ meV } /D \text{ [nm]} \tag{1.18}$$

where $a = 1.42$ Å is the carbon-carbon distance, R_i the radii of curvature, and $V_1 = 2.47$ eV, $V_2 = 6.33eV$, $V_{pp\sigma} = 5.38$ eV and $V_{pp\pi} = -2.24$ eV are band parameters in their tight-binding model. They use $\Delta = 12$ meV for the intra-atomic spin-orbit coupling.

Effective Hamiltonian for the π-band

Taking into account all s and p-orbitals using a tight-binding model and second-order perturbation theory, Huertas-Hernando *et al.* find the effective Hamiltonian for the π-bands of graphene near the K points, including the intrinsic, Rashba and curvature induced spin-orbit coupling [2]:

$$H_{graphene} = -i\hbar\nu_F[\sigma_x\partial_x + \tau_z\sigma_y\partial_y] + \Delta_{int}[\tau_z\sigma_y s_z] + \frac{\Delta_R}{2}[\sigma_x s_y + \tau_z\sigma_y s_x] \tag{1.19}$$

where $\tau = \pm 1$ denotes the K (K') Dirac point, $\boldsymbol{\sigma}$ are the Pauli matrices acting on the pseudospin and $\boldsymbol{s}$ are the Pauli matrices acting on the electronic spin. $\Delta_R = \Delta_{curv} + \Delta_E$ is the coupling due to curvature and Rashba (the expression for Δ_E is provided in table 1.2).

The first term is the unperturbed Dirac term (Eq. 1.9) and the second term is the intrinsic spin-orbit Hamiltonian which opens a small gap of size $2\Delta_{int}$. Note that it splits the $A\uparrow$,$B\downarrow$ states in energy from the $A\downarrow$,$B\uparrow$ states. This arises from the process depicted in Fig. 1.4B, which couples equal spins on second-nearest

neighbors. The third term can be written with raising and lowering operators as $-i\Delta_R(\sigma_+ s_+ - \sigma_- s_-)$ for the K point and $-i\Delta_R(-\sigma_+ s_- + \sigma_- s_+)$ for the K' point. Physically this describes a coupling between states which differ in spin and pseudospin, as visualized in Fig. 1.4C: starting with a $p_z \uparrow$ electron on the right atom ("$A \uparrow$"), curvature allows hopping to the $p_x \uparrow$-orbital on the left B atom, followed by a spin-flip process to the $p_z \downarrow$-orbital via intra-atomic spin-orbit coupling ("$B \downarrow$").

1.7 Spin-orbit coupling in nanotubes

Huertas-Hernando *et al.* extend their findings (Eq. 1.19) to carbon nanotubes [2]. Their results agree with those of DeMartino [5] and (after integration around the circumference of the nanotube) those of Ando [3]. Huertas-Hernando *et al.* do not assume specific chirality, and describe the eigenstates of the (unperturbed) nanotube by a longitudinal momentum $k_{||}$ and angular momentum $k_\perp$:

$$\epsilon = \pm\hbar\nu_F\sqrt{k_{||}^2 + k_\perp^2} \tag{1.20}$$

The effective Hamiltonian including spin-orbit interaction becomes

$$H_{CNT}\begin{pmatrix} A\tau \\ B\tau \end{pmatrix} = \begin{pmatrix} 0 & \hbar\nu_F(k_{||} - ik_\perp) + \tau i\Delta_R\pi s_z \\ \hbar\nu_F(k_{||} + ik_\perp) - \tau i\Delta_R\pi s_z & 0 \end{pmatrix}\begin{pmatrix} A\tau \\ B\tau \end{pmatrix} \tag{1.21}$$

where $\tau = \pm 1$ corresponds to the K (K') Dirac point and the basis states $A\tau$ and $B\tau$ are spinors in spin subspace where the matrix s_z acts on (refer to [2] for details). The z-direction is now defined parallel to the nanotube axis (and not perpendicular to the graphene sheet as usually done when describing graphene).

The contribution of the intrinsic spin-orbit coupling becomes zero after inte-

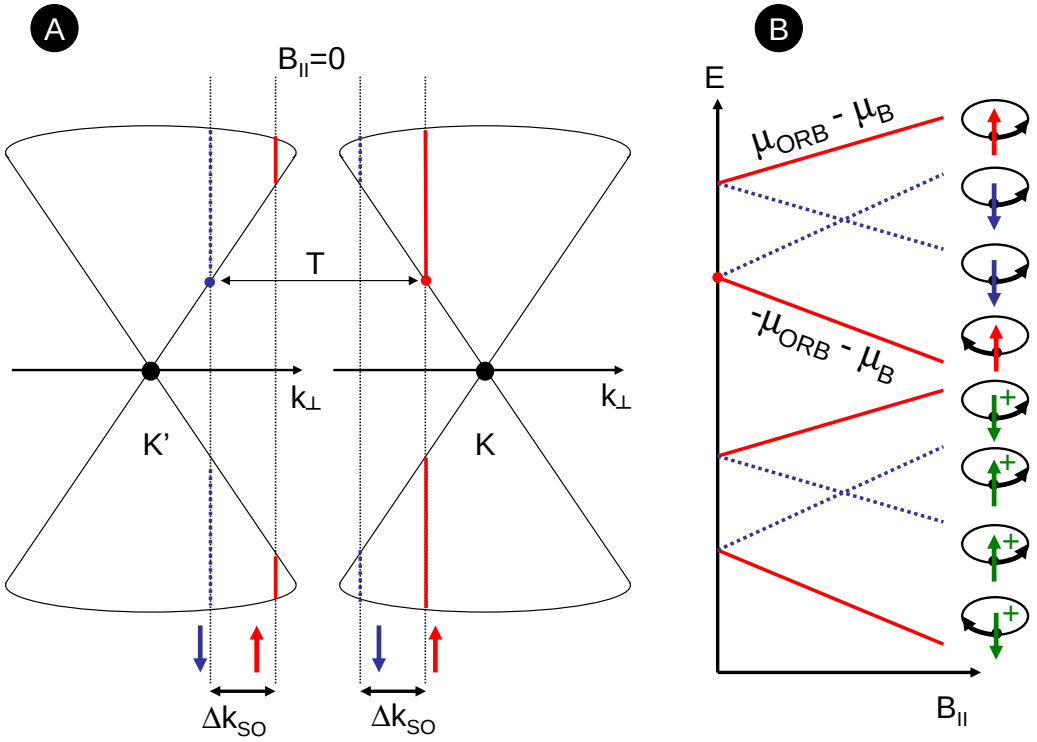


Figure 1.5: A) At zero magnetic field, spin-orbit interaction due to curvature modifies the quantization condition for spin-up electrons (red) and spin-down electrons (blue) by Δk_{SO}. Note that time reversal symmetry (T) is maintained, as exemplified by the red and blue dots. C) When applying a magnetic field parallel to the nanotube, each Kramer's doublet splits according to its spin and orbital magnetic moment. Top 4 levels: Each electronic state is schematically labeled by its spin magnetic momentum and cw or ccw orbital trajectory. The bottom 4 levels are identical, unless one prefers to think in terms of holes. Then the charge and spin need to be flipped, but not the orbital motion. This is shown for the bottom 4 levels.

grating over the circumference of the nanotube, and they obtain

$$\epsilon = \pm\hbar\nu_F\sqrt{k_{||}^2 + (k_\perp + s_z\Delta k_{SO})^2} \tag{1.22}$$

where $\Delta k_{SO} = \pi\Delta_R/\hbar\nu_F$ opens a gap $\pi\Delta_R$ due to spin-orbit interaction, in agreement with the results in [3,4]. This can be viewed as the consequence of the Berry phase gained by the electron and hole quasiparticles after completing a closed trajectory around the circumference of the nanotube [2]. Because this phase is different for spin-up and spin-down particles, the perpendicular quantization condition for a small-bandgap nanotube gets modified as schematically indicated in Figure 1.5A. At zero magnetic field, one obtains two-different quantization conditions at each Dirac point, separated by $\Delta k_{SO} = \pi\Delta_R/\hbar\nu_F$. Note that spin-orbit interaction preserves time-reversal symmetry, as schematically indicated by the Kramer's doublet marked with a red and blue dot. The magnetic field dependence in panel B follows by taking into account the Aharonov-Bohm shift (from Fig. 1.3E) as well as the Zeeman shift due to the spin (which is usually an order of magnitude smaller).

Our convention of spin and orbital assignments

What we measure experimentally is magnetic moments and not angular momenta, and hence when we assign spins we actually mean spin *magnetic* moments and not spin angular momenta[6]. For example, when we specify a "clockwise spin-up" electron in an external magnetic field pointing up, this corresponds to orbital magnetic moment pointing up as well as a spin magnetic moment pointing up (i.e.

[6]The spin angular momentum of an electron points opposite to its magnetic moment due to its negative g-factor. Semiclassically, this is due to its spinning negative charge. Note that the g-factor for an hole is negative as well, due to its negative spinning mass.

spin *angular* momentum pointing down).

In Figure 1.5 the total magnetic moment associated with the clockwise spin-up electron (red dot) is $\mu_{orb} + \frac{g}{2}\mu_B$ where $\mu_B = e\hbar/2m$ is the Bohr magneton, whereas that of its time-reversed partner (ccw spin-down) is exactly opposite. A different total magnetic moment is obtained for the cw spin-down electron ($\mu_{orb} - \frac{g}{2}\mu_B$) and its time reversed partner (ccw spin-up), see Fig. 1.5B. The assignment of spins and orbits is identical for both shells depicted in panel B. However, in the case of one missing electron in an otherwise fully occupied shell it may be more convenient to speak in terms of holes than electrons. Because a full shell neither carries spin angular momentum nor current, one needs to flip the spin and the charge, but not the cw or ccw motion when going from the electron assignment to the hole assignment. The hole assignment is given for the 4 bottom levels in panel B. Note that with this assignment (and the magnetic field pointing up), the ground state of the first electron is characterized by an orbital magnetic moment parallel to the magnetic field, and its spin magnetic momentum parallel as well. This is in contrast to the first hole, which has its orbital magnetic moment aligned parallel to the field, and its spin magnetic momentum aligned antiparallel.

Magnitude of the spin-orbit coupling

The calculated atomic spin-orbit splitting between the $J = 3/2$ and $J = 1/2$ states in carbon is $\sim$ 8.9 meV (using density functional theory [14]) or $\sim$ 12 meV (using relativistic Hartree-Fock calculations [30]). Dresselhaus *et al.* [11] list 8 meV for atomic carbon, 6 meV for the crystalline band splitting in diamond and 0.22 meV for graphite. Serrano *et al.* discuss spin-orbit splitting in diamond (13 meV [9]) and quote 11 meV for the free carbon atom. Numerical results for

Table 1.2: Theoretical results from various groups for spin-orbit interaction in graphene. Δ is the atomic spin-orbit coupling. [2] uses $\Delta = 12$ meV and band parameters $V_1 = 2.47$ eV, $V_2 = 6.33$ eV, $V_{pp\sigma} = 5.38$ eV and $V_{pp\pi} = -2.24$ eV. The distance between carbon atoms is $a = 1.42$ Å. [8] uses $\Delta = 6$ meV and $sp\sigma = 5.58$ eV. The estimates by Kane and Mele [2] are very crude.

intrinsic coupling Δ_{int}	$\frac{3}{4V_1}\frac{V_1^4}{V_2^4}\Delta^2$ $\sim$1 μeV	[2]
Rashba coupling Δ_E	$\frac{2\sqrt{2}}{3}\frac{\lambda eE}{V_2}\Delta$ $\sim$8 μeV	$z_0 \approx 0.264$ Å, $E = 50V/300$nm [2]
curvature coupling Δ_{curv}	$\frac{V_{pp\sigma}-V_{pp\pi}}{V_1}(\frac{a}{R1}+\frac{a}{R2})\frac{V_1^2}{V_2^2}\Delta \sim 8\mu$eV	$\sim$1.60 meV / D[nm] [2]
intrinsic coupling	$\frac{\lvert s\rvert}{18(sp\sigma)^2}\Delta^2$ $\sim$0.5 μeV	[8]
Rashba coupling	$\frac{z_0 eE}{3(sp\sigma)}\Delta$ $\sim$11 μeV	$z_0 \approx 1.86$ Å, $E = 50V/300$nm [8]
intrinsic coupling	$\frac{2\pi^2 e^2\hbar^2}{3m^2c^2a_B^3}$ $\sim$0.2 meV	[2]
Rashba coupling	$\frac{\hbar\nu_F eE}{4mc^2} \sim 0.04\mu$eV	$E = 50V/300$nm [2]
intrinsic coupling	0.05 meV	first principle calc. [6]
intrinsic coupling	0.8 μeV	[7]
carbon nanotube gap	$\sim$ 10 meV $\cdot$ nm/D	D=nanotube diameter [2]

graphene are based on assuming an intra-atomic splitting of 12 meV in the work of Huertas-Hernando [2] and 6 meV in the work of Min *et al.* [8]. Table 1.2 summarizes the numerical estimates reported from various groups for the intrinsic, Rashba and curvature induced spin-orbit coupling in graphene. For nanotubes, a gap of size $2\pi\Delta_{curv} \sim$10 meV$/D$ [nm] is obtained by Huertas-Hernando due to the curvature-induced spin-orbit coupling $\Delta_{curv} \sim 1.60$ meV/D[nm]. This corresponds to a splitting of 2 meV for a nanotube with diameter $D = 5$ nm, and is larger then the splitting we observe in Chapter 5 (0.4 meV for a nanotube with diameter ~ 5 nm).

Remark on geometric interpretation

The bandstructure of carbon nanotubes can be described by quantum phases which are due to the boundary conditions ("winding phase" [16]), due to curvature (acting like a vector potential on the charge of the Dirac particles [16]), or due to applied magnetic fields (Aharonov-Bohm phase [1]). It is tempting to consider the spin-orbit splitting as an Aharonov-Casher effect [12], i.e. due to the phase picked up by a spin-1/2 particle when encircling a rod of charge. Assuming that an electron with mass $m = 10^{-30}$ kg and velocity $\nu_F = 8 \times 10^5$ m/s is held on its orbit (diameter $D \sim 5$ nm) by a radial electric field

$$E = \frac{2m\nu_F^2}{eD} \sim 1.5 \text{ V/m} \tag{1.23}$$

then the line charge associated with this electric field is

$$\lambda = \pi D\epsilon_0 E = \pi\epsilon_0 m\nu_F^2 \approx 7^{-6}\lambda_0 \tag{1.24}$$

where $\lambda_0 = emc/\hbar g\alpha \approx 2 \cdot 10^5 e$/nm is the "line charge quantum" [12] similar to the flux quantum Φ_0 in the Aharonov-Bohm effect [13]. The phase associated with

the line charge λ translates into an energy shift of $\Delta E = \frac{2\hbar\nu_F\lambda}{D\lambda_0} \approx 1.5$ μeV (using Eq. 1.11 and 1.9 and $D = 5$ nm) which is much smaller than spin-orbit splitting due to curvature (Table. 1.2). However, the Rashba splitting found for an electric field of 50V/300nm by Huertas-Hernando *et al.* (8 μeV, see Table. 1.2) is two orders of magnitude larger then Kane and Mele's estimate for the same electric field and a "free" electron (0.04 μeV, see Table. 1.2). This suggests that a geometric interpretation of spin-orbit coupling in carbon nanotubes may be possible if the Aharonov-Casher phase is properly converted to a condensed matter version.

The interpretation of spin-orbit interaction in disordered conductors as a Aharonov-Casher effect was discussed in [19]. The spin-orbit coupling can be formulated as an effective vector potential, or alternatively, as a phase factor. Spin-orbit scattering and the Aharonov-Casher effect in mesoscopic systems is further discussed in [12,21]. For spin-orbit Berry phases in conducting rings (mostly of type Rashba) see [19, 22–27]. The connection between Thomas precession, spin-orbit interaction and Berry's phase is discussed in [47].

1.8 Probing quantum states with a single-electron transistor

Discrete states in quantum dots can be measured by connecting them to leads (called source and drain). Measuring the linear conductance gives information about the ground state of the quantum dot, whereas its excited states can be extracted from non-linear conductance measurements (at finite bias voltage). Furthermore, the number of electrons or holes on the quantum dot can be controlled by changing a gate voltage.

Spectroscopy of metal nanoparticles using single-electron transistors (SETs) has been discussed in detail by von Delft and Ralph [57]. The line shape and position of tunnel resonances reveal valuable information about electron-electron interactions, superconductivity, ferromagnetism and the lifetime of excitations. von Delft and Ralph also discuss the requirements needed in order to observe discrete states. Briefly, these are a temperature smaller than the level spacing in order to avoid thermal smearing, tunneling rates between the discrete states and the leads to be small enough to avoid tunneling-induced level broadening, and inelastic relaxation rates within the nanoparticle to be small enough such that the corresponding line width is less than the level spacing.

A good introduction to spectroscopy on carbon nanotube quantum dots and fullerenes can be found in the Ph.D. theses of Park [33] and Pasupathy [34]. A review on single-electron transistors was written by Kastner [35] (also see his non-technical version in Physics Today [36]). For the understanding of the double-dot characteristics observed in Chapter 5, the review on electron transport through double quantum dots by van der Wiel *et al.* is very helpful [37].

Coulomb-blockade and single-electron tunneling

Consider a small metal island connected via two weak tunnel barriers to a source and a drain electrode (see schematic in Figure 1.6). The phenomenon of sequential single-electron tunneling becomes relevant if the charging energy $E_C = e^2/2C$ associated with one excess electron on the island is much larger then the thermal energy of electrons in source and drain (here $C = C_S + C_D + C_G$ is the total capacitance). In fact, at low bias voltages no current flows at all due to the electrostatic energy cost of adding an electron to the island ("Coulomb blockade"),

unless the potential of the island is tuned to a degeneracy point by means of the gate voltage V_G. Alternatively, the bias voltage V can be increased until the Coulomb blockade is overcome and current I starts to flow. Typical I-V curves measured for a metal nanoparticle SET at 4.2 K is shown at the bottom of Fig. 1.6. Note the appearance of three partially visible "Coulomb blockade diamonds". Within each diamond the charge on the island is constant in time (i.e. no current flow) and differs from a neighboring diamond by exactly one electron. Hence the gate capacitance C_G can be extracted from the difference in gate voltage between two degeneracy points: $C_G = e/\Delta V_G$. The maximum bias voltage V_{max} at which current is blocked occurs when the gate voltage is tuned in between two degeneracy points, and is a measure for the charging energy $E_C = eV_{max}/2$.

The SET as a spectrometer

The single electron transistor turns into a spectrometer of the island if the electronic temperature is made smaller than the island's level spacing. Then, additional fine steps in the I-V curve near the degeneracy point can be observed (see Figure 1.7). They are due to the fact that the voltage source has to supply not only the electrostatic energy of order E_C, but also the energy of the discrete quantum levels to be occupied (see schematic in Figure 1.7).

Inspection of I-V curves measured near a degeneracy point reveals a discrete set of steps which conveniently show up as conductance peaks when plotting the derivative (bottom graph in Fig. 1.7). The first peak marked with an red arrow corresponds to a bias voltage which is just enough to allow electrons from the drain to tunnel onto the island and further to the source electron (one at a time due to the large charging energy associated if two electrons tunnel simultaneously).

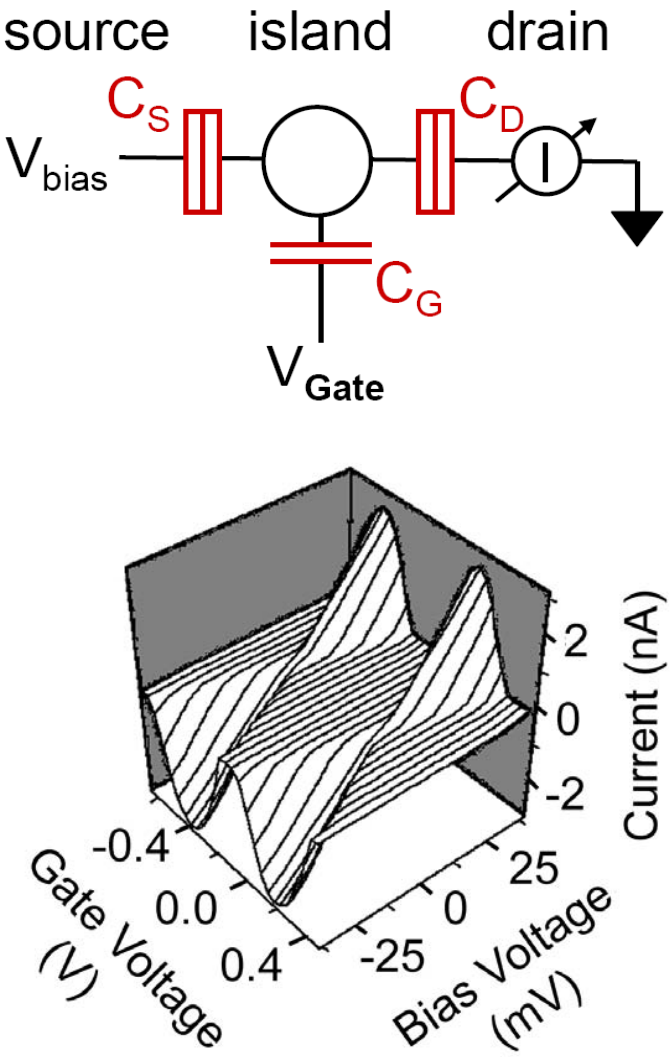


Figure 1.6: Top: Circuit schematic of a single electron transistor. Bottom: I-V curves for different gate voltages in the regime $\delta << k_bT << E_C$. The regions in which no current flows due to Coulomb blockade are known as Coulomb diamonds. From their geometry the capacitances C_S, C_D, C_G as well as the charging energy E_C can be extracted.

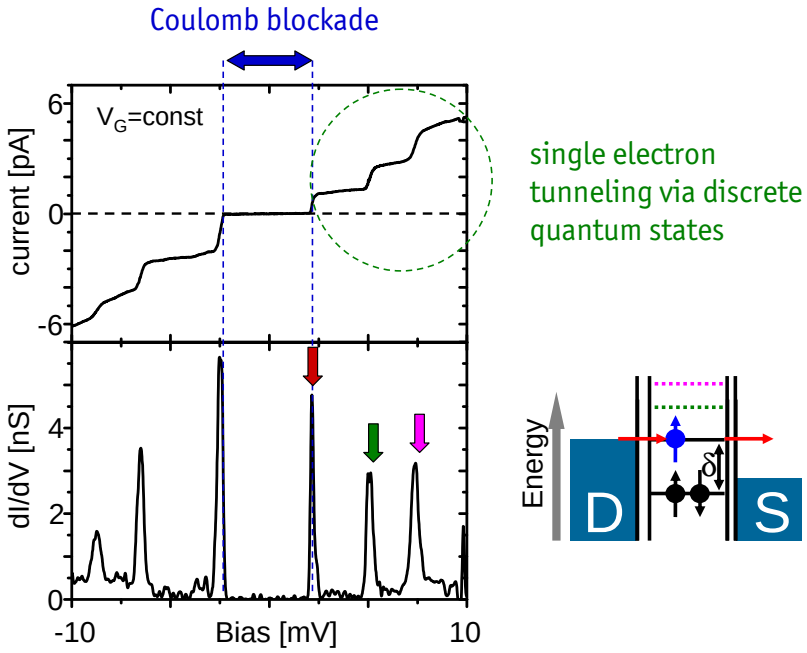


Figure 1.7: When the Coulomb blockade is tuned near a degeneracy point using the gate electrode, and $k_bT << \delta << E_C$, single-electron tunneling occurs through discrete states. The total current then depends on how many channels are accessed by the bias window and hence increases in discrete steps. Each step conveniently shows up as a conductance peak when plotting the differential conductance.

During this process, the nanoparticle makes a transition from its n electron ground state to its $n+1$ electron ground-state and back. This resonance therefore marks the ground-state to ground-state transition between the n electron quantum dot and the $n+1$ electron quantum dot. The next higher conductance peak (marked with a green arrow) occurs when the bias voltage is increased by a small increment ΔV, such that an electron may alternatively tunnel via the first excited state of the $n+1$ electron dot. This additional path of current flow is responsible for the steplike increase in current at this threshold. To obtain the level spacing δ one needs to convert ΔV to energy by taking into account the capacitance ratios C_S/C_G and C_D/C_G:

$$\delta = e \cdot \frac{C_S}{C_S + C_D + C_G} \cdot \Delta V \tag{1.25}$$

if the tunneling threshold is due to the drain electrode and

$$\delta = -e \cdot \frac{C_D + C_G}{C_S + C_D + C_G} \cdot \Delta V \tag{1.26}$$

if the tunneling threshold is due to the source electrode.

In Fig. 1.7 it was (correctly) assumed that the rate limiting barrier is the interface between the island and the drain electrode, and hence the resonances at positive bias correspond to thresholds of electrons tunneling onto the particle[7]. The determination of the capacitance ratios (as well as to determine which of the two equations to use) is most easily accomplished by measuring I-V curves as a function of V_G (Fig. 1.8). Then, each conductance feature with a positive slope $S_+ = dV_G/dV > 0$ corresponds to a threshold on the source, whereas negative slopes ($S_- = dV_G/dV < 0$) are associated with threshold conditions on the drain.

[7]If the source (drain) electrode reaches a threshold (i.e. lines up with an accessible quantum state) the current will only show a steplike increase if the rate limiting barrier is on the side of the source (drain) electrode.

We use the slopes S_+ and S_- to experimentally extract the capacitance ratios

$$\frac{C_D}{C_G} = S_+ - 1 \tag{1.27}$$

$$\frac{C_S}{C_G} = |S_-| \tag{1.28}$$

If the differential conductance is plotted as a function of bias and gate voltage, unoccupied states as well as occupied states can be extracted. This is depicted schematically in Fig. 1.8. Starting at a gate voltage below the degeneracy point (i.e. fixed number of N electrons at zero bias), the bias voltage is slowly increased until at point 1 an electron from the drain can tunnel onto the lowest unoccupied level and further to the source electrode, giving rise to a conductance peak. Increasing the bias further will eventually allow an electron to alternatively hop via the next higher unoccupied state. This increases the likelihood of current flow, and hence another conductance peak occurs at point 2. The bias difference between point 1 and point 2 therefore maps out the energy difference between the lowest two unoccupied states. A similar argument can be given for occupied states, by starting with $N + 1$ electrons and increasing the bias until at point 3 the ground-state to ground-state transition ($N + 1 \rightarrow N$) occurs. At point 4, instead of tunneling out an electron from the highest occupied state, the next lower electron may tunnel out, increasing the total current and hence giving rise to conductance peak in pink.

We will find in Chapter 2 that electron-electron interactions are small in gold nanoparticles, and hence we may indeed regard each resonance as a single particle level, with an energy in respect to the other levels given by simply recalibrating the bias axis into an energy axis. Depending on whether the rate limiting barrier is on the source or drain side, this calibration factor may be negative (see minus sign in Eq. 1.26). For instructive reasons, this situation is shown in the lower part of

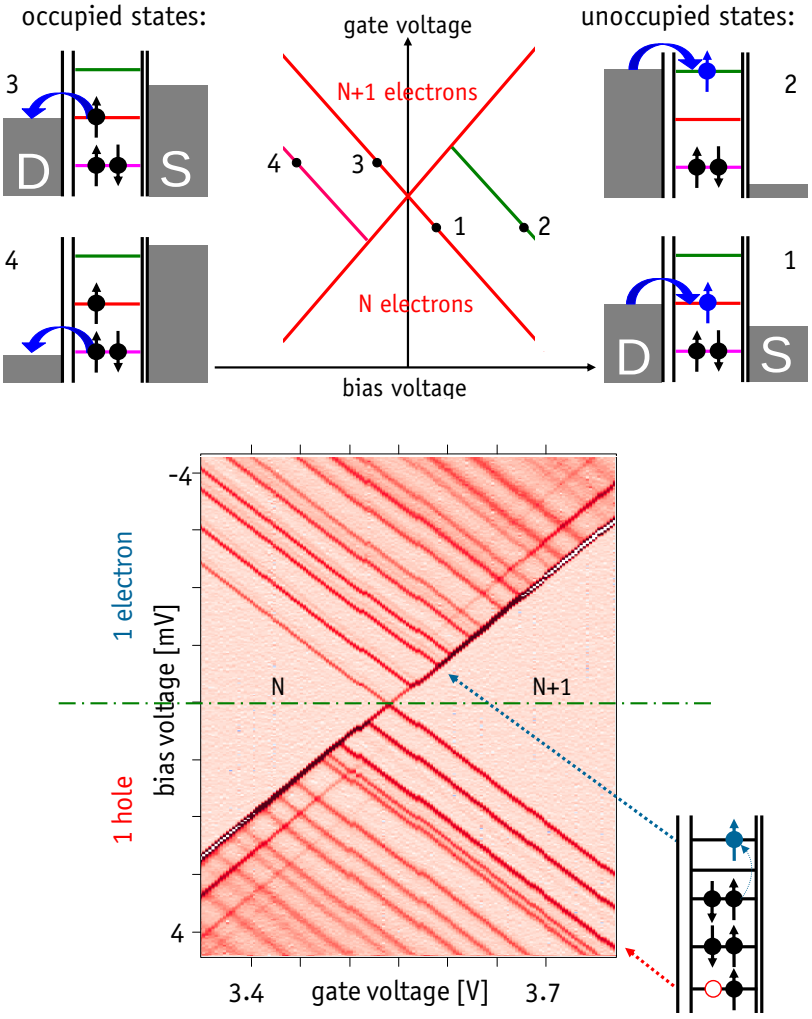


Figure 1.8: Top: Resonances typical for an SET with asymmetric tunneling coupling, operated near the degeneracy point between N and N+1 electrons at $k_bT << \delta << E_C$. The slopes of the conductance features are determined by the capacitances. The distance between lines corresponds to the energy difference between occupied or unoccupied states. Bottom: Real data for a device with asymmetric tunneling coupling. Under certain conditions each resonance may be associated with a configuration involving exactly one excited electron or hole.

Figure 1.8 for real data obtained from a chemically synthesized gold nanoparticle (now the bias axis is the vertical axis). Note that the bias voltage increases towards the bottom, whereas the energy of an electronic state increases toward the top. The resonances at the top can be view as involving an $N+1$ charge state with exactly one excited electron, whereas the lower resonances correspond to a N charge state with exactly one missing electron ("hole"). In Figure 1.8 the rate limiting barrier was on the source side. We will show data corresponding to almost equal tunnel barriers in Chapter 3.

DC versus AC measurements

Each peak in the differential conductance contains information about the tunnel matrix elements, the temperature (or lifetime broadening) and the energy of the associated quantum level. Because of this, it would make sense to measure the differential conductance directly, opposed to the current. This is done in the nanotube experiments by using a lock in amplifier. However, the raw data acquired for metal nanoparticle transistors is the DC current measured as a function of bias voltage (fast scan axis) and gate voltage (slow scan axis). This is because the high impedance of those devices ($\sim 100M\Omega$) in conjunction with the capacitive coupling between the metal wires in the cryostat (10s of pF) makes a lock-in detection impractical. The differential conductance is obtained by taking derivatives numerically.

REFERENCES

[1] D. Huertas-Hernando *et al.*, Phys. Rev. B **74**, 155426 (2006).

[2] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).

[3] T. Ando, J. Phys. Soc. Jpn. **69**, 1757 (2000).

[4] L. Chico *et al.*, Phys. Rev. Lett. **93**,176402 (2004).

[5] A. De Martino *et al.*, Phys. Rev. Lett. **88**, 207402 (2002).

[6] *et al.*, Phys. Rev. B **75**, 121402 (2007).

[7] Y. Yao *et al.*, Phys. Rev. B **75**, 041401 (2007).

[8] H. Min *et al.*, Phys. Rev. B **74**, 165310 (2006).

[9] J. Serrano *et al.*, Sol. State Comm. **113**, 411 (2000).

[10] G. Dresselhaus and M. S. Dresselhaus, Phys. Rev. **140**, A401 (1965).

[11] G. F. Dresselhaus *et al.*, Carbon **4**, 433 (1966).

[12] Y. Aharonov and A. Casher, Phys. Rev. Lett. **53**, 319 (1984).

[13] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 48 (1959).

[14] http://physics.nist.gov/PhysRefData/DFTdata/

[15] R. Saito *et al.*, Phys. Rev. B **46**, 1804 (1992).

[16] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **78**, 1932 (1997).

[17] E. Minot, *Tuning The Band Structure of Carbon Nanotubes*, Ph.D. thesis, Cornell University, 2004.

[18] E. D. Minot *et al.*, Nature **428**, 536 (2004).

[19] H. Mathur and A. D. Stone, Phys. Rev. Lett. **68**, 2964 (1992).

[20] Y. Oreg and O. Entin-Wohlman, Phys. Rev. B **46**, 2393 (1992).

[21] Y. Meir *et al.*, Phys. Rev. Lett. **63**, 798 (1989).

[22] P. Földi *et al.*, Phys. Lett. B **71**, 033309 (2005).

[23] A. G. Aronov and Y. B. Lyanda-Geller, Phys. Rev. Lett. **70**, 343 (1993).

[24] R. Capozza *et al.*, Phys. Rev. Lett. **95**, 226803 (2005).

[25] J. Nitta *et al.*, Appl. Phys. Lett. **75**, 695 (1999).

[26] Y. Lyanda-Geller, Phys. Rev. Lett. **71**, 657 (1993).

[27] T. Qian *et al.*, Phys. Rev. Lett. **72**, 2311 (1994).

[28] P. L. McEuen *et al.*, Phys. Lett. Lett. **83**, 5098 (1999).

[29] L. L. Foldy and S. A. Wouthuysen, PR 78, 29 (1950).

[30] F. Herman *et al.*, *Atomic Structure Calculations*, Prentice Hall, 1963 p. 2-6.

[31] R. J. Elliott, Phys. Rev. **96**, 266 (1954).

[32] R. J. Elliott, Phys. Rev. **96**, 280 (1954).

[33] J. Park, *Electron Transport in Single Molecule Transistors*, Ph.D. thesis, University of California, Berkeley, 2003.

[34] A. N. Pasupathy, *Electron Transport in Molecular Transistors*, Ph.D. thesis, Cornell University, 2004.

[35] M. A. Kastner, Rev. Mod. Phys. **64**, 849 (1992).

[36] M. A. Kastner, Phys. Today **46**, 24 (1993).

[37] W. G. van der Wiel *et al.*, Rev. Mod. Phys **75**, 1 (2003).

[38] A. Wachter and H. Hoeber, *Repetitorium Theoretische Physik*, Springer (Berlin), 1998.

[39] K. R. Kane *et al.*, Phys. Rev. B **51**, 945 (1995).

[40] G. Chen and N. Giordano, Phys. Rev. Lett. **66**, 209 (1991).

[41] U. Larsen, Phys. Rev. B **18**, 5014 (1978).

[42] J. W. Shih, Phys. Rev. **38**, 2051 (1931).

[43] J. Sone, J. Phys. Soc. Jap. **42**, 1457 (1977).

[44] B. Trauzettel *et al.*, Nature Physics **3**, 192 (2007).

[45] B. Trauzettel *et al.*, arXiv:0707.4622v1 [cond-mat.mes-hall].

[46] Y. Yafet, Phys. Rev. **85**, 478 (1952).

[47] H. Mathur, Phys. Rev. Lett. **67**, 3325 (1991).

[48] J. Friedel *et al.*, J. Phys. Chem. Solids **25**, 781 (1964).

[49] http://www.phys.ufl.edu/fermisurface/ (29Aug2007)

[50] M. G. Ramchandani,J. Phys. C, **3**, S1, 1970.

[51] http://en.wikipedia.org/wiki/Allotropes_of_carbon (29Aug2007)

[52] J. Fabian and S. Das Sarma,J. Vac. Sci. Technol. B **17**, 1708 (1999).

[53] K. A. Matveev *et al.*, Phys. Rev. Lett. **85**, 2789 (2000).

[54] P. Brouwer *et al.*, Phys. Rev. Lett. **85**, 369 (2000).

[55] D. A. Gorokhov and P. W. Brouwer, Phys. Rev. Lett. **91**, 186602 (2003).

[56] S. Adam *et al.*, Phys. Rev. B. **66**, 195412 (2002).

[57] J. von Delft and D. C. Ralph, Phys. Rep. **345**, 61 (2001).

[58] J. R. Petta and D. C. Ralph, Phys. Rev. Lett. **87**, 266801 (2001).

[59] D. G. Salinas *et al.*, Phys. Rev. B **60**, 6137 (1999).

[60] S. Guéron *et al.*, Phys. Rev. Lett **83**, 4148 (1999).

[61] D. C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **78**, 4087 (1997).

[62] M. M. Deshmukh, E. Bonet, A. N. Pasupathy, and D. C. Ralph, Phys. Rev. B **65**, 073301 (2002).

[63] N. E. Christensen *et al.*, Int. J. Quantum Chem. **25**, 233 (1984).

[64] P. Pyykko, Chemical Review **88**, 563 (1988).

[65] K. S. Novoselov *et al.*, Nature **438**, 197 (2005).

[66] M. I. Katsnelson and K. S. Novoselov, Solid State Commun. **143**, 3 (2007).

[67] D. Pettifor, *Bonding and structure of molecules and solids*, Oxford University Press (1995).

[68] J. Cao *et al.*, Nature Materials **4**, 745 (2005).

[69] A. H. MacDonald *et al.*, Phys. Rev. B **25**, 713 (1982).

Chapter 2

Fabrication of metal nanoparticle SETs with efficient gating

2.1 Motivation

In order to observe discrete quantum levels of a metallic single electron transistor, the metal island needs to be smaller than 20 nm and hence we cannot use the lithographic techniques which are routinely used to make micron-sized metallic SETs [3–5]. Whereas various techniques exist to fabricate metallic nanoparticles – thin film growth, cluster chambers, chemical synthesis – there is no easy way to incorporate them into a single-electron transistor geometry. One way is to locate a STM tip over a particle on a conducting surface [5, 7, 10, 12], but this is difficult to do at low temperatures and the lack of a gate electrode complicates interpretation of the data [5, 10]. The first observation of discrete energy levels in metallic nanoparticles was based on aluminum particles sandwiched in between two planar electrodes and involved nanoconstrictions in thin silicon nitride membranes [8]. This approach was further improved by incorporating a gate electrode [9], thus turning the device into a single-electron *transistor*. The other breakthrough in contacting nanostructures was the use of electromigration [14] to fabricate electrodes spaced close enough to each other to allow transport through single molecules [13] while maintaining a geometry which allows gating from a third electrode.

Here, we combine the technique of electromigration with the self assembly of evaporated nanoparticles to obtain single electron transistors in which the nanometer sized island is contacted via tunnel barriers by source and drain electrodes, and

coupled capacitively to an oxidized aluminum gate electrode (see [19]). The small size required to observe discrete states in a metal nanoparticle goes hand in hand with a large charging energy, and hence it's generally more difficult to add or remove electrons with a gate electrode than it is for quantum dots based on 2 dimensional semiconductors. The main benefit of the lateral device geometry described in this chapter is that the electrostatic coupling between the nanoparticle and the gate electrode is high enough to allow the addition or removal of several electrons from the island, while maintaining the stability needed to resolve the particle's discrete level spectrum. This allows one to study electron-electron interactions in a nanoparticle by comparing the spectrum associated with N-1, N, N+1,... electrons on the island.

In addition, in principal this geometry allows one to look at or modify the nanoparticle after the device is completed. For example, one can image the particle via SEM, intentionally add ferromagnetic impurities [15], or functionalize the particle for sensing applications). Another advantage of this lateral geometry over the early silicon nitride based devices is that source and drain electrode are not capacitively shorted to eachother. In the future this may allow, for instance, the study of magnetic dynamics in ferromagnetic nanoparticles by use use of high frequency techniques [17].

2.2 Fabrication of the electrodes

Each device - at the end of the fabrication process - is a three terminal device with two closely spaced gold electrodes (source and drain electrode) on top of an oxidized aluminum gate electrode. There's no easy way to incorporate the metal nanoparticle to be studied into the nanometer sized gap between source and drain

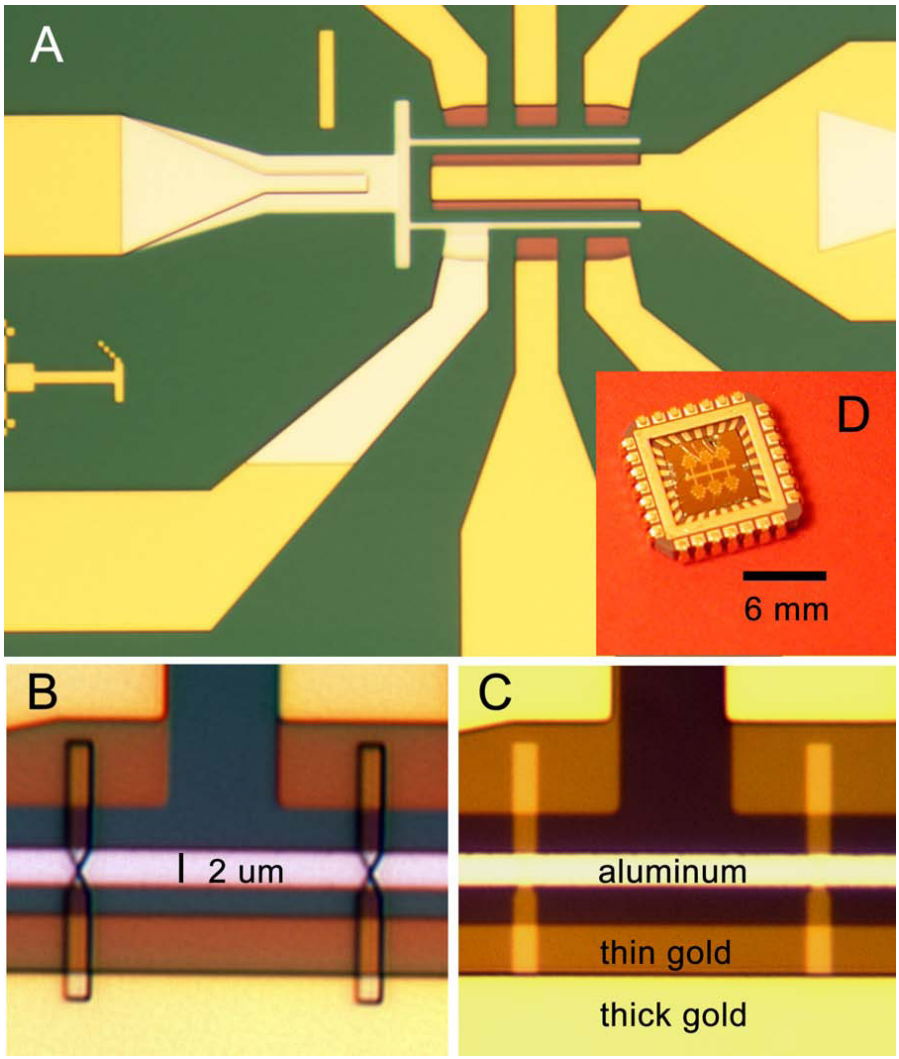


Figure 2.1: A) Optical image before e-beam lithography. The silicon oxide substrate looks greenish, the thin gold layer is brownish and the aluminum gate layer appears white. The thick gold layer connects to bigger bonding pads (not shown). The horizontal electrode in the center (coming from the right) is called the "common" electrode. B) Developed E-beam resist before depositing the break junctions. C) Finished gold break junctions. Because they are only 16 nm thick a thin gold layer is needed to assure good electrical connection. Note that each break junction has one side connected to the "common" (important for noise reduction). D) After dicing, each die is glued into a chip carrier and wirebonded for electromigration.

electrode, and hence the yield of working devices is low. Because of this, and in contrast to semiconducting quantum dots based on GaAs, our fabrication process is based on photo lithography to fabricate many devices in parallel on 100 mm silicon wafers. Electron beam lithography is typically used only in the final step to pattern bowtie shaped nanowires (see Fig. 2.3) which we refer to as "break junctions". The total area exposed by the electron beam is small enough such that a whole wafer containing several thousand break junctions can be exposed within a few hours using a Leica VB6 electron beam system. On each wafer we typically step 12×12 dies, each containing 30 break junctions. The term "break junction" is used because after the main fabrication (summarized in Figure 2.1) is completed, each bowtie shaped nanowire is broken into two closely spaced electrodes (source and drain electrode) using electromigration. Table 2.1 summarizes the main fabrication steps.

Fabrication Summary

We start with degenerately doped[1], $< 100 >$ silicon wafers, 100 mm in diameter, 500 μm thick with 200-300 nm thick thermal oxide. Deep etching of alignment marks is followed by stepping and deposition of layer "thin gold contacts" which will later allow the break junction layer to climb up the sidewalls and connect. Stepping and deposition of layer "thick gold bonding pads" provides low resistive connections to the wirebonds. All stepping is done using the 10x stepper and Shipley S1813 photoresist, in conjunction with an ammonia based image reversal process to obtain undercut. This is especially important for clean lift off of the gate electrode layer, which is deposited in the group's evaporator using a liquid

[1]important if wafers are intended for backgating instead of aluminum gates, such as in Chapter 3.

Table 2.1: Fabrication process starts with a single polished, 100 mm, $< 100 >$ Silicon wafer with 250 nm thermal oxide. Patterning using photo lithography, e-beam lithography, lift off, electromigration and evaporation of gold nanoparticles results in three terminal devices, some of which will show characteristics of a single-electron transistor. Most processing is done at the Cornell NanoScience Facility.

step	equipment	comments
alignment marks	HTG contact aligner	for stepper alignment
	PT72, Unaxis	oxide and silicon etch
thin gold layer	10x stepper	ammonia image reversal
	evaporator	16 nm thin gold
thick gold layer	10x stepper	ammonia image reversal
	evaporator	100 nm thick bonding pads
gate electrode layer	10x stepper	ammonia image reversal
	LN_2 cooled stage	16 nm thin Al, oxidize in situ
break junctions	Leica VB6	MMA/PMMA bilayer
	evaporator	16 nm thin gold lines
testing	probe station	verify gate leakage
dicing	Wafer saw	6 mm $\times$6 mm dies
electromigration	wirebonder	
	LHe dip stick	voltage ramp 30 mV/s
nanoparticles	evaporator	20 Å of gold
		optional oxide cap
testing	LHe dip stick	verify Coulomb diamonds
measurements	TLM dilution unit	measure DC I-V curves

nitrogen cooled stage. E-beam lithography is employed to pattern 16 nm thin gold break junctions. After dicing each die is wirebonded and electromigrated in liquid helium, after which nanoparticles are evaporated onto the chip. The main fabrication process is summarized in Figure 2.1.

Gate electrodes

We use photo lithography to define a 16 nm thick Al gate electrode with 2 nm of Ti as a sticking layer, on top of an oxidized Si substrate. Each gate electrode is nominally 2 μm wide (slightly larger due to overdeveloping during image reversal). The idea is to keep the gates wide enough such that global alignment suffices when exposing the break junctions via e-beam lithography, and narrow enough such that the area of overlap between break junctions and the native gate oxide (which may contain defects and lead to gate leakages) is kept small. Gate electrodes which are less than 2 μm wide are difficult to obtain with the ammonia based image reversal process, but may be an option when using lift-off resist.

Figure 2.2 shows the undercut obtained with the 10x i-line stepper using (g-line) photoresist Shipley S1813 and ammonia image reversal. If the 10x stepper is working properly it should give plenty of undercut ($\sim$ 500 nm), without fine tuning exposure times (1.5 s), focus offsets (0) and developing times (90 s MF321). On the other hand, parameters have to be determined very carefully for the 5x stepper, presumably due to different optics and response of the photoresist to the g-line. The best undercut obtained on the 5x stepper (after multiple focus and exposure arrays) is shown on the right hand of Fig. 2.2. It does not reliably lead to clean gate profiles and we do not recommend using it for gate electrodes. The AFM image (Figure 2.2) clearly shows rough flakes at the edges of the gate ("batman

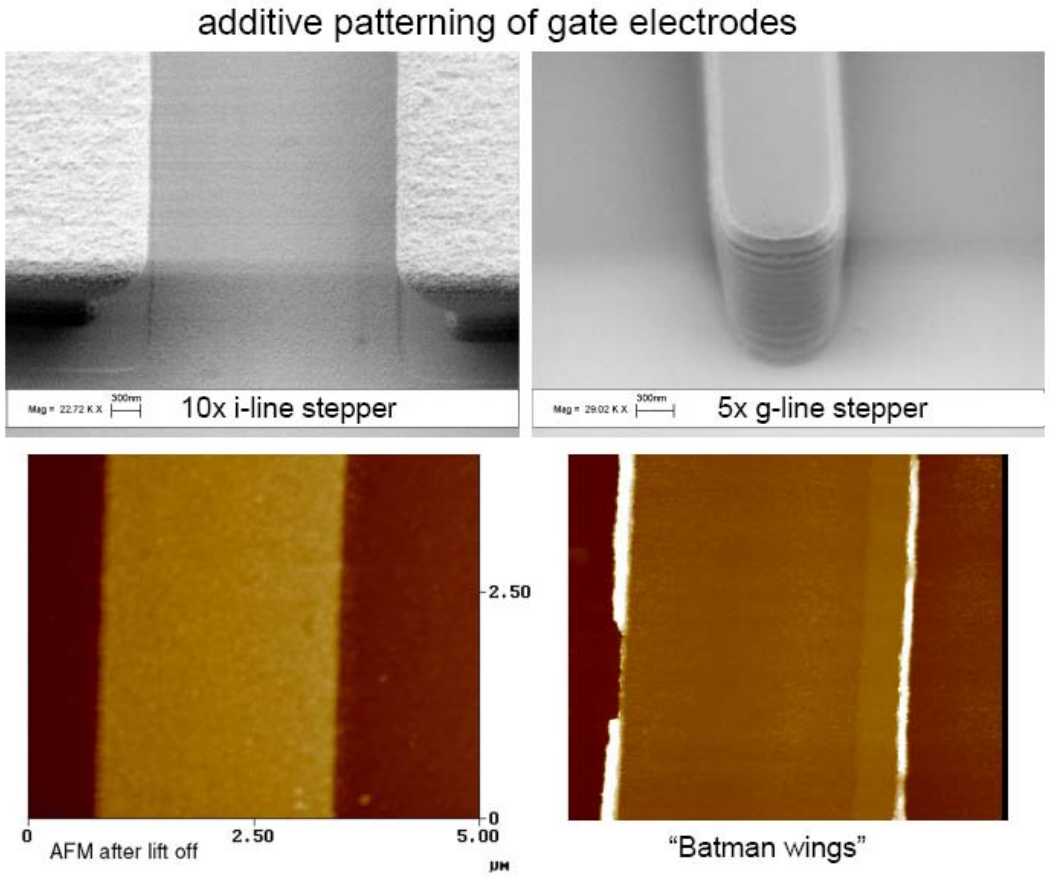


Figure 2.2: When using the 5x stepper instead of the 10x stepper, liftoff may result in rough gate edges due to insufficient undercut after ammonia image reversal. The SEM images were obtained before lift off, and the AFM images below after lift off.

wings"). Even if not as extreme as in this figure, they may cause electromigration to occur at the gate edge and not in the constriction (Fig. 2.7).

The gate electrodes are deposited with the substrate at liquid nitrogen temperature with the intention of obtaining a surface as smooth as possible. We used either thermal evaporation[2] or e-beam evaporation. The stage is warmed overnight to room temperature while in 50 mtorr of O_2 and then exposed to air. The aluminum film should appear clean when viewed with an optical microscope (just like in Fig. 2.1) and smooth when imaged by AFM.

Break junctions

Next we use electron-beam lithography on bilayer MMA/PMMA and liftoff to fabricate Au wires in a bowtie shape ("break junctions") with a thickness of 16 nm and a minimum width of 100-200 nm on top of the gate electrode. It is possible to fabricate narrower lines, but the chances of trapping a nanoparticle may be smaller. The exact width of the constriction depends on the dose. We typically exposed only a single line of pixels (i.e. a 5 nm narrow line) with an appropriate high dose to connect two triangles. Because of the proximity effect, this line would turn into a 100-200 nm wide constriction. The reason for the bowtie geometry (see Figure 2.3) is two-fold. First, it localizes the high current density needed for electromigration to a location where the nm-sized gap is intended to occur (namely above the gate electrode). Second, the widening of the gold lines guarantees a low resistance connection to the bonding pads. This widening is especially important where the gold lines transition from the gate electrode to the silicon substrate, because the edge of the gate electrode may have rough spots.

[2]Aluminum reacts with tungsten once the alumina coating of the boat develops cracks, so it may need to be replaced after each evaporation.

Because the layer "gate electrodes" may be misaligned with respect to the e-beam alignment marks patterned on the layer "thick gold" by a small rotation (shifts of $< 1\mu m$ across wafer), a profilometer is used to measure the offset between the gates and the thin gold layer on 4 corners of the wafer prior to spinning e-beam resist. This way, the break junctions can be placed on top of the gate electrodes using global alignment to 4 alignment marks on the thick gold layer (the contrast of the 16nm thick Al features is too low to allow alignment directly to the gate layer).

After developing and a short descum, 16 nm of clean gold is evaporated, usually without the use of a sticking layer. We use e-beam evaporation at a rate of 5 Å/s and a separate crucible which is dedicated to this step only. A "dummy evaporation" with shutter closed is employed prior to deposition to remove moisture and organic contamination from the gold source. Throughout this process the wafer is mounted on a metal stage at room temperature.

Avoiding contamination

It is advisable to check the deposited film underneath an optical microscope before lift off. In bright field mode black spots on the shiny gold film may indicate tiny gold spheres (spitting during evaporation?) or a contaminated gold source. In that case, the gold source can still be used for depositing the bonding pads etc, but a new source should be used for the break junctions. E-beam sources are easily contaminated in the process of rotating them into the hearth of the e-gun assembly, if the latter had not been cleaned for a while. Another plausible cause of contamination occurs if the venting pressure is set to high. Metal flakes which fall into the venting port at the bottom of the chamber can be stirred up and

contaminate the source. It is a good idea to clean the chamber frequently (including the venting port by an appropriate vacuum extension), to clean the hearth cover on a regular basis (top plate of the e-gun assembly has to be removed), and to protect the source during venting by rotating it into the (clean) hearth.

Dicing

Dicing constitutes the ticket out of the clean room. It also provides the opportunity to destroy weeks of work (and delay months of fun) very efficiently. We used to send out wafers for dicing to a company which did a good job, but required time and a minimum quantity of wafers to dice. We enthusiastically made use of a new dicing saw at the CNF with the result that whole wafers were unusable because of leakage between the gates and the source/drain electrodes. It turned out that the leakage was not caused by the cutting process itself (vibrations and heat produced by a fast spinning diamond blade) but by the jet of cooling water incident onto the wafer (which we protected by a layer of photoresist as usual). It presumably pushed the gold lines through the native oxide of the gate electrodes.

The solution to this problem is as follows. After completing lift off — by soaking in multiple baths of acetone, final rinse in IPA and drying with dry nitrogen — characterize four corners of the wafer on the probe station in terms of two criteria: electromigration and gate leakage. Electromigration tells you the resistances of the lines (they should be consistent with each other, say 60-70 Ω) as well as whether they break consistently at room temperature (typically around 1 V). Gate leakage is tested by ramping up a bias voltage between the common electrode and a gate electrode, with the gain of the current preamp set to 10^{-9} A/V and maximum filtering (100 ms). The current should be zero up to 2 V. As soon as the current

smoothly but positively turns on and reaches 20 pA, the bias voltage is ramped down to prevent damage of the gate oxide. The voltage at which 20pA are reached is our criteria whether gates are in good shape or not. It should be between 2 and 4 Volts. If it is much higher or does not occur at all, your gates may not be connected to the bonding pad. After the wafer is characterized a layer of S1827 is spun at 2000 RPM and baked at 115 for 60 s. Then another layer of S1075 is spun at 2000 RPM which is very thick and needs to be baked in the convection oven at 90℃for 30 minutes. Baking on the hotplate will result in a rippled surface. This provides a protection layer which is sufficient to withstand CNF's dicing saw.

After dicing verify that the gates previously checked leak at the same voltage. The rest of the chips are stored with photoresist until needed. The photoresist is removed by soaking the chips in multiple baths of acetone followed by a rinse in IPA. Never let acetone dry on the chip because of residues it will leave (which are hard to get rid of). After a chip is cleaned in oxygen plasma[3] it is glued into a chip carrier with silver paint (the vapors of super glue may contaminate the surface while drying) and wirebonded before cooldown in a liquid helium dipstick. Caution: Break junctions are especially susceptible to electrostatic discharge once they are wirebonded. Keep yourself grounded when handling the chip carrier. When removing the carrier from sockets (or when inserting) use tweezers which are grounded. Carry the chip carrier in a gel box with conducting case. Unground the lines to the socket only if needed for measurement, and keep the wires grounded at all other times (using a switchbox with a make-before-break rotating switch).

[3]with power settings high enough to remove a few nanometers of gold

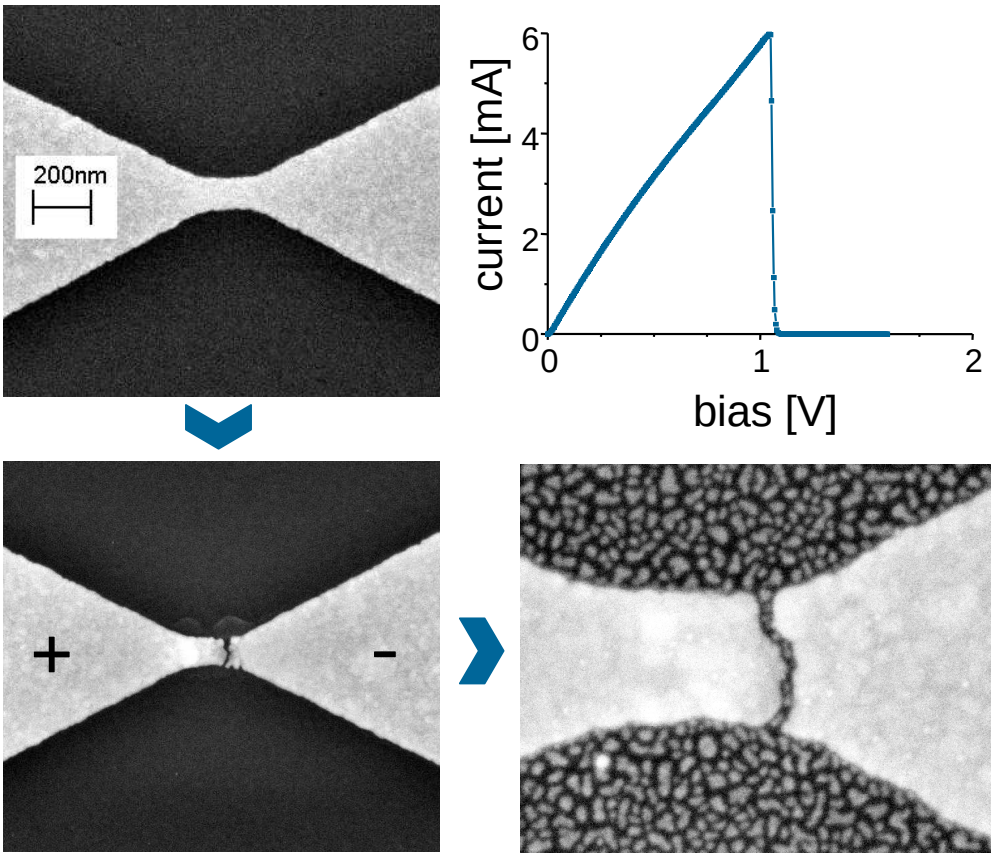


Figure 2.3: In each break junction a nm sized gap is induced using electromigration, by slowly ramping up a bias voltage until the wire fails. After evaporation of a small quantity of gold a single nanoparticle may bridge the gap, thereby completing the single electron transistor.

Electromigration

We submerge the wirebonded chip in liquid helium and break the wires using electromigration; a source-drain bias is slowly ramped up (30 mV/s) until the wire breaks and the conductance drops suddenly[4].

In most cases this happens at a bias of ∼1 V (see Fig. 2.3), and results in a gap about 5-10 nm wide after the sample is warmed to room temperature (also confer table of gap sizes in Table 2.2). A similar process of electromigration has been used to make single-molecule transistors [13].

[4]The potential of the gate electrode is left floating during this process.

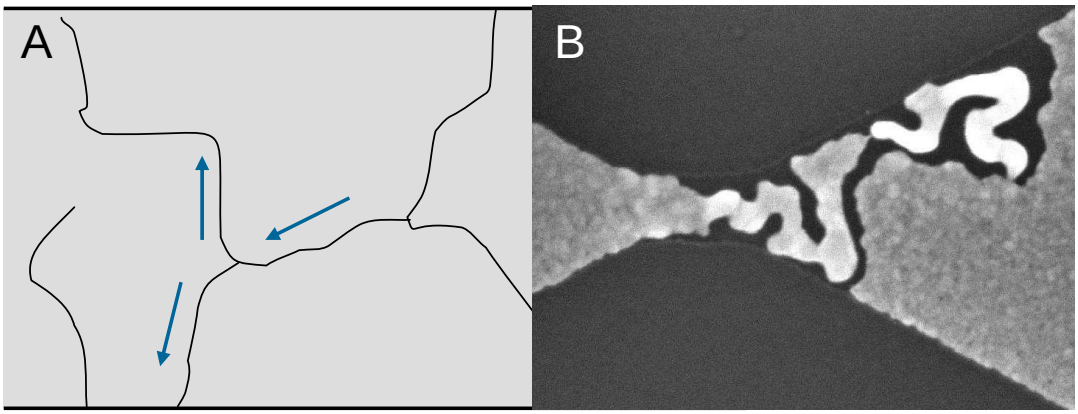


Figure 2.4: A) Illustration of a triple point turning into a void: more atoms move away than towards the triple point (assuming electrons flow from right to left). Electromigration occurs where momentum transfer from scattered electrons is largest and atoms are most mobile. Both conditions are met at grain boundaries, and depending on their geometry a complex transport of gold atoms may occur (B).

Because electromigration is the crucial step in establishing contact to a nm-sized object, a few comments should be made. For electromigration to happen one needs a force acting on the ion and a mobility of the ion. Both requirements are fulfilled simultaneously for atoms near vacancies and grain boundaries, because the driving force of electromigration is the momentum transfer from electrons which scatter off an ion ("wind force"). However, a lot of electrons are needed, and hence electromigration only occurs in metals or heavily doped semiconductors[5]. In our gold break junctions a current density of order 10^{12} A/m^2 is needed to initiate electromigration. Electromigration can occur underneath the break junction (and hence may be modified if a sticking layer is used) as well as on the surface (where the presence of thiols may influence electromigration). Elevated temperature will also facilitate both mobility and scattering. In gold break junctions we find that gold atoms are pushed in direction of the current flow, and we have studied the process of electromigration by taking videos while performing electromigration in a scanning electron microscope [16].

The mass flux induced by electromigration can lead to the formation of nm-sized gaps only if more ions are leaving than arriving. Such flux divergences occur where different grains meet. Consider for example the "triple point" in the schematic of Figure 2.4. Under the influence of electron flow from the right to the left, more atoms are leaving than arriving, and the triple point may develop into a void and eventually a nm-sized gap. In fact, electromigration in today's computer chips is prevented by engineering of the lines' grain structure. Another indication for the importance of grain structure is the fact that not all of our wafers behaved consistently. Some wafers would lead to nice gaps whereas other wafers caused

[5]The failures occur in different directions for p and n-type semiconductor

Table 2.2: Gap sizes estimated from SEM images for 11 nm and 16 nm thick gold and 10nm thick platinum break junctions, after performing electromigration (50 mV/s) in ambient conditions (LHe). The gold was deposited at 3 Å/s, the platinum at 0.8 Å/s, while cooling the stage with cold water.

Material	broken at RT	(in LHe)	# of break junctions
16 nm Au	8.0 ± 2.0 nm	(7.5 ± 2.0) nm	4 (6)
11 nm Au	5.4 ± 1.4 nm	(11.2 ± 2.4) nm	9 (7)
10 nm Pt	5.5 ± 1.8 nm	(2.8 ± 0.7) nm	9 (10)

problems, such as in Fig. 2.4. In periods of bad luck, we have tried different types of evaporations (electron beam evaporations vs. thermal evaporation), different types of evaporators (in Clark Hall and CNF) and different temperatures of the substrate while depositing the break junctions, until a wafer with satisfactory electromigration properties emerged.

We have also tried electromigration on aluminum break junctions, but found that it is more difficult to create a nm-sized gap (Fig. 2.5). This is probably due to the native oxide layer which forms on the aluminum. Electromigration did occur, and resulted in spherical droplets of aluminum popping up through the native oxide a couple of 100 nm away from the depleted region.

2.3 Evaporation of nanoparticles

After creating a nanometer sized gap between source and drain electrode, a metal nanoparticle has to be inserted to complete the single-electron transistor.

Figure 2.6 shows that nanoparticles of sizes appropriate for quantum dots are easily obtained for various metals, by simply evaporating a small quantity of metal

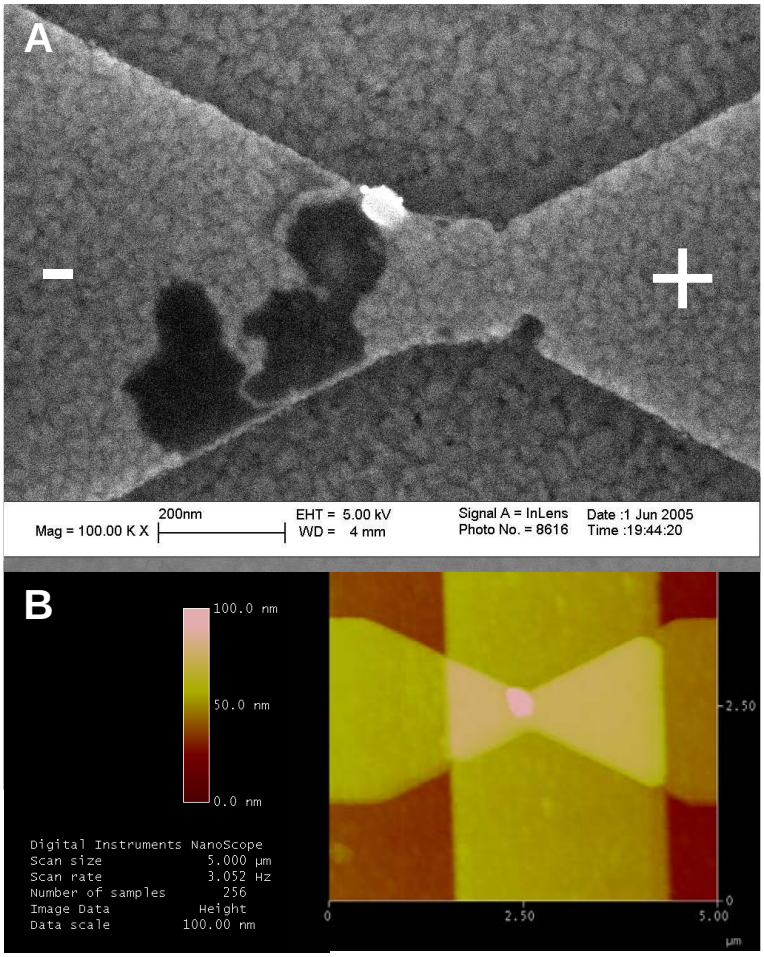


Figure 2.5: Electromigration performed on an aluminum break junction did not result in reliable gaps. Presumably Al atoms migrate underneath the native oxide layer and pop out at a weak spot (SEM image at top). Note the height of the protrusion (AFM image at bottom).

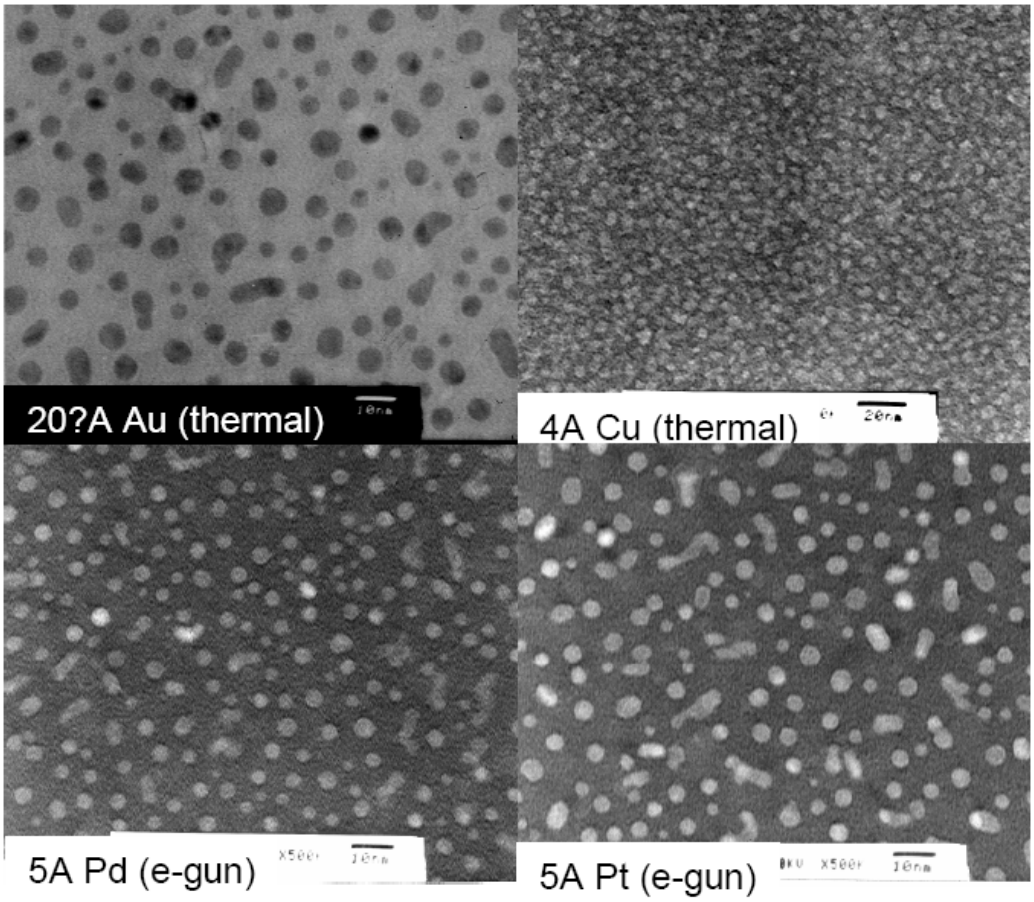


Figure 2.6: Evaporation of a small quantity of metal onto a silicon nitride membrane leads to the self assembly of nanoparticles, as viewed here with a TEM. Scale bars are 10, 20, 10 and 10 nm.

.

onto a suitable substrate. Instead of a continuous film, lots of nanoparticles self assemble under the influence of surface tension. Their size and shape are not very well controlled, but knowing how much metal we deposited we can estimate their aspect ratio from the TEM image. We find that it is consistent with a roughly hemispherical droplet shape. However, one should be cautious to infer from such TEM samples the shape and size of the nanoparticles in the actual device, as their nucleation and mobility will be affected by the substrate as well as the confined geometry (note the difference in size for particles inside and outside the gap in the SEM of Fig. 2.8). The demonstration of nanoparticles on the TEM grid (we actually used a 50 nm thick silicon nitride membrane) looked promising, but further refinement of the evaporation process was needed to obtain successful devices.

We achieved this by mounting the wirebonded chip on a stage of an evaporator which had been modified with a 24-pin military connector and a suitable "make before break" switch box. This way, the conductance of each break junction could be monitored safely while evaporating a small quantity of gold. A useful benchmark in determining how much metal to deposit to get nanoparticles is the percolation limit, which can be established by measuring a couple of lines in situ while depositing more and more metal.[6]

In order to get nanoparticles, one has to deposit a fraction of the percolation dose, and the exact amount has to be determined iteratively. For example, if none of the break junctions showed conductances <100 $G\Omega$ after deposition of 15 Å of gold, we simply evaporated another 5 Å and checked the resistances again. If none of the lines dropped below 1 $M\Omega$ or so after 30 Å, it was usually a good idea to

[6]To avoid the increased noise which is present during evaporation we would sometimes stop evaporation momentarily and measure the resistance of a few break junctions, and then continue evaporation until the majority of lines is shorted.

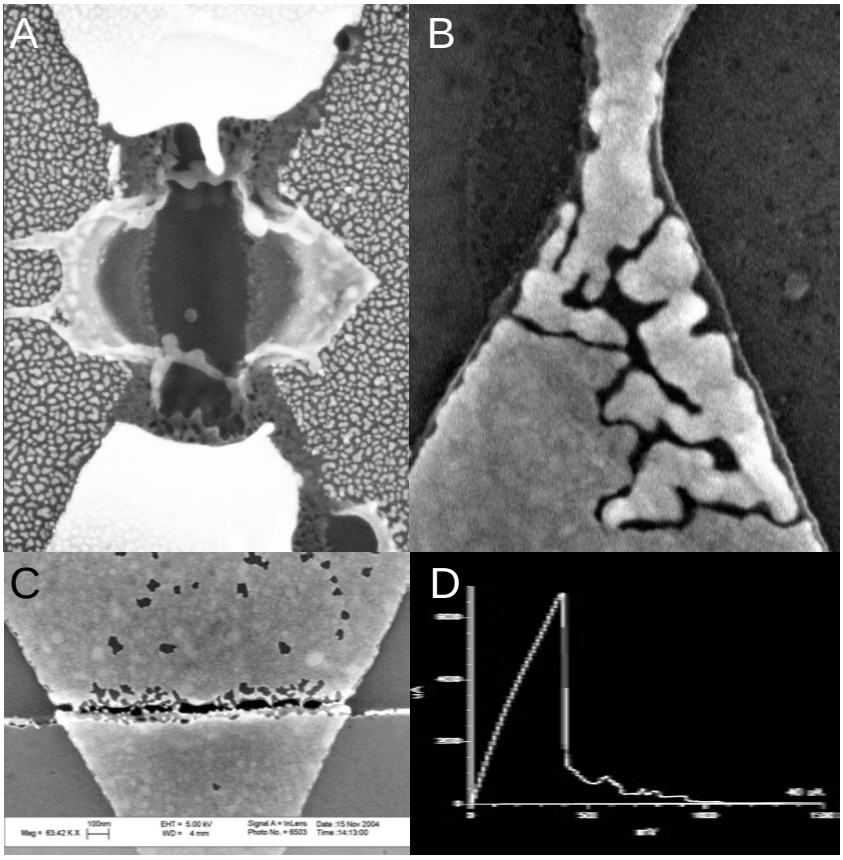


Figure 2.7: Some of the common failures can easily be identified via inspection by SEM and responded to accordingly. Electrostatic discharge (A, ground yourself, avoid arcing during evaporation), unfortunate course of electromigration (B, use different wafer) or electromigration at the wrong location (C and D, check gate profile via AFM)

.

skip the oxide capping and take the sample directly to a good scanning electron microscope. This can save a lot of time because it may immediately tell whether something went wrong during evaporation, electromigration, or while handling (Fig. 2.7).

For example, one problem we have encountered multiple times is that electromigration did not occur in the constriction of the nanowire, but at the edge of the aluminum gate where the thin gold wire climbs up the edge of the nominally 16

nm thick gate electrode. Sometimes this problem is apparent from the IV curves while performing electromigration, with a longer tail of current than usual (40 μA at 1.5 V in Fig. 2.7). This can be caused by imperfect liftoff when patterning the aluminum gates, and may depend on the location on the wafer from which the particular chip was taken[7]. More about the art of successful gates is addressed in section 2.5.

Successful devices

Successful devices were obtained from the following procedure: Electromigration was performed while the break junctions where in physical contact with the liquid helium. The dipstick was then removed from the bath and laid flat on a table in a position such that the chips surface was vertical. Ice and moisture will still build up on the chips surface, but to prevent lengthy exposure to water a gentle stream of hot air was directed towards the chip until it reached room temperature, but not longer. The chip was removed from the socket, and tiny droplets of water left on the chips surface were removed by a gentle stream of dry nitrogen. The chip was immediately loaded onto an appropriately wired stage of the evaporator, roughed out with a oil-based mechanical pump, and pumped to high vacuum with a cryo pump. The resistances of all lines (typically larger than detectable with a gain of $10^{-9}A/V$) were measured after performing dummy evaporations with shutter closed. Then 20 Å of gold was deposited at 1 Å/s, using thermal evaporation in a tungsten boat with alumina boundary. Thermal evaporation was preferred to e-beam evaporation (which would sometimes arc), but we have made successful

[7]If image reversal worked fine and undercut is plenty all chips on a 4" wafer should have a perfect square profile. However, if undercut is barely enough only the central part of a wafer may have clean gates, due to the finite working distance during evaporation

devices both ways. After evaporation, break junctions were ungrounded and their resistances measured. If 10 break junctions out of 24 showed resistances between 1 $M\Omega$ and 10 $M\Omega$ the deposition was considered a success and a capping layer of silicon oxide or aluminum oxide was deposited. We have subsequently reduced the thickness of the capping layer, and obtained good results even without capping at all. This allows one to image the device with a scanning electron microscope (Fig. 2.8).

Testing in liquid helium

The completion of the SET by simply evaporating a small amount of metal may seem like pure luck. It should be pointed out that if the dose is kept low enough such that the majority of devices remain open ($> 10G\Omega$) chances are good that some of the conducting lines are due to a single nanoparticle only. In such a device, there may be other nanoparticles in the same gap coupled to one or the other electrode, but as long as they don't carry tunnel current their only annoyance may be to cause jumps in the effective gate voltage whenever they change their charge state during a measurement (see Fig. 2.9). If two (distant) particles contribute tunnel current in parallel, it is very likely that their degeneracy points won't coincide. In that case, the spectrum of one particle can be measured at low temperature while transport through the other particle is Coulomb blocked.

Whether a device has potential for showing a clear level spectrum at mK temperatures can be judged from the shape and regularity of its Coulomb blockade diamonds at 4.2 K, as well as its overall conductance. The left sample in Fig. 2.8 has all characteristics to justify further study in a dilution unit: Regular Coulomb diamonds indicating a single particle, high gate coupling and an overall resistance

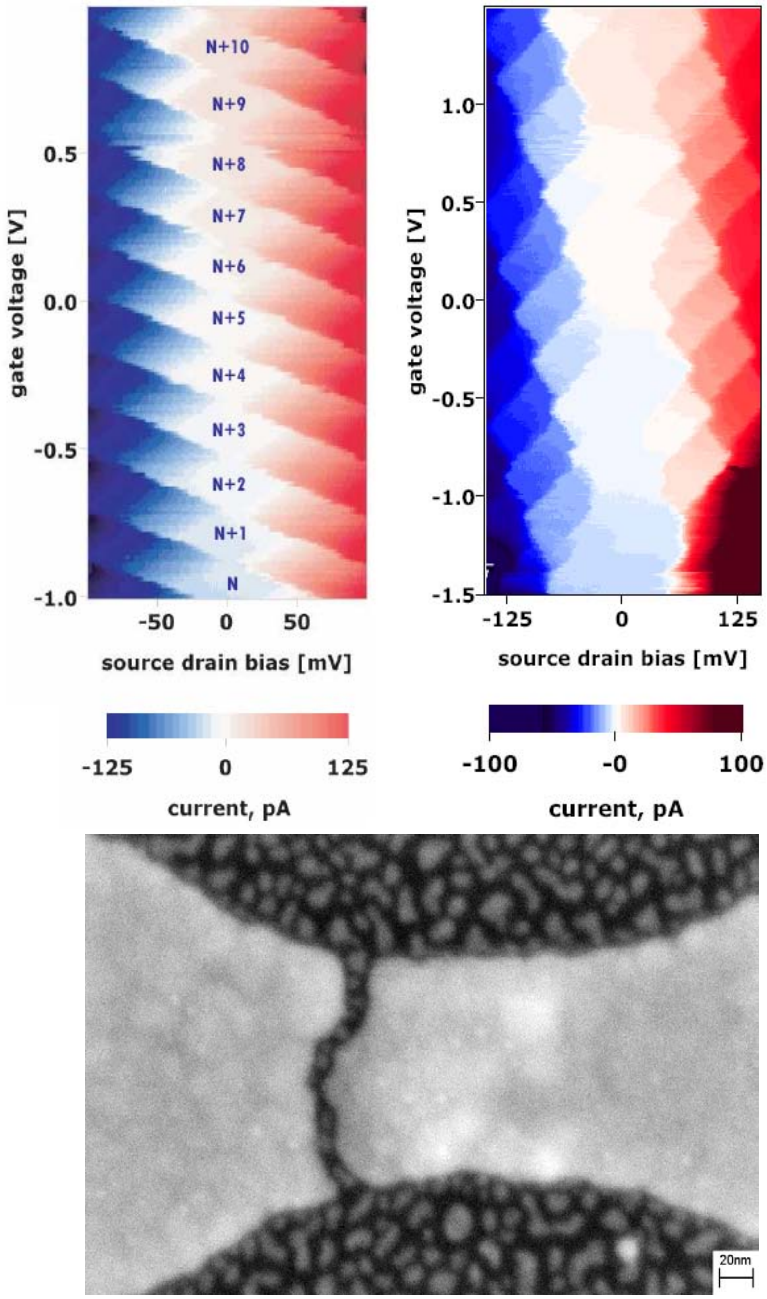


Figure 2.8: Coulomb diamonds of two SET devices based on evaporated gold nanoparticles. Left: Regular Coulomb diamonds with vanishing blockade at the degeneracy points, typical for current through an isolated nanoparticle. Right: Additional features indicate the presence of additional capacitances, probably due to a second nanoparticle participating in series (note the absence vanishing Coulomb blockade). Below: SEM image of the left device after measurements in a dilution refrigerator were completed.

which makes it likely that current levels will remain high enough (pA per quantum level) after warm up to room temperature and transfer into a TLM dilution refrigerator.

The regular geometry of a Coulomb blocked "diamond" is a good indication that an individual nanoparticle is measured. Without going into detail we note that further support comes from quantitative modeling of the current. Transport through a single nanoparticle at 4.2 K can be modeled by assuming a continuous density of states of the particle, and by specifying all three capacitances and the two tunnel resistances. Usually one tunnel resistance is much larger than the other due to the exponential dependence of the tunneling current on the distance between the nanoparticle and the electrode. Fig. 2.9a demonstrates that the IV curve obtained from a gold nanoparticle ("Coulomb staircase" at 4.2 K) can be reproduced accurately by simulation (offset for clarity). From the model, we obtain the values $C_S = 1.3$ aF, $C_D = 2.4$ aF, $C_G = 0.89$ aF. The ratio between the right and left tunnel resistance was indeed very high ($R_R = 2.7$ GΩ and $R_L = 0.08$ GΩ).

In panel d several jumps in the offset charge occurred at certain gate voltages.

Performing tunneling spectroscopy in a dilution refrigerator

The gate capacitances extracted from Fig. 2.9c are an order of magnitude larger than the gate capacitances achieved using the previous generation of devices based on silicon nitride membranes, and hence several electrons can be removed or added to the nanoparticle using the gate electrode. To show that discrete quantum states can be resolved in these devices, we performed transport measurements at dilution refrigerator temperatures. Filtering of all electrical connections and careful choice of grounds and data acquisition are crucial to obtain low electronic temperatures

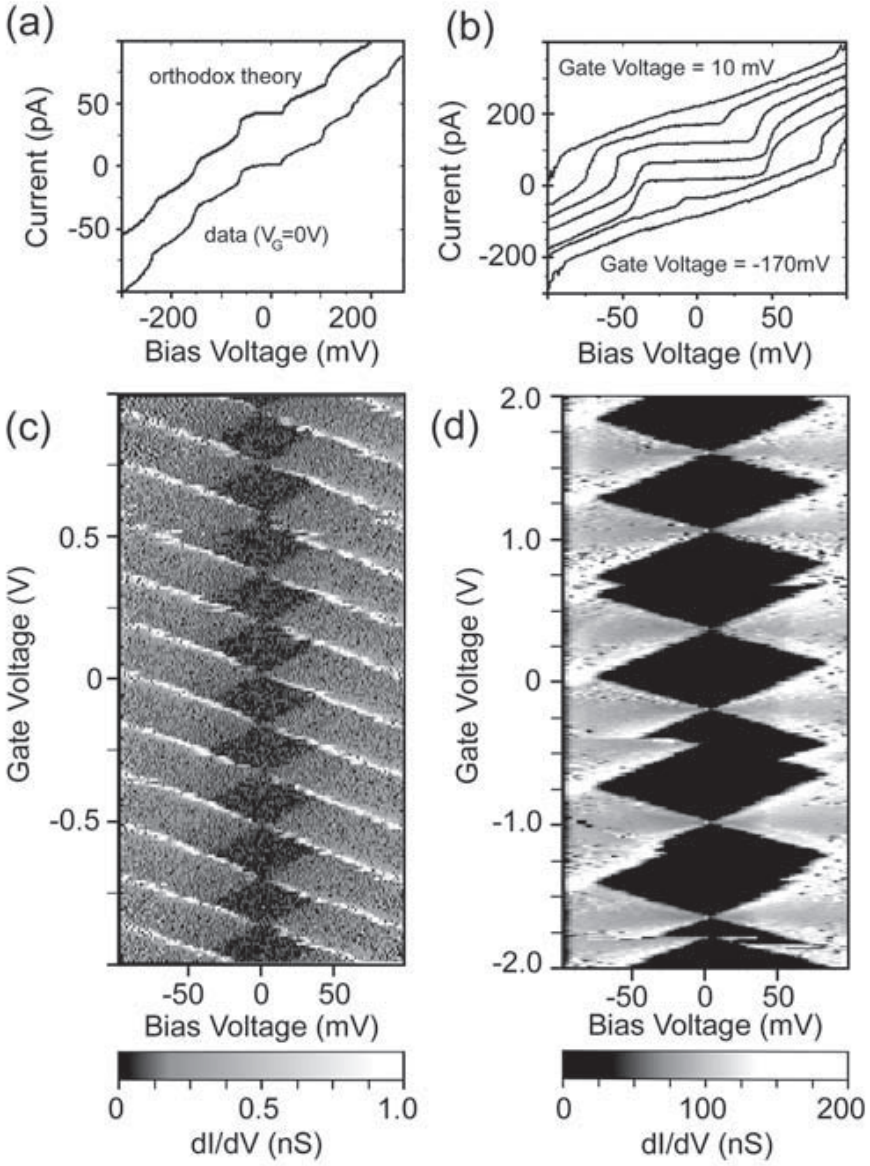


Figure 2.9: a) Coulomb staircase IV curve for a gold nanoparticle SET at 4.2 K, along with an orthodox model fit (offset for clarity). b) IV curves at 4.2 K, for equally spaced values of V_G. c) Gray-scale plot of dI/dV as a function of V_G and V at 4.2 K. Eleven degeneracy points separating twelve different charge states are visible within a 2 V range of V_G. d) Gray-scale plot of dI/dV as a function of V_G and V for a different device at 4.2 K, showing "Coulomb diamonds" as well as several abrupt changes in the background charge of the SET island as V_G is swept. From [19].

(≈ 100 mK) and high enough signal-to-noise ratios.

Data acquisition

We used a setup similar to the ones described in [5, 21]. Briefly, two GPIB-controlled wave generators (HP 3325A) were used as DC voltage sources for bias and gate voltage[8], with a π-filter inserted between the BNC cable and the break-out box of the probe. The bias voltage was amplified by a voltage amplifier and measured continuously by a differentially configured DAQ-card. The gate voltage was initialized and measured by a DAC-card once. During acquisition the gate voltage was disconnected from the DAC-card and inferred from the phase of the wave generator. The current was measured by a Ithaco current amplifier in close proximity to the break-out box. A 36" triax-cable (with outer shields shorted to each other for better shielding) was used to connect the input of the current amplifier to the break-out box without any filter. The output of the current amplifier was measured via a differentially configured DAC-channel. This was a pure DC measurement and no lock-in amplifier was involved. The differential conductance is obtained from the IV curves numerically. Ground loops were avoided by using isolation transformers. Humidifiers were sometimes employed to reduce the risk of electrostatic discharge when loading the sample. Some models were found to cause a considerable amount of electrical noise when running, and a smaller but detectable 60 Hz noise when turned off. Humidifiers were unplugged during measurements. Another source of 60 Hz magnetic pollution was associated with the resistance bridge in the area of its display[9] (which is in close proximity to the alu-

[8]Data in Chapter 3 was obtained using a Yokogawa voltage source on the gate electrode.

[9]A loop antenna, shielded to make it sensitive to magnetic fields only, was hooked up to a current amplifier and used as probe.

minum break-out-box). Because aluminum does not shield against low frequency magnetic fields, the unit was sometimes turned off.

Interchanging source and drain electrode

Another possibility to reduce noise is to interchange source and drain connections. This turned out to be very effective, and hence we shall describe in detail what we mean by it. In our design, all break junctions on a given chip have one electrode connected to each other (referred to as the "common electrode"). However, after cooling down it is unknown which of the potential break junctions is good and which ones are dead. The only way to find out is application of a bias voltage large enough to tell whether a device is dead or Coulomb blocked, but not larger. In order to not destroy a potentially good device while testing another, it is a good idea to initially connect the common electrode to the current preamplifier, and the individual side of the break junction to the bias voltage. That way all other devices connected to the common electrode don't feel the bias voltage. Once a good sample has been found the obtained data is of higher quality if bias voltage and current amplifier are exchanged. That way, the bias voltage is applied to the common electrode, and any current fluctuation produced by the other devices is not picked up by the current preamplifier. In addition, if all other wires are grounded they act as a shield.

Fig. 2.10 demonstrates that the particle's discrete states can be resolved in the vicinity of a degeneracy point. Using the gate electrode to add or remove electrons, the discrete level spectrum can be measured at various charge states of the nanoparticle, and these can be compared to each other.

There are two features in Fig. 2.10 which we will study further in the next

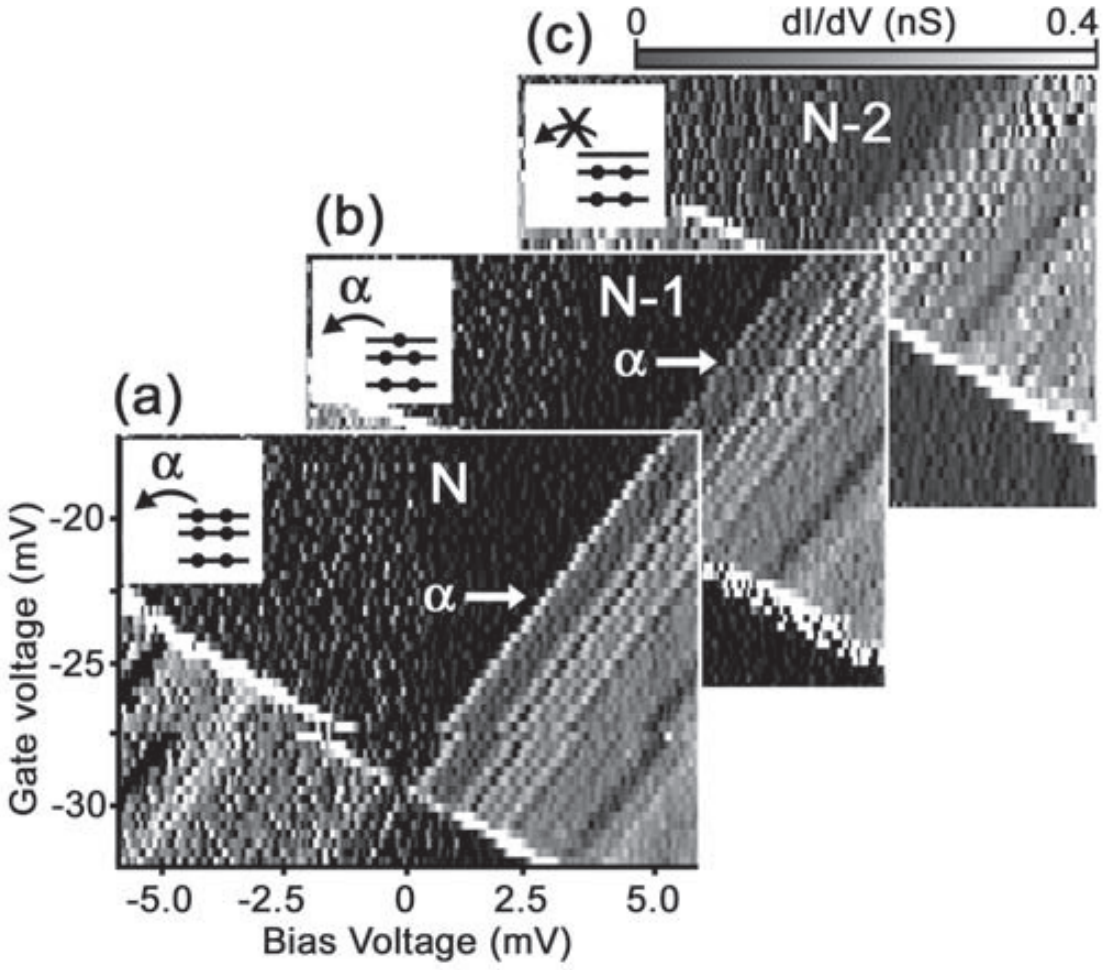


Figure 2.10: Gray-scale plots of dI/dV as a function of V_G and V, displaying the discrete electron-in-a-box level spectra within a gold nanoparticle, measured with an electron temperature less than 150 mK and zero magnetic field. Panels (a), (b), and (c) represent spectra for different numbers of electrons in the same nanoparticle. Panel (b) covers the gate voltage range from -95 to -110 mV and (c) the range from -180 to -195 mV. Insets: Energy-level diagrams illustrating the tunneling transitions that contribute to line α for different numbers of electrons on the particle. From [19].

section:(i) There's a correlation between the level spectra taken at different charge states. (ii) Each time the charge state is changed by two electrons, a resonance line (dis)appears.

2.4 Absence of electron-electron interactions

Motivation

In the constant interaction model of a quantum dot the electron-electron interaction is taken to be the classical charging energy [4], and quantum effects enter the model only through the specified single-particle levels, which may be simple standing waves or complicated chaotic wavefunctions. The constant interaction model neglects residual interactions which may, for example, determine the total spin of the many-body state. Denis A. Gorokhov and Piet W. Brouwer investigated the combined effect of spin-orbit scattering and electron-electron interactions on the probability distribution of g-factors in metal nanoparticles [23]. They describe the nanoparticle by the Hamiltonian

$$H = H_0 - J\vec{S}^2 - 2\mu_B B S_Z \tag{2.1}$$

where the first term describes the single-particle levels, the second term is the "constant exchange interaction" from the universal Hamiltonian for electron-electron interactions and the third term is the Zeeman coupling to a magnetic field B. Here $\vec{S}$ is the total spin of the particle. For many metals the ratio of the exchange constant J and the mean level spacing δ is between 0.2 and 0.4. In analogy to the bulk Stoner criterion one may derive the critical value of J/δ beyond which "mesoscopic magnetism" will occur (i.e. many-body ground states with high total

spin). Specifically, consider an even number of electrons occupying spin-degenerate single-particle levels such that the total spin is zero. To form a state with $S = 1$ one of the highest electrons needs to be promoted to the next higher level (Pauli exclusion principle). The kinetic energy required is the level spacing δ, whereas the exchange energy gained is $J \times S(S+1) = 2J$. We conclude that the criterion for a non-trivial spin-state ($J > \delta/2$) is weaker than the bulk Stoner criterion ($J > \delta$) [27]. Furthermore, level spacings in randomly shaped nanoparticles are non-uniform, and hence situations may occur in which the first excited state happens to be very close in energy, favoring a non-zero spin state even if the expected average ratio is $J/\delta < 0.5$. For gold, electron-electron interactions are predicted to be very weak ($J/\delta < 0.1$ [23, 34]), and in the following section we shall test this prediction experimentally.

Results

To experimentally show that exchange interactions are weak in gold nanoparticles, one has to either infer the total spin or measure energy shifts associated with the second term in Eq. 2.1. In principle, the addition spectrum reflects the magnetic moments associated with ground states, and may reveal the total spin. This is a practical technique for carbon nanotube quantum dots, but not for metal nanoparticles with their high charging energy and frequently occurring shifts in the offset charge when the gate voltage is swept by a large amount. However, non-zero spin states may reveal themselves also through a "spin-blockade" effect [8, 9]: If the ground states associated with two adjacent charge states have different total spin such that $\Delta S > 1/2$, the ground-state to ground-state transition is suppressed due to conservation of angular momentum when tunneling one spin-1/2 electron.

We have looked for spin blockade in evaporated gold nanoparticles, but have not encountered any. The second approach is to compare excitation spectra of different charge states. Without exchange interaction, no changes are expected between different charge states, whereas additional features associated with singlet and triplet states (split by 3 $J/4$) or other multiplets are expected for $J \neq 0$.

Figure 2.11 presents a sequence of level spectra obtained from different charge states of a device with a mean level spacing of 1.42 meV. The striking feature is the appearance of a set of resonances which is not modified by adding or removing electrons. However, one resonance disappears on the positive bias side and reappears at negative bias every time two electrons are added to the system. This is the main finding of this section and we shall walk through it in detail below. Also, the Coulomb blockade does not vanish between N and $N+1$ electrons, possibly due to the proximity of other nanoparticles inside the gap. This may also explain the excess of noise in parts of the spectra, and various jumps in the offset charge when changing the gate voltage.

Table 2.3 quantifies capacitances, level spacings and g-factors extracted for the device in Figure 2.11.

Next, let us describe in detail how the main features of Figure 2.11 are explained by the constant interaction model, in which two-fold degenerate states (as required by time-reversal symmetry are filled one after the other as in a non-interacting system. A closeup of three successive charge states is given in Fig. 2.12 along with additional data obtained with an applied magnetic field of 8 Tesla. We note that the rate limiting barrier in this device is between the nanoparticle and the drain electrode, as indicated by the appearance of predominantly negatively sloped resonances.

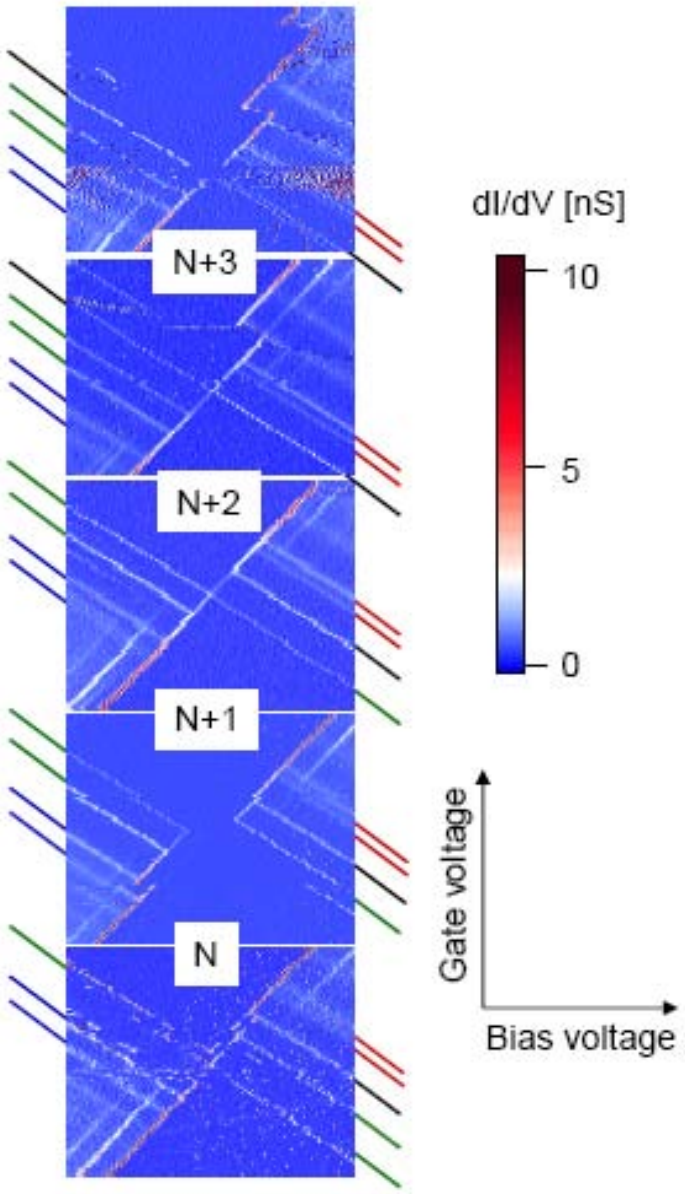


Figure 2.11: Even though the charging energy is much larger than the level spacing, the discrete level spectrum hardly changes when electrons are added to the dot. For actual scales refer to Figure 2.12.

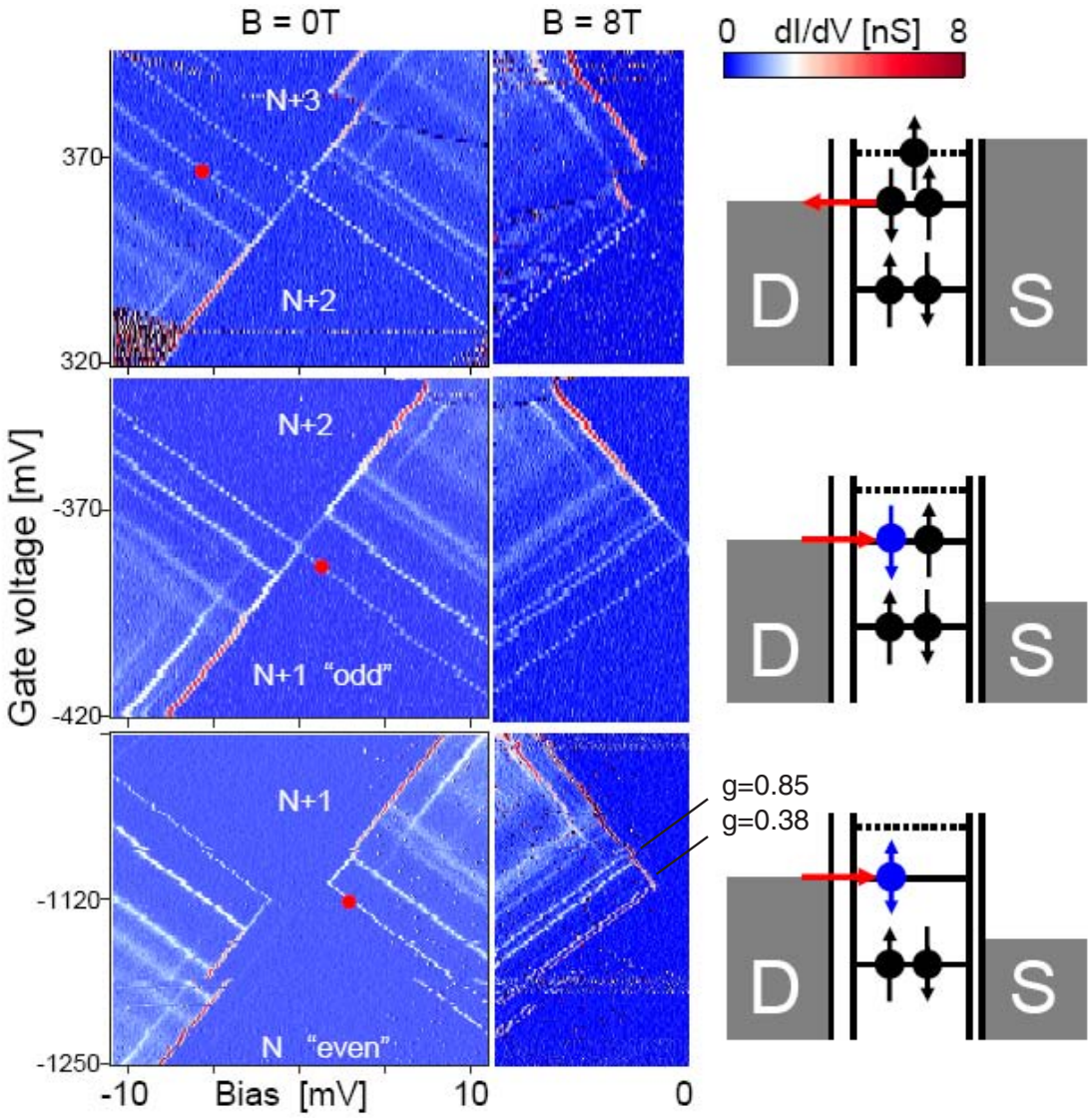


Figure 2.12: At $B = 0$ T spin degenerate states are filled as in a non-interacting model. A resonance associated with an unoccupied state gets weaker if one electron is added, and reappears at negative bias (occupied state) after two electrons are added. Zeeman splitting in a magnetic field confirm this even odd level filling. The g-factors associated with the two quantum levels are 0.35 and 0.85.

Table 2.3: Extracted parameters for the device in Fig. 2.11. The gate capacitance is 0.23 aF, the charging energy is 40 meV, the mean level spacing 1.42 meV.

lower charge state	N-1	N	N+1	N+2	N+3	N+4
C_S/C_g	3.3	3.0	3.1	3.0	3.0	3.1
C_D/C_g	4.4	4.3	4.3	4.2	4.3	4.1
δ [meV] (g-factor)				1.73	1.83	2.05
		2.2	2.1	1.7	1.71	2.31
	0.56	0.53	0.61	0.55	0.61	0.53
	1.51	1.48	1.58	1.54	1.26	1.47
	1.33 (0.76)	1.31 (0.85)	1.54 (0.75)	1.50 (0.88)	1.49 (0.87)	1.35
	1.48 (0.39)	1.34 (0.38)	1.30 (0.35)	1.30 (0.31)	1.30	2.0
	1.99 (0.65)	1.88 (0.62)	1.90 (0.56)	1.75	1.68	
	1.15 (0.66)	1.07 (0.56)		0.95	1.1	
	1.23	1.0				

Bottom panel

We assume that N is even and identify the resonance marked with a red dot with the transition between the ground state of N electrons and the ground state of $N + 1$ electrons. Because the bias window is small and positive, no other states are involved and electron flow occurs from the drain electrode to the source electrode. This is schematically depicted in the tunnel cartoon on the lower right. The red arrow indicates the rate limiting step, which is the tunneling of an electron through the drain barrier onto a completely empty orbit of the N-electron island. At $B = 0$ T this level is spin degenerate and hence either a spin-up electron or a spin-down electron may tunnel onto this state. When this happens, the island quickly returns to the N electron ground state by tunneling through the small barrier to the source electrode. If a magnetic field lifts the spin-degeneracy, the threshold voltage at which the "tunneling on" process occurs is different for spin-up electrons and spin-down electrons, and hence this resonance is expected to split. This is exactly what happens in the 8 T data. The splitting allows one to extract

a g-factor for this particular quantum state ($g = 0.38$).

Middle panel

The addition of one electron to the quantum dot by increasing the gate voltage brings us to the middle panel. Here, the red dot marks the resonance associated with the same quantum level as in the lower panel, but the conductance peak is much weaker. This is expected, because now one spin-state is already occupied and only an electron with opposite spin can tunnel through the drain barrier (see tunnel cartoon to the right). In the limit where the drain barrier is much larger than the source barrier, the resulting current is half of the previous current. Also, note that the resonance does not Zeeman split in the magnetic field, confirming that $N + 1$ is an odd integer.

Top panel

Upon addition of another electron the same quantum state is fully occupied and does not show up if electrons are tunneled onto the particle. However, starting with $N + 3$ electrons, if electrons are tunneled off the particle through the drain barrier (i.e. bias is negative) a resonance appears, as indicated by the red dot.

g-factors of other evaporated gold nanoparticles

Table 2.4 lists g-factors obtained for different devices using evaporated gold nanoparticles. The average g-factors are larger than g-factors reported previously for gold nanoparticles by J. Petta [2] and D. Davidovic [3] (those g-factors are listed in Table 3.2). At the end of Section 3.3 we argue that this is probably due to some difference in the particles' microscopic properties, and not related to the different level spacings.

Table 2.4: g-factors obtained from 3 other evaporated gold nanoparticles. $< g >$ is the mean g-factor and $< \delta >$ is the mean level spacing.

	g-factors	$< g >$	$< \delta >$ [meV]
Au#1	0.73	1.17	1.4
	0.74		
	1.44		
	1.77		
Au#2	1.27	1.59	1.0
	1.37		
	1.62		
	1.81		
	1.87		
Au#3	0.58	0.75	1.0
	0.60		
	1.07		

2.5 Remarks on this fabrication technique

Evaporated copper and gold-iron particles

Even though the creation of nanoparticles by simply evaporating a small quantity of metal is very simple in principle, it is limited to materials which ball up due to surface tension, and their size and shape are not well controlled. Even if a material balls up, it may not be suitable if the particle density is too high.

In order to established whether devices based on evaporated Cu or Pd particles are feasible (Pd is expected to show electron-electron interactions, and Cu(Mn) is interesting for Kondo physics) we performed percolation tests for Cu and Pd. For Cu we found that the resistance of break junctions (measured while depositing metal) dropped from $> 1M\Omega$ to less than 2 kΩ between 5 and 6 Å. Pd however "shorted" the break junctions already after 25 Å, indicating that evaporated Cu nanoparticle devices may work, but that other techniques like chemical synthesis may be necessary to study Pd nanoparticles.

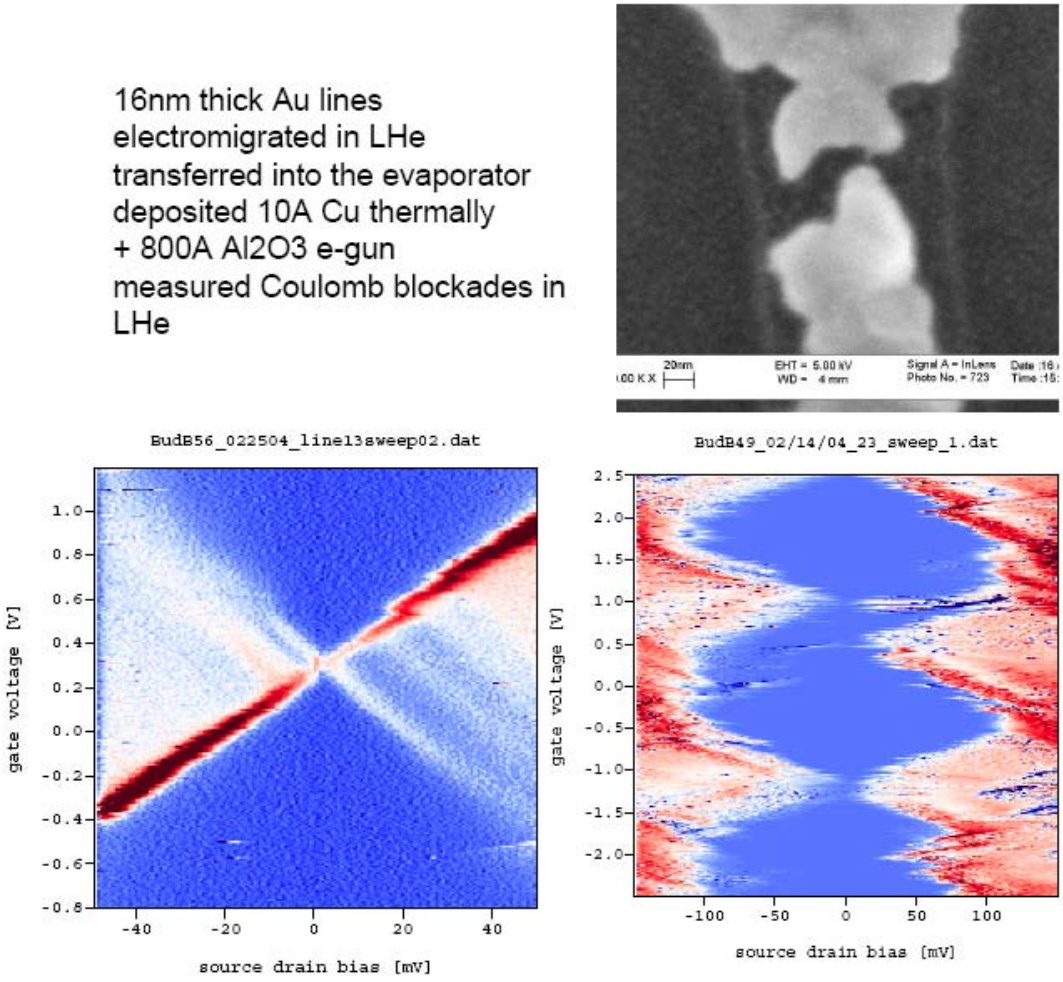


Figure 2.13: Copper nanoparticle devices fabricated by evaporation of 10 Å Cu capped with 800 Å of aluminum oxide. Cu nanoparticles are barely visible on the SEM. Spectral features at 4.2 K are consistent with a small nanoparticle size, but in most devices the fuzzy diamonds suggests that current does not flow through an isolated nanoparticle.

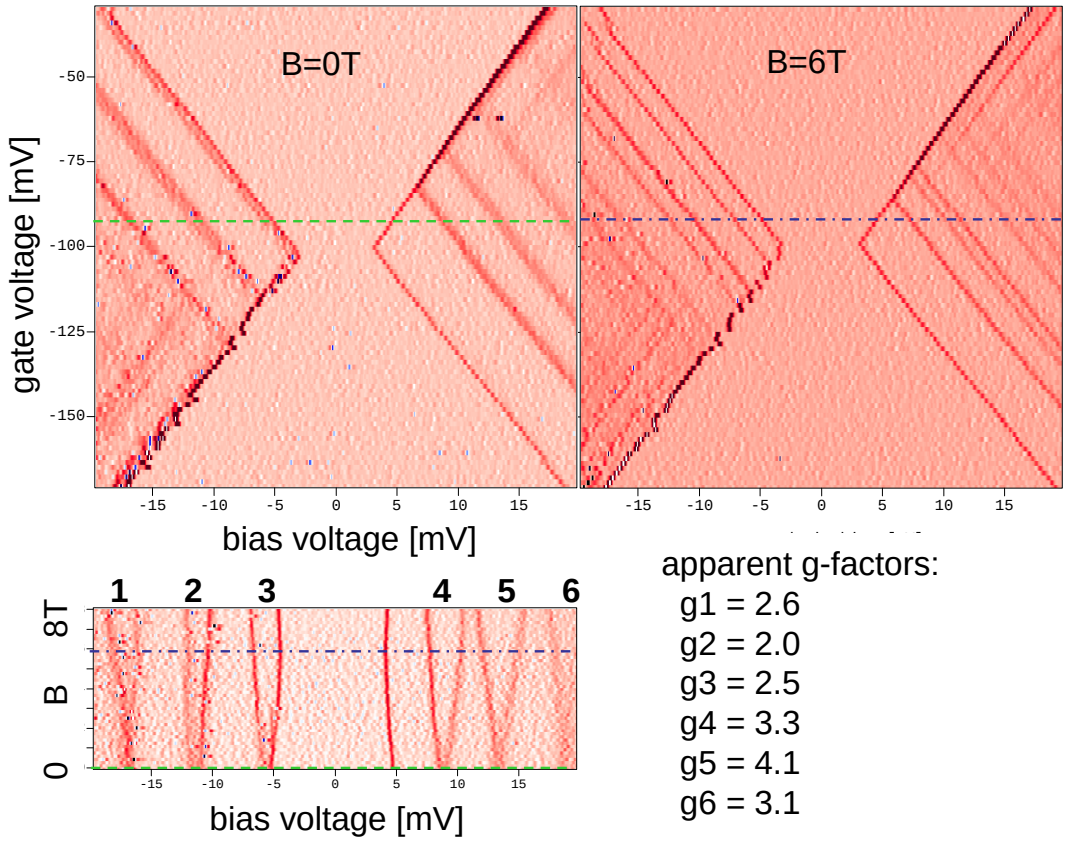


Figure 2.14: Spectrum of a device resulting from co-evaporation of gold and iron, with apparent g factors larger than two.

After further optimizing the amount of deposited Cu, we successfully made SETs which showed promising features when measured at 4.2 K (see Fig. 2.13). Optimization of these devices is more difficult than gold, because Cu particles don't show up easily on the SEM, and a capping layer of aluminum oxide or silicon oxide is necessary to prevent oxidation. We have cooled down multiple chips with Cu-particle devices in a Kelvinox dilution unit, without obtaining clean spectra. Instead we often obtained messy Coulomb diamonds which indicated that the copper particles where too small and too close to each other to get electrical transport through a single particle only.

The other approach is to stick with gold, but to modify its properties by in-

cluding a small quantity of ferromagnetic impurities. If the amount of doping is small, fabrication may work without any modifications to the pure gold procedure, while providing an electronically very different system to study. AuFe is interesting because in principle one should be able to make solid solutions from the dilute ferromagnetic regime (homogeneous solid solutions up to 15% percent iron by weight [25]) to the spin-glass regime [33], all the way down to the limit of the Kondo box: a single spin in one nanoparticle [26]. Au(Fe) is a classic Kondo alloy which has been well studied in bulk and thin films [31, 32], and therefore is an attractive canditate to experimentally address the Kondo effect in systems of reduced dimensionality [24].

We have fabricated Au(Fe) samples in two different ways: co-evaporation of Au and Fe in separate crucibles, and alloying of gold and iron in a quartz tube/furnace to be used as an evaporation source. In either case it is difficult to control the amount of iron impurities which ends up in the nanoparticles. XPS spectra obtained from continuous films deposited from the same evaporation sources indicated that the gold contained a few atomic percent of iron. The yield of good devices however was much lower than from pure gold evaporations, indicating that the presence of iron influences the nucleation or growth of nanoparticles. In Fig. 2.14 we show the spectrum of a device with g-factors apparently larger than 2. The g-factors listed where obtained by using the slopes of the resonances to convert their splittings into energy, ignoring the fact that the Coulomb blockade does not vanish at the degeneracy point. However, this gap may indicate that current flow occurs through a network of nanoparticles, and hence the actual conversion factor may be smaller. For a good discussion of Coulomb effects in a network of nanoparticles see [18] (however, they consider the regime k_BT larger than the level spacing).

Motivation for the following chapter

Using electromigrated break junctions on top of oxidized aluminum gates, we have demonstrated that high quality data on evaporated metal nanoparticles can be obtained, with a coupling to the gate electrode large enough to change the total number of electrons on the nanoparticle. However, the control over the nanoparticle's size and shape has not improved because we relied on the same mechanism to form these particles: self assembly due to surface tension while depositing a small quantity of metal in an evaporator. The average size of the nanoparticles inside the gap is different from outside the gap (Fig. 2.8), and some materials which because of their electronic structure would be interesting don't form nanoparticles of appropriate size, such as Pd.

The motivation of chapter 3 is to show that it is possible to incorporate chemically synthesized nanoparticles into a similar geometry as in this chapter. However, the use of aqueous solutions with low pH during processing damages the oxidized aluminum gate electrodes.

The way around aluminum gates is to use silicon backgates. This requires additional steps to grow a high quality thin oxides at the beginning of the fabrication and may allow high enough gate coupling. Fabrication of such substrates is described in [27]. The data presented in chapter 3 was obtaining with backgates separated by 30nm of thermal oxide from the source/drain electrodes. Leakage to the gate was never an issue, and the limiting factor to gating was loss of stability, not breakdown as it was in aluminum gates. We obtained discrete spectra at gate voltages as high as 14 Volts. Recently, we fabricated gate oxides 20nm thick, and believe that there should not be any problems reducing the gate oxide to 10-15 nm.

REFERENCES

[1] J. R. Petta and D. C. Ralph, Phys. Rev. Lett. **87**, 266801 (2001).

[2] D. Davidović and M. Tinkham, Phys. Rev. Lett. **83**, 1644 (1999).

[3] A. B. Zorin*et al.*, Phys. Rev. B, **53**, 13682 (1996).

[4] M. Zaffalon and B. J. van Wees, Phys. Rev. Lett. **91**, 186601 (2003).

[5] K. Segall*et al.*, Appl. Phys. Lett., **83**,4859 (2002).

[6] L. L. A. Adams, B. W. Lang, Yu Chen, and A. M. Goldman, Phys. Rev. B **75**, 205107 (2007).

[7] C. T. Black, Ph.D. thesis, Harvard University, 1996.

[8] D. C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **74**, 3241 (1995).

[9] D. C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **78**, 4087 (1997).

[10] L. L. A. Adams, B. W. Lang, and A. M. Goldman, Phys. Rev. Lett. **95**, 146804 (2005).

[11] L. L. A. Adams and A. M. Goldman, Rev. Sci. Instrum. **76**, 063907 (2005).

[12] J. G. A. Dubois *et al.*, Europhys. Lett. **33**, 279 (1996).

[13] J. Park, A. N. Pasupathy, *et al.*, Nature **417** 722 (2002)

[14] H. Park *et al.*, Appl. Phys. Lett. **75**, 301 (1999).

[15] A. A. Houck *et al.*, Nano Lett. **5**, 1685 (2005).

[16] T. Taychatanapat *et al.*, Nano Lett. **7**, 652 (2007).

[17] J. C. Sankey *et al.*, Phys. Rev. Lett. **96**, 227601 (2006).

[18] V. A. Danilov *et al.*, Phys. Rev. B **65**, 125312 (2002).

[19] K. I. Bolotin, F. Kuemmeth, A. N. Pasupathy, and D. C. Ralph, Appl. Phys. Lett. **84**, 3154 (2004).

[20] J. E. Grose, *Electron transport in single-molecule transistors based on high-spin molecules*, Ph.D. thesis, Cornell University, August 2007.

[21] A. N. Pasupathy, *Electron transport in molecular transistors*, Ph.D. thesis, Cornell University, August 2004.

[22] A. R. Champagne, *Mechanically adjustable single-molecule transistors and stencil mask nanofabrication of high-resolution scanning probes*, Ph.D. thesis, August 2005.

[23] D. A. Gorokhov and P. W. Brouwer, Phys. Rev. Lett. **91**, 186602 (2003).

[24] M. A. Blachly *et al.*, Phys. Rev. B **51**, 12537 (1995).

[25] J. W. Shih, Phys. Rev. **38**, 2051 (1931).

[26] W. B. Thimm *et al.*, Phys. Rev. Lett. **82**, 2143 (1999).

[27] E. C. Stoner, Proc. R. Soc. London, Ser. A **154**, 656 (1936).

[28] J. von Delft and D. C. Ralph, Phys. Rep. **345**, 61 (2001).

[29] A. K. Hüttel *et al.*, Europhys. Lett. **62**, 712 (2003).

[30] D. Weinmann *et al.*, Phys. Rev. Lett. **74**, 984 (1995).

[31] K. R. Kane *et al.*, Phys. Rev. B **51**, 945 (1995).

[32] G. Chen and N. Giordano, Phys. Rev. Lett. **66**, 209 (1991).

[33] U. Larsen, Phys. Rev. B **18**, 5014 (1978).

[34] A. H. MacDonald *et al.*, Phys. Rev. B **25**, 713 (1982).

Chapter 3

Spin-orbit interaction in chemically synthesized gold nanoparticles

3.1 Summary

We demonstrate that chemically synthesized metal nanoparticles can be incorporated into a single-electron-transistor geometry by means of an organic tunnel barrier. The motivation of this approach is to obtain discrete spectra of individual nanoparticles which are - by virtue of chemical synthesis - well defined in their composition, size and shape. We demonstrate this idea by measuring the discrete energy levels of gold nanoparticles varying in diameter from 5 nm to 15 nm. The average level spacings we measure correlate with the particles' sizes. Due to the random nature of an electron's wavefunction, we find that the tunnel rates are different for each quantum state and also vary between source and drain. However, the average tunnel coupling to source and drain can be very symmetric, indicating that high quality, homogeneous tunnel barriers are possible using organic monolayers.

We analyze the the magnetic field dependence of the spectra and find fluctuating g-factors and avoided level crossings, due to spin-orbit interaction and mesoscopic fluctuations [9].

3.2 Fabrication

Synthesis

Various techniques exist to synthesize monodisperse metallic nanoparticles (i.e. size variation $< 10\%$), and several parameters exist to also modify the shape of the particles (reaction rate, choice of ligands/solvent, excitation of quadrupole plasmons) [24–26]. Following reference [23] we synthesized almost spherical gold nanoparticles from an aqueous solution of tetrachloraurate at boil, controlling the size of the particles by the amount of reducing agent added (a mixture of tannic acid and trisodium citrate).

Fig. 3.1 demonstrates that with this technique gold nanoparticles in the size range of interest (5–15 nm) can be chemically synthesized in almost spherical shape and homogeneous size. The expected mean level spacing for diameters of 5, 10 and 15 nm are indicated on each SEM. The I-V curves obtained from devices using different sized colloid reflect the size dependence of the mean level spacings. We have also synthesized nanoparticles larger and smaller than the ones shown, but the quantum states of larger particles are too dense to be resolved, whereas nanoparticles smaller than 5 nm tended to self assemble with considerably lower density.

Colloidal suspensions of a few sizes (5 nm, 10 nm, 15 nm gold particles) could also be obtained from Ted Pella and Sigma, produced by a similar reduction method. These commercially available colloids may additionally contain potassium carbonate to increase the pH to $\sim$6 as well as sodium azide as preservative. In either case, due to Cl- ions partially replacing the citrate ligands, the gold particles carry a net negative charge which prevents aggregation and is used to electrostatically control the self assembly of particles onto a substrate [28].

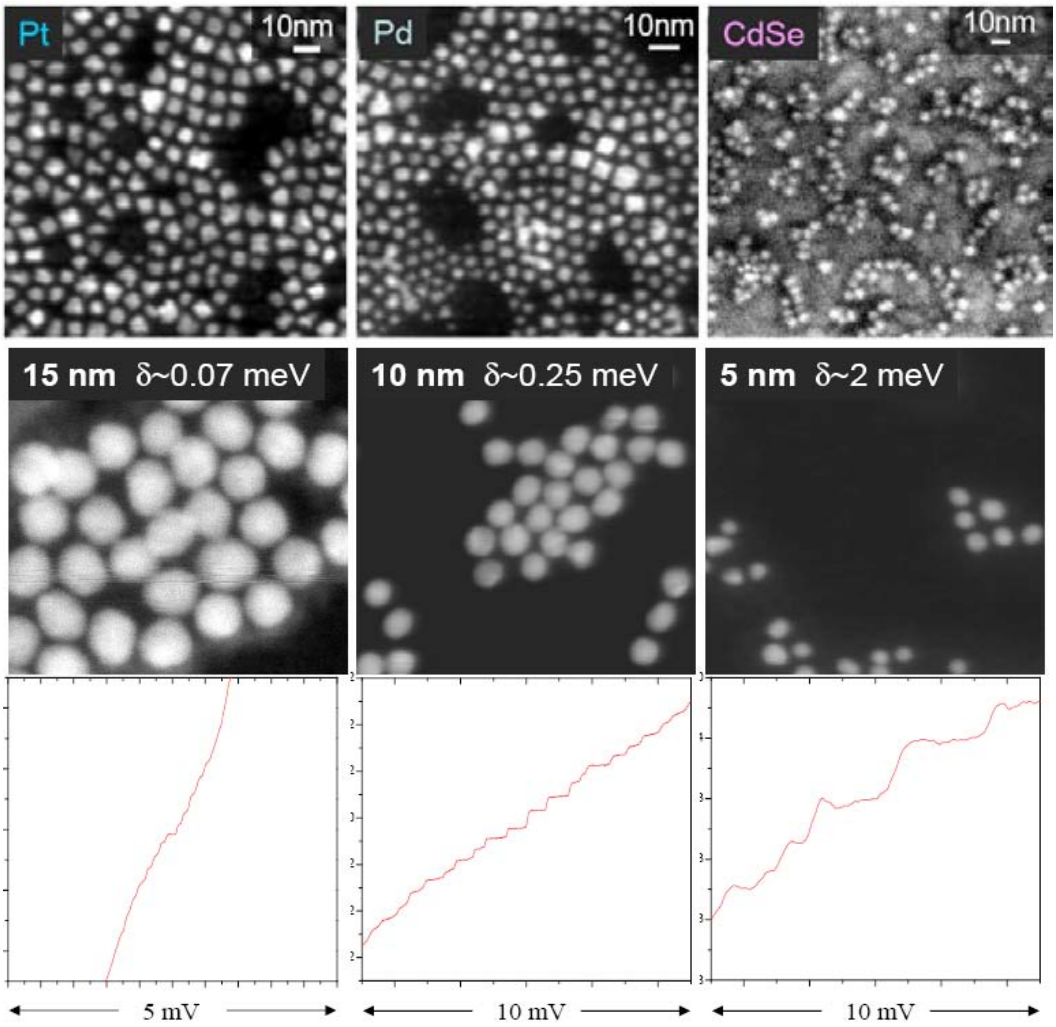


Figure 3.1: Chemical synthesis offers lots of different materials on the nanoscale. In particular gold nanoparticles of well defined size can be obtained. The size in turn determines the mean level spacing in a nanoparticle, as apparent from the steps in typical $I-V$ curves (the current is plotted on the vertical axis using arbitrary units).

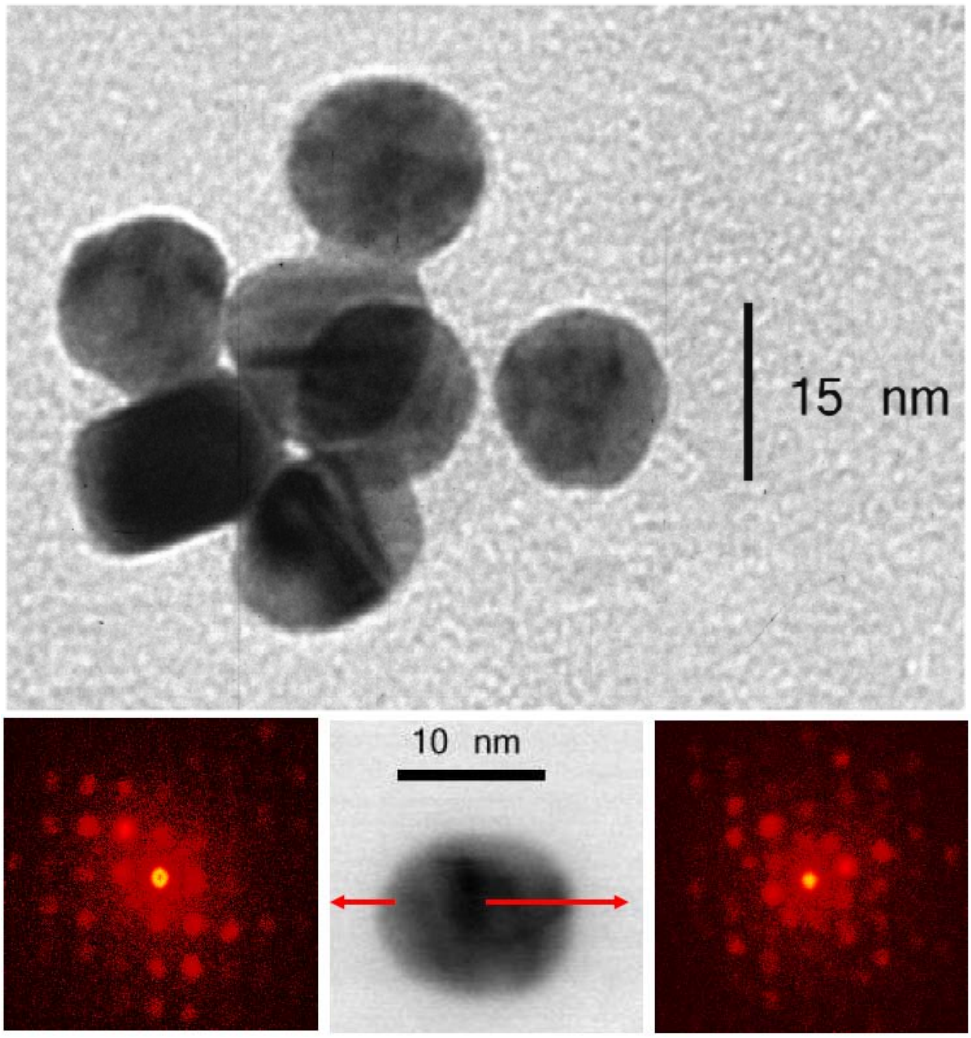


Figure 3.2: Inspection by HRSTEM reveals that nanoparticles are not perfectly spherical symmetric. Diffraction images obtained from two different locations of one particle indicates that nanoparticles are crystalline but may contain grain boundaries.

Assembly

To incorporate these gold particles of well defined size into a SET-geometry, gold break junctions are fabricated similar to the ones described in Chapter 2. The main difference is that we first grow 30 nm of silicon oxide on a degenerately doped silicon substrate which will serve as a backgate [27], instead of using aluminum gate electrodes. Electron lithography and electron-beam evaporation is used to define 30 nm thick gold nanowires - each with a constriction 50 nm wide - which are connected to thicker gold pads suitable for wirebonding by additional layers for photo lithography. After cleaning each chip with acetone, IPA and an oxygen plasma, the nanowires typically have a resistance of 60-100 Ohms[1]. Electromigration is employed to create a nm-sized gap within each of the 36 wires per chip, by slowly increasing the bias until the wire breaks (this typically occurs at 6-10 mA). After this step, some of the wires show a residual resistance of a few hundred MΩ or higher. The chip is then inserted into a 0.05% (w/v) aqueous solution of [(Aminoethylamino)propyl]trimethoxysilane ("APTS", Fluka#06668) for 10 minutes, rinsed in DI water and then baked for 30 minutes at 120℃, after which all lines typcially show no conductance (>50 GΩ). This produces a homogeneous organic monolayer which serves two purposes: It enables the electrostatic assembly of gold colloid on the substrate (see Fig. 3.3) and constitutes the tunnel barrier in the finished device.

More precisely, the silane end of APTS preferentially attaches to the silicon oxide (and gold) surface of the substrate, and a tunable positive surface charge can be created by (partial) protonation of the amino group [28]. This is achieved by applying an aqueous suspension of gold colloid whose pH has been lowered to

[1]Cleaned chips are stored in ethanol if not used immediately.

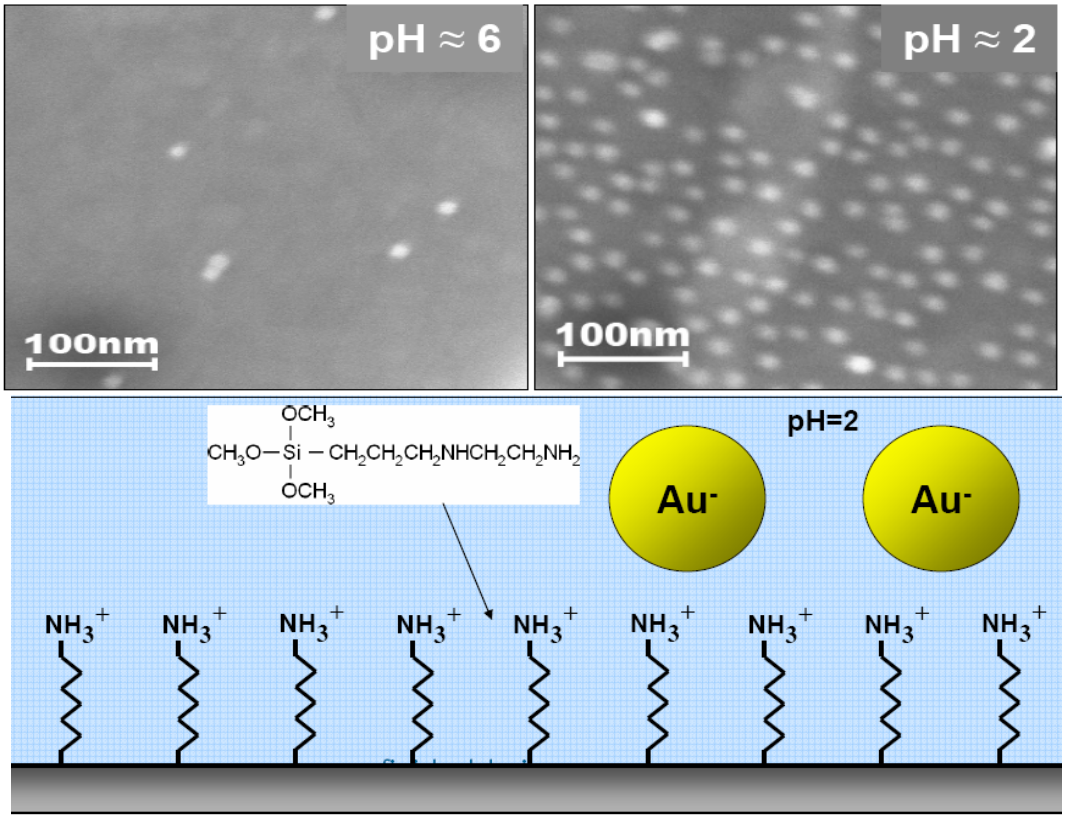


Figure 3.3: Top: The density of NPs which assemble on the APTS functionalized substrate depend on the pH of the NP solution: at low pH, the amino group of the APTS gets partially protonated (cartoon at the bottom) and therefore attracts the negatively charged NPs.

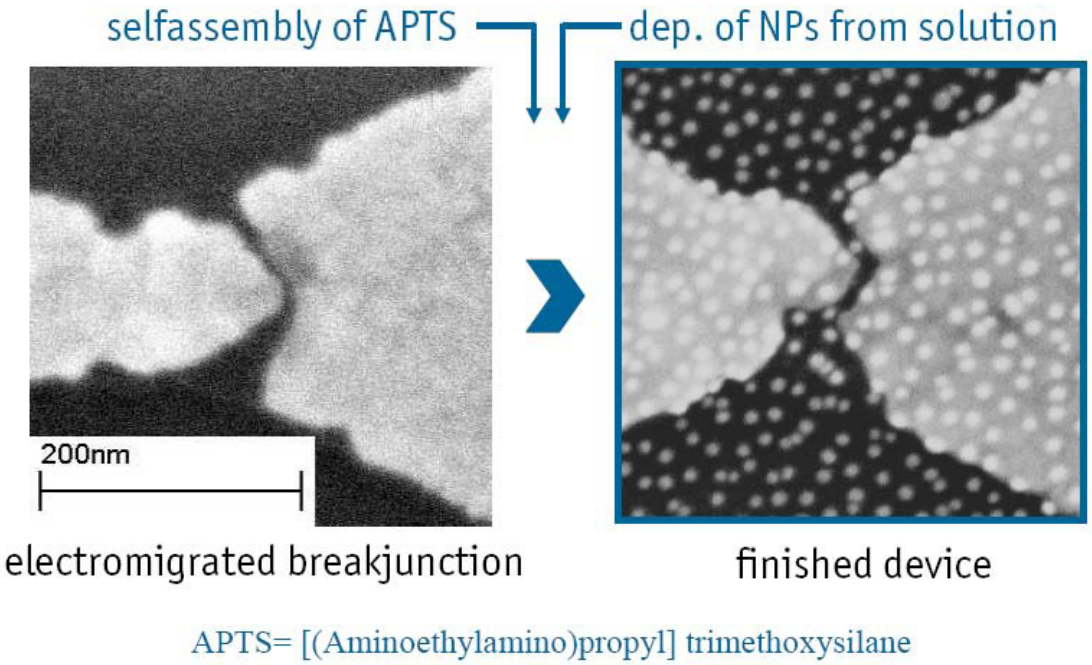


Figure 3.4: After electromigration and self assembly of APTS the chip is soaked in an acidic solution of gold nanoparticles which self assemble. After drying the chip the conductance of each break junction is tested to find those with nanoparticles inside the gap.

$\sim$2.5 by addition of citric acid. Note that pH$\sim$2.5 is below the $pK_{1/2}$ value of APTS ($\sim$5.3) but above the isoelectric point of the gold particles ($\sim$2) [29], hence at that pH a positive surface charge on the organic monolayer exists due to (partial) protonation of the amino groups while the gold particles are still negatively charged. We typically keep the chip in the acidic colloid solution more than 12 hours, during which self assembly of the gold particles onto the organic monolayer occurs (Fig. 3.4).

After rinsing the chip in deionized water for a few seconds, the room-temperature conductance of each electromigrated wire is briefly measured with a probe station up to a bias of $\sim$100 mV. Those lines which contain particles in their gaps (Fig. 1a) typically show linear I-V characteristics ($\sim$1 $G\Omega$) at room temperature due to

thermally lifted Coulomb blockade, and are wirebonded in preparation for tunnel spectroscopy at low temperatures.

Transport measurements

To measure the discrete energy spectra of the particles bridging the gaps, the same techniques as described in Chapter 2 are used. The wirebonded device is cooled down in a dilution refrigerator (<100 mK) and the DC tunnel current I is recorded as a function of the applied bias V and the potential of the backgate V_G. At small bias voltages, the tunnel current is suppressed due to Coulomb blockade unless the gate voltage is tuned such that two charge states of the particle are degenerate.

Fig. 3.5 shows that several such degeneracy points can be accessed despite the relatively thick gate oxide of 30 nm. From the geometry of the "Coulomb blockade diamonds" we extract a charging energy of 30 meV and capacitances of source, drain and gate to be $C_S \approx C_D \approx$1.3 aF and $C_G \approx$0.03 aF. The small gate capacitance is due to the rather thick gate oxide of 30 nm, and is expected to increase if a thinner gate oxide is chosen. In the low bias neighborhood of each degeneracy point, sequential tunneling occurs and the particle's discrete energy levels reveal themselves as conductance peaks when plotting dI/dV as a function of the applied bias. For this particular device monodisperse gold colloid of average diameter 9.1 nm was used (as specified by Sigma), with an estimated level spacing between spin-degenerate electron-in-a-box states of $\delta = 0.32$ meV (using Eq. 1.1). The experimental value we obtain from Fig. 3.5 is 0.31 meV, in good agreement with this estimate.

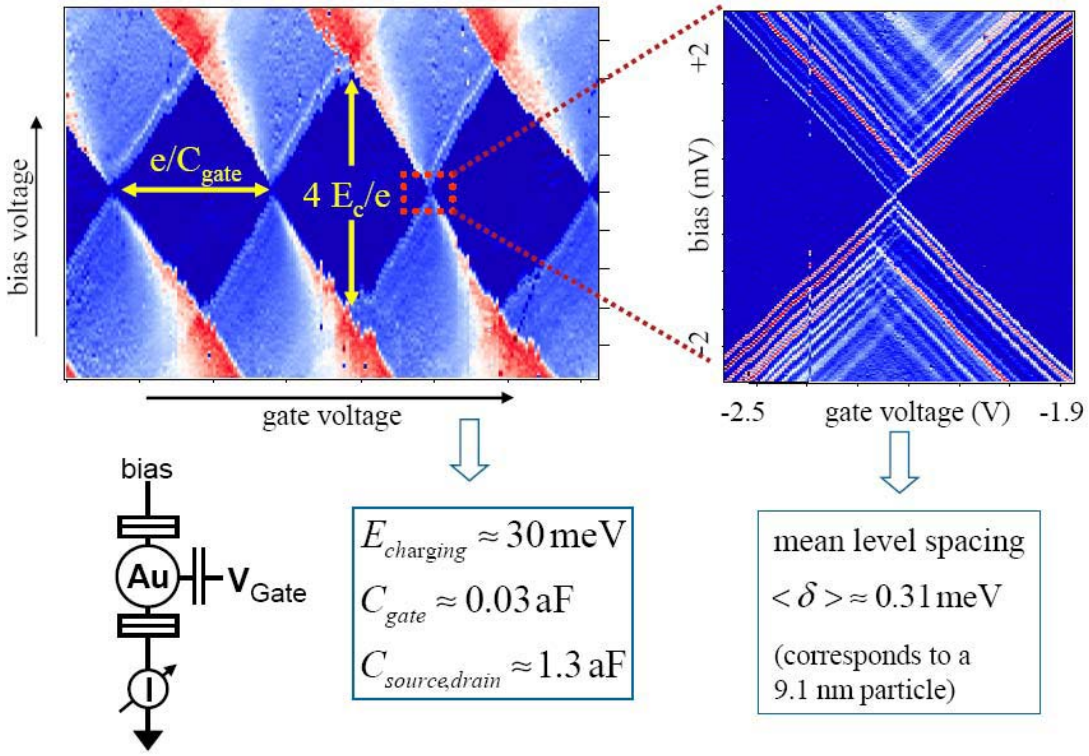


Figure 3.5: Differential conductance measured in a dilution refrigerator. The gate coupling is high enough to access several charge states, and the discrete level spectrum is measured near a degeneracy point. Note that the level spectrum appears twice: once as a set of lines with positive slope, and once with negative slope, indicating that the tunnel barrier on the source and drain side are comparable.

Tunnel coupling

Note that the spectrum in Fig. 3.5 is qualitatively different from spectra obtained from evaporated gold nanoparticles (for example Fig. 2.12), in that each transition shows up twice, namely as a conductance line with positive as well as negative slope with respect to the gate voltage. This indicates that the overall tunnel barrier of the source and drain are of equal strength. This symmetry is remarkable: unlike the capacitive coupling, the tunnel current is exponentially sensitive to the barrier thickness, suggesting that the monolayer of APTS constitutes a homogeneous tunnel barrier.

A more detailed look at the current flow via individual quantum states at four different bias regimes (as marked in the inset of Fig. 3.6) reveals that the tunnel matrix element is different for each quantum state. This is not surprising since near the Fermi energy each wavefunction is spatially highly oscillatory (with typical period λ_F) and orthogonal to all other wavefunctions, hence having its distinct overlap with the continuum on the source (drain) side of the tunnel barrier. If there were pointlike tunneling between the two mutually nearest points of the particle and electrode only, one would expect a Porter-Thomas distribution of conductances among many quantum states (cf. the low-temperature STM experiments on Pb nanoparticles reported in [5]). For a particle-electrode topology which is smooth on the length scale of the Fermi wavelength, the tunnel current will be a convolution between the spatially varying wavefunction and all contributing tunnel trajectories. Fig. 3.6 indicates that the tunnel matrix element of a quantum state (for example the one marked with *) to the source (red) is in general different from the tunnel matrix element to the drain (cyan), and the reverse is true for the next higher quantum state (marked with $\circ$). This reflects the variation in a wavefunction's

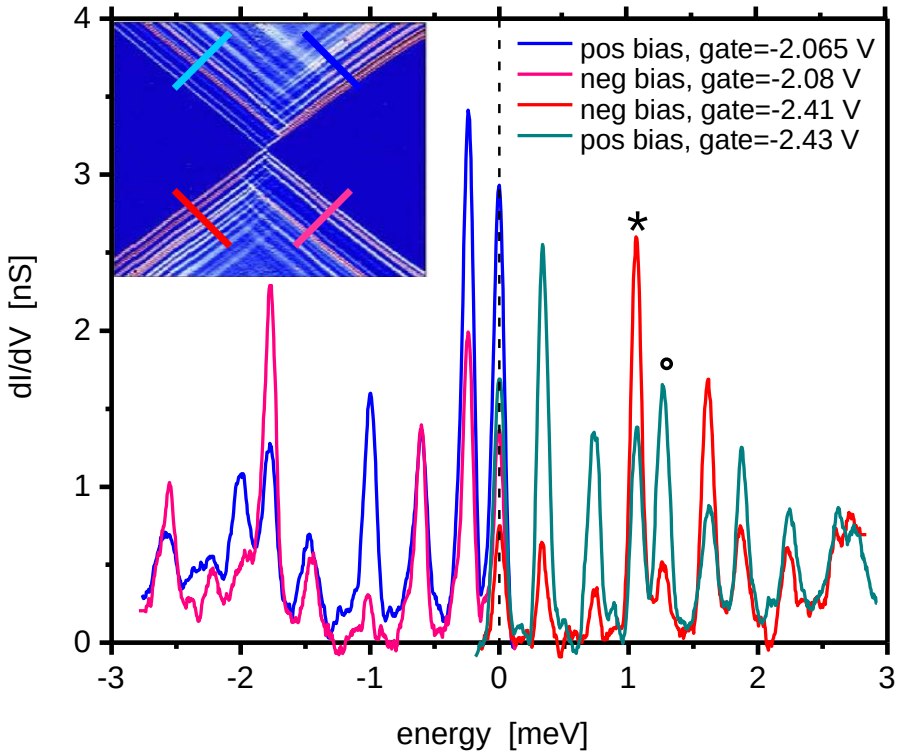


Figure 3.6: By plotting line traces from 4 different cuts as indicated in the inset (same sample as in Fig. 3.5) the tunnel couplings of occupied and unoccupied states can be compared. The red and blue traces reveal the tunnel couplings to the source electrode, whereas the cyan and pink traces reveal those to the drain electrode. For example, the quantum level marked with * has larger tunnel coupling to the source (red) than to the drain (cyan), while the opposite is true for the next quantum level marked with ∘. This reflects the spatial fluctuations within different wavefunctions.

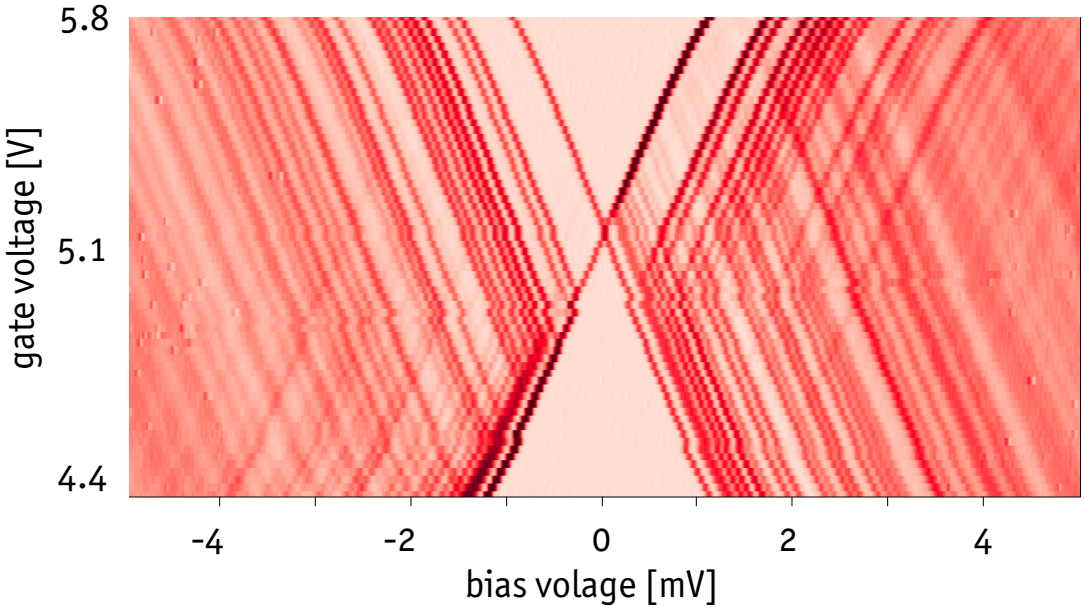


Figure 3.7: In this sample, more than 40 resonances are resolved, all without any lock-in techniques. This opens up the possibility to test the statistics within a single particle.

amplitude between the tunneling location on the source side and the drain side of the particle.

Gate coupling

Figure 3.5 indicates that the coupling to the gate electrode is high enough to access several charge states. Because the gate oxide was chosen rather thick (30 nm) the capacitance to the gate electrode is smaller than achieved in nanoparticles evaporated onto an oxidized aluminum gate. In this chapter we shall not make use of changing the total number of electrons, except for stating that electron-electron interactions are found to be weak in chemically synthesized gold nanoparticles,

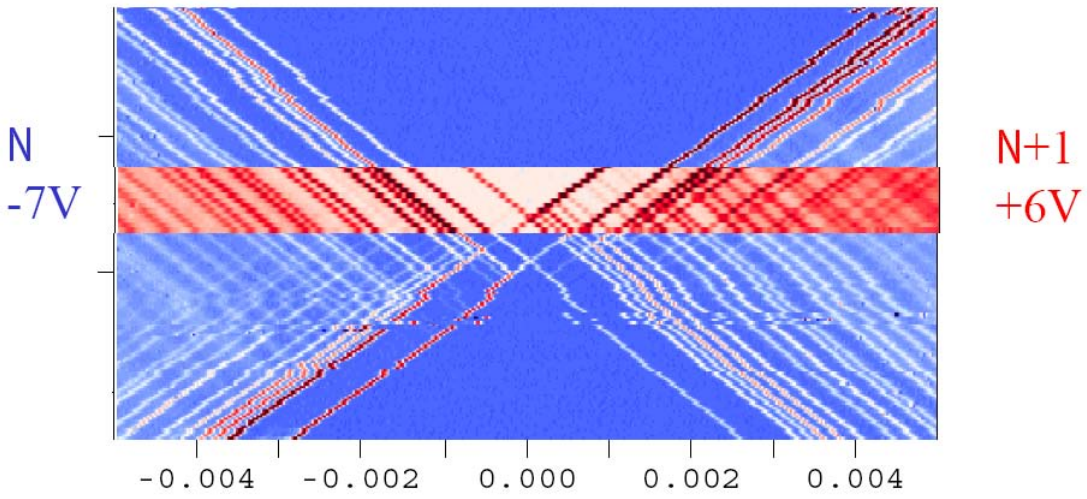


Figure 3.8: The excitation spectra associated with N and N+1 electrons are superimposed, with a slightly different color scale and a slight stretch of the gate scale to account for small ($\sim 10\%$) changes in the capacitance ratios. Note the match between N and N+1 excited states, indicating that electron-electron interactions are weak.

just as we found for evaporated nanoparticles.

This is demonstrated rather nicely in Figure 3.8, where two excitation spectra of the same particle, measured at gate voltages 13 Volts (!) apart, are superimposed onto each other. To contrast the two data sets, a different color scale has been used. The gate scale of the N+1 excitation spectrum has been stretched by approximately 10% to compensate for a small change in the capacitance ratios. Then the electronic levels match quite well. We also learn from this sample that even though a large gate voltage had to be applied, the noise level remained low enough to clearly resolve states. Hence the compatibility concerns associated with aluminum gates in a wet electrochemical environment (as needed for trapping colloid) may be avoided completely by using silicon back gates only, preferentially with a gate oxide thinner

than 30 nm to improve gate coupling.

Asymmetric tunnel coupling

Not all devices show characteristics of a symmetric tunnel barrier like the data shown in Fig. 3.5 and Fig. 3.7. This can have several reasons. First, if a nanoparticle is trapped in a gap slightly larger than its diameter it is likely to sit closer to one than to the other. One may argue that those cases rarely lead to a tunnel conductance at all. More likely a nanoparticle gets pulled into a gap by electrostatic forces or capillary forces when drying until it can't go further into the gap for mechanical reasons. In this case an incomplete layer of APTS may give rise to asymmetric tunnel coupling. We have not investigated whether the quality of the monolayer plays a role (i.e., does the APTS solution age over time? Did the solution get contaminated unintentionally by immersing chips?) for obtaining symmetric tunnel barriers, but two comments should be made: The highest quality monolayers of APTS are obtained on oxide surfaces not by self assembly from solution, but by vapor deposition in well controlled conditions [30]. The silane group is intended to bind to the oxide, and hence one should not expect to get perfect assembly on our gold electrodes. Thiol groups could be a good choice for gold, but we had good success with APTS. Second, a symmetric tunnel barrier may be regarded as a disadvantage, because twice as many conductance peaks arise from one IV curve with the same thermal broadening of each peak. Hence resolution is limited. Furthermore, when taking a magnetic field sweep at a constant gate voltage, it is more difficult to keep track of which resonances belong to a threshold associated with the electrochemical potential of the source, and which ones belong to the drain electrodes (each case is associated with a different conversion factor

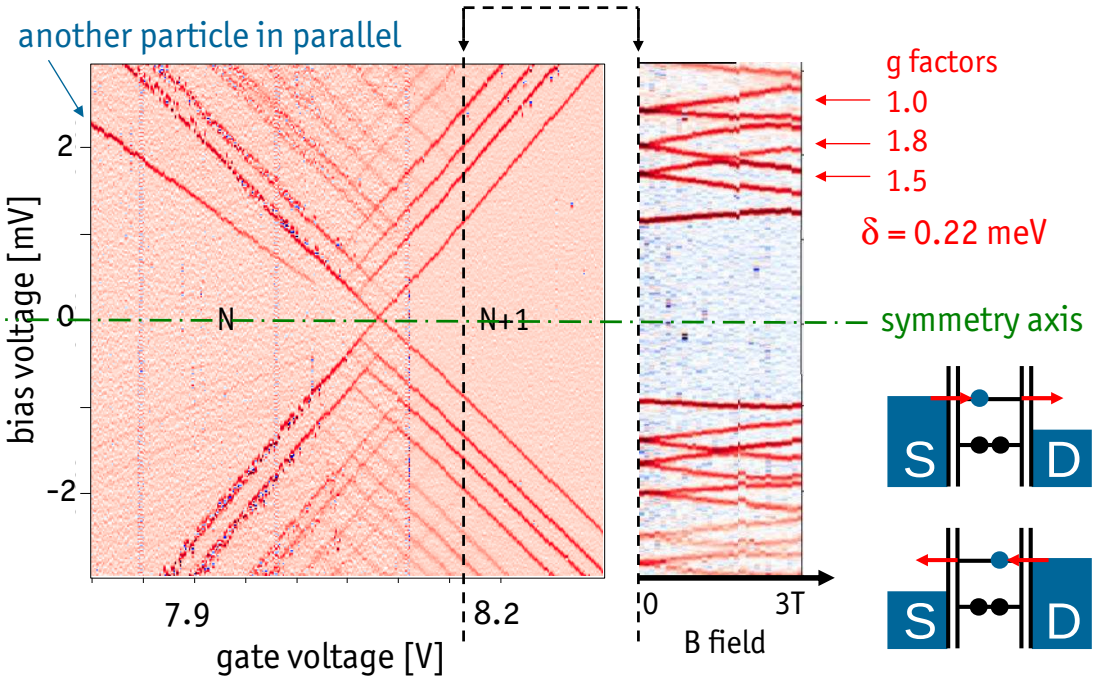


Figure 3.9: Symmetric tunnel coupling: reversing bias gives same current, and hence the spectrum looks symmetric about the zero bias axis (up to a distortion from the different capacitive coupling). For convenience, the magnetic field dependence is measured away from the degeneracy point to avoid confusion with the second set of resonances. Due to symmetry, each level and g-factor shows up twice.

to energy).

In such cases it helps to measure the magnetic field dependence at a gate voltage somewhat detuned from the degeneracy point. In Fig. 3.9 g-factors for the occupied states of a symmetrically coupled nanoparticle could be extracted reliably. However, to measure the Zeeman splitting of the unoccupied states a cut at lower V_G would have been needed. At that side of the degeneracy point the sample was noisy, possible due to a second nanoparticle nearby (see the faint set of lines with a different slope at the left).

3.3 Spin-orbit interaction in chemically synthesized gold nanoparticles

Absence of shell structure in colloidal metal nanoparticles

Shell structure occurs in the electronic spectra of many different systems, due to symmetries of the structure in real space. One example is the shell structure of carbon nanotubes expected from graphene's honeycomb lattice with two inequivalent (yet symmetrically related) atoms per unit cell. Another example is the rotational symmetry in atomic potentials which gives rise to $2(2l+1)$-fold degenerate shells[2] which can be labeled by n and l, where n is the number of nodes in the radial Schroedinger equation and l the orbital angular momentum number. Quantum dots can be considered as artificial atoms, and in fact electronic shell structure has been observed in semiconducting quantum dots where the confining potential has cylindrical symmetry (even though the host lattice doesn't[3]). It is a valid question to ask whether shell structure should be expected for the "artificial atoms" studied in this thesis, and the answer is different for the metallic quantum dots and the nanotube quantum dots.

Probability densities

The evaporated nanoparticles in Chapter 2 clearly have no spatial symmetry, leading to chaotic single-particle dynamics. The chemically synthesized gold nanoparticles discussed in this chapter appear spherical when imaged with a SEM, but show deviations from the spherical shape when inspected with a high resolu-

[2]For an infinitely high, spherical potential well one additionally finds a periodic bunching of shells known as "supershell" structure.

[3]this is because the "Fermi wavelength" of the confined electrons is much larger than the lattice

tion TEM. It is clear from Fig. 3.2 that the random surface fluctuations are larger than the Fermi wavelength in gold (0.5 nm) and hence remove the large orbital degeneracy ($\sim N^{1/3}$) which would otherwise result from the high symmetries of a spherical boundary condition. Consequently, the spectra of these nanoparticles differ between particles and are distributed according to a probability density. For example, if the level spacing in a metal nanoparticle was distributed purely random (with the constraint that the mean level spacing is $\delta \propto \text{Volume}^{-1}$), the probability density would read

$$P_n(\Delta) = \frac{1}{n!\delta}(\Delta/\delta)^n \exp(-\Delta/\delta) \tag{3.1}$$

This is the Poisson distribution and gives the probability density for finding an energy state with a separation Δ from a given one with precisely n levels in between. It is inappropriate because for $n = 0$

$$P_0(\Delta) = \frac{\exp^{-\Delta/\delta}}{\delta} \overset{\Delta\to 0}{\to} \frac{1}{\delta} \tag{3.2}$$

it predicts a constant probability density of finding two levels in the same particle arbitrarily close together. In a metal nanoparticle, levels repel each other. This is because in addition to the symmetries of the nanoparticles surface (or lack thereof) there are other fundamental symmetries which enter the Hamiltonian (such as spin rotation symmetry or time reversal symmetry) and determine the nature of the probability density. For example, in the absence of spin-orbit interaction and applied magnetic fields the electron Hamiltonian is invariant under space inversion and time reversal and it can be shown that the correct probability density decreases proportional to Δ for $\Delta \to 0$. The problem of random electronic levels in metallic nanoparticles has been treated successfully by the theory of random matrices.

Table 3.1: Energy level distributions depend on the magnitude of applied field ($\mu_B B$) and spin-orbit interaction ($\langle H_{SO} \rangle$) in comparison to the average level spacing δ. Adapted from [1].

Distribution	$P_n(\Delta \to 0)$	Magnetic field	Spin-orbit interaction
Poisson	$1/\delta$	large	small
orthogonal	Δ/δ^2	small	small
unitary	Δ^2/δ^3	large	large
symplectic	Δ^3/δ^4	small	large

Random Matrix theory

Briefly, random matrix theory is a sophisticated theory which can predict statistical properties of spin-orbit effects in various regimes of its strength. Specifically, it predicts different probability densities for the eigenvalues of a randomly generated matrix depending on whether the matrix is real, complex or symplectic. The corresponding transformations which diagonalize these matrices mathematically are orthogonal, unitary and symplectic matrices respectively, and hence the three basic classes of random matrices are referred to as "Gaussian Orthogonal Ensemble" (GOE), "Gaussian Unitary Ensemble" (GUE) and "Gaussian Symplectic Ensemble" (GSE). In the physical context of a metal nanoparticle, the transformation properties of the Hamiltonian depend on the relative strength of the spin-orbit interaction ($\langle H_{SO} \rangle$, which breaks spin-rotation symmetry) or the external magnetic field ($\mu_B B$, which breaks time reversal symmetry) as compared with the level spacing δ.

Table 3.1 indicates that level repulsion is strongest for a metal nanoparticle with strong spin orbit interaction and no external magnetic field applied. Random matrix theory is not limited to the limiting cases of Table 3.1 but can treat

crossovers between regimes as well [8, 9].

Comment on shell structure in metal clusters

Shell structure has been observed in metal nanoparticles much smaller than the ones used in our experiments[4]. In cluster physics, a shell can be associated with electronic shells as well as structural shells [11] — both of which can be detected by statistically analyzing the mass spectra of a suitable cluster source. Knight *et al.* [10] present mass spectra for clusters consisting of 4-100 sodium atoms. They explain the increased abundance for clusters of 8, 20, 40, 58 and 92 atoms by the stability of a cluster if an electronic shell is exactly full. The filling of electronic shells and subshells in sodium clusters can be observed up to 1500 atoms, and it was pointed out by Martin *et al.* [11] that the observed subshells may indicate the existence of a symmetry higher than spherical[5].

Shell structure associated with shells of atoms were observed for clusters consisting of 1500–22000 sodium atoms [11].

Previous work on Cu, Ag, Au

Random matrix theory can predict the level statistics resulting from any strength of spin-orbit interaction [8,9]. Experimentally, the dependence of g-factor distributions on the strength of SO can be studied by choosing metals with different atomic numbers. Jason Petta measured g-factor distributions in evaporated Cu, Ag, and Au nanoparticles and found that the mean g-factor decreases from $< g > \approx 1.2$

[4]Below 10000 atoms or so they are referred to as clusters rather than colloid.

[5]Experimentally, they found that $3n + l$ is an approximate energy quantum number for alkali clusters, similar to the role played by $n + l$ in the hydrogen atom (due to its $O(4)$ symmetry) and $2n + l$ in the spherical harmonic oscillator ($SU(3)$ symmetry).

Table 3.2: g-factors for evaporated gold nanoparticles from [2] and [3]. N is the number of resonances resolved, $< g >$ is the mean g-factor, σ is the experimental standard deviation, and δ is the mean level spacing.

Ref.	N	$< g >$	σ	$\delta[meV]$
[2] Au#1	7	0.12	0.06	0.10
Au#2	7	0.17	0.07	0.12
Au#3	5	0.45	0.27	0.27
[3] Au#4		0.28		1.0
Au#5		0.45		1.2

in Cu to $< g > \approx 0.25$ in Au [2]. This trend of g-factors is consistent with an increasing atomic number, but it should be noted that the details of scattering from impurities and surfaces determine the strength of spin-orbit scattering in a given grain. Evaporated Cu nanoparticles may be very different in terms of size, surface roughness and impurities than evaporated gold nanoparticles. Moreover, each nanoparticle made from the same material will be different, and indeed some variation of the mean g-factors among different samples was observed (Table 3.2).

Jason Petta managed to measure the g-factor distribution among several quantum states within a single nanoparticle. For Cu, Ag and Au he found good agreement with predictions of random matrix theory, but only if orbital contributions to the Zeeman splitting were neglected in the theory. However, for a ballistic nanoparticle and strong spin-orbit interaction there should be an orbital contribution to the g-factor of order 1 [12]. The small g-factors in Table 3.2 indicate that no orbital contribution is present, as if the particles were in the diffusive regime. This is a surprise, because TEM images of evaporated nanoparticles (on different substrates) indicate that they are crystalline. However, an assessment of the actual nanoparticles measured is difficult, as the nucleation process may be very different. This motivates the use of chemically synthesized nanoparticles which can be

characterized before trapping them into break junctions.

Size dependence of spin-orbit interaction?

Another motivation for better control of the particle's properties (in particular its size) arises from the apparent correlation in Table 3.2 between the mean level spacing and the mean g-factor. Consistent with this trend are the g-factors we obtained for evaporated gold nanoparticles with higher level spacing and higher g-factors (see Table 2.4). This brings up the question: Does the mean g-factor depend on the size of the nanoparticle? Another experiment claimed to observe suppression of spin-orbit interaction in highly disordered gold [13], so we may briefly argue why nm-sized particles fall into an interesting size regime in terms of the strength of spin-orbit scattering.

Whether spin-orbit interaction should be considered as weak or strong in a given nanoparticle depends on the ratio between the spin-orbit coupling and the level spacing [14]. Noting that in a particle of size L the mean level spacing δ may be expressed as

$$\delta \sim \frac{\nu_F \lambda_F^2}{L^3} \tag{3.3}$$

and assuming that scattering occurs mostly at the surface (ballistic nanoparticle)

$$\tau_{SO}^{-1} \sim \frac{\alpha \nu_F}{L} \tag{3.4}$$

(where α is the probability of flipping the spin during a momentum scattering event) Anaya *et al.* argued that there is a size below which spin orbit interaction becomes weak compared to the level spacing [13]. With $\alpha = 5 \times 10^{-3}$ for gold they estimate that the crossover occurs for $L \approx 7$ nm. Coincidently, this is exactly the size regime where discrete states are resolvable, and chemical synthesis possible.

To verify whether the difference between g-factors obtained for evaporated gold nanoparticles (such as in Fig. 2.12 and Table 3.2) originates from the particle's size or from the difference in fabrication, we focused on chemically synthesized nanoparticle with a similar level spacing as in Jason Petta's experiment. However, we have measured smaller sizes as well to verify that smaller particle's are consistent with the values we found for evaporated nanoparticles.

Results

As discussed earlier spin-orbit interaction modifies the discrete level spectrum of a randomly shaped metal nanoparticle in several ways: it modifies the level spacing statistics (enhanced level repulsion), reduces the average g-factor, and leads to avoided level crossings between different orbitals. Here we present data which supports all three signatures (Fig. 3.10 and 3.11).

Fig. 3.10 presents data of a chemically synthesized gold nanoparticle which is interesting for several reasons: First, the mean level spacing is δ=0.26 meV, and therefore allows direct comparison with an evaporated nanoparticle of equal size, but different fabrication technique (see sample Au#3 with δ= 0.27 in Table 3.2). Second, the data is of high enough quality such that many resonances can be resolved, allowing statistical analysis within a single particle. Third, the tunnel coupling happens to be very asymmetric, with a tunnel resistance on the source side much higher than on the drain side. Compared to symmetric devices, this guarantees well separated conductance peaks (as a function of bias voltage) which all belong to different quantum states. Therefore, we can measure the magnetic field dependence of occupied and unoccupied states simultaneously (Fig. 3.10B),

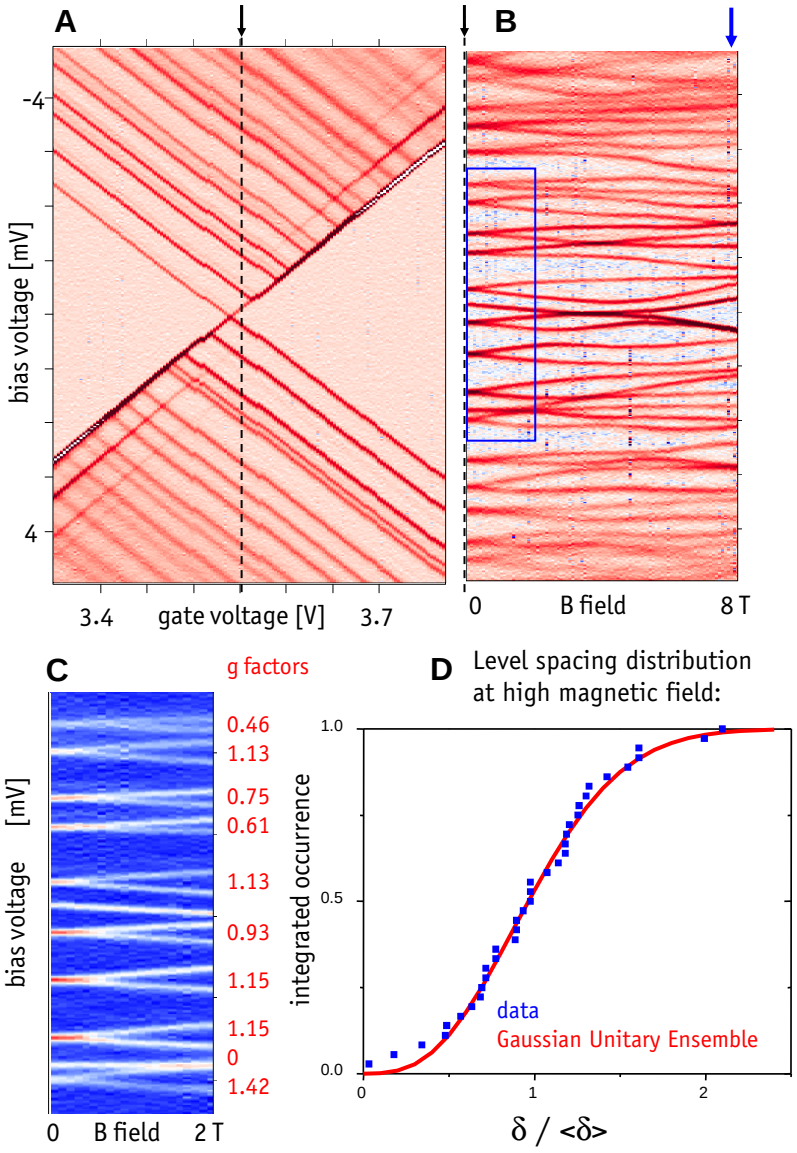


Figure 3.10: Top left: Excitation spectrum of an asymmetrically coupled gold nanoparticle at B=0 T. Top right: Magnetic field dependence measured at constant gate voltage near the degeneracy point. Bottom left: Linear Zeeman splitting at small fields (blue shaded area of top right) with extracted g-factors. Bottom right: integrated histogram of level spacings extracted at high B-field (blue arrow top right).

by sweeping the bias voltage at a fixed gate voltage near the degeneracy point[6]. We will analyze the low field and high field regime of Fig. 3.10B, and discuss the intermediate field regime in two other samples.

Low field regime

Note that every resonance which occurs at $B = 0$ is two-fold degenerate as required by Kramer's degeneracy (except the resonance marked with *). Application of a magnetic field breaks this degeneracy and hence each resonance splits into two. If there was no spin-orbit interaction this splitting would be the Zeeman splitting between the spin-up and spin-down states of an electron. In that case, the orbital wavefunction would carry no magnetic moment [9] and the measured g-factor should be 2 for each resonance. However, in the presence of spin-orbit interaction spin is not a good quantum number and the Kramer's doublet may be labeled by a pseudospin instead, to indicate that the magnetic moment contains orbital contributions as well[7]. From the linear splitting at low magnetic fields (see Fig. 3.10C) we can extract the g-factors associated with each pseudospin. Because of mesoscopic fluctuations among the wavefunctions the g-factors are different for each quantum state [9]. The g-factors in Fig. 3.10C range from 0.46 to 1.42, with one exception labeled "g=0". In this case the linear splitting near $B = 0$ is too small to be extracted. Two-fold degeneracy is still maintained, by counting the number of resonances at low field and high field in Fig. 3.10B. The interpretation of this resonance is that there's nothing special about it except that it arises from a quantum state which happens to have no orbital moment in the direction of

[6]Of course, the ground state transition still shows up twice (as a threshold for the drain and source), and hence the resonance marked with * has to be ignored.

[7]For the limit of strong spin orbit coupling spin is completely randomized and its contribution to the g-factor is negligible [12].

the magnetic field, due to an unfortunate interplay between level repulsion and mesoscopic fluctuations. It seems to be squeezed in between two nearby states, and only after one crosses around B=2 T does it split into two resonances. We therefore included the "g=0" state when calculating the average g-factor for this sample, which is $\delta = 0.87 \pm 0.42$. This is twice as big as the mean g-factor for a similar sized nanoparticle in Jason Petta's experiment.

High field regime

The difficulty in comparing experimental data with random matrix theory lies in the statistical nature of its predictions. Any distribution can be expanded by its moments, and hence for a probability distribution it may seem natural to quantify the mean, standard deviation, skewness, etc. This can be done if a lot of samples are available, say parked cars in London [19] or Coulomb oscillations in a nicely gated chaotic semiconducting quantum dot [18]. When only a limited number of samples are available, it makes more sense to plot the integrated probability distribution than to plot a histogram, because no binning is necessary and hence no information is lost.

To maximize the number of samples we analyzed the level spacing distribution at high magnetic field, where twice as many resonances are available than without magnetic field. At 8 Tesla neighboring levels have crossed each other roughly two times (g=0.87 and δ=0.26 meV) and hence we can argue that the crossover from GSE (B=0) to GUE is more or less completed. In Fig. 3.10D the integrated histogram of level spacings obtained at high field (blue arrow in panel B) is plotted, normalized by $< \delta >$=0.26 meV/2. The integrated probability predicted by random matrix theory for the Gaussian Unitary Ensemble [1] is plotted in red,

and matches the experimental data quite nicely. No parameters were adjusted. The only deviation is the increased probability we find for small level spacings. Those data points are less reliable because the level spacing cannot be accurately extracted if the spacing is smaller than the thermal broadening.

Intermediate field regime

Visually, the intermediate field regime is most interesting because there interacting quantum levels run into each other and start to mix significantly. The interaction is caused by the spin-orbit coupling, and the amount of mixing is different between different wave functions. This is in contrast to Al nanoparticles, where the absence of spin-orbit interaction generally leads to the crossing of neighboring energy levels without any interaction [15]. As apparent in Fig. 3.10B some states seem to become degenerate while others avoid each other ("avoided level crossing").

In Fig. 3.11 we show similar features obtained from two other samples with very different level spacings. Mixing between wavefunctions not only affects their energy, but also other properties which may be observable. One nice example is marked in the left data set. As the two levels undergo avoided level crossing spatial properties of their wavefunctions are exchanged, including their overlap with source and drain electrodes. This is evident from the height of the conductance peaks. In the right panel a dashed line marks the evolution of a hypothetical state with constant magnetic moment if it would not mix with other states. Because of mesoscopic fluctuations and spin orbit coupling we see instead a pronounced avoided crossing around B=3.8 T and a "normal" crossing at B=8 T.

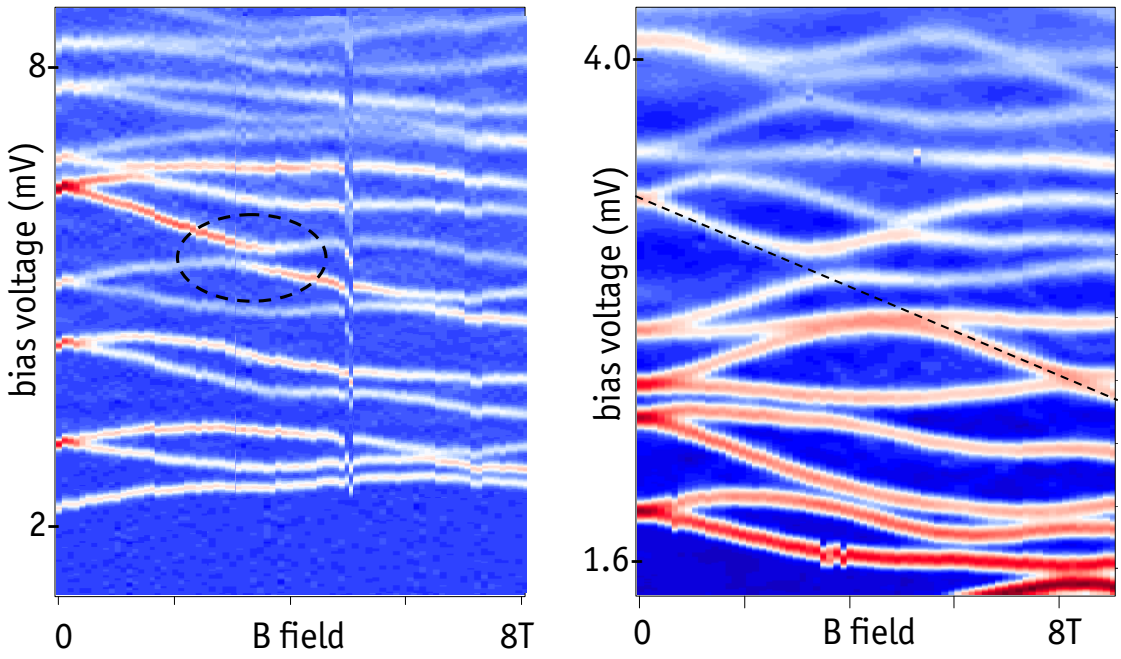


Figure 3.11: Two samples with very different level spacings, showing the same features of avoided level crossings. Left: The marked anticrossing nicely reflects how spatial properties of two wavefunctions are exchanged between two levels as they undergo avoided level crossing (in this case seen in the associated conductances). Right: Following a hypothetical single particle orbit (dashed line) from left to right illustrates mesoscopic fluctuations: strong avoided level crossing with a state near 3.8T and almost no mixing with another state near 8T.

Summary of g-factors in chemically synthesized gold nanoparticles

Our findings on g-factors are summarized as follows. We have measured g-factors for chemically synthesized gold nanoparticles 5-15 nm in diameter. We did neither find a correlation between g-factors and the size of the nanoparticle, nor a correlation between the g-factor of a state and the nearest neighbor in energy. In Fig. 3.12 we plot the g-factors of individual quantum states against the "nearest neighbor" (note logarithmic scale). With "nearest neighbor" we actually take the average between the energy difference to the next higher and next lower quantum level. The g-factor is different for each quantum level, but does not correlate with the proximity of a higher or lower energy level. The mean of all g-factors is 0.84 ± 0.39. This is consistent with theoretical calculations for ballistic nanoparticles and strong spin orbit coupling, which find that the orbital contribution is of order one and the contribution of spin to the g-factor is very small [9, 12]. For diffusive particles, the spin contribution would be the dominating contribution and would therefore lead to g-factors much smaller than 1 (see Fig. 3.12 right panel).

The g-factors we measure in chemically synthesized particles of various size, as well as evaporated nanoparticles (Table 2.4) are all larger than the ones measured by Jason Petta. From this we conclude that the origin for the small g-factors is not the size of the particles, but must be due to some other difference in the particles properties. In all chemically synthesized nanoparticles measured, we did not find any g-factors larger than 2, probably due to strong spin orbit coupling and weak exchange interaction in gold [8].

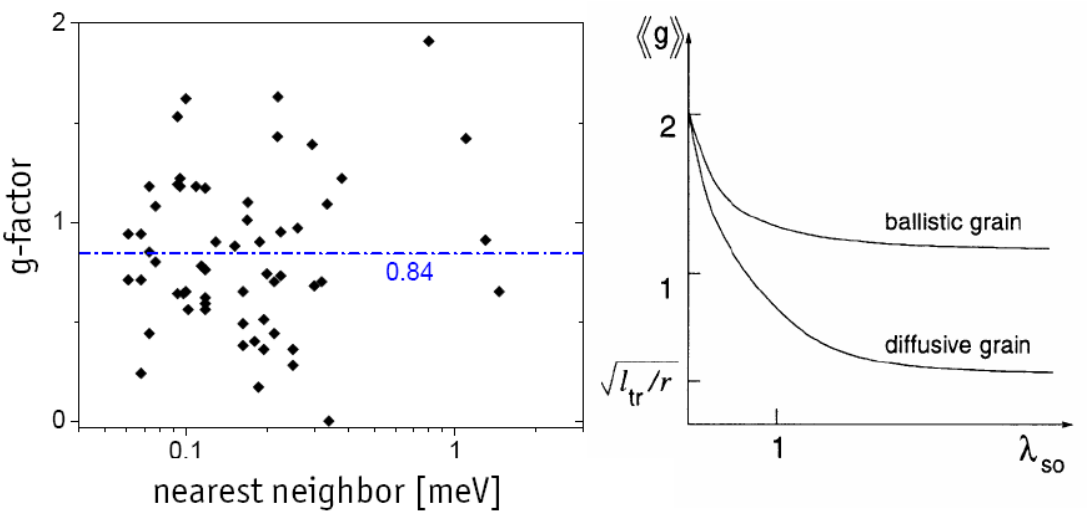


Figure 3.12: Left: g-factors of various samples fabricated from 5–15nm sized gold nanoparticles plotted against the nearest neighbor (see text). The mean value of all g-factors is 0.84 ± 0.39. Right: Dependence of g-factors on the strength of spin-orbit interaction, for diffusive and ballistic nanoparticles. From [12].

3.4 Outlook

We have demonstrated that chemically synthesized metal nanoparticles and break junctions can be combined to give functional single-electron transistors. Without a protective layer the devices were stable enough to allow tunneling spectroscopy in a dilution refrigerator. Even though oxidized aluminum gate electrodes could not be used, the coupling to a silicon backgate 30 nm away proved high enough to allow efficient gating.

For future experiments two improvements are desirable: Increasing the tunnel coupling, and adopting the fabrication scheme to metals other than gold. The tunnel coupling may be increased by simply using a shorter version of the molecule we used, or by using conjugated molecules. However, it is not clear whether high enough conductances can be achieved to observe Kondo effect or other higher-order processes phenomena. To achieve that, one may have to do without any organic monolayer, for example by using the technique based on "electrotrapping" [16,17,20,21]. Whereas the trapping technique described in this chapter was a fine tuned interplay between the nanoparticle charge and the protonation of an organic monolayer, electrotrapping has the advantage that it may work for a lot of different materials, including Pd, Pt and Au(Fe) nanoparticles, which we have already chemically synthesized in lab (see Fig. 3.13). Au nanoparticles with iron impurities were synthesized following [22]).

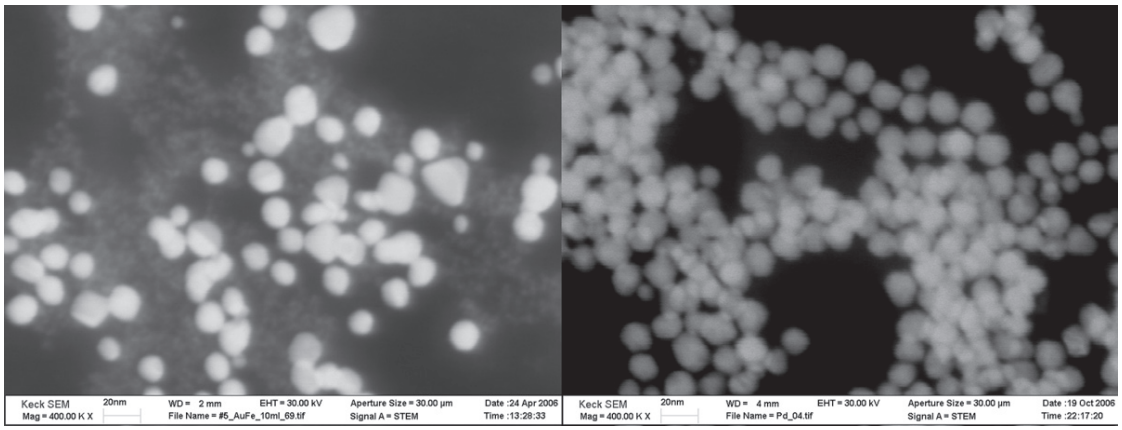


Figure 3.13: Chemically synthesized Au(Fe) nanoparticles (left) and Pd particles (right) are interesting candidates for future experiments to study magnetic interactions and electron-electron interactions.

REFERENCES

[1] W. P. Halperin, Rev. Mod. Phys. 1986

[2] J. R. Petta and D. C. Ralph, Phys. Rev. Lett. **87**, 266801 (2001).

[3] D. Davidović and M. Tinkham, Phys. Rev. Lett. **83**, 1644 (1999).

[4] J. von Delft and D. C. Ralph, Phys. Rep. **345**, 61 (2001).

[5] L. L. A. Adams, B. W. Lang, Yu Chen, and A. M. Goldman, Phys. Rev. B **75**, 205107 (2007).

[6] M. M. Deshmukh, E. Bonet, A. N. Pasupathy, and D. C. Ralph, Phys. Rev. B **65**, 073301 (2002).

[7] O. Agam *et al.*, Phys. Rev. Lett. **78**, 1956 (1997).

[8] D. A. Gorokhov and P. W. Brouwer, Phys. Rev. Lett. **91**, 186602 (2003).

[9] S. Adam *et al.*, Phys. Rev. B. **66**, 195412 (2002).

[10] W. D. Knight *et al.*, Phys. Rev. Lett. **52**, 2141 (1984).

[11] T. P. Martin *et al.*, Z. Phys. D **19**, 25 (1991).

[12] K. A. Matveev *et al.*, Phys. Rev. Lett. **85**, 2789 (2000).

[13] A. Anaya *et al.*, Phys. Rev. Lett. **93**, 246604 (2004).

[14] J. Fabian and S. Das Sarma,J. Vac. Sci. Technol. B **17**, 1708 (1999).

[15] D. G. Salinas *et al.*, Phys. Rev. B **60**, 6137 (1999).

[16] A. Bezryadin and C. Dekker, Appl. Phys. Lett. **71**, 1273 (1997).

[17] S. I. Khondaker and Z. Yao, Appl. Phys. Lett. **81**, 4613 (2002).

[18] R. A. Jalabert *et al.*, Phys. Rev. Lett. **68**, 3468 (1992).

[19] A. Y. Abul-Magd, arXiv:physics/0510136v1.

[20] T. Dadosh *et al.*, Nature **436**, 677 (2005).

[21] I. Amlani *et al.*, Appl. Phys. Lett. **80**, 2761 (2002).

[22] K. A. Brown *et al.*, Mater. Res. Soc. Symp. Proc. **900E** 0900-O07-01 (2006).

[23] J. W. Slot *et al.*, Eur. J. Cell Biol. **38**, 87 (1985).

[24] T. S. Ahmadi *et al.*, Science **272**, 1924 (1996).

[25] R. Jin *et al.*, Nature **425**, 487 (2003).

[26] R. Jin *et al.*, Science **294**, (2001).

[27] A. R. Champagne, *Mechanically adjustable single-molecule transistors and stencil mask nanofabrication of high-resolution scanning probes*, Ph.D. thesis, August 2005.

[28] T. Zhu *et al.*, Langmuir **15**, 5197 (1999).

[29] D. W. Thompson *et al.*, J. Colloid and Interface Science **152**, 197 (1992).

[30] B. Kannan *et al.*, J. Vac. Sci. Technol. B **23**, 1364 (2005).

Chapter 4

Fabrication of partially suspended, split-gate carbon nanotube devices

4.1 Motivation and summary of fabrication

In the past, a variety of transistor geometries have been developed to study electronic and optical properties of carbon nanotubes. In addition to top gates [2] and buried gates [3], suspended nanotube devices [4, 5] have attracted interest because they eliminate artifacts arising from contact with a disordered substrate and because they enable studies of a nanotube's mechanical degrees of freedom. Here, we describe the development of a new fabrication scheme, which allows complex gating of a partially suspended nanotube. The next chapter discusses the use of these devices to measure the electric polarizability of few-electron quantum states as well as their magnetic moment in an external magnetic field. The fabrication process has been developed with the goal of obtaining devices that are as clean as possible. Artifacts due to disorder on the substrate may still occur but are easier to identify with the flexibility provided by having two independent gate electrodes. This allows us to selectively measure clean samples and to observe effects which have not been seen in previous experiments. In Chapter 5 we present data showing a new type of shell splitting at $B = 0$ T. By measuring the magnetic field dependence very accurately we identify it as due to spin-orbit interaction.

Summary of fabrication

There are several challenges involved in the fabrication of suspended carbon

nanotube devices. Some of them have to do with choosing a silicon-on-insulator (SOI) platform, and some have to do with our intention of growing the nanotubes at the very last step[1]. Here is a summary for the reader who does not need to know details:

Fabrication starts with a highly doped SOI substrate. Dry etching of the device layer is employed to pattern electrically isolated gate electrodes spaced $0.1 - 2$ μm apart, which are later individually contacted using wire bonds. After isolating the gate electrodes with 100 nm of thermal oxide, all metal electrodes, including bonding pads, are deposited using liftoff techniques before growing the nanotubes. This avoids damage or contamination of the nanotubes by post-growth processing, but requires metals which are compatible with the high temperature chemical vapor deposition (CVD) process used for growing the nanotubes. For this reason we contact the gate electrodes using photo lithography and titanium/platinum leads, and we pattern source and drain electrodes using electron-beam lithography and a tungsten/platinum bilayer or a chromium/platinum bilayer. After applying wet catalyst and growing nanotubes in a CVD furnace, the devices are tested in a probe station and wirebonded for cooldown in a dilution refrigerator.

4.2 Device schematic and design criteria

The challenge is to fabricate single-electron-transistor devices made from an individual carbon nanotube suspended between two independent gate electrodes (which we will call left gate and right gate), as schematically sketched in Figure 4.1.

[1]The former requires spinning on a non-flat substrate and all its related problems. The latter requires the difficult patterning of metals compatible with carbon nanotube growth conditions.

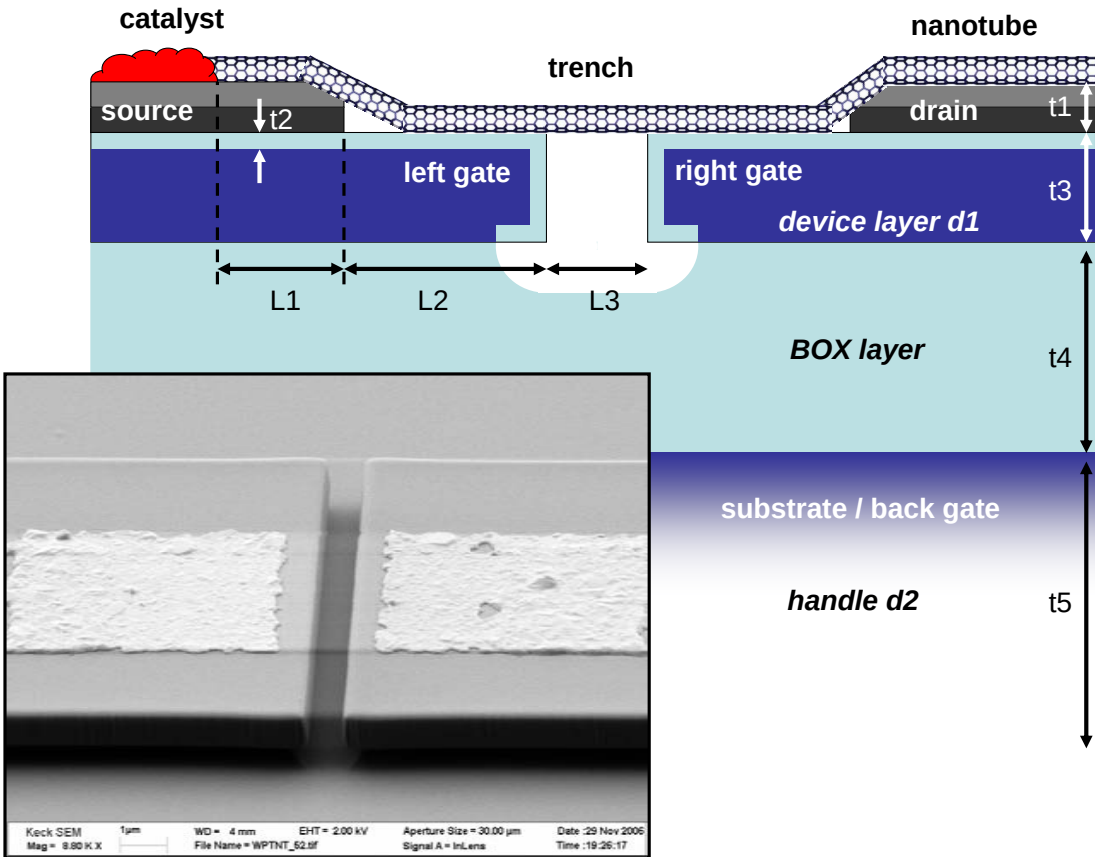


Figure 4.1: Device schematic of a carbon nanotube contacted by source and drain electrode and gated by two independent gate electrodes. The device is used at low temperatures as a single-electron transistor to measure discrete quantum states and their dependence on electric and magnetic fields. The various device parameters are discussed in the main text.

Physics considerations

The choice of device parameters determines what type of physics may be studied. For example, a device with $L_2 = 0$ will resemble the fully suspended, as-grown nanotube devices of H. Dai's group [12], with the new feature of having two independent gate electrodes. To avoid screening by the source and drain electrodes one may want to make L_2 very large. Then the device looks like two quantum dots in series, one localized above the left gate electrode and the other above the right gate electrode, resembling the double-dot devices reported in [7]. The challenge is to chose L_2 large enough that sufficient gating takes place, yet small enough to lead to a single dot localized on the suspended part of the nanotube. By applying gate voltages, we can then populate the dot controllably with any number of electrons ("n-type"), from zero to many tens. We can also apply negative gate voltages and populate the dot with holes ("p-type"). By changing the electric field component parallel to the nanotube (i.e. increasing the difference between gate potentials) we can measure the polarizability of individual quantum states. Increasing the electric field further will at some point lead to a pn-junction located between the gate electrodes, resulting in a p-type quantum dot in series with a n-type quantum dot. These possibilities sound all very interesting, but let us look at the limitations that fabrication imposes on the various device parameters.

Fabrication considerations

Even though silicon-on-insulator (SOI) substrates have become the industrial standard for high-frequency integrated circuits, it is not easy to obtain a wafer which fulfills all requirements. Briefly, SOI-wafers consist of three layers: the main layer ("handle", $t_5 = 300\ \mu$m) gives mechanical support for the "device layer,"

which is isolated from the handle by a buried oxide layer ("BOX", $t_4 = 2000$ μm). The device layer is a doped single crystal silicon wafer which has been thinned and polished by the manufacturer. There are two technologies for making SOI-wafers with a thin device layer, one based on bonding (e.g. ShinEtsu's "Unibond technology") and the other based on separation by implanted oxygen ("SIMOX"). Some manufactures don't like to do highly doped wafers, so it's necessary to look around [1].

Our main requirements are degenerate doping of the handle (d_2) and device layer (d_1) (to allow for gating at low temperatures) and a diameter of 100 mm (to allow the crucial silicon etch in CNF's PT770L automatic etcher). Furthermore, to create electric fields between left and right gate electrodes which resemble those of a parallel plate capacitor, one would like a high aspect ratio t_3/L_3, a thin gate oxide (t_2) and a backgate which is far away (large t_4). However, to obtain electrically isolated gate electrodes, one needs $t_3/L_3 < 10$ because of the finite side-wall steepness after etching. Large t_3 also require long etching times and thicker oxide masks, and therefore will lead to a larger undercut when removing the oxide mask. Also, a thicker oxide mask makes pattern transfer from resist layers more difficult, which are used as etch masks to pattern the oxide mask (the porous front walls in Figure 4.4b are artifacts from poor pattern transfer). If electron-beam lithography is used to pattern source and drain electrodes close to the trenches, then t_3 should be larger than 1 μm in order to get high enough contrast from *etched* alignment marks.

A wide trench on the other hand reduces the expected level spacing due to size quantization (Eq. 1.10). Electrical contact to the nanotube relies on van der Waals and wetting forces [11]. However, if the catalyst is too far away (large L_1)

nanotubes may terminate already on the source electrode. Initially, we thought it would be difficult to obtain bipolar contact if nanotubes are grown at the end of the process (for a discussion of Schottky barriers and a comparison between Pd and Pt contacted nanotubes see [11]). Therefore, a large enough L_2 in conjunction with a thin gate oxide is meant to efficiently turn the nanotube on, both in the "p-type" and "n-type" regimes. A thin gate oxide however is prone to leakage current from the gate to the source/drain electrode, especially after growing nanotubes at high temperature (900℃). We chose a rather thick gate oxide ($t_2 = 100$ nm) and encountered severe leakage, even if the titanium sticking layer was reduced to 1 Å. Because platinum itself does not stick to silicon oxide we used a bilayer of 5 nm tungsten and 25 nm platinum when depositing the source/drain electrodes. This successfully prevented gate leakage, and we think $t_2 = 50$ nm would be a good choice for future devices.

Tungsten is difficult to pattern via liftoff because high power settings are needed for e-beam evaporation, which lead to cracks in the resist. We recently made devices substituting chromium for tungsten, and did not observe any gate leakage. The fabrication of the source and drain electrodes ($t_1 = 30$ nm) is tricky if L_2 is too small for photo lithography. E-beam resist is very thin near the trenches, and special care has to be taken when choosing resist, spin speeds, descum parameters and the thickness of the metal.

Obtaining appropriate SOI substrates (at small quantities) turned out to be difficult. We researched various companies and explored possibilities to perform doping at CNF (spin-on doping) or remotely (implantation). Luckily our contact at Ultrasil put a few bargain wafers aside for us. The specifications are listed in Table 4.1.

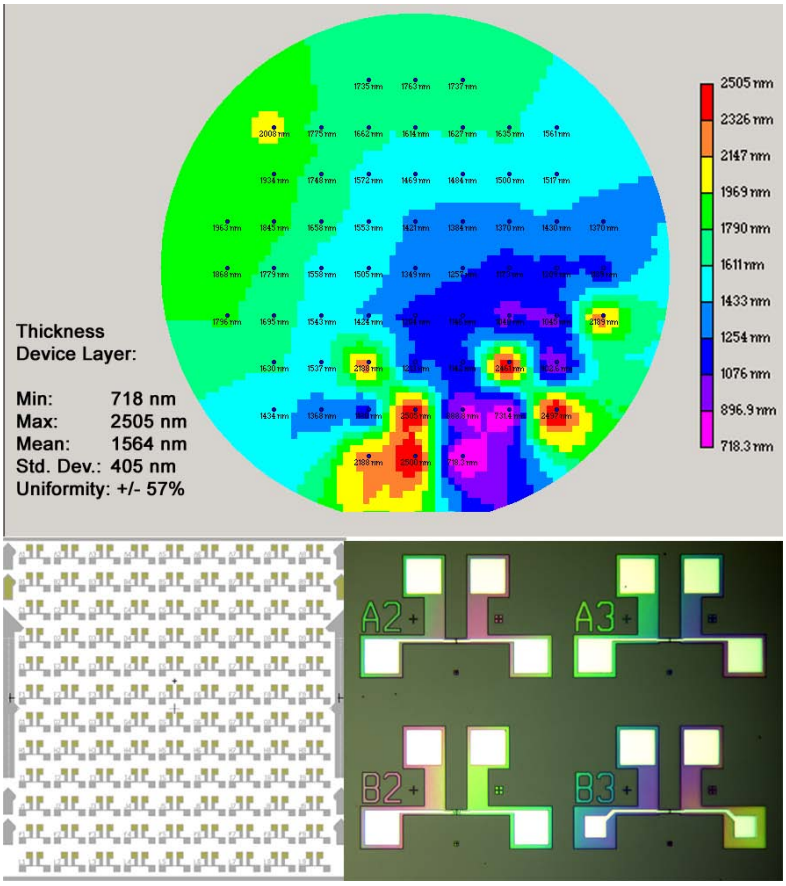


Figure 4.2: SOI-wafers have a non-uniform device layer thickness (top) and hence each die (see bottom left) is an array of many potential devices with different trench widths. Each device has 4 bonding pads, contacting left gate, right gate, source and drain electrodes. The colors of the optical image (bottom right) are due to the variation in the device layer thickness.

Table 4.1: Substrate and device specification, cf. Fig. 4.1. Trench widths may be somewhat different than listed due to swelling during thermal oxidation. In chapter 5 we show data obtained for the three devices listed.

Manufacturer/Lot:	Ultrasil Lot UD-5718
crystal orientation/diameter	Si $< 100 >$ polished, 100 mm
device layer	$t_3 = 1.5 \pm 0.5$ μm, As-doped, $d_1 < 4m\Omega \cdot cm$
BOX thickness	$t_4 = 2$ μm $\pm$ 5%
handle	$t_5 = 300 \pm 5$ μm, As-doped, $d_2 < 4m\Omega \cdot cm$
gate oxide thickness	$t_2 = 100$ nm
distance to catalyst	$L_1 = 0 - 1$ μm
distance to electrodes	$L_2 = 50 - 1600$ nm
trench widths [μm]	$L_3 = 0.1/0.2/0.3/0.5/0.7/1.0/1.2/1.5/2.0$
source/drain electrodes	5/25 nm W/Pt or 5/25 nm Cr/Pt
Device #1	$L_2 = 500$ nm, $L_3 = 500$ nm, 5/25 nm W/Pt
Device #2	$L_2 = 0$ nm, $L_3 = 1200$ nm, 5/25 nm Cr/Pt
Device #3	$L_2 = 200$ nm, $L_3 = 100$ nm, 5/25 nm Cr/Pt

The most important parameter to know before starting fabrication is the device layer thickness (t_3) because it will determine the etch times (and therefore the oxide thickness needed for an etch mask) and the minimum practical trench widths (L_3). It is also the parameter which is most uncertain (wafer specification will read something like 1.5 ± 0.5 μm), so it is a good idea to optically map out the layer thickness using the Filmetrics at CNF. Figure 4.2 indicates that the thickest regions are more than three times thicker than the thinnest parts. An evaluation of all etch times and masking requirements are needed before the first step of fabrication (thermal oxidation) can be done. To cope with the varying device layer thickness, we patterned on each die an array of many devices with varying trench widths, ranging from $L_3 = 100$ nm to $L_3 = 2$ μm. Some of the gate electrodes separated by nominally 100 nm were electrically connected ($k\Omega$ range) after etching the trenches. However, this leakage disappeared after growing the 100 nm thick gate oxide. The variation in the device layer thickness shows up as several periods of colorful fringes across the wafer (also note colors in Figure 4.2), and a rule of thumb is 70 nm thickness variation for each fringe.

4.3 Fabrication I: Whole wafer processing

We will explain the purpose of each step without providing tool settings, except the custom made recipe for the silicon etch. The tool settings which are needed to accomplish each step will depend on the specific wafer and the tools' conditions. It is highly recommended to buy a set of non-SOI "dummy wafers" from the same vendor, with identical crystal orientation, dopants and doping levels as the SOI-wafer. This way, furnace conditions, oxidations rates, etch rates, alignment procedures and much more can be established before performing a step on the

Table 4.2: Major equipment used for fabrication.

step	equipment	comments
thermal oxidation	oxide tube B2	formation of etch mask
alignment marks	HTG contact aligner	for global alignment
	PT72	oxide and silicon etch
oxide mask patterning	Jeol-9300	writing thin trenches
	10x stepper	writing big features
	PT72, Oxford 80	etching the oxide mask
	PT770L	anisotropic silicon etch
	BOE	removal of oxide mask
thermal oxidation	oxide tube B2	growth of gate oxide
gate contacts	10x stepper, BOE	opening windows in the gate oxide
Ti/Pt deposition	Sharon evaporator	using e-beam evaporation
S/D electrodes	Jeol-9300, evaporator	bilayer of 5/25 nm W/Pt (Cr/Pt)
S/D bonding pads	10x stepper	ammonia image reversal
nanotube growth	Plasma Cleaner/Sterilizer	Harrick Model PDC-32G
	Lindberg/Blue tube furnace	1.3kW Model: HTF55122A
wirebonding	MEI wirebonder	Model 1204W
characterization	Filmetrics	thickness of device layer, oxide
	P-10 Profilometer	etch depths
	Veeco Dimension 3100	atomic force microscope
	Keck FE-SEM	scanning electron microscope

actual SOI-wafer. This saves a lot of time in the long run, unless one forgets that SOI wafers may be thinner[2] than the dummy wafers (i.e. more fragile) and doubly polished (i.e. more slippery when handling). Each tool at CNF should be cleaned or conditioned before use. This includes running a dummy oxidation (with HCl, but without any wafers) when conditioning the oxide furnaces, or cleaning the chamber *and* running a dummy wafer prior to using reactive ion etchers.

First oxidation

The first thermal oxidation of the device layer serves two purposes: to create a high quality etch mask for the crucial silicon etch step and to calibrate oxidation rates for the second (and final) oxidation, which will form the gate oxide. Because oxidation rates are different for high concentrations of As dopants, we measured the sheet resistance of our device layers using CNF's 4-point automatic mapping tool and simulated oxidation rates using "Athena," which proved very accurate: shooting for 200 nm thermal oxide, we obtained 199.4 nm (center of wafer) and 201.1 nm (edge of wafer) by running the DRY HCL OXIDE recipe (3%) for 90 minutes at 1100℃, followed by a 20 min anneal.

Etching alignment marks

GCA keys for the steppers, as well as four global alignment marks (crosses consisting of 200 μm long and 4 μm thick lines) are etched all the way through the device layer as follows: Shipley S1813 photo resist is exposed using a contact aligner and hard baked for 10 minutes on a 115℃hotplate after developing. On the PlasmaTerm PT72, use recipe 1 (oxygen plasma, 1 min descum) followed by

[2]Place SOI-wafers on filter paper when stepping such that the height is within the focus range.

recipe 2 (oxide etch) and 5 (SF_6 silicon etch) to etch through the thermal oxide and device layer. Verify with profilometer. Run recipe 1 (oxygen plasma) for 1 minute to remove polymerized photoresist before removing the remainder of the resist using the AZ300T hotbath (30 minutes).

Patterning the oxide mask

The oxide mask is generated by removing the thermal oxide almost everywhere, except where the gate electrodes will be (all bonding pads will sit on the gate electrodes, as seen in Figure 4.2c). This is achieved in two steps, using e-beam lithography and photo lithography.

A PMMA layer thick enough to serve as an etch mask for the oxide layer is spun, exposed and developed as follows: spin acetone/IPA, then 495K PMMA A8 (8% in anisole) for 60 s at 3500 RPM (ramp speed 1000 RPM/s), bake for 15 minutes on 170℃hotplate. This gives 540 nm of PMMA (Filmetrics). On the Jeol-9300, align wafer to the four deep etched global alignment marks, and use the standard 2 nA beam current to expose trenches of desired width (and slightly longer than needed) on each die[3]. The schedule file and job file are reproduced at the end of this chapter. Develop in IPA:MIBK 3:1 for 2 minutes, rinse in IPA for 30 s, squirt with IPA for 30 s, then blow dry. Descum for 0.15 minutes using PT72 recipe 1 (oxygen plasma - will remove roughly 40 nm of PMMA). Hard bake PMMA on 100℃hotplate for 90 s.

Then, the oxide is etched using the CHF_3/O_2 recipe of the Oxford dry etcher. PMMA is a poor mask and will lose almost 200 nm in thickness. Remove the remainder of the PMMA using the Aura 1000 (recipe 9).

[3]In the future, we would include four additional dummy features per die, large enough to be used for "local alignment" when patterning S/D electrodes.

Photo lithography is used to define the bigger features of each device (Figure 4.2c). Spin acetone/IPA, P-20 and S1805 at 2000 RPM for 30 s, bake on 115℃hotplate for 60 s. This gives 600 nm thick resist. Expose 14×14 dies using the 10x stepper at a pitch of 4.9 mm. This way the die will fit into the 6 mm$\times$ 6 mm cavity of a chip carrier at the end of the process[4]. Develop in MIF300 for 60 s, descum for 0.33 minutes (PT72 recipe 1), hard bake on 115℃hotplate for 10 minutes. Figure 4.3a shows how the e-beam written features look underneath the photoresist at this stage of processing. Load into the Oxford dry etcher, include a 10 s oxygen descum before running the CHF_3/O_2 oxide etch. The oxide etch will also remove more than 100 nm photoresist and burn its surface, so we typically run a 1.5 min oxygen plasma in the PT72 (recipe 1) followed by the Aura 1000 (recipe 9) to remove the remainder of the photoresist[5]. Figure 4.3b shows the finished oxide mask right after this step.

Silicon etch and mask removal

The anisotropic silicon etch is accomplished using a Cl-based dry etch in the PT770L (this is a non-Bosch process). The left chamber is dedicated to silicon substrates only, and only oxide or FOX masks are allowed. To get steep side walls, it is important to condition the chamber by running a silicon etch on an oxidized wafer with partially exposed silicon. Michal Lipson's group perfected the art of fine tuning the etch parameters, and our recipe is based on theirs:

fk33.prc:
1) initialization
2) He stabilization, turbo 13 mtorr, 2:30 min, 0/0/0/Ar=30

[4]No dicing! It will be cleaved by hand.

[5]We suspect that this recipe may result in residue on the wafer which is hard to remove. Verify with an AFM.

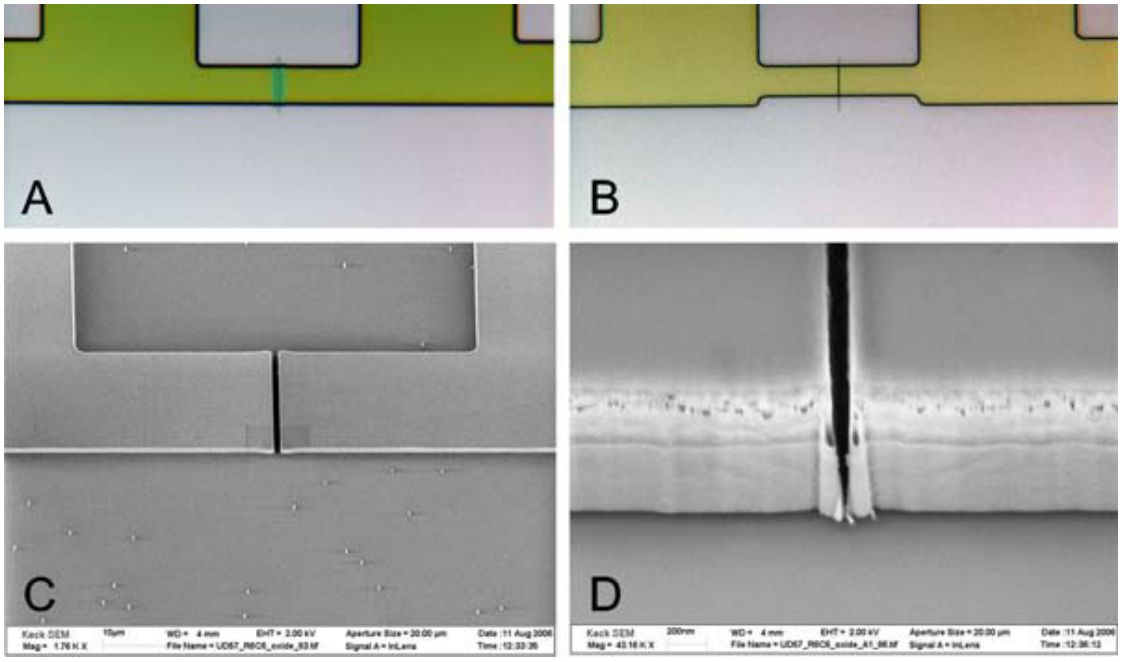


Figure 4.3: Etching of the device layer. (A) Trenches have already been etched into the thermal oxide layer. They are slightly longer than needed to allow for misalignment. Photoresist (greenish color) covers the trenches and the gate electrodes. (B) Finished oxide mask. 200 nm oxide = golden, exposed silicon = white. (C) Right after the anisotropic silicon etch. Note that masking "defects" give rise to lots of little pillars, which will cause trouble when released. (D) Close up view of a 150 nm wide trench before removing the oxide mask.

3) process gas stabilization, 13 mtorr, 15s, Cl2=10/BCl3=20/H2=3/Ar=0
4) light plasma, 13 mtorr, 10 s, 10/20/3/0, RF1=50/RF2=800
5) native oxide etch, 20mtorr, 10, 10/20/3/0, RF1=50/RF2=800
6) etch warm up a, 23 mtorr, 15 s, 25/15/2/0, RF1=50/RF2=850
7) etch warm up b, 25 mtorr, 15 s, 50/7/2/0, RF1=60/RF2=8508
8) Main etch step, 30 mtorr, 10 minutes, 70/2/2/0, RF1=65/RF2=850
9) pump it, 10E-2 torr, 20 s
10) default purge step, 10 s 10 mtorr
11) we're outta here! hold 1 min 10E-2 torr

Residues from the photoresist before etching the oxide mask, or defects in the oxide which did not get etched by the Oxford dry etcher, may lead to pillars all across the wafer, as seen in Figure 4.3c. Panel d shows a nominally 150 nm wide trench covered by the oxide mask. Note that the front wall of the device layer is very rough, probably due to the finite etch resistance of photoresist when patterning the oxide mask (the trench itself was defined by e-beam resist and looks clean, that's the one we care about).

The oxide mask is removed using buffered oxide etch (BOE 6:1) during which the pillars (visible in panel c) also get released (they are sitting on the BOX after all). They will cluster up near the trenches (much worse than in Figure 4.4a) if not taken care of properly. The trick is to keep the wafer wet after the BOE etch, and to perform a post-CMP brush clean (near the chemical mechanical planarization tool). It is a scary thought to run a mechanical brush over the devices, but that particular brush has been developed for polished silicon surfaces and uses an ammonia-based chemistry which helps in removing the particles. Panel b shows a 150 nm wide trench after the oxide mask is removed. At this stage a probe station is employed to verify that left and right gate electrodes are electrically disconnected.

Gate oxidation and gate contacts

Next, a 100 nm thick thermal oxide (blueish color) is grown using the DRY

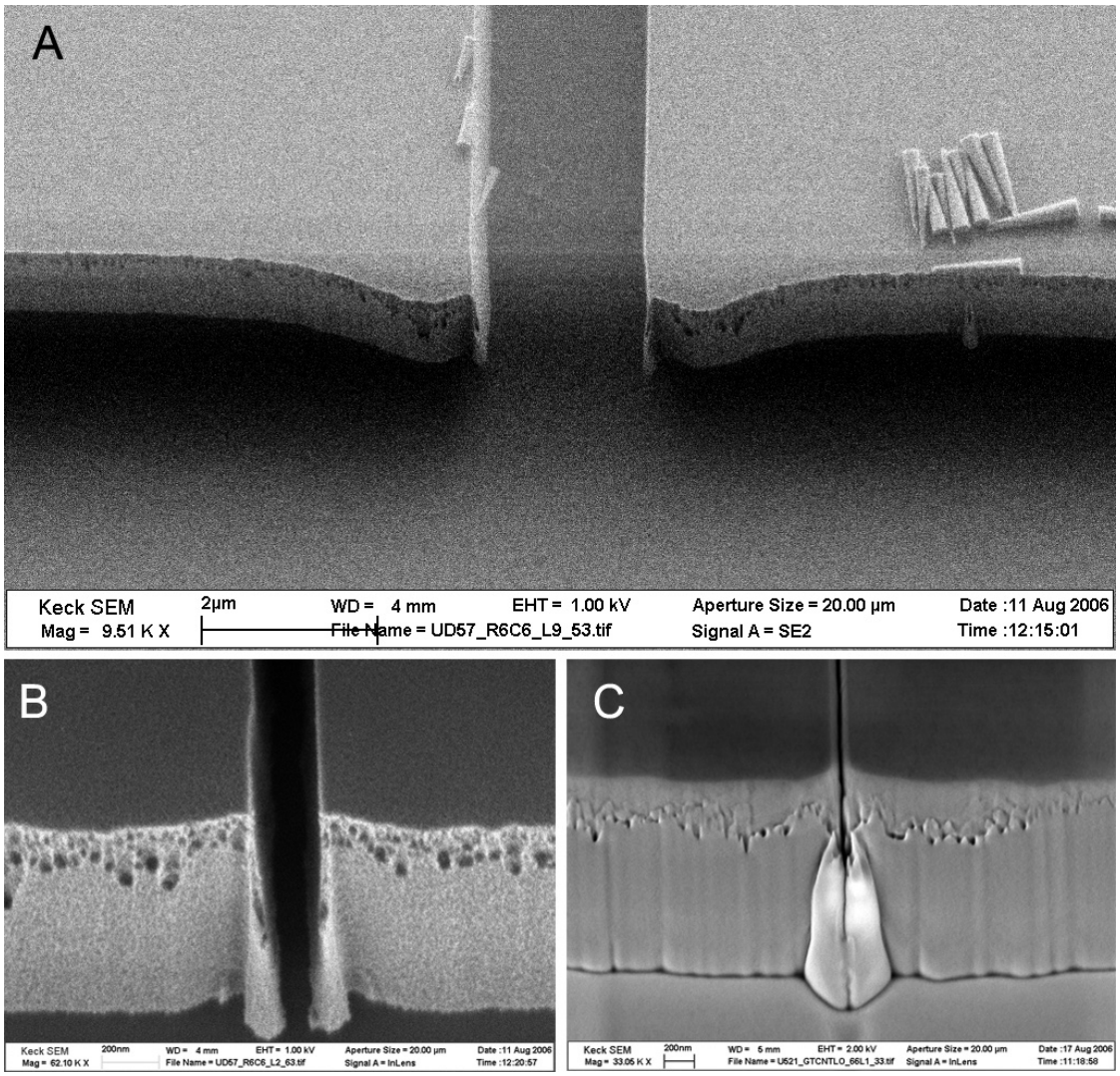


Figure 4.4: A) Removal of the oxide mask via a wet buffered oxide etch also releases the cone shaped pillars, which are removed with a brush developed for CMP processes. B) Close up of a 150 nm trench before growth of the gate oxide. Note the porous front wall resulting from photoresist, and the smooth trench walls resulting from e-beam resist. C) Growing a 100 nm thick thermal gate oxide will lead to a swelling of roughly 40 nm, and hence the trenches get smaller. This picture shows a nominally 100 nm wide trench.

HCL OXIDE recipe (3%), this time 50 minutes at 1050°Cfollowed by a 20 min anneal. The gate contacts can now be contacted using photolithograpy. Expose and image reverse S1813, followed by developing and descum in the Branson barrel etcher (1 min at 80°C). Then inspect the gate openings via AFM for scum and smoothness of thermal oxide, before performing a 2 min BOE 5:1 etch to remove the oxide. Load immediately into the evaporator and deposit 20 Å of Ti and 500 Å of Pt. This completes the bonding pads for the gate electrodes, which are roughly 100 μm $\times$ 100 μm in size (visible as white squares in Figure 4.5a).

E-beam patterning of source/drain electrodes

Electron-beam lithography is employed to position source and drain electrodes accurately and close to the trenches. Global alignment to the same four alignment marks would lead to inaccurate placement of the electrodes (shifts larger than 1 μm across the wafer) since the wafer changes its strain during processing. Therefore, a combination of global alignment and local alignment is used. The MMA/PMMA bilayer becomes very thin near the edges of the gate electrodes, which is a problem if the electrodes need to be placed close to the trenches (see the lack of any undercut in Figure 4.5b). Therefore, the resist needs to be spun thicker than needed and the descum clean (after developing) reduced to a minimum. This part is the most tricky one (see Figure 4.6), and in the future, it may be a good idea to explore the possibilities of planarizing the surface before spinning[6].

We typically clean the wafer before spinning by running an oxygen plasma for 2 minutes (PT72 recipe 1), then spin PMMA(8:5)MMA, 11%, 3:1 in ethyl lactate at

[6]For this purpose, there exist polymers which can be spun on and thinned down by dry etching until the device layer gets exposed, at which point the surface of the wafer is flat enough to allow spinning. The polymer can be removed afterward.

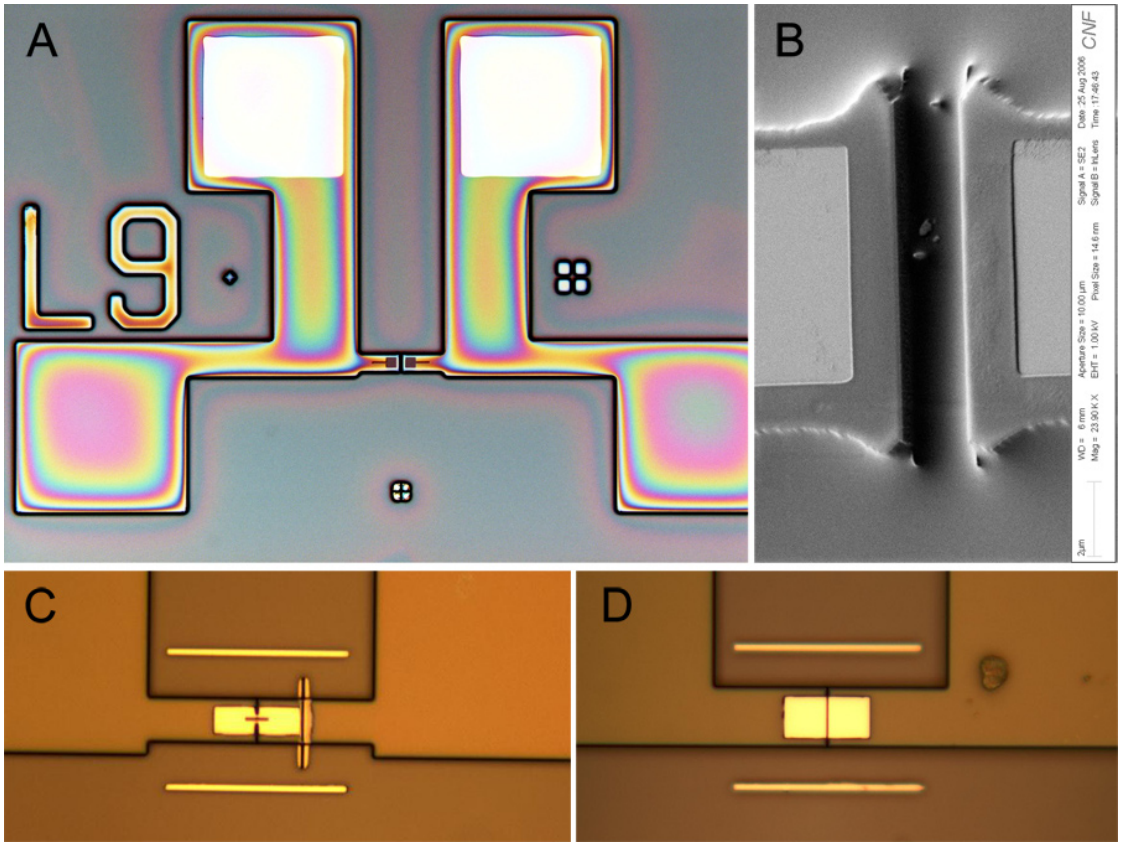


Figure 4.5: A) Optical image of a device after contacting the gate electrodes (white squares) and after developing the source/drain features in the MMA/PMMA bilayer. Color fringes indicate that the resist gets very thin near the trenches (and other edges), as seen in B. C) Alignment of source and drain electrodes via e-beam lithography is achieved by local alignment to 4 sacrificial devices per die. Scanning the beam horizontally over the trench and vertically over the right gate electrode is a convenient way of alignment. Devices which are not used for alignment look like D. The two horizontal bars above and below the device are explained in Figure 4.7 ("crack guides").

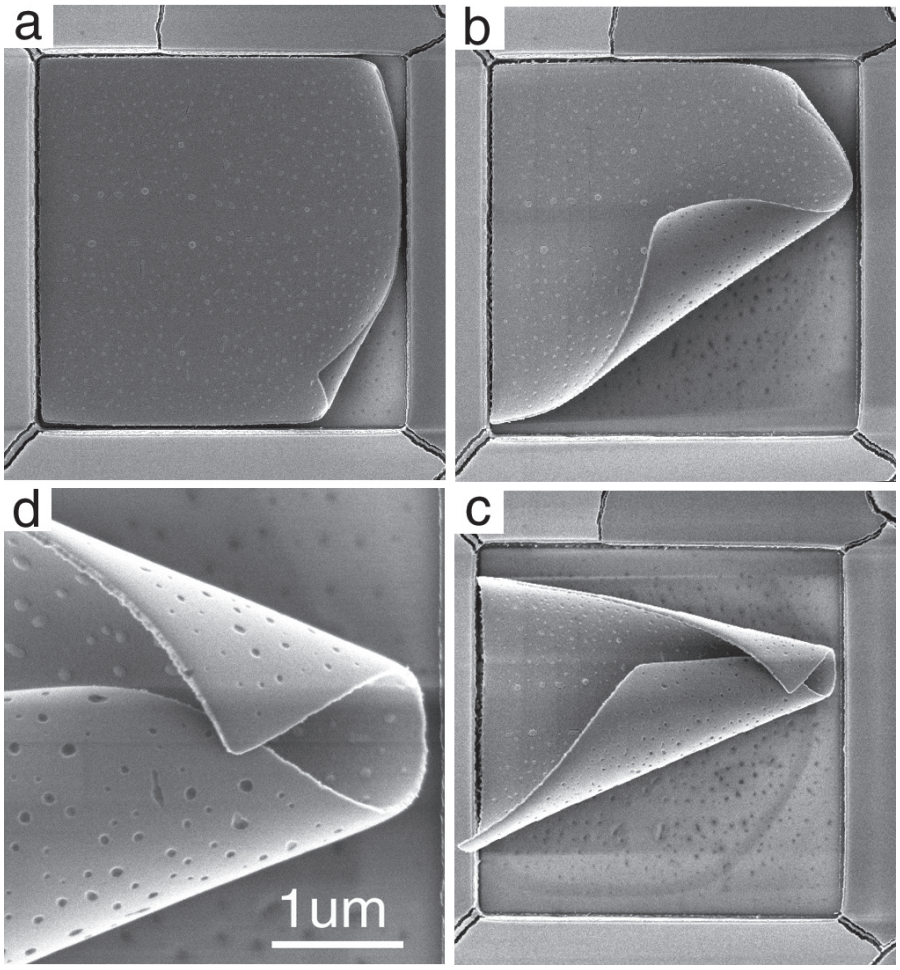


Figure 4.6: Thin e-beam resist and metallic bilayers like 5/30 nm W/Pt are difficult to pattern: In this sample an insufficient descum of the patterned substrate before metal deposition resulted in the peeling off of a (thermally stressed) electrode 9 μm $\times$ 9 μm during SEM imaging. The imprint of the scum on the metal film is clearly visible in panel d. Increasing the descum time will make the resist too thin, whereas a sticking layer like Ti will lead to gate leakage.

2000 RPM for 60 s and bake at 170℃for 15 minutes (this gives 350 nm thick resist on flat wafers). Then we spin and bake 5.5% 495K PMMA in anisole at same speed (equivalent to 370 nm on flat substrates). The electrodes are exposed in the Jeol-9300 using its capability to align to non-standard alignment marks. This is convenient, because in the "x-direction" alignment needs to be with respect to the trenches (first e-beam writing) and in the "y-direction" alignment is with respect to the photolithography layer which defined the gate electrodes. *Those two coordinate systems are not the same, due to misalignment of the stepper mask, as well as physical deformations of the wafer while processing. They differ not only by an offset or rotation, but also by a scaling factor leading to shifts larger than* 1 μm *across the wafer!* Therefore, in each die four devices are sacrificed to serve as local alignment marks. On such a device, the electron beam will scan horizontally across the trench to locate the "x-coordinate", and scan vertically (slightly offset) to locate the "y-coordinate", as seen in Figure 4.5c (panel d shows a real device which has not been used for alignment). To speed up the exposure, a beam current of 27 nA is used by simply opening and realigning the aperture. The schedule file and job file are reproduced at the end of this chapter. Finally, we develop for 90 s in 3:1 IPA:MIBK.

4.4 Fabrication II: Cleaving and final processing

It is best to cleave the wafer in such a way that it can be be re-exposed if needed (the Joel has limited capabilities of exposing wafer pieces). Those pieces selected for final processing should be large enough to allow stepping of the source and drain bonding pads, yet small enough to fit into the carbon nanotube growth furnace (4×4 dies works well). Thin SOI wafers are easy to cleave by hand, and a very short

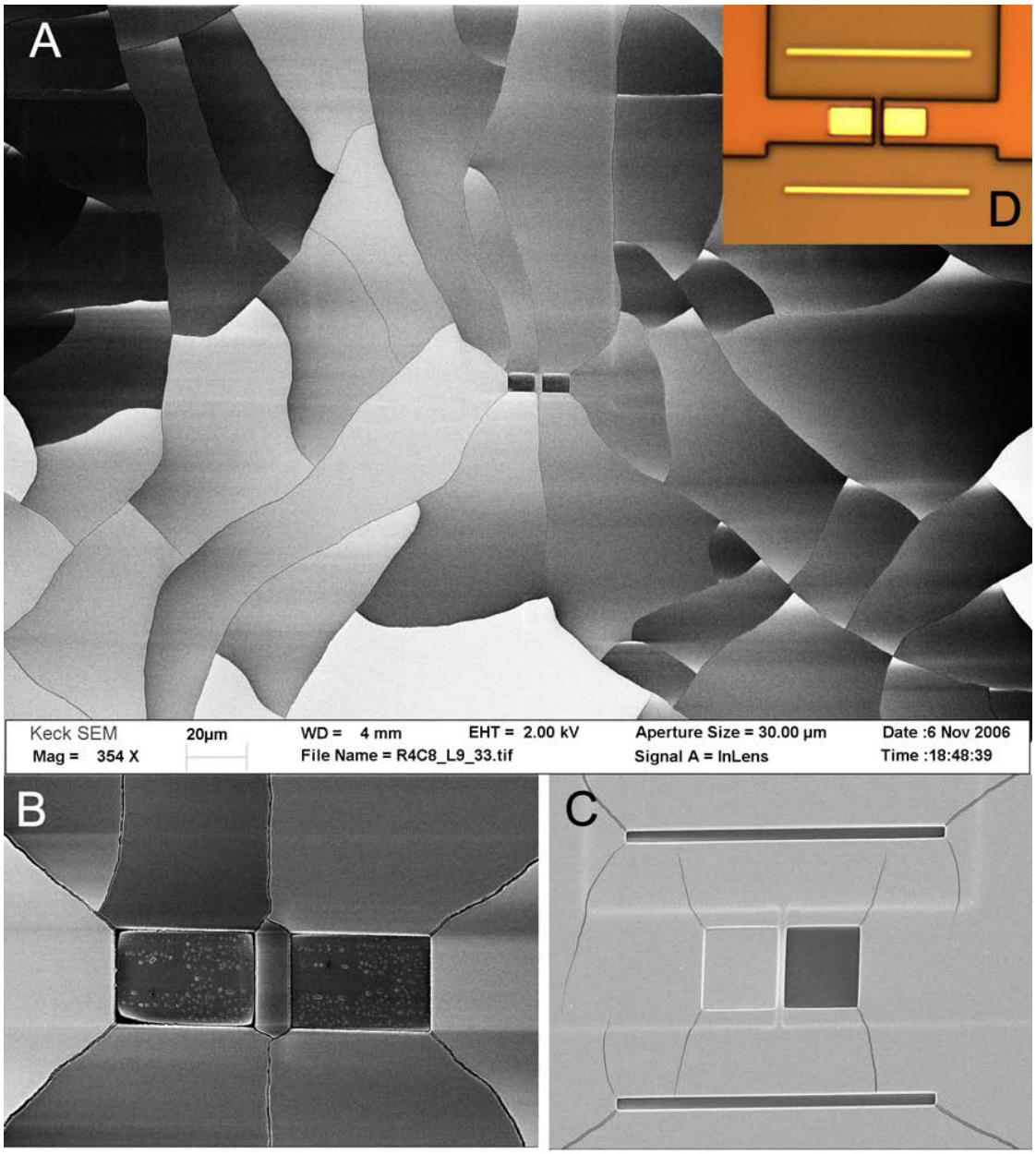


Figure 4.7: A) A bilayer of W/Pt before lift-off. Note that cracks developed during evaporation such that different flakes are electrically isolated from each other (the gray tones result from charging effects during SEM imaging. B) If no crack guides are used, the cracks are large and liftoff bad, resulting in difficulties to contact the electrodes in a later metal deposition (see for example left electrode in Figure 4.8b). C) Two electrodes (9 μm $\times$ 9 μm) patterned successfully on top of gate electrodes before liftoff. Note how the crack guides above and below the gate electrodes influence the formation of cracks in the resist, thereby reducing the retraction of resist and keeping the undercut large enough to result in good liftoff. D) Optical image taken after liftoff.

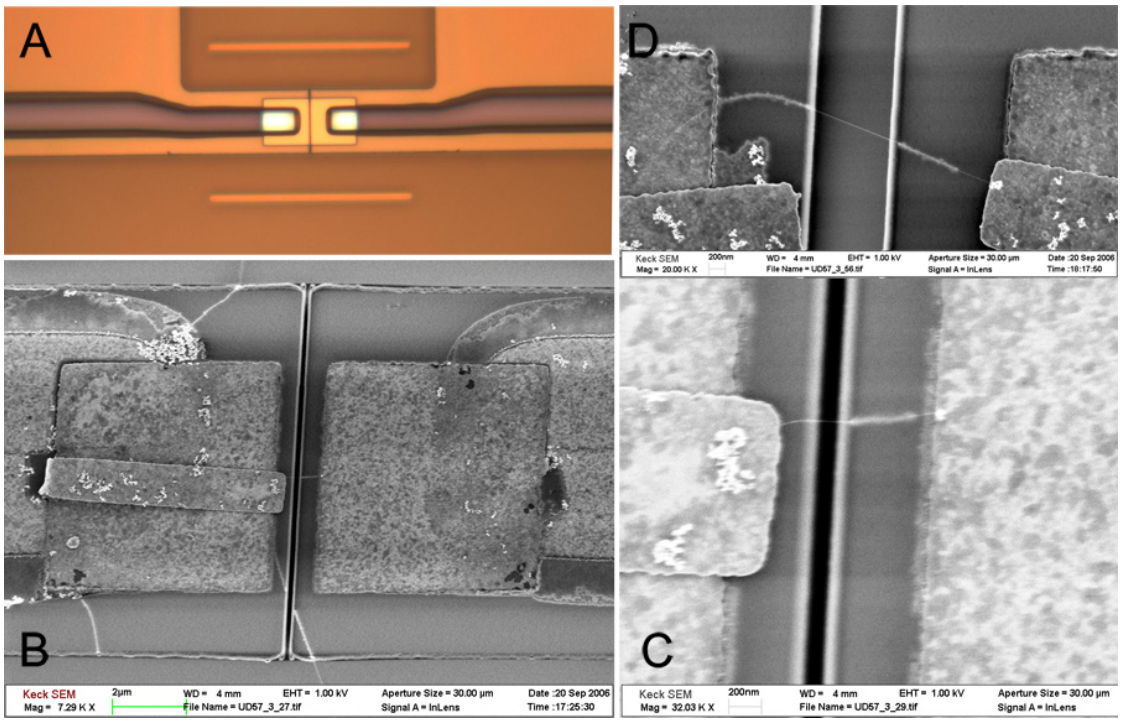


Figure 4.8: A) Source and drain electrodes are contacted using photo lithography (shown after developing). After developing but before liftoff, an aqueous solution is applied to the chip, followed by liftoff in warm acetone and growth in a CVD furnace. B) A large undercut (~ 1 μm, also visible in A) is used intentionally to bring the catalyst (bright spots) closer to the trenches. The tail of the left electrode got folded to the right unintentionally due to capillary forces during development (A shows a newer design without tails). C) and D) Finished devices with nanotubes crossing a 150 nm wide and 1000 nm wide trench. Nanotubes conveniently show up bright where they are in contact with the oxide.

descum plasma (3 s PT72 recipe 1) is applied only to the pieces which are ready for the evaporator. A bilayer of 5/25 nm of tungsten/platinum is deposited using the "flat stage" to get the most parallel metal beam. After liftoff in acetone/IPA S1813 is spun and the layer containing source and drain bonding pads is exposed on the 10x stepper, using the ammonia image reversal process.

This layer serves two purposes: It connects source/drain electrodes to larger bonding pads, and it is used to pattern the wet catalyst before growing nanotubes. The large undercut visible in Figure 4.8a (after developing) is used intentionally to

bring catalysts close the trenches. Here is how: After depositing 5/25 nm W/Pt using e-beam evaporation, we delay liftoff until *after* wet catalyst is applied to the chip's surface. The wet catalyst is based on a mixture of $Fe(NO_3)_3$, $MoO_2(acac)_2$ and Al_2O_3 similar to the one described in reference [13], with water substituted for IPA. To improve catalyst adhesion the chip's surface may be treated with a short oxygen plasma (15 s using a table-top plasma cleaner), prior to placing a few droplets of wet catalyst onto the chip. After a few minutes, the chip's surface is rinsed with DI water and dried with dry nitrogen. The W/Pt bilayer is lifted off in warm acetone (15 minutes). After gently rinsing the chip in IPA it is ready for growth.

Carbon nanotube growth

Growth conditions are optimized using chips from the dummy wafers. Inspection of the nanotubes via AFM and SEM also verifies whether the catalyst is working (some batches work better than others). We have had success using the following growth procedure: The 2" quartz tube is cleaned by heating to 900℃, and all lines are flushed with ultra high purity Ar, H_2 and C_2H_4 for a short time prior to loading the samples. After loading the samples, a flow of argon (0.8 SLM) is established and the temperature ramped up to 760℃after 5 minutes. Once the setpoint is reached, 0.25 SLM of hydrogen is added for 15 minutes, after which growth is initiated by adding 5.5 SCCM ethylene. Hydrogen and ethylene are stopped after 6 minutes, and the furnace is allowed to cool to room temperature. The flow of argon is stopped before removing the samples from the furnace.

We have also had success with the "flying catalyst" method, in which the tube is pulled through the furnace at 920℃.

4.5 Testing devices using a cryogenic probe station

Each chip has a hundred potential devices and only the four or five best ones will be wirebonded at the end. Therefore, an efficient way of screening all devices is needed. This is best accomplished using a full wafer cryogenic probe station (Desert Cryogenics), where the substrate can be moved from one device to the next without relocating each of the four probes needed for a single measurement. The first screening is done at room temperature with a parameter analyzer, by measuring the conductance between source and drain electrode while sweeping both gate voltages from -5 V to +5 V. Semiconducting devices will show a conductance on the p or n side, whereas nanotubes with a small band gap (these are the ones that we are interested in) may show a dip in conductance near $V_g = 0$ V (see left panel of Figure 4.9). If their p and n type conductance is more than 100 nA (using a 10 mV excitation) they are noted for further testing at 77 K.

At 77 K the good devices (R = 100 kΩ or less) are measured in exactly the same way. Additionally — because liquid nitrogen is cheap and liquid helium so short lived — two-dimensional conductance maps are measured of the best devices to obtain information about all four "corners" of the $V_{left} - V_{right}$ space (i.e. the nanotube's pp, nn, pn, and np conductance). This is accomplished with GPIB controlled voltage sources (Yokogawa 7651), a lock-in amplifier (Stanford Research 830) and a current preamp (Ithaco 1211). An example of a good device is given in Figure 4.10 (device #1), which shows all the features one would like to see at 77 K: a pronounced bandgap ("i"), a high conductance if the entire tube is turned p-type (PP corner) or n-type (NN corner), and an increasing conductance as one goes from fully intrinsic to "p-i-n" to "pn" (by increasing the electric field *parallel* to the nanotube (PN and NP corners). At this temperature, indications

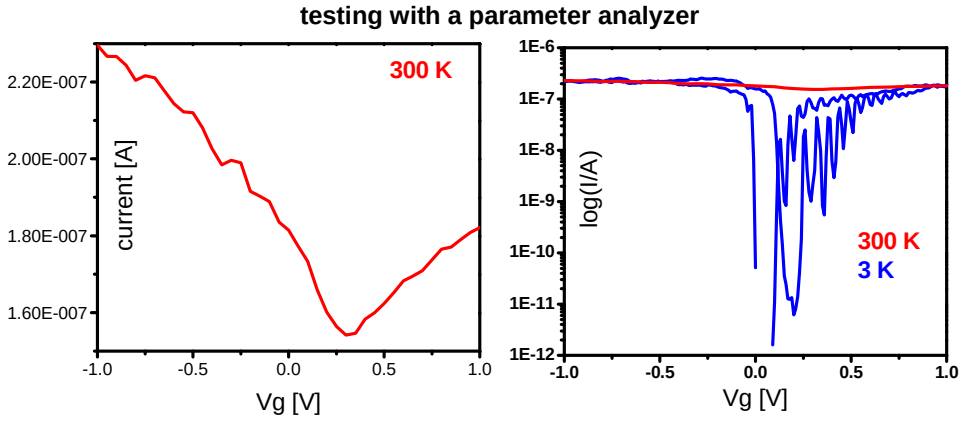


Figure 4.9: Left: After carbon nanotube growth a probe station and parameter analyzer are used to find "good" devices. The left shows the conductance of a $\sim 50k\Omega$ device (device #1) at room temperature as a function of a common voltage V_g applied to both gates simultaneously. A dip near V_g=0 V indicates the presence of a (small) bandgap. Right: same data on a logarithmic scale (red), along with data obtained from the same device at 3 K (blue) after it was loaded into a dilution refrigerator, and tested with a parameter analyzer before circulation was turned on. (Trace and retrace may not match each other if the gate voltage is swept fast.)

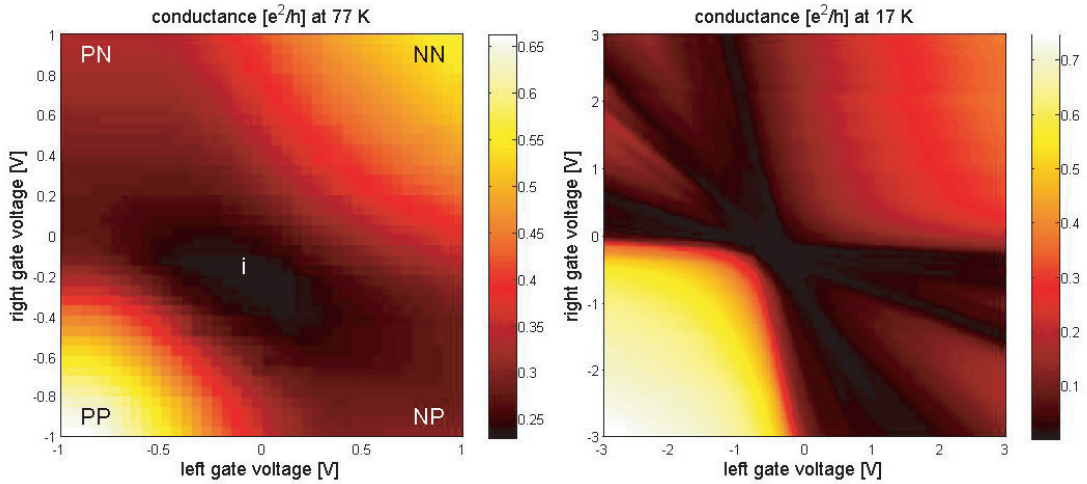


Figure 4.10: Left: Conductance through device #1 measured at 77 K using a cryogenic probe station. Note the high n-type and p-type conductance (NN and PP corners) as well as the low conductance when the nanotube is tuned to the intrinsic regime (i). The PN corner correspond to a nanotube where the left side is doped p-type and the right side n-type, with the transparency of the pn-junction increasing as one increases the electric field. Right: Same device at 17 K.

of Coulomb oscillations associated with a quantum dot may or may not show up. This is why we cool to 4 K using liquid helium before selecting the lucky devices.

Figure 4.11 shows 4 K data for a device which did not show any Coulomb oscillations at 77 K. At 4 K many Coulomb oscillations are visible. Interestingly, by increasing the right gate voltage the Coulomb oscillations change their character from a single dot (schematic C) to the honeycomb pattern of two quantum dots in series (schematic B). This may indicate that in this particular device the quantum dot forms not only on the suspended part, but on parts of the gate oxide as well. The data in panel (a) looks similar to data obtained from 2DEG double dots [6], where the interdot tunneling coupling could be tuned controllably from weak (B) to strong(D). Double-dots fabricated from carbon nanotubes have been reported by the groups of C. Marcus [9,10], L. P. Kouwenhoven [8] and C. Schönenberger [7].

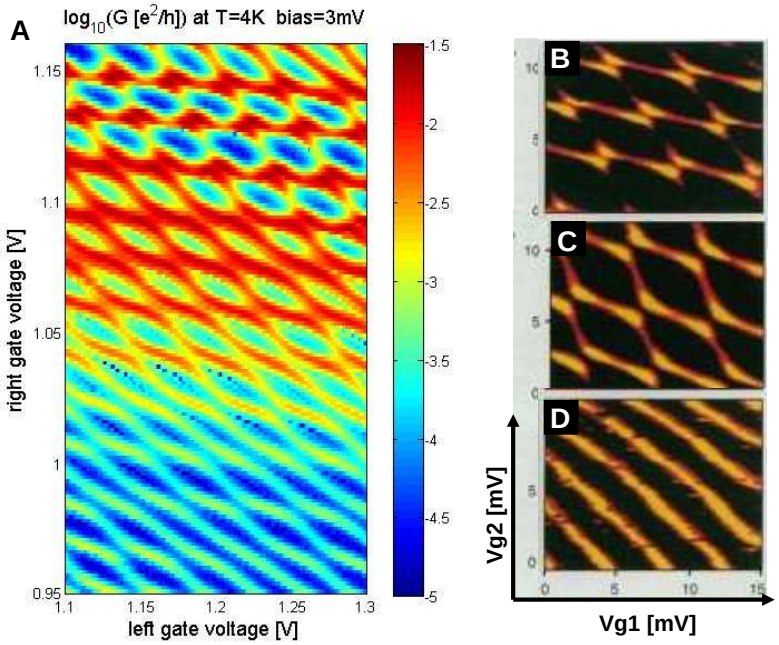


Figure 4.11: (A) Conductance through a sample measured at 4.2 K (with a constant bias of 3 mV). The honeycomb pattern near the top indicates that the charge is distributed between two interacting quantum dots, located presumably on the left and right gate electrode. By reducing the right gate voltage, tunnel couplings change and make the Coulomb oscillation look more like that of a single dot. (B),(C) and (D) show the conductance reported for a GaAs-based double dot with interdot couplings 1.3 e^2/h, 1.56 e^2/h and 1.96 e^2/h respectively [6].

4.6 Measurements in a dilution refrigerator

After identifying good devices using the cryogenic probe station, the wafer piece is cleaved (if necessary) and the die glued onto a chip carrier. Before wirebonding, power and force settings are optimized on a dummy chip to make sure that bonds stick to the bonding pads without penetrating the gate oxide.

The chip carrier is loaded into the chip socket on a Kelvinox dilution unit ("bottom loader") and the devices are tested using a parameter analyzer before cooldown (just like in Fig. 4.9a). We found it helpful to measure all devices at 1.5 K before turning on circulation.

How to reach 1.5 Kelvin

After establishing 4 K the exchange gas is pumped out of the inner vacuum jacket and the 1 K pot is started. Then an appropriate amount of mixture is let in from the dumps into the still line to establish a heat link between the 1 K pot and the sample. The sample can then be measured around 1.5 K without circulating the mixture.

At this temperature the bad samples can already be spotted and may be used to optimize the electronics. The setup consists of several GPIB controlled Yokogawa 7651 voltage sources. Two of them are connected to the left and right gate electrodes using π-filters (cut off frequency 1.9 MHz), and are operated in the lowest range needed. The third one (bias voltage) is added to the excitation voltage generated by a Stanford Research (SR 830) lock-in amplifier using a 10 Ω-10 kΩ resistor network. This reduces the excitation amplitude by a factor of ~ 1000. The smallest possible excitation is then 4 μV. Again we insert a π-filter between the BNC cable and the socket leading to the source electrode. The input of an Ithaco

1211 current preamplifier (usually used in the 10^{-7}-10^{-10} range) is connected to the drain electrode via a 1 $k\Omega$-resistor and a doubly shielded short (36") BNC cable. The output of the current amplifier is split and connected to the lock-in amplifier (to measure the differential conductance) as well as a GBIB-controlled digital multimeter to measure the DC current component (Agilent 34401A). It is important to eliminate all ground loops, which can be identified by monitoring the output of the current amplifier with a spectrum analyzer. This also helps for selecting a "quiet" lock-in frequency, usually between 30 and 300 Hz. At higher frequencies, the capacitive current component may be too dominating (especially if source and drain happen to be connected to the same twisted pair). We eliminate the capacitive current component by parking the gate voltages in a Coulomb valley (i.e. no current flows through the device) and measuring the phase. Then, data is acquired using a phase shifted by 90 degrees.

Measurements at base temperature

At 1.5 K good data can be obtained on the addition spectrum of a quantum dot. This is because at zero bias voltage the position of a Coulomb peak (and therefore the energy of a ground state transition) can be extracted very accurately. However, a lower temperature is of advantage when measuring the excitation spectra at finite bias voltage. The data presented in Section 5.5 (device #3) was taken at base temperature of the dilution unit. The electronic temperature inferred from the sharpest Coulomb peak widths was 130-180 mK.

Jeol-9300 schedule file and job file for writing trenches using global alignment:

(FKLX.v30 contains the actual pattern)

SOI01.sdf:

```
MAGAZIN 'ROB'

#1
%4A
JDF         'SOI01',1
ACC 100
CALPRM '100kv_2nA_JEOL'
DEFMODE 2          ;2_stage deflection
HSWITCH OFF,OFF
GLMDET S
CHIPAL 0
RESIST 1000
SHOT A,4
OFFSET (-620,500)

#1
%4B
JDF         'SOI01B',1
ACC 100
CALPRM '100kv_2nA_JEOL'
DEFMODE 2          ;2_stage deflection
HSWITCH OFF,OFF
GLMDET S
CHIPAL 0
RESIST 1000
SHOT A,4
OFFSET (-500,0)

END 1
```

SOI01.jdf:

```
JOB/W 'RALPH',4          ; 4 inch 1um square pattern
GLMPOS P=(-40000,0),Q=(40000,0)
GLMP 3.0,2000.0,0,0
GLMQRS 3.0,2000.0,0,0
PATH ROB
ARRAY ( -26950,12,4900 )/(26950,12,4900)
ASSIGN P(1) -¿ ((*,*),SHOT01)
```

```
AEND
PEND
;————————————————————————————-
;        Layer Definition
;————————————————————————————-
LAYER 1
P( 1 ) 'FKLX.v30'
SPPRM 4.0,16.0,,,1.0,1
SHOT01: MODULAT ((3,60),(4,60))
STDCUR 2.0;2nA
END
```

Jeol-9300 schedule file and job file for writing S/D electrodes using local alignment:

(FKSOI12.v30 contains the actual pattern of source and drain electrodes)

```
MAGAZIN 'ROB'

#1
%4A
JDF          'SOI13',1
ACC 100
CALPRM '100kv_2nA_JEOL'
DEFMODE 2          ;2_stage deflection
HSWITCH OFF,OFF
GLMDET S
CHMDET S
CHIPAL 4          ;chip alignment using 4 marks
RESIST 1000
SHOT A,4
OFFSET (37.5,529)

END 1

JOB/W 'RALPH',4          ; 4 inch 1um square pattern

GLMPOS          P=(-26950,2528), Q=(7350,2528)
;cross below F5, die R5C1 and R5C8
GLMP 1.0,4.0,0,0
GLMQRS 1.0,4.0,0,0
```

```
PATH ROB
ARRAY (-26950,8,4900)/(7350,2,4900) ;start R5C1
CHMPOS M1=(-2000,2153),M2=(-500,2151.5),M3=(-500,-2248.5),M4=(-2000,-2247)
;trench A1,A4, L4, L1
;CHMARK 1.0,4.0
ASSIGN P(1) -¿ ((*,*),SHOT00)
AEND


            PEND
;————————————————————————————-
;         Layer Definition
;————————————————————————————-
LAYER 1
P( 1 ) 'FKSOI12.v30'
SPPRM 4.0,16.0,,,1.0,1
SHOT00: MODULAT ((11,60),(13,60),(14,90),(15,60))

        STDCUR 2 ;set slightly larger than actual current
END
```

REFERENCES

[1] www.sehamerica.com (ShinEtsu), www.ankamerica.com (ANK America), www.addisonengineering.com, www.siwafers.com (Silcon Inc.), www.soitec.com (Soitec), and wafer brokers such as www.siliconquest.com, www.waferworld.com, www.universitywafers.com. Semiwafer in Taiwan (www.semiwafer.com) had a promising website, but never got back to us.

[2] S. Ilani *et al.*, Nature Physics **2**, 687 (2006).

[3] K. Bosnick *et al.*, Appl. Phys. Lett. **89**, 163121 (2006).

[4] V. Sazonova *et al.*, Nature **431**, 284 (2004).

[5] E. D. Minot *et al.*, Nature **428**, 536 (2004).

[6] C. Livermore *et al.*, Science **274**, 1332 (1996).

[7] M. R. Gräber *et al.*, Phys. Rev. B **74**, 075427 (2006).

[8] S. Sapmaz *et al.*, Semicond. Sci. Technol. **21**, S52 (2006).

[9] M. J. Biercuk *et al.*, Nano Lett. **4**, 2499 (2004).

[10] M. J. Biercuk *et al.*, Nano Lett. **5**, 1267 (2005).

[11] A. Javey *et al.*, Nature **424**, 654 (2003).

[12] J. Cao *et al.*, Nature Materials **4**, 745 (2005).

[13] J. Kong *et al.*, Nature **395**, 878 (1998).

Chapter 5

Few-electron carbon nanotube quantum dots

5.1 Significance of this work

A variety of previous experiments observed shell structure in carbon nanotube quantum dots by measuring addition spectra [10, 14, 23–25]. If the level spacing between different shells is not too small compared to the charging energy, a sequence of three smaller Coulomb diamonds followed by a larger one indicates the presence of a four-fold shell. Shell structure was observed for highly conducting devices [25], weakly coupled devices [11], as well as multi-wall carbon nanotubes [20]. It is of interest to understand the sequence in which the single-particle levels of a shell are filled, because deviations from a simple non-interacting shell filling indicate the presence of interactions such as K-K' scattering, exchange or spin-orbit interaction.

Quite different results have been obtained for the magnetic field dependence of addition spectra [17–19] which do not follow a single scheme. Other groups have analyzed excitation spectra obtained at different charge states [11,13]. Reference [13] claims the observation of an exchange splitting at zero magnetic field for two electrons ($J = 0.5$ meV, Δ=5.9 meV) whereas this is found to be small in [11]. Within the Hartree-Fock approximation [25] the other parameter which has to be taken into account is the excess Coulomb energy δU to put two electrons into a single level, and the interplay between δU, J and Δ then leads to a non-trivial shell filling. Liang *et al.* infer $J/\Delta \sim$ 0.1-0.3 and $\delta U/\Delta \sim$ 0.01-0.05 for devices

with level spacing $\Delta \sim 8$ meV [25]. Oreg *et al.* consider armchair nanotubes theoretically, and calculate the ratio between exchange coupling (J) and level spacing (Δ) for a (10/10) nanotube. They find $J/\Delta \approx 0.22$ and $\delta U/\Delta = 0.11$.

The contribution of our experiments to the understanding of shell filling is twofold. First, we have the capability to modify the electrostatic confinement spatially, by changing two gate voltages independently (Section 5.2). The surprising result is that the magnetic field dependence observed in the addition spectra (Section 5.3) as well as excitation spectra are dramatically modified by electrical fields parallel to the nanotube. This implies that previous experiments with less controlled gating may show effects due to the (unknown) electrostatic environment experienced by the quantum dot. Spin flips in the ground state induced by variation of an external potential have been predicted in [12]. We show that the configuration of a charge state can flip within a Coulomb valley, driven by electric fields as well as magnetic fields (Section 5.4). Second, we observe a new type of shell splitting which is present at zero magnetic field even if the quantum dot is occupied by only one electron or hole (Section 5.5). By measuring the magnetic moments very accurately, we conclude that it is caused by spin-orbit interaction [2].

5.2 Electric polarizability

In Figure 4.9 we observe the appearance of Coulomb oscillations as the temperature is lowered from room temperature to a few Kelvin. We can learn about the nature of this quantum dot by measuring the Coulomb oscillations as a function of the left gate voltage V_{left} and right gate voltage V_{right}, as shown in Figure 5.1a. This device (device #1) was contacted with 30 nm thick W/Pt electrodes, and its designed specification was $L_2 = 500$ nm and $L_3 = 500$ nm (SEM images suggest $L_3 \approx 420$

nm after oxidation).

Diagonal cut $V_{left} = V_{right}$

Let us first discuss the features along the "diagonal cut" defined by $V_g = V_{left} = V_{right}$. This is the familiar regime of previous quantum dot experiments with a single backgate. In our devices this is also the axis which defines "left-right symmetry" of a device (unless misalignment during fabrication or defects break the left right symmetry). Starting in the bottom left corner we find a high and constant conductance of the nanotube with negative voltages applied to both gates. This indicates that the entire nanotube is in its p-type "on-state". Raising the gate potentials along the diagonal we observe a sudden drop in conductance at $V_{left} = V_{right} = 0$, indicating that at least one segment of the nanotube turned off (i.e. the electrochemical potential is in the bandgap). Because of the left-right symmetry of the data, this segment must be either the suspended part of the nanotube or the left and right segment above the gate oxide. It is likely that the presence of oxygen atoms dopes a nanotube p-type where it is in contact with silicon dioxide, and hence we believe it is the suspended part which turns intrinsic at $V_g = 0$, while the segments on the oxide remain p-type. Because the suspended part is neutral at this point we may call it a quantum dot occupied with zero electrons. Increasing V_g further allows us to add one electron after another to this quantum dot, which are manifested as equally-spaced Coulomb oscillations in the conductance. In the language of energy bands, we may say that the suspended segment turns n-type (electrostatically) while the segments on the oxide remain p-type (chemically). In panel B (schematic 2) this is indicated with red and blue color. The intrinsic interface between the p-doped and n-doped segment then

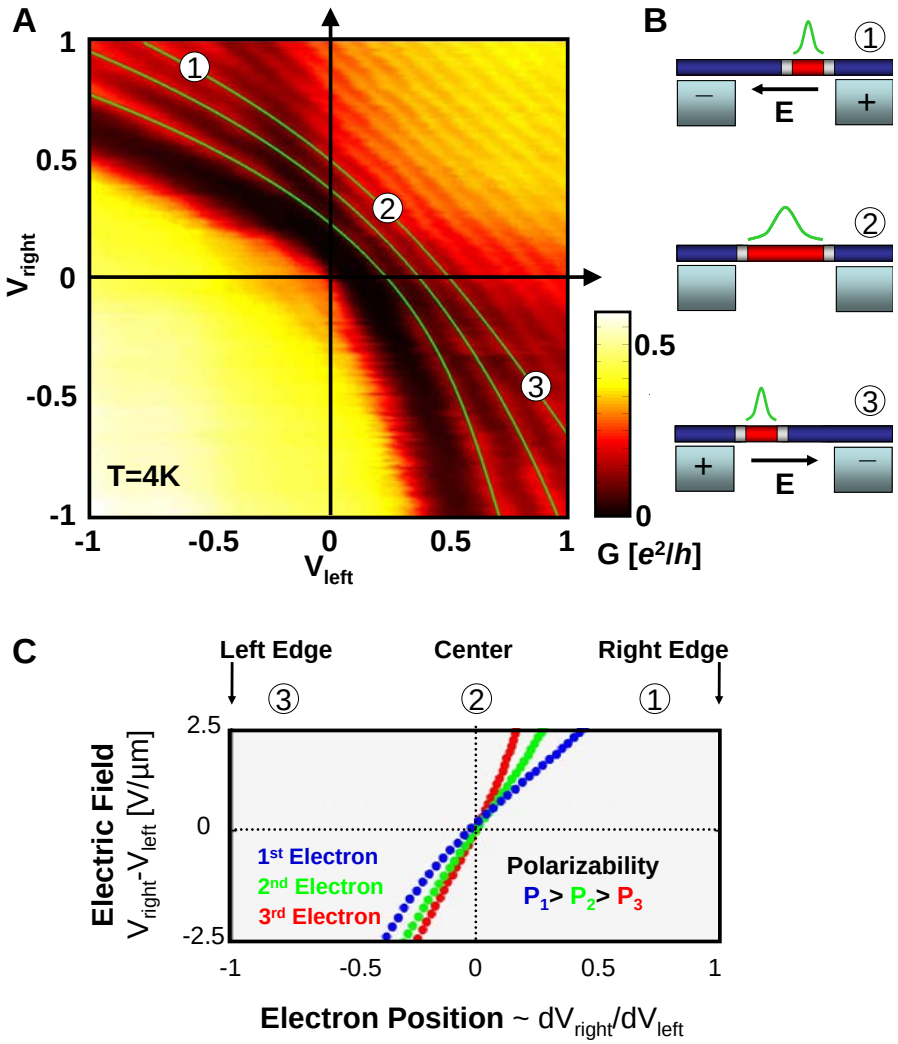


Figure 5.1: (A) Coulomb oscillations of device #1 reveal a quantum dot which is electrostatically confined between the left and right gate electrodes. The curvature of a given Coulomb oscillation reflects the spatial rearrangement under the influence of an electric field parallel to the nanotube. B) The electric field created by the gate electrodes pushes the quantum dot to the left (1) or to the right (3), thereby changing its capacitive coupling to the left and right gate electrodes. This is the origin of the curvature of a Coulomb oscillation. C) The location of the quantum dot's charge is extracted for the first three Coulomb oscillations (green lines in panel A) as a function of electric field between the gate electrodes. We find that the polarizability of the 1st electron is largest, with $P_1 = 1.8\ P_2 = 2.3\ P_3$.

.

constitutes the tunnel barrier which defines our quantum dot.

Turning on $\vec{E}$

Let us now vary the gate voltages separately ($V_{left} \neq V_{right}$). Then, each Coulomb peak can be traced as a function of the voltage difference between V_{left} and V_{right}. This difference generates an electric field parallel to the nanotube which we use to intentionally push the quantum dot toward the left or right gate electrode (Figure 5.1B). As a consequence, the capacitive coupling of the quantum dot to the gate electrodes is no longer symmetric, and hence Coulomb oscillations are observed to curve when increasing the electric field (see green lines in panel A). When measured at zero bias, the curvature of the nth Coulomb oscillation is a measure of the electric polarizability of the nth electron ground state. This is because the local slope dV_{right}/dV_{left} is a proxy for the ratio between $C_{g,left}$ and $C_{g,right}$, and a change in this ratio can be understood as a lateral shift of the charge distribution. Specifically $dV_{right}/dV_{left} = -1$ means that the charge is coupled equally to the left and right gate electrodes, from which we infer that the charge sits in the center. However, $dV_{right}/dV_{left} = 0$ means that the quantum dot is independent of the left gate potential and therefore must be located at the very left. In panel C we plot the electrons' "position" (inferred from dV_{right}/dV_{left}[1]) as a function of the electric field, and define the polarizability as the change in "position" per change in electric field. For simplicity we assumed that the applied electric field is that of a parallel plate capacitor $E = (V_{right} - V_{left})/L_3$, with L_3=420 nm for this device. For the first electron (blue data points) we obtain a shift of roughly 50 nm per V/μm. The polarizabilities associated with the second

[1]Specifically, we calculate $\phi = \arctan(-dV_{right}/dV_{left})$ and assign -1 to $\phi = 0$ and +1 to $\phi = 90°$.

and third electron are smaller. We obtain $P_1 = 1.8\ P_2 = 2.3\ P_3$.

Discussion

Unlike single particle levels in a harmonic potential (which all shift in space by an equal amount when applying a constant electric field), we find that an energy level in our quantum dot is less polarizable if other electrons are present on the dot. This indicates the importance of Coulomb interactions in suspended carbon nanotubes, and that screening takes place despite the one-dimensional confinement.

Our analysis in Figure 5.1 is simplistic in that we assume a constant electric field produced by the gate electrodes. We did not take into account the thickness of the gate oxide (100 nm), which cannot be neglected compared to the trench width (420 nm). The oxide thickness has at least two implications: if it is very thick, the potential of the quantum dot will always be determined by both gate electrodes, even if the charge is localized above one gate electrode only (geometric cross talk). If the gate electrode is very thin, the effective confining potential of a single electron may not be a tilted box due to attractive interactions between the electron and image charges in the highly doped gate electrodes. This means that if no electric field is applied the electron's charge distribution may be very different from that of a particle in a "square box". It would therefore be interesting to not only study the polarizability of ground states, but also of excited states[2]. The spacing between orbital states in this device is expected to be of order $\Delta E = \hbar\nu_F\pi$ / 420 nm=4 meV. Unfortunately we could not resolve excited states (cf. the Coulomb diamonds in Figure 5.2). However, the ground states proved interesting

[2]For a particle in an infinite square well, it follows from perturbation theory that the polarizability of the first excited state is of opposite sign than that of the ground state.

when subjected to magnetic fields. This brings us to the next section.

5.3 Magnetic polarizability

The intent of this Section is simply to state that addition spectra obtained from our devices are not explained by independent-particle models and simple shell-filling schemes.

We can obtain the magnetic moments associated with few-electron ground states by applying an external magnetic field parallel[3] to the nanotube and measuring the zero-bias Coulomb oscillations along $V_g = V_{left} = V_{right}$ [1]. Figure 5.2 shows data for device #1 obtained at $T = 1.5$ K. The spacing between adjacent Coulomb oscillations is roughly 50 mV (see panel B). It is ignored in what follows. The data in panel C show several striking features: (i) a modulation in the peak height, (ii) a large magnetic moment of the 1st electron, (iii) a reversal of the magnetic moment whenever one electron is added at low magnetic fields and (iv) non-linear magnetic moments in higher magnetic fields.

Discussion

The slope of the first Coulomb peak in Figure 5.2c is 4.1 mV/T, corresponding to a magnetic moment of $\mu_{orb} \approx 2$ meV/T. Using $\mu = e\nu_F D/4$ 1.12 we infer a diameter of 10 nm, which is larger than expected for the growth conditions we used. However, there are indications that the formula overestimates the diameter of the nanotube by approximately 50% [22]. The change of peak height at a kink is expected if quantum states with different tunnel couplings are involved. This can be associated with different longitudinal modes or when changing the orbital

[3] up to an unknown growth angle of the nanotube

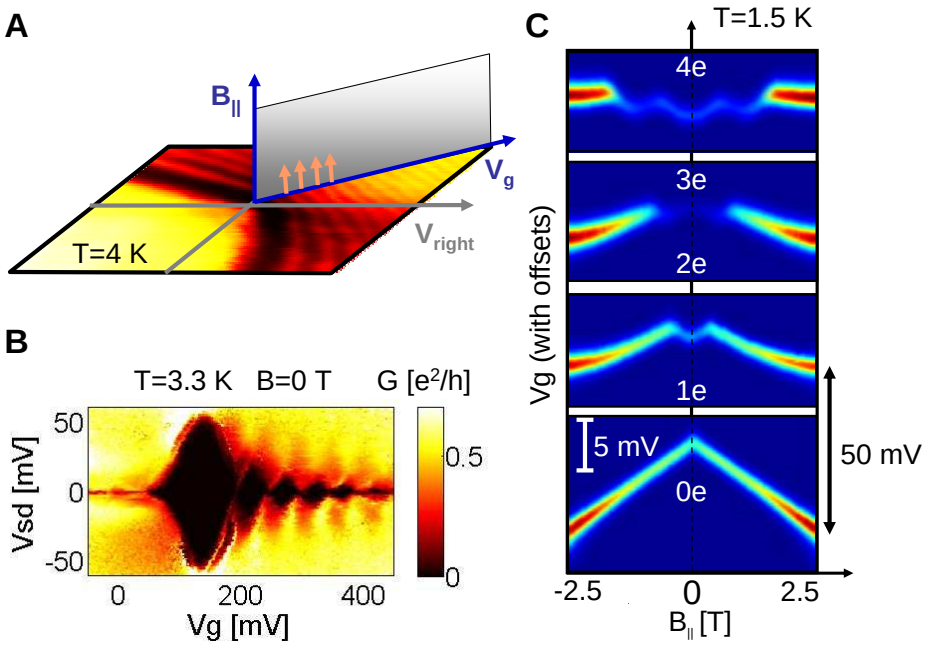


Figure 5.2: A) Limiting ourselves to the diagonal gate axis $V_g = V_{left} = V_{right}$ we measure the magnetic field dependence of the first few Coulomb oscillations of device #1. B) The first four Coulomb oscillations along V_g as a function of a magnetic field applied parallel to the nanotube axis. The spacing between Coulomb peaks is roughly 50 mV (see B). The slope of the first Coulomb oscillation in (C) is 4.1 mV/T, corresponding to a magnetic moment of ≈ 2 meV/T for the first electron. Note that at small magnetic fields the magnetic moment reverses every time an electron is added. This trend holds for all 7 Coulomb oscillations measured (not shown). A scale bar is provided in Fig. 5.4.

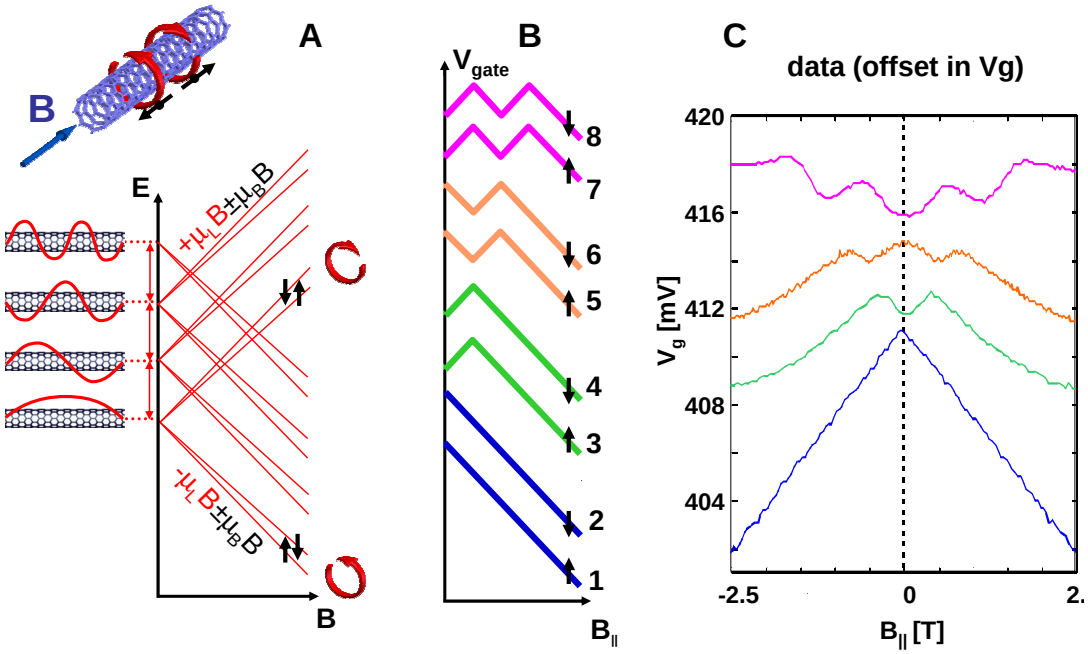


Figure 5.3: A) In a simple (incorrect) model the four-fold degeneracy of each orbit (spin-up, spin-down, clockwise, counterclockwise) is split by a magnetic field parallel to the nanotube. The splitting due to the orbital moment is much larger than the Zeeman splitting due to spin. B) Magnetic field dependence of the addition spectrum expected for the single-particle levels in panel A. At low field, the magnetic moment changes sign each time *two* electrons are added. C) Addition spectrum extracted from data in Fig. 5.2c. Note that the magnetic moment at low fields changes sign every time *one* electron is added.

direction (because the bandgaps at K and K' points differ if $B \neq 0$, see Fig. 1.3E). The increase in conductance of all four peaks at higher fields are consistent with a bandgap tuned by the magnetic field. Finally, the reduction of the magnetic moment as one goes to higher electron numbers is expected if orbits with non-zero longitudinal momentum are occupied. This is a direct consequence of the Dirac cone: particles move at constant speed v_F, and hence momentum parallel to the nanotube slows down the perpendicular velocity which is responsible for the magnetic moment.

Let us compare the data in Figure 5.2C with what we expect for four-fold degenerate shells (Figure 5.3A,B). It is convenient to label the single-particle levels of a shell by their spin (up, down) and orbital motion (clockwise, counterclockwise motion) because this is also how they split if a magnetic field parallel to the nanotube is applied. The splitting due to the orbital motion is larger than the Zeeman splitting due to spin, because of the large orbit of an electron encircling the nanotube. Therefore, we qualitatively expect an addition spectrum like the one shown in panel B. At low fields the magnetic moment is expected to reverse sign every time *two* electrons are added, in contrast to what we find in our data (panel C).

Specifically, we find that the second electron initially occupies a single particle level with opposite orbital motion. There are two plausible explanations.

First, the exchange interaction between the two electrons may align their spin parallel, preventing them from occupying the same orbital state. Moriyama *et al.* [13] reported modifications of a four-electron shell structure by exchange interaction. For two electrons at zero magnetic field they extract an exchange splitting of $J = 0.5$ meV. Oreg *et al.* theoretically found $J/\Delta \approx 0.22$ for a (10,10) nan-

otube [12] where Δ is the level spacing due to longitudinal confinement. There have been other groups which attempted to address the magnitude of exchange interaction in nanotubes experimentally ([10, 11]). Unlike our experiment, they used magnetic fields aligned perpendicular to the nanotube. Jarillo-Herrero *et al.* reported four-fold shell structure displaying singlet-triplet transitions, and the observation of the orbital Kondo effect using a magnetic field parallel to the nanotube [14, 15].

The second mechanism is spin-orbit interaction and has not yet been reported from any experimental group. It will be discussed for device #3 in Section 5.5. The direct measurement of the spin-orbit interaction in a two-electron InAs nanowire quantum dot has recently been reported from Samuelson's group [21].

Another indication that non-interacting, single-particle levels are insufficient to describe this data is the appearance of equally spaced Coulomb oscillations (Fig. 5.1a). Shell structure should show up in the addition spectrum [10] if the charging energy is not too large compared to the orbital level spacing Δ. In this sample the charging energy $E_C = e^2/2(C_G+C_S+C_D) \approx 10$ meV is just a few times larger than the level spacing ($\Delta = 4$ meV) expected from Eq. 1.10, and hence shell structure (2-fold or 4-fold) should in principle be detectable [10]. Another mystery is the size of the bandgap. From Figure 5.4 we infer that the Dirac point is crossed around 3 Tesla, suggesting a zero-field bandgap of order $\mu_{orb} \times 3$ T ≈ 6 meV. This is significantly smaller than inferred from the zero-electron Coulomb diamond in Figure 5.2b. Finally, when adding more electrons to the quantum dot and tuning the magnetic field near the Dirac point, we were able to observe artifacts arising from the left gate electrode and right gate electrode (see vertical and horizontal streaks in Fig. 5.4c). This is unfortunate because the Dirac point itself might be

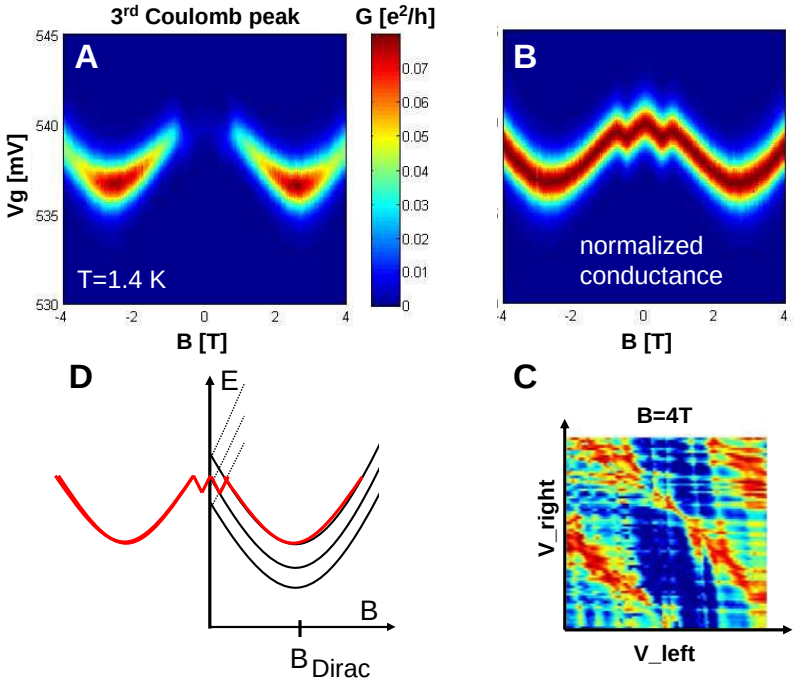


Figure 5.4: Details of the 3rd Coulomb peak of device #1. A) The peak height increases until a maximum is reached around 3T, at which point the magnetic moment reverses. B) For better visibility peak heights are normalized. D) Single particle levels expected for states with non-zero longitudinal momentum. Near the Dirac point a parabolic shape is expected. If one assumes that each single particle level is non-degenerate (except for the Kramer's degeneracy at $B = 0$ T) then the correct number of kinks is predicted for the 3rd Coulomb peak. C) At higher gate voltages and with the magnetic field tuned to the Dirac point, artifacts from the left gate (vertical streaks) and right gate (horizontal streaks) became apparent when plotting the conductance on a logarithmic scale.

interesting to study in detail. At the Dirac point, K electrons will experience a vanishing bandgap (and hence are sensitive to other potentials, like disorder on the gate electrodes) whereas K' electrons are confined by a large bandgap. One could think of this situation as two quantum dots, with one quantum dot very isolated in terms of tunnel coupling, but coupled strongly to the other dot capacitively.

Figure 5.5 shows data with similar alternation of magnetic moments, obtained for device #2 with a different trench width (1.2 μm) and different source and drain electrodes (bilayer of 5/25 nm thick Cr/Pt instead of W/Pt).

5.4 Configuration changes induced by magnetic and electric fields

How this section is organized

The configuration of a ground state is described by specifying which single-particle levels are occupied. Experimentally this is a difficult task, but the *change* of a configuration can be detected easily if it involves a change in the number of clockwise and counterclockwise moving electrons. This is accomplished by measuring changes in magnetic moments. In this section we first point out that the "kinks" in the magnetic field dependence also give rise to peculiar line shapes of the differential conductance vs gate voltage. We then provide evidence for orbital flips within a Coulomb valley, and describe the gating characteristics of device #3 in more detail. This will be helpful in understanding the data we present for the detuned two-electron dot, as well as the data we present for the one-electron dot and one-hole dot in Section 5.5.

Each kink in the magnetic field dependence of, say, the nth Coulomb peak

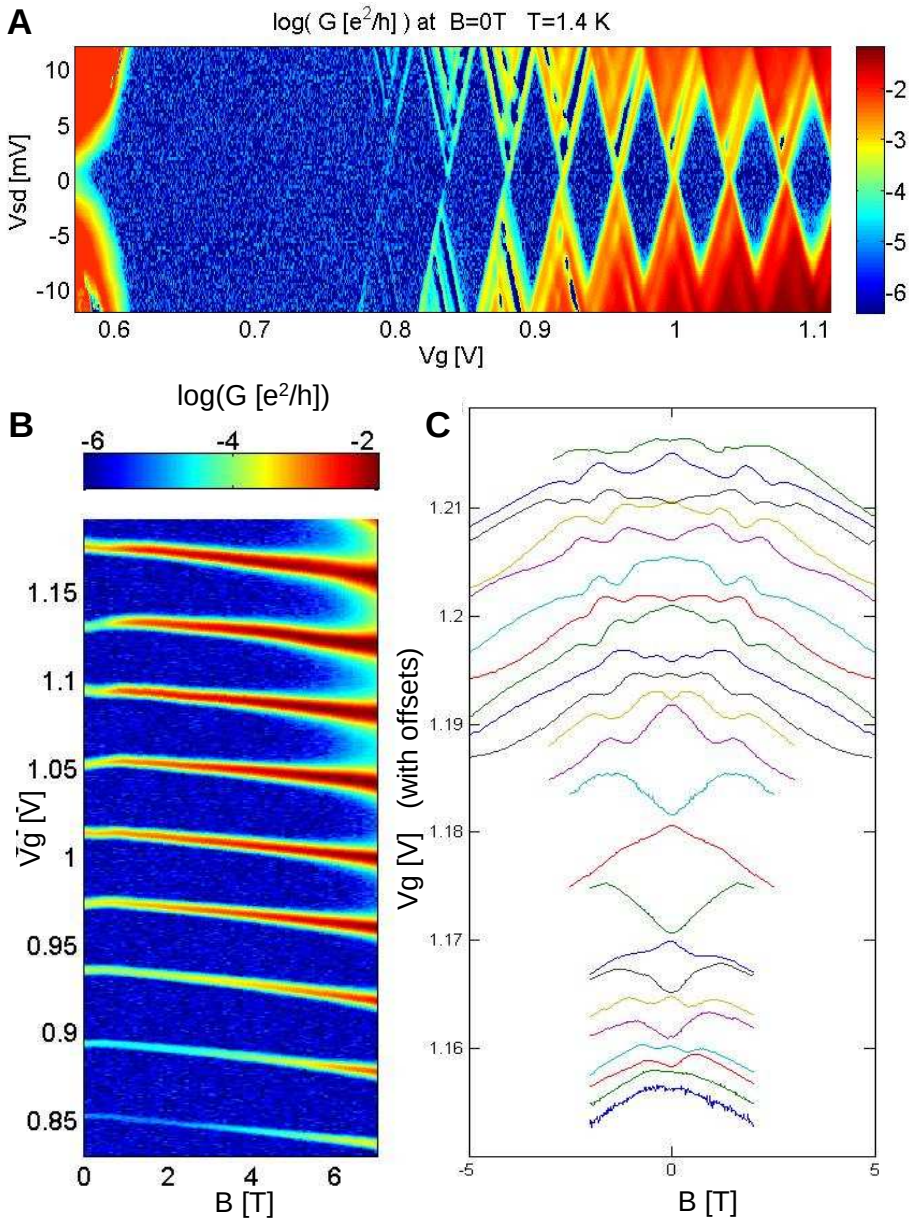


Figure 5.5: A) Coulomb diamonds of device #2 with 1.2 μm wide trenches. The total number of electrons is uncertain. We believe that the weakest Coulomb peak in panel B corresponds to adding the second electron to the quantum dot. C) Peak positions from B are extracted and offset along V_g to illustrate the alternating character of magnetic moments.

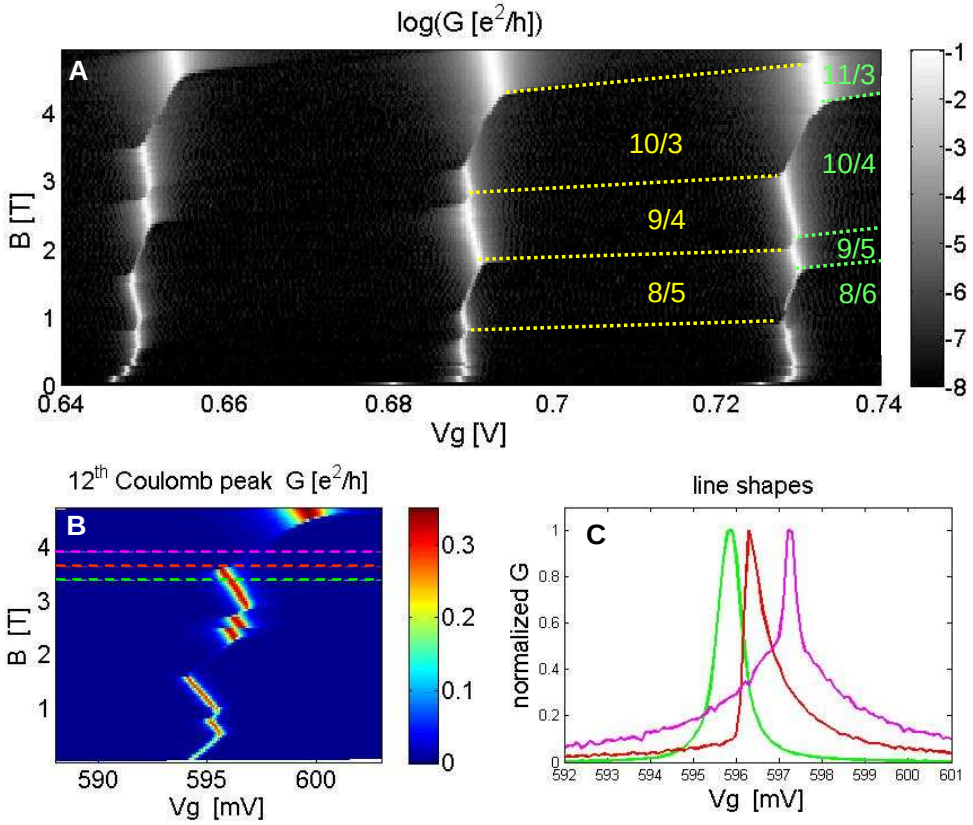


Figure 5.6: A) Ground states flips driven by a magnetic field for device #3. The left number indicates the number of K electrons, whereas the right number counts the K' electrons. Conductance is plotted using a logarithmic scale. B) Whenever the orbital direction of one electron is flipped the Coulomb peak alternates between high and low conductance. C) 3 normalized cuts taken between 3 and 4 Tesla for the 12th Coulomb peak. The asymmetric line shape (red) arises from the competition between a strongly coupled quantum state (green) and a nearby weakly coupled quantum state.

indicates a flip in the ground-state configuration of either n electrons or $n-1$. Let us first identify which one flips, and then argue that in our devices we can induce flips in the ground-state configurations not only by tuning the magnetic field, but also by tuning the gate voltages. We find that within one Coulomb valley the ground state may be flipped by changes along the gate diagonal ($V_g \approx V_{left} \approx V_{right}$), as well as by off-diagonal "detuning" ($dV \sim V_{left} - V_{right}$).

Number of kinks

If n electrons (n = even) form a ground state which is unpolarized at $B = 0$ T (in terms of orbital magnetic moments), then we expect $n/2$ flips of orbital moments when a sufficiently large external magnetic field is applied. Including the $(n-2)/2$ flips of the $n-1$ electron ground state we expect a total of $n-1$ kinks of the nth Coulomb peak. This is indeed what we observe for device #3 in Figure 5.6, even though the total number of kinks is not obvious from the data shown (see Fig. 5.7 instead). For the 14th Coulomb peak (the right peak in panel A) we have indicated how many clockwise (cw) and counterclockwise (ccw) single particle levels are occupied. We may also call them K and K' electrons. Each kink is associated with a configuration change of either 13 or 14 electrons. For example, the kink at 3 Tesla is caused by a change in the 13-electron ground state, going from 9 K electrons and 4 K' electrons to a configuration with 10 K and 3 K' electrons[4]. The kink around 4 Tesla is caused by flipping the orbital motion of one of the 14-electrons. Each time the magnetic moment is flipped, the conductance changes dramatically as evident in panel B (linear color scale). Each leg which

[4]We infer the number of K and K' electrons at a given magnetic field by assuming that at high magnetic field (not shown in Figure 5.6) all electrons are in a clockwise orbit, and then count the number of kinks backwards to infer how many electrons flipped their orbit.

moves toward higher V_g with increasing B is dark, whereas each down-moving peak is strong. This effect gets stronger with increasing field, and is explained by the different tunnel barriers felt by K and K' electrons: whereas the bandgap (and therefore the tunnel barrier) gets smaller for K electrons when increasing the magnetic field, the opposite is true for K' electrons. This leads to peculiar line shapes (panel C) in the regions where a weakly coupled ground state is close in energy to a strongly coupled resonance (like the life-time broadened green peak). The line traces in panel C were normalized by their maximum value, and the line shapes are qualitatively explained in the inset of Figure 5.7.

Peculiar line shapes

To explain the distinct line shapes, let us consider strongly coupled single particle levels near the K point (solid red lines in the inset of Fig. 5.7) and weakly coupled single particle levels near the K' point (dashed red lines). Turning on the magnetic field will lower the energy of the strongly coupled states and raise the weakly coupled states, inducing configuration changes as indicated by dots in the inset of Figure 5.7. The Coulomb peak is strong if it involves addition or removal of a K electron (black dots), and suppressed if it requires the addition or removal of a K' electron (gray dots). An asymmetric peak shape arises in regions where a weakly coupled electron flows through the device by forming an intermediate state on a strongly coupled (broadened) peak.

Flips between clockwise and counterclockwise orbits

The remainder of Figure 5.7 gives an overview over the first 10 Coulomb peaks of device #3 (the first Coulomb peak is omitted). Note that the nth Coulomb peak shows exactly n-1 kinks, indicating the the ground states at zero field have at most

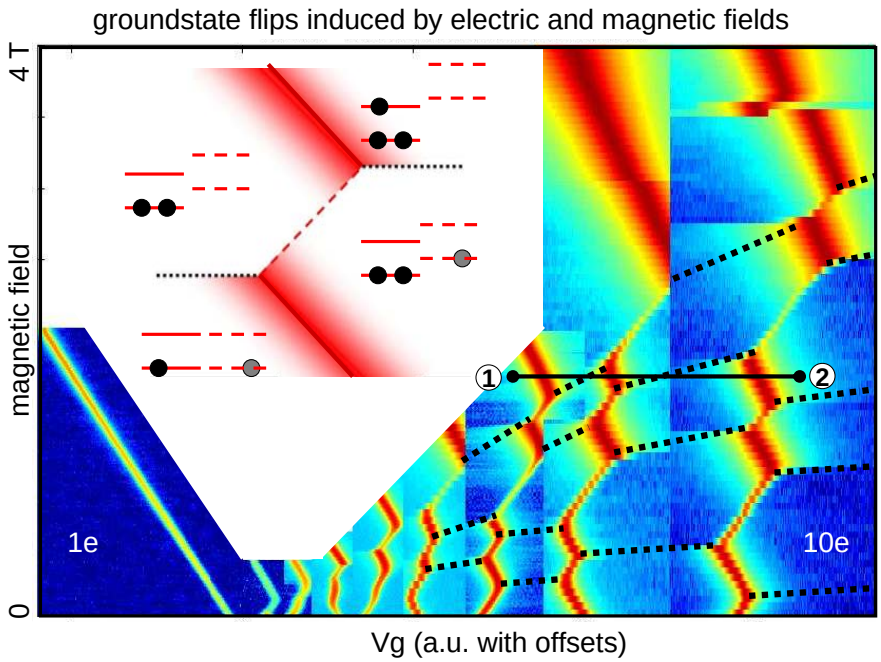


Figure 5.7: For better visibility adjacent Coulomb peaks have been moved closer toward each other along Vg, and plotted using a logarithmic color scale. Ground-state flips within a Coulomb valley are indicated by black dashed lines. The path 1→2 gives an example where a change in V_g induces a configuration change (every time the path intersects a black dashed line). Also note that the nth Coulomb peak shows $n-1$ kinks, consistent with a vanishing orbital moment at $B = 0$ T. Inset: schematic of weakly coupled (dashed red lines) and strongly coupled (solid red lines) single particle levels giving rise to bright and dark Coulomb peaks.

one orbital moment. Interestingly the kinks of adjacent Coulomb peaks occur at slightly different magnetic fields, usually at higher magnetic fields if V_g is increased. This implies that ground states can flip their magnetic moment within a Coulomb valley even if the magnetic field stays constant. This happens for example at the intersections between path 1→2 and the black dashed line, in Coulomb valley 8 and 9. These flips are induced by changing both gate voltages in the same direction (diagonal gate voltage $V_g = V_{left} = V_{right}$). From the magnitude of the changes in magnetic moment we conclude that the ground state changes its configuration in terms of its *orbital* magnetic moment. Whether those flips are associated with flips in the electronic spins as well is difficult to tell from the data. Note that Moriyama *et al.* observed "internal spin flips" in a nanotube oriented perpendicular to the magnetic field [13], and Oreg *et al.* studied spin configurations in nonuniform external potentials theoretically [12]. Next we demonstrate that an off-diagonal voltage change (i.e. creating electric field parallel to the nanotube) has dramatic effects as well.

Gating properties of device #3

The conductance of device #3 as a function of V_{left} and V_{right} is plotted in Figure 5.8a at a magnetic field lower than the Dirac point crossing. In addition to a mirror symmetry in respect to the $V_{left} = V_{right}$ axis (left-right symmetry) we observe another symmetry axis which is missing in Figure 5.1. This is the "electron-hole" symmetry defined by the $V_{left} = -V_{right}$ axis, and means that we can operate this device as a n-type or a p-type quantum dot[5]. This will be useful

[5]It is not clear to what extent the barrier which confines the quantum dot is "electron-hole" symmetric. However, this device has a larger bandgap than the device in Figure 5.1, and hence the Fermi level of source and drain electrode may be located somewhere in the intrinsic region of the nanotube. In that case, it is

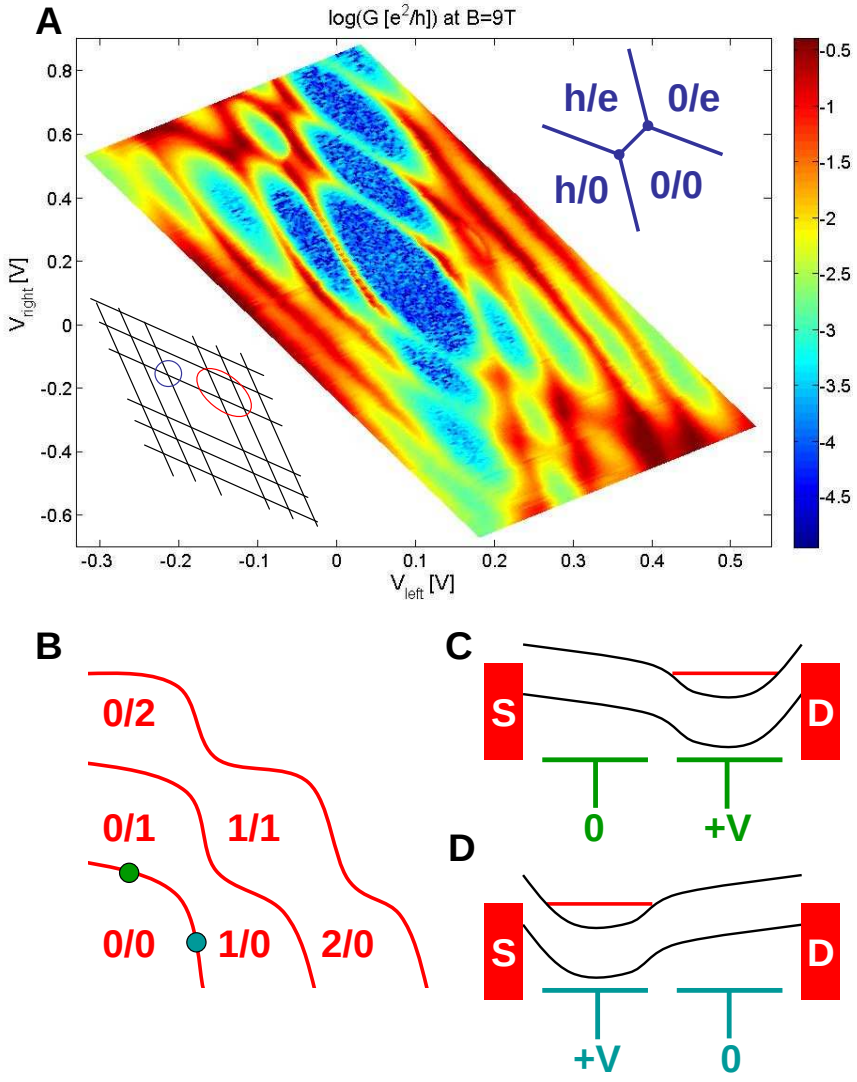


Figure 5.8: A) The pattern of Coulomb oscillations of device #3 can be understood in terms of a right and left quantum dot (black lines of inset). The blue circle corresponds to an electron dot on the right and a hole dot on the left separated by a pn-tunnel barrier. Hence the data shows honeycomb features at the top left. In the region marked with red however, the left charge combines with the right charge to form a single quantum dot. The stability diagram of that "molecular dot" looks like B, where 0/1 denotes "1 electron predominantly at the right" (C) and 1/0 denotes "1 electron predominantly at the left" (D).

in section 5.5 where we compare the shell splitting of one electron with that of one hole. The pattern of Coulomb oscillations can be understood as arising from two coupled quantum dots in series (see black lines of the inset), with one Coulomb valley larger than the others due the additional energy cost of the bandgap. The finite slope of the black lines arises from the cross-coupling between the right gate and the potential above the left gate electrode and vice versa[6]. The intersection between the first electron peak and first hole peak (encircled in blue) corresponds to a hole on the left side, separated by a pn-tunnel barrier from an electron on the right side. Accordingly, the data shows the familiar honeycomb structure of a double-dot (cf. the triple-point schematic in the blue inset, panel A).

In what follows we shall focus not on the blue region, but on the region marked in red (black inset of panel A) which corresponds to a n-type quantum dot extended from the left to the right. The strong tunnel coupling needed to turn it into a single dot is provided by the nanotube itself. There is no pn-barrier between the "left dot" and the "right dot" and hence the stability diagram schematically looks like panel B. Here, the numbers 0/1 indicate a region where the potential of the gate electrodes is such that the first electron is located predominantly on the right side (panel C). The transition to the 1/0 state of course is smooth. This transition corresponds to the polarizability experiments in Figure 5.1, except now the geometry is far from that of a parallel plate capacitor. Each Coulomb oscillation can be probed at many different locations, and in Section 5.5 we will study the first electron at two different detunings, indicated with green and cyan dots in panel B.

---

conceivable that a Schottky barrier confines both the n-type quantum dot as well as the p-type quantum dot.

[6]In this device the oxide thickness (100 nm) is comparable to the trench width ($\approx$ 135 nm)

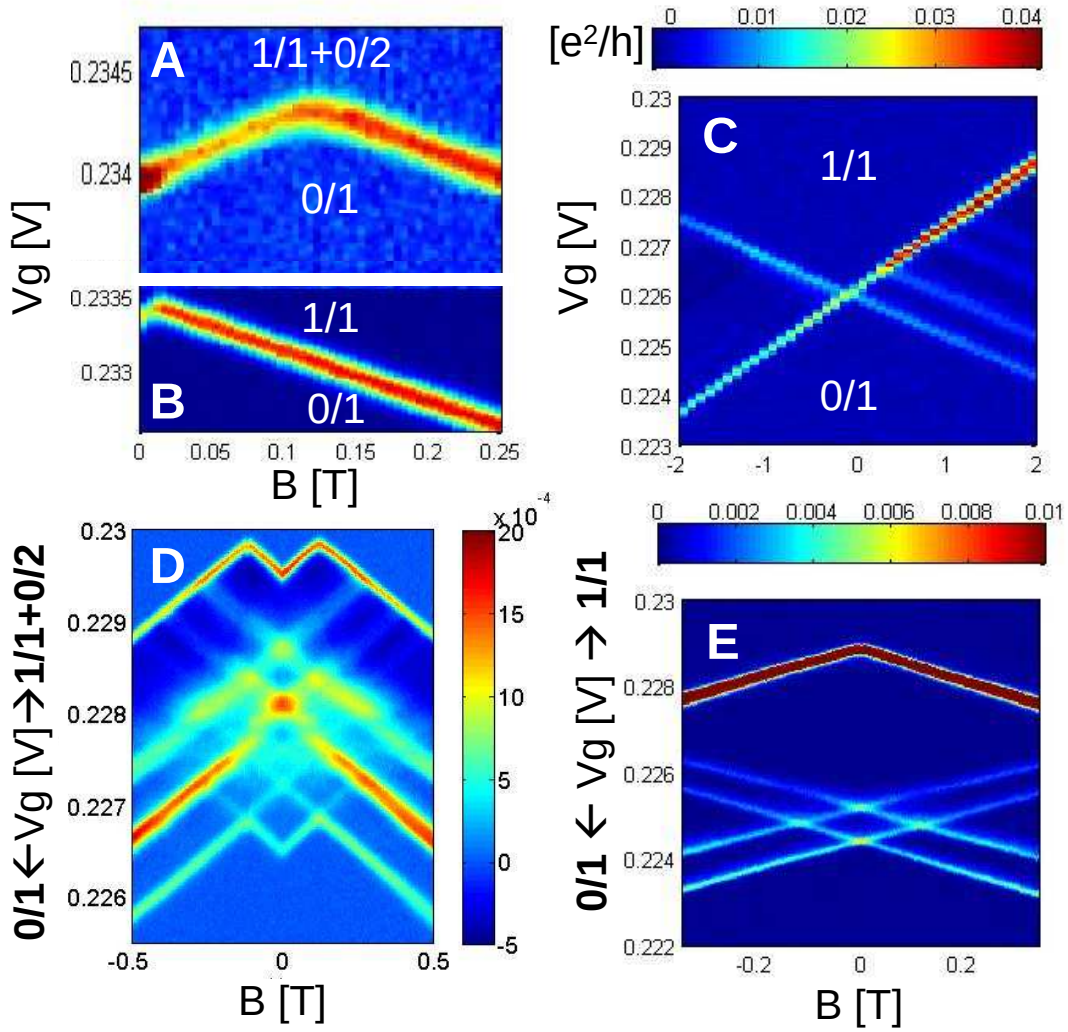


Figure 5.9: The second Coulomb peaks mimics the first Coulomb peak of Figure 5.11C if the quantum dot is tuned to the (1/1) configuration (B,C,E). If detuned toward the (0/2) configuration it looks very different (A,D) due to the Pauli exclusion principle and the presence of a nearby empty shell. Data in C was obtained at $B = 0.3$ T.

Detuning the second Coulomb peak

Let us briefly look at the second Coulomb oscillation at two different detunings: going from 0/1 to 1/1 (Figure 5.9B) and going from 0/1 to somewhere in between 1/1 and 0/2 (Figure 5.9A). What we find is a very different magnetic field dependence of the two ground states. Physically, this means that the ground-state configuration at $B = 50$ mT (for example) depends on the detuning of the quantum dot. We believe that for the detuning shown in A both electrons are occupying the same Kramer's doublet due to spin-orbit interaction ("antiferromagnetic coupling"), whereas for the detuning in panel B each electron occupies a different shell (with spins and orbital moments aligned parallel). Electrostatically speaking this is energetically possible because detuning is such that the 1st electron is predominantly at the right, while the second electron is predominantly on the left (0/1 to 1/1 transition). Indeed, the two-electron excitations measured at this detuning (shown for $B = 0.3$ T in panel C) look just like the excitations of a single electron (cf. Figure 5.11C), whereas many more excitations appear in the diamond corresponding to the detuning of panel A (not shown). For reference, the magnetic field dependence of excitation spectra are shown in panel D and E, where we note again the similarity between the two-electron spectrum in E and the one-electron spectrum in Figure 5.12 below. The lesson we learn from Figure 5.9 is that one has to be cautious when interpreting data if details about the electrostatic confinement are not known.

Exchange or Spin-orbit interaction?

We conclude this section by presenting the magnetic field dependence obtained for the first four electrons and the first two holes (device #3). To avoid complexity

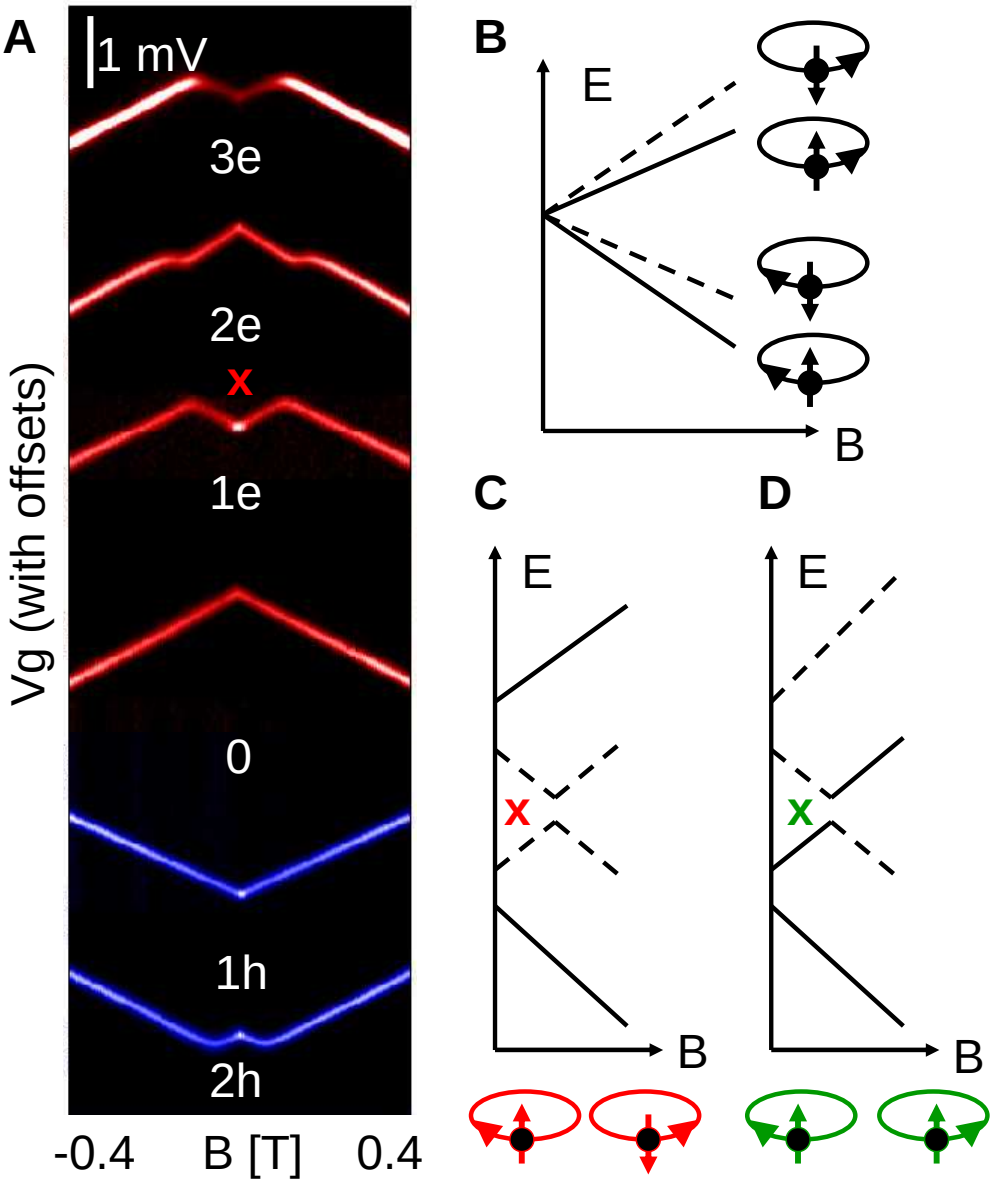


Figure 5.10: A)Addition spectrum obtained for the first 4 electrons and first 2 holes. B) Schematic spin assignment of the single particle levels in a shell. Solid lines indicate spin up, and dashed lines indicate spin down. The two-electron ground state marked with 'x' can be explained by spin-orbit interaction (C) or by exchange interaction (D).

when other shells come into play (cf. Figure 5.9) we show data in Figure 5.10A obtained for detunings (-2/0)→(-1/0)→(0/0)→(0/1)→(0/2)→(0/3)→(0/4). Of course, the next higher shell has to appear for high enough fields because each clockwise orbit can maximally accommodate 2 electrons, and hence the outermost kinks of the 3rd and 4th Coulomb peak are due to a different shell. Ignoring this minor detail let us discuss the two-electron configuration marked with a red cross.

From the slope of the second Coulomb peak we know that a clockwise and counterclockwise single particle level are occupied at small fields, and hence the generic four-fold degenerate shell depicted in panel B is not filled as in an non-interacting model (solid lines indicate spin up, whereas dashed lines indicate spin down). Exchange interaction as well as spin-orbit interaction both lead to qualitatively similar addition spectra but with different electron configurations. Panel C illustrates the addition spectrum of the two-electron configuration if spin-orbit coupling is causing the shell splitting (with both spins anti-parallel), whereas panel D illustrates the case of a ferromagnetic exchange coupling. To experimentally tell the difference between both scenarios, one can attempt to infer the spin magnetic moment from the slightly different slopes associated with the dashed and solid lines. However, the easiest way to test for exchange interaction is to remove the second electron altogether. This is the topic of the next section. Similar even-odd effects in the magnetic field dependence of addition spectra have been reported from various groups [10, 13, 17–19], as well as in multi-wall carbon nanotubes [20]. However, in most experiments the magnetic field was aligned perpendicular to the nanotube, presumably with a small angle of misaligned. Because of the large orbital magnetic moments in nanotubes, a small angle of misalignment may complicate the interpretation of such data.

5.5 Shell splitting in zero magnetic field

One electron in the dot

In Figure 5.11 we present excitation spectra obtained from the first Coulomb oscillation of device #3. Panel A shows a low lying excitation at 0.4 meV if no magnetic field is applied. Note that this resonance is an excitation of the first electron, and no excitations of the zero-electron state are present (the bandgap at $B = 0$ T is tens of meV). This data was obtained at a (0/1)-detuning, as indicated by the green dot in Figure 5.8B,C. As a consequence of this detuning the rate limiting barrier is on the side of the source electrode, consistent with the asymmetry in Figure 5.11A. We can exclude the possibility that this excitation results from artifacts in the source electrode by detuning the quantum dot toward the left (where we find exactly the same excitation). Excitation spectra obtained from the (0/1) and (1/0) configurations at finite magnetic field are compared in panel B and C. Clearly the rate limiting barrier in C is between the quantum dot and the drain electrode, and inspection of the slopes shows that the capacitance ratio between source and drain reverses as well. We find $C_S/C_D = 1/1.7$ (B) and $C_S/C_D = 2.2$ (C). The effective gate capacitance stays constant, $C_G = 1.3 \times (C_S + C_D)$ in B and $C_G = 1.2 \times (C_S + C_D)$ in C. Most importantly there are additional excitations at finite magnetic field, which we study in detail in Figure 5.12. We do this by measuring the magnetic field dependence of a cut at constant $V_{SD} = -2.2$ mV, as indicated by the white dashed line in Figure 5.11A.

Spin assignment

Figure 5.12A indicates that the ground state as well as the 0.4 meV excitation

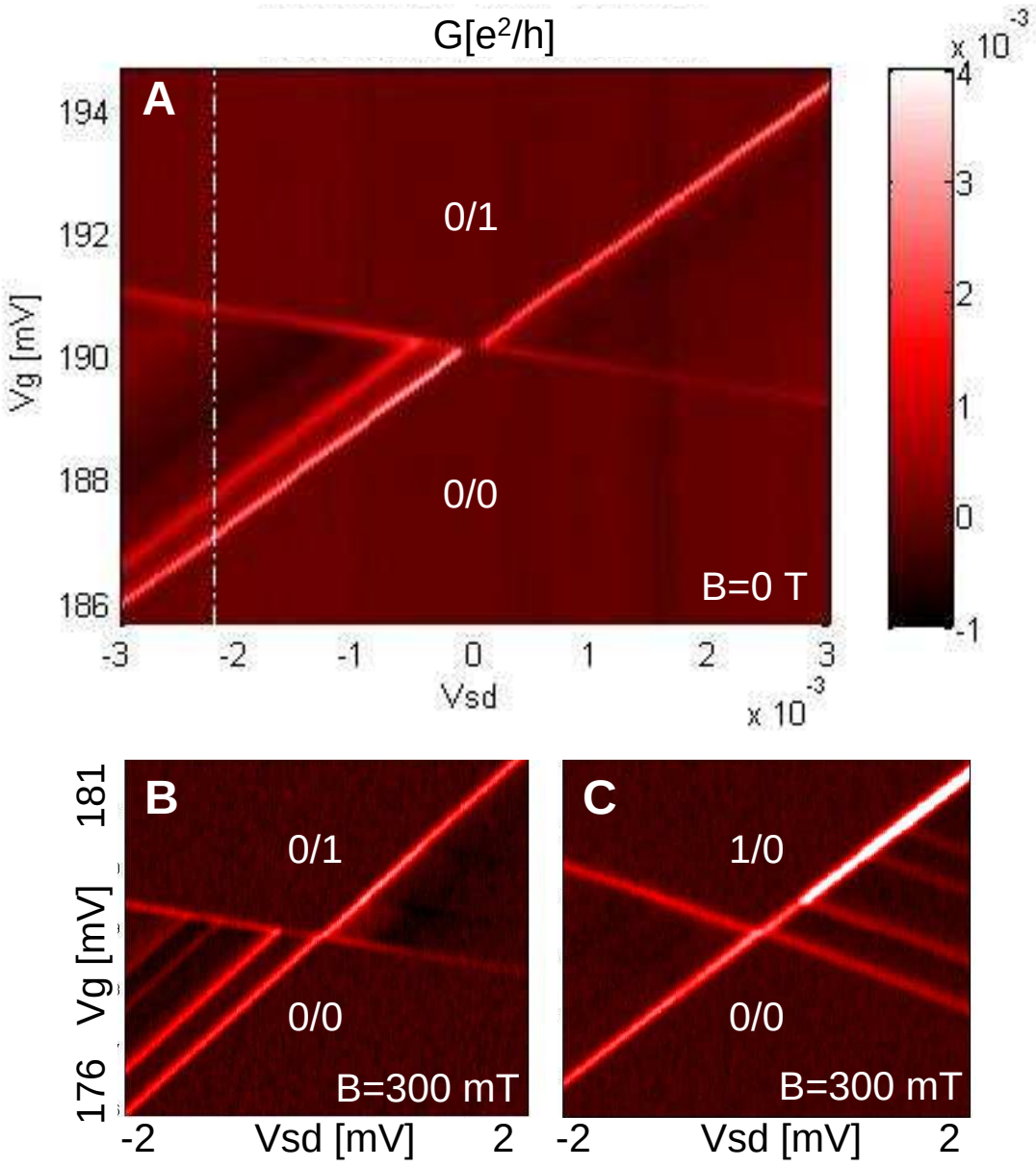


Figure 5.11: A) At zero field an excitation at 0.4 meV is observed for one electron in the dot. The ground state and excitation each split into two resonances when applying magnetic field parallel to the nanotube. This finding does not depend on detuning (B and C).

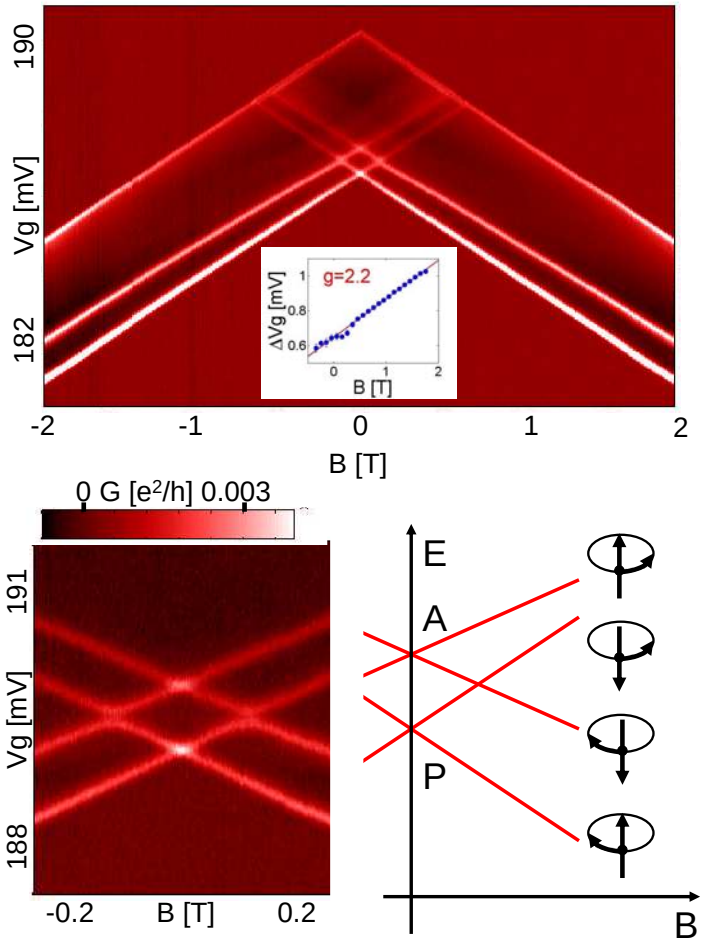


Figure 5.12: Magnetic field dependence of the one-electron dot. A) Each Kramer's doublet splits into a down moving resonance and an up-moving resonance. The splitting between the two down moving resonances increases with a g-factor of 2.2. B) Close up of the shell splitting. Note the avoided level crossing. C) Spin assignments based on the Zeeman splitting in A. Note that the ground state is a cw electron with spin up, i.e. the orbital magnetic moment is parallel to the spin and the magnetic field.

splits when applying a magnetic field parallel to the nanotube. The two resonances, which move down in energy, become stronger and move further apart with increasing magnetic field whereas. The time-reversed partners however become weaker and move closer to each other. The splitting as a function of magnetic field is plotted in the inset of panel A. Using capacitances extracted from Figure 5.11A we can convert the splitting into energy and obtain an effective g-factor of 2.2 ($\pm$10%). This corresponds very closely to the Zeeman splitting of an electronic spin, and hence we assign spins and orbits to all four states as indicated in panel C. In particular, we find that the ground state of the first electron has its spin magnetic moment aligned parallel to the orbital magnetic moment. Time reversal symmetry reverses the spin as well as the orbital magnetic moment, and hence at zero field we mark the lower Kramer's doublet by P ("parallel") and the excitation by A ("antiparallel"). This is an experimental statement which relates spin and orbital magnetic moments. At zero field the splitting between A and P states is 0.65 meV, and the orbital magnetic moments are $\pm$1.58 meV/T.

Observation of K-K' scattering

By carefully measuring the one-electron shell splitting at low fields we observe an additional interaction, evident as a small avoided level crossing at $B = 120$ mT (see Figure 5.12B). According to the spin assignments in panel C the two states involved have parallel spin and antiparallel orbital motions. This suggests that the physical origin of the avoided level crossing is K-K' scattering. A direct consequence of this assignment is the prediction that the two states which converge at high fields should eventually cross opposed to anti-cross (because they possess opposite spins). Experimentally we could not test this prediction for the first

electron, but we did confirm the crossing for the 1st hole (this crossing is partially visible in Figure 5.13C). Let us therefore turn to the same measurements performed on the first hole Coulomb peak (Fig. 5.13).

One hole in the dot

The diamond obtained at $B = 0$ T (panel A in Figure 5.13) looks identical to the one of the first electron, except the low lying excitation is now an excitation of the (-1/0) charge state. Also the excitation is somewhat lower in energy (0.35 meV) compared to the first electron (0.65 meV). The shell splitting measured at low magnetic fields again shows avoided level crossing (panel B). At higher magnetic fields (panel C) we observe that the splitting between ground state and excited state decreases, with the same g-factor as observed for the first electron (g=2.2). The orbital magnetic moment extracted from panel C is 1.55 meV/T. The spin assignment for the shell in B — if referenced to the electron — is identical to the one in Figure 5.12C. If one prefers to reference the spin assignments in respect to the hole, one has to flip the spin and charge, but not the orbital motion. This is because a full shell does not carry current nor spin angular momentum, and hence the orbital magnetic moment of hole ground state is aligned with the magnetic field, whereas its spin magnetic moment is aligned antiparallel[7].

Breaking of electron-hole symmetry

Jarillo-Herrero *et al.* reported electron-hole symmetry in semiconducting carbon nanotube quantum dots by comparing the excitation spectra associated with

[7]In this thesis, when we specify spin, we usually refer to the spin magnetic moment and not the spin angular momentum. Spin angular momentum and spin magnetic moment are antiparallel for both electrons and holes.

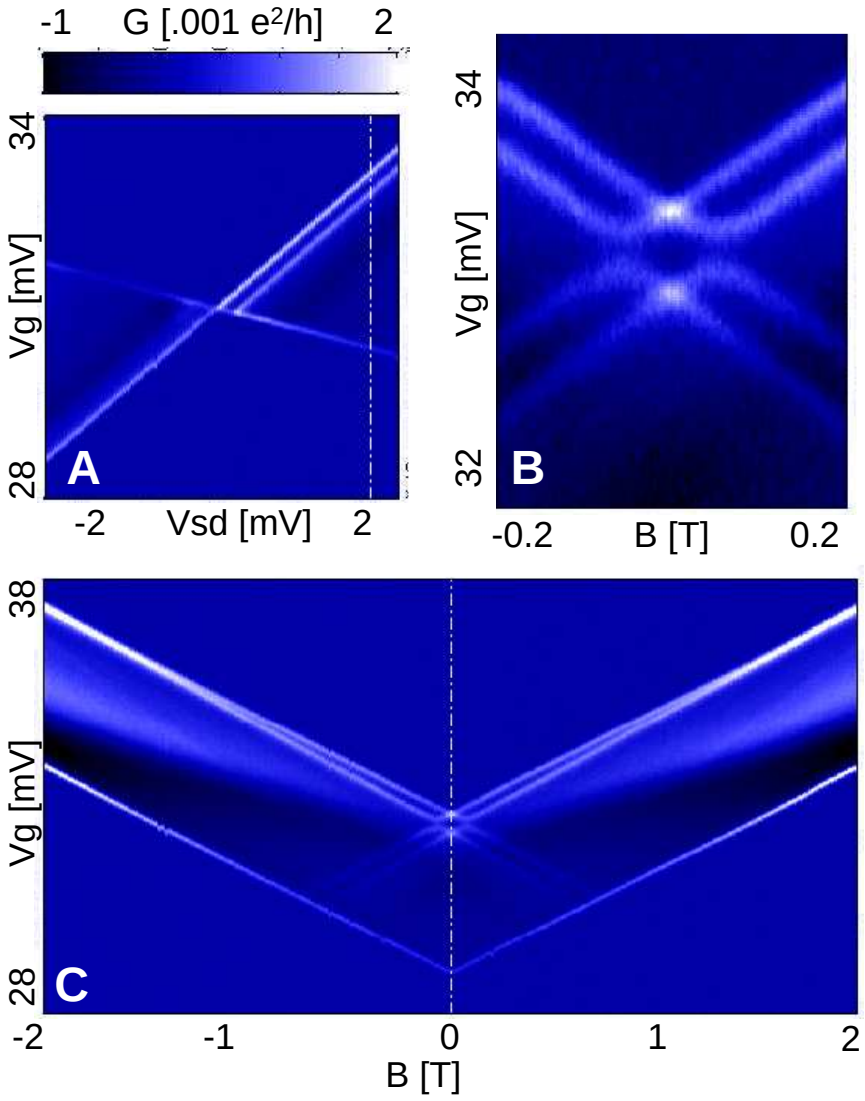


Figure 5.13: Magnetic field dependence of the one-hole quantum dot. A) Shell splitting at zero magnetic field. A cut at $V_{sd} = 2.1$ mV is plotted as a function of the magnetic field in C. B) Avoided level crossing due to K-K' scattering.

one, two, three electrons with those of one, two, three holes (at B=0T) [16]. However, the shell splitting we observe for the 1st hole (Fig. 5.13C) is not the same as if one reverses the energy axis of the 1st electron (Fig. 5.12A). We can phrase this observation concisely as follows. For a magnetic field aligned parallel to the nanotube, we find that the ground state of the first electron possesses an orbital magnetic moment parallel to the magnetic field and a spin magnetic momentum *parallel* to the magnetic field. The ground state of the first hole however possesses a orbital magnetic moment parallel to the magnetic field and a spin magnetic momentum aligned *antiparallel* to the magnetic field[8].

Note that the electronic structure of graphene is not electron-hole symmetric. After all, it is made up of light electrons and heavy protons, giving rise to energy bands which lack electron-hole symmetry. Electron-hole symmetry is only found in the π-band, and there it is localized near the K and K' points. Therefore the breaking of electron-hole symmetry near the K and K' points due to spin-orbit interaction can in some sense be traced back to the mixing with the non-symmetric σ-bands.

5.6 Outlook

We have provided a glimpse into the possibilities offered by our devices, for studying interactions among electrons in carbon nanotubes. We have seen that the excitations of a single electron confined in a dot are already very complex, when measured in an external magnetic field. Additionally, our devices allow an electric field parallel to the nanotube to be tuned, giving the opportunity to study polariz-

[8]This follows from the fact that there is no net current or spin if the nanotube is charge neutral. [3]

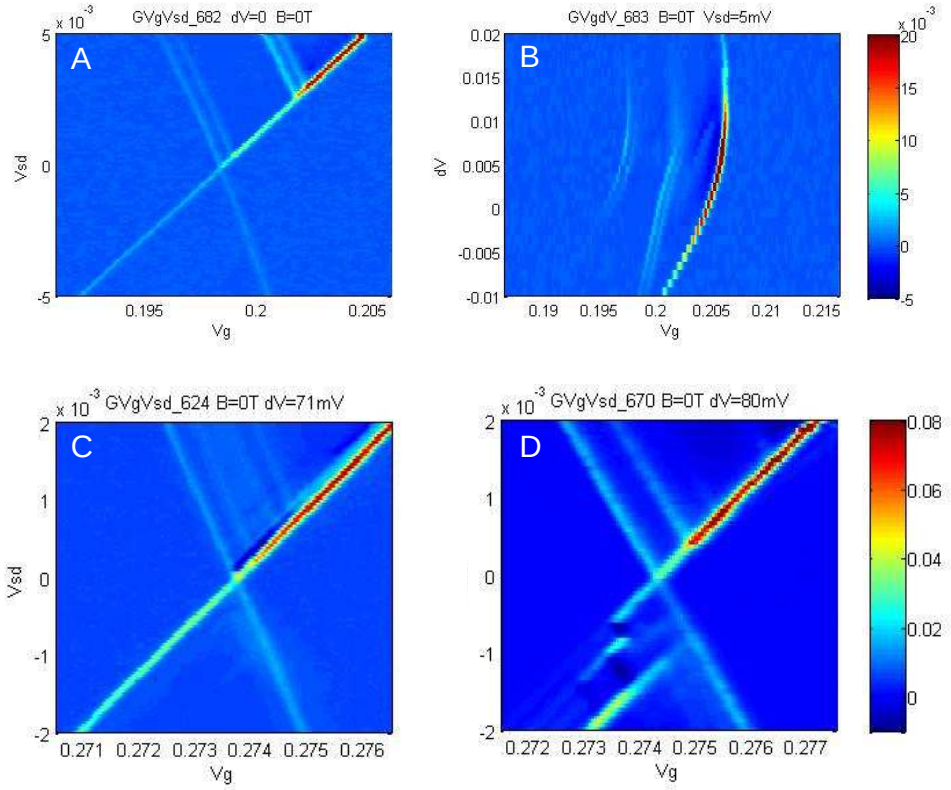


Figure 5.14: Our devices allow polarizabilities of ground states and excited states to be studied: the first excited shell of a one-electron quantum dot (visible in panel A at 3 meV, $B = 0$) shows opposite curvature (polarizability) when measured as a function of the detuning dV (panel B). C and D) Excitation spectra of the third Coulomb oscillation obtained at different detunings, showing negative differential-conductance features which we have not studied in detail yet.

abilities of ground states as well as excited states. Preliminary results for a single electron is shown in Figure 5.14A and B, where the bias window is large enough to observe the next higher shell. (At zero field this shell is split as well due to spin-orbit interaction; see the double excitation at approximately 3 meV in panel A). The different response of the ground state and the next higher shell to electric fields can be studied by changing the off-diagonal voltage dV(B). We observe that the excited state possesses curvature (polarizability) opposite to the ground state.

Another striking feature we observed is negative-differential conductance in the spectrum of the third Coulomb peak, with a strong dependence on the detuning parameter dV(panel C and D) and magnetic field (not shown). We believe that these features are caused by occupation of weakly coupled states, but in principle our devices should be suited to study blockade effects [4–9] due to spin or orbital selection rules as well. An interesting regime may be to study transport through a p-type quantum dot (on one gate electrode) in series with a n-type quantum dot (on the other gate electrode).

One regime which we have not yet explored is that of high electric fields. We typically limited gate voltages to 1 Volt or less, but in principle the gate oxide should be compatible with tens of Volts. This opens up experiments with semiconducting nanotubes by gating them in such a way that transport through suspended pn-junctions occurs. The width of the pn-junction as well as its location along the nanotube can then be controlled by the gate potentials. This feature may also be useful for studying electro-optical properties of nanotubes.

Lastly, our devices should be useful for studying and quantifying the effects of exchange interaction in nanotubes, by making use of the independent gate electrodes to change the number of electrons as well as the confining potential.

REFERENCES

[1] E. D. Minot *et al.*, Nature **428**, 536 (2004).

[2] D. Huertas-Hernando *et al.*, Phys. Rev. B **74**, 155426 (2006).

[3] N. Ashcroft and D. Mermin, *Solid State Physics*, Harcourt College Publishing, 1976, p. 225f

[4] B. Wunsch *et al.*, Physica E **26**, 464 (2005).

[5] D. Jacob *et al.*, Phys. Rev. B **79**, 081314 (2004).

[6] A. C. Johnson *et al.*, Phys. Rev. B **72**, 165308 (2005).

[7] *et al.*, Phys. Rev. Lett. **71**, 4019 (1993).

[8] A. K. Hüttel *et al.*, Europhys. Lett. **62**, 712 (2003).

[9] D. Weinmann *et al.*, Phys. Rev. Lett. **74**, 984 (1995).

[10] A. Makarovski *et al.*, Phys. Rev. B **74**, 155431 (2006).

[11] D. H. Cobden *et al.*, Phys. Rev. Lett. **89**, 046803 (2002).

[12] Y. Oreg *et al.*, Phys. Rev. Lett. **85**, 365 (2000).

[13] S. Moriyama *et al.*, Phys. Rev. Lett. **94**, 186806 (2005).

[14] P. Jarillo-Herrero *et al.*, Nature **434**, 484 (2005).

[15] P. Jarillo-Herrero *et al.*, Phys. Rev. Lett. **94**, 156802 (2005).

[16] P. Jarillo-Herrero *et al.*, Nature **429**, 389 (2004).

[17] D. H. Cobden *et al.*, Phys. Rev. Lett. **81**, 681 (1998).

[18] S. J. Tans *et al.*, Nature **394**, 761 (1998).

[19] H. W. Ch. Postma *et al.*, J. Low Temp. Phys. **118**,495 (2000).

[20] M. R. Buitelaar *et al.*, Phys. Rev. Lett. **88**, 156801 (2002).

[21] C. Fasth *et al.*, Phys. Rev. Lett. **98**, 266801 (2007).

[22] E. Minot, *Tuning The Band Structure of Carbon Nanotubes*, Ph.D. thesis, Cornell University, 2004.

[23] J. Cao *et al.*, Nature Materials **4**, 745 (2005).

[24] S. Sapmaz *et al.*, Phys. Rev. B **71**, 153402 (2005).

[25] W. Liang *et al.*, Phys. Rev. Lett. **88**, 126801 (2002).