

Electrical Properties  
of Atomic-Sized Metal Contacts

Kim Hansen

# Preface

This thesis is submitted to the Faculty of Science at the University of Aarhus, Denmark, in order to fulfill the requirements for obtaining the PhD degree in physics. The studies have been carried out in the scanning tunneling microscopy (STM) laboratory at the Institute of Physics and Astronomy from August 1995 to March 2000.

## Acknowledgments

It has been a real pleasure and a privilege to be a part of the STM group in Aarhus. The group thrives tremendously from a fruitful interplay between the three founding fathers Flemming Besenbacher, Ivan Stensgaard, and Erik Lægsgaard. My thanks extend to all of them. I admire Flemming for his drive, enthusiasm and visionary thoughts. His ceaseless encouragement and suggestions made me carry on, when things looked bad. Discussing physics with Ivan has always been a pleasure, and he has given me much well-considered advice. I appreciate his pragmatic approach to and long-standing experience in experimental surface science. Erik's ingenuity has surprised me over and over again. He is the person who can make elegant and easy solutions that actually *work*. Erik originally came up with the idea for the relay experiments presented in Chapter 3, and he has taught me most of what I know about electronics and programming.

Throughout the years it has been a great pleasure to collaborate with Mads Brandbyge, who is now at the Technical University of Denmark. Mads has always been extremely helpful whenever I have had questions regarding the theory of conduction. Several of the figures appearing in Chapter 2, which deals with this subject, have kindly been provided by him.

Following the relay studies, which were published in 1997 and form the basis of Chapter 3, Mads finally convinced me that it would be a good idea to investi-

gate the electrical properties of atomic-sized contacts at finite bias. This advice turned out to be most useful, and it triggered the work presented in Chapters 4–7. Mads has also contributed with numerical calculations of current-voltage curves in Chapter 5, thereby greatly strengthening the conclusions drawn in that Chapter. Some of the experimental results presented in Chapter 5 also initiated the investigation of the parabolic barrier model presented in Chapter 6. The first draft of this model was developed by Mads whereafter I ‘added the details’. I would also like to thank Mads for carefully reading most Chapters in this thesis. His corrections, comments, and suggestions have been invaluable.

Søren Kynde Nielsen joined my project in 1998 as a PhD student and helped completing the successful implementation of the setup described in Chapter 4. A significant part of the experimental results presented in Chapters 5 and 7 have skillfully been acquired by Søren, and he has also contributed to the data analysis. Ebbe Kruse Vestergaard is acknowledged for help with acquiring and analyzing some of the results presented in Chapter 7.

The professional assistance of Finn Sjøberg has been crucial during the build-up of the setup described in Chapter 4. He has had an important role in designing and constructing the fast current-voltage converter. This device has been vital in the experiments leading to the results presented in Chapters 5 and 7.

I have enjoyed the collaboration with the theory group at the Technical University of Denmark headed by Jens Nørskov. Especially, I have been inspired by discussions with Karsten W. Jacobsen, Jacob Schiøtz, Mads R. Sørensen, and Sune R. Bahn. Numerous people have passed through the STM group in the period during which I have been part of it. I would like to thank all of them. They have all contributed scientifically and socially to the unique atmosphere in the group.

During my studies, I have had many fruitful discussions and open-minded communications with other researchers in the field. These include J. Abellán, N. Agraït, G. R. Bollinger, A. M. Bratkovski, H. E. van den Brom, J. L. Costa-Krämer, U. Dürig, N. García, F. Grey, J. Hoekstra, S. P. Jarvis, M. Jonson, T. Junno, K. von Klitzing, B. Ludoph, U. Landman, J. Lunney, L. Montelius, H. Ohnishi, H. Ohshima, H. Olin, J. I. Pascual, C. Quate, J. M. van Ruitenbeek, J. J. Sáenz, A. Sakai, P. A. Serena, A. P. Sutton, K. Takayanagi, H. J. Tao, T. N. Todorov, J. A. Torres, J. Versluijs, and G. M. Whitesides.

I am indebted to Stig Helveg for his hospitality and for sharing many recent fast food meals. Also thanks to Lars Petersen and Jane H. Larsen for their useful comments to Chapters 2 and 6, respectively. Our librarian, Liv Vego, has been most helpful in providing me with numerous obscure and old references

published in journals I had never even heard of before. Also, I am grateful to Lone Niedziella for correcting the thesis for numerous linguistic errors and typos. Last, but not least, I am deeply indebted to Rikke, my wife, for her patience and support.

The STM group is part of the Center for Atomic-scale Materials Physics (CAMP), funded by the Danish National Research Foundation. Further funding has been obtained from the Danish Research Councils through the Center for Nanotribology, and through the Knud Højgaard Foundation.

Kim Hansen  
Aarhus, March 2000

Apart from a few corrected misprints and minor errors, this reprint is identical to the first version.

Kim Hansen  
Aarhus, May 2000



# List of Publications

In the thesis, the references listed below are referred to with capital roman numerals, whereas other references are cited using arabic numerals.

## Publications Related to Thesis

- [I] *Metallic nanowires: Formation and quantized conductance*,  
L. Olesen, K. Hansen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher,  
in *Nanowires* edited by P. A. Serena and N. García (Kluwer, Dordrecht,  
1997), pp. 191–210.
- [II] *Quantized conductance in relays*,  
K. Hansen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, *Phys. Rev.*  
*B* **56**, 2208–2220 (1997).
- [III] *Fast and accurate current-voltage curves of metallic quantum point con-*  
*tacts*,  
K. Hansen, S. K. Nielsen, E. Lægsgaard, I. Stensgaard, and F. Besen-  
bacher, *Rev. Sci. Instr.* **71**, 1793–1803 (2000).
- [IV] *Current-voltage curves of gold quantum point contacts revisited*,  
K. Hansen, S. K. Nielsen, M. Brandbyge, E. Lægsgaard, I. Stensgaard,  
and F. Besenbacher, Submitted to *Appl. Phys. Lett.*
- [V] *Current-voltage curves of gold quantum point contacts*,  
K. Hansen, S. K. Nielsen, M. Brandbyge, E. Lægsgaard, I. Stensgaard,  
and F. Besenbacher, To be submitted.

[VI] *Current-voltage characteristics of ultrathin insulating films: A simple parabolic barrier model,*

K. Hansen and M. Brandbyge, To be submitted.

[VII] *Current-induced disruptions of monatomic gold contacts,*

K. Hansen, S. K. Nielsen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, In preparation.

## **Publications in Danish**

[VIII] *Nanotråde og kvantemekanik: Hvad sker der, når man slukker for lyset?,*

K. Hansen, Kontakt Nr. 1, 1999.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Interconnects . . . . .	1
1.2	Moore and more transistors in the future? . . . . .	2
1.3	Motivation . . . . .	3
1.4	Structure of the thesis . . . . .	4
	References . . . . .	6
<b>2</b>	<b>Theory of Conductance</b>	<b>7</b>
2.1	Transport regimes . . . . .	8
2.2	Diffuse electron transport . . . . .	9
2.3	Ballistic electron transport . . . . .	14
2.4	The Landauer-Büttiker formalism . . . . .	19
2.5	The adiabatic approximation . . . . .	23
2.6	Quantized conductance . . . . .	25
2.7	Conductance eigenchannels . . . . .	26
2.8	Eigenchannels in experiments . . . . .	27
2.9	Realistic models of atomic-sized contacts . . . . .	30
2.10	Summary . . . . .	34
	References . . . . .	36
<b>3</b>	<b>Quantized Conductance in Relays</b>	<b>45</b>
3.1	Why study relays? . . . . .	46
3.2	Experimental details . . . . .	46
3.3	Results . . . . .	57
3.4	Discussion . . . . .	61
3.5	The resistance of the leads . . . . .	67
3.6	Conclusion . . . . .	71

3.7	Recent developements . . . . .	72
	References . . . . .	76
<b>4</b>	<b>Fast and Accurate <math>I</math>-<math>V</math> Curves</b>	<b>83</b>
4.1	Review . . . . .	84
4.2	Setup . . . . .	85
4.3	The current-to-voltage converter . . . . .	86
4.4	Model of the circuit . . . . .	94
4.5	Calibration . . . . .	96
4.6	Calculating $G$ , $I$ , and $V$ . . . . .	97
4.7	Timing the burst . . . . .	99
4.8	Performance on resistors . . . . .	102
4.9	Gold quantum point contacts . . . . .	106
4.10	Summary . . . . .	109
	References . . . . .	110
<b>5</b>	<b><math>I</math>-<math>V</math> Curves of Au Contacts</b>	<b>115</b>
5.1	Motivation . . . . .	117
5.2	Experimental details . . . . .	117
5.3	The stability criterion . . . . .	118
5.4	Experimental results . . . . .	120
5.5	Tight-binding calculations . . . . .	124
5.6	Discussion . . . . .	127
5.7	Summary . . . . .	134
	References . . . . .	136
<b>6</b>	<b>Tunneling Through a Thin Insulating Film</b>	<b>141</b>
6.1	Introduction . . . . .	142
6.2	Review . . . . .	143
6.3	A simplified geometry . . . . .	145
6.4	General formulas for the current . . . . .	145
6.5	Selecting a suitable model potential . . . . .	153
6.6	$I(V)$ for the parabolic barrier . . . . .	159
6.7	Summary . . . . .	171
	References . . . . .	172

<b>7</b>	<b>Current-Induced Disruptions</b>	<b>179</b>
7.1	Experimental details . . . . .	180
7.2	Results . . . . .	182
7.3	Discussion . . . . .	190
7.4	Outlook and conclusion . . . . .	192
	References . . . . .	194
<b>8</b>	<b>Outlook</b>	<b>197</b>
	References . . . . .	199
<b>9</b>	<b>Danish Summary</b>	<b>201</b>
<b>A</b>	<b>The Maxwellian Orifice</b>	<b>207</b>
A.1	The oblate spheroidal coordinate system . . . . .	207
A.2	Symmetry and boundary conditions . . . . .	208
A.3	Solving Laplace's equation . . . . .	208
A.4	Calculating the conductance . . . . .	209
<b>B</b>	<b>The Dilogarithm</b>	<b>213</b>
<b>C</b>	<b>Stratton's Result</b>	<b>217</b>
<b>D</b>	<b>The Kolmogorov-Smirnov Test</b>	<b>219</b>



# Chapter 1

## Introduction

In 1965 Gordon Moore<sup>1</sup> announced that the number of transistors that could be put on a single chip would double every year for the next ten years until reaching 65 000 components/chip [1]. The prediction turned out to be true — far more accurate than Moore had ever thought it would be. In fact, the complexity of a chip continued to double yearly for long after 1975. The rate of doubling has only recently slowed to about every 18 months (Moore's law [2]). Using state-of-the-art deep ultraviolet (0.18 micron) lithographic techniques, it is now possible to cram 30 million transistors onto a single chip.<sup>2</sup>

### 1.1 Interconnects

Not only the size of the transistors has decreased inexorably for the last three to four decades. So have the interconnects between the transistors. In previous generations of the lithographic technology, the interconnects have only been of minor importance to the performance and speed of the chips. However, with the current 0.18 micron design rule, the resistance of the interconnects has grown to such an extent that the main limitation to the speed is the  $RC$  relaxation time of the interconnects [3]. For the same reasons, the industry recently switched from Al to Cu as wire material [3] because Cu has a lower resistivity [4].

Typical current densities in the very thin interconnects used nowadays are of the order  $6 \times 10^5$  A/cm<sup>2</sup> [3]. At such high current densities, atoms at the

---

<sup>1</sup>Gordon Moore is one of the co-founders of Intel and a known industry sage.

<sup>2</sup>The Intel Pentium III processor.

grain boundaries of the polycrystalline interconnects begin to migrate due to the applied electric field [5, 6]. This effect, called electromigration, poses a serious limitation to long-term component reliability. In 2008, the linewidth is expected to reach 80 nm according to the International Technology Roadmap for Semiconductors (ITRS) [3]. At this point, current densities will have increased to  $2 \times 10^6$  A/cm<sup>2</sup> — a level at which the risk of copper interconnect failure reaches a catastrophic level. At the moment, no solutions are known to this problem. The industry is therefore (reluctantly) beginning to look for alternatives to Cu. One possibility that has been considered is Au [7, 8].

Albeit Au has a resistivity which is slightly larger than Cu [4], gold is more resistant to electromigration. In addition, gold is very inert [9]. Hence, it should be possible to fabricate interconnects of very high purity. This is important because the interconnect resistance is critically dependent on the impurity level. Unfortunately, Au is also much more expensive than, e.g., Cu. For this and other reasons, the industry will push the Cu technology as far as possible before considering Au or other materials as an alternative [3].

The semiconductor industry is facing another serious problem; as the size of the interconnects decreases, the classical continuum laws of physics breaks down, and new effects have to be considered. The detailed understanding of many of these effects is currently lacking [10] or not yet implemented in commercial design tools [3]. As a consequence, the present design tools used for modeling the interconnects are crude and inaccurate, and a pessimistic view of the performance has to be assumed in the design to assure that the devices will actually work [3]. It is estimated that the performance could be increased by 30% if the design tools were more accurate.

## 1.2 Moore and more transistors in the future?

Present-day interconnects are incredibly small compared to objects visible to the human eye. It would take 200 000 interconnects to make a wire bundle with a thickness comparable to a human hair! We might therefore ask ourselves: How long will this relentless miniaturization last? What is the smallest feature size that we can possibly obtain? The answer to the last question was given by Richard P. Feynman as early as 1959 in his visionary and famous talk ‘There’s Plenty of Room at the Bottom’ [11]. In his talk, Feynman emphasized that the only fundamental limit to miniaturization is the size of the atom. As a corollary to his remarks, we conclude that the thinnest wire, we can possibly fabricate will be a chain of single atoms. Likewise, the ultimate transistor will be based

on a single functional molecule. The technological obstacles in reaching this atomic-size limit may be gargantuan, but it is possible *in principle*. Since there are about 500 000 atoms in the cross-section of an 0.18 micron interconnect, we can still say today: There is plenty of room at the bottom!

If we boldly extrapolate Moore's law well into the future, the design rule will be 0.25 nm in the year  $2050 \pm 10$ , i.e., comparable to the diameter of a single atom. Needless to say, this prediction may not hold as the industry is facing several serious challenges in the near future. At present, Moore's law is expected to hold well into the 2010's [3], but severe physical limits are being approached with menacing haste. For instance, the gate oxide thickness in metal-oxide semiconductor field-effect transistors (MOSFETs) is expected to be as thin as five atoms in 2012 [3]. With such thin oxides the electrical insulation of the gate breaks down, and the transistor ceases to operate properly [12]. Therefore, the science community and the semiconductor industry have to come up with new solutions to avoid a bottle-neck in growth [12]. However, it is not the first time that apparent fundamental limits have been reached. In all previous cases, the dreaded showstops have been avoided.

### 1.3 Motivation

It is clear that more research into the fundamental properties of very small electrical conductors is needed to get a better physical understanding of these systems. The interconnects are now so small that surface electron scattering is becoming important, and in the future quantum mechanical effects have to be considered as well [3].

The results presented in this thesis hopefully add to the rapidly growing understanding of atomic-sized metal contacts. Such systems are prototypes for atomic-scale interconnects that may be mass-fabricated in 40–60 years. The results are certainly not directly applicable to the industry today, but they give glimpses of some of the new exciting phenomena and properties which characterize these systems. The goal of my studies has been to go to the ultimate limit of small atomic systems (Feynman's bottom so to say) and to probe some of the electronic transport properties of atomic-sized metal contacts with no more than a few atoms in the smallest cross-section. The work has primarily focused on gold contacts, which as previously mentioned is a very inert metal and hence quite easy to handle. The two other noble metals, Ag and Cu, have also been studied to some extent as well as a few transition metals.

## 1.4 Structure of the thesis

To provide the non-specialist with the necessary theoretical framework, a thorough introduction to the theory of conduction in small metal contacts is given in Chapter 2. Here we review how different electronic transport regimes come into play as we decrease the size of the contact. The first experimental results are presented in Chapter 3. Here we describe a particularly simple experimental method of forming atomic-sized contacts based on breaking relay contacts. The conductance of the transient contacts is investigated for the noble metals and a number of transition metals. An instrumental artifact called differential nonlinearity is also described, and a recipe for removing its influence on the results is given. Furthermore, a simple model for the resistance of the leads is discussed.

In Chapter 4 we describe a more advanced experimental setup, which is used for studying the current-voltage ( $I$ - $V$ ) characteristics of metastable contact geometries. The setup is based on a scanning tunneling microscope (STM) (see Ref. 13 and references therein) placed in an ultra-high vacuum chamber connected to a modified set of STM electronics. With the setup it is possible to accurately measure a single  $I$ - $V$  curve during a period of down to  $10 \mu\text{s}$ , which is four to seven orders of magnitude faster than previous measurements.

Using the new setup, we have measured  $I$ - $V$  characteristics of atomic-sized gold contacts. The results of these measurements are presented in Chapter 5. We show that the obtained  $I$ - $V$  curves depend on sample cleaning procedures and contact formation techniques. Our results indicate that clean atomic-sized gold contacts have (almost) linear (Ohmic)  $I$ - $V$  characteristics, which is also in agreement with tight-binding calculations. In contrast, nonlinear curves are correlated with a contaminated Au-Au interface. We suggest that the nonlinearities result from tunneling through an ultra-thin contaminant layer squeezed between the Au electrodes. This scenario is discussed in detail in Chapter 6, where we present a parabolic barrier model for tunneling through the layer. The model is described by three parameters, and good agreement with experimental curves is found.

As a final application of the fast  $I$ - $V$  curve setup, we show in Chapter 7 preliminary results regarding current-induced disruptions of monatomic gold contacts. Here we demonstrate that it is possible to pass a current of up to  $190 \mu\text{A}$  through a single gold atom. To our knowledge, this is the largest current that has ever been measured through a single atom. The equivalent current density is  $2.8 \times 10^{11} \text{ A/cm}^2$ ! This dazzling number is five orders of magnitude larger than the current density leading to long-term failure in semiconductor interconnects. Admittedly, the QPCs are not stable for long under such extreme

conditions. We show that they break on a time-scale of a few microseconds.

In Chapter 8, we briefly mention various promising avenues of future research. Finally, in Chapter 9, a summary in Danish is given for the interested non-expert. Four appendices are dedicated to auxiliary calculational details.<sup>3</sup>

The equations in the thesis are not equally important. Certain definitions and formulas are used repeatedly throughout the thesis. These formulas are enclosed in a box to catch the eye of the reader. The same goes for main results of lengthy calculations. For convenience, references are found at the end of each Chapter.

---

<sup>3</sup>Despite the appendices, some readers may still find that there are too many calculational details in Chapters 2, 4, and 6. This reflects a flaw in my personality: I *love* equations, and I think they look *great* when typeset in L<sup>A</sup>T<sub>E</sub>X.

## References

- [1] G. E. Moore, “Cramming more components onto integrated circuits,” *Electronics* **April 19**, 114–117 (1965).
- [2] G. E. Moore, “Progress in digital integrated electronics,” In *International Electron Devices Meeting.*, Technical Digest pp. 11–13 (New York, 1975).
- [3] Semiconductor Industry Association, “The International Technology Roadmap for Semiconductors.”, 1999, available on the web at [http://notes.sematech.org/1999\\_SIA\\_Roadmap/Home.htm](http://notes.sematech.org/1999_SIA_Roadmap/Home.htm).
- [4] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [5] P. S. Ho and T. Kwok, “Electromigration in metals,” *Rep. Prog. Phys.* **52**, 301–348 (1989).
- [6] R. S. Sorbello, “Theory of Electromigration,” *Solid State Physics* **51**, 159–231 (1998).
- [7] G. Kamarinos and P. Felix, “How will physics be involved in silicon microelectronics,” *J. Phys. D: Appl. Phys.* **29**, 487–500 (1996).
- [8] B. Roberts, A. Harrus, and R. L. Jackson, “Interconnect metallization for future device generations,” *Solid State Technology* **38**, 69–78 (1995).
- [9] B. Hammer and J. K. Nørskov, “Why gold is the noblest of all the metals,” *Nature* **376**, 238–240 (1995).
- [10] C. Durkan, M. A. Schneider, and M. E. Welland, “Analysis of failure mechanisms in electrically stressed Au nanowires,” *J. Appl. Phys.* **86**, 1280–1286 (1999).
- [11] R. P. Feynman, “There’s Plenty of Room at the Bottom,” 1959, talk at the annual meeting of The American Physical Society at California Institute of Technology. Available on the web at <http://www.zyvex.com/nanotech/feynman.html>.
- [12] M. Schulz, “The end of the road for silicon?,” *Nature* **399**, 729–730 (1999).
- [13] G. Binnig and H. Rohrer, “Scanning tunneling microscopy — from birth to adolescence,” *Rev. Mod. Phys.* **59**, 615–625 (1987).

## Chapter 2

# Theory of Conductance

In this Chapter we will see how different length scales have to be considered when trying to calculate the resistance  $R$ , or equivalently the conductance  $G = 1/R$ , of a metallic constriction connecting two macroscopic electrodes. As the constriction is reduced in size from macroscopic to atomic dimensions, the electronic transport mechanisms go through a series of different regimes as illustrated in Fig. 2.1.

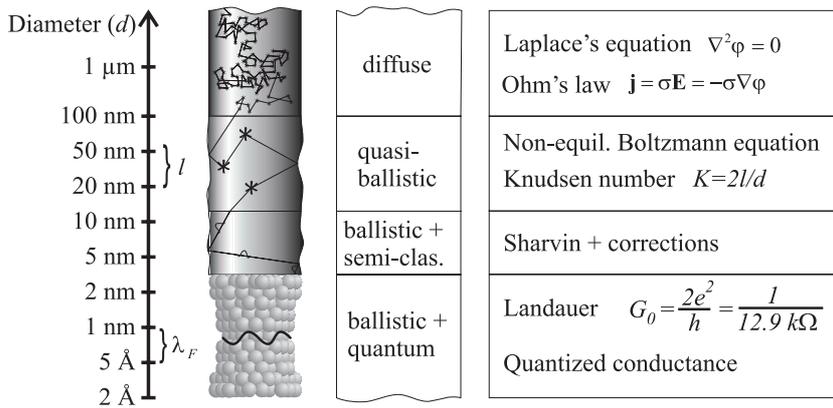


Figure 2.1: The different electronic transport regimes as the size of the metal contact is reduced.

## 2.1 Transport regimes

For macroscopic constrictions,  $G$  will depend on the geometry of the constriction as well as on the conductivity  $\sigma$  of the metal.<sup>1</sup> This is the familiar diffuse (or Maxwellian) transport regime, where the differential form of Ohm's law

$$\mathbf{j} = \sigma \mathbf{E} \quad (2.1)$$

is valid.<sup>2</sup> Here  $\mathbf{j}$  is the current density and  $\mathbf{E}$  is the electric field. The word diffuse is used because the diameter of the contact  $d$  is much larger than the mean free path  $l$  of the electrons in this regime ( $d \gg l$ ). The (bulk) mean free path at room temperature is 200–600 Å for the alkali and noble metals,<sup>3</sup> but smaller for the transition metals ( $\approx 100$  Å) [1].

When  $d$  becomes comparable to  $l$ , the differential form of Ohm's law is no longer valid and we enter a quasi-ballistic regime, where  $G$  depends on the Knudsen number<sup>4</sup>  $K = 2l/d$ . In this regime, scattering from the boundary of the constriction has to be considered along with scattering processes inside and near the constriction.

As we diminish the size of the constriction even further ( $K \gg 1$ ), scattering processes within the constriction can be neglected. This is called the ballistic (or Sharvin) regime as the electrons will move ballistically through the constriction and only scatter at the boundary. Thus, the conductance will be independent of the mean free path and solely governed by the geometry.

For constriction sizes comparable to the Fermi wavelength  $\lambda_F$  of the conducting electrons ( $d \sim \lambda_F$ ), the quantum mechanical wave nature of the electrons has to be considered. For the free-electron metals, the Fermi wavelength is

<sup>1</sup>The conductivity  $\sigma$  [length<sup>-1</sup> × conductance] is the reciprocal of the resistivity  $\rho$  [length × resistance]. Strictly speaking, the conductivity is a tensor. However, if we restrict ourselves to crystals with cubic symmetry, the tensor will be diagonal with  $\sigma = \sigma_{xx} = \sigma_{yy} = \sigma_{zz}$ , corresponding to a direction independent scalar conductivity [1].

<sup>2</sup>Ohm's law actually states that  $\mathbf{j} = \sigma \mathbf{f}$ , where  $\mathbf{f}$  is the force per unit charge. In principle, the force that drives the charges to produce the current could be anything — chemical, gravitational, electric. For our purpose, however, it will always be the Lorentz force  $\mathbf{E} + \mathbf{v} \times \mathbf{B}$  that does the job. Since typical drift velocities in conductors are vanishing, the  $\mathbf{v} \times \mathbf{B}$  term can be neglected, and we end up with  $\mathbf{f} = \mathbf{E}$  and Eq. (2.1) [2].

<sup>3</sup>Recent experiments by Ludoph and Ruitenbeek [3] indicate that the (elastic) mean free path in the vicinity of atomic-scale noble and alkali metal contacts at liquid helium temperatures is as small as 30–70 Å.

<sup>4</sup>Named after the Danish physicist Martin Knudsen, a pioneer in kinetic gas theory, who studied the transport properties of gases with  $K \gg 1$  (the Knudsen limit).

direction independent and given by

$$\lambda_F = \frac{2\pi}{k_F}, \quad (2.2)$$

where  $k_F$  is the magnitude of the Fermi wave vector. The Fermi wavelength is approximately 5 Å for the noble metals and 6–10 Å for the alkali metals [1]. For these atomic-scale constrictions, the conductance is determined by transmission probabilities of electronic wave functions at the Fermi energy  $E_F$ . If the constriction is sufficiently smooth, the transverse part of the wave functions will form standing waves in the constriction. This results in a quantization of the conductance in units of

$$G_0 = \frac{2e^2}{h} \simeq 77.481 \mu\text{S} \simeq \frac{1}{12\,906 \Omega}, \quad (2.3)$$

where  $e$  is the elementary charge and  $h$  is Planck's constant.  $G_0$  is often referred to as the *fundamental quantum unit of conductance*<sup>5</sup> and will be used extensively throughout this thesis.

The calculational methods used in the different regimes will be outlined in the following, using various important model constriction geometries.

## 2.2 Diffuse electron transport

### Basic theory

In the diffuse electron transport regime ( $d \gg l$ ), the differential form of Ohm's law given by Eq. (2.1) is valid and can be used when calculating the conductance. Consider the continuity equation

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0,$$

where  $\rho(\mathbf{r}, t)$  is the charge density,  $\mathbf{r} = (x, y, z)$  is the position, and  $t$  is the time. If we restrict ourselves to calculating the DC conductance (as we will do throughout this Chapter), we have  $\partial \rho / \partial t = 0$ , and therefore  $\nabla \cdot \mathbf{j} = 0$ . Using Ohm's law  $\mathbf{j} = \sigma \mathbf{E}$ , this implies that  $\nabla \cdot \mathbf{E} = 0$ . From Gauss' law  $\nabla \cdot \mathbf{E} = \rho / \epsilon_0$  it then follows that  $\rho(\mathbf{r}, t) = 0$  everywhere inside the conductor.

<sup>5</sup>A physicist from the quantum Hall community would probably find the quantity  $G_0/2 = e^2/h$  more fundamental.

In steady state it follows from Faraday's law that  $\nabla \times \mathbf{E} = 0$ . Since  $\nabla \times (\nabla f(\mathbf{r})) = 0$  for any scalar function  $f(\mathbf{r})$ , it is therefore convenient to write the electric field as (minus) the gradient of a scalar function  $V(\mathbf{r})$  called the electrostatic potential

$$\mathbf{E} = -\nabla V. \quad (2.4)$$

With this definition, Gauss' law ( $\nabla \cdot \mathbf{E} = 0$ ) leads directly to Laplace's equation

$$\nabla^2 V = 0. \quad (2.5)$$

In order to obtain a unique solution to Laplace's equation in some bounded region of space (in this case within a metal object), we have to specify either  $V$  (Dirichlet boundary conditions) *or*  $\partial V/\partial n$  (Neumann boundary conditions) everywhere at the boundary of the surface [4]. Here  $\partial V/\partial n$  denotes the derivative of the potential along the normal of the boundary, which is equivalent to the normal component of the electric field.

### The conductance of arbitrary contact geometries

When dealing with the conductance of some object with a finite conductivity, it is important to specify the surface regions between which the conductance is to be measured or calculated. This corresponds to specifying the Dirichlet boundary conditions; for instance, we may apply a potential  $V = U$  at one end of the object, and at the other end we may ground it at  $V = 0$ . The remaining parts of the surface are assumed to be in contact with an isolating medium like air or vacuum. Hence, no current can flow into or out of these parts of the surface. This in turn implies that  $\mathbf{j}$  will be parallel to the surface in those regions. As a consequence of Ohm's law, it follows that  $\mathbf{E}$  will be parallel as well, and thus  $\partial V/\partial n = 0$ . This is the Neumann boundary condition, whereby we have specified the (mixed) boundary conditions needed to obtain a unique solution of Laplace's equation [4].

Once Laplace's equation is solved, the electric field can be found from Eq. (2.4). The current density is subsequently calculated from Ohm's law, and the total current  $I$  flowing between the electrodes is found by calculating the flux

$$I = \oint_S \mathbf{j} \cdot d\mathbf{S} \quad (2.6)$$

through a surface  $S$  placed somewhere between the electrodes and attached to the boundary of the region. Finally, the conductance can be found by using the

geometry dependent version of Ohm's law

$$I = GU. \quad (2.7)$$

### The conductance of a wire

To illustrate how this procedure works, we will first consider the simple example of a straight, circular wire of length  $L$  and area  $A$ . The ends of the wire are placed along the  $z$ -axis at  $z = 0$  and  $z = L$ , and we apply a voltage difference  $U$  between the ends thus imposing the Dirichlet boundary conditions  $V(z=0) = U$  and  $V(z=L) = 0$ . In addition, we have the Neumann boundary conditions  $\partial V/\partial x = \partial V/\partial y = 0$  everywhere on the cylindrical surface.

The next step is to solve Laplace's equation. For simple geometries like this, the solution is most easily found by guessing a solution, which fulfills all the boundary conditions and obeys Laplace's equation everywhere in the interior [2]. If we succeed in guessing the potential, we know from the uniqueness theorem that this is the *only* solution [2, 4]. In this particular case, the potential is of course

$$V(x, y, z) = U(1 - z/L).$$

The electric field is now found from Eq. (2.4)

$$\mathbf{E} = -\nabla V = \frac{U}{L}\hat{\mathbf{z}},$$

where  $\hat{\mathbf{z}}$  is a unit vector in the  $z$ -direction. Using the differential form of Ohm's law, we then find

$$\mathbf{j} = \frac{\sigma U}{L}\hat{\mathbf{z}},$$

and the total current is found by multiplying by the area

$$I = \frac{\sigma A}{L}U.$$

Finally, from the geometry dependent version of Ohm's law [Eq. (2.7)], we find the conductance by dividing by  $U$

$$\boxed{G_{\text{wire}} = \frac{\sigma A}{L}}. \quad (2.8)$$

This is a very well known result,<sup>6</sup> which could have been found using less rigorous arguments (and space). However, in this derivation, care is taken to keep track of *where* the external voltage is applied. In the following, we will see that this is important. We note that the conductance of a straight wire is proportional to the area and inversely proportional to the length. Thus the wires obey the normal sum rules for resistors in parallel and in series.

### The conductance of an orifice

In the derivation of Eq. (2.8), it is assumed that the voltage is applied right at the ends of the wire. This is easily realized when dealing with long and thin wires, but when the wires become very short we have to consider *how* the wire is connected with the macroscopic world through the connecting electrodes. To illustrate this, we can consider a cylindrical wire of diameter  $d$  and let the length go to zero, thereby turning the ‘wire’ into a circular orifice. The conductance of the orifice ‘itself’ will be infinitely large as can be seen from Eq. (2.8). However, this is not what is normally measured because we have to connect the ‘wire’ with an externally applied potential through electrodes with finite conductivity and measure the current.

Such a point contact could for example be modeled by two semi-infinite half spaces of conductivity  $\sigma$ , separated by a planar diaphragm and connected by a circular orifice of diameter  $d$  as illustrated in Fig. 2.2. This model was first considered by Maxwell in 1892, who calculated the corresponding conductance  $G_M$  [5]. Details of a (rather lengthy) derivation based on the previously outlined technique are given in Appendix A, where the following simple result is found

$$\boxed{G_M = \sigma d.} \tag{A.14}$$

We note that the conductance is *not* proportional to the area as was the case for the wire [cf. Eq. (2.8)]. Since the diameter is the only length scale in this problem and the conductivity has physical dimension of conductance  $\times$  length<sup>-1</sup>, the conductance *has* to be proportional to  $d^1$ .

It is interesting to note from Fig. 2.2 that disturbances in the potential persist quite far away from the orifice. At a distance of  $\approx 3d$  from the orifice, 10% of the total potential drop still remains. One may say, that the ‘resistance’ of the orifice is concentrated in a rather large region around the orifice. This clearly demonstrates the importance of considering the electrodes as well as the

---

<sup>6</sup>The analogous expression  $R_{\text{wire}} = \rho L/A$  for the resistance is used in high school physics classes.

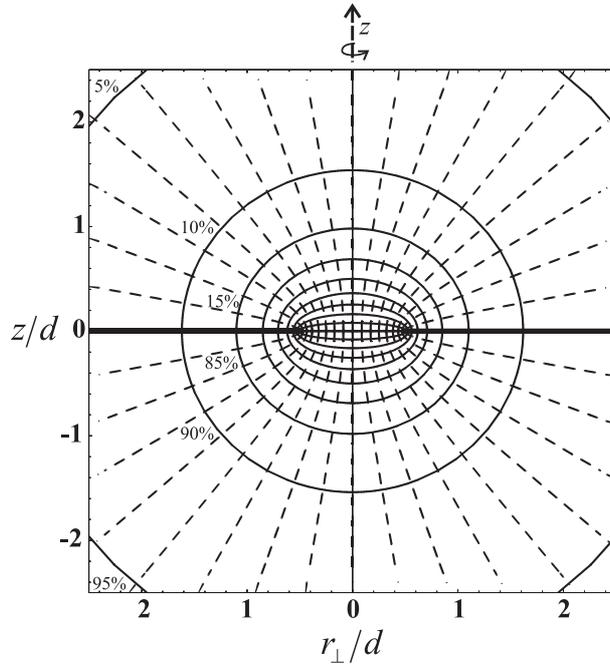


Figure 2.2: Equipotential surfaces (solid lines) and electric field lines (dashed) for the diffuse transport of electrons through a circular orifice of diameter  $d$  placed coaxially with the  $z$ -axis in the  $z = 0$  plane. Rotational symmetry prevails around the  $z$ -axis. The spacing between two equipotential surfaces corresponds to 5% of the applied voltage. In terms of the oblate spheroidal coordinates discussed in Appendix A, equipotential surfaces correspond to constant values of  $\xi$ , and electric field lines correspond to constant values of  $\eta$ .

constriction when calculating the conductance. From the density of field lines in Fig. 2.2, we also see that the electric field close to the edge of the orifice is very large. From Eq. (A.13) on page 210 in Appendix A, it actually follows that the current density goes to infinity as  $r_{\perp} \rightarrow d/2$ .

## 2.3 Ballistic electron transport

Let us consider the orifice geometry in the ballistic regime ( $d \ll l$ ). We rotate the geometry in Fig. 2.2 on page 13 by  $90^\circ$  so that the electrodes are positioned to the left and right as indicated in Fig. 2.3. We ground the electrode to the left and apply a small positive voltage  $V$  deep inside the electrode to the right, thus lowering the Fermi energy in this electrode by  $eV$ . The electric field is concentrated in a small region around the orifice of size  $\approx d$ .<sup>7</sup> Therefore, electrons passing through the orifice from left to right will be accelerated within a single mean free path. Only the electrons in the left electrode with energies in the range  $[E_F - eV; E_F]$  have to be considered as these have sufficient energy to go into unoccupied states in the positively biased electrode (neglecting temperature effects).

### A general formula

The scenario is shown in Fig. 2.3. The semiclassical equations of motion state that the velocity of an electron with wave vector  $\mathbf{k}$  is [1]

$$\mathbf{v}_n(\mathbf{k}) = \frac{1}{\hbar} \nabla_{\mathbf{k}} E_n(\mathbf{k}), \quad (2.9)$$

where  $E_n(\mathbf{k})$  is the  $n$ th energy band. Since Fermi surfaces (FSs) are constant energy surfaces, the velocity is perpendicular to the surface as shown in Fig. 2.3. Only electrons with a velocity component to the right will contribute to the current. In mathematical terms, this means that the  $z$ -component of the unit vector  $\hat{\mathbf{n}}(\mathbf{k})$  normal to the FS should be positive. This implies that only states in the shaded region ( $\Omega$ ) of Fig. 2.3 have to be considered. The current density

---

<sup>7</sup>In the ballistic regime, the electrostatic potential can be found by solving the collisionless Boltzmann equation for the orifice geometry [6]. The resulting potential is qualitatively similar to the diffuse potential depicted in Fig. 2.2. However, the potential drop is located in a smaller region around the orifice (a 90% drop within a distance of  $\simeq d$  as compared to  $\simeq 3d$  for the Maxwellian orifice).

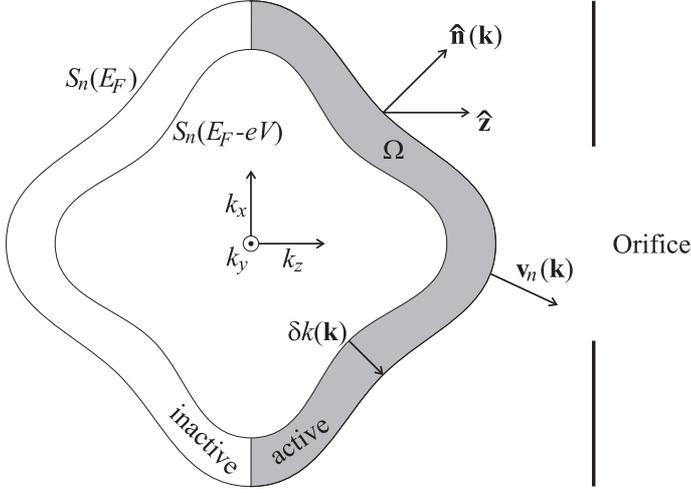


Figure 2.3: Cross section through the  $k_y = 0$  plane of the Fermi surfaces  $S_n(E_F)$  and  $S_n(E_F - eV)$  of the  $n$ th energy band of an arbitrary metal constituting the left electrode. The shaded region of  $\mathbf{k}$  points (denoted by  $\Omega$ ) indicates electronic states with a velocity component to the right and sufficient energy to go into unoccupied states in the electrode to the right.

of the active electrons in the  $n$ th band  $\mathbf{j}_n$  incident on the orifice is then given by an integration over  $\Omega$  [1]

$$\mathbf{j}_n = (-e) \int_{\Omega} \frac{d\mathbf{k}^3}{4\pi^3} \mathbf{v}_n(\mathbf{k}).$$

If we introduce the quantity  $\delta k(\mathbf{k})$  as the distance between the two FSs (see Fig. 2.3),  $\mathbf{j}_n$  can be rewritten to a surface integral over  $S_n(E_F - eV)$

$$\mathbf{j}_n = (-e) \oint_{S_n(E_F - eV)} \frac{dS_n}{4\pi^3} \delta k(\mathbf{k}) \mathbf{v}_n(\mathbf{k}). \quad (2.10)$$

where only parts of the surface fulfilling  $n_z(\mathbf{k}) > 0$  should be included. Using that [1]

$$\delta k(\mathbf{k}) = \frac{eV}{|\nabla_{\mathbf{k}} E_n(\mathbf{k})|},$$

we see from Eq. (2.9) that  $\delta k(\mathbf{k})\mathbf{v}_n(\mathbf{k}) = eV/\hbar \hat{\mathbf{n}}(\mathbf{k})$ , whereby all explicit references to the band structure cancels. When this relation is inserted in Eq. (2.10), we find for  $eV \ll E_F$  that the  $z$  component of the total current density  $j_z = \sum_n \mathbf{j}_n \cdot \hat{\mathbf{z}}$  incident on the orifice is given by

$$j_z = -\frac{e^2V}{2\pi^2\hbar}S(\hat{\mathbf{z}}), \quad (2.11)$$

where  $S(\hat{\mathbf{z}})$  is the area of the FS, projected on the plane normal to  $\hat{\mathbf{z}}$  [7]

$$S(\hat{\mathbf{z}}) = \sum_n \int_{n_z > 0} dS_n \hat{\mathbf{n}}(\mathbf{k}) \cdot \hat{\mathbf{z}}. \quad (2.12)$$

If we assume that none of the electrons which pass the orifice to the right-hand electrode is backscattered<sup>8</sup> into the left-hand electrode, the current  $I$  through an orifice with area  $A$  is simply given by

$$I = -Aj_z \quad (2.13)$$

(minus since the current is running from right to left). By inserting Eq. (2.11) in Eq. (2.13) and dividing with  $V$  we find a general expression for the ballistic conductance [6, 8, 11], which is written in units of  $G_0$  as [7]

$$G_{\text{ballis}} = \frac{AS(\hat{\mathbf{z}})}{4\pi^2} G_0. \quad (2.14)$$

### The Sharvin conductance

For a free-electron metal, the FS is simply a sphere with a radius of the Fermi wave vector  $k_F$  [1]. In that case we have  $S(\hat{\mathbf{z}}) = 4\pi^3/\lambda_F^2$ . Upon insertion in Eq. (2.14), a particularly simple formula for the conductance is obtained, which was first derived by Sharvin in 1965 [12]

$$G_S = \frac{\pi A}{\lambda_F^2} G_0. \quad (2.15)$$

---

<sup>8</sup>Due to the finite value of  $l$ , some backscattering in real point contacts will always exist. The amount of backscattering changes with voltage due to the energy dependence of  $l$ . This leads to non-zero values of  $d^2I/dV^2(V)$ . It can be shown that  $d^2I/dV^2(V)$  is roughly proportional to  $\alpha^2 F(eV)$ , where  $\alpha^2$  is the square of the electron-phonon matrix element and  $F(\omega)$  is the phonon density of states [6, 8–10]. This quantity is measured in point-contact spectroscopy.

A direction independent Fermi wavelength cannot be defined for metals with non-spherical Fermi surfaces. In that case it is better to write the Sharvin conductance in terms of the free electron density  $n = k_F^3/(3\pi^2)$  [1], whereby we can approximate Eq. (2.14) by

$$G_S \simeq \frac{\sqrt[3]{3^2} \pi}{4} n^{2/3} A G_0 \quad (2.16)$$

without taking the detailed shape of the Fermi surface into account.

Unlike the expression for the Maxwellian orifice given by Eq. (A.14), the Sharvin conductance does not depend on the conductivity (which is proportional to the mean free path). The only material dependent parameter appearing in Eq. (2.15) is the Fermi wavelength, which is the important length scale in this regime. If the area is measured in units of  $\lambda_F^2$ , the conductance is conveniently expressed in units of  $\pi G_0$ . In contrast to the result for the Maxwellian orifice, the conductance is again proportional to the area.

### The perimeter correction

The Sharvin formula [Eq. (2.15)] is only asymptotically correct when  $A/\lambda_F^2 \gg 1$  because it neglects quantum finite size effects, which suppress the conductance as the area is decreased [13]. This can be taken into account by making a so-called perimeter correction to the Sharvin formula [13], which may be written as [14]

$$G_S = \left( \frac{\pi A}{\lambda_F^2} - \alpha \frac{P}{\lambda_F} \right) G_0, \quad (2.17)$$

where  $P$  is the perimeter of the orifice, and  $\alpha$  is a parameter which depends on the shape of the orifice. For a circular orifice  $\alpha = 1/4$  [13].

### Soft-wall corrections

Further refinements can be done by including a ‘soft-wall’ correction, which considers the electronic tails in the classically forbidden region near the perimeter [14]. The smaller the ratio between the work function and the Fermi energy, the longer the tails will be. This can be taken into account by decreasing the value of  $\alpha$  in Eq. (2.17) to a value around 0.1, thereby damping substantially the hard-wall perimeter suppression of the conductance [14].

Alternatively, one can use  $\alpha = 1/2$  (corresponding to the perimeter corrected hard-wall conductance of a cylindrical wire [13]) and substitute  $A$  and  $P$  in

Eq. (2.17) with larger effective values  $A^* = \pi(R + \delta R)^2$  and  $P^* = 2\pi(r + \delta R)$ , where  $\delta R$  is a measure of the extension of the decaying tail of the wavefunction into the vacuum [14, 15]. Important examples of actual values are  $\delta R_{\text{Au}} = 0.68 \text{ \AA}$  and  $\delta R_{\text{Na}} = 0.93 \text{ \AA}$  [15].

When the perimeter and ‘soft-wall’ corrections are included, the Sharvin formula can be used to make a reliable estimate of the contact area provided the conductance is above 3–4  $G_0$  [14–17]. However, fine details caused by scattering from the constriction boundaries are not revealed by the Sharvin formula. To include these quantum interference effects, one has to use the Landauer-Büttiker formalism, which will be described in Sec. 2.4.

### How can a ballistic conductor have a finite conductance?

The finite value of the conductance of the ballistic orifice must imply that a power  $P = UI = I^2/G$  is dissipated when a current  $I$  flows through the constriction. This may seem contradictory since we have stated explicitly that we operate in the ballistic regime, where all scattering (energy loss) mechanisms can be neglected.

However, if we re-examine the arguments leading to Eqs. (2.14) and (2.15) we have assumed that none of the electrons passing the orifice will be backscattered. This corresponds to regarding the interior of the right electrode as a ‘black body’, which absorbs all incoming electrons: when the electron enters the electrode to the right, it has an energy which exceeds the Fermi energy by  $eV$  due to the work done by the electric field in the vicinity of the orifice. This energy is gradually lost in a series of inelastic processes, which at room temperature will be dominated by electron-phonon scattering. In each scattering event an energy of the order  $\hbar\omega_D \sim 10 \text{ meV}$  is lost, where  $\omega_D$  is the Debye frequency [1]. Finally, the electron equilibrates in the electrode with an energy close to the Fermi energy (on a scale of  $k_B T$ ).

So although we have not explicitly included dissipation in the derivation, the neglect of backscattering implicitly corresponds to including dissipative electron reservoirs. This, once again, stresses the importance of considering the role of the electrodes.

### A quasiballistic interpolation formula

By considering the non-local Boltzmann equation in the circular orifice geometry, Wexler has derived a formula for the conductance  $G_{\text{quasi}}$  in the quasiballistic

regime, where  $d \sim l$  [18]. This can be written as an interpolation formula between the diffuse [Eq. (A.14)] and the ballistic [Eq. (2.15)] conductance as [9]

$$G_{\text{quasi}} = [G_S^{-1} + \Gamma(K)G_M^{-1}]^{-1}, \quad (2.18)$$

where  $\Gamma(K)$  is a smoothly varying function of the Knudsen number with limiting values  $\Gamma(K=0) = 1$  (the diffuse limit) and  $\Gamma(K \rightarrow \infty) = 0.694$  (the ballistic limit). Since  $G_S \propto A$  and  $G_M \propto \sqrt{A}$ , we get  $G_{\text{quasi}} = G_S$  in the ballistic limit, and  $G_{\text{quasi}} = G_M$  in the diffuse limit as expected. Since  $\Gamma$  is close to one for most values of the Knudsen number, *the resistance of a quasiballistic contact is roughly given by the sum of the diffuse and ballistic resistances.*

This rule of thumb also holds for other contact geometries. For instance, de Jong calculated the resistance of a wire with a diameter  $d \ll l$  and an arbitrary length  $L$  [19]. The exact calculation shows that the naive procedure of summing the Sharvin and diffusive wire resistances ( $G_S^{-1} + G_{\text{wire}}^{-1}$ ) is correct within 3.5% for all values of  $l/L$ .

## 2.4 The Landauer-Büttiker formalism

The Landauer-Büttiker formalism [21–42] gives a very convenient, transparent and intuitive description of the electronic transport in the quantum regime, where electrons move elastically (phase-coherent) through a nanoscale contact. The first pioneering work was done by Landauer already in 1957 [21], but more than twenty years passed before the ‘Landauer-Büttiker’ formula [Eq. (2.21) below] was first derived by Fisher and Lee in 1981 [27].<sup>9</sup> A vivid debate followed [28–31, 38], but finally in the late 1980’s, thirty years after Landauer’s first contribution, a general consensus was finally reached regarding the applicability of Eq. (2.21) [32, 35, 37–40].

### The one-dimensional case

Consider first a one-dimensional (1D) conductor as shown in Fig. 2.4 with a potential barrier characterized by the transmission probability  $T$ . The conductor is connected to two electron reservoirs (the electrodes). We connect a current

---

<sup>9</sup>Although Fisher and Lee were the first to derive the Landauer-Büttiker formula, it has become widespread practice in the literature to refer to a later article from 1985, co-authored by Büttiker and Landauer [31] (hence the name). In my opinion, this is misleading since Eq. (2.21) does not appear in that paper!

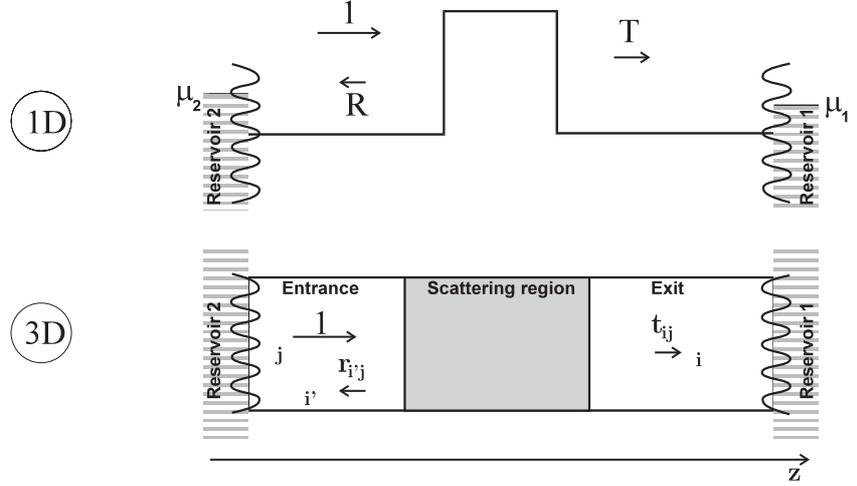


Figure 2.4: The Landauer model. (From: [20].)

source to the electrodes and pass a small current  $I$  from right to left through the conductor. The current is carried by electrons traveling from left to right, driven by a finite gradient of the chemical potential  $\mu$  in the vicinity of the conductor. Deep inside the electrodes, the chemical potential settles at its equilibrium values  $\mu_2$  (right) and  $\mu_1$  (left), with  $\mu_2$  infinitesimally larger than  $\mu_1$ .

The current running through the conductor will be carried by electrons in the energy window  $[\mu_1; \mu_2]$  (at zero temperature). In analogy with the discussion leading to the Sharvin result, we can relate the current to the difference in the chemical potential as

$$I = e \frac{dn}{dE} (\mu_2 - \mu_1) T v,$$

where  $dn/dE = (dn/dk)(dk/dE) = 1/[\pi(dE/dk)]$  is the 1D density of states,<sup>10</sup> and  $v = (1/\hbar)dE/dk$  is the velocity [cf. Eq. (2.9)]. Again, all explicit reference to the band structure  $E(k)$  cancels in the expression for the current and we find

$$I = \frac{2e}{h} (\mu_2 - \mu_1) T. \quad (2.19)$$

<sup>10</sup>The 1D density of  $k$ -points is  $1/2\pi$  [1]. The density of  $k$ -states  $dn/dk$  is twice that value due to spin degeneracy.

The voltage drop  $V$  between the deep interior of the electrodes is given in terms of the chemical potential difference as  $eV = (\mu_2 - \mu_1)$ , and we find the two-terminal conductance [25, 32, 35, 38]

$$\boxed{G_{1D} = G_0 T.} \quad (2.20)$$

At first it may seem surprising that the conductance of a perfect conductor ( $T = 1$ ) is finite. However, as we also saw for the diffuse and ballistic orifice, this is due to the coupling between the electrodes and the constriction. As an alternative, we could (in principle) make a four-terminal measurement of the conductance by placing two additional (non-invasive) voltage probes between the current injectors right next to the barrier. In such a four-probe scenario, the conductance *due to the barrier* is given by [22, 24, 25, 29–32, 38]

$$G_{1D}^{\text{barrier}} = G_0 \frac{T}{1 - T}.$$

We emphasize that this is the conductance of the barrier *itself*, defined as the ratio between the current  $I$  through it and the voltage drop  $V$  *across* it [32]. From a Sharvin type argument, it can be shown that the ‘contact resistance’ between the reservoirs and the conductor is  $\simeq 1/G_0$  [32]. If this resistance is coupled in series with the  $1/G_{1D}^{\text{barrier}}$ , we arrive at the two-terminal conductance given by Eq. (2.20).

### The three-dimensional case

The arguments leading to the two-terminal conductance for a 1D conductor [Eq. (2.20)] can be generalized in three dimensions (3D), see Fig. 2.4. Unlike the 1D case, we now have to consider complete sets of eigenstates (channels) at the Fermi energy on both sides of the constriction (conductor). In the 3D case, we therefore have to examine how a channel  $j$  incident on the constriction is scattered elastically into the outgoing channel  $i$  on the other side of the scattering region and calculate the corresponding transmission amplitude  $t_{ij}$ . The two-terminal conductance will then be given by a straight-forward generalization of Eq. (2.20), and we obtain the celebrated *Landauer-Büttiker formula* [27, 32]

$$\boxed{G_{3D} = G_0 \text{Tr}(\mathbf{t}^\dagger \mathbf{t}),} \quad (2.21)$$

where  $\text{Tr}$  denotes the trace and  $\mathbf{t}$  is a transmission matrix containing the elements  $t_{ij}$ . (The dagger symbol denotes the Hermitian conjugate.)

The total transmission

$$T_{\text{tot}} \equiv \text{Tr}(\mathbf{t}^\dagger \mathbf{t}) = \sum_{i,j} |t_{ij}|^2 \quad (2.22)$$

depends on the actual constriction geometry and will change if the geometry is changed by, e.g. elongating and thinning the constriction. Examples will be given in some of the following sections.

The Landauer-Büttiker formula clearly states that “conductance is transmission” [42], and it can be applied to a wide range of problems *provided the transport is elastic*. In a metallic quantum point contact (QPC), the Landauer-Büttiker formula can be used for calculating the conductance of a constriction, which is smaller than the decoherence (phase-breaking) length of the electrons. In metallic QPCs, the decoherence length is roughly given by the inelastic mean free path, which is the typical path traversed by the electron between inelastic scattering events where the phase memory is lost. The most important inelastic processes are electron-phonon and electron-electron scattering.

In experiments, the QPC is typically connected with the macroscopic world through diffuse leads, and the conductance is measured between points which exceed the decoherence length by many orders of magnitude. To use the Landauer-Büttiker formula, one should therefore take the resistance of the diffuse leads into account [37, 38].

## Finite temperatures and voltages

The transmission matrix, which is the central ingredient in the Landauer-Büttiker formula, should be evaluated at the Fermi energy, which is the energy of the conducting electrons in the  $T = 0$  and  $V = 0$  limit. However, in experiments, the temperature and bias voltage are always finite. Therefore, if  $\mathbf{t}(E)$  varies substantially on a scale of  $\max(k_B T, eV)$ , Eq. (2.21) is invalid.

Fortunately, the Landauer-Büttiker formula can easily be generalized to finite temperatures and voltages by integrating the transmission over the Fermi-Dirac distributions of the carriers [31, 32]. Thus, a general finite bias and temperature formula for the current can be found<sup>11</sup> [32]

$$I = \frac{2e}{h} \int_0^\infty dE \left\{ [f(E) - f(E + eV)] \times \text{Tr} [\mathbf{t}^\dagger(E, V) \mathbf{t}(E, V)] \right\}, \quad (2.23)$$

---

<sup>11</sup>See also Eq. (2.19) on page 20.

where

$$f(E) = \frac{1}{1 + e^{(E-E_F)/k_B T}} \quad (2.24)$$

is the Fermi-Dirac distribution. In Eq. (2.23) we have explicitly written that  $\mathbf{t}$  not only depends on  $E$ , but also on the applied bias  $V$ . The reason is that an applied bias results in an electric field in the constriction, which deforms the potential landscape felt by the transmitted electrons.

Eq. (2.23) is a very versatile formula. For instance, in paper [IV] Brandbyge uses it for calculating the  $I$ - $V$  curve for one- and two-atom Au contacts. Likewise, in Chapter 6 on page 147, we use the formula as a starting point for calculating the  $I$ - $V$  characteristics for tunneling through a thin insulating film. Actually, in the limit of wide contacts (and after doing certain auxiliary approximations), Eq. (2.23) takes a form [39], which was discussed already in 1930 by Frenkel [43] in relation to metal-insulator-metal tunneling. (See Chapter 6 for further details and a review.)

## 2.5 The adiabatic approximation

As mentioned earlier, the total transmission defined in Eq. (2.22) [and thus the conductance, cf. Eq. (2.21)] depends on the geometry of the constriction. To get a better feeling of the interplay between geometry and transmission, the *adiabatic approximation* is introduced [44–46].

Consider a constriction coupled to free-electron electrodes and placed along the  $z$ -axis. The potential energy  $V(x, y, z)$ , which confines the electrons to stay within the constriction, is determined by the geometry. We restrict ourselves to electrons at the Fermi energy since only electrons in a small energy region around  $E_F$  have to be considered when the applied voltage is small. The one-electron wave function  $\psi$ , which is a solution to the Schrödinger equation, is now expanded at each cross section  $z_0$  in the following basis set [46]

$$\psi(x, y, z_0) = \sum_i \chi_{z_0, i}(x, y) \phi_i(z_0), \quad (2.25)$$

where the wave functions  $\chi_{z_0, i}$  are the *adiabatic channels*, which are the two-dimensional (2D) eigenstates of the Schrödinger equation,

$$\left[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + V(x, y, z_0) \right] \chi_{z_0, i}(x, y) = E_i^\perp(z_0) \chi_{z_0, i}(x, y) \quad (2.26)$$

evaluated at the cross section  $z_0$ .

When this expansion is inserted in the 3D Schrödinger equation, we get the following set of coupled channel equations for the motion along  $z$

$$\left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial z^2} + E_i^\perp(z) \right] \phi_i(z) + \sum_j \hat{F}_{ij} \phi_j(z) = E_F \phi_i(z), \quad (2.27)$$

where the operators  $\hat{F}_{ij}$  couple the different  $z$  solutions  $\phi_i$ . They have the form

$$\hat{F}_{ij} = -\frac{\hbar^2}{2m} \left\{ \left\langle \chi_i(z) \left| \frac{\partial^2}{\partial z^2} \chi_j(z) \right. \right\rangle + 2 \left\langle \chi_i(z) \left| \frac{\partial}{\partial z} \chi_j(z) \right. \right\rangle \frac{\partial}{\partial z} \right\}. \quad (2.28)$$

It is seen that  $\hat{F}_{ij}$  depends on derivatives of  $\chi_i$  with respect to  $z$ . Therefore, if the variation of  $\chi_i$  with  $z$  in Eq. (2.26) is sufficiently smooth (on a length scale of  $\lambda_F$  [44]), the coupling between individual channels  $i$  in Eq. (2.27) becomes vanishingly small. In this limit — the adiabatic approximation [44] — we are left with one equation for each  $i$ , which has the form of a 1D Schrödinger equation with an effective potential barrier  $E_i^\perp(z)$ .

What is the physical interpretation of  $E_i^\perp$ ? If we for simplicity consider a hard-wall potential which is zero inside and infinite outside the constriction, Eq. (2.26) takes the form of the simple 2D particle-in-a-box problem. The energy  $E_i^\perp(z)$  is then the kinetic energy due to transverse motion in the contact. As is well known, this energy is inversely proportional to the area due to boundary conditions at the perimeter of the cross section [in this case  $\chi_{z_0,i}(x, y) = 0$ ]. This is schematically shown in Fig. 2.5, where  $E_i^\perp(z)$  is plotted for four different channels along a smoothly varying constriction.

If the adiabatic channels are chosen as the basis set of incoming and outgoing channels in the Landauer-Büttiker formula [cf. Eq. (2.21)], then the off-diagonal elements in the transmission matrix  $\mathbf{t}$  will be zero as there is no mixing between the channels. Thus,  $\text{Tr}(\mathbf{t}^\dagger \mathbf{t}) = \sum_i |t_{ii}|^2 = \sum_i T_i$ , where  $T_i = |t_{ii}|^2$  is the transmission probability for traversing the effective potential barrier  $E_i^\perp(z)$ . The 3D Landauer-Büttiker formula is therefore simplified in the adiabatic approximation to

$$G_{3D}^{\text{ad}} = G_0 \sum_i T_i, \quad (2.29)$$

in close analogy with the corresponding 1D formula [Eq. (2.20)].

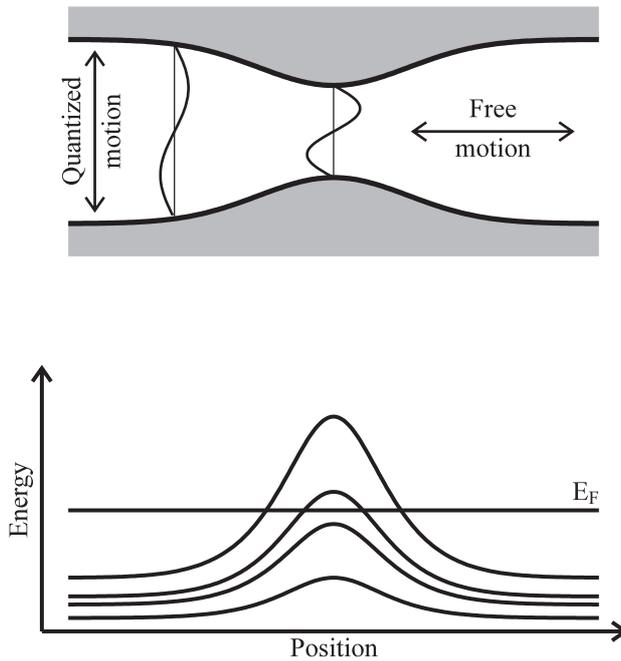


Figure 2.5: The adiabatic picture. (From: [46].)

## 2.6 Quantized conductance

In the strict adiabatic case, the potential barriers  $E_i^\perp(z)$  caused by the transverse kinetic energy will be very wide on the scale of  $\lambda_F$ . This implies that the transmission probability  $T_i$  of an adiabatic channel will be either one or zero depending on whether or not  $E_i^\perp(z)$  stays below  $E_F$  throughout the constriction. This is illustrated in the lower panel of Fig. 2.5, in which two of the channels stay below  $E_F$  (open channels), whereas the two other cross the Fermi energy (closed channels). Consequently, the conductance is  $G = 2 G_0$ .

Now, imagine what happens if the smallest cross-section is made even smaller. The transverse confinement will then be increased, leading to a larger transverse kinetic energy and increased effective potential barriers  $E_i^\perp(z)$ . One by one the maxima of  $E_i^\perp(z)$  will cross the Fermi energy, and the corresponding channels will be closed, leading to a sharp, stepwise decrease in the conductance

in units of  $G_0$ . In this highly idealized picture, the conductance will therefore be strictly quantized in units of  $G_0$ .

The quantization follows from the formation of well defined transverse modes or standing waves in the constriction.<sup>12</sup> In the hard-wall 2D geometry with 1D cross sections shown in Fig. 2.5, the transverse modes are similar to the fundamental and higher harmonics of e.g. a violin string.<sup>13</sup>

In 3D geometries with 2D cross sections, some of the transverse modes may be degenerate in certain simple geometries<sup>14</sup> due to an additional degree of freedom, but the qualitative picture remains unaltered even though the 2D wave functions become more complicated [13, 14, 45–48].

## 2.7 Conductance eigenchannels

Within the independent electron approximation, the coupled channel equations [Eq. (2.27)] can be used for calculating the elastic transmission amplitudes  $t_{ij}$  for any potential  $V$  if the operator  $\hat{F}_{ij}$  introduced in Eq. (2.27) is included. In general, this leads to inter-channel scattering (off-diagonal elements in  $\mathbf{t}$ ), which makes it harder to interpret the calculations in simple physical terms.

The understanding can be restored by making suitable unitary transformations of the basis set of incoming and outgoing states in Fig. 2.4 and thereby define a new set of channels — the so-called *eigenchannels* — which *do not mix* [36, 50–52] no matter how strong the (elastic) scattering is! This implies that all off-diagonal elements in the transformed transmission amplitude matrix  $\bar{\mathbf{t}}$  are zero. Since the trace of a matrix is invariant under unitary transformations, this in turn implies that

$$\text{Tr}(\mathbf{t}^\dagger \mathbf{t}) = \text{Tr}(\bar{\mathbf{t}}^\dagger \bar{\mathbf{t}}) = \sum_i |\bar{t}_{ii}|^2.$$

Thus, the two-terminal, many-channel Landauer-Büttiker formula [Eq. (2.21)]

---

<sup>12</sup>This is in close analogy to the propagating modes in an optical fiber or in a microwave guide.

<sup>13</sup>When a violin string is excited, there is no limit to the number of higher harmonics. This is not the case for a constriction because no modes with energies larger than  $E_F$  are allowed. In the violin picture, this would correspond to low-pass filtering the sound, much like listening to a violin concert on an old transistor radio.

<sup>14</sup>For a cylindrical wire, this leads to the sequence  $T_{\text{tot}} = 1, 3, 5, 6, 8, \dots$  as the radius is gradually increased [13, 14, 45, 47–49]. For a wire with a quadratic cross-section, the sequence is changed to  $T_{\text{tot}} = 1, 3, 4, 6, 8, \dots$  [14, 46].

can be simplified to

$$\boxed{G_{3D} = G_0 \sum_i T_i}, \quad (2.30)$$

where  $T_i = |\bar{t}_{ii}|^2$  is the *eigenchannel transmission probability* of the  $i$ th eigenchannel. This transformation is advantageous since the behavior of the conductance is then, again, determined by the transmission probabilities of single, independent channels as was the case in the adiabatic approximation [cf. Eq. (2.29)].<sup>15</sup>

## 2.8 Eigenchannels in experiments

The eigenchannel concept is a powerful tool for visualizing and analyzing the electron transmission through complicated constriction geometries in model calculations [20, 51–59]. Interestingly, the eigenchannels turn out to be more than mathematical abstractions; they can actually be related directly to experiments! In the following, we will mention four newly developed experimental techniques which can be used for extracting information about the eigenchannel transmission probabilities [3, 55, 60–64].

### *I-V* curves of contacts between superconducting electrodes

The *full set* of eigenchannel transmission probabilities  $\{T_i\}$  can actually be deduced from measurements of current-voltage (*I-V*) curves on atomic-scale contacts formed between superconducting electrodes as has been shown by Scheer *et al.* [55, 60]. Hereby, the electronic transport properties of atomic-scale Al, Pb, and Nb contacts have been unraveled in great detail. In contrast, measurements of the total conductance only give information about  $\sum_i T_i$ .

Measurements on atomic-scale contacts between superconducting electrodes are, however, very demanding [63]. This is especially true for free-electron-like metals such as Na, K, Cu, Ag, and Au since these metals are not superconductors. This problem may be circumvented by depositing a very small volume of a free-electron metal (such as Au) between superconducting electrodes (such as Al) [55]. Hereby, the Au develops a gap in its quasiparticle density of states due

---

<sup>15</sup>In the adiabatic limit, the eigenchannels are, in fact, identical to the adiabatic channels [52].

to the proximity effect [65], which is a prerequisite for performing the experiments.<sup>16</sup>

### Shot noise

There are other transport phenomena which are related to the eigenchannel transmission probabilities and do not require superconductivity. One such quantity is the shot noise spectral density  $P_I$  for the current, which arises due to the discreteness of the electron charge. For classical systems, the transfer of electrons is a Poisson process. For such systems, Schottky showed in 1918 [66] that the spectral density of fluctuations in the current  $P_I^{\text{class}}$  is independent of frequency and given by  $2e\langle I \rangle$ , where  $\langle I \rangle$  is the mean current.

In a QPC, the noise  $P_I^{\text{QPC}}$  will be smaller than  $2e\langle I \rangle$  because fully closed or open eigenchannels leave no room for fluctuations in the occupation numbers. For  $eV \gg k_B T$  (where  $V$  is the applied voltage), it can be shown that the shot noise will be suppressed by a factor  $s_I = P_I^{\text{QPC}}/P_I^{\text{class}}$ , which depends on the eigenchannel transmission probabilities in the following simple way [50, 67–69]

$$s_I = \frac{\sum_i T_i(1 - T_i)}{\sum_i T_i}. \quad (2.31)$$

In general,  $s_I < 1$ . It is particularly interesting to note that if all channels are either fully open or closed, we get  $s_I = 0$ .

Simultaneous measurements of both  $s_I$  and  $G$  (or equivalently  $\sum_i T_i$ ) in atomic-size Au and Al contacts have recently been performed by van den Brom and Ruitenbeek [62]. One of their findings is that the current noise in a  $1 G_0$  (monatomic) Au contact is only a few percent of the Schottky noise, proving that the current is carried by a single, almost fully open eigenchannel in good agreement with calculations [20, 52, 53, 59]. In contrast, results from monatomic Al contacts indicate that several partially open eigenchannels contribute to the total transmission. This latter result is also consistent with tight-binding [53, 55] and first-principles [56, 57] calculations as well as measured  $I$ - $V$  curves on superconducting Al QPCs [60], which all show that *three* partially open channels do in fact exist.

---

<sup>16</sup>Until now, this has only been done for Au [55], but unfortunately the conductance in these samples did not exhibit the almost perfect quantization in units of  $G_0$ , which is usually found. This suggests, that the Au sample was disordered in this experiment, as has also been pointed out by the authors [55].

### Conductance and thermopower fluctuations

Another technique focuses on measuring small, random fluctuations in the differential conductance  $G_d(V) = dI/dV(V)$  in a series of stable contact geometries [3, 61, 63]. The origin of the fluctuations is due to the coupling between the ballistic constriction and the diffuse banks, which connects the bare contact with the electrodes.

The effect can be explained qualitatively by the following picture: Electron waves which propagate away from the constriction into the diffuse banks have a finite probability of being backscattered (coherently) to the bare contact by diffusing paths. Back at the constriction, the waves are partially reflected and interfere with the directly transmitted wave and thereby modify the total conductance. The interference will be destructive or constructive depending on the phase accumulated during the diffuse excursion into the banks.

The accumulated phase is determined by the ratio between the path length and the (energy dependent) Fermi wavelength. When a voltage  $V$  is applied, the energy of the transmitted electrons increases by  $eV$ , leading to a small decrease in the Fermi wavelength. Since the path length is usually much larger than the Fermi wavelength, this results in a substantial change in the accumulated phase and thus a measurable change in  $G_d$  (fluctuations of  $\approx 0.0$ – $0.1 G_0$  when the voltage is changed by  $\sim 10$  mV) [63].

These interference effects will occur only for partially open channels ( $0 < T_n < 1$ ): First of all, we only get a directly transmitted wave if  $T_n > 0$ . Secondly, the reflectivity should also be finite ( $1 - T_n > 0$ ) for the secondary wave to backscatter at the constriction and interfere.

A more detailed analysis of this problem shows that the standard deviation of the differential conductance  $\sigma_{GV}$  will depend on the eigenchannel transmission probabilities in the following way [3, 61, 63]

$$\sigma_{GV} \propto \sqrt{\sum_i T_i^2 (1 - T_i)}. \quad (2.32)$$

This formula closely resembles the expression for the shot noise spectral density in Eq. (2.31) in the sense that  $\sigma_{GV} = 0$  when all eigenchannels are either fully open or closed. Measurements of  $\sigma_{GV}$  have been done by Ludoph and Ruitenbeek [3, 61, 63] on Au, Ag, Cu, Na, Nb, and Fe QPCs.

The mechanism used for describing the fluctuations in the conductance also produces fluctuations in the thermopower  $S$ , which is the proportionality constant between temperature difference and induced voltage [61]. The standard

deviation  $\sigma_S$  of these fluctuations depend on  $\{T_i\}$  in the same way as  $\sigma_{GV}$ , see Eq. (2.32) [3].

### Number of eigenchannels in a monatomic contact

By combining all the measurements of eigenchannel transmissions through monatomic contacts, a close connection to the electronic structure of the apex atom has been found: The number of contributing eigenchannels in a monatomic contact is most often equal to and always limited by the number of valence electrons in the atom [55]. This is also in agreement with quantitative tight-binding calculations [53, 70, 71]. Another simple rule that holds so far for the materials that have been considered is the following: The number of contributing eigenchannels in a monatomic contact is equal to the number of bands crossing the Fermi energy in the 1D band structure of the infinite metal chain [20].

## 2.9 Realistic models of atomic-sized contacts

In experiments, the conductance or other transport properties are often measured for a series of constriction geometries formed by, e.g., elongating a metal contact between two metal electrodes. A prerequisite for performing model calculations that can be compared with experiment is therefore an understanding of *how* the constriction geometry changes during such a deformation. Pioneering work on this problem was done in 1990 by Landman *et al.* [72], who used molecular dynamics (MD) simulations for studying the atomistic adhesion and fracture of a Ni tip indenting into and retracting from a Au substrate.

Since then, the sophistication level of modeling and system sizes have grown substantially, and many new results have appeared (see e.g. Ref. 16, 46, 52, 58, 73–88). One of the landmarks was the work by Todorov and Sutton in 1993, who were the first to actually use the atomic configurations from an MD simulation for calculating the conductance during indentation and retraction of a metal tip in a metal surface [74]. Their transport calculations were based on a simple 1s tight-binding (TB) model, in which the electrons are transported by hopping among nearest-neighbor sites [16, 74].

A completely different approach was taken by Olesen *et al.* in 1994 [76], who modeled the QPC as a free-electron gas confined in a hard-wall potential and calculated the conductance from the channel equations in the adiabatic approximation. Unlike the ‘discrete’ TB calculations, this model resulted in clear conductance quantization as the wire was stretched. The quantization

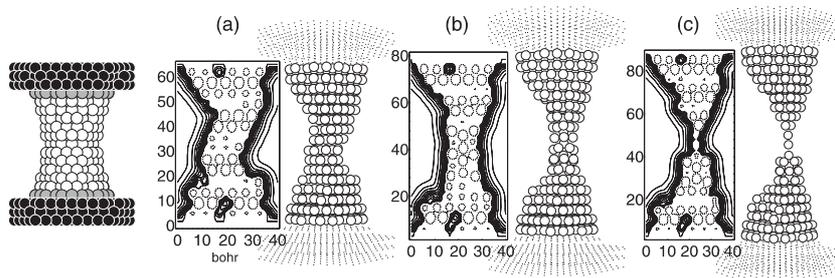


Figure 2.6: Snapshots (a-c) from a low temperature ( $kT = 1$  meV) MD simulation of the elongation of a gold nanocontact performed by M. R. Sørensen. In the initial configuration (left), the atoms are directed along the  $[111]$  direction and bounded by static atomic planes (black). The white atoms are used for constructing an effective one-electron potential, which is shown in cross-sectional cuts through the wire at different stages of the elongation. Constant energy contours are at 0, 1,  $\dots$ , 10 eV. The dashed contours correspond to  $-15$  eV and indicate the positions of the atomic cores closest to the cross-section. (From: [20].)

was subsequently argued to be a natural consequence of the use of the adiabatic approximation and a complete neglect of atomic structure in the one-electron potential [89]. In recent years, improved free-electron calculations have emerged [20, 52, 54] in which the confining potential includes the atomic corrugation inside the contact, and the transmission probabilities are resolved into individual eigenchannels.

### Free-electron calculations on gold contacts

We end this Chapter by showing selected results from such calculations on Au QPCs. Further details and results are given elsewhere [20, 52, 54]. We will refer to them again, when discussing the experimental results on Au QPCs.

Snapshots from an MD simulation of a Au nanocontact which is being elongated are shown in Fig. 2.6. The interatomic forces used in the simulations are calculated within the effective-medium theory [90]. From the atomic coordinates an effective one-electron potential is constructed using the local-density approximation [52]. Cross-sectional contours of the potential are shown for each snapshot in Fig. 2.6.

The structure is initially strained until it becomes mechanically unstable.

This results in rapid atomic rearrangements, which release the strain. The deformation proceeds through a series of elastic and yielding stages during elongation, thereby thinning the wire until the point where only a single atom remains in the smallest cross-section [cf. Fig. 2.6(c)]. Finally, the junction is disrupted. The existence of rapid variations in the potential on a length scale, which is comparable to the Fermi wavelength [ $\lambda_F(\text{Au}) \approx 10$  Bohr] is evident from the contours of the potential in Fig. 2.6. Hence the adiabatic picture is expected to break down.

This is reflected in the calculated eigenchannel transmission probabilities  $\{T_i\}$ , which are shown in the lower panel of Fig. 2.7 as a function of the elongation. When the wire is still fairly thick, as in the initial configuration shown in Fig. 2.6, many eigenchannels contribute to the conductance. As the wire is elongated, the channels are gradually closed. At any given elongation between 0 and 20 Bohr, several eigenchannels are partially open. Thus, when the wire is still fairly thick, the conductance is *not* quantized as this would require that the transmission probabilities were all very close to either one or zero. In a narrow regime between 21 and 24 Bohr, quite a sudden change in the transmission of the eigenchannels occurs, which all become fairly close to being either open or closed. This is again followed by a region with partially open channels. Finally, when just a single or a few atoms remain in the smallest cross-section [see Fig. 2.6(c)], only the eigenchannel with the lowest transverse kinetic energy contributes to the conductance with a transmission value close to one.

If we look at the total transmission, which is proportional to the conductance shown in the middle panel of Fig. 2.7, the conductance is seen to decline rather smoothly during the first 20 Bohr although small jumps occur occasionally due to sudden atomic rearrangements in the constriction. Around the point (a), a plateau close to  $6 G_0$  is seen in the conductance. We might say that this is due to conductance quantization as all transmission probabilities are close to either one or zero at this point.

Features faintly resembling plateaus are seen to develop close to 5, 3, and 2  $G_0$  as the wire is elongated further. However, not all of these are signatures of quantized conductance. For instance, the ‘plateaus’ at 2 and 5  $G_0$  are not due to the presence of an integer number of well defined transverse modes in the wire. Instead, several partially open channels add up to give a total transmission, which happens to be close to an integer. In the end, however, a very well defined plateau occurs close to 1  $G_0$ , which persists for several Bohr of the elongation. This plateau is caused by the almost perfect transmission of a single eigenchannel.

From these calculations, we learn that plateaus close to integer values of  $G_0$

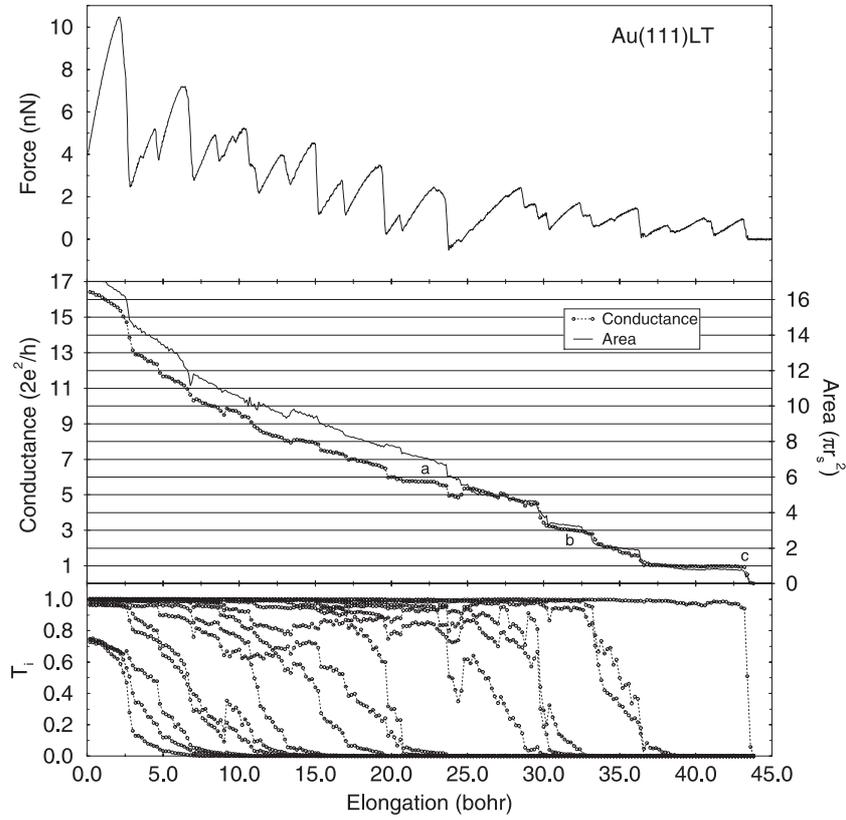


Figure 2.7: Calculated force, conductance, cross-sectional area and eigenchannel transmission probabilities during elongation of the gold nanocontact shown in Fig. 2.6. The points a, b, and c on the conductance curve correspond to the snapshots in the MD simulation. The conductance is calculated using a recursion transfer matrix method. (From: [20].)

are not necessarily caused by quantum size effects. Sometimes they are simply due to the discrete nature of sudden atomic rearrangements. The observation of ‘apparent’ conductance quantization from single conductance curves can therefore not automatically be attributed to ‘real’ quantized conductance caused by the perfect transmission of a limited number of eigenchannels.

## 2.10 Summary

The size of a metal contact determines the method, to be used when calculating the conductance: If the size is much larger than the mean free path of the electrons (the diffuse regime), the differential form of Ohm’s law  $\mathbf{j} = \sigma \mathbf{E}$  is valid from which the conductance can be calculated after solving Laplace’s equation. In this regime, the conductance is proportional to the conductivity of the material. For a circular orifice, the conductance is proportional to the diameter.

When the size of the contact becomes much smaller than the mean free path (the ballistic regime), the electrons are accelerated within a single mean free path in the vicinity of the contact. The conductance will now be proportional to the contact area as well as the area of the Fermi surface projected on a plane normal to the flow of electrons. It will no longer depend of the conductivity. For free-electron metals, the simple Sharvin formula can be used. In the quasiballistic regime, where the size of the contact is comparable to the mean free path, the (inverse) conductance is roughly given by the simple sum of the diffuse and ballistic resistances.

As the size of the contact is decreased towards the Fermi wavelength, the Sharvin result gradually breaks down due to quantum size effects. This effect can be included in the Sharvin formula by making so-called perimeter and soft-wall corrections whereby the modified formula can be used for *estimating* the contact area from the conductance, provided the conductance is larger than  $3\text{--}4 G_0$ .

For contacts with a size comparable to the Fermi wavelength, interference effects from scattering at the boundary of the constriction have to be considered in more detail. In this regime the two-terminal Landauer-Büttiker formula is well suited for calculating the conductance. The Landauer-Büttiker formula relates the conductance (which is conveniently measured in units of  $G_0 = 2e^2/h$ ) to quantum mechanical transmission amplitudes of electronic wave functions.

Within the independent electron approximation, the transmission can be found by, e.g., solving the coupled channel equations. If cross-sections of the one-

electron potential normal to the direction of current flow varies smoothly on the scale of the Fermi wavelength — the adiabatic limit, the conductance becomes quantized in units of  $G_0$ . Even when strong (elastic) scattering is present, defining a set of independent conduction channels — the so-called eigenchannels, is possible. A transmission probability  $T_i$  can be ascribed to each eigenchannel. From measurements of non-linear  $I$ - $V$  characteristics on superconducting QPCs, the extraction of the full set  $\{T_i\}$  is possible. From measurements on the shot noise and fluctuations in the differential conductance and thermopower, it can be tested whether the total transmission is due to an underlying set of fully open or closed eigenchannels, or due to a set of several partially open eigenchannels.

Realistic models of Au QPCs with only a few atoms in the smallest cross-section, indicate that the current tends to be carried by a set of almost fully open or closed eigenchannels. For a single atom Au contact the transmission is especially seen to be mediated by a single, almost fully open eigenchannel in good agreement with recent measurements of shot noise and fluctuations in the differential conductance and thermopower.

## References

- [1] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [2] D. J. Griffiths, *Introduction to Electrodynamics*, 2nd ed. (Prentice Hall, New Jersey, 1989).
- [3] B. Ludoph and J. M. van Ruitenbeek, “Conductance fluctuations as a tool for investigating the quantum modes in atomic size metallic contacts,” *Phys. Rev. B* **61**, 2273–2285 (2000).
- [4] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (John Wiley, New York, 1975).
- [5] J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Oxford University Press, London, 1892), Vol. 1, p. 426–434.
- [6] A. N. Omelyanchuk, I. O. Kulik, and R. I. Shekhter, “Contribution to the theory of nonlinear effects in the electric conductivity of metallic junctions,” *Pis'ma Zh. Eksp. Teor. Fiz.* **25**, 465–469 (1977), [*JEPT Lett.* **25**, 437 (1977)].
- [7] D. Sánchez-Portal, C. Untiedt, J. M. Soler, J. J. Sáenz, and N. Agraït, “Nanocontacts: Probing Electronic Structure under Extreme Uniaxial Strains,” *Phys. Rev. Lett.* **79**, 4198–4201 (1997).
- [8] I. O. Kulik, A. N. Omel'yanchuk, and R. I. Shekhter, “Electrical conductivity of point contact microbridges and phonon and impurity spectroscopy in normal metals,” *Fiz. Nizk. Temp.* **3**, 1543–1558 (1977), [*Sov. J. Low Temp. Phys.* **3**, 740 (1977)].
- [9] A. G. M. Jansen, A. P. van Gelder, and P. Wyder, “Point-contact spectroscopy in metals,” *J. Phys. C* **13**, 6073–6118 (1980).
- [10] C. Untiedt, G. Rubio, S. Vieira, and N. Agraït, “Fabrication and characterization of metallic nanowires,” *Phys. Rev. B* **56**, 2154–2160 (1997).
- [11] S. Krompiewski, “Conductance of mesoscopic magnetic systems,” *Acta Phys. Pol. A* **97**, 15–24 (2000).
- [12] Y. V. Sharvin, “A possible method for measuring Fermi surfaces,” *Zh. Eksp. Teor. Fiz.* **48**, 984–985 (1965), [*Sov. Phys. JETP* **21**, 655 (1965)].

- [13] J. A. Torres, J. L. Pascual, and J. J. Sáenz, “Theory of conduction through narrow constrictions in a three-dimensional electron gas,” *Phys. Rev. B* **49**, 16581–16584 (1994).
- [14] A. García-Martín, J. A. Torres, and J. J. Sáenz, “Finite size corrections to the conductance of ballistic wires,” *Phys. Rev. B* **54**, 13448–13451 (1996).
- [15] J. A. Torres and J. J. Sáenz, “Conductance and Mechanical Properties of Atomic-Size Metallic Contacts: A Simple Model,” *Phys. Rev. Lett.* **77**, 2245–2248 (1996).
- [16] A. M. Bratkovsky, A. P. Sutton, and T. N. Todorov, “Conditions for conductance quantization in realistic models of atomic-scale metallic contacts,” *Phys. Rev. B* **52**, 5036–5051 (1995).
- [17] E. N. Bogachek, A. G. Scherbakov, and U. Landman, “Shape effects on conductance quantization in three-dimensional nanowires: Hard versus soft potentials,” *Phys. Rev. B* **56**, 1065–1068 (1997).
- [18] G. Wexler, “The size effect and the non-local Boltzmann transport equation in orifice and disk geometry,” *Proc. Phys. Soc. London* **89**, 927–941 (1966).
- [19] M. J. M. de Jong, “Transition from Sharvin to Drude resistance in high-mobility wires,” *Phys. Rev. B* **49**, 7778–7781 (1994).
- [20] M. Brandbyge, “Conductance of Metal Nanocontacts,” PhD thesis, Technical University of Denmark, 1997.
- [21] R. Landauer, “Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction,” *IBM J. Res. Dev.* **1**, 223–231 (1957).
- [22] R. Landauer, “Electrical Resistance of Disordered One-dimensional Lattices,” *Philos. Mag.* **21**, 863–867 (1970).
- [23] M. Y. Azbel, “Generalized Landauer formula,” *Phys. Lett.* **78A**, 410–412 (1980).
- [24] P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, “New method for a scaling theory of localization,” *Phys. Rev. B* **22**, 3519–3526 (1980).
- [25] E. N. Economou and C. M. Soukoulis, “Static Conductance and Scaling Theory of Localization in One Dimension,” *Phys. Rev. Lett.* **46**, 618–621 (1981).

- [26] P. W. Anderson, “New method for scaling theory of localization. II. Multichannel theory of a “wire” and possible extensions to higher dimensionality,” *Phys. Rev. B* **23**, 4828–4836 (1981).
- [27] D. S. Fisher and P. A. Lee, “Relation between conductivity and transmission matrix,” *Phys. Rev. B* **23**, 6851–6854 (1981).
- [28] M. Y. Azbel, “Quantum  $\delta$ -dimensional Landauer formula,” *J. Phys. C* **14**, L225–L230 (1981).
- [29] R. Landauer, “Can a length of perfect conductor have a resistance?,” *Phys. Lett.* **85A**, 91–93 (1981).
- [30] D. C. Langreth and E. Abrahams, “Derivation of the Landauer conductance formula,” *Phys. Rev. B* **24**, 2978–2984 (1981).
- [31] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, “Generalized many-channel conductance formula with application to small rings,” *Phys. Rev. B* **31**, 6207–6215 (1985).
- [32] Y. Imry, “Physics of mesoscopic systems,” in *Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenko, eds., (World Scientific, Singapore, 1986), pp. 101–163.
- [33] M. Büttiker, “Role of quantum coherence in series resistors,” *Phys. Rev. B* **33**, 3020–3026 (1986).
- [34] M. Büttiker, “Voltage fluctuations in small conductors,” *Phys. Rev. B* **35**, 4123–4126 (1987).
- [35] R. Landauer, “Electrical Transport in Open and Closed Systems,” *Z. Phys. B* **68**, 217–228 (1987).
- [36] M. Büttiker, “Coherent and sequential tunneling in series barriers,” *IBM J. Res. Dev.* **32**, 63–75 (1988).
- [37] A. D. Stone and A. Szafer, “What is measured when you measure a resistance? — The Landauer formula revisited,” *IBM J. Res. Dev.* **32**, 384–413 (1988).
- [38] R. Landauer, “Conductance determined by transmission: probes and quantised constriction resistance,” *J. Phys. Cond. Matt.* **1**, 8099–8110 (1989).

- [39] M. Büttiker, “Transmission, reflection and the resistance of small conductors,” In *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, J. M. Chamberlain, L. Eaves, and J.-C. Portal, eds., Series B: Physics **231**, 51–73 (Plenum, New York, 1990).
- [40] M. Yosefin and M. Kaveh, “Conduction in Curvilinear Constrictions: Generalization of the Landauer Formula,” *Phys. Rev. Lett.* **64**, 2819–2822 (1990).
- [41] X.-G. Zhang and W. H. Butler, “Landauer conductance in the diffusive regime,” *Phys. Rev. B* **55**, 10308–10318 (1997).
- [42] Y. Imry and R. Landauer, “Conductance viewed as transmission,” *Rev. Mod. Phys.* **71**, S306–S312 (1999).
- [43] J. Frenkel, “On the electrical resistance of contacts between solid conductors,” *Phys. Rev.* **36**, 1604–1618 (1930).
- [44] L. I. Glazman, G. B. Lesovik, D. E. Khmel’nitskii, and R. I. Shekhter, “Reflectionless quantum transport and fundamental ballistic-resistance steps in microscopic constrictions,” *Pis’ma Zh. Eksp. Teor. Fiz.* **48**, 218–220 (1988), [*JETP Lett.* **48**, 238 (1988)].
- [45] É. N. Bogachek, A. N. Zagoskin, and I. O. Kulik, “Conductance jumps and magnetic flux quantization in ballistic point contacts,” *Fiz. Nizk. Temp.* **16**, 1404–1411 (1990), [*Sov. J. Low Temp. Phys.* **16**, 796 (1990)].
- [46] M. Brandbyge, J. Schiøtz, M. R. Sørensen, P. Stoltze, K. W. Jacobsen, J. K. Nørskov, L. Olesen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, “Quantized conductance in atom-sized wires between two metals,” *Phys. Rev. B* **52**, 8499–8514 (1995).
- [47] J. M. Kras, J. M. van Ruitenbeek, V. V. Fisun, I. K. Yanson, and L. J. de Jongh, “The signature of conductance quantization in metallic point contacts,” *Nature* **375**, 767–769 (1995).
- [48] A. G. Scherbakov, E. N. Bogachek, and U. Landman, “Quantum electronic transport through three-dimensional microconstrictions with variable shapes,” *Phys. Rev. B* **53**, 4054–4064 (1996).
- [49] M. Brandbyge, K. W. Jacobsen, and J. K. Nørskov, “Scattering and conductance quantization in three-dimensional metal nanocontacts,” *Phys. Rev. B* **55**, 2637–2650 (1997).

- [50] T. Martin and R. Landauer, “Wave-packet approach to noise in multichannel mesoscopic systems,” *Phys. Rev. B* **45**, 1742–1755 (1992).
- [51] M. Brandbyge and K. W. Jacobsen, “Quantum transmission channels in perturbed 3D nanowires,” In *Proceedings of the NATO Advanced Research Workshop on Nanowires, Madrid 1996*, P. A. Serena and N. García, eds., NATO ASI pp. 61–78 (Kluwer, Dordrecht, 1997).
- [52] M. Brandbyge, M. R. Sørensen, and K. W. Jacobsen, “Conductance eigenchannels in nanocontacts,” *Phys. Rev. B* **56**, 14956–14959 (1997).
- [53] J. C. Cuevas, A. L. Yeyati, and A. Martín-Rodero, “Microscopic Origin of Conducting Channels in Metallic Atomic-Size Contacts,” *Phys. Rev. Lett.* **80**, 1066–1069 (1998).
- [54] M. Brandbyge and M. Tsukada, “Local density of states from transmission amplitudes in multichannel systems,” *Phys. Rev. B* **57**, R15088–R15091 (1998).
- [55] E. Scheer, N. Agraït, J. C. Cuevas, A. L. Yeyati, B. Ludoph, A. Martín-Rodero, G. R. Bollinger, J. M. van Ruitenbeek, and C. Urbina, “The signature of chemical valence in the electrical conduction through a single-atom contact,” *Nature* **394**, 154–157 (1998).
- [56] N. Kobayashi, M. Brandbyge, and M. Tsukada, “Conductance through Atoms: Dot or Channel?,” *Jpn. J. of Appl. Phys.* **38**, 336–338 (1999), part 1.
- [57] N. Kobayashi, M. Brandbyge, and M. Tsukada, “Transmission channels through Na and Al atom wire,” *Surf. Sci.* **433–435**, 854–857 (1999).
- [58] A. Nakamura, M. Brandbyge, L. B. Hansen, and K. W. Jacobsen, “Density Functional Simulation of a Breaking nanowire,” *Phys. Rev. Lett.* **82**, 1538–1541 (1999).
- [59] M. Brandbyge, N. Kobayashi, and M. Tsukada, “Conduction channels at finite bias in a single-atom gold contact,” *Phys. Rev. B* **60**, 17064–17070 (1999).
- [60] E. Scheer, P. Joyez, D. Esteve, C. Urbina, and M. H. Devoret, “Conduction Channel Transmissions of Atomic-Size Aluminum Contacts,” *Phys. Rev. Lett.* **78**, 3535–3538 (1997).

- [61] B. Ludoph, M. H. Devoret, D. Esteve, C. Urbina, and J. M. van Ruitenbeek, “Evidence for Saturation of Channel Transmission from Conductance Fluctuations in Atomic-Size Point Contacts,” *Phys. Rev. Lett.* **82**, 1530–1533 (1999).
- [62] H. E. van den Brom and J. M. van Ruitenbeek, “Quantum Suppression of Shot Noise in Atom-Size Metallic Contacts,” *Phys. Rev. Lett.* **82**, 1526–1529 (1999).
- [63] B. Ludoph, “Quantum Conductance Properties of Atomic-size Contacts,” PhD thesis, Rijksuniversiteit Leiden, 1999.
- [64] B. Ludoph and J. M. van Ruitenbeek, “Thermopower of atomic-size metallic contacts,” *Phys. Rev. B* **59**, 12290–12293 (1999).
- [65] W. Belzig, C. Bruder, and G. Schön, “Local density of states in a dirty normal metal connected to a superconductor,” *Phys. Rev. B* **54**, 9443–9448 (1996).
- [66] W. Schottky, “Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern,” *Ann. Phys. (Leipzig)* **57**, 541–567 (1918).
- [67] M. Büttiker, “Scattering Theory of Thermal and Excess Noise in Open Conductors,” *Phys. Rev. Lett.* **65**, 2901–2904 (1990).
- [68] M. Büttiker, “Scattering theory of current and intensity noise correlations in conductors and wave guides,” *Phys. Rev. B* **46**, 12485–12507 (1992).
- [69] J. Bürki and C. A. Stafford, “Comment on “Quantum Suppression of Shot Noise in Atomic-Size Metallic Contacts”,” *Phys. Rev. Lett.* **83**, 3342 (1999).
- [70] C. Sirvent, J. G. Rodrigo, S. Vieira, L. Jurczyszyn, N. Mingo, and F. Flores, “Conductance step for single-atom contact in the scanning tunneling microscope: Noble and transition metals,” *Phys. Rev. B* **53**, 16086–16090 (1996).
- [71] A. L. Yeyati, A. Martín-Rodero, and F. Flores, “Conductance quantization and electron resonances in sharp tips and atomic-size contacts,” *Phys. Rev. B* **56**, 10369–10372 (1997).
- [72] U. Landman, W. D. Luedtke, N. A. Burnham, and R. J. Colton, “Atomistic Mechanisms and Dynamics of Adhesion, Nanoindentation, and Fracture,” *Science* **248**, 454–461 (1990).

- [73] A. P. Sutton and J. B. Pethica, “Inelastic flow processes in nanometre volumes of solids,” *J. Phys. C* **2**, 5317–5326 (1990).
- [74] T. N. Todorov and A. P. Sutton, “Jumps in Electronic Conductance Due to Mechanical Instabilities,” *Phys. Rev. Lett.* **70**, 2138–2141 (1993).
- [75] A. P. Sutton and T. N. Todorov, “Mechanical and electrical properties of metallic contacts at the nanometer scale,” *J. Phys. Chem. Solids* **55**, 1169–1174 (1994).
- [76] L. Olesen, E. Lægsgaard, I. Stensgaard, F. Besenbacher, J. Schiøtz, P. Stoltze, K. W. Jacobsen, and J. K. Nørskov, “Quantized Conductance in an Atom-Sized Point Contact,” *Phys. Rev. Lett.* **72**, 2251–2254 (1994).
- [77] U. Landman, W. D. Luedtke, B. E. Salisbury, and R. L. Whetten, “Reversible Manipulations of Room Temperature Mechanical and Quantum Transport Properties in Nanowire Junctions,” *Phys. Rev. Lett.* **77**, 1362–1365 (1996).
- [78] M. Heinemann and R. A. de Groot, “Aluminum break-point contacts,” *Phys. Rev. B* **55**, 9375–9378 (1997).
- [79] R. N. Barnett and U. Landman, “Cluster-derived structures and conductance fluctuations in nanowires,” *Nature* **387**, 788–790 (1997).
- [80] M. R. Sørensen, M. Brandbyge, and K. W. Jacobsen, “Mechanical deformation of atomic-scale metallic contacts: Structure and mechanisms,” *Phys. Rev. B* **57**, 3283–3294 (1998).
- [81] O. Gülseren, F. Ercolessi, and E. Tosatti, “Noncrystalline Structures of Ultrathin Unsupported Nanowires,” *Phys. Rev. Lett.* **80**, 3775–3778 (1998).
- [82] U. Landman, R. N. Barnett, and W. D. Luedtke, “Nanowires: size evolution, reversibility, and one-atom contacts,” *Z. Phys. D* **40**, 282–287 (1997).
- [83] H. Mehrez, S. Ciraci, C. Y. Fong, and Ş. Erkoç, “An atomistic study on the stretching of nanowires,” *J. Phys. C* **9**, 10843–10854 (1997).
- [84] Ç. Kılıç, H. Mehrez, and S. Ciraci, “Quantum point contact on graphite surface,” *Phys. Rev. B* **58**, 7872–7881 (1998).
- [85] G. Taraschi, J.-L. Mozos, C. C. Wan, H. Guo, and J. Wang, “Structural and transport properties of aluminum atomic wires,” *Phys. Rev. B* **58**, 13138–13145 (1998).

- [86] G. Bilalbegović, “Structure and stability of finite gold nanowires,” *Phys. Rev. B* **58**, 15412–15415 (1998).
- [87] J. A. Torres, E. Tosatti, A. Dal Corso, F. Ercolessi, J. J. Kohanoff, F. D. Di Tolla, and J. M. Soler, “The puzzling stability of monatomic gold wires,” *Surf. Sci.* **426**, L441–L446 (1999).
- [88] H. Häkkinen, R. N. Barnett, and U. Landman, “Gold Nanowires and their Chemical Modifications,” *J. Phys. Chem. B* **103**, 8814–8816 (1999).
- [89] J. M. Kras, C. J. Muller, N. van der Post, F. R. Postma, A. P. Sutton, T. N. Todorov, and J. M. van Ruitenbeek, “Comment on “Quantized Conductance in an Atom-Sized Point Contact”,” *Phys. Rev. Lett.* **74**, 2146 (1995).
- [90] K. W. Jacobsen, J. K. Nørskov, and M. J. Puska, “Interatomic interactions in the effective-medium theory,” *Phys. Rev. B* **35**, 7423–7442 (1987).



## Chapter 3

# Quantized Conductance in Relays

This Chapter covers the material which was published in papers [II] and [I].<sup>1</sup> These papers were submitted more than three years ago at a time when the theoretical concept of conductance eigenchannels (discussed on page 26 in Chapter 2) was just emerging [1–4], but before it was realized that eigenchannels could be related directly to experiments [5].

At this time the widely (ab)used term ‘quantized conductance’ (QC) had a less rigorous meaning than today. Most researchers regarded the conductance as ‘quantized’ if the conductance had a statistical preference for being close to an integer value of  $G_0$ . The concept of conductance histograms [6] was widely used, where measurements of the conductance in a large ensemble of constriction geometries were accumulated into probability distributions. A series of peaks close to integer values of  $G_0$  was then taken as evidence for QC.

These experiments, however, only showed that the total transmission  $T_{\text{tot}}$  was close to an integer value. Nowadays, the conductance is normally said to be quantized if all eigenchannel probabilities  $T_i$  are very close to either zero or one [7]. Of course, such a conclusion cannot solely be inferred from the observation that  $T_{\text{tot}}$  is close to an integer since this could be a coincidence: There may be certain preferred contact geometries, which are characterized by a set of partially open eigenchannels which *happen* to have a total transmission close to an integer [7]. Today we know that this is actually the case for Al [5, 7–10].

---

<sup>1</sup>The preliminary results in Paper [I] form a subset of the results in paper [II].

### 3.1 Why study relays?

Our studies of the conductance of relay contacts were inspired by a very interesting article published by Costa-Krämer *et al.* in 1995 [11]. Here, the authors showed that QC could be observed simply by recording the conductance through breaking contacts between macroscopic wires in a simple table top experiment.

At first, we were sceptical about their approach but most of all, we were provoked by several statements in their paper such as: “The formation of these nanowires and the associated quantized conductance is a universal phenomenon that occurs when any two metals get in and out of contact . . . ” and when discussing a conductance curve for Au “ . . . for Au all the curve is quantized up to the 40 quanta level!”.

The latter statement was clearly not consistent with a conductance histogram for Au published at about the same time from our group [12]. This particular histogram was accumulated from point contacts formed using an STM under clean UHV conditions and did not contain any peaks beyond  $4 G_0$ . The experimental findings were supplemented by extensive molecular dynamics simulations from which a similar conductance histogram could be constructed.

We found it hard to understand how results obtained under ambient conditions could display a much clearer quantization than UHV experiments. Thus, we wanted to reproduce the results by Costa-Krämer *et al.* However, instead of making and breaking the contacts by ‘tapping the table top’, we decided to use electro-mechanical relays instead because these contacts could be broken automatically by applying a periodic signal to the relay coil.

In the following sections we present the results from these studies in a form which closely resembles the original work (paper [II]). At the end of this Chapter, we will discuss how the results and views presented in paper [II] relates to later work.

### 3.2 Experimental details

#### Setup

Our experimental setup is shown in Fig. 3.1. Commercial and home-built electro-mechanical relays represented by interchangeable modules can be studied using the same equipment for breaking the contact and studying the time-evolution of the conductance. In the case of commercial relays, a contact is closed and opened by a magnetic force from a coil in the relay [13]. The driving

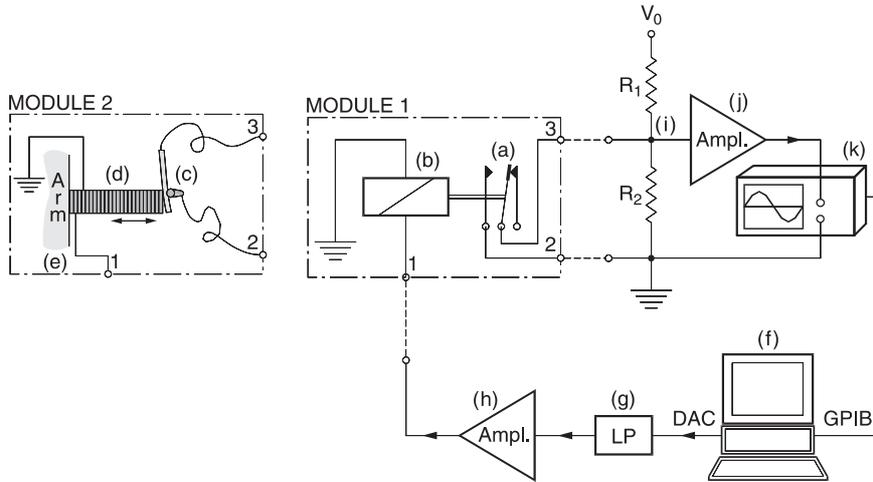


Figure 3.1: Schematic diagram of the experimental setup. Two types of contacts corresponding to the modules 1 and 2 can be studied interchangeably. In the commercial relay (module 1) a normally open contact (a) can be closed and reopened by changing the voltage on the relay coil (b). In the home-built relay (module 2) two wires (c) get in and out of contact by means of a piezo-electric actuator (d), which is mounted on an adjustable lever (e). The switch state of the contact in the commercial relay (module 1) as well as the length of the actuator in the home-built relay (module 2) is computer controlled from a Personal Computer (f) via a digital to analog converter, a low pass filter (g), and an amplifier or buffer (h). Together with  $R_1$  and  $R_2$  the chosen contact [(a) or (c)] form a voltage divider (i) with supply voltage  $V_0$  (15 V). The divided voltage is amplified (j) and fed into a digital storage oscilloscope (k).

coil voltage is set via a 12 bit digital to analog converter (DAC) installed in a Personal Computer (PC). The DAC voltage passes a low pass filter and an amplifier to produce smooth voltage ramps.

In the home-built relay, the contact is made of two thin, pure and polycrystalline metal wires (0.25 mm diameter, 99.99% purity) mounted in a crossed cylinder geometry. One of the wires is held in a fixed position, whereas the other is resting against an isolated, low voltage (0–100 V) piezo-electric actuator (Tokin NLA-2×3×9 mm<sup>3</sup>). Adjusting the gap between the wires to less than 1  $\mu\text{m}$  is possible by means of a mechanical coupling utilizing a precision micrometer. Finer adjustments are performed by applying a suitable voltage to

the actuator, which can be elongated up to  $7 \mu\text{m}$ . The smoothed DAC voltage is amplified and connected to the actuator, thereby making it possible to move the wires very gently in and out of point contact.

Individual measurements are carried out slightly differently depending on the chosen module. In Fig. 3.2 we show the procedure when a commercial relay (module 1) is under investigation. The switch state of the relay depends on whether the voltage applied to the relay coil  $V_{\text{coil}}$  is above the pull-in voltage  $V_{\text{close}}$  or below the drop-out voltage  $V_{\text{open}}$ . The normally open contact is closed by raising  $V_{\text{coil}}$  above  $V_{\text{close}}$ . After a transfer time of less than 1 ms, the electrodes collide with a speed of a few m/s [14]. In the closed position, typical

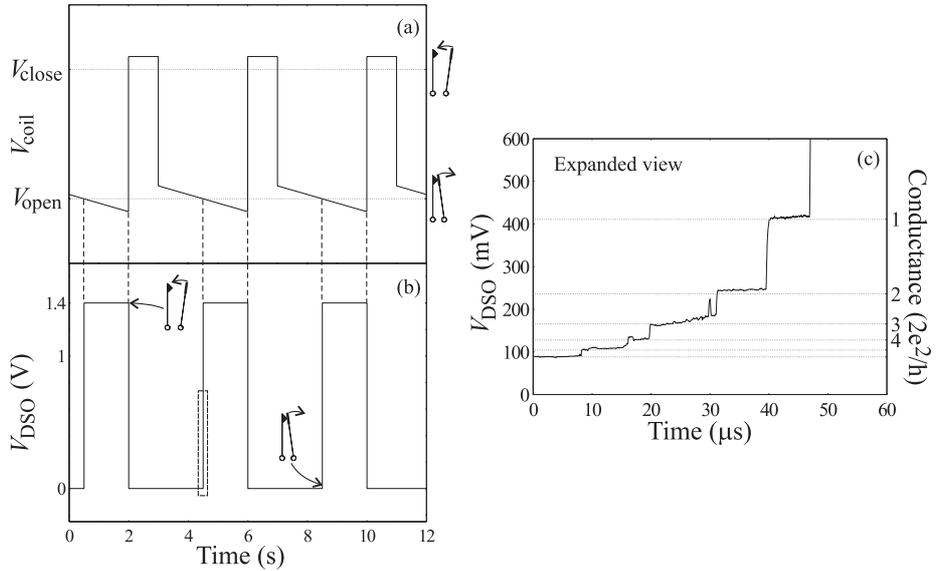


Figure 3.2: Contact-stretch-disruption cycles for the study of breaking point contacts in commercial relays. Panel (a) shows schematically the time-evolution of the voltage  $V_{\text{coil}}$  being applied to the relay coil. The pull-in and drop-out voltages ( $V_{\text{close}}$  and  $V_{\text{open}}$ ) are shown as horizontal lines. Panel (b) depicts the resulting voltage drop across the contact  $V_{\text{DSO}}$ . Part of an up-going flank, indicated by a dashed box in panel (b), has been expanded with a factor of  $2 \times 10^5$  in time in panel (c). The horizontal grid lines indicate the voltages corresponding to integer values of the conductance measured in units of  $2e^2/h$ , as calculated from Eq. (3.1).

contact forces are approximately 0.1 N, corresponding to a contact area of approximately  $100 (\mu\text{m})^2$  and a contact resistance of less than  $1 \text{ m}\Omega$  [13, 14]. The contact is gently reopened by ramping the coil voltage slowly below  $V_{\text{open}}$ .

When the contact is open, the voltage measured by the digital storage oscilloscope (DSO)  $V_{\text{DSO}}$  is high with a value determined by the voltage divider  $R_1$  and  $R_2$  [Fig. 3.1(i)] and the amplifier [Fig. 3.1(j)]. When the contact is closed, it short circuits  $R_2$ , implying  $V_{\text{DSO}} = 0 \text{ V}$ . If the voltage signal  $V_{\text{DSO}}$  is monitored on a time-scale of several seconds, no interesting features are observed [see Fig. 3.2(b)]. However, if the time-varying voltage signal is expanded dramatically in time on the rising voltage flanks (stretch of contact), the voltage is seen to increase in steps on a time-scale of less than  $100 \mu\text{s}$  before the contact is finally broken [see the expanded view in Fig. 3.2(c)].

When a home-built relay (module 2) is under investigation, basically the same procedure is used. Now, instead of using a magnetic coil for controlling the switch state of the contact, an amplified voltage is applied to a piezoelectric actuator. The contact is closed by raising the piezo voltage, thereby extending the actuator  $\approx 100 \text{ nm}$  beyond the recorded position of the previous disruption, followed by a contraction done with a speed of  $\approx 100 \text{ nm/s}$ . In the home-built relays, the contact is stretched during 0.1–10 ms, depending on how the wires are mounted and which contact material is being used.

Independent of the chosen module, a DSO is set to trigger when the contact breaks, and the voltage trace is transferred to the PC and stored for later analysis. Usually there are 1000 data points in a single trace. It takes 3–4 s to complete one contact-stretch-disruption cycle, and the time-evolution of the conductance in approximately 25 000 breaking point contacts can be recorded during 24 h.

By measuring the voltage drop across the contact, the conductance is found. The current through the contact is limited to 4–40  $\mu\text{A}$  by suitable choice of  $R_1$ . The resistor  $R_2$  limits the voltage across the contact to avoid saturation problems in the amplifier and to decrease the rise-time of the circuit to approximately  $1 \mu\text{s}$ . Typical values of the resistances are  $R_1 = 400 \text{ k}\Omega$  and  $R_2 = 40 \text{ k}\Omega$ .

The output from the amplifier is fed into a 10-bit, 100 Msamples/s, four-channel Yokogawa DL 4100 DSO. The DSO is fully computer controlled via a General Purpose Interface Bus (GPIB). We have used a 10-bit DSO instead of the more conventional 8-bit models in order to obtain the highest possible resolution in our conductance measurements. We have abstained from using enhanced resolution techniques to increase the true bit resolution, as a compensation for differential non-linearity is thereby rendered impossible which will be explained on page 52.

The critical components in the DSO are four built-in analog-to-digital converters (ADCs) which are used for digitizing the measured voltages on the four input lines. Each of these ADCs has a resolution of 10 bits. Thus, on every input line, the measured voltages are binned into  $n = 2^{10} = 1024$  different values  $V_i$  ( $1 \leq i \leq n$ ). In the further analysis of the traces, the  $V_i$  values are transformed into corresponding conductance values  $G_i$  using the equation

$$G_i = \frac{I_0}{V_i - V_{\text{off}}} - G_R, \quad (3.1)$$

where  $V_{\text{off}}$  is a small offset voltage that may accumulate in the voltage amplifier and DSO. The current  $I_0$  and the conductance  $G_R$  can be derived from the circuit in Fig. 3.1 as  $I_0 = AV_0/R_1$  and  $G_R = (1/R_1 + 1/R_2)$ , where  $A$  is the gain of the voltage amplifier. The values of  $I_0$ ,  $G_R$ , and  $V_{\text{off}}$  are determined accurately by replacing the wires with calibrated conductances and fitting the measured voltages to Eq. (3.1). To obtain a high resolution in conductance over a wide range of conductances, we make simultaneous measurements on two input lines on the DSO with different V/div settings. Later, on page 57 we will demonstrate how conductance histograms can be used for determining the calibration accuracy and the instrumental broadening in a very direct way.

In Fig. 3.3, we show two conductance traces acquired during the opening of the contact in a commercial AuCo5 relay.<sup>2</sup> In Fig. 3.3(a) is depicted a type of conductance trace which is often found, whereas in Fig. 3.3(b) we show a specially selected trace with a clear staircase structure and sharply defined plateaus very close to integer values of  $G_0$ . The detailed structure of the conductance traces differs a lot, but the general trend is that quantization is observed only for low integer values of  $G_0$ , as in Fig. 3.3(a). Further discussion of the physics underlying the detailed structure of the traces will be deferred to Sec. 3.3. For now, we concentrate on how to deal with a large ensemble of traces in general.

### The histogram method

The formation of nanowires is inherently connected with plastic deformation and a change of the geometry in the contact region. This results in new geometries every time a constriction is made, meaning that a new trajectory in the phase

---

<sup>2</sup>The relay is manufactured by Matsushita Automation Controls. According to the manufacturer, the contact material is a special Au/Co alloy with 5% Co (AuCo5). To protect company know-how, further details could not be given. We note that Au and Co do *not* alloy in equilibrium at room temperature, meaning that the alloy is segregated into almost pure Au and Co phases *or* the alloy is a quenched non-equilibrium structure.

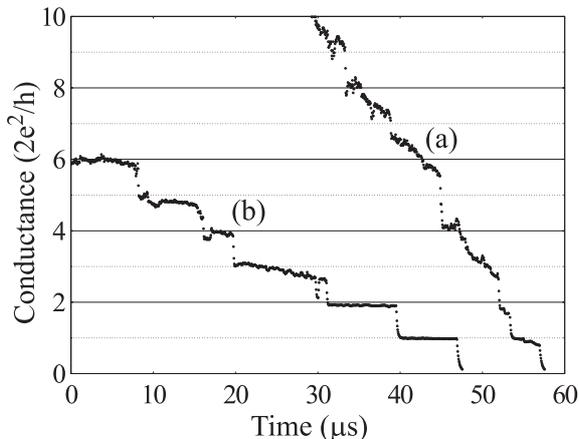


Figure 3.3: Two examples of the time evolution of the conductance through a breaking contact in a commercial AuCo5 relay with arbitrary zero point in time. A trace (a) with low degree of quantization and an exceptional trace (b) showing almost perfect quantization up to  $6 G_0$ . The exceptional trace is equivalent with the voltage trace shown in Fig. 3.2(c).

space of constriction geometries is probed in every contact-stretch-disruption cycle. A specific geometry results in a particular value of the conductance, but since single trajectories may follow very different paths in phase space, this results in very diverse appearances of single conductance traces.

On the other hand, these trajectories will not be distributed randomly in phase space as the mechanical properties of the constriction may lead to certain preferred geometries especially in the last stages before the contact is broken, as has been indicated in molecular dynamics simulations [12, 15–18]. Therefore some values of the conductance may occur more often than others; not only because of the non-uniform distribution or bundling of the trajectories in phase space caused by stable constriction geometries, but also because a wide range of constriction geometries may correspond to essentially the same conductance value. The extreme case is here a free-electron QPC with an ideal, adiabatic geometry. As the width of the constriction is varied continuously, the conductance settles exactly at integer values of  $G_0$ , whereas in a more realistic geometry in which the effects of surface roughness, impurities and a larger opening angle are taken into account, the sharp staircase structure will be blurred, and the

positions of the plateaus will be shifted downwards [3, 19, 20].

Because of the very diverse behavior of single conductance traces, it has become a widespread practice to supplement investigations of single traces with histograms accumulated from many traces in order to reveal the general conductance behavior of the nanowires formed [6, 11, 12, 19, 21–24]. The reason is that an analysis based only on single representative traces is insufficient, because defining what *representative* means is difficult in this context. The idea behind the histogram technique is to project many traces onto the conductance axis and make a histogram showing the probability distribution of the conductance. As a starting point, we have chosen to include *all* traces in the histogram.

We define histogram values  $H(G_i)$ , corresponding to each of the conductance values  $G_i$  in analogy with a previously published method as follows [6, 12]

$$H(G_i) \equiv \frac{N_i}{N\Delta G_i}, \quad (3.2)$$

where  $N_i$  is the number of counts in bin number  $i$  accumulated from all traces in a single experiment,  $N$  is the total number of data points, and  $\Delta G_i$  is the conductance bin width.

In Fig. 3.4(a) we show a histogram generated from measurements on a commercial AuCo5 relay. The histogram is accumulated from 50 000 conductance traces measured during 25 000 contact-stretch-disruption cycles (one trace per V/div setting per disruption) corresponding to 50 million individual measurements of the conductance. Distributing the 50 million measurements in the  $2 \times 1024$  bins implies that the average number of counts in a single bin  $i$  is more than 20 000. Thus, the root mean square (RMS) fluctuations due to counting statistics for a single histogram point  $H(G_i)$  should be less than 1%. Surprisingly, the histogram is very noisy (30% RMS) and the peaks in the histogram are not well defined. We have realized that the noise is associated with differential non-linearity of the fast, high-resolution ADC's, that are used in the DSO. Different aspects of differential non-linearity will be described in detail below — most importantly, we shall see how the noise can be removed in a proper way.

## Differential non-linearity

### Digital storage oscilloscopes

Differential non-linearity is a measure of the standard variation of the incremental voltage bin widths  $\Delta V_i$  of each bin  $i$  on an ADC [25, 26] that could be used in, e.g., an input line of a DSO. To illustrate in a simplified way how differential

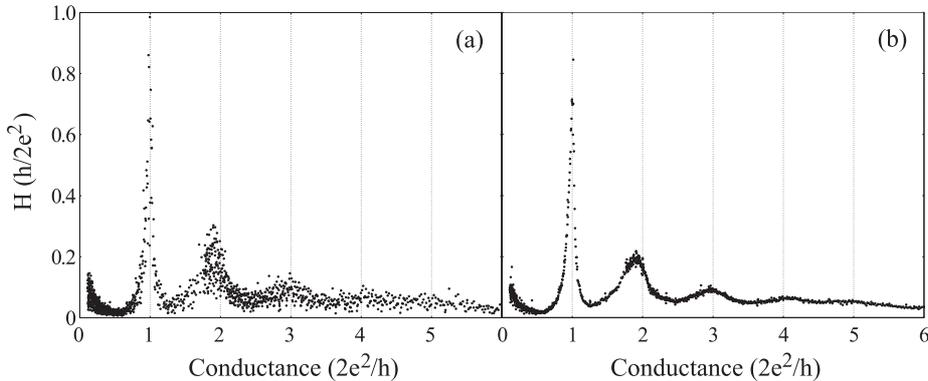


Figure 3.4: Conductance histogram for the commercial AuCo5 relay accumulated from 25 000 breaking point contacts. In panel (a), no compensation has been made for differential non-linearity, and the histogram is very noisy because it is erroneously assumed that the incremental voltage bin widths are constant ( $\Delta V_i = \Delta V$ ). In panel (b), the same data has been compensated for differential non-linearity. The fluctuations are reduced tremendously as compared with panel (a).

non-linearity affects the histograms, we consider what happens on a single input line of a very simple model DSO. For illustrative purposes, we assume that the ADC used for digitizing the voltage signal has only a resolution of 3 bits ( $2^3 = 8$  bins), and it samples the voltage 16 times at a fixed time period  $\tau$  during the acquisition of a single trace.

In Fig. 3.5 we show two examples of the way in which such an ADC could digitize a linear voltage ramp (solid line) covering the full input voltage range. The digitized ADC output is a sequence of 16 bin numbers as indicated by shaded rectangles. If the ADC is ideal (left) with zero differential non-linearity, that is all the voltage bin widths are the same ( $\Delta V_i = \Delta V$ ), then the same number of samples will appear in each bin, as can be seen from the projected histogram. In other words, the histogram indicates that all voltages occurred with equal probability, which is in accordance with the original constant-rate ramp.

If, however, we do the same experiment using an ADC with finite differential non-linearity (right), we find that the number of samples found in different bins changes irregularly. This is a simple consequence of the fact that the incremental voltage bin widths  $\Delta V_i$  are not equally large. So, if ideal differential linearity is

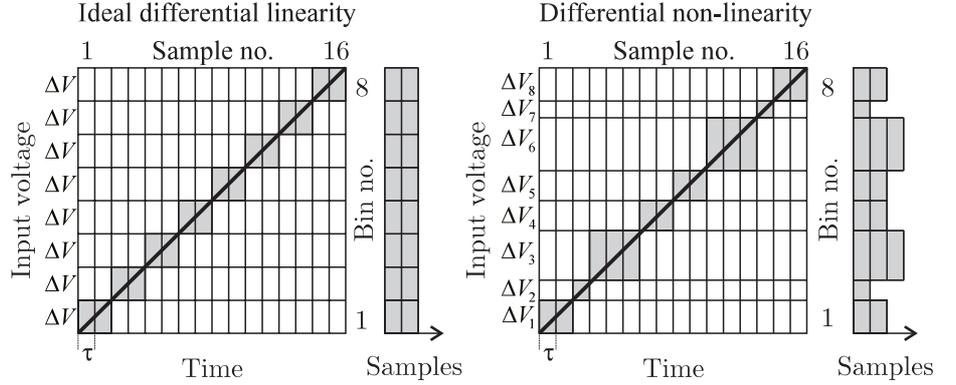


Figure 3.5: Schematic representation of the conversion of a linear voltage ramp (solid line) in an ideal (left) and a realistic (right) 3-bit ADC. It is assumed that the voltage is sampled 16 times with a fixed time interval  $\tau$  during a single ramp. The number of samples in a single bin  $i$  is proportional to the incremental voltage bin widths  $\Delta V_i$ , as seen in the projected histograms showing the number of samples versus bin number. This information can be used for eliminating the noise found in conductance histograms [cf. Fig. 3.4(b)].

erroneously assumed ( $\Delta V_i = \Delta V$ ), the (apparent) statistical weight of different bins will not be the same, as can be seen in the rightmost histogram in Fig. 3.5 at variance with the original constant-rate ramp. This means that the weight of wide bins like  $i = 6$  in the right panel of Fig. 3.5 will be overestimated in histograms similar to the one depicted in Fig. 3.4(a) unless the true  $\Delta V_i$  values are measured and used in a compensation scheme. The effects of differential non-linearity have not been discussed previously in connection with these kinds of experiments, but differential non-linearity is a well known problem in for instance nuclear spectroscopy [25, 26].

### Measuring differential non-linearity

If the correct  $\Delta V_i$  values can be found, then the correct  $\Delta G_i$  values, which are used in the calculation of individual histogram points [cf. Eq. (3.2)], can be calculated straightforwardly. This correction will reduce the noise in the histograms considerably, as e.g. over-represented histogram values  $H(G_i)$  will be dampened accordingly with a larger  $\Delta G_i$  value and *vice versa*.

Fortunately, the correct  $\Delta V_i$  values can be measured directly using the voltage ramp technique illustrated previously in Fig. 3.5. A voltage ramp is applied repeatedly, and the total number of samples recorded in a single bin  $N_i$  is monitored. For large  $N$ , the number of samples in each bin  $N_i$  divided by the average number  $N/n$  will converge towards the normalized voltage bin width  $\Delta V_i/\Delta V$ . The normalized widths have to be measured independently for each input line on the DSO as no two ADC's are the same. Just like a finger print.

In Fig. 3.6 we show measurements of the normalized voltage widths of bin numbers 512 to 640 on one of the input lines of the DSO. This interval has been picked out at random — only 128 of the 1024 bins are shown for clarity. Very large fluctuations are present as the widths range from 0.3 to 1.8. That is, some bins are six times wider than others!

The measured standard deviation is as large as 34%, which in fact roughly equals the differential non-linearity of the simple model ADC shown in the right panel of Fig. 3.5. The measured differential non-linearity of the DSO is much larger than the 1% or better encountered in ADC's used in nuclear spectroscopy [25, 26]. However, typical ADC's used in nuclear spectroscopy

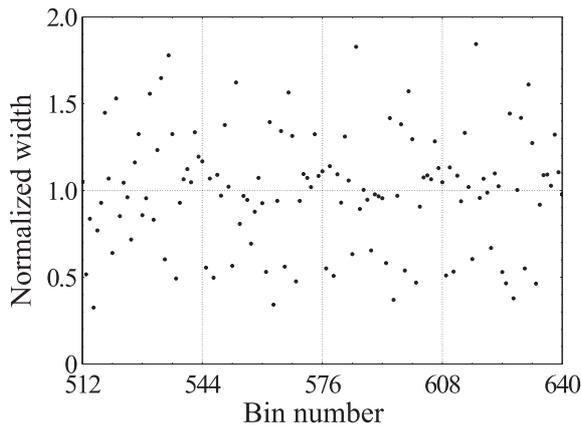


Figure 3.6: Differential non-linearity for a sequence of 128 bins on an input line on a Yokogawa DL 4100 DSO. The normalized width  $\Delta V_i/\Delta V$  of single bins is shown as a discrete function of channel number. The standard deviation is 34%. The widths are to some extent periodic with a period of 32 bins.

normally need  $\sim 10 \mu\text{s}$  to sample one voltage value, whereas for the Yokogawa DSO used in our measurements the conversion time is only 10 ns. We may stress that this is a general problem for most DSOs. For instance, we have measured 16% differential non-linearity on an 8-bit LeCroy 9361 DSO. In general, the non-constancy of incremental voltage bin widths increases with increasing sample rate and bit resolution. This is why the effect is so large in the Yokogawa DSO (100 Msamples/s, 10 bit). However, this does not pose a serious problem as it is measurable, and we can correct for it.

The normalized widths  $\Delta V_i/\Delta V$  are nearly unaffected by the DSO settings. We have checked this by measuring the normalized widths using different V/div, time/div and samples/trace settings. Dividing the two discrete functions of normalized widths obtained under the different conditions yields values centered around 1 with a standard deviation of only 2%. This gives an estimate of how much of the noise can be removed by doing proper compensation for differential non-linearity, i.e. it should be reduced to approximately 2%.

### Compensating for differential non-linearity

As the effects of differential non-linearity is very large for our 10-bit DSO, we have calculated the  $V_i$  values from an algebraic sum of the experimentally found  $\Delta V_i$  values and corrected the corresponding  $\Delta G_i$  values in the following. In Fig. 3.4(b) we show the histogram for the commercial AuCo5 relay using the proper  $\Delta G_i$  values. The fluctuations are now almost gone. The noise remaining in the histogram is mainly due to systematic errors in the measured  $\Delta V_i/\Delta V$  values. We emphasize that these are raw data, and no smoothing has been applied. A discussion of the exact positions of the peaks in the histogram will be deferred to Sec. 3.3.

An alternative and very tempting approach to solving the problem with noise in the histograms is to use enhanced resolution (ER) techniques, as they are readily available on many DSO's. The idea is to increase the true bit resolution of the DSO by folding a trace with a sliding data window, thus giving the average voltage. This tends to decrease the effect of differential non-linearity because contributions from bins with different widths are added together. If the averaging window is large enough, the effects of differential non-linearity will no longer be present. As a consequence of the smearing, real structure can be lost in the histogram, and much information about the detailed shape of the peaks will be lost [27]. This may lead to unreliable determinations of the peak positions.

We may add that the fluctuations in voltage bin width as a function of

bin number are *not* random. Some correlation between the width and the bin number exists. This can be seen by inspecting Fig. 3.6 carefully, in which a sequence of widths has a tendency to repeat itself after a period of 32 bins. This is related to the way in which the ADC converts a voltage into a bin number in binary representation. The periodicities are strongly related to similarities in the bit patterns of different bin numbers and may differ significantly from one ADC to the other. These periodicities may remain after ER is employed and emerge as artificial peaks in a histogram! Because of these complications we have desisted from using ER.

### Calibration histograms

On page 46 we described how the measured voltages  $V_i$  were converted to the corresponding conductance values  $G_i$  using Eq. (3.1). The three parameters in this equation were determined accurately by inserting calibrated conductances in the circuit and fitting the measured voltages to Eq. (3.1). We can check how well the fitting procedure works in a very direct way by making a conductance histogram from measurements on calibrated conductances. We have selected conductances with integer values of  $G_0$  (calibrated within 0.1 %). Calibration conductance histograms with sharp peaks centered at integer values of  $G_0$  is indeed obtained, as can be seen in Fig. 3.7. The two input lines have different V/div settings to cover the interesting ranges of the conductance most efficiently.

The widths of the peaks give a direct measure of the instrumental broadening in the histogram caused by electronic and digitizing noise. The peaks have widths that are negligible as compared to the widths of the peaks encountered in the real conductance histograms [cf. Fig. 3.4(b)], and the offsets of the peak positions from integer values of  $G_0$  are also negligible compared to the accuracy with which the peak positions (or other features) in the conductance histograms can be determined. In the following we can therefore neglect contributions to uncertainties originating from the calibration procedure.

## 3.3 Results

### The commercial AuCo5 relay

In Fig. 3.8, we have plotted the conductance histogram for the commercial AuCo5 relay on a logarithmic scale to enhance the higher-order peaks [27], which are barely visible when plotted on a linear scale as in Fig. 3.4(b). We

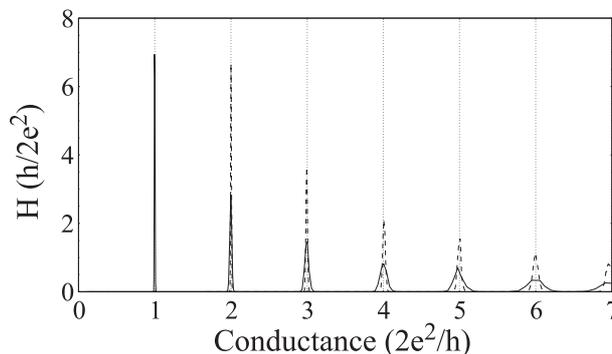


Figure 3.7: Calibration conductance histograms from two input lines on the DSO generated from measurements of inserted calibrated conductances. The solid line is the calibration histogram from an input line used for measuring conductances in the range  $0.1$ – $1.7 G_0$ , whereas the dashed line is for another input line used for measuring conductances larger than  $1.7 G_0$ .

have also eliminated some of the noise in the histogram by a removal of the highest Fourier components in the sequence of compensated histogram data points, effectively corresponding to a smoothing over 12 bins [28]. Five peaks with decreasing weight/increasing width are seen in the histogram. The peak maxima are positioned slightly below the integer values, and the shifts seem to be roughly proportional to the peak number. Interestingly, we note that the steepest descent of the peaks is positioned almost exactly at the integer values of  $G_0$ .

It has become a widespread practice to subtract a constant residual resistance of a few hundred Ohms from the measured values to bring the peaks close to the integer values [12, 19, 29]. In mechanically controllable break junction (MCBJ) experiments, it has also been noted that a limited number of traces can be mapped exactly onto integer values by subtraction of a constant resistance [24]. It has been argued that the residual resistance may compensate for, e.g., the resistance in the leads connecting the ballistic contact to the macroscopic world [12, 29] and the average effects of backscattering in the QPC [19]. However, the detailed physical justification for this empirical procedure is rather weak. It is not at all clear why a constant residual resistance should be able to account for the average effects of backscattering. Furthermore, it is difficult to justify the actual value of several hundred Ohms.

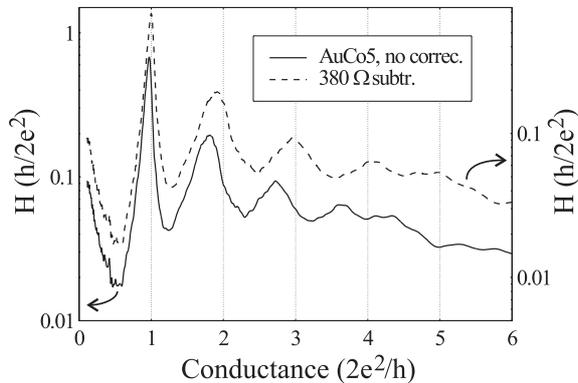


Figure 3.8: Continuous conductance histograms for the commercial AuCo5 relay without (solid line, left vertical scale) or with (dashed line, right vertical scale) a subtraction of a serial resistance of  $380 \Omega$ . The histograms have been smoothed over 12 bins and are plotted on a logarithmic scale.

Nevertheless, this empirical procedure seems to work fine in our experiment as well. In Fig. 3.8, we tentatively show the original histogram after a subtraction of a constant residual resistance of  $380 \Omega$  along with the uncorrected histogram. Now all the peaks are positioned at integer values, although statistical significant deviations are still present. The subtracted residual resistance is of the same order as previously reported values for histograms based on measurements of the conductance in Au nanowires at room temperature. Examples are the UHV-STM study by Brandbyge *et al.* ( $150 \pm 50 \Omega$ ) [12] and a macroscopic wire experiment by García *et al.* ( $490 \Omega$ ) [19]. Further discussion of the residual resistance will be given later on page 67.

Other types of commercial relays have also been studied with e.g. W or Rh as contact material, but no structures were revealed in the corresponding histograms. This may indicate that the choice of contact material is crucial. To investigate this in further detail, we have performed analogous studies with the home-built relay using only pure metal wires.

### Au and Co wires used in the home-built relay

The presence of Co in the commercial AuCo5 relay may influence the conductance of the nanowires formed, which complicates a direct comparison with

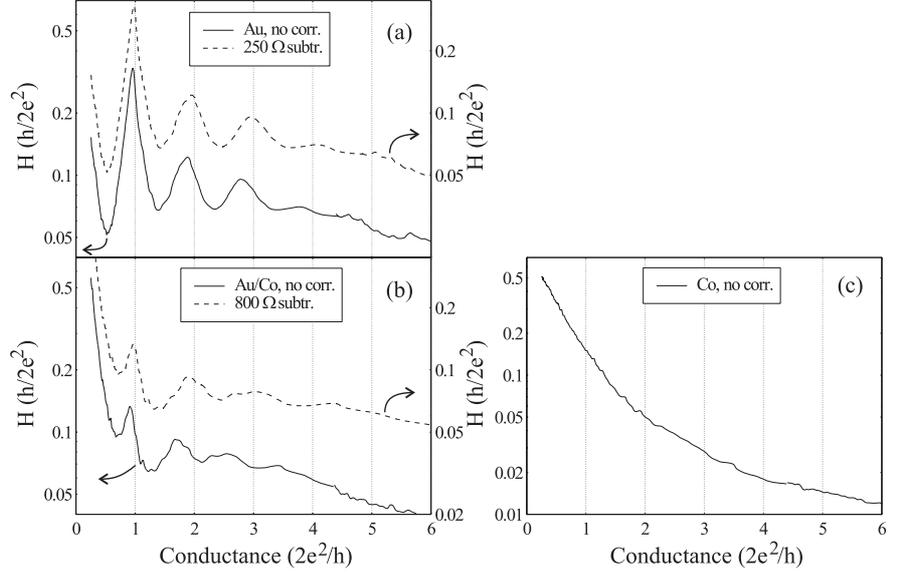


Figure 3.9: Conductance histograms based on measurements with the home-built relay using different combinations of Au and Co wires as contact materials. Each histogram is accumulated from 25 000 breaking point contacts, compensated for differential non-linearity, smoothed over 12 bins and plotted on a logarithmic scale. Histograms with subtracted residual resistances (dashed curves, right vertical scale) of 250  $\Omega$  for (a) Au/Au and 800  $\Omega$  for (b) Au/Co are plotted along with the original histograms (solid curves, left vertical scale). In the Co/Co histogram (c), no resistance has been subtracted from the data.

previous studies on Au. To elucidate this, we have studied what happens when the following combinations of pure metal wires are used in the home-built relay: (i) two Au wires, (ii) one Au and one Co wire, and (iii) two Co wires. The conductance histograms for the three combinations are shown in Fig. 3.9.

The resulting conductance histogram for the home-built relay mounted with two Au wires [Fig. 3.9(a)] has again up to five peaks displaced slightly below the integer values of  $G_0$ . As in the AuCo5 case (Fig. 3.8), the integer values of  $G_0$  coincide very closely with the points of steepest descent of the peaks. It is seen that a correction for a residual resistance of 250  $\Omega$  is able to shift the peak maxima very close to the integer values of  $G_0$ .

When one of the wires is interchanged with a Co wire [Fig. 3.9(b)], the peaks become less well defined, and the smoothly varying background is increased dramatically for low conductance values ( $<1 G_0$ ), which may be indicative of impurities in the nanowires. Finally, when two Co wires are used [Fig. 3.9(c)], only a smoothly varying background is observed with an increasing amplitude towards low conductance values. On the basis of the histograms in Fig. 3.9, we conclude that the presence of Co under ambient conditions tends to destroy the quantization of the conductance.

### Ag and Cu as contact material

Since Au is very free-electron-like, it is of interest to perform similar experiments on other metals with an almost spherical Fermi surface like, e.g., Ag and Cu. The resulting conductance histograms are shown in Fig. 3.10, and it is seen that the quantization is not as distinct as for the Au case. Again, the integer values of  $G_0$  are close to the points of steepest descent, and a correction for a constant residual resistance is able to shift the peaks close to integer values of  $G_0$ . Clear peaks are observed at  $1 G_0$ , but the higher-order peaks are less well defined, and no structure is seen above  $3 G_0$ .

### Transition metals

The transition metal Co exhibits no sign of quantization of the conductance under ambient conditions [cf. Fig. 3.9(c)]. To investigate whether this is a general trend for transition metals, we have also performed similar experiments using Ti, Fe, Ni, and Pt as contact materials. In contrast to the free-electron-like metals, no peaks are observed in any of the corresponding histograms. They all qualitatively resemble the Co results.

## 3.4 Discussion

### Commercial vs. home-built relays

When one or two Co wires are used in the home-built relay [see Fig. 3.9(b) and (c)], conductance quantization is either very weak or completely absent, and, moreover, the smoothly varying background has a very large amplitude for low conductance values. This is in contrast to the corresponding histogram obtained from the commercial AuCo5 relay (Fig. 3.8), which has clear indications of quantized conductance as well as a low background for small values of

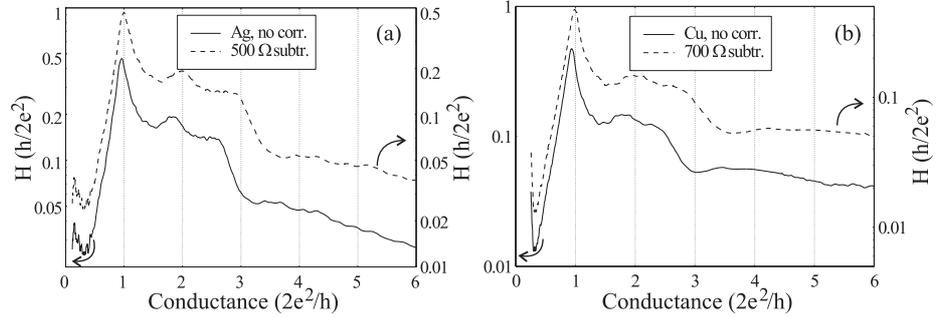


Figure 3.10: Conductance histograms accumulated from measurements with the home-built relay using Ag and Cu wires as contact materials. Each histogram has been compensated for differential non-linearity, smoothed over 12 bins and plotted on a logarithmic scale. Histograms corrected for serial resistances of 500  $\Omega$  for Ag (a) and 700  $\Omega$  for Cu (b) are shown as dashed curves on the right vertical scale along with the original histograms (solid curves, left vertical scale). The Ag histogram is accumulated from 14 000 breaking point contacts, whereas the Cu histogram is based on 25 000 breaking point contacts.

the conductance. In fact, the peaks are even sharper than those obtained using only pure Au wires in the home-built relay (Fig. 3.9). From these observations, we are led to conclude that the presence of Co in the commercial relay plays no significant role in these experiments and that nanowires formed in the AuCo5 relay therefore mainly consist of Au.

At first it may seem surprising that the quantization is so prominent in the commercial relay. However, it has other properties that improve the possibilities of forming nanowires which exhibit quantized conductance: In the commercial AuCo5 relay, the electrodes are enclosed in a contact chamber, which before assembly is evacuated (to approximately  $10^{-8}$  bar; see Ref. 13) and heated to 120  $^{\circ}$ C. This procedure cleans the contact surfaces. Furthermore, the contact chamber contains a permanent magnet (BaOFe), which works like a getter after the chamber is sealed [14]. As a consequence of the heat and vacuum treatment, pores with diameters  $>100$  nm are formed on the surface of the getter, which makes it very well suited for binding harmful substances which occur in the contact chamber over a long period of time (aromatic polymers) [13]. Otherwise, the organic molecules will polymerize and form highly resistive contamination layers on the contact areas [13]. In the home-built relay, the atmosphere is not

as well controlled, and it is plausible that a larger concentration of impurities is present, thus impeding the formation of pure wires.

Additionally, the design of the commercial AuCo5 relay is such that the contacts are self cleaning [13]. This is accomplished by sliding the electrodes against each other during the closing action, thus partly rubbing any contaminant layers away from the contact region. This is not expected to happen in the home-built relay as the contacts are brought very gently into contact, and the externally applied contact force is much smaller than the 0.1 N applied in the commercial relay.

### Au and other free electron metals

As the experiments with the home-built relay are carried out under ambient conditions, the reactivity of the metals will have an important impact on the presence of impurities in the nanowires. These impurities introduce electronic scattering centers, which could destroy the conductance quantization [4, 19]. Since Au is the noblest of all metals [30], the effect of impurities is expected to be at a minimum for such nanowires, whereas wires of Ag and Cu will be more affected due to e.g. oxidation. This may be the reason why conductance quantization is more prominent in Au nanowires than in wires of Ag and Cu.

Quantization of the conductance in Au has been observed in many previous studies. Those that incorporated conductance histograms are reviewed here.

It was shown in a UHV-STM study by Brandbyge *et al.* carried out at room temperature (RT) (Ref. [12]) that 3–4 peaks could be resolved in the histogram after subtraction of a constant residual resistance of  $150 \pm 50 \Omega$  based on 227 traces selected from a total of 429 traces. Using the same experimental setup, it has been shown by Olesen that it is not necessary to employ the selection criteria in the case of Au [29]. In this study, four peaks were found in the histogram positioned at integer values of  $G_0$  after a subtraction of  $230 \pm 50 \Omega$ . At lower (170 K) and higher (390 K) temperatures, qualitatively similar histograms were obtained, each containing three clear peaks [29]. Each of these experiments was based on more than 350 conductance traces.

Gai *et al.* also carried out RT UHV-STM studies and found four peaks at integer values of  $G_0$  based on 964 out of 2500 traces [22]. Sirvent *et al.* have studied Au nanowires using an STM in a He atmosphere at high vacuum at three different temperatures (4.2 K, 77 K and 293 K) [21]. The three histograms all had three peaks slightly offset below the integer values of  $G_0$ , and the varying temperature seemed to have little effect on the results. The histograms were based on selected curves from thousands of indentations.

Gold has also been studied using the MCBJ technique in high vacuum, both at 4.2 K [31] and at RT [24]. On the basis of 65 conductance traces, Krans found only the first peak in the histogram at 4.2 K from which he concluded that conductance quantization is more or less absent at this temperature [31]. However, if one considers the histogram by Krans in detail, peaks at 1 and 2  $G_0$  may be present. At RT, MCBJ results for Au are more conclusive as shown by Muller *et al.* [24]. A histogram accumulated from 72 traces showed 5 peaks close to integer values of  $G_0$ .

Turning to the macroscopic wire experiments, a histogram based on 103 selected [27] curves obtained under ambient conditions has been published by Costa-Krämer *et al.* [11]. At least five peaks were found close to integer values. A tremendous leap concerning the amount of statistics was subsequently published by García *et al.* [19], where a histogram built from 12000 conductance traces was shown with peaks very similar to those presented here with the home-built relay [Fig. 3.9(a)]. In that study, a constant residual resistance of 490  $\Omega$  was subtracted and all data were included.

All the previous results as well as those presented here clearly demonstrate that quantization of the conductance in Au nanowires is very robust and independent of temperature, atmosphere and chosen experimental technique. We therefore conclude that Au is a bench-mark system for demonstrating quantized conductance, independent of the conditions.

Our experiments with Ag and Cu as contact materials (Fig. 3.10) are in accordance with analogous macroscopic wire experiments by Costa-Krämer *et al.* [27], in which all traces are included. Cu has also been studied previously in MCBJ experiments [23, 24]. At 4.2 K, a conductance histogram based on 90 traces had clear peaks close to 1 and 3  $G_0$ , but also minor persistent peaks at non-integer values [23]. The analogous RT histogram based on 78 traces had, on the other hand, 4–5 peaks at integer values of  $G_0$  [24].

### Skewness of histogram peaks

In previous studies, the positions of individual peak maxima in the conductance histograms have been used as the relevant quantities. On the basis of their offsets slightly below integer values of  $G_0$ , the discrepancies have often been explained by the presence of a constant residual resistance of a few hundred  $\Omega$  [12, 19]. This interpretation is in qualitative agreement with our experiments [cf. Figs. 3.8, 3.9(a) and (b), and 3.10]. However, a close inspection of the raw data, like those presented for the commercial AuCo5 relay in Fig. 3.4(b), shows that the resistances needed for shifting the different peaks to integer values are

not completely identical. For instance, in the particular case of the AuCo5 relay, they are 500, 750, 500, 330 and 400  $\Omega$ , respectively for the five peaks, all with an uncertainty of approximately 100  $\Omega$ . This indicates that we still lack a detailed understanding of how to interpret the conductance histograms.

In fact, it is not entirely clear how relevant quantities can be extracted from conductance histograms. First of all, it is not evident that the peak maxima are the most important quantities, even though they indicate conductances that occurred with maximum probability during the experiment. Other characteristics of the histogram may be of equal importance. For instance, we note that peaks in the Au histograms are skew [Figs. 3.8 and 3.9(a)]. That is, they have asymmetric tails extending out towards lower conductance values. This is reflected in different values of the half width at half maximum (HWHM) of the low and high conductance tails.

The skewness is particularly prominent for the first peak, which is shown in the expanded view in Fig. 3.11 for the commercial AuCo5 relay. Here HWHM values of 0.065 and 0.035  $G_0$  are found for the low/high part of the peak. For comparison, values of 0.09 and 0.06  $G_0$  are found for the home-built relay when Au wires are used as contact material [cf. Fig. 3.9(a)].<sup>3</sup> This high end cutoff is probably related to the fact that transmission probabilities for single channels can never exceed one.

The shape of histogram peaks has been discussed by Torres and Sáenz [20], who pointed out that the peak width depends critically on the opening angle in the contact and that the shape results from an interplay between scattering in the contact region and contact shape evolution. As an example, their calculations show that the peak at 3  $G_0$  for Na has contributions from contacts with three, four, and five atoms in the smallest cross-section; the different kinds of contacts contribute differently to the statistics in the peak.

In the last stages before a nanowire is disrupted, it can be assumed that one channel is being closed at a time, even though this depends on the eccentricity of the cross-section [4, 12]. However, this rule of thumb only seems valid for the last 3–4 transmitting eigenchannels, as suggested by the free-electron calculations of transmission probabilities through Au nanowires with geometries obtained from molecular dynamics simulations by Brandbyge *et al.* [2, 3].<sup>4</sup>

When this picture is valid, the conductance is restrained from surpassing the integer values of  $G_0$ .<sup>5</sup> Thus, for the first few peaks, the cutoff values, or the

---

<sup>3</sup>In paper [II], all the quoted HWHM values were unfortunately a factor of ten too large!

<sup>4</sup>See the review of selected results on pages 31–34 in Chapter 2.

<sup>5</sup>This can be seen from the eigenchannel version of the Landauer-Büttiker formula [Eq. (2.30) on page 27 in Chapter 2].

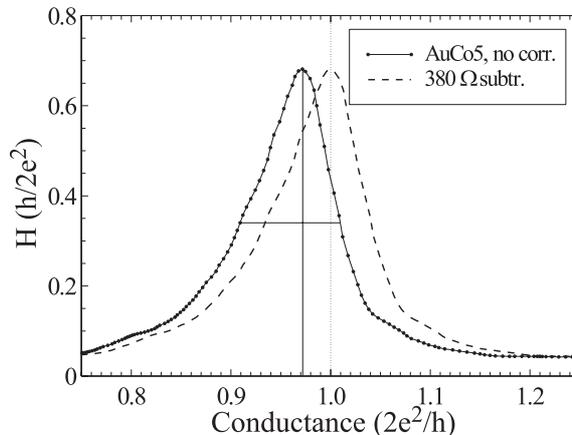


Figure 3.11: Expanded view of the first peak of the AuCo5 conductance histogram. The solid curve is the original data from Fig. 3.4(b) smoothed over 12 points, whereas the dashed curve shows the peak after a subtraction of a serial resistance of  $380 \Omega$ . The peak is very skew with  $\text{HWHM} \approx 0.65$  and  $0.35 G_0$  for the lower and upper tails. The original peak has a cutoff value, which is very close to  $1 G_0$ .

positions of steepest descent, are expected to be close to the integer values of  $G_0$ . Displacements of the cutoff values below integer values can then be interpreted as due to the resistance originating from diffuse (Ohmic) electron transport in the leads.

In all of the conductance histograms presented in this Chapter, we observe that the points of steepest descent are very close to the integer values of  $G_0$  for the first three peaks [cf. the solid curves in Figs. 3.8, 3.9(b) and (c), 3.10, and 3.11]. We see no statistical significant displacements below the integer values. We therefore conclude that the resistance in the leads is small and probably less than  $100 \Omega$ .

The fact that the peaks are asymmetric with a short tail on the high conductance side provides yet another indication that the Au histograms are indicative of conductance quantization and not an area effect caused by discrete changes in contact area. If each peak should correspond to a particular stable configuration of the atoms in the smallest cross-sectional area [32], one would expect the peaks to be symmetric at variance with e.g. the peak shown in Fig. 3.11.

### 3.5 The resistance of the leads

We will now discuss a simple model for the resistance of the leads connecting the ballistic contact with the macroscopic world. In our setup, we measure the total conductance between the calibration points in our circuit. This part of the circuit only consists of macroscopic wires and the contact itself. The resistance of the macroscopic wires can be neglected completely, but this may not be the case for the final part of the wires which connects to the QPC.

We model this by dividing the contact into three parts: two diffuse leads and a ballistic constriction<sup>6</sup> as shown in Fig. 3.13. For simplicity, the leads are assumed to be identical, semi-infinite cones, truncated at a diameter of  $d$ . The truncated cones are characterized by an opening angle  $\alpha$  and a uniform resistivity  $\rho$ .

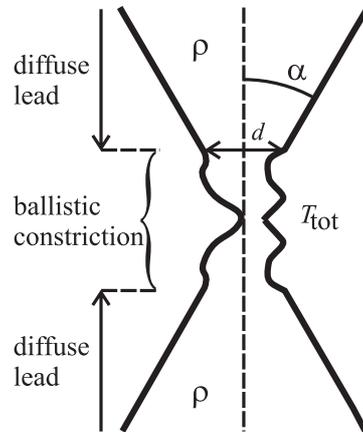


Figure 3.12: Model used for estimating the resistance of the leads. The diffuse leads have a uniform resistivity  $\rho$  and are modeled as semi-infinite cones, truncated at a diameter of  $d$ . The leads are characterized by an opening angle  $\alpha$  and are connected by a ballistic constriction with total transmission  $T_{\text{tot}}$ .

<sup>6</sup>In principle, we cannot simply divide the system into strictly diffuse and ballistic parts. The transport will be quasi-ballistic in the two regions between the truncated cones and the constriction. However, as discussed on page 19 in Chapter 2, such a crude division works surprisingly well.

Let us consider the diffuse resistance of the two truncated cones  $R_{\text{TC}}(\alpha, d)$ . As discussed on page 10, the evaluation of this resistance requires that we specify the boundary conditions for the electrostatic potential at the connections at the ends of the truncated cones. Far away from the constriction, the equipotential surfaces will be spherical, but at the connection to the constriction, we do not know the detailed shape of the equipotential surface due to the unknown geometry of the constriction. Thus, we are in principle unable to specify the exact boundary conditions. Instead, we will *assume* that the equipotential surfaces are also spherical close to the constriction and calculate the resistance. Afterwards, we estimate the error we may introduce using this assumption. It is now easily found that

$$R_{\text{TC}}(\alpha, d) = f(\alpha)R_{\perp}(d), \quad (3.3)$$

where

$$\boxed{f(\alpha) = \frac{\sin \alpha}{1 - \cos \alpha}} \quad (3.4)$$

is a geometric factor describing the dependence of the opening angle  $\alpha$ . A plot of  $f(\alpha)$  is shown in Fig. 3.13. It diverges as  $\alpha$  goes to zero (an infinite wire) and approaches 1 as  $\alpha$  goes to  $90^\circ$ .

The other factor  $R_{\perp}(d)$ , which occurs in Eq. (3.3), is the resistance of the leads at an opening angle of  $90^\circ$ . It is given by

$$R_{\perp}(d) = \frac{2\rho}{\pi d}. \quad (3.5)$$

This resistance can be compared with the resistance  $R_M = \rho/d$  of the Maxwellian orifice [33], which was discussed on page 12 in Chapter 2 and calculated in Appendix A. The formula for  $R_{\perp}$  in Eq. (3.5) is smaller than  $R_M$  with a factor of 0.64. This is because we have omitted the resistance of a sphere of diameter  $d$  and assumed that equipotential surfaces are spherical. As demonstrated in Appendix A and shown in Fig. 2.2 on page 13, they are, in fact, oblate spheroidal for the Maxwellian orifice.

It is interesting to note that in the ballistic Sharvin limit ( $l \gg d \gg \lambda_F$ ) the resistance  $R_S$  of the same geometry (an orifice of diameter  $d$ ) is given by [34]<sup>7</sup>

$$R_S(d) = \frac{16\rho l}{3\pi d^2}. \quad (3.6)$$

---

<sup>7</sup>This follows by combining Eq. (2.16) on page 17 in Chapter 2 with Eq. (3.8).

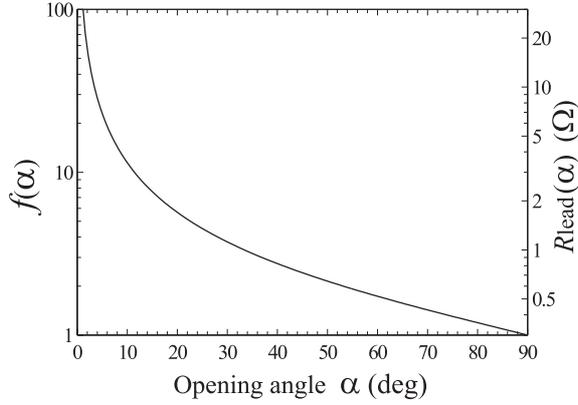


Figure 3.13: The geometric factor  $f$  (left axis) and the resistance of the leads  $R_{\text{lead}}$  (right axis) as a function of opening angle  $\alpha$ . The resistance of the leads is estimated from Eq. (3.9) using  $\rho = 2 \mu\Omega \text{ cm}$  and  $n = 6 \times 10^{22} \text{ cm}^{-3}$ . The vertical scale is logarithmic.

In the quasi-ballistic regime ( $d \sim l$ ), it has been shown that there is a smooth transition from the diffuse ( $R_M$ ) to the ballistic ( $R_S$ ) regime, and the interpolation formula Eq. (2.18) on page 19 has been derived [35], but here it is sufficient to note that  $R_{\perp}(l) = \frac{3}{8}R_S(l)$ . Thus, Eq. (3.5) is expected to be valid within, say, a factor of 3, even if  $d = l$ . We can therefore proceed and estimate the resistance of the leads  $R_{\text{lead}}$  with the following formula

$$R_{\text{lead}}(\alpha) \approx R_{\text{TC}}(\alpha, l). \quad (3.7)$$

Thereby, we define the leads to be the parts of the contact that are wider than  $l$ . In a free-electron model, the mean free path  $l$ , which appears in Eq. (3.7), can be written in terms of  $G_0$ ,  $\rho$  and the free electron density  $n$  as [36]

$$l = \sqrt[3]{\frac{3}{\pi G_0 \rho n^2/3}}. \quad (3.8)$$

Inserting this expression in Eq. (3.5), combined with Eqs. (3.3) and (3.7), gives

$$\boxed{R_{\text{lead}}(\alpha) \approx f(\alpha) \frac{2}{\sqrt[3]{3\pi^2}} G_0 \rho^2 n^{2/3}}, \quad (3.9)$$

which is our main result.

Typical bulk resistivities for metals like Au, Ag, and Cu at RT are  $\rho \approx 2 \mu\Omega \text{ cm}$ . If this value is inserted in Eq. (3.9) along with a typical free electron density  $n = 6 \times 10^{22} \text{ cm}^{-3}$ , we obtain  $R_{\text{lead}}(\alpha) \approx 0.3 f(\alpha) \Omega$ . This estimate of  $R_{\text{lead}}(\alpha)$  is shown in Fig. 3.13.

For realistic opening angles ( $4^\circ$ – $60^\circ$ ), the resistance of the leads is expected to be  $0.5$ – $10 \Omega$ , which is small compared to the residual resistances of  $250$ – $800 \Omega$  that are needed to shift the peaks in the conductance histograms up to integer values of  $G_0$ . The estimate is, however, fully consistent with our previous argument: That the cutoff values of the peaks are the relevant quantities for determining the resistance of the leads. As the cutoffs were found to be placed almost at the integer values of  $G_0$ , we concluded on page 66 that the resistance of the leads most probably was below  $100 \Omega$  in perfect agreement with the estimate.

In the estimate presented above, we used a typical bulk resistivity for a pure metal, which may be somewhat lower than the actual resistivity of the metals used in these experiments near the contact region. During the repeated switching, impurities and films of highly resistive material may be embedded in the surface, thus increasing the resistivity substantially [13]. Since  $R_{\text{lead}} \propto \rho^2$  [Eq. (3.9)], this may lead to a significant increase of the estimate of the lead resistance, but this issue is not completely resolved.

By separating the resistance of the leads from the rest of the system, we have implicitly assumed that the resistance of the contact can be calculated by adding the resistance of each part in series. However, additivity of resistances is based on the assumption of phase-incoherence [37]. That is, an incoming electron wave in the constriction should be completely phase-randomized in inelastic scattering processes in the constriction. This assumption may be questionable when the typical dimension of the constriction becomes of the order of  $l$ .

The shapes of the peaks as well as the positions of the peak maxima are governed by the evolution of the transmission through the constriction as the contact is stretched. This and previous studies have shown that the concept of a constant residual resistance can explain offsets of peak maxima below the integer values of  $G_0$ . However, no realistic calculations have yet been able to explain why this is the case.

The cleanliness of the system seems to be correlated with the residual resistance. In experiments with Au nanowires under ambient conditions, the residual resistance is quite low [ $250 \Omega$ , Fig. 3.9(a)], whereas residual resistances of  $500$  and  $700 \Omega$  are needed in the case of Ag and Cu [Fig. 3.10]. The increase of the resistances from Au to Ag and Cu may be a simple consequence of the larger

chemical reactivity of Ag and Cu as compared to Au.

In general, the metal reactivity scales proportionally to the filling of the  $d$ -band, which makes transition metals very susceptible to impurities. This may also be one of the reasons why we have not been able to observe quantized conductance under ambient conditions in the transition metal contacts. Also, transition metals have unfilled  $d$ -bands and non-spherical Fermi surfaces, which may impede the quantization [38, 39].

It is normally assumed that the constrictions are ballistic. That is, the mean free path of the electrons is larger than the size of the nanowires formed. Pascual *et al.* have shown that Au nanowires with a length up to 400 Å can be formed with an STM under ambient conditions [40], which is approximately equal to the bulk value of the mean free path for Au, Ag, and Cu at RT [36]. As mentioned previously,  $l$  may be even smaller close to the contact region due to the presence of impurities. For the transition metals Ni, Pt, and Fe, which were shown not to exhibit quantized conductance under ambient conditions,  $l$  is even smaller than 100 Å. At the present state, it is not clear how this affects the conductance through the constriction.

### 3.6 Conclusion

We have shown that it is relatively easy to observe quantized conductance under ambient conditions in certain metals. We have proven this by using various contact materials in commercial and home-built relays. Au, Ag, and Cu exhibit quantization in contrast to the transition metals Ti, Fe, Co, Ni, W, Rh, and Pt. This is explained in terms of the higher reactivity of the transition metals. It may, however, also be correlated with their rather complicated electronic structures, which are far from free-electron-like. Furthermore, the mechanical properties of the transition metals may impede the formation of adiabatic nanowires. We have shown that Au is a bench-mark system for observing quantized conductance as the quantization is very robust, i.e., almost independent of environment and experimental technique.

Peaks in the Au conductance histograms are shown to be asymmetric. The rapid peak cutoffs close to integer values of  $G_0$  are interpreted as a further indication of conductance quantization. Deviations from perfect quantization are explained in terms of the resistance of the leads and the average effects of backscattering in the constriction. We argue that the high end cutoff values for the histogram peaks are relevant for determining the resistance of the leads. A simple model of the contact suggests that this resistance is small and less than

100  $\Omega$ , in full agreement with the experimental conductance histograms. Finally, we have shown that it is of important to compensate the observed conductance histograms for differential non-linearity of the DSO.

### 3.7 Recent developments

As promised in the beginning of this Chapter, we will now discuss how the results and views presented in paper [II] relate to later work.

#### Other relay studies

Independent of our work, a paper by Yasuda and Sakai appeared in an issue prior to paper [II] in *Phys. Rev. B* [41]. Here, the authors also presented conductance histograms accumulated from breaking gold contacts in commercial electromechanical relays using a setup, which was very similar to ours. Their primary interest was to investigate how the shape of the histogram changes under high applied bias voltages — a highly interesting idea, which is intimately related to the following Chapters. They also presented histograms acquired under low bias voltages, and these looked very similar to those presented here (Figs. 3.3 and 3.4). Since then, further relay studies from Sakai and co-workers have appeared, focusing on Au [42–44] and Pt [45] relay contacts under high applied bias voltages.

The observation of QC in relays was re-discovered by Ott and Lunney in 1998 [46], who, at first [47], had overlooked the existence of paper [II] and Ref. 41. They pointed out that such relay experiments are well suited for under-graduate laboratories.<sup>8</sup> A simpler one-to-one relationship between the output voltage and the conductance was also accomplished by using a simple transimpedance setup (to be discussed in Chapter 4). In another report, Ott *et al.* mounted samples of either a ferromagnetic oxide or Fe on commercial relay contacts. Surprisingly, the conductance histogram for the ferromagnetic oxide had clear peaks at integer values of up to 7  $G_0$ !

#### Compensating for differential nonlinearity

The scheme presented in paper [II] for how to correct conductance histograms for differential non-linearity has been used in several later reports by the groups

---

<sup>8</sup>This idea has later been pursued by Foley *et al.* using macroscopic wires [48].

of García [49, 50], Sakai [43–45], and Tao [51].<sup>9</sup> The effects of differential non-linearity have also been discussed by de Heer’s group [53].

### The noble metals

In paper [II] it was emphasized that Au seems to be a bench-mark system for observing QC in metal QPCs independent of temperature and experimental conditions. This has been confirmed in a number of recent studies [8, 10, 42–46, 48–50, 54–57] although there have been a few exceptions [53, 58, 59]: (i) Fractional quantum conductance has been reported by de Heer [53], but has not been confirmed by others; (ii) Scheer *et al.* found a conductance well below  $1 G_0$  for the smallest Au contacts [58] in their experiments. A result, which has been explained as due to disorder (see footnote 16 on page 28 in Chapter 2); (iii) using a dynamic atomic force microscope setup with a carbon contaminated Pt tip on a Au sample, Jarvis *et al.* found peaks at half integer values of  $G_0$  in the histogram [59]. However, it remains unclear whether the tip composition influenced these results.

Turning to Cu and Ag, recent reports [10, 60, 61] are also in agreement with the results of paper [II] (Fig. 3.10). Thus, Li and co-workers [60, 61] have shown that it is possible to form very stable Cu QPCs by electrochemical deposition and dissolution. A histogram built from repeated deposition/dissolution cycles has well defined peaks at integer values of up to  $5 G_0$  [60]. Using the MCBJ technique at 4.2 K, Ludoph and Ruitenbeek also acquired similar well resolved histograms for Cu and Ag [10].

In paper [II] it was suggested that the points of steepest descent in the noble metal conductance histogram correspond to a set of fully transmitting eigenchannels. This suggestion has been beautifully confirmed by the experiments of Ludoph and Ruitenbeek [9, 10, 62], who complemented measurements of the total conductance with measurements of the standard deviation of the voltage dependent differential conductance  $\sigma_{GV}$ . As explained on page 29 in Chapter 2  $\sigma_{GV}$  will be zero when all eigenchannels are fully closed or open. Indeed, the experiments show that  $\sigma_{GV}$  drops dramatically at the highend cutoffs of the histogram peaks for Au, Ag and Cu, confirming that this interpretation is correct. For Au, the same conclusions can be drawn from measurements of shot noise and thermopower fluctuations [8, 63].

---

<sup>9</sup>Confusingly, these groups have adopted the term ‘nonideal differential linearity’ instead of the technically correct ‘differential nonlinearity’ [25, 52].

### Skewness of histogram peaks

A few interesting papers which discuss the shape and positions of the histogram peaks have recently appeared [64–66]. García-Mochales and Serena [64] used a disordered 2D 1s tight binding model for simulating conductance histograms. The resulting histogram peaks were skew and looked qualitatively similar to the experimental histograms. Empirically, it was also found that the peak positions could be shifted to near integer values by subtracting a classical resistance of a few hundred Ohms.

Using a 2D hard-wall free-electron model with impurities represented by random distributions of delta functions in the scattering potential, Bürki *et al.* have obtained similar results [65], which are furthermore in excellent quantitative agreement with simultaneous conductance and shot noise measurements on Au by van den Brom and Ruitenbeek [8, 66].

These simple models suggest that the basic ingredients needed to explain the conductance of Au QPCs are quantum confinement and coherent backscattering from imperfections in the QPC [66]. However, it should be noted that the models are based on 2D geometries, which naturally includes the  $2 G_0$  peak, which is absent in 3D geometries with cylindrical symmetry [20]. Until now, the only 3D calculations which have been able to reproduce the  $2 G_0$  peak with a realistic height and width have been based on molecular dynamics simulations and simplified free-electron calculations [12]. However, these calculations resulted in almost symmetrical peaks.

### The resistance of the leads

In paper [II] it was found that the deviations of the histogram cutoffs from integer values of  $G_0$  were small. The shifts corresponded to a series resistance (of the leads), which was *probably* lower than  $100 \Omega$ . A truncated cone model of the leads resulted in a lead resistance value  $R_{\text{lead}}$  around  $0.5\text{--}10 \Omega$ . This estimate was based on bulk resistivities  $\rho$  of the noble metals. It was warned, however, that  $R_{\text{lead}} \propto \rho^2$  could be much larger due to contamination or disorder in the contact region.

The measurements of  $\sigma_{GV}$  by Ludoph and Ruitenbeek has shed new light on these speculations [9, 10, 62]. They observe that minima in  $\sigma_{GV}$  occurs at a total contact conductance slightly (but significantly) below integer values of  $G_0$ , corresponding to (in the case of Au) a series resistance of around  $130 \Omega$  [9]. From the  $\sigma_{GV}$  data, a value for the elastic mean free path around  $70 \text{ \AA}$  could also be deduced. This corresponds to a mean free path (and local conductivity)

close to the constriction, which is 5–6 times smaller than the bulk value even at RT (where inelastic events would normally dominate). Since  $R_{\text{lead}}$  (for not too small opening angles) is determined mainly by the resistivity in a region close to the constriction, the estimate for  $R_{\text{lead}}$  will increase by a factor of  $\sim 30$ . If we furthermore use a plausible opening angle of  $\alpha = 35\text{--}50^\circ$  [10, 57, 67] in Eq. (3.9), we obtain a value for  $R_{\text{lead}}$  around  $250 \Omega$ . Considering the crudeness of the model and our assumptions, this estimate is in fair agreement with the value of  $130 \Omega$  found by Ludoph and Ruitenbeek.

## References

- [1] M. Brandbyge and K. W. Jacobsen, “Quantum transmission channels in perturbed 3D nanowires,” In *Proceedings of the NATO Advanced Research Workshop on Nanowires, Madrid 1996*, P. A. Serena and N. García, eds., NATO ASI pp. 61–78 (Kluwer, Dordrecht, 1997).
- [2] M. Brandbyge, M. R. Sørensen, and K. W. Jacobsen, “Conductance eigenchannels in nanocontacts,” *Phys. Rev. B* **56**, 14956–14959 (1997).
- [3] M. Brandbyge, “Conductance of Metal Nanocontacts,” , PhD thesis, Technical University of Denmark, 1997.
- [4] M. Brandbyge, K. W. Jacobsen, and J. K. Nørskov, “Scattering and conductance quantization in three-dimensional metal nanocontacts,” *Phys. Rev. B* **55**, 2637–2650 (1997).
- [5] E. Scheer, P. Joyez, D. Esteve, C. Urbina, and M. H. Devoret, “Conduction Channel Transmissions of Atomic-Size Aluminum Contacts,” *Phys. Rev. Lett.* **78**, 3535–3538 (1997).
- [6] L. Olesen, E. Lægsgaard, I. Stensgaard, F. Besenbacher, J. Schiøtz, P. Stoltze, K. W. Jacobsen, and J. K. Nørskov, “Reply to ‘Comment on ‘Quantized conductance in an Atom-Sized Point Contact’’,” *Phys. Rev. Lett.* **74**, 2147 (1995).
- [7] A. I. Yanson and J. M. van Ruitenbeek, “Do Histograms Constitute a Proof for Conductance Quantization?,” *Phys. Rev. Lett.* **79**, 2157 (1997).
- [8] H. E. van den Brom and J. M. van Ruitenbeek, “Quantum Suppression of Shot Noise in Atom-Size Metallic Contacts,” *Phys. Rev. Lett.* **82**, 1526–1529 (1999).
- [9] B. Ludoph, M. H. Devoret, D. Esteve, C. Urbina, and J. M. van Ruitenbeek, “Evidence for Saturation of Channel Transmission from Conductance Fluctuations in Atomic-Size Point Contacts,” *Phys. Rev. Lett.* **82**, 1530–1533 (1999).
- [10] B. Ludoph and J. M. van Ruitenbeek, “Conductance fluctuations as a tool for investigating the quantum modes in atomic size metallic contacts,” *Phys. Rev. B* **61**, 2273–2285 (2000).

- [11] J. L. Costa-Krämer, N. García, P. García-Mochales, and P. A. Serena, “Nanowire formation in macroscopic metallic contacts: quantum mechanical conductance tapping a table top,” *Surf. Sci.* **342**, L1144–L1149 (1995).
- [12] M. Brandbyge, J. Schiøtz, M. R. Sørensen, P. Stoltze, K. W. Jacobsen, J. K. Nørskov, L. Olesen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, “Quantized conductance in atom-sized wires between two metals,” *Phys. Rev. B* **52**, 8499–8514 (1995).
- [13] H. Sauer, *Modern Relay Technology*, 2nd ed. (Hüthig, Heidelberg, 1986).
- [14] Ernst Weber, EURO-Matsushita Electric Works AG, Germany (private communication).
- [15] U. Landman, W. D. Luedtke, N. A. Burnham, and R. J. Colton, “Atomistic Mechanisms and Dynamics of Adhesion, Nanoindentation, and Fracture,” *Science* **248**, 454–461 (1990).
- [16] T. N. Todorov and A. P. Sutton, “Jumps in Electronic Conductance Due to Mechanical Instabilities,” *Phys. Rev. Lett.* **70**, 2138–2141 (1993).
- [17] A. M. Bratkovsky, A. P. Sutton, and T. N. Todorov, “Conditions for conductance quantization in realistic models of atomic-scale metallic contacts,” *Phys. Rev. B* **52**, 5036–5051 (1995).
- [18] U. Landman, W. D. Luedtke, B. E. Salisbury, and R. L. Whetten, “Reversible Manipulations of Room Temperature Mechanical and Quantum Transport Properties in Nanowire Junctions,” *Phys. Rev. Lett.* **77**, 1362–1365 (1996).
- [19] N. García and J. L. Costa-Krämer, “Quantum-Level Phenomena in Nanowires,” *Europhys. News* **27**, 89–91 (1996).
- [20] J. A. Torres and J. J. Sáenz, “Conductance and Mechanical Properties of Atomic-Size Metallic Contacts: A Simple Model,” *Phys. Rev. Lett.* **77**, 2245–2248 (1996).
- [21] C. Sirvent, J. G. Rodrigo, N. Agrait, and S. Vieira, “STM study of the atomic contact between metallic electrodes,” *Physica B* **218**, 238–241 (1996).

- [22] Z. Gai, Y. He, H. Yu, and W. S. Yang, "Observation of conductance quantization of ballistic metallic point contacts at room temperature," *Phys. Rev. B* **53**, 1042–1045 (1996).
- [23] J. M. Krans, J. M. van Ruitenbeek, V. V. Fisun, I. K. Yanson, and L. J. de Jongh, "The signature of conductance quantization in metallic point contacts," *Nature* **375**, 767–769 (1995).
- [24] C. J. Muller, J. M. Krans, T. N. Todorov, and M. A. Reed, "Quantization effects in the conductance of metallic contacts at room temperature," *Phys. Rev. B* **53**, 1022–1025 (1996).
- [25] J. Krugers, in *Instrumentation in Applied Nuclear Chemistry* (Plenum, New York, 1973), pp. 279–281.
- [26] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments* (Springer-Verlag, Berlin, 1987), p. 283.
- [27] J. L. Costa-Krämer and N. García (private communication).
- [28] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal* (Cambridge University Press, Cambridge, 1989).
- [29] L. Olesen, "Experimental Investigation of the Junction in a Scanning Tunneling Microscope," PhD thesis, University of Aarhus, 1996.
- [30] B. Hammer and J. K. Nørskov, "Why gold is the noblest of all the metals," *Nature* **376**, 238–240 (1995).
- [31] J. M. Krans, "Size Effects in Atomic-Scale Point Contacts," PhD thesis, Rijksuniversiteit Leiden, 1996.
- [32] C. J. Muller, J. M. van Ruitenbeek, and L. J. de Jongh, "Conductance and Supercurrent Discontinuities in Atomic-Scale Metallic Constrictions of Variable Width," *Phys. Rev. Lett.* **69**, 140–143 (1992).
- [33] J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Oxford University Press, London, 1892), Vol. 1, p. 426–434.
- [34] Y. V. Sharvin, "A possible method for measuring Fermi surfaces," *Zh. Eksp. Teor. Fiz.* **48**, 984–985 (1965), [*Sov. Phys. JETP* **21**, 655 (1965)].
- [35] G. Wexler, "The size effect and the non-local Boltzmann transport equation in orifice and disk geometry," *Proc. Phys. Soc. London* **89**, 927–941 (1966).

- [36] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [37] M. Büttiker, “Coherent and sequential tunneling in series barriers,” IBM J. Res. Dev. **32**, 63–75 (1988).
- [38] C. Sirvent, J. G. Rodrigo, S. Vieira, L. Jurczyszyn, N. Mingo, and F. Flores, “Conductance step for single-atom contact in the scanning tunneling microscope: Noble and transition metals,” Phys. Rev. B **53**, 16086–16090 (1996).
- [39] J. M. Krans, C. J. Muller, I. K. Yanson, T. C. M. Govaert, R. Hesper, and J. M. van Ruitenbeek, “One-atom point contacts,” Phys. Rev. B **48**, 14721–14724 (1993).
- [40] J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. M. Baró, N. García, U. Landman, W. D. Luedtke, E. N. Bogachek, and H.-P. Cheng, “Properties of Metallic Nanowires: From Conductance Quantization to Localization,” Science **267**, 1793–1795 (1995).
- [41] H. Yasuda and A. Sakai, “Conductance of atomic-scale gold contacts under high-bias voltages,” Phys. Rev. B **56**, 1069–1072 (1997).
- [42] H. Yasuda and A. Sakai, “Observation of Quantum Conductance in Commercial Gold-Contact Relays,” J. Kor. Phys. Soc. **31**, S54–S57 (1997).
- [43] K. Itakura, K. Yuki, S. Kurokawa, H. Yasuda, and A. Sakai, “Bias dependence of the conductance of Au nanocontacts,” Phys. Rev. B **60**, 11163–11170 (1999).
- [44] K. Yuki, A. Enomoto, and A. Sakai (unpublished).
- [45] K. Yuki, S. Kurokawa, and A. Sakai (unpublished).
- [46] F. Ott and J. Lunney, “Quantum Conduction: a Step-by-Step Guide,” Europhys. News **29**, 13–16 (1998).
- [47] F. Ott and J. Lunney, “Retracing quantum steps,” Europhys. News **29**, 55 (1998).
- [48] E. L. Foley, D. Candela, K. M. Martini, and M. T. Tuominen, “An undergraduate laboratory experiment on quantized conductance in nanocontacts,” Am. J. Phys. **67**, 389–393 (1999).

- [49] J. L. Costa-Krämer, N. García, and H. Olin, “Conductance quantization histograms of gold nanowires at 4 K,” *Phys. Rev. B* **55**, 12910–12913 (1997).
- [50] J. L. Costa-Krämer, N. García, P. García-Mochales, P. A. Serena, M. I. Marqués, and A. Correia, “Conductance quantization in nanowires formed between micro and macroscopic metallic electrodes,” *Phys. Rev. B* **55**, 5416–5424 (1997).
- [51] C. Shu, C. Z. Li, H. X. He, A. Bogozi, J. Bunch, and N. J. Tao (unpublished).
- [52] *LeCroy Test & Measurement Products Catalog*, 1997, see glossary of technical terms.
- [53] W. A. de Heer, S. Frank, and D. Ugarte, “Fractional quantum conductance in Gold nanowires,” *Z. Phys. B* **104**, 469–473 (1997).
- [54] T. Junno, S.-B. Carlsson, H. Xi, L. Montelius, and L. Samuelson, “Fabrication of quantum devices by Ångström-level manipulation of nanoparticles with an atomic force microscope,” *Appl. Phys. Lett.* **72**, 548–550 (1998).
- [55] H. Ohnishi, Y. Kondo, and K. Takayanagi, “Quantized conductance through individual rows of suspended gold atoms,” *Nature* **395**, 780–783 (1998).
- [56] A. I. Yanson, G. R. Bollinger, H. E. van den Brom, N. Agraït, and J. M. van Ruitenbeek, “Formation and manipulation of a metallic wire of single gold atoms,” *Nature* **395**, 783–785 (1998).
- [57] W. B. Jian, C. S. Chang, W. Y. Li, and T. T. Tsong, “Geometrical dependence of conductance quantization in metal point contacts,” *Phys. Rev. B* **59**, 3168–3172 (1999).
- [58] E. Scheer, N. Agraït, J. C. Cuevas, A. L. Yeyati, B. Ludoph, A. Martín-Rodero, G. R. Bollinger, J. M. van Ruitenbeek, and C. Urbina, “The signature of chemical valence in the electrical conduction through a single-atom contact,” *Nature* **394**, 154–157 (1998).
- [59] S. P. Jarvis, M. A. Lantz, H. Ogiso, H. Tokumoto, and U. Dürig, “Conduction and mechanical properties of atomic scale gold contacts,” *Appl. Phys. Lett.* **75**, 3132–3134 (1999).

- [60] C. Z. Li and N. J. Tao, “Quantum transport in metallic nanowires fabricated by electrochemical deposition/dissolution,” *Appl. Phys. Lett.* **72**, 894–896 (1998).
- [61] C. Z. Li, A. Bogozi, W. Huang, and N. J. Tao, “Fabrication of stable metallic nanowires with quantized conductance,” *Nanotech.* **10**, 221–223 (1999).
- [62] B. Ludoph, “Quantum Conductance Properties of Atomic-size Contacts,” PhD thesis, Rijksuniversiteit Leiden, 1999.
- [63] B. Ludoph and J. M. van Ruitenbeek, “Thermopower of atomic-size metallic contacts,” *Phys. Rev. B* **59**, 12290–12293 (1999).
- [64] P. García-Mochales and P. A. Serena, “Disorder as Origin of Residual Resistance in Nanowires,” *Phys. Rev. Lett.* **79**, 2316–2319 (1997).
- [65] J. Bürki, C. A. Stafford, X. Zotos, and D. Baeriswyl, “Cohesion and conductance of disordered metallic point contacts,” *Phys. Rev. B* **60**, 5000–5008 (1999).
- [66] J. Bürki and C. A. Stafford, “Comment on “Quantum Suppression of Shot Noise in Atomic-Size Metallic Contacts”,” *Phys. Rev. Lett.* **83**, 3342 (1999).
- [67] C. Untiedt, G. Rubio, S. Vieira, and N. Agraït, “Fabrication and characterization of metallic nanowires,” *Phys. Rev. B* **56**, 2154–2160 (1997).



## Chapter 4

# Fast and Accurate Current-Voltage Curves

The relay experiments described in the previous Chapter used a simple setup to demonstrate that the conductance is quantized in certain commercial electro-mechanical relays. In our opinion, such a setup is useful because it gives a clear demonstration of the quantum-mechanical wave nature of conducting electrons in confined geometries. Notwithstanding, it also has severe limitations of which we mention:

- The contacts are disrupted during very short time-scales ( $< 100 \mu\text{s}$ , see Fig. 3.3 on page 51).
- The supply voltage  $V_0$  is fixed in the setup (see Fig. 3.1). Hence, it is impossible to study the voltage dependence of the conductance for a certain contact geometry. The very short time which is available for such measurements is also a severe hindrance.
- The cleanliness of the system is unknown and beyond our control.

Most of these shortcomings can be solved by using alternative and more sophisticated setups, such as a scanning tunneling microscope (STM) working under ultra-high vacuum (UHV) conditions.

Such a setup has previously been used in our group by Olesen, who achieved several interesting results (see Refs. 1–4 and paper [I]). This setup was based

(with some modifications) on a standard set of STM electronics and was therefore best suited for measuring very small currents in the nA regime. The time-resolution of the setup was also limited to  $\sim 10 \mu\text{s}$  due to a fairly low bandwidth of the low-noise current preamplifier and reasonably long sampling times for the high resolution analog-to-digital converter monitoring the tunnel current. The setup could therefore not be used to measure, e.g. current-voltage ( $I$ - $V$ ) characteristics<sup>1</sup> of (often transient) contact geometries because the current would saturate the preamplifier and the data acquisition would be too slow.

We therefore decided to modify the electronics and software in one of our STM setups to facilitate fast and accurate acquisition of  $I$ - $V$  curves from transient contact geometries. Details of this setup and its performance are described in paper [III] and will be the subject of this Chapter.

## 4.1 Review

Recently, there has been a growing interest in measuring the  $I$ - $V$  curves of QPCs to obtain further information on their electrical properties [5–12]. At liquid He temperatures or below, the QPCs can be very stable for a long period of time. Under such conditions, highly nonlinear  $I$ - $V$  curves on QPCs formed between superconducting electrodes at bias voltages extending up to the mV regime have been recorded and used to successfully determine the transmission of each mode contributing to the current [5, 6]. Measurements of fluctuations in the differential conductance  $dI/dV$  of noble metal QPCs up to 100 mV at 4.2 K have shown that the conduction eigenchannels tend to close one by one during rupture of the contacts [7].

At room temperature (RT), the QPCs tend to be much more unstable due to thermal fluctuations. In spite of this, there has been a number of reports on the formation of stable Au QPCs at RT [8–13]. For these contacts, nonlinear  $I$ - $V$  curves extending up to 1 V have been recorded during a time-scale of seconds [8–12].

When working in a clean ultra-high vacuum (UHV) environment at RT, we find that Au QPCs tend to break spontaneously over a period of less than one ms. During the break, the conductance goes down in a series of metastable steps, each corresponding to a specific atomic configuration [14]. The duration of a single step is typically less than  $\approx 100 \mu\text{s}$ , whereas the size of a step is close to  $G_0$ .

To capture the  $I$ - $V$  curves of these metastable states, we have therefore developed a novel setup which is capable of acquiring a full  $I$ - $V$  curve during a

---

<sup>1</sup>Except for very small voltage sweeps.

period of down to  $10 \mu\text{s}$  with an acquisition rate of 100 Msamples/s.

## 4.2 Setup

A block diagram of the setup can be seen in Fig. 4.1. A Hewlett Packard 33120A, 15 MHz arbitrary waveform/function generator (FG) supplies a bias voltage to a QPC. A home-built  $I$ - $V$  converter is used to transform the resulting current  $I$  into a voltage signal  $V_{\text{out}}$ , which in turn is measured with a 10-bit true resolution, 100 Msamples/s, four-channel Yokogawa DL 4100 digital storage oscilloscope (DSO). Individual voltage traces, each consisting of up to 100k data points, are transferred from the DSO to a personal computer (PC) using a general purpose interface bus (GPIB). Each trace is stored for later analysis. All relevant settings on the DSO and FG are stored along with the trace.

Normally, the FG supplies a constant bias voltage to the QPC (the offset voltage). However, when the FG receives a trigger signal from the DSO, the FG can add a burst signal to the offset voltage. In our experiments, the burst has a predefined waveform (triangular or sinusoidal), number of cycles (1 or more), amplitude (0.05–5 V), and frequency ( $\leq 100$  kHz). All the settings on the FG and the DSO are controlled from the PC via the GPIB.

The trigger signal to the FG is sent out when the DSO itself is triggered. This occurs when (i) the operator triggers the DSO manually from the PC

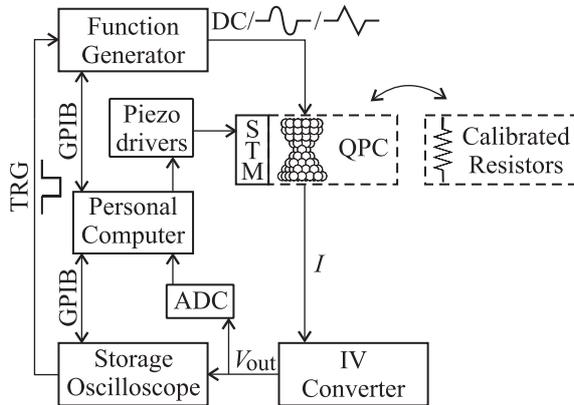


Figure 4.1: Block diagram of the experimental setup.

via the GPIB, or (ii) the voltage signal  $V_{\text{out}}$  from the  $I$ - $V$  converter fulfills a predefined trigger criterion. A simple example could be when  $V_{\text{out}}$  drops below a preset voltage corresponding to a specific value of the conductance.

The QPCs are formed by indenting the tip of a home-built scanning tunneling microscope (STM) into a metal surface. The STM has been described elsewhere [15, 16]. Three piezo drivers controls the lateral ( $x$ ,  $y$ ) and vertical ( $z$ ) movement of the scanner tube on which the tip is mounted. Three other piezo drivers are used for controlling an inchworm motor, which is used for coarse approach of the tip to the sample. Each piezo driver consists of a 16-bit digital-to-analog converter (DAC) and a low-noise, high-voltage amplifier. The drivers are all part of a home-built set of STM electronics. The DAC, which is normally used for setting the current via a loop amplifier in constant-current mode, is instead connected directly to the  $z$ -piezo amplifier. In this manner, the *full* motion of the tip is under direct computer control.

A 16-bit analog-to-digital converter (ADC), which is normally intended for reading the output from the STM tunnel current pre-amplifier, is instead connected to the output of the  $I$ - $V$  converter. The ADC has a sampling time of 30  $\mu\text{s}$ , much shorter than the time-scale on which information about the current value of  $V_{\text{out}}$  can be transferred from the DSO to the PC ( $\approx 500$  ms). The ADC readings are needed when we attempt to make QPCs with a preselected value of  $G$ : A feedback loop is implemented in the PASCAL program which controls the motion of the tip. When  $V_{\text{out}}$  as read by the ADC is higher (lower) than a predefined value of  $G$ , the tip is retracted from (approached to) the sample in order to make the cross-sectional area of the QPC smaller (larger) and thus reduce (increase)  $G$ ,  $I$ , and  $V_{\text{out}}$ . The ADC is also read regularly while indenting (retracting) the tip to record the  $z$ -positions at which contact between the tip and sample is made (lost).

During calibration, the STM is replaced by a set of high-precision resistors. By acquiring voltage traces from the DSO while different constant bias voltages are applied, the detailed response of the setup can be measured and represented by a set of calibration parameters. Details of the calibration procedure will be presented in Sec. 4.5.

### 4.3 The current-to-voltage converter

Transimpedance and shunt setups (see Fig. 4.2) have frequently been used for constant bias voltage measurements of the conductance in QPCs. A common feature of the setups is a bias voltage supply  $V_b$ , which drives a current  $I$  through

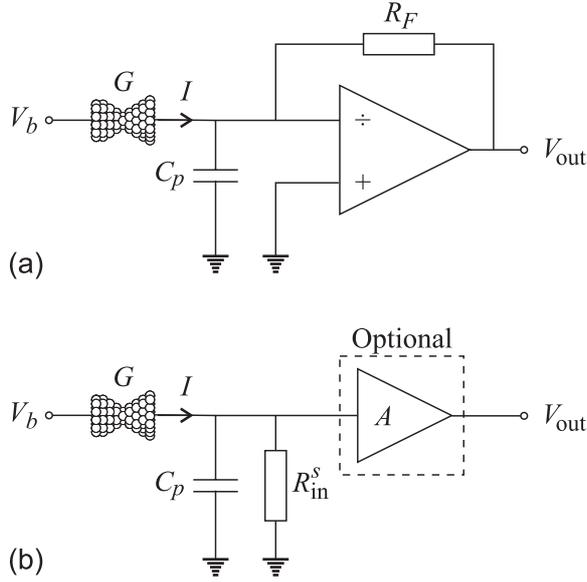


Figure 4.2: (a) Transimpedance and (b) shunt method for measuring the conductance in a QPC.

a QPC with conductance  $G$  (usually,  $G = 1-10 G_0$  corresponding to  $1-10 \text{ k}\Omega$ ). When dealing with short response time setups, the parasitic capacitance  $C_p$  following the QPC also has to be considered.

In a transimpedance setup [Fig. 4.2(a)], the current is converted to a voltage signal  $V_{out}$  by passing it through a measuring resistor  $R_F$ , placed in the feedback of an operational amplifier (OA) [13, 17–20]. In a shunt setup [Fig. 4.2(b)], a measuring resistor  $R_{in}^s$  is placed in series with the QPC, and the induced voltage drop in  $R_{in}^s$  is measured either directly [21–23] or after amplification [24].

In the following, we will comment on the strengths and weaknesses of these two basic setups. Other sophisticated types of setups also exist which are useful for measuring the conductance of a QPC over many orders of magnitude (from tunneling to point contact) [3, 4, 25]. Very accurate measurements of the differential conductance and its derivative can also be made with lock-in techniques. However, the bandwidth in those setups tends to be much smaller than what is required here.

### Transimpedance converters

The input impedance of the simple transimpedance amplifier [see Fig. 4.2(a)] is given by

$$R_{\text{in}}^t = \frac{R_F}{1 + A(f)}, \quad (4.1)$$

where  $A(f)$  is the frequency-dependent open loop gain of the OA. For a typical OA, such as the OP07 used by Foley *et al.* [20],  $A \approx 5 \times 10^5$  from DC to a frequency of 5 Hz whereafter it drops 20 dB per decade (unity gain at 1 MHz) [26]. In typical setups ( $R_F = 10 \text{ k}\Omega$ – $1 \text{ M}\Omega$ ,  $G = 1$ – $10 \text{ G}_0$ ) [9, 13, 17–20],  $R_{\text{in}}^t$  can be neglected at low frequencies, whereas for higher frequencies  $R_{\text{in}}^t$  may become comparable to  $G^{-1}$ .

The transfer equation for frequency components of the input current  $I(f)$  and output voltage  $V_{\text{out}}(f)$  is given by

$$V_{\text{out}}(f) = -\frac{A(f)}{1 + A(f)} R_F I(f). \quad (4.2)$$

When the time-dependent current signal  $I(t)$  is composed of a wide spectrum of Fourier components, the frequency-dependent transimpedance tends to degrade the quality of the signal. For instance, when a QPC is stretched,  $G$  usually drops in abrupt steps due to sudden atomic rearrangements in the contact region. The resulting output voltage loses some of its high-frequency components in the transimpedance converter as can be seen from Eq. (4.2). This leads to smearing or ringing at the steps in the conductance curves.

If the OA is assumed to be ideal ( $A \gg 1$ , and thus  $R_{\text{in}}^t \ll G^{-1}$ ), the relation between  $G$  and  $V_{\text{out}}$  is particularly simple: With an applied bias voltage  $V_b$ , the current will be  $I = GV_b$ . Thus, it follows from Eq. (4.2) that  $V_{\text{out}} = -R_F V_b G$ . This simple relation makes the transimpedance setup particularly useful in undergraduate laboratory experiments [19, 20]. Furthermore, when the OA is ideal, the QPC will be virtually grounded. Therefore, if  $V_b$  is kept constant, the voltage drop  $V$  across the QPC will be independent of  $G$  [ $V(G) = V_b$ ]. Hereby the relative energy shift  $eV$  of the Fermi energies in the electrodes which connect the QPC to the macroscopic world, is kept constant. This is advantageous since theoretical calculations are often performed under this assumption.

### Shunt converters

In the shunt setup [see Fig. 4.2(b)],  $I$  is converted to a voltage by passing it through a measuring resistor  $R_{\text{in}}^s$  with a value of typically  $2 \text{ k}\Omega$  or less [21–

24]. By using a passive resistor to convert the current, the input impedance becomes frequency-independent (neglecting the shunt capacitance of  $\approx 0.2$  pF) in sharp contrast to the transimpedance amplifier, cf. Eq. (4.1). By placing an optional low-gain, low-noise, high-input impedance, flat frequency response voltage amplifier immediately after  $R_{\text{in}}^s$ ,  $C_p$  will be kept at a minimum. The voltage signal will increase, and long transmission lines between the output terminal and the DSO can be used.

The QPCs are often formed by first indenting the STM tip deep into the sample, whereby  $G^{-1}$  decreases to a value lower than  $200 \Omega$  [3]. The QPC is then formed by retracting the tip. In these initial, low-resistance stages of the contact formation,  $I$  is naturally limited by  $R_{\text{in}}^s$ , which also dissipates most of the power. Heating in the contact region is hereby reduced. In contrast to the transimpedance amplifier,  $V(G) \neq V_b$  since  $R_{\text{in}}^s$  is placed in series with the QPC. However, in the most interesting conductance regime ( $1-10 G_0$ ),  $R_{\text{in}}^s \ll G^{-1}$ , which means that  $V(G) \lesssim V_b$ , and  $V$  will only depend weakly on  $G$ .

Under these conditions,  $C_p$  will mainly be discharged via  $R_{\text{in}}^s$ . The overall response time of the setup will therefore be limited by the time constant of this  $RC$  circuit

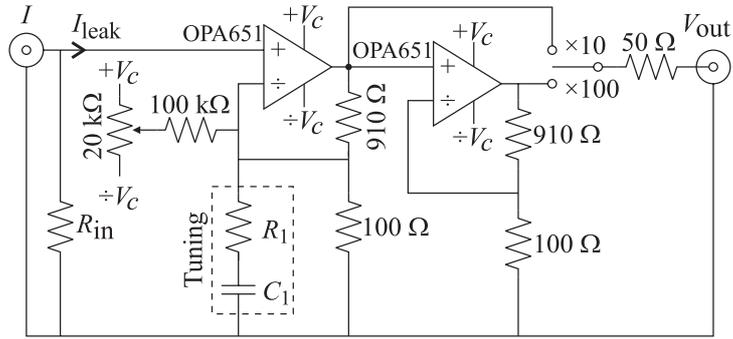
$$\tau \approx R_{\text{in}} C_p.$$

In our case,  $C_p \approx 50$  pF due to the wiring from the STM to the  $I$ - $V$  converter. It cannot be reduced since the STM and the converter are placed on opposite sides of a UHV flange, 25 cm apart. With  $R_{\text{in}}^s = 500 \Omega$ , we therefore get  $\tau \approx 25$  ns, corresponding to a cutoff frequency of 6 MHz.

For a transimpedance setup, this time constant depends on the frequency characteristics of the OA, cf. Eq. (4.1). Normally,  $\tau$  will be comparable to that of the shunt setup.

## Diagram and construction

In the experiments, we want to apply short voltage bursts and measure the current response. As we will see later, the burst consists of a wide range of Fourier frequency components. To obtain high accuracy, it is therefore essential that the frequency response of the  $I$ - $V$  converter is flat over a wide range of frequencies. This requires an input impedance and an amplification which are independent of frequency. Following the previous discussion, we have therefore based the design of our  $I$ - $V$  converter on the shunt setup in Fig. 4.2(b), followed by a flat-frequency response, low-noise, high-bandwidth voltage amplifier.

Figure 4.3: Diagram of the home-built  $I$ - $V$  converter.

A diagram of the converter is shown in Fig. 4.3. The voltage amplifier is based on two identical, non-inverting OPA651 (Burr-Brown) operational amplifier stages [27]. Each stage has a gain of  $A \simeq 10$ , as determined by the  $100\ \Omega$  and  $910\ \Omega$  surface mounted feedback resistors. Initially, the converter is assembled without the tuning components  $R_1$  and  $C_1$ . Based on detailed measurements of the frequency response of the converter, the tuning components are then selected and mounted. This procedure will be discussed on page 91.

The impedance seen by the two inputs on an OA is not necessarily the same. It depends on the value of the interchangeable metal-film shunt resistor  $R_{in}$  ( $0.2$ – $2\ \text{k}\Omega$ ), which is mounted in a socket. Since the OA has a sizeable input bias current of approximately  $4\ \mu\text{A}$  [27], this mismatch in impedance leads to an offset voltage of several mV on the output of the first stage. Consequently, a  $20\ \text{k}\Omega$  multiturn potentiometer connected to the supply voltages  $\pm V_c$  ( $V_c = 5\ \text{V}$ ) is used together with a  $100\ \text{k}\Omega$  resistor for adjusting the offset voltage whenever  $R_{in}$  is changed.

In most applications, only the first amplifier stage is used. For low  $V_b$ , low  $R_{in}$  studies, a high-frequency (HF) quality switch can be used to activate the second stage, whereby the overall gain of the voltage amplifier is increased to  $A \simeq 100$ . The output of the amplifier is series terminated with  $50\ \Omega$  to match the impedance of a 4 m coaxial cable, which is used as a transmission line to the DSO. To preserve signal integrity [27], the cable is also  $50\ \Omega$  terminated at the DSO at the cost of a 6 dB attenuation loss. Thus, the effective gain of the voltage amplifier is 5 (50) for one (two) stage(s). To assure the best HF performance of the amplifier, all design recommendations have been followed [27].

### Frequency response and tuning

We use the FG and the DSO for measuring the transfer characteristics of the converter up to 15 MHz. A harmonic voltage signal from the FG is measured before and after amplification. The input of the converter is coupled directly to an input line of the DSO using BNC connectors. We insert  $R_{\text{in}} = 100 \Omega$  to decouple the input bias current. The output of the converter is connected with a 30 cm long, unterminated coaxial cable to the DSO.

For each frequency, a snapshot of the signal at the two input lines is acquired and transferred to the PC. The voltage traces are corrected for the differential nonlinearity (DNL) of the DSO (see pages 52–57 in Chapter 3) and fitted to functions of the type  $V(t) = V_0 \sin(2\pi ft + \phi) + V_{\text{off}}$  with the time  $t$  as the dependent parameter, and the amplitude  $V_0$ , phase  $\phi$ , and offset voltage  $V_{\text{off}}$  as free parameters. The frequency  $f$  is fixed. When calculating the phase shift of the converter, the phase difference between the two signals is compensated for the finite signal speed in the cables. The voltage gain is simply the ratio between the output and input amplitudes.

These measurements can be used for tuning the feedback. This is illustrated in Fig. 4.4, where fine details of the gain and phase for the first amplifier stage before (open circles) and after (filled squares) insertion of  $R_1$  and  $C_1$  are shown. Note that the gain scale is linear and extends only from 9.5 to 10. Although the frequency response prior to the tuning is quite flat, a clear 1.2% (0.11 dB) decrease occurs in the gain between the lower frequencies (1–10 kHz) and the higher frequencies (30 kHz–1 MHz).<sup>2</sup> Such minute changes depend on details of the board layout and the performance of the individual OA and can, in general, not be predicted. However, a compensation can be made by tuning the gain of the feedback, such that it gradually increases by 1.2% around 20 kHz. This is approximately achieved by mounting a tuning resistor  $R_1 = 6.5 \text{ k}\Omega$  and capacitor  $C_1 = 1 \text{ nF}$  in the feedback, as shown in Fig. 4.3.

After tuning, the gain is now constant within 0.5% (1.0%) for frequencies up to 800 kHz (11 MHz). The phase shift is also lowered to a value  $\ll 0.5^\circ$  for frequencies in the range 10–100 kHz. The second amplifier stage could be tuned in a similar way. However, we have not found this necessary.

In Fig. 4.5 we show the final transmission characteristics for one (filled symbols) and two (open symbols) amplifier stages. The gain is now plotted on the usual logarithmic scale. The gains are close to 10 (100) for one (two) amplifier

---

<sup>2</sup>Above 1 MHz the gain increase again up to 3 MHz, where the cutoff sets in. However, since the burst frequency is never larger than 100 kHz, tuning of this frequency regime is not required.

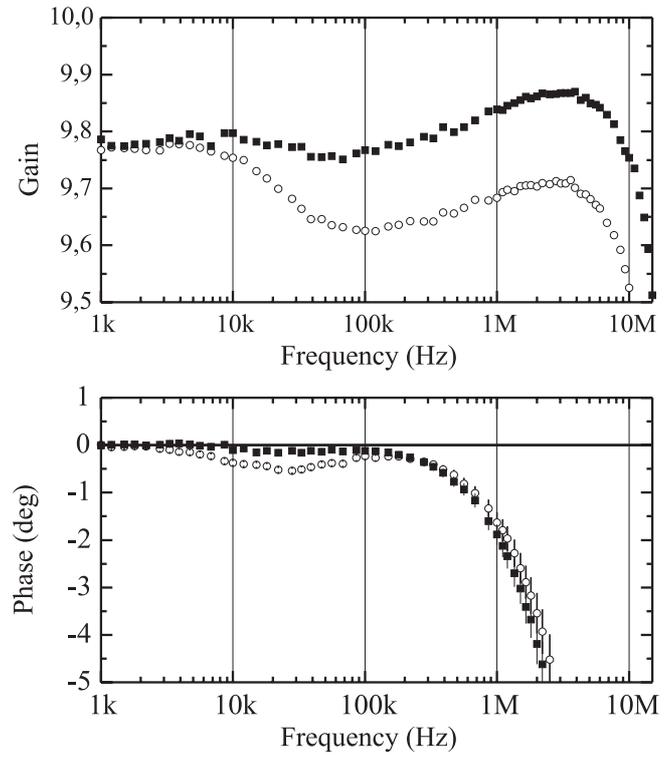


Figure 4.4: The detailed frequency response of the first amplifier stage in the  $I$ - $V$  converter before ( $\circ$ ) and after ( $\blacksquare$ ) insertion of the tuning components  $R_1 = 6.5 \text{ k}\Omega$  and  $C_1 = 1 \text{ nF}$  in the feedback of the circuit shown in Fig. 4.3.

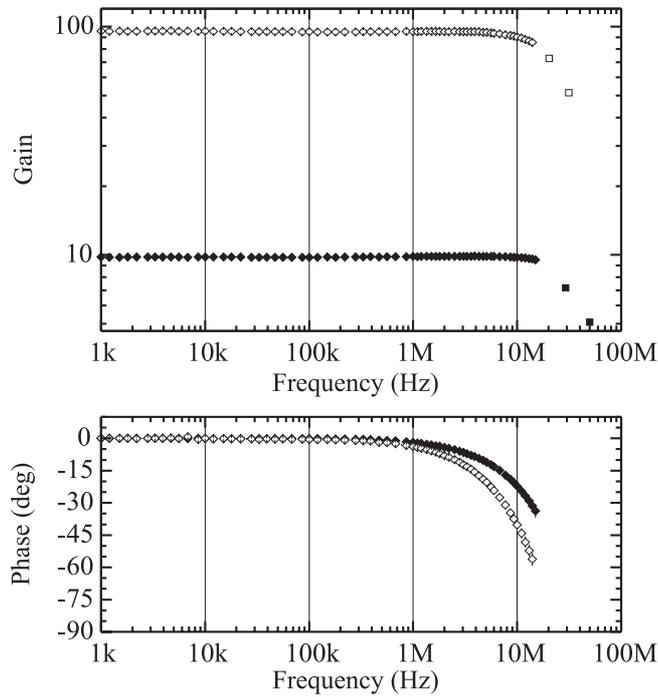


Figure 4.5: Frequency response of the  $I$ - $V$  converter after tuning when one ( $\diamond$ ,  $\square$ ) or two ( $\blacklozenge$ ,  $\blacksquare$ ) OPA651 amplifier stages are used. Most measurements ( $\diamond$ ,  $\blacklozenge$ ) were done using the FG and DSO from the setup. Four measurements ( $\square$ ,  $\blacksquare$ ) were done using a spectrum analyzer.

stage(s) as expected up to at least 10 MHz, whereafter they gradually decrease. Noticeable phase shifts are observed only above 1 MHz.

Measurements of the 3 and 6 dB bandwidths (BW) are also included (as squares) in the gain panel of Fig. 4.5. These measurements were done using a Hewlett Packard 4195A spectrum analyzer. The 3 dB BW for a single stage is 29 MHz, which is only slightly lower than the optimum value of 34 MHz [27]. When two amplifier stages are used, the BW decreases to 20 MHz.

As will be shown in Sec. 4.9,  $I$ - $V$  curves of QPCs can be almost Ohmic (linear). Since the voltage bursts extend over a finite number of cycles, even sinusoidal bursts are composed of a wide spectrum of Fourier components. The

widest spectra occur for single period bursts. To detect small nonlinearities in the measured  $I$ - $V$  curves, it is therefore essential that a large majority of the Fourier components are transmitted with the same gain and phase.

In Fig. 4.6 we show the calculated power spectral density (PSD) (the norm square of the Fourier transform) for single period sinusoidal and triangular bursts. The PSDs for the two bursts are quite similar and dominated by contributions from zero frequency to twice the burst frequency. Above three times the burst frequency, the PSD is negligible. In our experiments, the burst frequency is never larger than 100 kHz, which requires a flat frequency response up to 300 kHz. As can be seen from Fig. 4.5, this requirement is easily met with the  $I$ - $V$  converter. 300 kHz is also well below the cutoff frequency determined by  $R_{\text{in}}$  and  $C_p$  ( $\lesssim 6$  MHz).

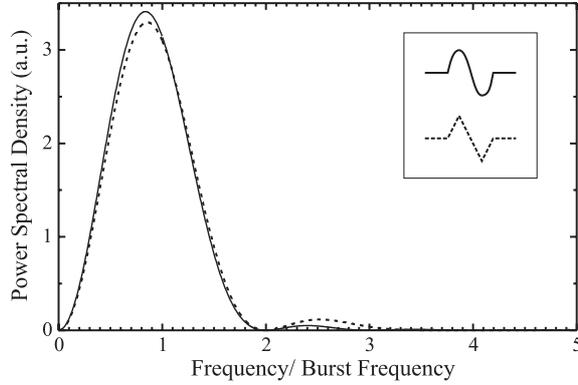


Figure 4.6: Power spectral density of a single period sinusoidal (solid) or triangular (dashed) burst.

#### 4.4 Model of the circuit

As indicated in Fig. 4.1, we calibrate the setup by interchanging the QPC with a set of calibrated resistors and measuring  $V_{\text{out}}$  as a function of bias voltage for each resistor. Hereby a set of parameters, such as  $R_{\text{in}}^s$  and  $A$  in Fig. 4.2(b), which accurately describe the behavior of the circuit, can be found.

However, for our setup, more than two parameters are needed for an accurate calibration. A realistic model based on five parameters is shown in Fig. 4.7. This

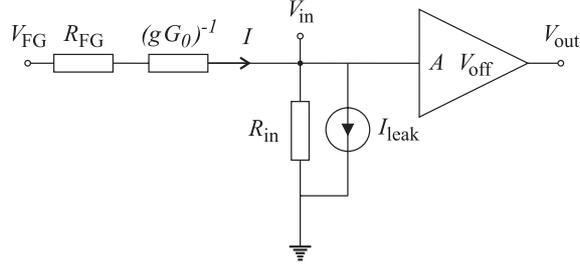


Figure 4.7: The model used for calculating the transfer characteristics of the setup.

is a DC model, where all parasitic inductances and capacitances are disregarded. This is justified since we always operate in the flat frequency response regime. The finite output impedance of the FG ( $\approx 50 \Omega$ ) is represented by  $R_{\text{FG}}$ . We write the conductance of a calibrated resistor/the QPC as  $G = gG_0$  implying that  $g$  is the conductance measured in quantum units. The current  $I$  that flows through the QPC is divided between the input impedance  $R_{\text{in}}$  and the OA input bias current, which is represented by the current sink  $I_{\text{leak}}$ . The voltage amplifier is represented by a frequency-independent gain  $A$  and an offset voltage  $V_{\text{off}}$ , such that

$$V_{\text{out}} = AV_{\text{in}} + V_{\text{off}}. \quad (4.3)$$

where  $V_{\text{in}}$  is the input voltage on the  $I$ - $V$  converter. The 6 dB attenuation loss due to the doubly terminated transmission line between the converter and the DSO as well as the voltage scaling error on the DSO are absorbed in  $A$ . Similarly, the voltage axis offset error on the DSO is absorbed in  $V_{\text{off}}$ .

From the model in Fig. 4.7, we get

$$V_{\text{in}} = \frac{R_{\text{in}}}{R_{\text{FG}} + (gG_0)^{-1} + R_{\text{in}}} V_{\text{FG}} - \left\{ R_{\text{in}}^{-1} + \left[ R_{\text{FG}} + (gG_0)^{-1} \right]^{-1} \right\}^{-1} I_{\text{leak}}. \quad (4.4)$$

When Eq. (4.4) is inserted in Eq. (4.3),  $V_{\text{out}}$  may (with some effort) be written as

$$V_{\text{out}} = f_1(g)V_{\text{FG}} + f_2(g), \quad (4.5)$$

where

$$f_i(g) = \frac{\alpha_i g}{g + \beta_i} + \gamma_i, \quad i = 1, 2 \quad (4.6)$$

are nonlinear functions of  $g$ . The parameters  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  are given by

$$\alpha_1 = \frac{A}{1 + R_{\text{FG}}/R_{\text{in}}}, \quad (4.7a)$$

$$\alpha_2 = \frac{AI_{\text{leak}}R_{\text{in}}}{1 + R_{\text{FG}}/R_{\text{in}}}, \quad (4.7b)$$

$$\beta_1 = \beta_2 = \frac{1}{G_0(R_{\text{FG}} + R_{\text{in}})}, \quad (4.7c)$$

$$\gamma_1 = 0, \quad (4.7d)$$

$$\gamma_2 = -AI_{\text{leak}}R_{\text{in}} + V_{\text{off}}. \quad (4.7e)$$

## 4.5 Calibration

To avoid thermal drift, it is important to warm up the equipment at least one hour before the calibration. A calibration is only reliable when the V/div setting on the DSO is kept constant throughout the calibration and experiment. Normally, our DSO auto calibrates every 30 minutes after warm up. This feature is turned off during calibration and experiments to avoid sudden changes in the offset and scaling errors. Also, to reduce the noise in voltage traces, a 20 MHz bandwidth filter is used on the DSO.

The setup is calibrated using a set of 12 calibrated resistors with conductances  $g_j$ ,  $j = 1 \dots 12$  around 0, 0.5, 1, 2, 3.5, 6, 9, 12, 25, 50, 100, and 400 (in units of  $G_0$ ). For each resistor 10–20 different values of  $V_{\text{FG}}$  are applied by adjusting the offset voltage of the FG. Then the DSO is triggered from the PC, whereby a constant voltage trace is acquired. Finally, the trace is transferred to the PC, corrected for DNL, and  $V_{\text{out}}$  is calculated as its mean value.

Two of the  $V_{\text{FG}}$  values are chosen such that the resulting  $V_{\text{out}}$  is close to the minimum and maximum values that can be measured by the DSO (except of course for  $g=0$ , where  $V_{\text{out}}$  is independent of  $V_{\text{FG}}$ ). The remaining voltages are chosen *randomly* between these extremes (to average out systematic errors).

For each conductance  $g_j$ , the corresponding function values  $f_1(g_j)$  and  $f_2(g_j)$  can now be found by making a linear fit to Eq. (4.5). The parameters in Eq. (4.6)

are subsequently found by fitting the twelve known values of  $f_1(g_j)$  and  $f_2(g_j)$  (and their uncertainties) to Eq. (4.6) using the Levenberg-Marquandt method for minimizing  $\chi^2$  [28]. Initial guesses of the parameters are obtained from Eq. (4.7) and the approximate circuit parameters in Fig. 4.7. Although  $\beta_1 = \beta_2$  according to Eq. (4.7c), we relax this restriction in the fitting procedure. Thus, we determine five free parameters, which is also the number of free parameters in our original model, cf. Fig. 4.7 ( $\gamma_1 = 0$  is fixed during fitting).

Care should be taken when fitting to a functional form like Eq. (4.6). Usually  $\beta_i$  in Eq. (4.7c) has a value of 6–50. When  $g \ll \beta_i$ , the parameters  $\alpha_i$  and  $\beta_i$  will be fully dependent [29] since  $f_i$  will only depend on the fraction  $\alpha_i/\beta_i$ . Similarly, when  $g \gg \beta_i$ , there will be full dependency between  $\alpha_i$  and  $\gamma_i$ . Therefore, to keep the dependencies between all parameters at a reasonable level, it is important to calibrate with several resistors that span a wide range of conductances. With our set of calibrated resistors, the dependency between any two parameters is always smaller than 90% and typically around 50%. This is sufficient for our purpose.

In Table 4.1 we give an example of the fitted parameters that result from a typical calibration. The fitted values are in good agreement with the rough initial guesses, and the most important parameters  $\alpha_1$  and  $\beta_1$  are determined with good accuracy. The goodness-of-fit is high, indicating that the order of magnitude of the stated uncertainties is reliable [28].

## 4.6 Calculating the conductance, current and voltage

The calibration procedure allows us to calculate  $V_{\text{out}}(g, V_{\text{FG}})$  using Eqs. (4.5) and (4.6). However, in experiments, we are interested in finding  $g(V_{\text{FG}}, V_{\text{out}})$ . This is done as follows: If we insert Eq. (4.6) in Eq. (4.5) and solve for  $g$ , we obtain a quadratic equation with the usual solution

$$g = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (4.8)$$

With  $\gamma_1 = 0$  and independent values for  $\beta_1$  and  $\beta_2$ , the coefficients in Eq. (4.8) are given by

$$a = \alpha_2 + \gamma_2 + \alpha_1 V_{\text{FG}} - V_{\text{out}}, \quad (4.9a)$$

$$b = (\beta_1 + \beta_2)(\gamma_2 - V_{\text{out}}) + \alpha_2 \beta_1 + \alpha_1 \beta_2 V_{\text{FG}}, \quad (4.9b)$$

$$c = \beta_1 \beta_2 (\gamma_2 - V_{\text{out}}). \quad (4.9c)$$

Parameter in Eq. (4.6)	Initial guess	Fitted value
$\alpha_1$	4.53	$4.5817 \pm 0.0005$
$\beta_1$	24.1	$23.906 \pm 0.004$
$\alpha_2$ (mV)	8.13	$9.30 \pm 0.08$
$\beta_2$	24.1	$23.7 \pm 0.7$
$\gamma_2$ (mV)	0.0	$0.33 \pm 0.04$

Table 4.1: The fitted parameters that result from a typical calibration. Approximate specifications for the circuit in Fig. 4.7 are  $R_{\text{FG}} = 50 \Omega$ ,  $R_{\text{in}} = 486 \Omega$ ,  $I_{\text{leak}} = 4 \mu\text{A}$ ,  $A = 5$ , and  $V_{\text{off}} = 9.7 \text{ mV}$ . The expressions in Eq. (4.7) are used to calculate initial guesses for the parameters in Eq. (4.6). Ten data points are used for each linear fit, and the vertical sensitivity on the DSO is  $166 \text{ mV/div}$  ( $1.66 \text{ mV/LSB}$ ). The dependency between the fitted values of  $\alpha_1$  and  $\beta_1$  is 90% [29]. The dependencies between the remaining parameters are lower than 60%. For the nonlinear fits to Eq. (4.6) we obtained a  $\chi^2$  of 6.7 for  $f_1$  and 13.8 for  $f_2$ . This corresponds to a goodness-of-fit of 75% and 13%, respectively [28].

The sign in Eq. (4.8) depends in a rather complicated way on the sizes and signs of the voltages and parameters in Eq. (4.9). The correct sign can be found by checking which solution solves Eq. (4.5).

When  $g$  is known, the current  $I$  can be found from the model in Fig. 4.7 as

$$I = \frac{V_{\text{FG}} + R_{\text{in}} I_{\text{leak}}}{R_{\text{FG}} + (gG_0)^{-1} + R_{\text{in}}}.$$

This expression can be written in terms of the known voltages and calibration parameters in Eq. (4.6) by using Eqs. (4.7a-c)

$$I = \frac{V_{\text{FG}} + \alpha_2/\alpha_1}{\beta_1^{-1} + g^{-1}} G_0. \quad (4.10)$$

According to Eq. (4.7c) we are free to use either  $\beta_1$  or  $\beta_2$  in Eq. (4.10). We use  $\beta_1$  due to its very small uncertainty, cf. Table 4.1. We have checked the validity of Eq. (4.10) by inserting an ammeter and measuring the current through calibrated resistors under DC conditions (its inner resistance is absorbed in  $R_{\text{FG}}$  during calibration).

Finally, using Ohms law, the voltage drop across the QPC can be calculated

from Eqs. (4.8) and (4.10)

$$V = \frac{I}{gG_0}. \quad (4.11)$$

## 4.7 Timing the burst

In our experiments, we only measure  $V_{\text{out}}(t)$ . To calculate the conductance of and current through a QPC, we also need to know  $V_{\text{FG}}(t)$  according to Eqs. (4.8), (4.9) and (4.10). This is no problem during the calibration procedure since  $V_{\text{FG}}$  here has a constant, known value (the offset voltage). When we work with voltage bursts, this is no longer the case. Measuring  $V_{\text{FG}}(t)$  directly with another input channel on the DSO is impossible since  $V_{\text{FG}}$  is the voltage to which the FG is internally set, cf. Fig. 4.7. The voltage that can be measured at the output terminal is different due to the voltage drop in  $R_{\text{FG}}$ . Furthermore,  $R_{\text{FG}}$  cannot be found from the calibration parameters in Eq. (4.7) (four equations with five unknowns).

Instead, we *assume* that the FG sends out an ideal signal with the requested properties (frequency, amplitude, etc.) and calculate  $V_{\text{FG}}(t)$ . Later, we will see that this is a very good assumption for our FG provided the burst frequency is no larger than 100 kHz. However, the proper synchronization of the calculated FG signal  $V_{\text{FG}}(t)$  with  $V_{\text{out}}(t)$  still remains. This is best illustrated with an example.

In Fig. 4.8 is shown a measured voltage trace of  $V_{\text{out}}(t)$  when a triangular voltage burst is applied to a calibrated resistor. This particular voltage trace will be used as an example in this section and in the following section. At a given time the operator gives the DSO a trigger signal corresponding to  $t = 0$  on the time axis. In response to this, the DSO sends out a trigger signal to the FG (Fig. 4.1). This happens within 150 ns. Upon receipt of the trigger signal, the FG spends another  $1\,300 \pm 25$  ns before the burst is sent out (the burst latency). Adding to this a propagation time of  $\approx 50$  ns through the cables, the burst signal will be delayed by up to  $1.5\ \mu\text{s}$  as compared to the DSO trigger time.

This can be seen in Fig. 4.8, where a  $1\text{--}2\ \mu\text{s}$  delay occurs between the DSO trigger time and the burst onset. When calculating  $V_{\text{FG}}(t)$ , we have to take the burst delay very carefully into account.

In Fig. 4.9 we demonstrate how sensitive the calculated conductance traces are to the burst delay, which is used when calculating  $V_{\text{FG}}(t)$ . The calculations are based on the voltage trace in Fig. 4.8. Since this measurement is performed on a calibrated resistor, the conductance should ideally be independent of the

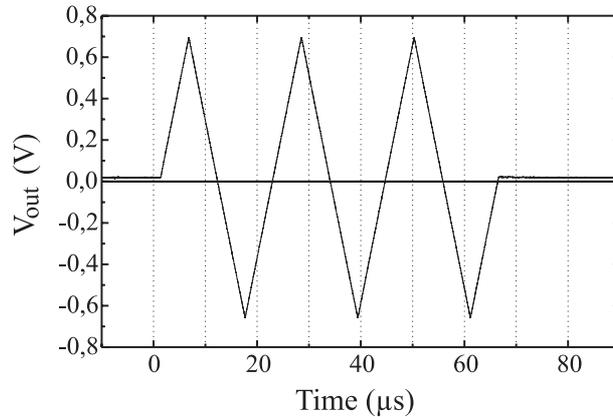


Figure 4.8: The measured output voltage  $V_{\text{out}}(t)$  when a three cycle, 46 kHz triangular burst with an amplitude of 1.15 V, superimposed on a 0.03 V offset voltage is applied to a resistor with a conductance of  $g = 3.5373$ . Setup specifications are identical to those given in Table 4.1. The voltage trace consists of 10k data points (100 Msamples/s). The DSO is triggered at  $t = 0$ .

applied voltage, and thus on the time. This is basically what we observe in Fig. 4.9(b), where a burst delay  $t_{bd} = 1349$  ns is used: Before the onset of the burst, only a small fraction of the dynamic voltage range on the DSO is used (see Fig. 4.8), resulting in noise in the calculated conductance value. During the burst, the noise decreases with increasing signal magnitude. When the burst voltage changes sign (a crossing), the signal is lost, and the noise level increases dramatically. So, although the noise level changes during the burst, the average value of the conductance is almost constant and close to the calibrated value (3.5373).

When the timing between  $V_{\text{FG}}(t)$  and  $V_{\text{out}}(t)$  is changed by only 30 ns (2% of  $t_{bd}$ ), the conductance becomes very asymmetric around the zero voltage crossings as illustrated in Fig. 4.9(a) and (c). Since there is a jitter of 50 ns in the FG burst latency, it is necessary to adjust the burst delay of every single voltage trace. Normally, we have several hundred traces in a single data set, so it is quite time-consuming to make these adjustments manually using trial and error and visual inspection.

To ease the delay adjustments, we have developed an automatic procedure for finding the optimum delay. The basic idea is to construct a suitable figure-

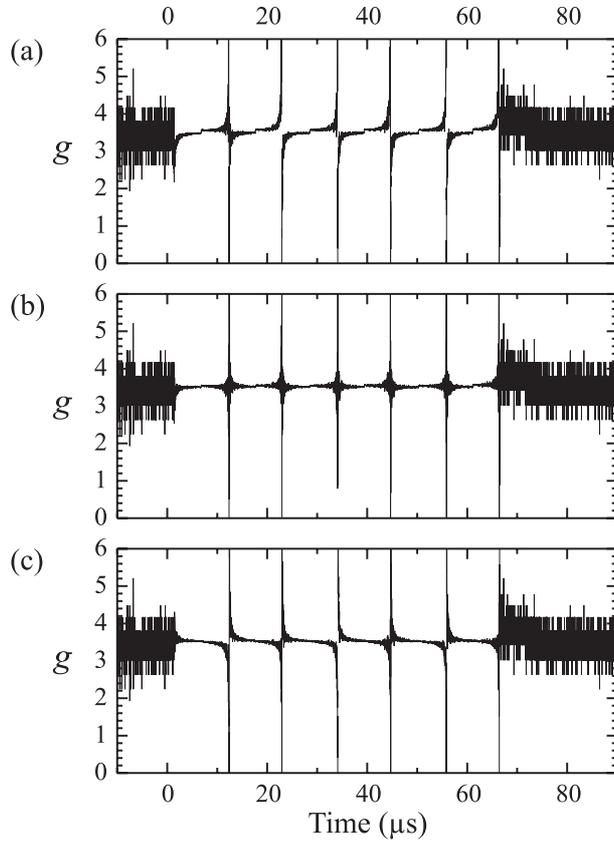


Figure 4.9: Calculated time evolution of the conductance  $g(t)$  of a calibrated resistor ( $g = 3.5373$ ) for three different delays of the FG burst signal [(b) 1 349 ns, (a) 30 ns shorter, and (c) 30 ns longer]. The conductance is calculated using Eqs. (4.8) and (4.9). To calculate the time evolution of the coefficients in Eq. (4.9), we use the measured voltage trace in Fig. 4.8, the calibration parameters in Table 4.1 and  $V_{FG}(t)$ . The latter is calculated based on the burst parameters in Fig. 4.8 and the chosen burst delay.

of-merit  $F_{\text{asym}}(t_{bd})$ , which increases with increasing asymmetry of  $g$  around the crossings. The optimum delay  $t_{bd}^0$  is then found by minimizing  $F_{\text{asym}}(t_{bd})$  with respect to  $t_{bd}$ .

To construct  $F_{\text{asym}}$ , we calculate the total conductance at a crossing  $g_c$  from the slope of the calculated  $I$ - $V$  curve in the vicinity of the crossing. In addition, we find the crossing time  $t_c$ , where  $V_{\text{FG}}(t_c) = 0$ . Now, the sum

$$\sum_i (g_i - g_c) (t_i - t_c)$$

over all data points  $(g_i, t_i)$  near the crossing is a measure of the asymmetry around the crossing. It is negative (positive) when the delay is too short (long), cf. Fig. 4.9(a) and (c). If we improve the statistics by extending the summation to all crossings and square the result, we obtain a figure-of-merit which increases with increasing asymmetry. We therefore define

$$F_{\text{asym}}(t_{bd}) \equiv \ln \left( \left\{ \sum_{c,i} [g_i(t_{bd}) - g_c] [t_i - t_c(t_{bd})] \right\}^2 \right), \quad (4.12)$$

where the logarithm is included because the square of the sum tends to vary over many orders of magnitude. Usually, we include 100 data points around each crossing when evaluating  $F_{\text{asym}}$ .

A plot of  $F_{\text{asym}}(t_{bd})$  is shown in Fig. 4.10 for the voltage trace in Fig. 4.8. It has one sharp minimum at  $t_{bd}^0 = 1\,349$  ns, which is identical to the delay used for calculating the conductance trace in Fig. 4.9(b).

The scheme presented here for finding the optimum burst delay is very robust when the conductance is constant, as is the case when we insert calibrated resistors in the circuit. However, it also works quite well for measurements on QPCs with rapidly fluctuating values of  $g(t)$ . For about 80% of these traces, the optimal delay can be found automatically. For the remaining traces, very narrow and large glitches appear in  $F_{\text{asym}}$  whereby a wrong burst delay is found. Such a (small) glitch can actually be seen at a burst delay around 1 340 ns in Fig. 4.10.

## 4.8 Performance on resistors

The performance of the setup after calibration and optimization of the burst delay can conveniently be tested using calibrated resistors. In Fig. 4.11(a) we

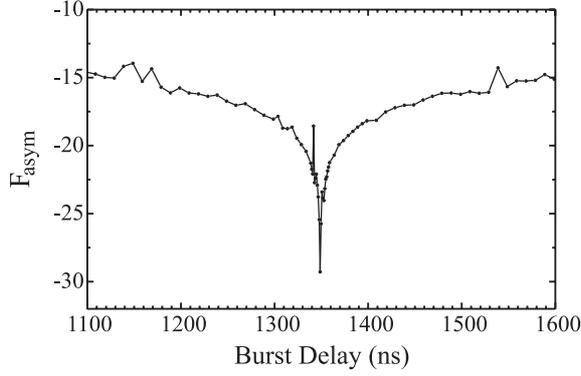


Figure 4.10: Plot of the asymmetric figure-of-merit  $F_{\text{asym}}$  [Eq. (4.12)] as a function of burst delay for the voltage trace in Fig. 4.8. 100 data points were used around each crossing. The optimum burst delay is  $t_{bd}^0 = 1349$  ns.

show the final, calculated  $I$ - $V$  curve for a  $g = 3.5373$  resistor based on the voltage trace in Fig. 4.8. Note, that although the voltage burst consists of three cycles, all the curves collapse onto one straight line. The acquisition time is  $65 \mu\text{s}$ , cf. Fig. 4.8.

To check for nonlinearities in the  $I$ - $V$  curve, we fit the calculated curve to the following third order polynomial in  $V$

$$I(V) = I_0 + g^{(0)} G_0 V + \frac{1}{2} g^{(1)} G_0 V^2 + \frac{1}{6} g^{(2)} G_0 V^3, \quad (4.13)$$

where  $I_0$  is a constant offset current. The other terms are related to the voltage-dependent *differential* conductance measured in quantum units  $g_d(V)$ , which is given by

$$\begin{aligned} g_d(V) &\equiv \frac{1}{G_0} \frac{\partial I}{\partial V} \\ &= g^{(0)} + g^{(1)} V + \frac{1}{2} g^{(2)} V^2. \end{aligned} \quad (4.14)$$

From this equation, it is clear that  $g^{(0)}$  is the differential (and total) conductance at zero bias measured in quantum units, whereas  $g^{(1)}$  and  $g^{(2)}$  are the first and second partial derivatives of  $g_d$  with respect to  $V$ .

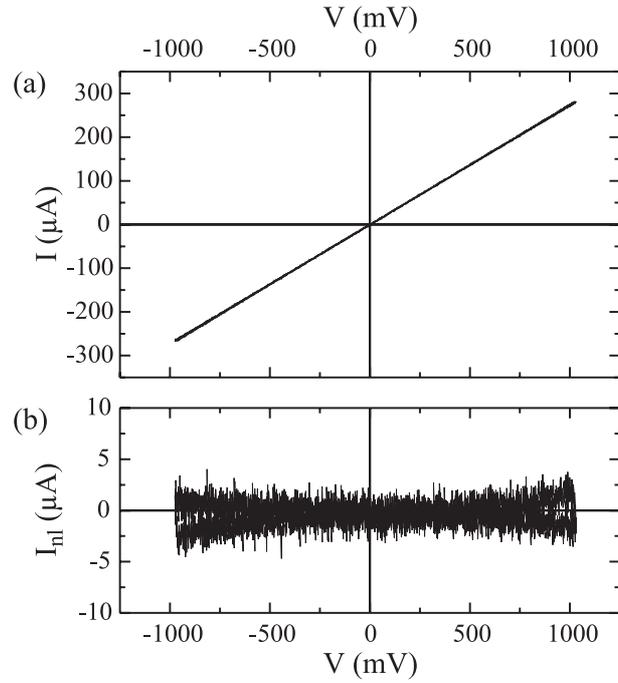


Figure 4.11: (a)  $I$ - $V$  curve of a calibrated resistor with conductance  $g = 3.5373$  calculated from Eqs. (4.10) and (4.11) using the burst part of the measured voltage trace in Fig. 4.8, the calibration parameters in Table 4.1, and the optimized burst delay  $t_{bd}^0$  found in Fig. 4.10. When the  $I$ - $V$  curve is fitted to the third order polynomial in Eq. (4.13), we get  $g^{(0)} = 3.5252 \pm 0.0008$ . (b) The residual part of the calculated  $I$ - $V$  curve after subtracting the fitted linear term.

$I_0$ (nA)	$-330\pm 30$
$g^{(0)}$	$3.5252\pm 0.0008$
$g^{(1)}$ ( $V^{-1}$ )	$-0.0002\pm 0.0002$
$g^{(2)}$ ( $V^{-2}$ )	$0.030\pm 0.008$

Table 4.2: Parameters for third-order polynomial fit [Eq. (4.13)] to  $I$ - $V$  curves on a  $g = 3.5373$  resistor (Fig. 4.11). The uncertainties include only random measurement errors. The dependency between  $g^{(0)}$  and  $g^{(2)}$  is 92% and  $-72\%$  between  $I_0$  and  $g^{(1)}$  [29].

The fitting parameters for the  $I$ - $V$  curve in Fig. 4.11(a) are given in Table 4.2. The dominant linear term  $g^{(0)}$  is 0.3–0.4% below the calibrated value. The residual current signal after subtracting the linear term from Fig. 4.11(a) is shown in Fig. 4.11(b). It is dominated by noise. On close inspection, the residual current tends to split into two signals at large magnitudes of the voltage. This originates from the different voltage sweeps, which do not overlap completely. In this case a rising (falling) voltage flank gives a curve with a small negative (positive)  $g^{(1)}$ . This is probably due to a not 100% correct burst delay. However, if data from an integer number of full burst cycles are included in the fit (as is the case here), the curvature tends to average to zero, as can be seen in Table 4.2, where  $g^{(1)}$  is very small and zero within the uncertainty of the fit.

The voltage trace in Fig. 4.8 was acquired several hours after the setup was calibrated, after more than 200  $I$ - $V$  curves on Au QPCs had been acquired. Thus, a small current offset error due to thermal drift is to be expected. This is confirmed in Table 4.2, where a small statistically significant  $I_0$  of about  $-300$  nA is found from the fits. Finally, the third order term  $g^{(2)}$  has a small, but statistically significant, positive value of about  $0.03$   $V^{-2}$ . This is not surprising since we have assumed that (i) the gain of the setup is frequency-independent; (ii) the calculated FG signal is an ideal signal; and (iii) the burst delay is correct. These assumptions are not strictly valid and will inevitably lead to small, systematic nonlinearities in the  $I$ - $V$  curves.

For the  $I$ - $V$  curve in Fig. 4.11(a), the maximum nonlinear contribution to the  $I$ - $V$  curve occurs around  $-1.0$  V [since  $g^{(1)}$  is negative]. Here, the second and third order terms contribute less than 0.2% to the total current. If we decrease the burst amplitude, the nonlinear constants  $g^{(1)}$  and  $g^{(2)}$  tend to have larger magnitudes. However, the maximum nonlinear contribution to the current is kept at the same low level since these terms depend on higher orders in  $V$ .

Tests on other calibrated resistors with conductances between 0.5 and 400  $G_0$  have also been performed using both sinusoidal and triangular bursts with frequencies in the range 1–100 kHz. In all cases, we have found that (i) the linear term  $g^{(0)}$  can be determined with an accuracy better than 1% (normally  $\approx 0.5\%$ ); and (ii) the nonlinear contributions from  $g^{(1)}$  and  $g^{(2)}$  to the  $I$ - $V$  curves are smaller than 1% (normally  $\approx 0.3\%$ ). We have also checked how the fitting parameters depend on the burst delay. As long as  $t_{bd}$  is changed by only a few percent [like in Fig. 4.9(a), where it is clear that the delay is inadequate] the change in the parameters is negligible.

## 4.9 Gold quantum point contacts

We finally demonstrate how the setup can be successfully applied to measuring  $I$ - $V$  curves on Au QPCs formed between an STM Au tip and a Au(111) surface in UHV at RT. The Au(111) surface is cleaned by repeated sputtering with 1.5 keV Ar ions at  $45^\circ$  incidence followed by annealing to 550 °C. The Au tip is made by cutting a 0.25 mm diameter, 99.99% pure Au wire. The tip is ‘cleaned’ prior to the experiments by indenting it repeatedly into the surface to a depth of 500 Å.

The topography of the Au(111) surface is imaged in constant current mode using a conventional set of STM electronics. Images displaying atomic resolution and the so-called ‘herringbone’ reconstruction [30] are shown in Fig. 4.12. The images are clean on the atomic scale, and we take this as an indication of a well-prepared surface. After the initial STM inspection, the tip is retracted a few  $\mu\text{m}$ , and the fast electronics for acquiring fast  $I$ - $V$  curves (Fig. 4.1) is connected to the STM.

By means of the inchworm motor, the tip is approached to the surface until the conductance increases above 1  $G_0$ . The computer-controlled feedback loop is then enabled wherein  $V_{\text{out}}$  is acquired repeatedly from the ADC (cf. Fig. 4.1), and the corresponding conductance is calculated using Eq. (4.8). In this manner it is quite easy to stabilize the Au contact at a conductance larger than 10  $G_0$ . However, smaller contacts tend to be quite unstable, and we cannot stabilize them using the computer controlled feedback since it has a rather long response time of  $\approx 10$  ms.

Quite often we observe that contacts with a conductance smaller than 10  $G_0$  tend to break spontaneously. This is probably due to a spontaneous ‘diffusion out of contact’. External vibrations do not seem to cause the breaks since we have been able to acquire images with atomic resolution [see Fig. 4.12(a)].

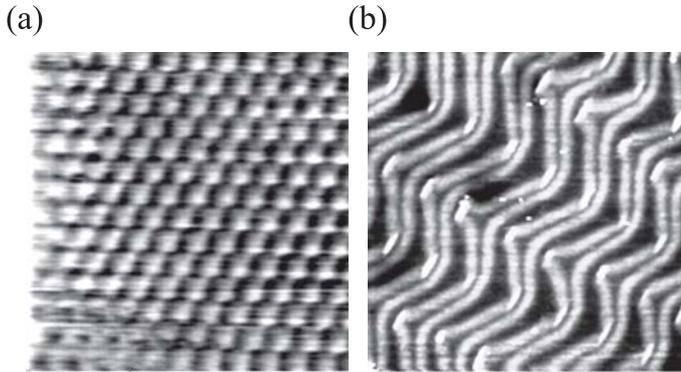


Figure 4.12: STM images of the clean Au(111) surface using a Au tip. (a) The atomically resolved surface ( $30 \times 30 \text{ \AA}^2$ ). (b) The ‘herringbone’ reconstruction ( $500 \times 500 \text{ \AA}^2$ ) [30].

The DSO is triggered by events in which the conductance goes below  $\approx 3 G_0$  such that a voltage burst is initiated immediately before the contact is broken. Hereby it is possible to study the  $I$ - $V$  curves of several metastable configurations in the contact during the final stages of the breaking process. An example of this is given in Fig. 4.13, where a three-cycle, 10 kHz triangular burst with an amplitude of 1.1 V, superimposed on a 40 mV offset voltage, is applied to a Au QPC with an initial value of the conductance around  $3 G_0$  [see inset in Fig. 4.13(a), only the first two cycles are shown].

The calculated time evolution of the total conductance is shown in Fig. 4.13(a). The conductance trace is composed of four very well defined plateaux close to  $3, 2, 1,$  and  $0 G_0$ . The noise in the signal is large before the onset of the burst and diverges at the crossings of  $V_{\text{FG}}$  as discussed in relation to Fig. 4.9. The conductance on a plateau is almost independent of the bias voltage, and the transition between plateaux occurs very quickly. Glitches in the conductance occur immediately before B and  $10\text{--}15 \mu\text{s}$  before E. An intermediate plateau around  $1.7 G_0$  is observed for a very short period of time ( $0.6 \mu\text{s}$ ) at D. The overall behavior of the trace in Fig. 4.13(a) is quite similar to what is usually observed in Au QPCs under constant bias voltage conditions at RT [3, 8, 9, 14, 31–36] (see also Fig. 3.3 on page 51 in Chapter 3).

In Fig. 4.13(b), we show the calculated  $I$ - $V$  curve for the conductance trace in Fig. 4.13(a). It consists of segments of almost straight lines through the

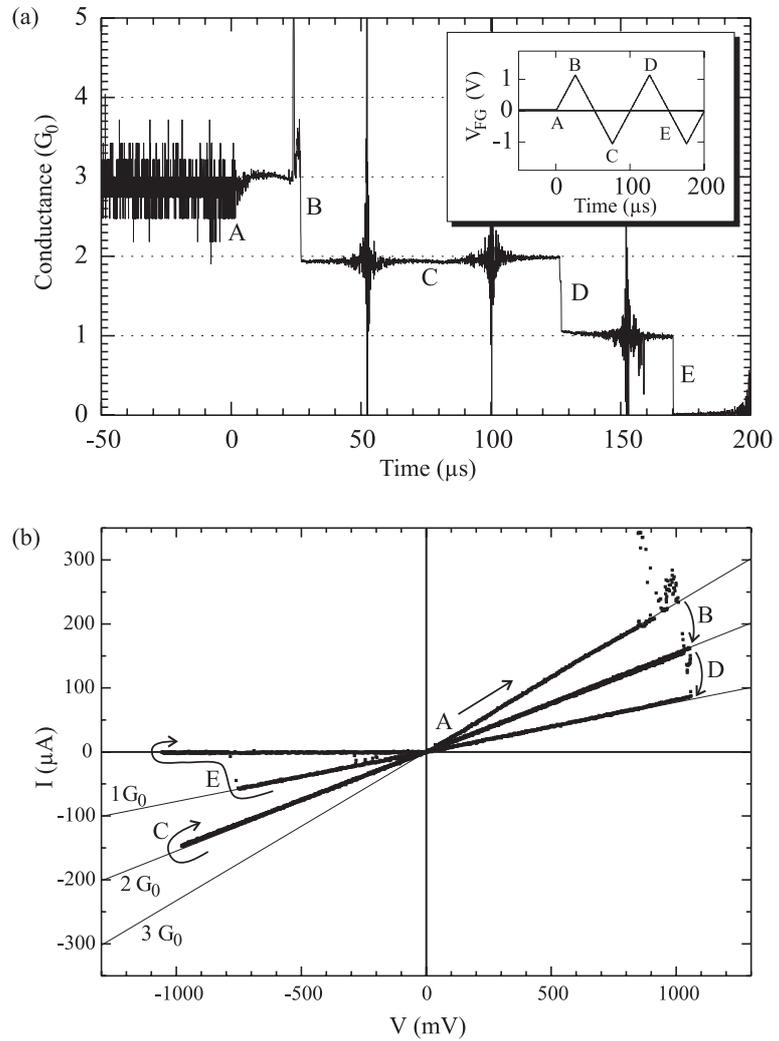


Figure 4.13: (a) Calculated total conductance on a breaking Au QPC under an applied voltage burst (inset). (b) The corresponding calculated  $I$ - $V$  curve. Lines corresponding to integer values of the conductance quantum have been inserted to guide the eye. There are 5 000 data points in the plot (50 ns/point).

origin. The segments closely follow lines corresponding to quantized and voltage-independent values of the conductance. This is at variance with all previous  $I$ - $V$  curves recorded on *stable* Au contacts at RT [8–12]. The origin of these apparent discrepancies will be discussed in the next two Chapters.

## 4.10 Summary

We have constructed a setup which is capable of acquiring accurate  $I$ - $V$  curves on metallic QPCs during a period which may be as short as 10  $\mu$ s. Tests on resistors show that nonlinear contributions to the  $I$ - $V$  curves are typically smaller than 0.3% (always better than 1%). From the slope of the  $I$ - $V$  curve, the conductance can be found with an accuracy of typically 0.5% (always better than 1%). The time resolution of the setup is typically 25–100 ns. Initial tests on metastable Au QPCs at RT reveal almost Ohmic  $I$ - $V$  curves with nonlinearities much smaller than reported for stable contacts [8–12].

## References

- [1] L. Olesen, E. Lægsgaard, I. Stensgaard, F. Besenbacher, J. Schiøtz, P. Stoltze, K. W. Jacobsen, and J. K. Nørskov, “Quantized Conductance in an Atom-Sized Point Contact,” *Phys. Rev. Lett.* **72**, 2251–2254 (1994).
- [2] L. Olesen, E. Lægsgaard, I. Stensgaard, F. Besenbacher, J. Schiøtz, P. Stoltze, K. W. Jacobsen, and J. K. Nørskov, “Reply to ‘Comment on ‘Quantized conductance in an Atom-Sized Point Contact’’,” *Phys. Rev. Lett.* **74**, 2147 (1995).
- [3] M. Brandbyge, J. Schiøtz, M. R. Sørensen, P. Stoltze, K. W. Jacobsen, J. K. Nørskov, L. Olesen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, “Quantized conductance in atom-sized wires between two metals,” *Phys. Rev. B* **52**, 8499–8514 (1995).
- [4] L. Olesen, “Experimental Investigation of the Junction in a Scanning Tunneling Microscope,” PhD thesis, University of Aarhus, 1996.
- [5] E. Scheer, P. Joyez, D. Esteve, C. Urbina, and M. H. Devoret, “Conduction Channel Transmissions of Atomic-Size Aluminum Contacts,” *Phys. Rev. Lett.* **78**, 3535–3538 (1997).
- [6] E. Scheer, N. Agraït, J. C. Cuevas, A. L. Yeyati, B. Ludoph, A. Martín-Rodero, G. R. Bollinger, J. M. van Ruitenbeek, and C. Urbina, “The signature of chemical valence in the electrical conduction through a single-atom contact,” *Nature* **394**, 154–157 (1998).
- [7] B. Ludoph, M. H. Devoret, D. Esteve, C. Urbina, and J. M. van Ruitenbeek, “Evidence for Saturation of Channel Transmission from Conductance Fluctuations in Atomic-Size Point Contacts,” *Phys. Rev. Lett.* **82**, 1530–1533 (1999).
- [8] J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. M. Baró, N. García, U. Landman, W. D. Luedtke, E. N. Bogachek, and H.-P. Cheng, “Properties of Metallic Nanowires: From Conductance Quantization to Localization,” *Science* **267**, 1793–1795 (1995).
- [9] J. L. Costa-Krämer, N. García, P. García-Mochales, P. A. Serena, M. I. Marqués, and A. Correia, “Conductance quantization in nanowires formed between micro and macroscopic metallic electrodes,” *Phys. Rev. B* **55**, 5416–5424 (1997).

- [10] J. Abellán, R. Chicón, and A. Arenas, “Properties of nanowires in air. Controlled values of the conductance,” *Surf. Sci.* **418**, 493–501 (1998).
- [11] B. Rodell, V. Korenivski, J. Costa, and K. V. Rao, “Quantized conductivity in STM drawn 100Å nanowires of Au and Ni,” *Nanostruct. Mater.* **7**, 229–235 (1996).
- [12] A. Correia, J.-L. Costa-Krämer, Y. W. Zhao, and N. García, “Non-linear contribution to intensity-voltage characteristics of gold nanowires,” *Nanostruct. Mater.* **12**, 1015–1020 (1999).
- [13] H. Ohnishi, Y. Kondo, and K. Takayanagi, “Quantized conductance through individual rows of suspended gold atoms,” *Nature* **395**, 780–783 (1998).
- [14] G. Rubio, N. Agraït, and S. Vieira, “Atomic-Sized Metallic Contacts: Mechanical Properties and Electronic Transport,” *Phys. Rev. Lett.* **76**, 2302–2305 (1996).
- [15] L. Eierdal, F. Besenbacher, E. Lægsgaard, and I. Stensgaard, “Interaction of oxygen with Ni(110) studied by scanning tunneling microscopy,” *Surf. Sci.* **312**, 31–53 (1994).
- [16] M. Ø. Pedersen, S. Helveg, A. Ruban, I. Stensgaard, E. Lægsgaard, J. K. Nørskov, and F. Besenbacher, “How a gold substrate can increase the reactivity of a Pt overlayer,” *Surf. Sci.* **426**, 395–409 (1999).
- [17] D. P. E. Smith, “Quantum Point Contact Switches,” *Science* **269**, 371–373 (1995).
- [18] F. Ott and J. Lunney, “Quantum Conduction: a Step-by-Step Guide,” *Europhys. News* **29**, 13–16 (1998).
- [19] F. Ott, S. Barberan, J. G. Lunney, J. M. D. Coey, P. Berthet, A. M. de Leon-Guevara, and A. Revcolevschi, “Quantized conductance in a contact between metallic oxide crystals,” *Phys. Rev. B* **58**, 4656–4659 (1998).
- [20] E. L. Foley, D. Candela, K. M. Martini, and M. T. Tuominen, “An undergraduate laboratory experiment on quantized conductance in nanocontacts,” *Am. J. Phys.* **67**, 389–393 (1999).

- [21] U. Landman, W. D. Luedtke, B. E. Salisbury, and R. L. Whetten, “Reversible Manipulations of Room Temperature Mechanical and Quantum Transport Properties in Nanowire Junctions,” *Phys. Rev. Lett.* **77**, 1362–1365 (1996).
- [22] H. Yasuda and A. Sakai, “Conductance of atomic-scale gold contacts under high-bias voltages,” *Phys. Rev. B* **56**, 1069–1072 (1997).
- [23] W. A. de Heer, S. Frank, and D. Ugarte, “Fractional quantum conductance in Gold nanowires,” *Z. Phys. B* **104**, 469–473 (1997).
- [24] H. Oshima and K. Miyano, “Spin-dependent conductance quantization in nickel point contacts,” *Appl. Phys. Lett.* **73**, 2203–2205 (1998).
- [25] U. Dürig, L. Novotny, and B. Michel, “Logarithmic current-to-voltage converter for local probe microscopy,” *Rev. Sci. Instrum.* **68**, 3814–3816 (1997).
- [26] Data sheet for the OP07, Analog Devices operational amplifier. (Can be downloaded from <http://www.analog.com/pdf/op07.pdf> in pdf format.)
- [27] Data sheet for the OPA651, Burr-Brown operational amplifier. (Can be downloaded from <http://www.burr-brown.com/download/DataSheets/OPA651.pdf> in pdf format.) The OPA651 is stable in gains  $\geq 2V/V$ , has a gain bandwidth product of 340 MHz, a 12-bit settling time in 16 ns, an input bias current of 4  $\mu A$  and an input voltage noise of 4.6 nV/ $\sqrt{Hz}$  above 10 kHz.
- [28] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal* (Cambridge University Press, Cambridge, 1989).
- [29] The covariance matrix  $C_{ij}$  for the fitted parameters can be estimated from the fitting procedure [28]. The uncertainty of a parameter  $i$  is given by  $\sqrt{C_{ii}}$ . The dependency between two parameters  $i$  and  $j$  is given by  $C_{ij}/\sqrt{C_{ii}C_{jj}}$ .
- [30] J. V. Barth, H. Brune, G. Ertl, and R. J. Behm, “Scanning tunneling microscopy observations on the reconstructed Au(111) surface: Atomic structure, long-range superstructure, rotational domains, and surface defects,” *Phys. Rev. B* **42**, 9307–9318 (1990).
- [31] J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. M. Baró, N. García, and V. T. Binh, “Quantum Contact in Gold Nanostructures by Scanning Tunneling Microscopy,” *Phys. Rev. Lett.* **71**, 1852–1855 (1993).

- [32] J. L. Costa-Krämer, N. García, P. García-Mochales, and P. A. Serena, “Nanowire formation in macroscopic metallic contacts: quantum mechanical conductance tapping a table top,” *Surf. Sci.* **342**, L1144–L1149 (1995).
- [33] C. J. Muller, J. M. Krans, T. N. Todorov, and M. A. Reed, “Quantization effects in the conductance of metallic contacts at room temperature,” *Phys. Rev. B* **53**, 1022–1025 (1996).
- [34] Z. Gai, Y. He, H. Yu, and W. S. Yang, “Observation of conductance quantization of ballistic metallic point contacts at room temperature,” *Phys. Rev. B* **53**, 1042–1045 (1996).
- [35] N. García and J. L. Costa-Krämer, “Quantum-Level Phenomena in Nanowires,” *Europhys. News* **27**, 89–91 (1996).
- [36] W. B. Jian, C. S. Chang, W. Y. Li, and T. T. Tsong, “Geometrical dependence of conductance quantization in metal point contacts,” *Phys. Rev. B* **59**, 3168–3172 (1999).



## Chapter 5

# Current-Voltage Curves of Gold Quantum Point Contacts

Ohm's law  $I = GV$  is known to hold for many everyday conductors such as macroscopic wires and carbon resistors. The linear relationship between the current and the voltage is, however, not of universal validity as shown in Fig. 5.1.

A seemingly trivial example of this is the light bulb: At low applied voltages the conductance is high because the filament is cold. With increasing voltage, the filament heats up, whereby the conductivity (and hence the conductance)

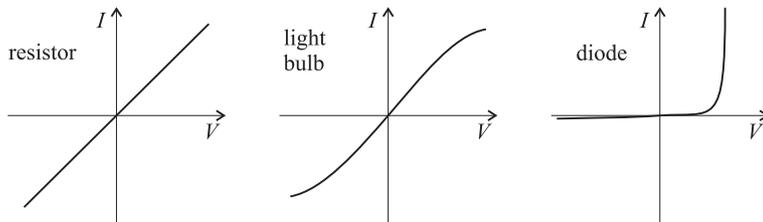


Figure 5.1: Schematic current-voltage curves of some common electrical components.

decreases. This leads to a rather nonlinear current-voltage (*I-V*) curve, where the current tends to saturate at voltages where the filament begins to glow. But Ohm's law is still obeyed for a light bulb in the sense that the explicit dependence between the current and the voltage is linear:  $I = G(T)V$ . The nonlinearities arise because the filament temperature  $T$  depends on the voltage  $V$ . If the temperature of the light bulb could be kept constant (for example by sweeping the voltage very fast), the *I-V* curve would be linear.

A real breakdown of Ohm's law is found in many electronic components. A familiar example is the diode, which obeys the highly nonlinear current-voltage relation  $I \propto [\exp(\beta eV) - 1]$ , where  $\beta \sim (k_B T)^{-1}$  [1]. This relation implies that a forward biased diode is a very good conductor. On the other hand, a reversely biased diode basically blocks the current.

The basic differences between a diode and a resistor (or light bulb) are; (i) the magnitudes of the electric fields are much larger in a diode because it is concentrated in the thin depletion layer of the *p-n* junction, and (ii) the *p-n* junction is inhomogeneous and breaks inversion symmetry. The latter results in a polarity dependence of the current (rectification).

In the previous Chapter on page 108 in Fig. 4.13 we showed that voltages up to 1 V can be applied to a Au quantum point contact (QPC) at room temperature (RT). The QPC conductance in Fig. 4.13(a) goes through a series of transient states with a conductance close to 3, 2, and 1  $G_0$ , meaning that only a few atoms are found in its smallest cross-section. Thus, the voltage drop is concentrated in a nanometer-sized region around the QPC (see footnote 7 on page 14 in Chapter 2 and Refs. 2, 3). As a consequence, the field strength reach values of the order V/nm, i.e., comparable to the field in the vacuum gap of a scanning tunneling microscope (STM), and much larger than the field in a *p-n* junction. Unlike a diode, a QPC formed between similar electrodes is not expected to exhibit much rectification due to its homogeneous composition. There may, however, still be a weak polarity dependence due to dissimilarities between the contact geometry on each side of the constriction.

Despite this, the *I-V* curve shown in Fig. 4.13(b) is almost linear (by the piece). As discussed briefly on page 109, this linear behavior is in contrast to all previous measurements on *I-V* curves acquired from Au contacts at RT. To further elucidate this apparent puzzle, we have performed more systematic studies of Au *I-V* curves under different experimental conditions. The outcome of this work is presented in this Chapter and forms the basis of papers [IV] and [V].

Our experiments confirm that *I-V* curves on (clean) Au QPCs are indeed (almost) linear. Self-consistent *spd* tight-binding calculations on one- and two-

atom contacts (performed by Mads Brandbyge) are shown to be in excellent agreement with the experiments. Only when the sample has not been sputtered and annealed do we observe stable contacts, which exhibit clear nonlinearities for voltages larger than 0.1–0.5 V. We suggest that the nonlinear curves in our experiments are due to tunneling through a thin contaminant layer of organic molecules in the tip-sample interface. A detailed discussion of these curves are deferred to Chapter 6.

## 5.1 Motivation

For QPCs, the main focus has so far been on the low bias voltage ( $V < 0.25$  Volt) regime of the conductance. For high bias new important questions concerning nonlinear conductance [2, 4–14], electromigration [15–19] and heating [20] arise. At the present stage, however, the available experimental results [4–10, 17–19] on finite bias conduction in Au QPCs are contradictory, and the theoretical interpretation of the results is unclear [2, 8, 12–14].

Sakai and co-workers [17–19] studied conductance histograms acquired from transient Au contacts in a relay and found that the  $1 G_0$  peak did not change its position as the bias voltage was gradually increased to 2 V. This observation suggests that the conductance of monatomic Au contacts (which have a conductance close to  $G_0$  [21–23]) is almost independent of voltage. This is in contrast to several reports [4–10] of  $I$ - $V$  characteristics measured on stable Au contacts at room temperature (RT), which display clear nonlinearities for voltages larger than 0.1–0.5 V.

## 5.2 Experimental details

Previous measurements of  $I$ - $V$  curves on Au contacts [4–10] were done over a period of 0.1–50 s, thus restricting the measurements to contacts which are stable over relatively long time-scales. Instead, we use the fast setup already described in Chapter 4, whereby we decrease the acquisition time by four to seven orders of magnitude.

### Contact formation techniques

The contacts are formed between a polycrystalline Au tip mounted in a scanning tunneling microscope (STM) and a Au(110) single crystal surface under ultra-high vacuum (UHV) conditions in two different ways:

- *Hard* indentations. The tip is first indented 40–100 Å into the surface and retracted at a speed of  $\approx 100$  Å/s. When the conductance has dropped below a user defined preset, a voltage burst is automatically applied to the contact.
- *Soft* indentations. The indentation is stopped the first time the conductance surpasses a user defined preset (1–20  $G_0$ ). In a computer controlled feedback loop, the vertical position of the tip is then continuously adjusted to keep the conductance as stable as possible. When (if) stabilized, a voltage burst is applied to the contact upon request from the user.

Contacts with stable conductance values are favored by the latter contact formation technique because it relies on human response. For the hard indentations, on the other hand, there is a delay of only  $1.5 \mu\text{s}$  between the fulfilment of the trigger criterion and the onset of the burst (see the discussion beginning on page 99 in Chapter 4).

### Cleaning procedures

Besides varying the contact formation technique, we have investigated how the cleanliness of the Au(110) sample influences the measured *I-V* curves by using two different sample preparation techniques:

- A clean surface (verified by STM imaging with atomic-scale resolution, cf. Fig. 4.12) is produced by sputtering and annealing under UHV conditions as explained on page 106 in Chapter 4.
- The sample is rinsed in water and ethanol in air before being placed in a UHV chamber.

### 5.3 The stability criterion

As we saw in Fig. 4.13, the contact geometry may change during a single period of a voltage burst and go through several transient states. Such mechanical deformations complicate the interpretation of the *I-V* characteristics. For example, from the *I-V* curve in Fig. 4.13(b) it is evident that the points B, D, and E correspond to major, abrupt changes in the atomic configuration in the contact. However, as shown in several low temperature experiments with mechanically controllable break junctions (MCBJs) [23–25], the conductance of a

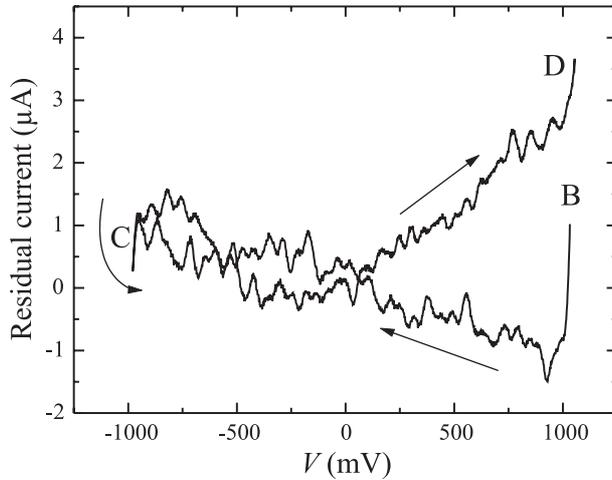


Figure 5.2: The B–C–D segment of the  $I$ - $V$  curve in Fig. 4.13(b) after subtraction of  $1.95 G_0 V$  and smoothing over 20 data points. The segment was acquired during  $22 \mu s$ . The voltage sweeps are seen to split upon approaching points B and D. Thus, the contact geometry must have changed slightly during the burst cycle and it does not meet our selection criterion.

Au QPC may also change *continuously* in between these points due to elastic deformations of the metallic bonds.

If such elastic deformations take place during a burst, ‘false’ nonlinearities emerge in the  $I$ - $V$  curve. We illustrate this point in Fig. 5.2, where we have taken the B–C–D segment of the  $I$ - $V$  curve in Fig. 4.13, subtracted an  $I$ - $V$  curve corresponding to a  $1.95 G_0$  conductor, and smoothed over 20 points to remove some of the noise. Neglecting the rapid fluctuations,<sup>1</sup> it is clearly seen that the residual current signal for the segment B–C differs from the C–D segment. Especially, points B and D do not meet. This indicates that the geometry of the contact has changed slightly during the  $22 \mu s$ , which have passed from point B to D.

In our analysis of the  $I$ - $V$  curves, we have therefore only included curves,

---

<sup>1</sup>Since these wiggles do not ‘retrace’ themselves upon going forth and back close to point C, they are probably not a signature of the kinds of conductance fluctuations that have been observed by Ludoph and Ruitenbeek at low temperatures [26, 27]. Instead, instrumental measurement noise is probably the origin of the fluctuations.

which fulfill the following criterion: *The  $I-V$  curve should retrace itself when the voltage is swept back and forth.* This judgment is done ‘by eye’ since there is no way of clearly distinguishing random noise from systematic deviations. If we for example focus alone on the  $V < 0$  part of Fig. 5.2, we would say that there is a fair agreement between the two sweeps. However, due to the large discrepancies for  $V > 0$ , this particular curve is discarded.

### Anomalies due to instrumental artifacts

Finally, we note that the residual current curve in Fig. 5.2 is very steep close to the maxima and minima of the voltage burst. In general, we find such features in the curves independent of the amplitude of the burst and the stability of the contact. Hence, we can conclude they must be due to instrumental artifacts. Since these points in the curve are close in time to the turning points of the calculated (triangular) burst voltage, the observed anomalies may be due to the finite bandwidth of the setup, which smears such sharp features. Small errors in the burst delay will also affect the shape of the  $I-V$  curve close to the turning points due to a lack of timing between the calculated and measured turning points.

Whatever the cause might be for these anomalies, they can affect the non-linear terms that may come out of a least-squares fitting to, say, a polynomial. To avoid this, we will in the following always neglect the 80 data points that are closest to a turning point in a full burst cycle. (For Fig. 5.2 this would correspond to neglecting 20 data points at B, 40 data points at C, and 20 data points at D.)

## 5.4 Experimental results

### Current-voltage curves

Examples of  $I-V$  curves acquired under different experimental conditions are shown in Fig. 5.3. All curves are acquired at RT, except curve (b), where the sample is cooled to  $-100$  °C (the base of the tip is still close to RT). This is necessary because contacts formed at RT by soft indentations on sputtered and annealed samples tend to break spontaneously when the conductance is lower than about  $5G_0$ . Only by cooling the sample is it possible to stabilize the contact on a ‘human response’ time scale. Contacts formed by hard indentations in UHV also break spontaneously when the conductance is only a few quantum units,

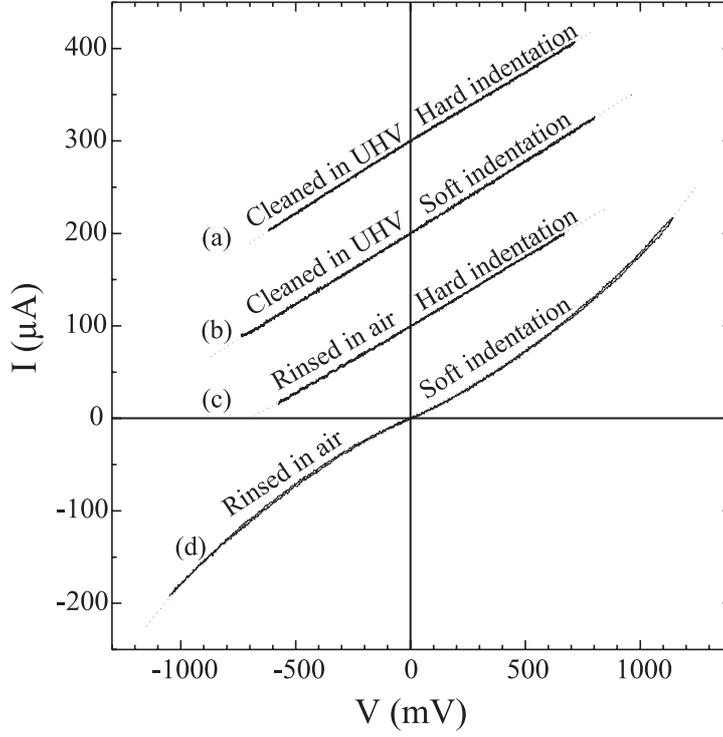


Figure 5.3: Examples of measured  $I$ - $V$  curves of Au contacts using different sample preparation/contact formation techniques. Curves (a), (c), and (d) are acquired at RT during  $20 \mu\text{s}$  (a single period of a 50 kHz triangular voltage burst) and consist of 2000 data points each. The sample temperature is  $-100 \text{ }^\circ\text{C}$  for curve (b), which is acquired during  $22 \mu\text{s}$  (a single period of a 46 kHz triangular voltage burst) and consists of 2200 data points. Best fits to Eq. (5.1) (dashed lines) give the following values of  $(G, G', G'')$  in units of  $(G_0, G_0/\text{Volt}, G_0/\text{Volt}^2)$ : (a) (1.95,  $-0.08$ , 0.07); (b) (1.99, 0.03, 0.00); (c) (1.91, 0.07,  $-0.05$ ); (d) (1.72,  $-0.02$ , 0.59). Individual  $I$ - $V$  curves are shifted by  $100 \mu\text{A}$  for clarity.

but under these circumstances the setup triggers automatically as explained previously in Sec. 5.2.

The curves in Fig. 5.3 are almost linear if the sample is sputtered and annealed *or* the contacts are formed by hard indentations [curves (a), (b), and (c)]. In contrast, nonlinear *I-V* curves are almost exclusively found when the sample is rinsed in air *and* the indentations are soft [curve (d)].

We find that contacts exhibiting nonlinear *I-V* curves are more robust than their linear counterparts: Reproducible *I-V* curves can be obtained at voltages up to  $\pm 2$  V, and the contacts can be stable for hours. In contrast, contacts corresponding to linear *I-V* curves tend to destabilize for voltages larger than  $\approx 0.75$  V (at RT) and contacts with a conductance smaller than  $5 G_0$  typically break spontaneously within milliseconds.

### Polynomial fits

Independent of the contact formation technique and the cleaning procedure, we find that the *I-V* curves can be fitted with good accuracy (see dashed lines in Fig. 5.3) to a third-order polynomial in the applied voltage [7–10]

$$I(V) = GV + G'V^2 + G''V^3, \quad (5.1)$$

where  $G$  is the low bias conductance, and  $G'$  is a constant describing the polarity dependence of the current due to asymmetries in the contact region. For most curves  $|G'| < 0.1 G/\text{Volt}$ . In the following we will focus on the nonlinear parameter  $G''$  and its dependence on  $G$ .

Fitted values for  $G$  and  $G''$  based on *I-V* curves obtained from 166 different contacts with a low bias conductance lower than  $4 G_0$  are shown in Fig. 5.4. For *I-V* curves formed in hard indentations or on a sputtered and annealed sample, 80% of the points collapse onto two well-defined regions which have a small third-order term ( $|G''| < 0.2 G_0/\text{Volt}^2$ ) and a low bias conductance close to  $1 G_0$  or slightly below  $2 G_0$ . About 80% of the *I-V* curves obtained from soft indentations on a rinsed sample have non-linear terms which exceed  $0.2 G_0/\text{Volt}^2$  and arbitrary values of  $G$ . The remaining 20% are found in the above-mentioned regions.

In Fig. 5.5 we have extended the graph in Fig. 5.4 to values of the low bias conductance up to  $12 G_0$ . For contacts formed by soft indentations on a rinsed sample,  $G''$  increases steadily and roughly linearly with  $G$  reaching a maximum value of about  $8 G_0/\text{Volt}^2$ . In contrast, contacts formed on the sputtered and annealed sample have much smaller values of  $G''$ . For  $G < 4 G_0$ , we saw in

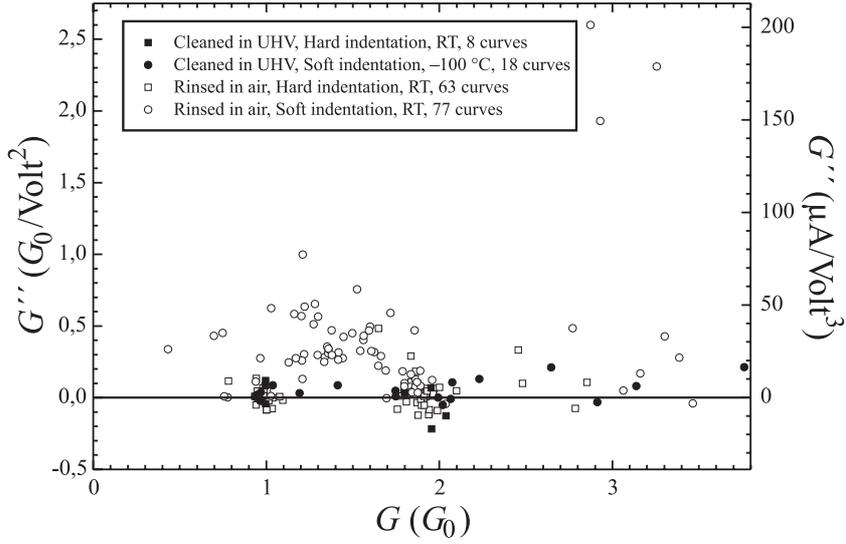


Figure 5.4: The nonlinear term  $G''$  vs. low bias conductance  $G$  obtained by fitting measured  $I$ - $V$  curves to Eq. (5.1) for different sample preparation/contact formation techniques. Only fits to nearly polarity independent  $I$ - $V$  curves ( $|G'| < 0.05 G/\text{Volt}$ ) with small uncertainties of the cubic term ( $\sigma_{G''} < 0.01 G/\text{Volt}^2$ ) are shown.

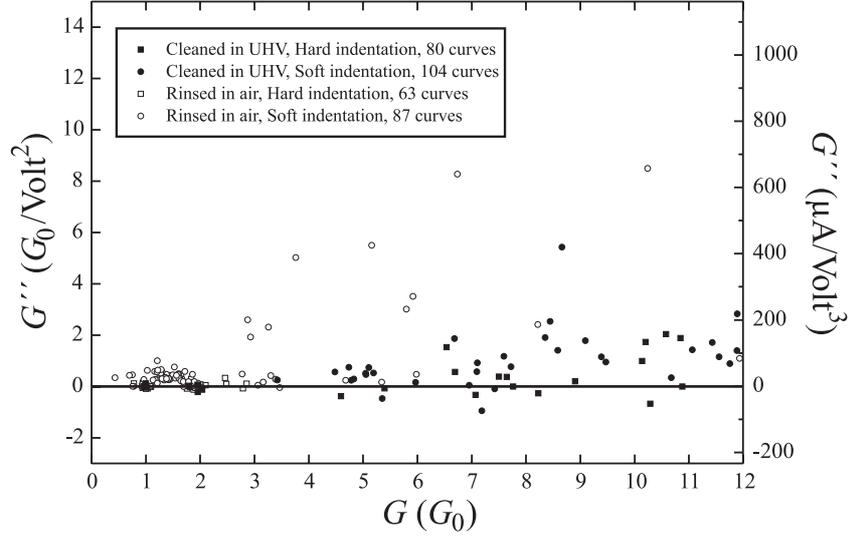


Figure 5.5: An extension of Fig. 5.4 to larger values of the low bias conductance. Instead of the low temperature data in Fig. 5.4 for soft indentations on a sample cleaned in UHV, we show the equivalent room temperature data. Thus, all data in the figure are taken at room temperature.

Fig. 5.4 that  $|G''| < 0.3 G_0/\text{Volt}^2$  for these contacts. However, as  $G$  increases towards  $12 G_0$ ,  $G''$  becomes positive for most curves with a mean value around  $1 G_0/\text{Volt}^2$ . Also, the clustering of  $G$  near integer values of  $G_0$  in Fig. 5.4 is evidently not present for larger values of  $G$ .

## 5.5 Tight-binding calculations

### Contact Geometry

To further elucidate our experimental findings for the smallest contacts, we have performed transport calculations for the two simple contact geometries on the (110) surface shown in the inset of Fig. 5.6. The contacts have either one (a) or two (b) atoms in their smallest cross-section. The geometries are basically identical to a model proposed by Ohnishi *et al.* [22] in relation to the observation

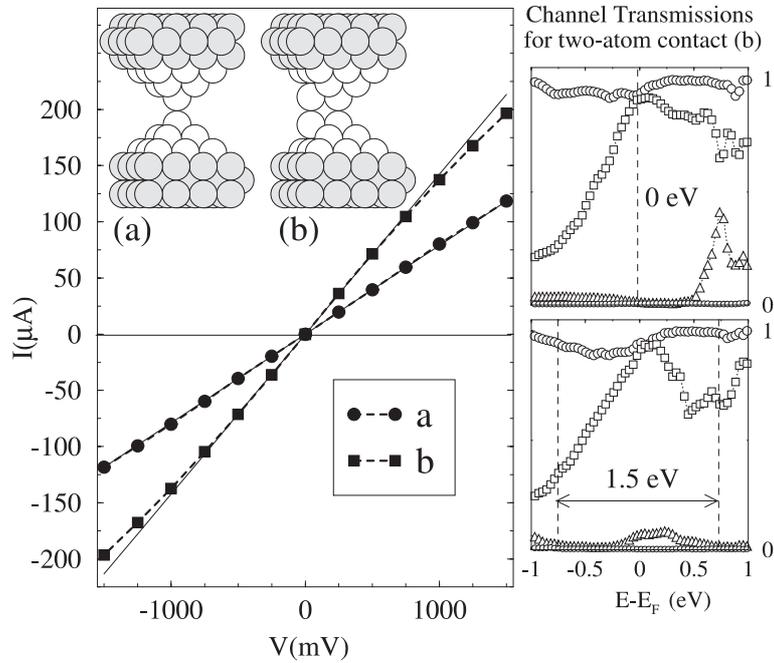


Figure 5.6: Left:  $I$ - $V$  characteristics calculated using a self-consistent tight-binding model for the one- (a), and two-atom (b) contacts shown in the insert. All interatomic distances are taken to be equal to the bulk value. Solid lines are zero-bias extrapolations corresponding to (a)  $G = 1.02 G_0$ , (b)  $G = 1.88 G_0$ . The  $I$ - $V$  curve for (b) has  $G'' = -0.07 G_0/\text{Volt}^2$ . Right: The eigenchannel transmissions  $\{T_n(E - E_F, V)\}$  for the two-atom contact (sorted by size and labeled by  $\{\circ, \square, \triangle, \circ\}$ ) are shown for  $V = 0$  Volt and  $V = 1.5$  Volt. The vertical dashed lines define the energy window [cf. Eq. (5.2)] used when calculating the current.

of single-atom Au chains in high-resolution transmission electron microscope (TEM) images. However, the chains used here are one atom shorter.

### Computational details

A non-orthogonal tight-binding (TB) model (with  $s$ ,  $p$ , and  $d$  orbitals) is used for calculating the electronic structure and current self-consistently at finite bias (see Ref. 2 for details and further references). Parameters for the TB Hamiltonians are based on fits to *ab initio* bulk band structures obtained from density functional calculations. All interatomic distances in the contact are fixed at their bulk values.

All atoms are assumed to be charge neutral. This corresponds to assuming that the electron screening length is no longer than one nearest neighbor distance. Charge neutrality is accomplished by adjusting the local potential on each atom in the contact region (white atoms in Fig. 5.6) self-consistently *also* at finite bias. The atoms in the contact region are connected to semi-infinite (110) gold electrodes (gray atoms). The current is calculated from the zero temperature limit of the Landauer-Büttiker form [Eq. (2.23)], which we encountered on page 22 in Chapter 2<sup>2</sup>

$$I = \frac{2e}{h} \int_{E_F - eV/2}^{E_F + eV/2} dE \operatorname{Tr} [\mathbf{t}^\dagger(E, V) \mathbf{t}(E, V)] , \quad (5.2)$$

where the transmission amplitude matrix  $\mathbf{t}(E, V)$  is calculated using the non-equilibrium Green-function formalism. Using the same formalism, the local density of states projected onto TB orbitals can also be found. At a fixed voltage  $V$ , the set of eigenchannel transmission probabilities  $\{T_n(E, V)\}$  are given by the eigenvalues of  $\mathbf{t}^\dagger(E, V) \mathbf{t}(E, V)$ .

### Computational results

For both contact geometries, mainly the  $6s$  valence electrons conduct. In agreement with earlier calculations for Au contacts between (100) and (111) surfaces [2], we find a conductance of the one-atom contact close to  $G_0$ : At zero bias, the individual channel transmissions ( $> 0.005$ ) are  $T_1 = 1.00$  and  $T_2 = 0.02$ . For comparison, the total transmission through a three-atom chain is dominated

---

<sup>2</sup>In Brandbyge's calculations, the electrodes to the left and right are shifted in energy by  $\pm eV/2$ . A different convention was used in Chapter 2. That is why the integration limits of the zero temperature version of Eq. (2.23) differ from Eq. (5.2) given above.

by a single eigenchannel with transmission  $T_1 = 0.98$ . The total transmission for the longer chain goes down by 0.04 due to loss of tunneling contributions through neighboring atoms [13].

The eigenchannel transmissions vs. energy  $T_n(E, V)$  for the two-atom contact are shown in the panels to the right in Fig. 5.6 for  $V = 0$  Volt and  $V = 1.5$  Volt. At zero bias, we find a conductance slightly below  $2G_0$ , which is dominated by two eigenchannels with high transmission probabilities ( $T_1 = 0.95$  and  $T_2 = 0.91$ ). In addition, there are two weakly conducting eigenchannels ( $T_3 = 0.01$  and  $T_4 = 0.01$ ). We find that the  $I$ - $V$  curve is very linear for the one-atom contact and has a slight down bend for the two-atom contact, which is caused mainly by the decrease of one of the main eigenchannel transmissions below  $E_F$ . Although there are substantial variations in the transmissions, only modest variations are seen in the  $I$ - $V$  curve since the total transmission is found by averaging the sum of individual eigenchannel transmissions over the voltage window  $[E_F - eV/2; E_F + eV/2]$ , see Eq. (5.2).

## 5.6 Discussion

### When Au is exposed to air

Although Au is the noblest of all metals [28], a layer of contaminant organic molecules with a thickness of a few Å actually forms spontaneously on the clean surface when exposed to ambient conditions [29, 30]. Minuscule amounts of organic molecules in the laboratory atmosphere is sufficient to rapidly cover the surface. For example, with a concentration of only one part-per-billion (ppb) at atmospheric pressures, the surface will be covered in about 1 s if the sticking probability is one. When using a sample which has been rinsed in ethanol and exposed to atmospheric pressures before being transferred to the UHV chamber, this layer is likely to hinder the formation of a clean metal-metal contact. This is especially the case if the contact is formed by soft indentation. In that case the current could be due to tunneling through an insulating film instead of conduction through an atomic-size metal contact. As will be discussed thoroughly in the next Chapter, this leads to  $I$ - $V$  curves with a sizeable positive value of  $G''$  in agreement with curve (d) in Fig. 5.3 and the open circles in Figs. 5.4 and 5.5.

Since transmission coefficients in tunneling are orders of magnitudes smaller than those typically encountered in metallic conduction, the contact area needed to reach a conductance of a few  $G_0$  will be much larger than atomic dimensions.

This readily explains the high stability of these contacts, the continuous distribution of  $G$  in Fig. 5.4, and the fact that they can sustain larger voltages than metallic contacts [cf. curve (d) in Fig. 5.3]. In hard indentations, the thin contaminant layer is likely to be destroyed, as evidenced by the preference for quantized values of the low bias conductance in Fig. 5.4.

For the sputtered and annealed samples all  $I$ - $V$  curves are almost linear in perfect agreement with the TB calculations in Fig. 5.6 and the high bias conductance histograms obtained from transient relay contacts by Sakai and co-workers [17–19].

### Previous measurements of $I$ - $V$ curves

In the past, there have been several reports regarding  $I$ - $V$  characteristics of Au contacts [4–10]. These studies have all been done at RT on contacts that were stable over a period of  $\gtrsim 0.1$  s. The reported  $I$ - $V$  curves are all qualitatively similar to the nonlinear curve (d) in Fig. 5.3, but quantitatively the size of  $G''$  varies from study to study. Although our experiments strongly suggest that stable contacts and nonlinear  $I$ - $V$  curves are signatures of a contaminated Au-Au interface, this does not automatically imply that all previously reported  $I$ - $V$  curves are nonlinear for the same reasons. For instance, it has been suggested that the nonlinearities could be due to electron-electron interactions in long and thin wires [4, 7, 8, 14] or Coulomb charging [4].

It is, however, relevant to review the experimental techniques used in the previous studies (when specified) and to test whether the nonlinearities could possibly be related to contamination. Such a survey is given in the following.

A majority of the previous studies of  $I$ - $V$  curves on Au have been done under ambient conditions [4–6, 8–10]. Only in a few other studies have other environments been used, namely UHV [7, 8], high vacuum [10], or contacts embedded in a polymer [10]. In several reports, the contacts were formed by using an STM to indent a metal tip into and retract it from a gold substrate [4–6, 8, 9]. When specified, the tip material was either W [6], PtIr [4], or Au [9], and several different kinds of gold substrates were used such as (i) gold evaporated on mica [4] or silicon [6], and (ii) Au(110) single crystals [4].

In most studies done under ambient conditions, it was fairly easy to obtain stable contacts with low bias conductances ranging continuously from 0–15  $G_0$  [4–6, 9]. The  $I$ - $V$  curves obtained in these air experiments are qualitatively similar to what we have found for contacts formed by soft indentations on rinsed samples [curve (d) in Fig. 5.3, and the open circles (o) in Figs. 5.4 and 5.5]. Abellán *et al.* [9] quantified the nonlinearities by fitting curves to Eq. (5.1) and

obtained  $G'' = 1.2 \pm 0.4 G_0 / \text{Volt}^2$  as typical values in agreement with our results.

However, in the air experiments by Pascual *et al.* [4] and Rodell *et al.* [5], the tip was retracted  $\sim 100$  Å while the conductance gradually decreased. These rather long excursions are comparable to the displacements used in our hard indentations, where we mostly observe transient contact geometries exhibiting almost linear  $I$ - $V$  characteristics [see curve (c) in Fig. 5.3 and the open squares ( $\square$ ) in Figs. 5.4 and 5.5]. These two studies are therefore not directly comparable to our rinsed in air/soft indentation experiments, which we explain in terms of tunneling through a thin insulating film squeezed between clean metal electrodes. The authors propose that the observed suppression of the low bias conductance far below  $G_0$  is due to localization effects [4, 5, 31] or transport of correlated electrons through a long, thin wire with impurities [4] (a dirty Luttinger liquid [32]).

Various other experimental techniques have been tested by Costa-Krämer *et al.* [7, 8, 10] such as home-made MCBJs in air [10] or thinning of 100 nm wide Au nanowires deposited on a  $\text{SiO}_2$  substrate by electrical heating [10]. However, these techniques have only been described very briefly in the literature, and we will instead focus on their UHV experiments [7, 8]. In these experiments, the contacts were formed between a thin wire loop and a loosely mounted wire [7]. To clean the surfaces, the wire loop was heated to glowing temperatures (600–700 °C) by passing a current through it [33]. The loosely mounted wire was then contacted with the loop a couple of times, and the setup was left to bake for about one hour [33]. Since the wires were not sputtered, we note in passing that such a cleaning procedure does not guarantee atomically clean surfaces.

In their experiments, contacts were formed by gently moving the wire loop and the loosely mounted wire in and out of contact. We would therefore expect that their results should be comparable to our cleaned in UHV/soft indentations experiments [curve (b) in Fig. 5.3, filled circles ( $\bullet$ ) in Figs. 5.4 and 5.5]. However, this is definitely not the case. Our results differ in at least three ways:

1. We can not stabilize contacts with  $G < 5 G_0$  formed on sputtered and annealed samples at RT; to acquire an  $I$ - $V$  curve, we have to cool the sample to  $-100$  °C. In contrast, Costa-Krämer *et al.* could stabilize the contacts for hours [7].
2. We find values of  $G$  close to  $1 G_0$ , and slightly below  $2 G_0$ . For larger values of  $G$ , the preference for quantized values is gradually lost, in agreement with for instance the conductance histogram shown in Fig. 3.4(b) on page 53 in Chapter 3 (and the conductance histogram in Ref. [7] for that matter). In contrast, Costa-Krämer *et al.* finds that [7]: “In all

cases, the dynamical conductance (the slope of the  $I$ - $V$  characteristic at zero bias) agrees within a few percent with the expected quantized conductance value  $[n \times (2e^2/h)]$ ." Since they show  $I$ - $V$  curves for  $n = 1, 2, \dots, 10$  in the paper, this statement implies that the conductance is quantized for much larger values of  $n$  than what is observed in, for instance, the above-mentioned histograms.

3. For contacts formed by soft indentations on a sputtered and annealed sample [filled circles ( $\bullet$ ) in Figs. 5.4 and 5.5] we find  $G'' = 0-2G_0/\text{Volt}^2$ . In contrast, Costa-Krämer *et al.* find much larger nonlinearities, ranging from 10 to  $25G_0/\text{Volt}^2$ . Even for soft indentations on rinsed samples ( $\circ$ ) we always find values smaller than  $10G_0/\text{Volt}^2$ .

The high stability of the contacts and the large values of  $G''$  found in the experiments by Costa-Krämer *et al.* fit well with the ‘tunneling through a thin insulating film’ scenario suggested by our experiments. However, if this was the underlying physical explanation for their results, we would expect a continuous distribution of low bias conductance values. Instead, an unusually strong preference for quantized values of the conductance is found ( $G = nG_0$  within a few percent for  $n = 1, \dots, 10$  [7]), which is clearly inconsistent with the tunneling picture.

It has been suggested that the observed large nonlinearities could be a signature of Luttinger liquid effects [8, 14], and this is certainly an interesting idea to pursue. However, in these models, all transmission probabilities are assumed to be very close to zero or one (otherwise the conductance would not be quantized). For real contacts, the validity of such an adiabatic model is highly questionable for  $G \gtrsim 4G_0$  (see, e.g. Fig. 2.7 on page 33 in Chapter 2 and Refs. 26, 34, 35). Therefore, in our opinion, the underlying physical explanation for the results obtained by Costa-Krämer *et al.* remains an open question.

### Contact stability

As mentioned previously, we cannot stabilize contacts with  $G \lesssim 5G_0$  at RT in UHV on ‘human response’ time scales. To achieve such a stability, we have to cool the sample to  $-100^\circ\text{C}$ . This observation coupled with the fact that we can obtain STM images with atomic resolution with the same tip [cf. Fig. 4.12 on page 107 in Chapter 4] suggest that contact is lost due to diffusion processes and not due to external vibrations. However, this interpretation is in apparent disagreement with work done by Takayanagi and co-workers [22].

At RT, they bombarded a thin gold film in the specimen stage of a UHV electron microscope with a strong electron beam, whereby holes were bored in the film. A free-standing linear chain of four gold atoms was produced by further thinning the material between adjacent holes as documented by TEM imaging. Such chains could be kept stable for two minutes or more before breaking. Moreover, the spacing between the Au atoms was found to be as large as 3.5–4.0 Å. For comparison, the nearest neighbor distance in bulk Au is 2.89 Å and the separation between the nuclei in a gold dimer (Au<sub>2</sub>) is 2.68 Å.

A possible explanation for the anomalously long Au–Au distance observed in the TEM images has been published by Sánchez-Portal *et al.* [36]. Using DFT calculations they found that monatomic gold wires exhibit a local energy minimum for a zigzagged configuration. When such a wire is placed in between the gold electrodes, it will rotate, leaving every second atom in a static position, which can be imaged with TEM. According to Takayanagi [37], however, spinning Au atoms would lead to a detectable increase of the background in between the atoms, and such an increase is not observed in the images.

Another explanation for the unreasonably large Au–Au distance in the chain could be that it is surrounded by impurity atoms such as C, which have a much smaller scattering cross-section in the TEM. It remains unclear whether such C structures would lead to a detectable signal in the TEM [37]. At least it is known that certain important geometrical changes are undetectable in the TEM images. For instance, in Fig. 3 of Ref. 22, measurements of the conductance of a spontaneously breaking contact were shown along with TEM images. In between two video frames, the last remaining row of atoms disappeared in the TEM image [38], but the conductance stayed *finite* at a value of  $\approx 0.1 G_0$  [38]! Three seconds later, the conductance suddenly jumped to zero without any apparent changes in the TEM images [38].

A thorough theoretical study on the stability of monatomic gold wires has been published by Torres *et al.* [39]. Using classical many-body force simulations, local density and generalized gradient electronic structure calculations on infinite wires, they found that such a chain would collapse spontaneously when the equilibrium distance between atoms was increased beyond about 2.8 Å. Moreover, *ab initio* molecular dynamics simulations of a three-atom chain suspended by Au electrodes with an initial interatomic spacing of 3.5 Å resulted in a broken wire within less than a picosecond. Using thermodynamical arguments, Torres *et al.* also found that wires suspended by electrodes will be thinned spontaneously by diffusion processes driven by a decreasing surface free energy. In the continuum limit the rate of decrease of the wire is (roughly) inversely proportional to both the radius and the temperature. For a 10 Å di-

ameter Au wire, this simple model gives a thinning rate of the order 1 Å/s at room temperature.

Considering these results, it makes sense that we cannot stabilize small gold contacts at RT. The underlying explanation for the remarkable stability of these contacts and the unphysically large apparent atomic spacing remains an open question.

### Limitations of the tight-binding model

The TB model used in this Chapter resulted in almost linear *I-V* curves, which are in excellent quantitative agreement with the experimental curves concerning the size of  $G$  and  $G''$ . Since simple free-electron models [12, 13] also result in an almost linear response, we should not consider the agreement between our experiment and the calculations as a proof of the validity of the TB model in all systems.

Here we take the opportunity to comment on some of the basic limitations and possible refinements of the model that could be implemented in the future. Some of these issues have also been discussed in Ref. 2.

- Although the TB model employed here is quite sophisticated in the sense that it includes non-orthogonal,  $s$ ,  $p$ , and  $d$  orbitals fitted to *ab initio* band structures, it is still of an approximate nature. A better method is to use self-consistent *ab initio* methods, such as the ones used by Lang for calculating the nonlinear conductance of single atom S-Al-Al [40], Xe [41], Al-Br-Br-Al [42], and C [43] chains. Unfortunately, such calculations are computationally very demanding, even for very small systems.
- Bulk distances and bulk TB parameters are used in the model. A better scheme would probably be to use TB parameters fitted to *ab initio* calculations of the specific system.
- Electron-electron correlations are not included in the model, thus automatically excluding the Luttinger liquid effects suggested by Jonson and co-workers as a possible source of nonlinearities [8, 14]. A Mott-Hubbard potential could be incorporated in future refinements of the model to investigate this issue further.
- In the TB model, it is assumed that all atoms are charge neutral, analogous to assuming that the electron screening length is at most a single nearest neighbor distance. Although this is a good approximation in the bulk, this

is not necessarily the case when the coordination of the atoms is lowered. To go beyond the charge neutrality approximation, Poisson's equation would have to be solved.

- The considered contact geometries have inversion symmetry, which automatically results in polarity independent  $I$ - $V$  curves [ $G'$  in Eq. (5.1) is identical to zero]. Although  $G'$  is usually quite small in our experiments, it sometimes has significant non-zero values. It could therefore be interesting to calculate  $I$ - $V$  curves for contacts which do not have inversion symmetry to investigate whether the obtained polarity dependencies are comparable to those observed in experiments.
- The model contacts are sandwiched between (110) surfaces because we have used a Au(110) sample. However, in the experiments the tip is polycrystalline, and we do not know its crystallographic orientation. This may influence the finite bias conductance as shown in similar calculations for single-atom contacts sandwiched between (100) and (111) surfaces [2]. These calculations showed an almost voltage independent transmission close to one for a three-atom chain between (100) surfaces. In contrast, the transmission of a three-atom chain between (111) surfaces dropped from almost one at zero bias to  $\sim 0.65$  at a bias voltage of 2 Volt, corresponding to  $G'' \approx -0.1 G_0/\text{Volt}^2$ .

## Electrostatic forces

When the voltage is swept back and forth during the acquisition of an  $I$ - $V$  curve, large electric fields ( $\sim \text{V}/\text{nm}$ ) build up and disappear in the contact region. This leads to a fluctuating electrostatic stress on the atoms in the contact region, which results in an attractive force between the electrodes, quadratic in the voltage. As a consequence, the contact geometry is strained whereby metallic bonds are bent, stretched, or compressed as the voltage increases. This may lead to changes in the transmission whereby artificial features reminiscent of a gradually changing contact geometry may appear in the  $I$ - $V$  curve.

Is this an important effect for the voltages used here ( $\lesssim 1$  Volt for the almost linear curves)? We think this is an open question since no studies have yet appeared trying to investigate this effect in quantitative detail. However, in an STM, where the electrodes are separated by vacuum, a recent continuum study by Hansen *et al.* [44] has shown that for realistic tip-surface separations and geometries, the tip snaps into mechanical contact with the surface when the

voltage exceeds a few Volt. This indicates that mechanical deformations caused by electrostatic forces *can* be very important in these atomic-scale systems.

### Current-induced effects

There are (at least) two additional effects, which may lead to contact deformations during a voltage burst. These are related to the extreme current-densities ( $10^8$ – $10^9$  A/mm<sup>2</sup>), which are found in the atomic-scale region.

- When the current is transmitted through a set of partially open eigenchannels, a backscattering of electrons occurs at the entrance of the constriction. Thus, there will also be a momentum transfer, which leads to ‘electron wind’ forces on the atoms [3, 17]. A simple estimate [17, 45] shows that these forces are in the nN regime when the bias voltage is  $\sim 1$  Volt, i.e., comparable to the force needed to break a Au-Au bond [21]! This recoil force is linear in the voltage.
- The metallic bonds will be weakened when a large current flows through them [3, 45]. The bond weakening is due to a depopulation of electronic states, with electrons below the Fermi energy being promoted to states above the Fermi energy with a more anti-bonding character. This effect is of second order in the applied voltage. An estimate based on a simple TB model gives  $10^{-3}$  nN for the bond weakening in an atomic chain biased at 0.1 Volt [3, 20].

## 5.7 Summary

We have shown that *I-V* characteristics of transient atomic-size gold contacts formed on sputtered and annealed samples *or* via hard indentations are almost linear. In addition, we have presented self-consistent tight-binding calculations of *I-V* curves which are in excellent quantitative agreement with our experimental findings. The results are also in perfect agreement with the high bias conductance histograms published by Sakai and co-workers [17–19].

All previously published *I-V* curves of stable gold contacts at room temperature [4–10] are nonlinear with a slope which increases with increasing voltage in apparent disagreement with our results. However, when our samples are rinsed in air *and* the contacts are formed in soft indentations, we also get nonlinear curves. The nonlinearities in our experiments are suggested to be due to tunneling through a thin insulating film of contaminant organic molecules.

When working with sputtered and annealed samples, we always check the cleanliness of the surface by atomic-scale STM imaging. In none of the previous experiments have the samples been sputtered and annealed, *nor* has the cleanliness been checked. Since our experiments demonstrate a clear link between surface cleaning procedures and the shape of the  $I$ - $V$  curves, there are reasons to suspect that previous measurements were influenced by a contaminated Au-Au interface. This would also explain the puzzling stability of these contacts at room temperature.

## References

- [1] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [2] M. Brandbyge, N. Kobayashi, and M. Tsukada, “Conduction channels at finite bias in a single-atom gold contact,” *Phys. Rev. B* **60**, 17064–17070 (1999).
- [3] T. N. Todorov, J. Hoekstra, and A. P. Sutton, “Current-induced forces in atomic-scale conductors,” *Philos. Mag. B* **80**, 421–455 (2000).
- [4] J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. M. Baró, N. García, U. Landman, W. D. Luedtke, E. N. Bogachek, and H.-P. Cheng, “Properties of Metallic Nanowires: From Conductance Quantization to Localization,” *Science* **267**, 1793–1795 (1995).
- [5] B. Rodell, V. Korenivski, J. Costa, and K. V. Rao, “Quantized conductivity in STM drawn 100Å nanowires of Au and Ni,” *Nanostruct. Mater.* **7**, 229–235 (1996).
- [6] V. V. Dremov and S. Y. Shapoval, “Quantization of the conductance of metal nanocontacts at room temperature,” *Pis'ma Zh. Eksp. Teor. Fiz.* **61**, 321–324 (1995), [*JEPT Lett.* **61**, 336 (1995)].
- [7] J. L. Costa-Krämer, N. García, P. García-Mochales, P. A. Serena, M. I. Marqués, and A. Correia, “Conductance quantization in nanowires formed between micro and macroscopic metallic electrodes,” *Phys. Rev. B* **55**, 5416–5424 (1997).
- [8] J. L. Costa-Krämer, N. García, M. Jonson, I. V. Krive, H. Olin, P. A. Serena, and R. I. Shekter, “Low voltage nonlinear conductance of gold nanowires: Room temperature Luttinger-liquid effect?,” In *Nanoscale science and technology: Proceedings of the NATO advanced research workshop, Toledo 1997.*, N. García, M. Nieto-Vesperinas, and H. Rohrer, eds., E: Applied Sciences **348**, 1–10 (Kluwer, Dordrecht, 1998).
- [9] J. Abellán, R. Chicón, and A. Arenas, “Properties of nanowires in air. Controlled values of the conductance,” *Surf. Sci.* **418**, 493–501 (1998).
- [10] A. Correia, J.-L. Costa-Krämer, Y. W. Zhao, and N. García, “Non-linear contribution to intensity-voltage characteristics of gold nanowires,” *Nanostruct. Mater.* **12**, 1015–1020 (1999).

- [11] K. Yuki, S. Kurokawa, and A. Sakai (unpublished).
- [12] J. I. Pascual, J. A. Torres, and J. J. Sáenz, “Nonlinear ballistic conductance in atomic-scale metallic contacts,” *Phys. Rev. B* **55**, R16029–R16032 (1997).
- [13] A. García-Martín, T. López-Ciudad, J. A. Torres, A. J. Caamaño, J. I. Pascual, and J. J. Sáenz, “Differential conductance in atomic-scale metallic contacts,” *Ultramicroscopy* **73**, 199–203 (1998).
- [14] M. Jonson, P. Sandström, R. I. Shekhter, and I. V. Krive, “Nonlinear conductance of nanowires: a signature of Luttinger liquid effects?,” *Superlattices and Microstructures* **23**, 957–967 (1998).
- [15] P. S. Ho and T. Kwok, “Electromigration in metals,” *Rep. Prog. Phys.* **52**, 301–348 (1989).
- [16] R. S. Sorbello, “Theory of Electromigration,” *Solid State Physics* **51**, 159–231 (1998).
- [17] H. Yasuda and A. Sakai, “Conductance of atomic-scale gold contacts under high-bias voltages,” *Phys. Rev. B* **56**, 1069–1072 (1997).
- [18] K. Itakura, K. Yuki, S. Kurokawa, H. Yasuda, and A. Sakai, “Bias dependence of the conductance of Au nanocontacts,” *Phys. Rev. B* **60**, 11163–11170 (1999).
- [19] K. Yuki, A. Enomoto, and A. Sakai (unpublished).
- [20] T. N. Todorov, “Local heating in ballistic atomic-scale contacts,” *Philos. Mag. B* **77**, 965–973 (1998).
- [21] G. Rubio, N. Agraït, and S. Vieira, “Atomic-Sized Metallic Contacts: Mechanical Properties and Electronic Transport,” *Phys. Rev. Lett.* **76**, 2302–2305 (1996).
- [22] H. Ohnishi, Y. Kondo, and K. Takayanagi, “Quantized conductance through individual rows of suspended gold atoms,” *Nature* **395**, 780–783 (1998).
- [23] A. I. Yanson, G. R. Bollinger, H. E. van den Brom, N. Agraït, and J. M. van Ruitenbeek, “Formation and manipulation of a metallic wire of single gold atoms,” *Nature* **395**, 783–785 (1998).

- [24] J. M. Krans, C. J. Muller, N. van der Post, F. R. Postma, A. P. Sutton, T. N. Todorov, and J. M. van Ruitenbeek, “Comment on “Quantized Conductance in an Atom-Sized Point Contact”,” *Phys. Rev. Lett.* **74**, 2146 (1995).
- [25] H. E. van den Brom, A. I. Yanson, and J. M. van Ruitenbeek, “Characterization of individual conductance steps in metallic quantum point contacts,” *Physica B* **252**, 69–75 (1998).
- [26] B. Ludoph, M. H. Devoret, D. Esteve, C. Urbina, and J. M. van Ruitenbeek, “Evidence for Saturation of Channel Transmission from Conductance Fluctuations in Atomic-Size Point Contacts,” *Phys. Rev. Lett.* **82**, 1530–1533 (1999).
- [27] B. Ludoph and J. M. van Ruitenbeek, “Conductance fluctuations as a tool for investigating the quantum modes in atomic size metallic contacts,” *Phys. Rev. B* **61**, 2273–2285 (2000).
- [28] B. Hammer and J. K. Nørskov, “Why gold is the noblest of all the metals,” *Nature* **376**, 238–240 (1995).
- [29] G. S. Ferguson, M. K. Chaundhury, G. B. Sigal, and G. M. Whitesides, “Contact Adhesion of Thin Gold Films on Elastomeric Supports: Cold Welding Under Ambient Conditions,” *Science* **253**, 776–778 (1991).
- [30] G. W. Whitesides, (private communication).
- [31] P. W. Anderson, “Absence of Diffusion in Certain Random Lattices,” *Phys. Rev.* **109**, 1492–1505 (1958).
- [32] M. Ogata and H. Fukuyama, “Collapse of Quantized Conductance in a Dirty Tomonaga-Luttinger Liquid,” *Phys. Rev. Lett.* **73**, 468–471 (1994).
- [33] J. L. Costa-Krämer, (private communication).
- [34] H. E. van den Brom and J. M. van Ruitenbeek, “Quantum Suppression of Shot Noise in Atom-Size Metallic Contacts,” *Phys. Rev. Lett.* **82**, 1526–1529 (1999).
- [35] B. Ludoph and J. M. van Ruitenbeek, “Thermopower of atomic-size metallic contacts,” *Phys. Rev. B* **59**, 12290–12293 (1999).

- [36] D. Sánchez-Portal, E. Artacho, J. Junquera, P. Ordejón, A. García, and J. M. Soler, “Stiff Monatomic Gold Wires with a Spinning Zigzag Geometry,” *Phys. Rev. Lett.* **83**, 3884–3887 (1999).
- [37] K. Takayanagi, (private communication).
- [38] H. Ohnishi, (private communication).
- [39] J. A. Torres, E. Tosatti, A. Dal Corso, F. Ercolessi, J. J. Kohanoff, F. D. Di Tolla, and J. M. Soler, “The puzzling stability of monatomic gold wires,” *Surf. Sci.* **426**, L441–L446 (1999).
- [40] N. D. Lang, “Resistance of atomic wires,” *Phys. Rev. B* **52**, 5335–5342 (1995).
- [41] A. Yazdani, D. M. Eigler, and N. D. Lang, “Off-Resonance Conduction Through Atomic Wires,” *Science* **272**, 1921–1924 (1996).
- [42] N. D. Lang, “Negative differential resistance at atomic contacts,” *Phys. Rev. B* **55**, 9364–9366 (1997).
- [43] N. D. Lang and P. Avouris, “Carbon-Atom Wires: Charge-Transfer Doping, Voltage Drop, and the Effect of Distortions,” *Phys. Rev. Lett.* **84**, 358–361 (2000).
- [44] O. Hansen, J. T. Ravnkilde, U. Quaade, K. Stokbro, and F. Grey, “Field-Induced Deformation as a Mechanism for Scanning Tunneling Microscopy Based Nanofabrication,” *Phys. Rev. Lett.* **81**, 5572–5575 (1998).
- [45] A. P. Sutton and T. N. Todorov, “Mechanical and electrical properties of atomic-scale metallic contacts — experiment, simulation and future directions,” In *Boundaries and Interfaces in Materials: The David A. Smith Symposium*, R. C. Pond, W. A. T. Clark, A. H. King, and D. B. Williams, eds., pp. 189–197 (Warrendale, Pennsylvania, 1998).



## Chapter 6

# Tunneling Through a Thin Insulating Film

*... And when a plate of golde shall be bonded with a plate of silver, or joyned thereto, it needeth to beware of three things, of dust, of winde, and of moysture, for if any hereof come betweene golde and silver, they may not be joyned together the one with the other; and therefore it needeth to bond these two mettals together in a full cleane place and quiet ...*

These words might very well describe the exceptional degree of cleanliness needed to make metal bonding between two metal surfaces. They come from the Chapter on gold in a massive encyclopedia, *De Proprietatibus Rerum* — “The Properties of Things” — compiled around 1250 by the Franciscan friar Bartholomaeus Anglicus, originally written in Latin (for a review, see Ref. 1). The quotation given here is a slightly modernized version of the English translation made in the year 1398.

Although this medieval source discusses the problem that arises when trying to cold weld two metals by hammering, the remarks are also highly relevant in the context of contact formation. For instance, we have already mentioned in the previous Chapter that when a freshly prepared Au surface is exposed to ambient conditions, a thin layer of weakly adsorbed organic molecules with a thickness of a few Ångström forms spontaneously within a short period of time on the exposed surface [2, 3]. This layer has to be destroyed in order to make a good metal-metal contact. In cold welding of macroscopic objects, this is accomplished by hammering or rolling.

## 6.1 Introduction

As demonstrated in the previous Chapter, clear nonlinearities in recorded  $I$ - $V$  curves on Au point contacts are clearly coupled to the cleanliness of the system and the method used to form the contacts. Only when the sample has not been sputtered and annealed *and* the contacts are formed by soft indentations, do we see sizeable nonlinearities in the  $I$ - $V$  curves with a faster than linear increase of the current with increasing voltage.

It is therefore quite certain that these  $I$ - $V$  curves do not correspond to pure atomic-scale metal-metal contacts. Instead we may envision that the surface is covered with a thin layer of a contaminant, insulating material as illustrated in Fig. 6.1 [2]. As the tip is gently pressed against the layer in a soft indentation, a relatively large contact may be formed where the dielectric material is squeezed between the tip and sample and hinders the formation of metallic contacts.

The current is now due to tunneling through the layer in a large contact instead of conduction through an atomic-scale metallic contact. These contacts can sustain larger voltages than the metallic contacts because the current is distributed over a much larger area. The improved stability is also a natural consequence of the much larger contact. Changing the contact area by a few atomic cross-sections does not change the conductance dramatically. Thus, the conductance can be varied almost continuously.

It is therefore interesting to investigate, whether tunneling through a thin insulating film results in nonlinear  $I$ - $V$  curves similar to those observed in ex-

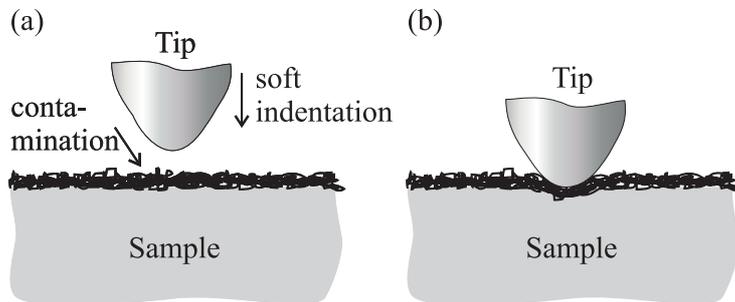


Figure 6.1: (a) When a gold sample is exposed to atmospheric conditions, it may be covered with a few molecular layers of contaminant molecules. (b) If the tip is indented softly into the surface, the dielectric layer is squeezed between tip and sample.

periments. This problem has been investigated by numerous authors in the past, and various general current-voltage formulas were derived in a series of classical papers in the early 1960's. A review of this work is given in Sec. 6.2 below. However, these treatments were mainly intended for calculating  $I$ - $V$  curves through fairly thick layers (say, 15 Å or thicker), where the WKB<sup>1</sup> approximation is valid. Since we suspect that the contaminant layer on the Au surface is ultra-thin (down to a single monolayer), these formulas may be of limited value.

Instead, we have been inspired by the methods which are used to evaluate transmission probabilities in metallic conduction. As discussed in Chapter 2, this problem has many similarities with classical tunneling problems, but in contrast to thick tunneling barriers, the transmission probabilities are of the order unity. In these problems, the WKB approximation is clearly inadequate. As an alternative, many authors [5–10] have instead used Kemble's method [4, 11, 12] or variations thereof for calculating the transmission probability. This method gives much more reliable transmission probabilities for almost transparent barriers, and in some cases it is *exact*.

In the following, we will use Kemble's method in deriving analytic formulas for the current-voltage relationship for tunneling through a (possibly very thin) parabolic barrier between planar free-electron electrodes. A parabolic barrier model is chosen because it is simple and gives a better description of the real barrier than the often used square barrier. We show that the larger the thickness, the more nonlinear the IV curves will be. By relating the model parameters to experimental curves we estimate that the insulating film in the junctions varies from 5 to 20 Å. In the opaque barrier limit, the formulas are shown to be a special case of a general  $I$ - $V$  formula, which was derived within the WKB approximation by Stratton in 1962 for arbitrary symmetric barriers [13].

## 6.2 Review

Pioneering work on the related problem of calculating the voltage dependent current density  $j(V)$  due to tunneling between metal electrodes separated by vacuum was done by Frenkel in 1930 [14]. By considering the velocity distribution of the electrons in free electron electrodes, Frenkel derived a general integral

---

<sup>1</sup>Named after Wentzel, Kramers, and Brillouin, who independently treated this problem in 1926, see Ref. 4 and references therein.

formula for the current,<sup>2</sup> which has been used widely since then (see, e.g. Refs. 13, 16–23). As an application of the general formula, he modeled the vacuum by a rectangular barrier with a height (relative to the Fermi energy) equal to the work function of the electrodes. From the barrier an approximate expression for the tunneling probability was calculated and used to evaluate  $j(V)$  to first order in  $V$ . He also argued that the slope of the  $j(V)$  curve should increase at higher voltages.

However, Frenkel's physical model did not include the electron-image potential, which arises due to the attractive force existing between the positive-polarized metal surfaces and the negative-charged electron in the gap. A first attempt to include this refinement was done by Sommerfeld and Bethe in 1933 [16], who showed that the image potential leads to a lowering of the barrier with decreasing distance between the electrodes. They calculated  $j(V)$  (to first order in  $V$ ) by (i) approximating the image potential with a parabolic barrier; and (ii) by using the WKB approximation to calculate the tunneling probability. A closer parabolic fit to the image potential was given by Holm and Kirchstein in 1935 [17].

Formulas for the nonlinear current response of the tunneling gap were calculated within a rectangular barrier model by Holm in 1951 [18]. These results were generalized to thin insulating films and arbitrary barriers by Stratton in 1962 [13]. He discussed how screening in the dielectric film would lead to a damping of the image potential by a factor of  $K$ , where  $K$  is an appropriate high-frequency dielectric constant approximately equal to the square of the refractive index. He also pointed out that the classical image potential is invalid very close to the surface of the electrodes. Thus, for computational simplicity, he again approximated the image potential by a parabola (in a slightly different manner than done by both Sommerfeld and Bethe [16] and Holm [17]) and calculated the transmission using the WKB approximation.

In 1963 Simmons presented an alternative arbitrary barrier approximation within WKB for calculating the tunneling current density between similar [19] and dissimilar [20] (free electron) electrodes. This time, the image potential was fitted with a hyperbolic function, which, especially at high voltages, is a much better fit to the true potential than any of the parabolic approximations [13, 16, 17, 19].<sup>3</sup>

---

<sup>2</sup>In the following, we will see that this formula [basically Eq. (6.9)] is a special case of a more general Landauer-Büttiker type formula [Eq. (6.4)] as discussed by Büttiker [15].

<sup>3</sup>The image potentials and all formulas based on it in Simmons papers from 1963 [19, 20] are in error by a numerical factor of two as noticed by Simmons in a later paper [21]. Unfortunately, the uncorrected forms are still in widespread use as discussed by Miskovsky *et*

All transmission probabilities mentioned so far were calculated using the WKB approximation [13, 16–23] or other approximations [14]. However, the WKB approximation breaks down when the barrier varies rapidly on a length scale of  $\lambda_F$  [24]. Thus, it is invalid for very thin films with a thickness of one or a few molecular layers as we suspect may be the case.

In the following, we present a new modified model based on a parabolic barrier, which is valid for these very thin films and gives a simple relationship between the thickness of the film and the observed nonlinearities.

### 6.3 A simplified geometry

To get a better feeling of how tunneling at finite bias gives rise to an increasing slope of the  $I$ - $V$  curve with increasing voltage, we need a simple geometrical model to describe the contaminated system. We do not know any details about the geometry in the real system so we employ the much simpler geometry shown in Fig. 6.2, where tip and sample are replaced by rectangular rods of similar metals oriented along the  $z$ -axis.<sup>4</sup> The cross-section of the rods is bounded by  $0 \leq x \leq L_x$  and  $0 \leq y \leq L_y$ . The cross-sectional area  $A = L_x L_y$  is considered as an effective area and is expected to be comparable to the real contact area in Fig. 6.1(b). The contaminant layer is modeled as an insulating slab of uniform thickness  $d$ , which separates the left ( $z < 0$ ) and right ( $z > d$ ) electrodes.

### 6.4 General formulas for the current

In the following, we will derive various general formulas for the voltage dependent current due to tunneling. We start out with the most general expressions and proceed with a series of reasonable simplifications. Most of these formulas have been derived previously in the literature.

#### Are inelastic tunneling processes important?

When an electron tunnels through a dielectric, it may lose energy in inelastic scattering processes (inelastic tunneling). The most important loss process is caused by the electron-phonon coupling [26, 27]. However, this coupling can be

---

*al.* [23]. There are also other errors in these frequently cited papers [22].

<sup>4</sup>Hereby, we exclude any polarity dependence (rectification) of the current in the model due to geometrical asymmetry (see [23, 25] and references therein) or dissimilar metals [14, 20].

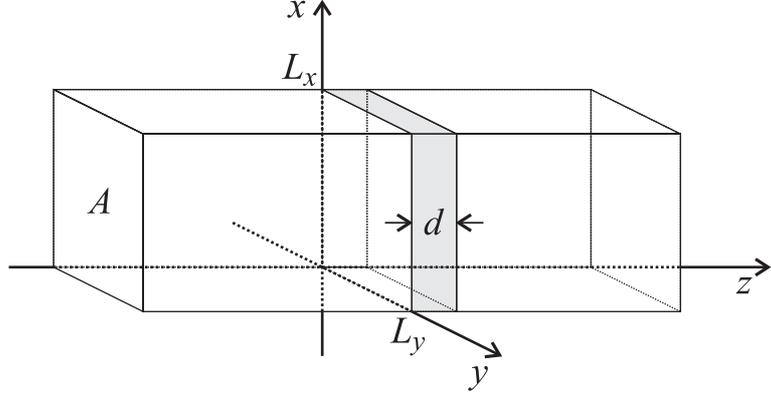


Figure 6.2: Model geometry used for calculating the voltage dependent tunnel current through a thin insulating film.

neglected if the tunneling time [28–31]  $\tau$  of the electron is much smaller than the typical stretch times  $\omega_d^{-1}$  of the dielectric [26, 27] ( $\omega_d$  is a typical phonon frequency of the dielectric). Thus, we can assume that the tunneling is elastic if  $\xi \ll 1$  [26, 27], where

$$\xi \equiv \omega_d \tau. \quad (6.1)$$

The tunneling time can be estimated from the formula [28–31]

$$\tau = \sqrt{\frac{m}{2}} \int_{z_1}^{z_2} \frac{dz}{\sqrt{V(z) - E}}, \quad (6.2)$$

where  $m$  is the electron mass,  $V(z)$  is the barrier,  $E$  is the energy of the electron, and  $z_1, z_2$  are the classical turning points [ $V(z_1) = V(z_2) = E$ ]. Eq. (6.2) is valid in the WKB limit for energies below the potential barrier.

Assuming we have a rectangular barrier of height  $U_0 = V - E$  and width  $d$ , the integral in Eq. (6.2) is trivial. Upon inserting  $\tau$  in Eq. (6.1) we get

$$\begin{aligned} \xi &= \sqrt{\frac{m}{2}} \frac{\omega_d d}{\sqrt{U_0}} \\ &\approx 0.0017 \frac{\omega_d [10^{13} \text{ s}^{-1}] d [\text{\AA}]}{\sqrt{U_0 [\text{eV}]}} \end{aligned} \quad (6.3)$$

and we see that for reasonably thin barriers, say  $d < 20 \text{ \AA}$ , typical phonon frequencies ( $\sim 10^{13} \text{ s}^{-1}$ ) and not too small barrier heights ( $\sim 1 \text{ eV}$ ), we have  $\xi \ll 1$  and inelastic tunneling processes can be neglected to a first approximation.

### Elastic tunneling of independent electrons

A general energy diagram corresponding to the geometrical model in Fig. 6.2 is shown in Fig. 6.3. When no voltage is applied [Fig. 6.3(a)], the Fermi energies of the two electrodes are aligned at  $E = E_F$ . An arbitrary one-dimensional (1D) potential  $\phi(z)$  for the  $z$ -motion of the electrons is used to model the barrier which the electrons experience in the film [13, 19–21]. At finite bias the Fermi energies are displaced by  $eV$ , thus allowing electrons from the left electrode to tunnel into unoccupied states to the right.

Within the independent electron approximation, the general (elastic) current-voltage relation is given by the Landauer-Büttiker form [32, 33] [see Eq. (2.23) on page 22 in Chapter 2], which we write as

$$I(V) = \frac{2e}{h} \int_0^\infty dE [f(E) - f(E + eV)] \times \sum_{k_x, k_y} T_{k_x, k_y}(E, V), \quad (6.4)$$

where  $f(E)$  is the Fermi-Dirac distribution defined by Eq. (2.24), and the states in the left electrode are labeled according to the components  $k_x$  and  $k_y$  of the wave vector  $\mathbf{k}_\perp$  perpendicular to the  $z$ -axis [ $\mathbf{k}_\perp = (k_x, k_y, 0)$ ]. It is assumed that  $eV < E_F$  so that the bottom of the Fermi sea of the left electrode stays below the Fermi energy of the right electrode [34, 35].  $T_{k_x, k_y}(E, V)$  is the probability for an incoming electron in state  $(k_x, k_y)$  with total energy  $E$  to tunnel through the insulator.

The Fermi-Dirac functions  $f(E) - f(E + eV)$  define an energy window of ‘active’ electrons which can tunnel from the left to the right electrodes as indicated in Fig. 6.3. At zero temperature  $f(E) - f(E + eV) = 1$  if  $E_F - eV \leq E \leq E_F$  and zero for all other values of  $E$ . At finite temperatures, the steps are smeared over an energy of the order  $k_B T$ .

### Field-induced modification of the barrier

If we neglect charge build-up inside the molecules constituting the insulating layer [36], the zero voltage barrier  $\phi(V=0; z)$  is modified by  $-eVz/d$  due to the

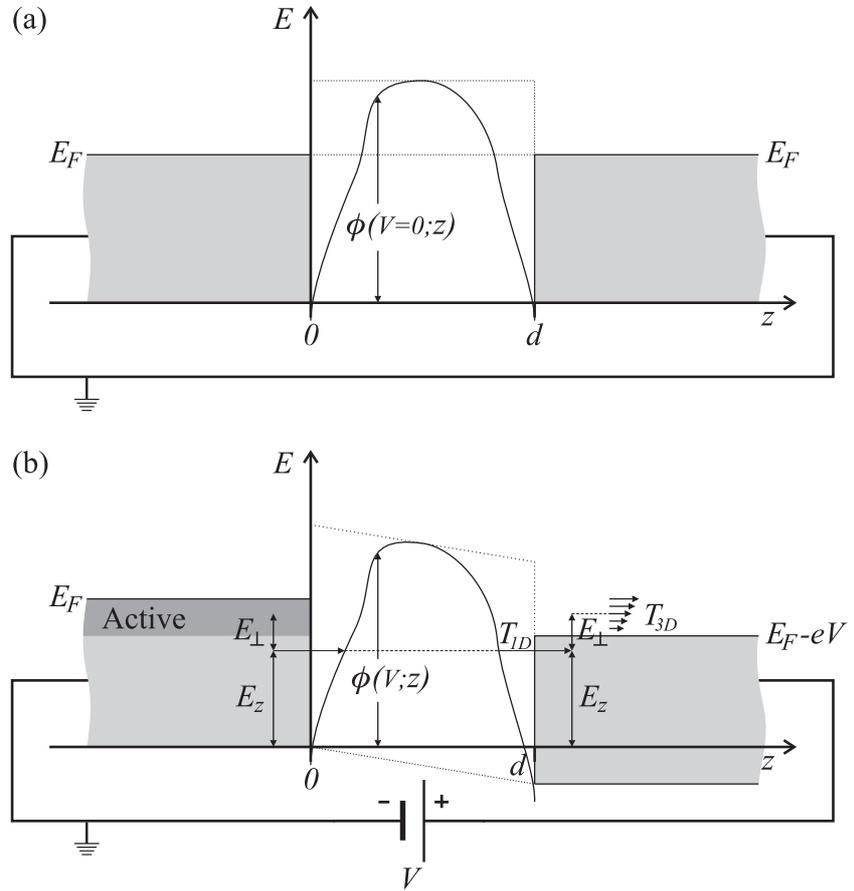


Figure 6.3: Energy diagram for tunneling through a thin insulating film corresponding to the geometry in Fig. 6.2 when (a) no bias voltage is applied, (b) a positive potential  $V$  is applied to the electrode to the right. The thin dotted lines illustrate how the generated electric field in the film deforms the zero voltage barrier.

electric field in the film when a bias voltage  $V$  is applied [13, 14, 16–19, 21, 36]

$$\boxed{\phi(V; z) = \phi(0; z) - eV \frac{z}{d}.} \quad (6.5)$$

This leads to the explicit voltage dependence of  $T_{k_x, k_y}$  in Eq. (6.4).

### Large cross-sections

If the size of the cross-section is large on a scale of  $\lambda_F$  ( $L_x, L_y \gg \lambda_F$ , the Sharvin limit [37]), the distance between the allowed values of  $\mathbf{k}_\perp$  will be so small that the sum in Eq. (6.4) can be approximated by the integral [15, 38]

$$\sum_{k_x, k_y} \dots = \int \frac{dk_x}{2\pi/L_x} \int \frac{dk_y}{2\pi/L_y} \dots = \frac{A}{4\pi^2} \iint dk_x dk_y \dots$$

Upon insertion in Eq. (6.4), we get

$$I(V) = \frac{Ae}{2\pi^2\hbar} \int_0^\infty dE [f(E) - f(E + eV)] \times \iint dk_x dk_y T_{k_x, k_y}(E, V), \quad (6.6)$$

where the integration over  $k_x$  and  $k_y$  are restricted to values which conserve  $E$ ,  $k_x$  and  $k_y$  in the transition such that both the initial and final states belong to the energy band. Eq. (6.6) was derived by Price and Radcliffe in 1959 [39] and holds for arbitrary dispersion relations (band-structures).

### Free-electron electrodes

For Au the top of the  $d$ -bands are located approximately 1.4 eV below the Fermi energy [40]. At higher energies the density of states (DOS) is dominated by the free-electron like  $sp$ -band. If we restrict ourselves to voltages less than, say, 1.2 V, it is consequently a good approximation to assume that the electrodes are free-electron-like. This will be done in the following.

Since the barrier is translationally invariant along  $x$  and  $y$ , the transmission probability depends only on the kinetic energy  $E_z$  along  $z$ , and we may write

$$T_{k_x, k_y}(E) = T_{1D}(E_z) = T_{1D}(E - E_\perp), \quad (6.7)$$

where

$$E = E_{\perp} + E_z, \quad (6.8)$$

and

$$E_{\perp} = \frac{\hbar^2(k_x^2 + k_y^2)}{2m} = \frac{\hbar^2 k_{\perp}^2}{2m}$$

is the kinetic energy due to motion perpendicular to the direction of the current flow.  $T_{1D}(E_z, V)$  is simply the probability for tunneling through the 1D potential<sup>5</sup> given by Eq. (6.5), which can be found by solving the corresponding 1D Schrödinger equation or by more approximate methods such as the WKB approximation [24].

We see that the transmission now only depends on  $k_x$  and  $k_y$  through the energy  $E_{\perp}$ . By using the identity

$$dk_x dk_y = 2\pi k_{\perp} dk_{\perp} = \pi dk_{\perp}^2 = \frac{8\pi^3 m}{h^2} dE_{\perp},$$

along with Eq. (6.7), the integrals over  $k_x$  and  $k_y$  in Eq. (6.6) are simplified to

$$\begin{aligned} \iint dk_x dk_y T_{k_x, k_y}(E) &= \frac{8\pi^3 m}{h^2} \int_0^E dE_{\perp} T_{1D}(E - E_{\perp}) \\ &= \frac{8\pi^3 m}{h^2} \int_0^E dE_z T_{1D}(E_z), \end{aligned}$$

where the last equality follows from Eq. (6.8). By inserting this result in Eq. (6.6), we obtain a *general formula* for the current due to tunneling between free-electron electrodes

$$\boxed{I(V) = A \frac{4\pi m e}{h^3} \int_0^{\infty} dE [f(E) - f(E + eV)] \times \int_0^E dE_z T_{1D}(E_z, V)}. \quad (6.9)$$

A similar formula (for the current density) was first derived by Frenkel in 1930 [14] and later re-derived by several others [13, 16, 18, 19].

<sup>5</sup>Although  $k_x$  and  $k_y$  are conserved during the tunneling process, this is not necessarily so for  $E_{\perp}$  (and thus  $E_z$ ) as is implicitly assumed here [13]. For  $E_{\perp}$  to be conserved, the momentum-energy dispersion in the conduction band of the insulator has to be parabolic with an effective mass equal to the free electron mass. Otherwise  $\phi$  will also be a function of  $E_{\perp}$  [13].

To better understand the general behavior of Eq. (6.9), we define the mean transmission  $T_{3D}$  for an incoming electron with energy  $E$  by averaging over all possible values of  $E_z$

$$T_{3D}(E, V) \equiv \frac{1}{E} \int_0^E T_{1D}(E_z, V) dE_z, \quad (6.10)$$

whereby Eq. (6.9) is written as

$$I(V) = A \frac{4\pi m e}{h^3} \int_0^\infty dE [f(E) - f(E + eV)] E T_{3D}(E, V). \quad (6.11)$$

$T_{3D}$  increases with increasing energy since here the electrons with  $E_z$  close to  $E$  penetrate a smaller potential barrier.<sup>6</sup> This is illustrated by the arrows in Fig. 6.3(b). Since  $T_{1D}$  increases with increasing  $V$  due to the field-induced lowering of the barrier [cf. Eq. (6.5)],  $T_{3D}$  will also increase with increasing  $V$ . Thus, the slope of the  $I$ - $V$  curve given by Eq. (6.11) increases with increasing voltage [14].

### Recovering the Sharvin result

For very thin barriers ( $T_{3D} \rightarrow 1$ ), the voltage drop in the electrodes has to be considered as well and the physical picture used in Fig. 6.3 breaks down. However, it is interesting to note that in the limit of small voltages ( $eV \ll E_F$ ), the integral in Eq. (6.11) is simply given by  $E_F eV$ . Thus, the current is given by

$$I(V) = \pi A \frac{2E_F m}{h^2} \frac{2e^2}{h} V = \frac{\pi A}{\lambda_F^2} G_0 V = G_S V,$$

where  $G_S$  is the Sharvin conductance, which we derived earlier [cf. Eq. (2.15) on page 16 in Chapter 2].

### A Sharvin type current-voltage relation

The Sharvin conductance gives an upper limit to the conductance through the film. It is therefore natural to rewrite the expression for the current such that

<sup>6</sup>We assume that the 1D barrier has only a single maximum, thus disregarding the possibility for resonant tunneling.

$G_S$  is used as a prefactor and we arrive at the final expression

$$\boxed{I(V) = G_S \mathcal{T}(V) V}, \quad (6.12)$$

where

$$\boxed{\mathcal{T}(V) = \frac{\int_0^\infty dE [f(E) - f(E + eV)] E T_{3D}(E, V)}{eV E_F}} \quad (6.13)$$

is the mean transmission probability averaged over all electrons in the energy window  $eV$  below the Fermi energy in the left electrode. In general  $0 < \mathcal{T} < 1$ , and for realistic barriers  $\mathcal{T} \ll 1$ . By writing the current in the form of Eq. (6.12) we see that the interesting physics such as nonlinearities is contained in the voltage dependence of  $\mathcal{T}(V)$ .

Quite often the integral over the Fermi-functions in Eq. (6.13) has a rather complicated structure. However, since  $T_{3D}(E, V)$  is most often a smoothly varying function of  $E$  (on the scale of  $k_B T$ ), we may instead use the Sommerfeld expansion. To second order in the temperature  $T$ , the expansion is [38]

$$\int_{-\infty}^{\infty} H(E) f(E) dE = \int_{-\infty}^{E_F} H(E) dE + \frac{\pi^2}{6} (k_B T)^2 H'(E_F) + O\left(\frac{k_B T}{E_F}\right)^4, \quad (6.14)$$

where  $H(E)$  is a function which varies smoothly on an energy scale of  $k_B T$ . Thus, we may write Eq. (6.13) as

$$\boxed{\mathcal{T}(V) \simeq \mathcal{T}_0(V) + \Delta\mathcal{T}(V)}, \quad (6.15)$$

where

$$\boxed{\mathcal{T}_0(V) = \frac{1}{eV E_F} \int_{E_F - eV}^{E_F} dE E T_{3D}(E, V)} \quad (6.16)$$

is the zero temperature mean transmission, and

$$\boxed{\Delta\mathcal{T}(V) = \frac{\pi^2}{6} \frac{(k_B T)^2}{eV E_F} [T_{1D}(E_F, V) - T_{1D}(E_F - eV, V)]} \quad (6.17)$$

is the second order temperature correction to the mean transmission. Note the steps needed to calculate  $\mathcal{T}$ :

1. Calculate the voltage and energy dependent 1D transmission probability  $T_{1D}(E_z, V)$  from the 1D barrier given by Eq. (6.5) by solving the 1D Schrödinger equation or by more approximate methods, such as the WKB approximation.
2. Calculate the voltage dependent average transmission  $T_{3D}(E, V)$  for an incoming electron with total energy  $E$  by averaging  $T_{1D}(E_z, V)$  over all possible values of  $E_z$  [cf. Eq. (6.10)].
3. Finally, calculate  $\mathcal{T}(V)$  by averaging  $T_{3D}(E, V)$  over the window of active electrons by evaluating Eq. (6.13) directly or by using the Sommerfeld expansion [Eqs. (6.15), (6.16) and (6.17)].

In the following, we will see how this can be done analytically using a simple (but reasonable) model potential.

## 6.5 Selecting a suitable model potential

In our case, we do not have any specific information about the insulating film. Therefore we have no chance of knowing the exact form of the 1D potential barrier  $\phi(0; z)$  which is needed for evaluating  $T_{1D}(E_z, V)$ . Instead we will use a simple model potential, which is described by a minimum of free parameters such as the height and thickness of the barrier and gives a reasonable description of the real barrier. Some of the most widely used model barriers are shown schematically in Fig. 6.4.

### The rectangular barrier

The simplest barrier of choice is the widely used [13, 14, 16, 18–23] rectangular barrier, which is shown in Fig. 6.4(a). At zero voltage, the barrier has a constant height  $\phi_0$  (relative to the bottom of the Fermi sea). We may loosely think of the barrier as coinciding with the bottom of the conduction band of the insulator [13, 19] or the lowest unoccupied molecular orbital (LUMO).<sup>7</sup> Usually this level is positioned below the vacuum level leading to smaller barriers than in metal-vacuum-metal tunneling [41–44]. At finite bias, the barrier becomes

---

<sup>7</sup>If the valence band of the insulator [or equivalently the highest occupied molecular orbital (HOMO) for a molecule] is closer to  $E_F$  than the conduction band/LUMO, the electrical current will actually be dominated by the tunneling of holes from right to left through the valence band (or HOMO). However, the physics of this process is completely analogous to tunneling of electrons from left to right.

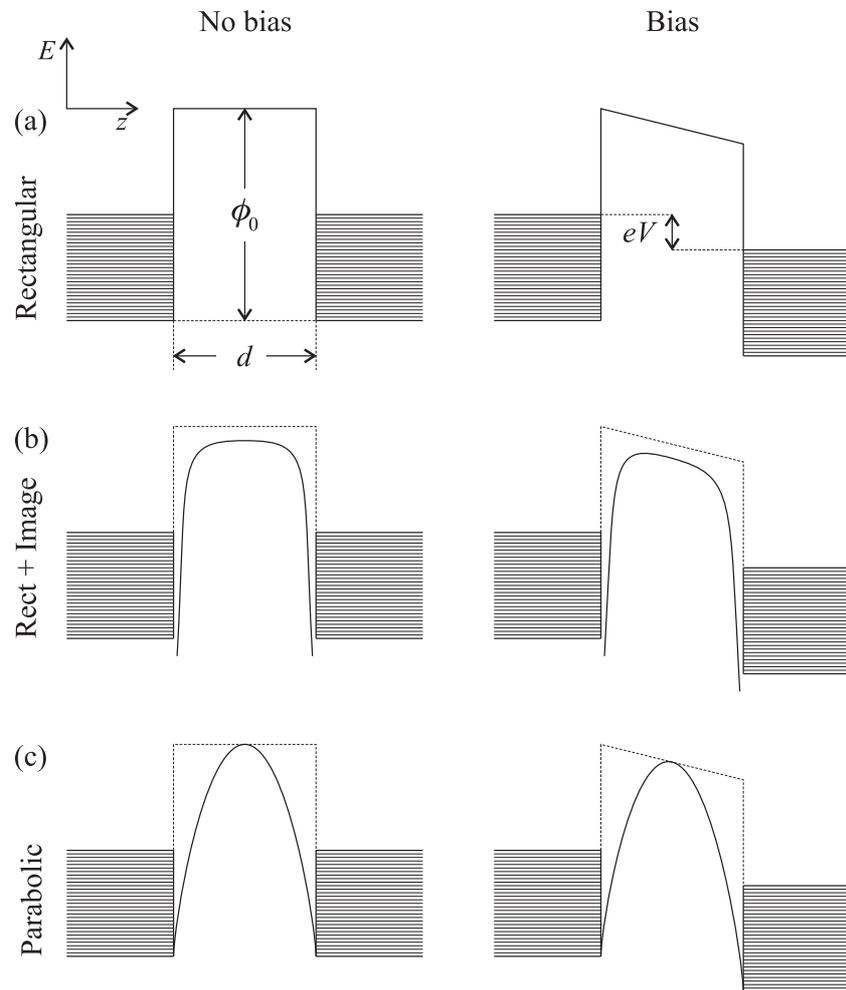


Figure 6.4: Examples of 1D model symmetric potential barriers that may be used when calculating the current through a thin insulating film; (a) rectangular barrier, (b) rectangular barrier with image potential, and (c) parabolic barrier. To the left (right) is shown the potential without (with) an applied bias voltage, cf. Eq. (6.5).

trapezoidal due to the generated electric field in the layer [cf. Eq. (6.5)] with barrier height  $\phi_0 - eVz/d$ .

The rectangular barrier is a reasonable model barrier for thick films. However, strictly speaking, it is unphysical. The discontinuities in the potential at the surface planes lead to infinite forces on the tunneling electrons. Furthermore, a pure rectangular barrier results in cusps in the  $I$ - $V$  curve at voltages of  $V = \pm(\phi_0 - E_F)/e$  [19, 22] when we go from a normal tunneling regime with a constant tunneling distance to the field-emission regime [45], where the tunneling distance decreases with increasing magnitude of voltage. Such discontinuities are never seen in our experiments.

A more realistic model takes into account that the edges of the barrier will be rounded due to quantum mechanical interactions between the tunneling electron and electrons in the electrodes. In the following, we will name this interaction with its classical analogue — the *image potential*. This modification becomes increasingly important as the thickness of the film is decreased. However, as will be outlined in the following, it is quite complicated to take this effect into account in a realistic manner.

### The image potential

For simplicity, we will first discuss a classical electrostatic system consisting of a grounded semi-infinite metal electrode placed in the half space  $z \leq 0$ . If we take an electron and place it at rest in  $z > 0$ , the metal surface responds by reorganizing its surface charge distribution so that it screens the external field of the electron. Since the metal plate is grounded, this will be done in such a way that the electrostatic potential  $\phi$  fulfills the boundary condition  $\phi(x, y, z=0) = 0$ . The induced surface charge density generates an electric field, which exerts an attractive force on the electron. According to classical electrostatics, this force will be identical to that mediated by a virtual image charge  $+e$  placed in  $-z$ :

$$F = \frac{1}{4\pi\epsilon_0} \frac{e^2}{(2z)^2}.$$

The associated electrostatic potential energy is

$$W_i = -\frac{1}{4\pi\epsilon_0} \frac{e^2}{4z}. \quad (6.18)$$

If we place another grounded semi-infinite metal electrode in  $z \geq z_0$ , an electron in the vacuum gap  $0 < z < z_0$  will polarize both surfaces. This modifies the

classical image potential in Eq. (6.18) to an infinite sum. This sum can be expressed in closed form as [16]

$$W_i = -\frac{e^2}{8\pi\epsilon_0 z_0} \left[ \psi(1) - \frac{1}{2}\psi\left(\frac{z}{z_0}\right) - \frac{1}{2}\psi\left(1 - \frac{z}{z_0}\right) \right], \quad (6.19)$$

where  $\psi(x) = d[\ln\Gamma(x)]/dx$  is the digamma function,<sup>8</sup> and  $\Gamma(x)$  is the gamma function [46]. The expression in square brackets gives the shape of the barrier. The relative importance of the image barrier increases with decreasing distance between the electrodes since the prefactor is proportional to  $z_0^{-1}$ .

The image potential is negative for all values of  $z$ , thus lowering the effective barrier when superimposed on a rectangular barrier as shown schematically in Fig. 6.4(b). The maximum value  $W_i^{\max}$  is found in the middle of the vacuum gap. Using the identity  $\psi(1/2) = \psi(1) - 2\ln 2$  [46], we find

$$W_i^{\max} = W_i(z_0/2) = -\frac{e^2 \ln 2}{4\pi\epsilon_0 z_0} \approx -\frac{10 \text{ eV}}{z_0[\text{\AA}]}. \quad (6.20)$$

According to this expression, the tunneling barrier will actually collapse when the electrodes are only a few Å apart [16]. However, as will be outlined in the following, it is not correct to use the classical expressions for the image potential [Eqs. (6.18) and (6.19)] in actual tunneling problems.

### Corrections at the surface due to exchange and correlation

The classical expressions for the image potential are singular at the metal surfaces. This means that an infinite amount of work is to be exerted on the electron to move it from the surface to infinity. However, this work should be comparable to the work function of the surface, which is finite [47]. This anomaly is an artifact of the continuum description of matter in classical electrostatics, which does not take into account the atomic-scale structure of the metal surface [47] and the quantum mechanical nature of the electron [48–50].

Close to the surface, the electron will interact with the decaying tail of electron density, which modifies the potential [49]. Within density functional theory (DFT), the image potential can be identified with the exchange-correlation potential, which represents the interaction between an electron and its exchange-correlation hole [49, 50]. As an electron moves from the bulk of the metal out through the surface, the exchange-correlation hole around it lags behind and stays on the surface as an image charge [50].

<sup>8</sup>Also sometimes called Gauss' psi [16, 46].

The electron density decays exponentially into the vacuum as  $\exp(-z/l)$  with a decay length  $l$  on the order of 1 Å [48]. Within the local density approximation (LDA), which is reasonably accurate close to the surface, the exchange-correlation potential is roughly proportional to the cube root of the electron density leading to a somewhat longer range [ $\exp\{-z/(3l)\}$ ] of the interaction [50]. Inside the metal, the potential connects smoothly to the bottom of the Fermi sea [48, 49]. If a non-local formulation of DFT is used instead, it can be shown that the asymptotic  $z^{-1}$  behavior of the image potential in Eq. (6.18) is restored 2–3 Å outside the surface [51, 52].

### Dynamical corrections

Tunneling is a *dynamic* process. The classical expressions are only valid in the quasistatic regime, where the induced surface charge density is fully relaxed in response to the changing field of the moving electron. We already identified the image potential with the potential due to exchange and correlation. Alternatively, we may think of the image potential as primarily due to interactions between the external charge and the metal surface plasmons [53–56]. The relevant response time for the metal is then  $\omega_s^{-1}$ , where  $\omega_s$  is a typical surface plasmon frequency. This should be compared with the traversal time for tunneling  $\tau$  [28–31] given by Eq. (6.2).

Following the discussion about inelastic tunneling on page 145, we substitute  $\omega_d$  with  $\omega_s$  in Eq. (6.3) to find

$$\xi_s = \omega_s \tau \approx 1.7 \frac{\omega_s [10^{16} \text{ s}^{-1}] d [\text{Å}]}{\sqrt{U_0 [\text{eV}]}}.$$

If  $\xi_s \gg 1$ , the traversal time for tunneling is much larger than the response time of the surface plasmons, and we can use the quasistatic expression in Eq. (6.19) (suitably modified near the surface due to exchange and correlation). If on the other hand,  $\xi_s \ll 1$ , the surface charge density does not equilibrate at all, and the tunneling process is highly dynamic. Inserting a typical surface plasmon frequency of  $10^{16} \text{ s}^{-1}$  and barrier height  $U_0 = 2 \text{ eV}$  in Eq. (6.3), we find that  $\xi_s \approx d[\text{Å}]$ . From this estimate, we see that dynamical corrections to the quasistatic result become increasingly important as the thickness of the barrier is reduced.

### Quantum mechanical corrections

Tunneling is a *quantum mechanical* process, which has no analogue in classical physics [57, 58]. The quasiparticle states have a finite lifetime. Hence, the

image potential moves into the complex plane [57]. Since the image barrier is caused by the tunneling electron itself, it has been argued that the real part of the potential is actually a fluctuating barrier in contrast to the classical expressions [58]. Furthermore, the barrier becomes slightly asymmetric [59–61].

Calculations which include a quantum mechanical description of the tunneling electron and the dynamic response of the surface plasmons indicate that the dynamic tunneling barrier is damped in comparison with the classical result [26, 58–64]. Although the qualitative picture of these processes is well understood, there is no consensus about quantitative numbers and the validity of the methods used [26, 58–64].

For instance, Persson and Baratoff have estimated that the static image potential overestimates the polarization contribution to the tunneling *exponent* by 20–30% in a typical STM experiment [58]. Similar trends are shown in recent work by Tokić *et al.* [63], who found that the maximum of the image barrier for two W electrodes separated by  $z_0 = 5 \text{ \AA}$  is  $W_i^{\max} \approx -1.2 \text{ eV}$ , only 60% of the classical result [cf. Eq. (6.20)]. Recent work by Ness and Fisher indicates that small values of  $\xi_s$  lead to very flat barriers resembling a rectangular barrier [61].

### Corrections due to the dielectric response of the insulator

In a metal-insulator-metal (MIM) system, where the vacuum gap is replaced by a dielectric, the classical electrostatic image potential in Eq. (6.19) is reduced by a factor of the dielectric constant of the insulator [19, 21]. In the 1960's it was suggested that this formula could be adapted to tunneling problems by using a suitable high-frequency value of the dielectric constant such as the square of the refractive index [13, 53]. Modern treatments, which include quantum and dynamic effects, show that the dielectric response leads to further fluctuations of the barrier around its mean value [27]. Since the transmission probability depends exponentially on energy, this leads to an increase of the mean transmission by about 10% for typical barrier parameters [27].

### The parabolic barrier

From the discussion given above, it is clearly a complicated task to take the image potential accurately into account. This is especially true in our case since we do not know anything about the dielectric response of the contaminant layer. Since we want to keep the number of free parameters in our model at an absolute minimum, we should therefore turn to simpler barriers, which still have the most essential features of the ‘real’ barrier. An obvious choice is

the parabolic barrier [see Fig. 6.4(c)], which, unlike the rectangular barrier is continuous at the electrodes, thereby removing the unphysical infinite forces at the surface and the cusps in the  $I$ - $V$  curves. This will be used in the following.

## 6.6 Current-voltage formula for the parabolic barrier

In the past, there have been several attempts to fit the *classical* image potential in Eq. (6.19) with a parabolic barrier [13, 16, 17]. These approximations were criticized by Simmons in 1963 [19], who showed that (especially at finite bias) the parabolic approximations give a bad fit. In light of the improved understanding of the ‘real’ image potential, which have evolved during the last three decades, this critique seems less relevant nowadays. Of course, a parabolic barrier cannot mimic the singularities in the classical image potential at the surfaces. On the other hand, corrections due to exchange and correlation show, that there are no such singularities in the real barrier! And corrections due to dynamic, quantum-mechanical and dielectric effects further modify the detailed shape of the image potential. It therefore seems relevant to revisit the parabolic barrier.

### The parabolic barrier model

In our model, we place the parabolic barrier in the middle of the gap between the metal electrodes as shown by the thick solid line in Fig. 6.5 and write the barrier at equilibrium as

$$\phi(0; z) = \begin{cases} 0 & \text{if } z < 0 \text{ or } z > d, \\ \phi_0 - \frac{1}{2}K \left(z - \frac{d}{2}\right)^2 & \text{if } 0 \leq z \leq d. \end{cases} \quad (6.21)$$

The zero voltage barrier height  $\phi_0$  is closely related to the maximum of the classical image barrier  $W_i^{\max}$  [cf. Eq. (6.20)] superimposed on the energy  $E_c$ , corresponding to the bottom of the conduction band/LUMO in the insulator. However, as argued above, the lowering of the barrier predicted by classical electrostatics is probably too large. Therefore,  $\phi_0$  should be chosen such that  $E_c + W_i^{\max} < \phi_0 < E_c$ .

We select the curvature  $K$  in Eq. (6.21) such that the barrier connects continuously to the bottom of the Fermi seas at the surfaces [ $\phi(0; 0) = \phi(0; d) = 0$ ]. This choice is in the spirit of the corrections of the image potential due to exchange and correlation [49, 51, 52, 65]. Thus, the curvature can be expressed in

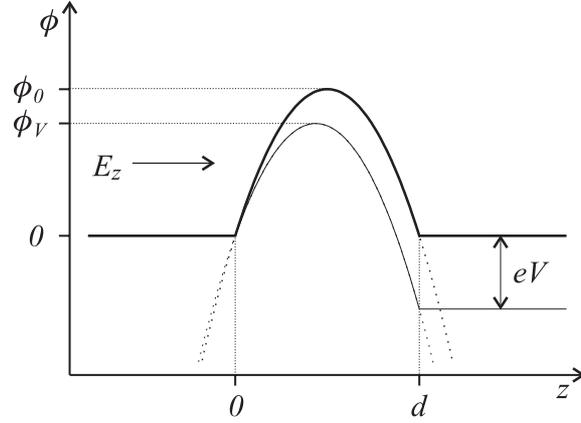


Figure 6.5: Solid lines: The truncated parabola model without (thick line) and with (thin line) an applied bias voltage. The dotted lines show the continuations of the parabolas.

terms of the barrier height and thickness<sup>9</sup> as

$$K = \frac{8\phi_0}{d^2}, \quad (6.22)$$

whereby the shape of the barrier is fully described by the two parameters  $\phi_0$  and  $d$ . If we insert  $K$  in Eq. (6.21), we can write the barrier at finite bias (shown as a thin solid line in Fig. 6.5) by using Eq. (6.5)

$$\begin{aligned} \phi(V; z) &= \phi_0 - \frac{4\phi_0}{d^2} \left( z - \frac{d}{2} \right)^2 - eV \frac{z}{d} \\ &= \phi_0 \left( 1 - \frac{1}{4} \frac{eV}{\phi_0} \right)^2 - \frac{4\phi_0}{d^2} \left( z - \frac{d}{2} \left[ 1 - \frac{1}{4} \frac{eV}{\phi_0} \right] \right)^2. \end{aligned} \quad (6.23)$$

<sup>9</sup>Self-consistent calculations on jellium [48, 49] and Al(111) [52] surfaces show that the image potential coincide with the bottom of the Fermi sea 0.5–1 Å behind the edge of the uniform, positive background for the jellium surface or the geometric edge of the aluminum surface. The distance  $d$  in our model is therefore a few Å larger than the thickness of the insulating film.

As can be seen from Fig. 6.5, the presence of an electric field in the film effectively lowers the barrier height. The voltage-dependent barrier height  $\phi_V$  is identified as the first term in Eq. (6.23)

$$\boxed{\phi_V = \max(\phi(V; z)) = \phi_0 \left(1 - \frac{1}{4} \frac{eV}{\phi_0}\right)^2.} \quad (6.24)$$

To first order in  $eV/\phi_0$ , the barrier is lowered by  $eV/2$  as expected. At higher voltages, however, the maximum is displaced away from the middle of the gap to

$$z_{\max} = \frac{d}{2} \left(1 - \frac{1}{4} \frac{eV}{\phi_0}\right)$$

as can be seen directly from Eq. (6.23). Unlike rectangular and image barriers, the parabolic barrier does not change its shape (curvature) when a voltage is applied [compare Fig. 6.4 (a), (b), and (c)]. Only the barrier height and the maximum position change. In the following, we will see that this is a very convenient feature of the model.

### Evaluating the 1D transmission

Following the ‘recipe’ we gave on page 152, we now proceed by calculating the 1D transmission  $T_{1D}(E_z, V)$  through the truncated parabolic barrier.

#### The WKB transmission

Within the WKB approximation, the transmission (for  $E_z < \phi_V$ ) is given by [24]

$$T_{1D}^{\text{WKB}}(E_z, V) = \exp \left\{ -2 \int_{z_1}^{z_2} dz \sqrt{(2m/\hbar^2)[\phi(V; z) - E_z]} \right\}, \quad (6.25)$$

where  $z_1$  and  $z_2$  are the classical turning points. Upon insertion of Eq. (6.23) in the integral, the WKB transmission is evaluated to the simple expression

$$T_{1D}^{\text{WKB}}(E_z, V) = \exp[-\gamma(\phi_V - E_z)], \quad \phi_V > E_z, \quad (6.26)$$

where  $\gamma$  is given in terms of the barrier height and width as

$$\boxed{\gamma = \frac{\sqrt{2}\pi^2}{h} \sqrt{\frac{m}{\phi_0}} d.} \quad (6.27)$$

In practical units,  $\gamma$  is given by

$$\boxed{\gamma [(eV)^{-1}] \simeq 0.805 \frac{d[\text{\AA}]}{\sqrt{\phi_0[\text{eV}]}}}. \quad (6.28)$$

The WKB approximation is quite accurate when dealing with opaque barriers. To our knowledge, it has been the method of choice in all previous applications of the parabolic barrier in metal-metal tunneling [13, 16, 17]. However, when the barrier height is only slightly larger than the Fermi energy, as is often the case in metal-insulator-metal tunneling (in particular, when the barrier is lowered further due to a large applied voltage), the validity of the WKB approximation becomes questionable. This is especially the case when the insulating layer is extremely thin (comparable to the Fermi wavelength).

### Exact transmission of an extended parabolic barrier

Alternatively, we may consider a parabolic barrier which extends to  $z = \pm\infty$  instead of being truncated at the metal surfaces as indicated by the dotted curves in Fig. 6.5. For this potential, the corresponding 1D Schrödinger equation can be rewritten to a differential equation, which can be solved analytically [4]. The resulting wavefunctions are products of confluent hypergeometric and exponential functions. These can be combined into left and right moving states, and from the asymptotic behavior at  $z = \pm\infty$ , the transmission  $T_{1D}^P$  can be found. The result is (miraculously) simple [4, 11, 12]

$$T_{1D}^P(E_z, V) = \frac{1}{1 + \exp[\gamma(\phi_V - E_z)]}. \quad (6.29)$$

Unlike the WKB result derived in Eq. (6.26), this formula is also valid for  $E_z > \phi_V$ . In the tunneling regime, it is instructive to rewrite Eq. (6.29) to [6]

$$T_{1D}^P = \frac{1}{1 + 1/T_{1D}^{\text{WKB}}}, \quad E_z < \phi_V. \quad (6.30)$$

In the opaque barrier limit ( $T_{1D}^{\text{WKB}} \ll 1$ ), the WKB transmission is identical to the extended parabolic barrier transmission. For more transparent barriers, the WKB approximation gradually breaks down. In the extreme case  $E_z = \phi_V$ , where  $T_{1D}^{\text{WKB}} = 1$ , the parabolic result is  $T_{1D}^P = 0.5$ . By refining the WKB approximation, it can actually be shown that a better estimate of the

transmission probability close to the top of an *arbitrary* barrier is given by Eq. (6.30) [11, 66, 67].

Although the expression for  $T_{1D}^P$  in Eq. (6.29) gives more reliable values for the transmission close to the top of the barrier, one might still question its validity at energies approaching the bottom of the Fermi sea. Since the asymptotic wavefunctions for the extended parabolic barrier bare no resemblance to the plane waves found for the truncated parabolic barrier, this could give very different results. On the other hand, small values of  $E_z$  correspond to opaque barriers, where the WKB approximation is known to hold fairly well. Since the WKB transmission depends only on the shape of the classically forbidden region [cf. Eq. (6.25)] (which is the same whether we truncate the parabola or not), the deviations may not be so large.

### Numerical results for the truncated parabolic barrier

To fully elucidate this point, we have used a recursion method [68] to perform exact numerical calculations of the transmission  $T_{1D}^{TP}$  of the truncated parabolic barrier for different barrier parameters and energies. Representative results are shown in Fig. 6.6 and compared with the analytical expressions for  $T_{1D}^P$  [Eq. (6.29)] and  $T_{1D}^{WKB}$  [Eq. (6.26)]. For all curves, the zero bias barrier height is chosen to  $\phi_0 = 6$  eV, only  $\approx 0.5$  eV larger than the Fermi energy for Au (5.53 eV) [38].

At zero bias voltage, the transmissions gradually saturate to one as the energy increases towards  $\phi_0$ . When a voltage of 2 V is applied, the barrier is lowered by  $\approx 1$  eV [Eq. (6.24)], effectively shifting the curves by  $\approx 1$  eV to lower energies. For the numerically calculated transmission  $T_{1D}^{TP}$ , the applied voltage also results in a slight change in the shape of the curve, which can be seen in the upper left panel by comparing  $T_{1D}^{TP}$  with  $T_{1D}^P$  for the two voltages.

As expected, the WKB approximation overestimates the transmission close to the top of the barrier. At lower energies, on the other hand,  $T_{1D}^{TP} > T_{1D}^{WKB}$  [68] as can be seen in the lower left panel. In general,  $T_{1D}^P$  closely follows  $T_{1D}^{TP}$ . Only for the very thin 8 Å barrier (left panels) do we observe small differences. For thicker barriers, like the 16 Å barrier (right panels), the difference is barely visible. It is therefore a very good approximation to use the analytical expression for  $T_{1D}^P$  in Eq. (6.29) for the transmission through the truncated parabolic barrier.

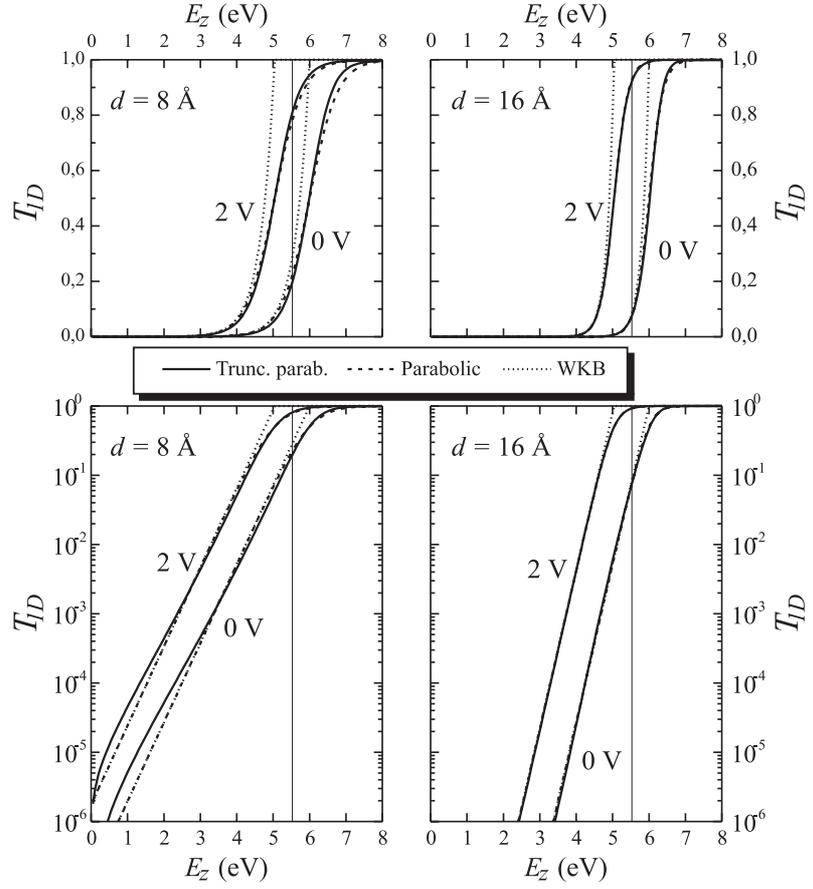


Figure 6.6: The energy dependence of the 1D transmission through a truncated parabolic barrier with barrier height  $\phi_0 = 6$  eV. The transmissions are plotted for barrier thicknesses of 8 Å (left panels) and 16 Å (right panels) using a bias voltage of 0 V and 2 V as indicated in the graphs. For each choice of barrier parameters, the transmission is calculated using three different methods: (i) the exact transmission  $T_{1D}^{\text{TP}}$  using a recursion method [68] (solid lines); (ii) the transmission  $T_{1D}^{\text{P}}$  through an extended parabolic barrier [Eq. (6.29)] (dashed lines); (iii) the transmission  $T_{1D}^{\text{WKB}}$  calculated within the WKB approximation [Eq. (6.26)] (dotted lines). The transmissions are shown on both linear (upper panels) and logarithmic (lower panels) scales. The Fermi energy for Au [38] is indicated by a thin vertical line on each panel.

### Evaluating the 3D transmission

We can now proceed to step two in our recipe on page 152 and calculate the average transmission  $T_{3D}(E, V)$  for an incoming electron with total energy  $E$

$$T_{3D}(E, V) \equiv \frac{1}{E} \int_0^E T_{1D}(E_z, V) dE_z. \quad (6.10)$$

Inserting Eq. (6.29) in Eq. (6.10) we get

$$\begin{aligned} T_{3D}^P(E, V) &= \frac{1}{E} \int_0^E \frac{dE_z}{1 + \exp[\gamma(\phi_V - E_z)]} \\ &= \frac{1}{\gamma E} \ln\{1 + \exp[-\gamma(\phi_V - E)]\}, \end{aligned} \quad (6.31)$$

where we have omitted a constant term  $\ln\{1 + \exp[-\gamma\phi_V]\}$  in the last line since it is negligible in comparison to  $\ln\{1 + \exp[-\gamma(\phi_V - E)]\}$  for realistic values of the barrier parameters.<sup>10</sup>

The energy dependence of  $T_{3D}^P$  for the same barrier parameters that were used for calculating  $T_{1D}^P$  in Fig. 6.6 is shown in Fig. 6.7. The barrier is lowered with increasing bias voltage, leading to an increased transmission in analogy with the related graphs for  $T_{1D}^P$  in Fig. 6.6. For total energies approaching  $\phi_0$ , the transmissions gradually roll off from an exponentially increasing regime (the WKB regime) and saturate at a value below one. The energy windows of active electrons [cf. Eq. (6.13)] are accentuated with thicker lines.

### Evaluating the voltage averaged mean transmission

Finally, we can calculate the mean transmission averaged over the active voltage window  $\mathcal{T}^P(V)$ . Using the Sommerfeld expansion, we first calculate the zero temperature mean transmission by inserting Eq. (6.31) in Eq. (6.13) and solving the integral. The result is

$$\mathcal{T}_0^P(V) = \frac{1}{\gamma^2 E_F eV} \left( \text{Li}_2\left\{-\exp[-\gamma(\phi_V + eV - E_F)]\right\} - \text{Li}_2\left\{-\exp[-\gamma(\phi_V - E_F)]\right\} \right), \quad (6.32)$$

where  $\text{Li}_2(z)$  is the dilogarithm, see Appendix B.

<sup>10</sup>Ignoring this term corresponds to extending the Fermi sea to all negative energies.

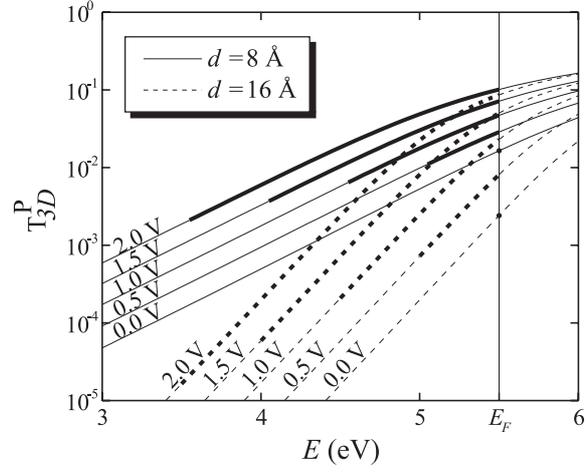


Figure 6.7: The energy dependence of the average transmission  $T_{3D}^P(E)$  for an electron incident on a parabolic barrier with barrier height  $\phi_0 = 6.0$  eV. The transmissions are plotted for a barrier thickness of 8 Å (solid lines) and 16 Å (dashed lines) and applied voltages ranging from 0 V to 2 V in steps of 0.5 V. The Fermi energy for Au is shown as a solid vertical line. The thick parts of the curves indicate the active energy window of the electrons, see Eq. (6.32).

The temperature correcting transmission term  $\Delta\mathcal{T}^P$  can be found immediately by inserting Eq. (6.29) in Eq. (6.17)

$$\Delta\mathcal{T}^P(V) = \frac{\pi^2 (k_B T)^2}{6 E_F e V} \left\{ \frac{1}{1 + \exp[\gamma(\phi_V - E_F)]} - \frac{1}{1 + \exp[\gamma(\phi_V - E_F + eV)]} \right\}. \quad (6.33)$$

Adding the correction to the zero temperature term yields the total transmission [cf. Eq. (6.15)]

$$\mathcal{T}^P(V) \simeq \mathcal{T}_0^P(V) + \Delta\mathcal{T}^P(V). \quad (6.34)$$

In Fig. 6.8 we show examples of the voltage dependence of  $\mathcal{T}^P$  at room temperature. We have used the same barrier height and Fermi energy as in the

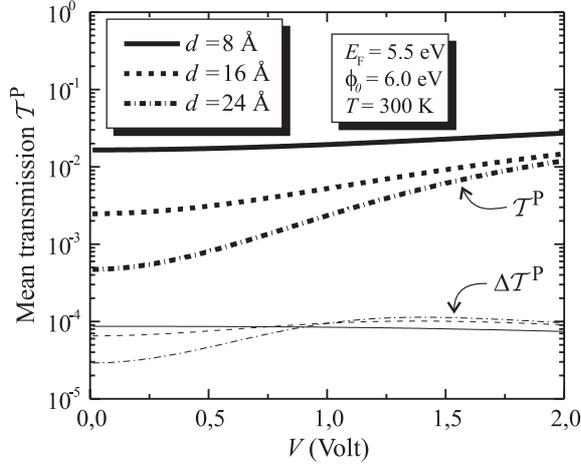


Figure 6.8: Thick lines: Examples of the voltage dependence of the mean transmission  $\mathcal{T}^P$  [cf. Eq. (6.34)] for the parabolic barrier calculated using Eqs. (6.32) and (6.33). Results for three different barrier thicknesses are shown. Thin lines: The temperature correction to the mean transmission  $\Delta\mathcal{T}^P$  calculated using Eq. (6.33).

previous Figs. 6.6 and 6.7; besides the 8 and 16  $\text{\AA}$  barriers used previously, we also show the transmission for a 24  $\text{\AA}$  barrier.

With these parameters, the zero voltage transmission increases exponentially with decreasing distance with a decay length of approximately 1.1  $\text{\AA}$ . When the voltage is increased, the transmission generally increases. Recalling that  $I^P(V) = G_S \mathcal{T}^P(V)V$ , we see that this leads to an increase in the conductance with increasing voltage, as is also observed in experiments. We also note that the thicker the barrier, the larger are the nonlinearities.

To get a better feeling of the effect of finite temperatures, we have plotted the temperature correcting term  $\Delta\mathcal{T}^P$  for  $T = 300 \text{ K}$  with thin lines in the figure. For the thin 8  $\text{\AA}$  barrier (solid lines),  $\Delta\mathcal{T}^P$  is two to three orders of magnitude lower than  $\mathcal{T}^P$ , and it can be neglected. However, when the thickness is increased, the temperature correction becomes increasingly important, and for the 24  $\text{\AA}$  barrier (dashed-dotted lines), it gives a 5% contribution to the total transmission. This is because the energy dependence of  $T_{3D}^P$  increases with increasing thickness as seen from Fig. 6.7. Thus, electrons which are excited by an energy of the order  $k_B T$  at the Fermi energy will have a significantly

enhanced transmission, and this is exactly what  $\Delta\mathcal{T}^P$  accounts for. The relative importance of finite temperature corrections for opaque barriers will be treated in more quantitative detail in the next section.

In Fig. 6.9 we demonstrate that the parabolic barrier model can be fitted with good accuracy to experimental  $I$ - $V$  curves obtained by soft indentations on rinsed Au(110) samples. This is hardly surprising since we have three parameters in the model, namely  $d$ ,  $\phi_0$ , and  $A$  (we fix the Fermi energy  $E_F$  at 5.55 eV and the temperature  $T$  at 300 K). The area  $A$  is just a scaling factor, and it turns out, that  $\phi_0$  also works mostly like a scaling factor. Therefore the dependency between  $A$  and  $\phi_0$  is almost 100%, and we should fix one of them at a reasonable value. We choose to fix the barrier height at  $\phi_0 = 7.0$  eV, which is between the Fermi energy and the vacuum level. We emphasize that this is a guessed value. The exact energy could easily be an eV higher or lower. Since  $\phi_0$  and  $A$  are closely interrelated, this implies that the fitted values for  $A$  are very unreliable. They could be off by an order of magnitude or more. The barrier thickness, on the other hand is quite robust against changes in  $\phi_0$  — it only changes by 5–10%, when  $\phi_0$  is changed by an eV.

The fitted values of the thickness for the three curves in Fig. 6.9 are of the order 10 Å, which seems to be a reasonable value. The (very uncertain) contact areas correspond to contact radii of 6–25 nm, which are plausible values. In general the nonlinearities in the  $I$ - $V$  curves are consistent with a thicknesses ranging from 5 to 20 Å.

### The opaque barrier limit

For opaque barriers, the exponentials in Eq. (6.32) will be much less than one. Since  $\text{Li}_2(x) \simeq$  for  $x \ll 1$  [Eq. (B.3) in Appendix B], we can then simply replace the dilogarithms in Eq. (6.32) with their arguments. If we furthermore use the expression for  $\phi_V$  given by Eq. (6.24), we find in the opaque barrier limit

$$\mathcal{T}_0^P \simeq \frac{\exp[-\gamma(\phi_0 - E_F)]}{\gamma E_F} \exp\left[-\frac{\gamma}{16\phi_0}(eV)^2\right] \frac{2 \sinh(\gamma eV/2)}{\gamma eV}. \quad (6.35)$$

Using similar arguments, we may simplify the expression for  $\Delta\mathcal{T}^P$  in Eq. (6.33) to a formula closely resembling Eq. (6.35). The two expressions differ only by a simple temperature and barrier shape dependent factor

$$\frac{\Delta\mathcal{T}^P}{\mathcal{T}^P} = \frac{\pi^2}{6} \gamma^2 (k_B T)^2 \approx 7.9 \times 10^{-9} \frac{d[\text{Å}]^2 T[\text{K}]^2}{\phi_0[\text{eV}]}. \quad (6.36)$$

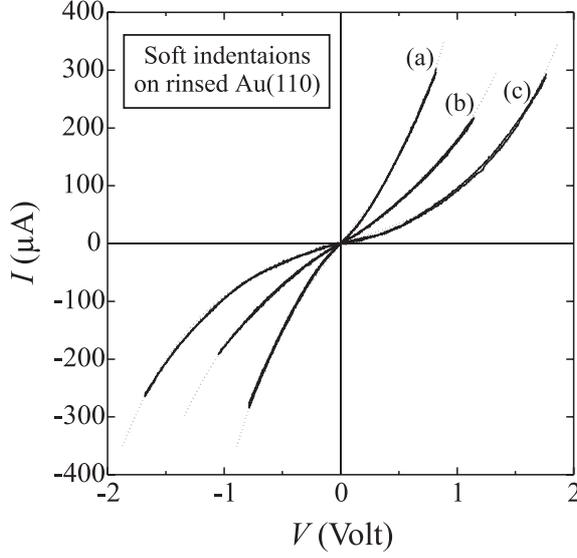


Figure 6.9: Solid lines: Three different  $I$ - $V$  curves taken from the soft indentations on rinsed samples experiments described in Chapter 5. Dotted lines: Best fits to  $I(V) = G_S T^P(V)V$ , where  $G_S$  is defined by Eq. (2.15) and  $T^P(V)$  is defined by Eqs. (6.34), (6.32), and (6.33). Fixed parameters:  $T = 300$  K,  $E_F = 5.55$  eV, and  $\phi_0 = 7.0$  eV. Free parameters: area  $A$  and barrier width  $d$ . The obtained fitting parameters for the three curves: (a)  $d = 12.7$  Å,  $A = 1750$  nm<sup>2</sup>; (b)  $d = 9.7$  Å,  $A = 184$  nm<sup>2</sup>; (c)  $d = 10.7$  Å,  $A = 147$  nm<sup>2</sup>.

We see that the relative importance of the temperature correction increases with the square of the barrier thickness, confirming the trends observed for the exact expressions in Fig. 6.8. For 100 Å thick barriers at room temperature, the factor  $\gamma k_B T$  becomes of order unity, and the Sommerfeld expansion breaks down. The formulas in Eqs. (6.35) and (6.36) are therefore only valid when  $d$  is well below 100 Å.

It is interesting to compare the opaque barrier results in Eqs. (6.35) and (6.36) with earlier work. For example, Stratton gave in his classical paper from 1962 a general formula for the  $I$ - $V$  curve of an *arbitrary* symmetric barrier within the WKB approximation [13]. As shown in Appendix C this formula has the same general form as Eqs. (6.35) and (6.36),<sup>11</sup> and for a parabolic barrier

<sup>11</sup>This could indicate that the exact expressions for the transmission in Eqs. (6.32) and

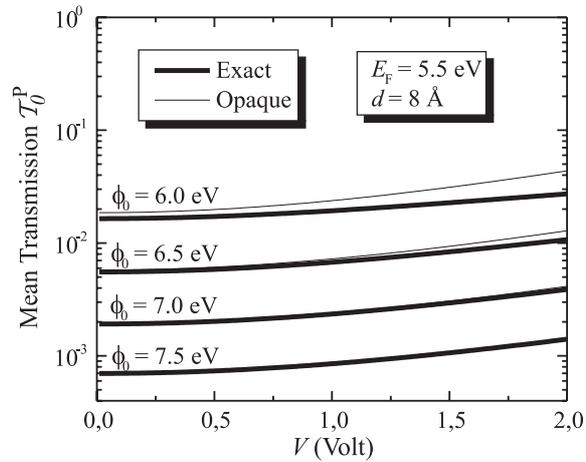


Figure 6.10: The voltage dependence of the zero temperature mean transmission  $\mathcal{T}_0^P$  using both the exact [Eq. (6.32), thick lines] and opaque barrier approximation [Eq. (6.35), thin lines]. Results for an 8 Å thick barrier using four different barrier heights are shown.

the formulas are identical!

### The validity of the WKB approximation

Since the WKB approximation has been used very extensively in the past in tunneling models, it is of relevance to investigate the conditions under which this approximation is valid. To illustrate this, we have plotted the voltage dependence of the zero temperature mean transmission  $\mathcal{T}_0^P$  using both the exact [Eq. (6.32), thick lines] and opaque barrier [Eq. (6.35), thin lines] expressions in Fig. 6.10.

We use an 8 Å thick barrier and increase the barrier height from 6.0 eV to 7.5 eV in steps of 0.5 eV. For the 7.5 eV barrier, the two expressions are basically identical for the range of voltages shown. For the 7.0 eV barrier, the curves also coincide except at voltages approaching 2 V, where the opaque expression slightly overestimates the transmission. This trend is seen more

(6.17), which go beyond the WKB approximation, are of a more general nature and also hold for other symmetric barriers which resemble the parabolic barrier (such as the true potential).

clearly as the barrier is lowered to 6.5 eV and 6.0 eV. This is because the voltage dependent barrier height  $\phi_V \simeq \phi_0 - eV/2$  decreases with increasing voltage, thus lowering the barrier towards  $E_F$ , where the WKB approximation breaks down [cf. Fig. 6.6]. However, when  $\phi_0$  is only 1 eV higher than  $E_F$ , there is still fair agreement between the two expressions, and even for the 6.0 eV barrier, the curves look qualitatively similar.

## 6.7 Summary

We have derived an analytic result for the current-voltage relation for tunneling of electrons through a thin uniform insulating layer described by a free-electron energy-momentum dispersion. The qualitative features of the classical image potential corrected for exchange and correlation are taken into account by using a truncated parabolic barrier, which connects to the bottom of the Fermi seas at the surface of the electrodes. Our model goes beyond the usual WKB approximation by using a more accurate formula for the transmission. This makes the model well suited for calculating  $I$ - $V$  curves for thin barriers with small barrier heights.

In the opaque barrier limit, the model agrees with previous findings. A comparison between the exact and opaque barrier expressions reveals good qualitative agreement even for quite transparent barriers. It is shown that temperature effects become increasingly important as the thickness of the barrier is increased. Nonlinearities in the  $I$ - $V$  curves are shown to be clearly coupled with the thickness of the barrier. The thicker the barrier, the larger the nonlinearity. Even without knowing the barrier height, it is still possible to deduce a fairly accurate value for the thickness of the insulator from experimental  $I$ - $V$  curves. It is, however, very difficult to say anything quantitatively about the barrier height and the contact area from measured  $I$ - $V$  curves as these two parameters are highly dependent.

## References

- [1] L. B. Hunt, “Bartholomaeus Anglicus on Gold. The modern views of a mediaeval encyclopaedist,” *Gold Bull.* **8**, 22–27 (1975).
- [2] G. S. Ferguson, M. K. Chaundhury, G. B. Sigal, and G. M. Whitesides, “Contact Adhesion of Thin Gold Films on Elastomeric Supports: Cold Welding Under Ambient Conditions,” *Science* **253**, 776–778 (1991).
- [3] G. W. Whitesides, (private communication).
- [4] B. S. Jeffreys, “The Asymptotic Approximation (AA) Method,” in *Quantum Theory*, D. R. Bates, ed., (Academic Press, New York, 1961), Vol. 1, Chap. 7, pp. 240–242.
- [5] J. A. Torres, J. L. Pascual, and J. J. Sáenz, “Theory of conduction through narrow constrictions in a three-dimensional electron gas,” *Phys. Rev. B* **49**, 16581–16584 (1994).
- [6] M. Brandbyge, J. Schiøtz, M. R. Sørensen, P. Stoltze, K. W. Jacobsen, J. K. Nørskov, L. Olesen, E. Lægsgaard, I. Stensgaard, and F. Besenbacher, “Quantized conductance in atom-sized wires between two metals,” *Phys. Rev. B* **52**, 8499–8514 (1995).
- [7] A. G. Scherbakov, E. N. Bogachek, and U. Landman, “Quantum electronic transport through three-dimensional microconstrictions with variable shapes,” *Phys. Rev. B* **53**, 4054–4064 (1996).
- [8] E. N. Bogachek, A. G. Scherbakov, and U. Landman, “Magnetic switching and thermal enhancement of quantum transport through nanowires,” *Phys. Rev. B* **53**, R13246–R13249 (1996).
- [9] E. N. Bogachek, A. G. Scherbakov, and U. Landman, “Shape effects on conductance quantization in three-dimensional nanowires: Hard versus soft potentials,” *Phys. Rev. B* **56**, 1065–1068 (1997).
- [10] T. López-Ciudad, A. García-Martín, A. J. Caamaño, and J. J. Sáenz, “Conductance histograms and conducting channels in atomic-scale metallic contacts,” *Surf. Sci.* **440**, L887–L890 (1999).
- [11] E. C. Kemble, in *The Fundamental Principles of Quantum Mechanics with Elementary Applications* (Dover, New York, 1937), pp. 109–112.

- [12] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics. Non-relativistic theory*, Vol. 3 of *Course of Theoretical Physics*, second revised ed. (Pergamon, London, 1965), see §50, problem 4.
- [13] R. Stratton, "Volt-current characteristics for tunneling through insulating films," *J. Phys. Chem. Solids* **23**, 1177–1190 (1962).
- [14] J. Frenkel, "On the electrical resistance of contacts between solid conductors," *Phys. Rev.* **36**, 1604–1618 (1930).
- [15] M. Büttiker, "Transmission, reflection and the resistance of small conductors," In *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, J. M. Chamberlain, L. Eaves, and J.-C. Portal, eds., Series B: Physics **231**, 51–73 (Plenum, New York, 1990).
- [16] A. Sommerfeld and H. Bethe, "Elektronentheorie der Metalle," in *Aufbau der Zusammenhängende Materie*, No. 24/2 in *Handbuch der Physik*, 2nd ed., H. Geiger and K. Schell, eds., (Julius Springer-Verlag, Berlin, 1933), Chap. 3, pp. 443–454.
- [17] R. Holm and B. Kirschstein, "Über den Widerstand dünnster Fremdschichten in Metallkontakten," *Z. Tech. Phys.* **16**, 488–494 (1935).
- [18] R. Holm, "The Electric Tunnel Effect across Thin Insulator Films in Contacts," *J. Appl. Phys.* **22**, 569–574 (1951).
- [19] J. G. Simmons, "Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film," *J. Appl. Phys.* **34**, 1793–1803 (1963).
- [20] J. G. Simmons, "Electric Tunnel Effect between Dissimilar Electrodes Separated by a Thin Insulating Film," *J. Appl. Phys.* **34**, 2581–2590 (1963).
- [21] J. G. Simmons, "Potential Barriers and Emission-Limited Current Flow Between Closely Spaced Parallel Metal Electrodes," *J. Appl. Phys.* **35**, 2474–2481 (1964).
- [22] T. E. Hartman, "Tunneling Through Asymmetric Barriers," *Appl. Phys. Lett.* **35**, 3283–3294 (1964).
- [23] N. M. Miskovsky, P. H. Cutler, and T. E. Feuchtwang, "The Multiple-Image Interactions and the Mean Barrier Approximation in MOM and MVM Tunneling Junctions," *Appl. Phys. A* **27**, 139–147 (1982).

- [24] S. Gasiorowicz, in *Quantum Physics* (John Wiley & Sons, New York, 1974), pp. 469–471.
- [25] N. M. Miskovski, S. J. Shephard, P. H. Cutler, T. E. Sullivan, and A. A. Lucas, “The importance of geometry, field, and temperature in tunneling and rectification behavior of point contact junctions of identical metal,” *Appl. Phys. Lett.* **35**, 560–562 (1979).
- [26] K. L. Sebastian and G. Doyen, “Tunneling in presence of coupling to other modes: Application to scanning tunneling microscopy,” *J. Chem. Phys.* **99**, 6677–6687 (1993).
- [27] D. Rostkier-Edelstein, M. Urbakh, and A. Nitzan, “Electron tunneling through a dielectric barrier,” *J. Chem. Phys.* **101**, 8224–8237 (1994).
- [28] M. Büttiker and R. Landauer, “Traversal Time for Tunneling,” *Phys. Rev. Lett.* **49**, 1739–1742 (1982).
- [29] M. Büttiker, “Larmor precession and the transversal time for tunneling,” *Phys. Rev. B* **27**, 6178–6188 (1983).
- [30] M. Büttiker and R. Landauer, “Traversal Time for Tunneling,” *Phys. Scr.* **32**, 429–434 (1985).
- [31] M. Büttiker and R. Landauer, “Traversal time for tunneling,” *IBM J. Res. Dev.* **30**, 451–454 (1986).
- [32] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, “Generalized many-channel conductance formula with application to small rings,” *Phys. Rev. B* **31**, 6207–6215 (1985).
- [33] Y. Imry, “Physics of mesoscopic systems,” in *Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenko, eds., (World Scientific, Singapore, 1986), pp. 101–163.
- [34] L. Martín-Moreno, J. T. Nicholls, N. K. Patel, and M. Pepper, “Non-linear conductance of a saddle point constriction,” *J. Phys. C* **4**, 1323–1333 (1992).
- [35] J. I. Pascual, J. A. Torres, and J. J. Sáenz, “Nonlinear ballistic conductance in atomic-scale metallic contacts,” *Phys. Rev. B* **55**, R16029–R16032 (1997).

- [36] S. Datta, W. Tian, S. Hong, R. Reifenberger, J. I. Henderson, and C. P. Kubiak, "Current-Voltage Characteristics of Self-Assembled Monolayers by Scanning Tunneling Microscopy," *Phys. Rev. Lett.* **79**, 2530–2533 (1997).
- [37] Y. V. Sharvin, "A possible method for measuring Fermi surfaces," *Zh. Eksp. Teor. Fiz.* **48**, 984–985 (1965), [*Sov. Phys. JETP* **21**, 655 (1965)].
- [38] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [39] P. J. Price and J. M. Radcliffe, "Esaki Tunneling," *IBM J. Res. Dev.* **3**, 364–371 (1959).
- [40] N. E. Christensen and B. O. Seraphin, "Relativistic Band calculation and the Optical Properties of Gold," *Phys. Rev. B* **4**, 3321–3344 (1971).
- [41] D. Lamoen, P. Ballone, and M. Parrinello, "Electronic structure, screening, and charging effects at a metal/organic tunneling junction: A first-principles study," *Phys. Rev. B* **54**, 5097–5105 (1996).
- [42] M. P. Samanta, W. Tian, S. Datta, J. I. Henderson, and C. P. Kubiak, "Electronic conduction through organic molecules," *Phys. Rev. B* **53**, R7626–R7629 (1996).
- [43] H. Ishii, K. Sugiyama, E. Ito, and K. Seki, "Energy Level Alignment and Interfacial Electronic Structures at Organic/Metal and Organic/Organic Interfaces," *Adv. Mater.* **11**, 605–625 (1999).
- [44] E. Emberly and G. Kirczenow, "Electrical conductance of molecular wires," *Nanotech.* **10**, 285–289 (1999).
- [45] R. H. Fowler and L. Nordheim, "Electron Emission in Intense Electric Fields," *Proc. R. Soc. London, Ser. A* **119**, 173–181 (1928).
- [46] *Handbook of Mathematical Functions*, Vol. 55 of *Applied Mathematics Series*, M. Abramowitz and I. A. Stegun, eds., (National Bureau of Standards, Washington, D.C., 1967), sixth printing.
- [47] W. Schottky, "Über den Einfluß von Strukturwirkungen, besonderes der Thomsonschen Bildkraft, auf die Elektronenemission der Metalle," *Phys. Z.* **15**, 872–878 (1914).

- [48] N. D. Lang and W. Kohn, “Theory of metal Surfaces: Charge Density and Surface Energy,” *Phys. Rev. B* **1**, 4555–4568 (1970).
- [49] N. D. Lang and W. Kohn, “Theory of Metal Surfaces: Induced Surface Charge and Image Potential,” *Phys. Rev. B* **7**, 3541–3550 (1973).
- [50] N. D. Lang, “Apparent barrier height in scanning tunneling microscopy,” *Phys. Rev. B* **37**, 10395–10398 (1988).
- [51] A. G. Eguluz, M. Heinrichsmeier, A. Fleszar, and W. Hanke, “First-Principles Evaluation of the Surface Barrier for a Kohn-Sham Electron at a Metal Surface,” *Phys. Rev. Lett.* **68**, 1359–1362 (1992).
- [52] I. D. White, R. W. Godby, M. M. Rieger, and R. J. Needs, “Dynamic Image Potential at an Al(111) Surface,” *Phys. Rev. Lett.* **80**, 4265–4268 (1998).
- [53] P. Schnupp, “The tunneling time of an electron and the image force,” *Thin Solid Films* **2**, 177–183 (1968).
- [54] R. Ray and G. D. Mahan, “Dynamical Image Charge Theory,” *Phys. Lett.* **42A**, 301–302 (1972).
- [55] M. Šunjić and G. Toulouse, “Dynamical corrections to the image potential,” *Solid State Commun.* **11**, 1629–1631 (1972).
- [56] J. Heinrichs, “Response of Metal Surfaces to Static and Moving Point Charges and to Polarizable Charge Distributions,” *Phys. Rev. B* **8**, 1346–1364 (1973).
- [57] M. Jonson, “The dynamical image potential for tunneling electrons,” *Solid State Commun.* **33**, 743–746 (1980).
- [58] B. N. J. Persson and A. Baratoff, “Self-consistent dynamic image potential in tunneling,” *Phys. Rev. B* **38**, 9616–9627 (1988).
- [59] D. B. T. Thoai and M. Šunjić, “Dynamical effects in electron tunneling: nonlocal image potentials,” *Solid State Commun.* **77**, 955–959 (1991).
- [60] M. Šunjić and L. Marušić, “Dynamical effects in electron tunneling,” *Solid State Commun.* **84**, 123–126 (1992).
- [61] H. Ness and A. J. Fisher, “A matrix method for treating the coupling between an electron and a surface plasmon: a dynamical image potential in model tunnelling junctions,” *J. Phys. Cond. Matt.* **10**, 3697–3718 (1998).

- [62] N. Klipa and M. Šunjić, “Dynamical effective potentials in electron tunneling: Path-integral study,” *Phys. Rev. B* **52**, 12408–12413 (1995).
- [63] B. Tokić, D. Šestović, L. Marušić, and M. Šunjić, “Dynamical effects and conductance asymmetry in metal-insulator-metal systems with different electrodes,” *Phys. Rev. B* **60**, 8368–8372 (1999).
- [64] K. L. Sebastian and G. Doyen, “Dynamical image interaction in scanning tunneling microscopy,” *Phys. Rev. B* **47**, 7634–7636 (1993).
- [65] N. D. Lang and W. Kohn, “Theory of Metal Surfaces: Work Function,” *Phys. Rev. B* **3**, 1215–1223 (1971).
- [66] S. C. Miller, Jr. and R. H. Good, Jr., “A WKB-Type Approximation to the Schrödinger Equation,” *Phys. Rev.* **91**, 174–179 (1953).
- [67] L. I. Glazman, G. B. Lesovik, D. E. Khmel’nitskii, and R. I. Shekhter, “Reflectionless quantum transport and fundamental ballistic-resistance steps in microscopic constrictions,” *Pis’ma Zh. Eksp. Teor. Fiz.* **48**, 218–220 (1988), [*JETP Lett.* **48**, 238 (1988)].
- [68] H. Q. Nguyen, P. H. Cutler, T. E. Feuchtwang, and N. Miskovsky, “Investigation of a new numerical method for the exact calculation of one-dimensional transmission coefficients: Application to the study of limitations of the WKB approximation,” *Surf. Sci.* **160**, 331–352 (1985).



## Chapter 7

# Current-Induced Disruptions of Monatomic Gold Contacts

In Chapter 5, we presented current-voltage ( $I$ - $V$ ) curves of atomic-sized, gold contacts formed in a clean, ultra-high vacuum (UHV) environment at room temperature (RT). Here we made two important observations concerning the mechanical stability of the contacts: (i) Only by cooling the sample could we obtain stable few-atom contacts; and (ii) if the maximum voltage applied to the contact during a voltage burst was larger than  $\sim 0.75$  Volt, the contacts became mechanically unstable. [See also Fig. 4.13(b) on page 108.]

The first observation indicates that the contact breaks primarily because of thermally activated processes and not due to, e.g., external mechanical vibrations. Instead, the spontaneous thinning of the contact is probably due to diffusion processes driven by a decreasing surface free energy [1]. The second observation tells us that the diffusion processes are activated by the applied voltage. As discussed on pages 130–132 in Chapter 5, several different underlying physical effects may, in combination, lead to more unstable contacts. These include heating [2], electron-wind forces [3, 4], bond weakening [4, 5], and mechanical stress caused by electrostatic forces [6].

At the moment very little is known about the relative importance of the different effects, and a detailed understanding of break mechanism is lacking. To further elucidate these interesting questions, we present in this Chapter

preliminary results from a new type of experiment, which enables us to study current-induced disruptions of monatomic gold contacts.

## 7.1 Experimental details

Observing current-induced disruptions of monatomic contacts requires that the gold sample is cooled. Otherwise, the non-activated contact break rate becomes so large that we cannot unravel disruptions caused by RT thermal processes from current-induced effects. Here we will briefly describe how the cooling is accomplished.

### Cooling the Aarhus STM

A cross-sectional sketch of the variable-temperature Aarhus STM is shown in Fig. 7.1. To minimize thermal drift, only the temperature of the top-plate (C) is varied. Cooling is accomplished through two Cu braids connecting the top-plate with two Al blocks, which act as heat reservoirs (not shown). Prior to the experiments, the Al blocks are cooled by contacting them with a Cu dewar cooled with liquid N<sub>2</sub>. After cooling, the Al reservoir warms up at a temperature rate of 5–10 K/hour [7]. A set of thermocouple wires are connected to the top-plate, whereby the sample temperature can be measured with an accuracy of a few K. The top-plate is thermally isolated from the STM housing by quartz balls (E). The STM housing, STM (G, I, and H), as well as the tip shaft (F) are kept at RT. The STM is mounted on a stainless steel cradle, which is mechanically isolated from the surroundings via suspension springs. A slow decrease in the temperature of the STM is avoided by counter-heating the cradle by means of zener diodes.

### Contact formation

A Au(111) sample is sputtered and annealed before being placed in the STM. The cleanliness of the sample is checked by STM imaging with atomic resolution, see Fig. 4.12 on page 107. Following this, the conventional set of STM electronics is interchanged with the fast electronics described in Chapter 4. As a first step towards forming a stable atomic-sized gold contact in UHV, we cool the tip apex by indenting it  $\sim 100$  Å into the surface and leaving it for a few seconds. Subsequently, the tip is retracted with a speed of about 100 Å/s. During the retraction, a constant bias voltage of 150 mV is applied across the contact

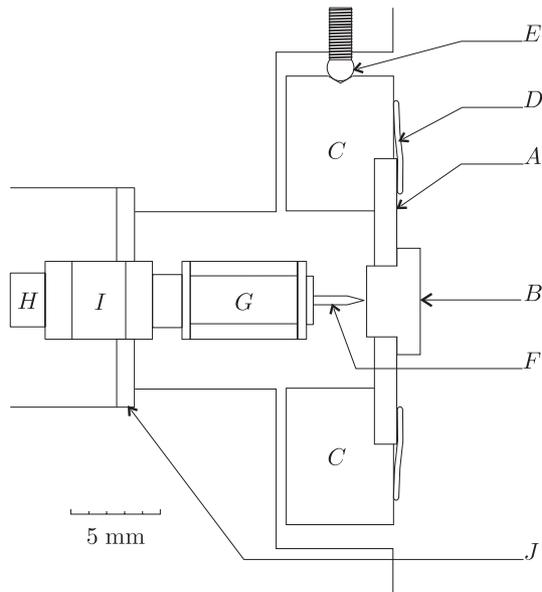


Figure 7.1: Cross-sectional sketch of the Aarhus STM. The sample holder (A) with the sample (B) is clamped against the top-plate of the STM (C) by two springs (D). The top-plate is thermally isolated from the STM housing by three quartz balls (E). The STM tip (F) is mounted on the scanner tube (G), which again is glued onto a shaft (H). Coarse positioning of the tip is accomplished via the inchworm (I). (Adapted from Ref. 7.)

and the measuring resistor ( $R_{\text{in}} = 750 \Omega$ , cf. Fig. 4.7 on page 95). Once the conductance reaches a value close to  $1 G_0$  (corresponding to a monatomic gold contact [8–11]), a large (3 Volt) triangular voltage burst is applied to the contact.

## 7.2 Results

### Current-induced disruptions

In the upper panel of Fig. 7.2, we show the time-evolution of the output voltage  $V_{\text{out}}$  from the  $I$ - $V$  converter in response to a 3 V amplitude, 22 kHz triangular voltage burst superimposed on a 150 mV offset voltage (shown in the inset).  $V_{\text{out}}(t)$  closely follows the function generator signal  $V_{FG}(t)$  until about  $7 \mu\text{s}$  after the onset of the burst, at which it drops to zero. The drop seems abrupt — at least on a time-scale of the relaxation time of the setup, which is about 70 ns in this experiment. In the lower panel, we show the corresponding calculated conductance trace for the voltage trace. As we have seen several times [cf. Figs. 4.9 and 4.13(a) in Chapter 4], the calculated conductance is noisy prior to the burst because only a small fraction of the dynamic range on the oscilloscope is used here. The total conductance is close to  $1 G_0$  independent of the voltage until the break occurs.

The calculated  $I$ - $V$  curve for the burst segment of the measurements in Fig. 7.2 is shown in Fig. 7.3. The segment is almost linear except for a bulge around 1.7 Volt. At a voltage of 1.92 Volt and a current of  $145 \mu\text{A}$ , the break occurs. From the graphs in Figs. 7.2 and 7.3, one is tempted to conclude that the loss of contact is due to the large current running through the gold atom. However, a single measurement is not enough to infer such a conclusion. It could be a coincidence that the break happened during the burst.

### Varying the voltage rate and sample temperature

To further elucidate the underlying mechanism for the breaks, we have performed a series of experiments, varying the temperature of the sample from 140–240 K and using three different burst frequencies (10, 22, and 100 kHz). Changing the burst frequency (at a constant burst amplitude) corresponds to changing the voltage ramp rate (voltage rate hereafter). The chosen burst frequencies correspond to voltage rates of 0.11, 0.25, and  $1.1 \text{ V}/\mu\text{s}$ , respectively.<sup>1</sup>

---

<sup>1</sup>The voltage rate across a  $1 G_0$  contact is about 5% lower than the voltage rate applied from the function generator due to the voltage drop through  $R_{FG}$  and  $R_{\text{in}}$ , cf. Fig. 4.7 on page 95.

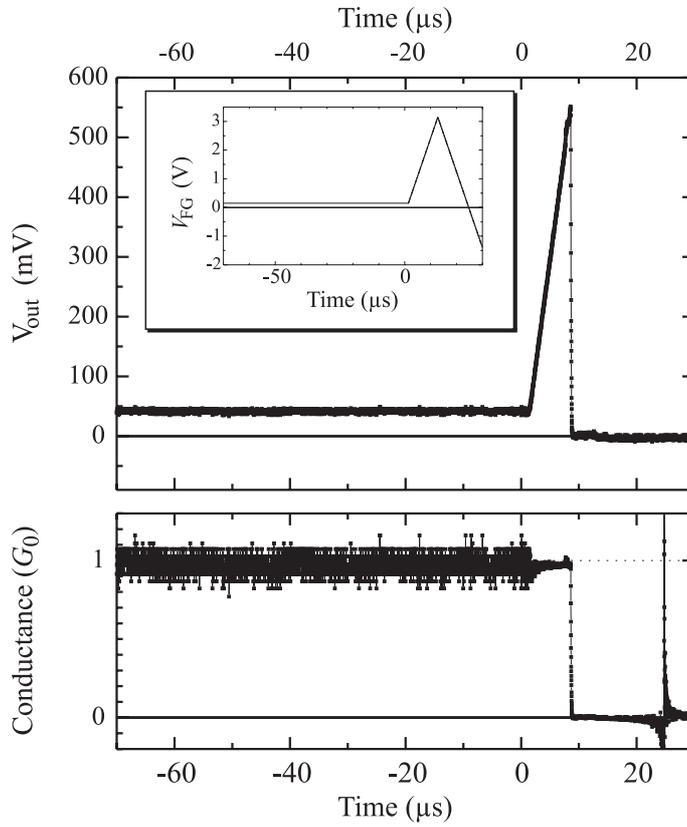


Figure 7.2: Current-induced disruption of a monatomic gold contact. Upper panel: Measured time-evolution of the  $I$ - $V$  converter output voltage under an applied voltage burst (inset). Lower panel: The corresponding calculated time-evolution of the conductance, assuming the burst delay is  $1.481 \mu\text{s}$ . The conductance goes to zero  $7.15 \mu\text{s}$  after the onset of the burst. There are 10 000 data points in each plot (10 ns/point). The temperature of the sample is 162 K.

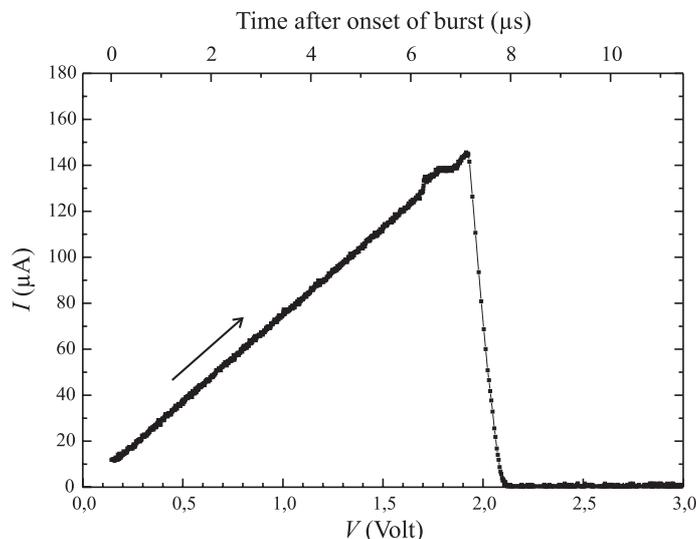


Figure 7.3: Calculated  $I$ - $V$  curve for the burst part of the measurement shown in Fig. 7.2. The lower axis is the voltage drop across the contact, whereas the upper axis is the time that have passed after the onset of the burst. Contact is lost at  $V_{\text{br}} = 1.92$  V and  $I_{\text{br}} = 145$   $\mu\text{A}$ .

### Selecting data corresponding to monatomic disruptions

In a significant number of the traces, we find small or large steps in the conductance prior to the burst indicating that the formed contact is unstable. These traces are disregarded in the analysis. Instead we focus on conductance traces, where the conductance stays close to  $1 G_0$  at least  $40$   $\mu\text{s}$  before the burst.

For contacts which are stable before the onset of the burst, we have found that electrical contact is always lost before the maximum voltage of  $3$  Volt is reached. However, the contacts break in many different ways. Quite often, for instance, the conductance suddenly jumps from  $1 G_0$  up to a value close to  $2 G_0$  during the burst. Shortly thereafter it breaks — either directly or via one or more transitions. In other cases, the conductance goes from  $1 G_0$  to an intermediate subquantum value of  $0.1$ – $0.5 G_0$  before contact is lost, and there are other more complicated types of behavior. Nevertheless, a substantial fraction of the curves are qualitatively similar to Fig. 7.3, i.e., they exhibit a

clean transition from  $1G_0$  to  $0G_0$ , and deviations from  $1G_0$  during the burst are small. Since these transitions can be interpreted unambiguously as due to rupture of monatomic gold contacts, we will focus on these types of breaks in the following.

### The break voltage

For each clean break of a monatomic contact the last data point on the  $I$ - $V$  curve before the sudden drop is found, and the corresponding break voltage  $V_{\text{br}}$  and break current  $I_{\text{br}}$  are registered. For instance,  $V_{\text{br}} = 1.92$  V and  $I_{\text{br}} = 145$   $\mu$ A for the burst shown in Fig. 7.3. A majority of the contacts have a total conductance of  $0.9$ – $1.0G_0$  at the time of break. Therefore one quantity can be estimated from the other by using the relation  $I_{\text{br}} \simeq 0.95G_0V_{\text{br}}$ . We will therefore focus on only one of these quantities in the following, namely  $V_{\text{br}}$ .

In Fig. 7.4 we show recorded values of  $V_{\text{br}}$  as a function of sample temperature for the three voltage rates (upper panel:  $0.11$  V/ $\mu$ s; middle panel:  $0.25$  V/ $\mu$ s; and lower panel:  $1.1$  V/ $\mu$ s). Each panel corresponds to a single day of data taking in the laboratory, where the sample temperature slowly increased after the initial cooling. Thus, the data points are in chronological order from left to right. There are regions on the temperature axis without data points. Each of these gaps correspond to an interruption in the experiments (re-calibrations of the setup, lunch breaks, etc.)

Several interesting observations can be extracted from the graphs.

- For all sample temperatures and voltage rates, we find  $0.58$  V  $< V_{\text{br}} < 2.65$  V. Out of 271 disruptions, none occur close to the beginning of the burst, where the voltage is  $\simeq 0.15$  V. This is a clear proof that almost 100% of the disruptions are induced by the burst.
- $V_{\text{br}}$  only depends very weakly on the sample temperature. If we (tentatively) fit the widely scattered values of  $V_{\text{br}}$  to straight lines (shown in the figure), we find a slope of  $\simeq -2$  mV/K independent of the voltage rate.
- $V_{\text{br}}$  seems to depend on the voltage rate. For break voltages recorded with  $0.11$  V/ $\mu$ s and  $0.25$  V/ $\mu$ s voltage rates, the mean values of the break voltage are almost identical;  $1.46 \pm 0.04$  V and  $1.48 \pm 0.04$  V, respectively. For the  $1.1$  V/ $\mu$ s voltage rate experiments, the mean break voltage increases by about a quarter of a Volt to  $1.71 \pm 0.04$  V.

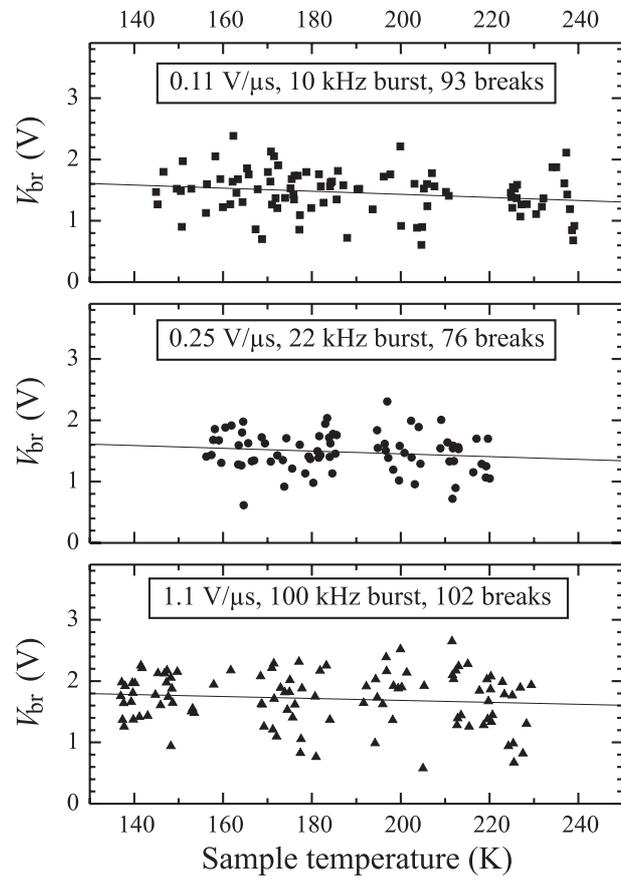


Figure 7.4: Scatter graphs: Measured break voltages for monatomic gold contacts as a function of sample temperature for three different voltage rates. Solid lines: Best linear fits.

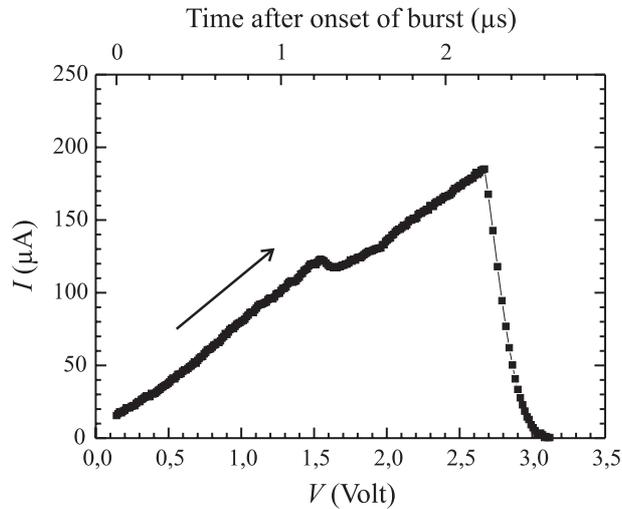


Figure 7.5: Calculated  $I$ - $V$  curve for the burst part of a measurement on a monatomic gold contact formed on a Au(111) crystal cooled to 212 K. The axes have the same meaning as in Fig. 7.3. The voltage rate is  $1.1 \text{ V}/\mu\text{s}$ . The break voltage and current is record-high:  $V_{\text{br}} = 2.7 \text{ V}$  and  $I_{\text{br}} = 190 \mu\text{A}$ .

### Extreme current densities

In a recent experiment, Yanson *et al.* succeeded in passing a current of up to  $80 \mu\text{A}$  through a chain of gold atoms formed at liquid He temperatures [11]. The break current of  $145 \mu\text{A}$  deduced from Fig. 7.3 almost doubles this value, but it is possible to go a little higher. In Fig. 7.5 we show the burst segment of the most enduring contact we have encountered during the experiments. Here the fastest voltage rate of  $1.1 \text{ V}/\mu\text{s}$  is applied.  $2.3 \mu\text{s}$  after the onset of the burst, the current has reached a value of  $180 \mu\text{A}$ . This is the highest current that have ever been measured through a single atom. Such a current corresponds to an extremely high transient current density of  $2.8 \times 10^{11} \text{ A}/\text{cm}^2$ . This is more than five orders of magnitudes larger than typical current densities in present-day interconnects used in microchips [12]. The peak power deposited in the contact region is about  $500 \mu\text{W}$  just before rupture. An electron passes through the atom once every fs at a speed of about  $1000 \text{ km/s}$  (slightly less than the Fermi velocity) separated by an average distance of  $1\text{--}2 \text{ \AA}$ . This would never

have worked out in classical physics! Such extreme current densities are only possible because the monatomic contact is ballistic and intimately connected to large heat reservoirs with good thermal conductance.

### Cumulative break probabilities

Clearly, the break voltages are distributed over a rather wide range of voltages. The mean break voltages given above indicate that  $V_{\text{br}}$  tends to increase with increasing voltage rate. However, the mean value only gives information about the first moment of a distribution, and we may lose information by solely focusing on this quantity. A better alternative is to sort the continuous values of break voltages, bin them into intervals of a fixed size and plot the resulting histograms. The histograms for the different voltage rates can then be compared. However, we still lose information by following such a procedure since the generated histograms will depend on the bin size: Increasing the bin size gives less noise in the histogram due to improved counting statistics, but features are also lost and *vice versa*.

The best solution to this problem consists of generating an unbiased estimator  $S_N(V)$  of the cumulative break probability (CBP) of the underlying distribution from which the break voltages were drawn [13]. We can think of  $S_N(V)$  as an ‘integrated’ break voltage histogram. If the  $N$  break voltages have values  $V_{\text{br}}^i$ ,  $i = 1, \dots, N$ , then  $S_N(V)$  is the function giving the fraction of the measured break voltages which are lower than  $V$ . This function is obviously constant between consecutive  $V_{\text{br}}^i$ ’s and jumps by  $1/N$  at each  $V_{\text{br}}^i$ . The limiting values of  $S_N(V)$  are of course zero when  $V \lesssim 0.15$  V (the offset voltage) and one as  $V \rightarrow \infty$ .

The CBP of monatomic gold contacts for the three voltage rates are shown in Fig. 7.6. It is seen that the CBPs are almost identical for the two lowest voltage rates. In contrast, the CBP for the high  $1.1$  V/ $\mu\text{s}$  voltage rate experiments is shifted to higher voltages by a few tens of a Volt. However, we should not forget that these distributions are constructed from a limited number of data points. The shown distributions are therefore only approximations to the true underlying distributions. Although the  $0.11$  V/ $\mu\text{s}$  and  $0.25$  V/ $\mu\text{s}$  CBPs look almost similar, they do differ somewhat between  $0.7$  and  $1.3$  V. It is difficult to judge by eye whether this is due to a ‘real’ difference between the underlying distributions or random fluctuations caused by the limited number of data points.

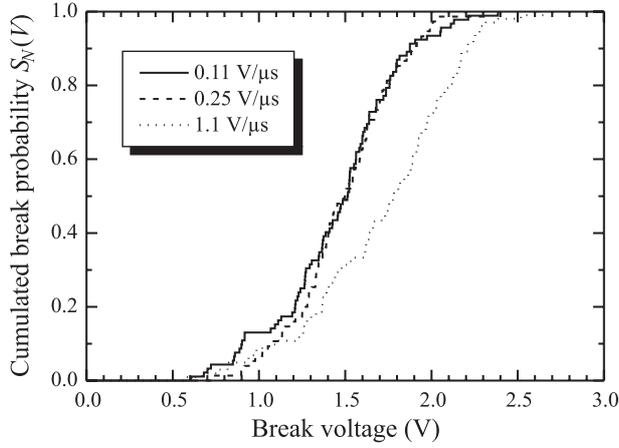


Figure 7.6: Cumulated break probabilities of monatomic gold contacts for three different voltage rates.

### The Kolmogorov-Smirnov test

To better quantify these and other questions, we have used the Kolmogorov-Smirnov (KS) test for comparing the different CBPs (see Appendix D for the mathematical details). The KS test provides an answer to the following question:

Assuming that two CBPs  $S_{N_1}(V)$  and  $S_{N_2}(V)$  are drawn from the same underlying distribution, what is the probability  $P(\text{disagreement} > \text{observed})$  of finding a worse agreement between the CBPs if the experiments are repeated?

In short,  $P$  is calculated from the maximum value of the absolute deviation between  $S_{N_1}(V)$  and  $S_{N_2}(V)$  as well as the number of data points  $N_1$  and  $N_2$  from which the two probability functions are generated. If  $S_{N_1}(V)$  and  $S_{N_2}$  are drawn from the same underlying distribution,  $P$  should be distributed randomly between zero and one. The lower  $P$  is, the worse is the agreement between the two probability functions. If  $P < 0.01$ , it is most likely that the underlying distributions are *different* [13].

In Table 7.1, the KS probability is tabulated for the three different voltage rate combinations. When the break voltage distribution for the fast 1.1 V/ $\mu$ s voltage rate experiments is compared with either of the two other distributions,

$V/\mu\text{s} \leftrightarrow V/\mu\text{s}$	$P$
0.11 $\leftrightarrow$ 0.25	0.76
0.11 $\leftrightarrow$ 1.1	$7 \times 10^{-6}$
0.25 $\leftrightarrow$ 1.1	$3 \times 10^{-5}$

Table 7.1: Kolmogorov-Smirnov probabilities  $P$  between break voltage distributions accumulated from different voltage rates. See text for details.

exceedingly small KS probabilities are found. This indicates with great significance that the underlying distribution for the 1.1 V/ $\mu\text{s}$  experiments is indeed different from the 0.1 and 0.25 V/ $\mu\text{s}$  experiments. In contrast, the KS probability is large for the two smallest voltage rates. As noted in Appendix D, this does *not* automatically imply that the underlying distributions are identical. The KS test can only be used for testing whether two distributions are different.

### 7.3 Discussion

#### Transition state theory

In our experiments, we cannot resolve any structure in the conductance as the contact breaks during a voltage burst as seen in Figs. 7.3 and 7.5. This means that the transition time during rupture is shorter than the experimental relaxation time ( $\simeq 70$  ns). This is not surprising as we expect the transition to happen on a time-scale of the vibrational frequency of the atom (pico-seconds). Since the time-scale for actually breaking the contact (micro-seconds) is much longer than this, we are in a rare-event regime. In this case, the break rate can be described within transition state theory [14]. In this theoretical picture, the transition takes place along a minimum energy path on the potential energy surface between the initial and final states. It follows that the contact break rate  $\Gamma$  is given by the well known Arrhenius rate equation

$$\Gamma = \nu \exp(-E_b/k_B T_{\text{loc}}), \quad (7.1)$$

where the activation barrier  $E_b$  is the energy difference between the maximum energy along the minimum energy path on the potential energy surface (the transition state) and the initial state, and  $T_{\text{loc}}$  is the local temperature in the contact. The prefactor  $\nu$  is related to the curvatures of the potential energy

surface in the initial and transition states and is expected to be of the order of the phonon frequency ( $\nu \sim 10^{12} \text{ s}^{-1}$ ). Within the harmonic approximation,  $\nu$  is independent of the temperature. The activation barrier depends on the contact geometry, which is unknown. This clearly complicates the physical picture.

When a voltage is applied to the contact, the activation barrier is expected to go down due to electron wind forces [3, 4], bond weakening [4, 5], and possibly electrostatic forces [6]. At the same time, the local temperature in the contact rises due to the current [2]. Thus, the exponent in Eq. (7.1) is damped with increasing voltage and the break rate increases. As the voltage exceeds 2.3 V,  $\Gamma$  goes well beyond the typical time-scale for the voltage ramp ( $\sim 10^6 \text{ s}^{-1}$ ). This can clearly be seen from the cumulated break probabilities in Fig. 7.6, where nearly all contacts have broken at this voltage.

### Local temperature prior to the burst

Due to the design of our STM and the rather large constant bias voltage applied across the contact and  $R_{\text{in}}$ , the local temperature in the contact region  $T_{\text{loc}}$  prior to the voltage burst is quite uncertain. When the tip is still deeply indented into the sample, the voltage drop across the contact is very small because it is placed in series with  $R_{\text{in}}$ , which has a much larger resistance than the contact at this point. Thus the effects of Ohmic heating in the contact region are expected to be very small at this stage of the contact formation. We expect that  $T_{\text{loc}}$  is between sample and room temperatures. Since the thermal conductivity of gold increases with decreasing temperature and the geometry of the flat sample is better suited for conducting heat than a cone-shaped tip, we also expect that  $T_{\text{loc}}$  is somewhat closer to the sample temperature as compared to RT.

Things get increasingly complicated when the tip is retracted. As the contact conductance decreases, the voltage drop across the contact increases towards  $\simeq 150 \text{ mV}$ , and Ohmic heating may become important. In addition, the thermal conductivity of the contact decreases in response to a decreasing cross-sectional area, and a large temperature gradient probably builds up in the contact region [15]. In the end, when the smallest cross-section consists of only a few atoms, the temperature on the tip side of the contact rapidly increases. The rate of increase is a competition between how rapidly the contact thins down and the rate of heat transfer between the tip and the heat reservoirs.

### Rate of heating

The electron wind and bond weakening effects set in instantaneously (on an experimental time-scale) when the burst is applied. This is not necessarily the

case for the local temperature as it depends on the time it takes for the system to reach local thermal equilibrium. At the moment it is not known whether it can be assumed that the local temperature reaches an instantaneously steady state value or whether it lags behind during the voltage burst. This possibility of an integrated effect on the time-evolution of  $T_{\text{loc}}$  clearly complicates the interpretation of the results.

### Local temperature at finite bias

Todorov recently calculated the local temperature in an atomic-sized contact in the low temperature limit within a simple 1s tight binding model [2]. At zero ambient temperatures, it was estimated that the local temperature would rise by something like 60 K under a constant applied voltage of 0.1 V. In the model, it was assumed that the heat capacity  $c_V$  is proportional to  $T^3$ . That is, the temperature should be well below the Debye temperature, which is 170 K for Au [16]. When this assumption is valid,  $T_{\text{loc}} \propto \sqrt{V}$  [2]. At finite temperatures  $T_0$  of the environment (but still in the  $c_V \propto T^3$  regime), the local temperature in the contact is given by [17]

$$T_{\text{loc}}^4 = T_{\text{ind}}^4 + T_0^4, \quad (7.2)$$

where  $T_{\text{ind}}$  is the current-induced temperature due to the presence of a current (identical to the zero ambient temperature expression for  $T_{\text{loc}}$ ).

The validity of Todorov's formulas in our experiments is questionable. First of all, the temperature of the environment is comparable to the Debye temperature of Au or higher. Secondly, the applied voltages are typically a factor of 15 larger than the estimate leading to a rise in temperature of 60 K. Thus,  $T_{\text{loc}}$  is probably around room temperature or larger. Clearly, the current understanding of current-induced local heating in the contact is still in its infancy, and more theoretical and experimental work is needed in the future.

## 7.4 Outlook and conclusion

Current-induced mechanical deformations of atomic-sized metal contacts is a rather unexplored field. Until now only a few experimental studies have addressed the issue of current- or voltage-induced effects. These include an investigation of the switch rate in two-level fluctuations [18] and the dependence of mechanical hysteresis loops [19] as a function of bias voltage. Also from a theoretical point of view useful results are only beginning to emerge [2, 4, 5].

The preliminary results presented in this Chapter represents the first controlled study of current-induced disruptions of atomic-sized contacts. The break voltage of monatomic contacts is shown to be weakly dependent on the sample temperature and a clear correlation between the voltage rate and the break voltage is also demonstrated by means of the Kolmogorov-Smirnov test. From the data obtained so far, we conclude that the current-induced break rate is of the order of a few microseconds for an applied voltage of 1.5–2.0 Volt. In the most extreme case a current of up to  $190 \mu\text{A}$  was measured through a single atom, which more than doubles the previous ‘record’ [11].

We are right now planning a new series of experiments, where we change the burst waveform from triangular to square. Then, instead of recording the break voltage, we will measure the life-time of contacts for varying amplitudes of the burst. The outcome of such experiments are more easily compared with the Arrhenius form for the contact break rate appearing in Eq. (7.1). From these experiments we hope to get a better understanding of how the local contact temperature depends on the applied voltage, and whether integrated heating effects are important on our experimental time-scales.

## References

- [1] J. A. Torres, E. Tosatti, A. Dal Corso, F. Ercolessi, J. J. Kohanoff, F. D. Di Tolla, and J. M. Soler, “The puzzling stability of monatomic gold wires,” *Surf. Sci.* **426**, L441–L446 (1999).
- [2] T. N. Todorov, “Local heating in ballistic atomic-scale contacts,” *Philos. Mag. B* **77**, 965–973 (1998).
- [3] H. Yasuda and A. Sakai, “Conductance of atomic-scale gold contacts under high-bias voltages,” *Phys. Rev. B* **56**, 1069–1072 (1997).
- [4] T. N. Todorov, J. Hoekstra, and A. P. Sutton, “Current-induced forces in atomic-scale conductors,” *Philos. Mag. B* **80**, 421–455 (2000).
- [5] A. P. Sutton and T. N. Todorov, “Mechanical and electrical properties of atomic-scale metallic contacts — experiment, simulation and future directions,” In *Boundaries and Interfaces in Materials: The David A. Smith Symposium*, R. C. Pond, W. A. T. Clark, A. H. King, and D. B. Williams, eds., pp. 189–197 (Warrendale, Pennsylvania, 1998).
- [6] O. Hansen, J. T. Ravnkilde, U. Quaade, K. Stokbro, and F. Grey, “Field-Induced Deformation as a Mechanism for Scanning Tunneling Microscopy Based Nanofabrication,” *Phys. Rev. Lett.* **81**, 5572–5575 (1998).
- [7] L. Olesen, “Experimental Investigation of the Junction in a Scanning Tunneling Microscope,” PhD thesis, University of Aarhus, 1996.
- [8] G. Rubio, N. Agrait, and S. Vieira, “Atomic-Sized Metallic Contacts: Mechanical Properties and Electronic Transport,” *Phys. Rev. Lett.* **76**, 2302–2305 (1996).
- [9] M. Brandbyge, M. R. Sørensen, and K. W. Jacobsen, “Conductance eigenchannels in nanocontacts,” *Phys. Rev. B* **56**, 14956–14959 (1997).
- [10] H. Ohnishi, Y. Kondo, and K. Takayanagi, “Quantized conductance through individual rows of suspended gold atoms,” *Nature* **395**, 780–783 (1998).
- [11] A. I. Yanson, G. R. Bollinger, H. E. van den Brom, N. Agrait, and J. M. van Ruitenbeek, “Formation and manipulation of a metallic wire of single gold atoms,” *Nature* **395**, 783–785 (1998).

- [12] Semiconductor Industry Association, “The International Technology Roadmap for Semiconductors.”, 1999, available on the web at [http://notes.sematech.org/1999.SIA\\_Roadmap/Home.htm](http://notes.sematech.org/1999.SIA_Roadmap/Home.htm).
- [13] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal* (Cambridge University Press, Cambridge, 1989).
- [14] G. Wahnström, “Rate Equations, Rate Constants, and Surface Diffusion,” in *Interaction of atoms and molecules with solid surfaces, Physics of solids and liquids*, V. Bortolani, N. H. March, and M. P. Tosi, eds., (Plenum, New York, 1990), Chap. 16, pp. 529–566.
- [15] B. Ludoph and J. M. van Ruitenbeek, “Thermopower of atomic-size metallic contacts,” *Phys. Rev. B* **59**, 12290–12293 (1999).
- [16] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [17] T. N. Todorov, (private communication).
- [18] C. J. Muller, J. M. van Ruitenbeek, and L. J. de Jongh, “Conductance and Supercurrent Discontinuities in Atomic-Scale Metallic Constrictions of Variable Width,” *Phys. Rev. Lett.* **69**, 140–143 (1992).
- [19] H. E. van den Brom, A. I. Yanson, and J. M. van Ruitenbeek, “Characterization of individual conductance steps in metallic quantum point contacts,” *Physica B* **252**, 69–75 (1998).



## Chapter 8

# Outlook

What are the most promising avenues of future research in the field of atomic-sized metal contacts? This is not an easy question to answer but let me in the following give a few examples of where the field might be heading.

### **The Interplay between electronic structure, conductance, and mechanical properties**

There has recently been an increased focus on the possible interplay between the electronic structure within the contact, the tensile force as well as the conductance [1–13]. The simplest jellium models based on either non-interacting [1–4] or first-principles [5, 6] calculations of energies, conductance and tensile forces have shown that the closing of individual conductance channels is correlated with large fluctuations in the tensile force, which are of the same order of magnitude as the forces measured in experiments [14, 15]. Although the simple jellium models elude to possibly important interconnections between electronic structure, mechanical properties, and the conductance, they probably overestimate the effect because they completely neglect the discrete atomic structure of the material. There have been attempts to simulate more realistic contact geometries [7–10], but so far these systems have been too small to answer these questions in a satisfactory way.

An indication of a possible interplay between electronic structure and conductance has been observed in many-atom contacts of Na [13]. Here, a series of peaks in the histogram of conductances up to  $120 G_0$  have been observed and

interpreted as being due to the formation of closed electronic shells resembling those known from metal clusters. It will be interesting to investigate whether the electronic structure has observable effects on other and less free-electron-like systems in the future.

## Electronic transport at finite bias

As is evident from Chapters 4–7, there are still several unsolved puzzles regarding  $I$ - $V$  curves of atomic-sized contacts. There is a need for improved theoretical models for calculating the conductance self-consistently at finite bias. More realistic models of current-induced effects such as heating, electron-wind forces, bond weakening, and electrostatic forces are also needed.

In Chapter 5, we found that clean Au  $I$ - $V$  curves are almost linear. It could be interesting to chemisorb different kinds of molecules on the contacts and investigate how this would influence the  $I$ - $V$  curves. The first theoretical steps in this direction have already been taken [16].

## One-dimensional physics in atomic chains

At liquid He temperatures, it is possible to form long chains (at least four atoms long) that are stable for several hours [17]. Such chains are interesting model systems for 1D metals. Due to their reduced dimensionality, they may exhibit new exciting properties that are unknown from ordinary metals.

## References

- [1] C. A. Stafford, D. Baeriswyl, and J. Bürki, “Jellium Model of Metallic Nanocoherence,” *Phys. Rev. Lett.* **79**, 2863–2866 (1997).
- [2] J. M. van Ruitenbeek, M. H. Devoret, D. Esteve, and C. Urbina, “Conductance quantization in metals: The influence of subband formation on the relative stability of specific contact diameters,” *Phys. Rev. B* **56**, 12566–12572 (1997).
- [3] S. Blom, H. Olin, J. L. Costa-Krämer, N. García, M. Jonson, P. A. Serena, and R. I. Shekter, “Free-electron model for mesoscopic force fluctuations in nanowires,” *Phys. Rev. B* **57**, 8830–8833 (1998).
- [4] F. Kassubek, C. A. Stafford, and H. Grabert, “Force, charge and conductance of an ideal metallic nanowire,” *Phys. Rev. B* **59**, 7560–7574 (1999).
- [5] C. Yannouleas and U. Landman, “On Mesoscopic Forces and Quantized Conductance in Model Metallic Nanowires,” *J. Phys. Chem. B* **101**, 5780–5783 (1997).
- [6] C. Yannouleas, E. N. Bogachek, and U. Landman, “Energetics, forces, and quantized conductance in jellium-modeled metallic nanowires,” *Phys. Rev. B* **57**, 4872–4882 (1998).
- [7] R. N. Barnett and U. Landman, “Cluster-derived structures and conductance fluctuations in nanowires,” *Nature* **387**, 788–790 (1997).
- [8] U. Landman, R. N. Barnett, and W. D. Luedtke, “Nanowires: size evolution, reversibility, and one-atom contacts,” *Z. Phys. D* **40**, 282–287 (1997).
- [9] G. Taraschi, J.-L. Mozos, C. C. Wan, H. Guo, and J. Wang, “Structural and transport properties of aluminum atomic wires,” *Phys. Rev. B* **58**, 13138–13145 (1998).
- [10] A. Nakamura, M. Brandbyge, L. B. Hansen, and K. W. Jacobsen, “Density Functional Simulation of a Breaking nanowire,” *Phys. Rev. Lett.* **82**, 1538–1541 (1999).
- [11] C. A. Stafford, J. Bürki, and D. Baeriswyl, “Comment on “Density Functional Simulation of a Breaking Nanowire”,” *Phys. Rev. Lett.* **84**, 2548 (2000).

- [12] A. Nakamura, M. Brandbyge, L. B. Hansen, and K. W. Jacobsen, “Reply to ‘Comment on ‘Density Functional Theory Simulation of a Breaking nanowire’”,” *Phys. Rev. Lett.* **84**, 2549 (2000).
- [13] A. I. Yanson, I. K. Yanson, and J. M. van Ruitenbeek, “Observation of shell structure in sodium nanowires,” *Nature* **400**, 144–146 (1999).
- [14] G. Rubio, N. Agrait, and S. Vieira, “Atomic-Sized Metallic Contacts: Mechanical Properties and Electronic Transport,” *Phys. Rev. Lett.* **76**, 2302–2305 (1996).
- [15] A. Stalder and U. Dürig, “Study of yielding mechanics in nanometer-sized Au contacts,” *Appl. Phys. Lett.* **68**, 637–639 (1996).
- [16] H. Häkkinen, R. N. Barnett, and U. Landman, “Gold Nanowires and their Chemical Modifications,” *J. Phys. Chem. B* **103**, 8814–8816 (1999).
- [17] A. I. Yanson, G. R. Bollinger, H. E. van den Brom, N. Agrait, and J. M. van Ruitenbeek, “Formation and manipulation of a metallic wire of single gold atoms,” *Nature* **395**, 783–785 (1998).

## Kapitel 9

# Dansk resumé

Jeg vil i dette kapitel forsøge at forklare hovedresultaterne i min ph.d.-afhandling på et niveau der er forståeligt for den fysikinteresserede ikke-ekspert. Afhandlingens titel er på dansk: *Elektriske egenskaber af metalkontakter bestående af få atomer*. Groft sagt har mine studier gået ud på at måle den elektriske modstand af minimale punktkontakter dannet mellem to stykker metal. For det meste har metallet været guld, men også sølv og kobber samt visse overgangsmetaller er blevet studeret. Formålet med mit arbejde har været at bidrage til forståelsen af de grundlæggende fysiske egenskaber af minimale metalkontakter. På langt sigt kan man håbe at mit og andres arbejde som helhed kan finde nytte i praktiske anvendelser. I kapitel 1 forklarer jeg blandt andet hvordan mine resultater har relevans for udviklingen af bedre ledningsforbindelser mellem transistorerne i integrerede kredsløb.

### Teori

I kapitel 2 gennemgås den grundlæggende teori for beregning af elektrisk modstand i punktkontakter/ledninger. Det viser sig at beregningsmetoden helt afhænger af, hvor stor kontakten er i forhold til de to vigtige længdeskalaer kaldet *den middelfri vejlængde* og *Fermibølgelængden*. Dette er illustreret i figur 2.1 på side 7. Den middelfri vejlængde er et mål for hvor langt en elektron kan bevæge sig inden i metallet uden at kolliderer. For gode elektriske ledere som ædelmetallerne kobber, sølv og guld er den middelfri vejlængde 20–50 nm.

Vi tænker ofte på elektroner som punktpartikler, men kvantemekanikken

fortæller os, at vi også betragte dem som bølger.<sup>1</sup> Man kan følgelig tilskrive de strømbærende elektroner en bølgelængde, kaldet Fermibølgelængden. For de førnævnte ædelmetaller er denne bølgelængde omkring 5 Å.

For punktkontakter bestående af blot nogle få atomer er diameteren meget mindre end den middelfri vejlængde (en ballistisk punktkontakt) og sammenlignelig med Fermibølgelængden (en kvantepunktkontakt). For sådanne kontakter udregnes modstanden (eller den reciprokke størrelse kaldet ledningsevnen) vha. *Landauer-Büttiker* formelen [lign. (2.21) på side 21]. I denne formel optræder størrelsen  $G_0 = 2e^2/h$ , hvor  $e$  er elementarladningen og  $h$  er Plancks konstant.  $G_0$  kaldes også for den fundamentale kvanteenhed for ledningsevne og svarer til en elektrisk modstand på ca. 13 kΩ. Den anden vigtige ingrediens i formelen er kvantemekaniske transmissionsamplituder for elektrontilstande, som relaterer til transmissionssandsynligheden af de forskellige egensvingningstilstande af elektronens bølgefunktion [de såkaldte *egenkanaler*, se lign. (2.30) på side 27]. For en typisk enatomar guldkontakt er transmissionen domineret af en enkelt egenkanal med en transmissionssandsynlighed tæt på 100%. En enatomar guldkontakt har derfor en modstand der er tæt på 13 kΩ.

## Resultater

Hvis man skal studere de elektriske egenskaber af metalliske punktkontakter bestående af få atomer, er det selvfølgelig en forudsætning at man er i stand til at fabricere dem. I kapitel 3 er beskrevet en særlig simpel fremstillingsmetode som tager udgangspunkt i at afbryde en kontakt i et kommercielt eller hjemmebygget elektromekanisk relæ. Lige inden kontakten endeligt afbrydes, vil forbindelsen mellem de to elektroder i et kort øjeblik — af størrelsesordenen 1/10000 sekund — udelukkende bestå af en enkelt punktkontakt. Denne strækkes og gøres tyndere, mens kontaktankeret i relæet bevæger sig væk fra modelektroden for til sidst at knække. Under kontaktafbrydelsen kan vi måle ledningsevnen.

Eksempler herpå er vist i figur 3.3 på side 51, som stammer fra målinger på et kommercielt relæ der har guldbelagte elektroder. Af figuren ses at ledningsevnen for det meste har en ledningsevne tæt på et helt antal  $G_0$ . Man siger at ledningsevnen er *kvantiseret*, hvilket skyldes at elektronbølgerne har en tendens til at danne et antal veldefinerede stående bølger på tværs af kontakten. Kvantiseringen er dog ikke helt skarp, hvilket ses tydeligt hvis man beregner

---

<sup>1</sup>Dette er helt analogt til vores forstilling om lys; en regnbue afspejler f.eks. tydeligt bølgeegenskaberne, men lys kan også opfattes som partikler, kaldet fotoner.

sandsynlighedsfordelingen for ledningsevnen akkumuleret fra et stort antal kontaktabbrydelser. En sådan fordeling kaldes også for et ledningsevnehistogram, se f.eks. figur 3.4(b) på side 53. Man bemærker at kvantiseringen gradvist bryder sammen, jo større ledningsevnen er.

Laver man lignende afbrydelser af sølv- og kobberkontakter i luft, er ledningsevnen også kvantiseret, men ikke så tydeligt som for guld. Kvantiseringen forsvinder helt for et udvalg af overgangsmetallerne. Resultaterne fortolkes i termer af metallernes forskellige reaktivitet, elektronstruktur og mekaniske egenskaber.

I kapitel 3 beskrives tillige en instrumentel fejl kaldet differentiell ikke-linearitet, der fører til støj og andre fejl i ledningsevnehistogrammerne. Vi angiver en procedure til at korrigere de indsamlede data så støjen stort set forsvinder. Dette fremgår af figur 3.4 på side 53, der viser histogrammet før (a) og efter (b) korrektionen. Denne procedure er sidenhen taget i brug af flere andre grupper inden for feltet.

Ud over relæeksperimenterne har vi foretaget studier af guldkontakter dannet vha. et såkaldt skanning tunnel mikroskop (STM). Et STM består af en meget spids metalnål (en tip) hvis bevægelse kan kontrolleres meget fint vha. nogle piezoelektriske komponenter, der deformerer på velkontrolleret vis når de påtrykkes en spænding. Normalt benyttes et STM til at afbilde overfladen af elektrisk ledende materialer ved at føre tippet meget tæt hen over overfladen og aftaste dens bevægelse (se figur 4.12 på side 107 for eksempler på STM-billeder af en guldoverflade). Man kan dog også ‘dyppe’ tippet ned i overfladen og derved danne en punktkontakt på kontrolleret vis. Ved samtidig at placere STMet i et ultrahøjt vakuum (UHV) kammer udstyret med teknikker til rensning af overfladen kan vi sikre os at vi får dannet *rene* guldkontakter.

Almindelige ledninger og modstande adlyder Ohms lov  $V = RI$ , hvor  $V$  er spændingen,  $R$  er den konstante modstand, og  $I$  er strømmen. Der er dog ingen garanti for at f.eks. en kvantepunktkontakt adlyder Ohms lov; faktisk er der blevet publiceret adskillige målinger der viser at loven ikke gælder for stabile guldkontakter dannet i luft.

Med vores opstilling finder vi at rene kontakter dannet i UHV bryder spontant på en tidsskala af milisekunder, når de er tilstrækkeligt små. Hvis vi skal teste Ohms lov på disse metastabile kontakter er vi derfor nødt til at kunne udmåle kontaktens strøm-spændings ( $I$ - $V$ ) karakteristik meget hurtigt. Opbygningen af en sådan opstilling er beskrevet i kapitel 4. Med dette udstyr kan vi måle en  $I$ - $V$  karakteristik i løbet af blot  $10 \mu s$ , hvilket er fire til syv størrelsesordener hurtigere end i alle tidligere eksperimenter.

Et eksempel på en  $I$ - $V$  karakteristik af en guldkontakt der spontant afbrydes,

er vist i figur 4.13(b) på side 108.  $I$ - $V$  karakteristikken er stykkevis lineær (altså Ohmsk). Desuden svarer hældningen til kvantiserede værdier af ledningsevnen, hvilket klart viser, at vi har at gøre med en kvantepunktkontakt. Sådanne  $I$ - $V$  karakteristikker er aldrig før blevet målt for guld-kvantepunkt-kontakter, og er med få undtagelser i tilsyneladende modstrid med alle tidligere eksperimenter.

Et omfattende studium af guld  $I$ - $V$  karakteristikker er præsenteret i kapitel 5, hvor forskellige overfladepræpareringsmetoder og kontaktdannelsessteknikker er blevet afprøvet. Her er det vist (se figur 5.3 på side 121) at  $I$ - $V$  karakteristikkerne er Ohmske, hvis kontakterne er dannet på en overflade der er renset i UHV, *eller* kontakterne er dannet ved såkaldte hårde indtastninger.

Hvis kontakterne derimod dannes ved bløde indtastninger på en overflade der er renset i luft, får man ikke-lineære  $I$ - $V$  karakteristikker der ligner tidligere eksperimenter. Vores eksperimenter indikerer at urenheder i kontaktområdet medfører at kontakterne bliver stabile *samt* at  $I$ - $V$  karakteristikkerne bliver ikke-lineære. Vi konkluderer herudfra at (rene) guld-kvantepunkt-kontakter (stort set) er Ohmske.

Vores målinger og konklusioner er understøttet af beregninger (se figur 5.6 på side 125) af  $I$ - $V$  karakteristikker for en- og to-atomare guldkontakter, idet disse er (næsten) lineære. Vi foreslår desuden en sandsynlig forklaring på de ikke-lineære  $I$ - $V$  karakteristikker: Det er sandsynligt at der spontant adsorberes et ultratyndt lag af organiske molekyler på guldoverfladen når den renses i luft. Når tippen herefter indenteres blødt i overfladen, klemmes det dielektriske lag af molekyler mellem tippen og overfladen som vist i figur 6.1 på side 142. Dette isolerende lag virker som en tunnelbarriere for elektronerne.

Dette fysiske billede forfølges i kapitel 6, hvor vi opstiller en simpel fysisk model for tunnelbarrieren som lader sig løse eksakt. Ud fra modellen kan vi beregne teoretiske  $I$ - $V$  karakteristikker baseret på tre parametre: tykkelsen af laget, kontaktarealet og barrierehøjden. De beregnede karakteristikker er i god overensstemmelse med de eksperimentelle kurver som vist i figur 6.9 på side 169.

I kapitel 7 viser vi de foreløbige resultater af en ny type eksperimenter, hvor vi vha. elektrisk strøm kan inducere ændringer i kontaktgeometrien af guldkontakter betående af få atomer. Vi fokuserer på studiet af strøminducerede brud af enatomare kontakter. I figur 7.5 på side 187 er f.eks. vist  $I$ - $V$  karakteristikken af en enatomar guldkontakt mens spændingen over atomet øges til 2–3 Volt i løbet af nogle få  $\mu$ s. Kontakten brydes pludseligt, men inden da er strømmen gennem atomet nået op på  $190 \mu$ A. Dette er den største strøm, der nogensinde er blevet målt gennem et atom. Hvis man tager atomets meget lille tværsnitsareal i betragtning, svarer det til en strømtæthed på omtrent  $3 \times 10^{11}$  A/cm<sup>2</sup>, hvilket er mere end fem størrelsesordener højere end de ellers meget store strømtætheder

man kan finde i de små ledningsforbindelser i integrerede kredsløb. Afhandlingen afrundes i kapitel 8, hvor der gives nogle bud på den fremtidige udvikling inden for feltet.



# Appendix A

## The Maxwellian Orifice

### A.1 The oblate spheroidal coordinate system

It is not convenient to solve Laplace's equation analytically for the orifice geometry shown in Fig. 2.2 on page 13 in the usual cartesian coordinate system. Instead we will use the more exotic oblate spheroidal coordinate system  $(\xi, \eta, \phi)$ , which is a curvilinear orthogonal system [1, 2]. The oblate spheroidal coordinates are related to the usual cartesian coordinates  $(x, y, z)$  by the formulas [2]

$$x = a \cosh \xi \cos \eta \cos \phi, \quad (\text{A.1})$$

$$y = a \cosh \xi \cos \eta \sin \phi, \quad \text{and} \quad (\text{A.2})$$

$$z = a \sinh \xi \sin \eta, \quad (\text{A.3})$$

where

$$\xi \geq 0, \quad -\pi/2 \leq \eta \leq \pi/2, \quad \text{and} \quad 0 \leq \phi < 2\pi, \quad (\text{A.4})$$

and  $a$  is a length, which we set equal to the radius of the orifice. Clearly,  $\xi$  is a measure of the distance from the orifice. When  $\xi = 0$ , we are in the  $z = 0$  plane in the orifice. When  $\eta$  is positive, we are situated in the upper electrode and *vice versa*. As  $|\eta| \rightarrow \pi/2$ , we approach the  $z$ -axis. The last coordinate  $\phi$  is just the aximuthal angle.

## A.2 Symmetry and boundary conditions

Since the orifice geometry has rotational symmetry around the  $z$ -axis, the electrostatic potential will not depend on  $\phi$ . Thus we may in this problem combine Eqs. (A.1) and (A.2) to

$$r_{\perp} = \sqrt{x^2 + y^2} = a \cosh \xi \cos \eta. \quad (\text{A.5})$$

We now apply bias voltages  $U/2$  ( $-U/2$ ) infinitely far away from the orifice in the  $z > 0$  ( $z < 0$ ) half spaces. By symmetry the potential in the  $z = 0$  plane will then be zero for  $\rho < a$ , and we will in the following only consider the upper halfspace ( $0 < \eta \leq \pi/2$ ). Here, the Dirichlet boundary conditions are

$$V = \begin{cases} U/2 & \text{for } \xi \rightarrow \infty, \\ 0 & \text{for } \xi = 0. \end{cases} \quad (\text{A.6})$$

The Neumann boundary conditions  $\partial V / \partial n = 0$  for  $z \rightarrow 0^+$  and  $\rho > a$  correspond to

$$\frac{\partial V}{\partial \eta} = 0 \text{ for } \eta \rightarrow 0^+, \quad (\text{A.7})$$

whereby we have formally stated the boundary conditions needed to find a unique solution to Laplace's equation. From the Neumann boundary condition in Eq. (A.7), we see that the potential is independent of  $\eta$  at the diaphragm separating the two electrodes. This could indicate that  $V$  is independent of  $\eta$  everywhere inside the electrodes. We will therefore pursue a solution to Laplace's equation, which only depends on  $\xi$

$$V(\xi, \eta, \phi) = V(\xi). \quad (\text{A.8})$$

## A.3 Solving Laplace's equation

If the ansatz in Eq. (A.8) is correct, the only term in the expression for the Laplace operator, which has to be considered is the one which depends on derivatives with respect to  $\xi$  [2]

$$\begin{aligned} \nabla^2 V &= \frac{1}{a^2 (\sinh^2 \xi + \sin^2 \eta) \cosh \xi} \\ &\quad \times \frac{\partial}{\partial \xi} \left( \cosh \xi \frac{\partial V}{\partial \xi} \right) + \dots \end{aligned}$$

Since  $\nabla^2 V = 0$  for all values of  $\xi$  and  $\eta$ , this implies that

$$\cosh \xi \frac{\partial V}{\partial \xi} = C, \quad (\text{A.9})$$

where  $C$  is a constant which will be determined later using the Dirichlet boundary conditions in Eq. (A.6). Proceeding to find  $V(\xi)$ , we get [2]

$$V(\xi \rightarrow \infty) - V(\xi) = C \int_{\xi}^{\infty} \frac{d\xi'}{\cosh(\xi')} = 2C [\arctan(\exp\{\xi'\})]_{\xi}^{\infty}.$$

Using Eq. (A.6), we can substitute  $V(\xi \rightarrow \infty)$  with  $U/2$ . By rearranging the terms, we then get

$$V(\xi) = \frac{U}{2} + 2C \left[ \arctan(\exp\{\xi\}) - \frac{\pi}{2} \right],$$

from which we see that

$$V(0) = \frac{U}{2} + 2C \left[ \frac{\pi}{4} - \frac{\pi}{2} \right] = \frac{U}{2} - C \frac{\pi}{2}.$$

According to Eq. (A.6), the Dirichlet boundary conditions require that  $V(0) = 0$ . This implies that

$$C = \frac{U}{\pi}. \quad (\text{A.10})$$

The unique solution to Laplace's equation is therefore<sup>1</sup>

$$V(\xi) = U \left[ \frac{2}{\pi} \arctan(\exp\{\xi\}) - \frac{1}{2} \right]. \quad (\text{A.11})$$

## A.4 Calculating the conductance

Since the potential only depends on  $\xi$ , the gradient will be parallel to the direction of increase of  $\xi$  defined by the unit vector  $\hat{\xi}$ . The electric field will therefore

<sup>1</sup>At first, this result seems to be at variance with the formula

$$V(\xi) = U \left[ \frac{1}{2} - \frac{2}{\pi} \arctan(1/\xi) \right],$$

which is normally given in the literature [3, 4]. However,  $\xi$  is defined differently in those references. If this is taken into account, it can be shown that the two expressions are identical.

be given by  $\mathbf{E} = -\nabla_{\xi} V \hat{\boldsymbol{\xi}}$ , where  $\nabla_{\xi}$  is the  $\xi$  component of the gradient, given by [1]

$$\nabla_{\xi} = \frac{1}{a\sqrt{\sinh^2 \xi + \sin^2 \eta}} \frac{\partial}{\partial \xi}. \quad (\text{A.12})$$

From Eqs. (A.9) and (A.10), we immediately see that

$$\frac{\partial V}{\partial \xi} = \frac{U}{\pi \cosh \xi},$$

and we find the electrical field

$$\mathbf{E}(\xi, \eta) = -\frac{U}{\pi a \cosh \xi \sqrt{\sinh^2 \xi + \sin^2 \eta}} \hat{\boldsymbol{\xi}},$$

and the corresponding current density from Ohm's law

$$\mathbf{j}(\xi, \eta) = -\frac{\sigma U}{\pi a \cosh \xi \sqrt{\sinh^2 \xi + \sin^2 \eta}} \hat{\boldsymbol{\xi}}.$$

The next step is to calculate the current flux through a surface placed between the electrodes. For convenience we chose the  $\xi = 0$  surface, which is a disc of radius  $a$ , covering the orifice between the electrodes. At this surface, the unit vector  $\hat{\boldsymbol{\xi}}$  points along the  $z$ -axis. Thus, the  $z$  component of the current density is

$$j_z(0, \eta) = -\frac{\sigma U}{\pi a} \frac{1}{\sin \eta}.$$

Using Eq. (A.5), we may also express  $j_z$  in terms of  $r_{\perp}$  as

$$j_z(\rho) = -\frac{\sigma U}{\pi} \frac{1}{\sqrt{a^2 - r_{\perp}^2}}. \quad (\text{A.13})$$

Next, the current can be found

$$\begin{aligned} I &= \int_0^{a^2} |j_z(r_{\perp})| \pi dr_{\perp}^2 \\ &= \sigma a U \int_0^1 \frac{dx}{\sqrt{1-x}} \\ &= 2\sigma a U, \end{aligned}$$

and finally, we obtain the conductance  $G_M = I/U$  of the Maxwellian orifice

$$\boxed{G_M = \sigma d}, \quad (\text{A.14})$$

where  $d = 2a$  is the diameter.

## References

- [1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, International series in pure and applied physics* (McGraw-Hill, New York, 1953).
- [2] M. R. Spiegel, *Mathematical handbook of formulas and tables, Schaum's Outline Series* (McGraw-Hill, New York, 1992).
- [3] G. Wexler, "The size effect and the non-local Boltzmann transport equation in orifice and disk geometry," *Proc. Phys. Soc. London* **89**, 927–941 (1966).
- [4] A. G. M. Jansen, A. P. van Gelder, and P. Wyder, "Point-contact spectroscopy in metals," *J. Phys. C* **13**, 6073–6118 (1980).



## Appendix B

# The Dilogarithm

The dilogarithm  $\text{Li}_2(z)$  is defined by<sup>1</sup> [1]

$$\text{Li}_2(z) \equiv - \int_0^z dz \frac{\ln(1-z)}{z}. \quad (\text{B.1})$$

This definition holds for general complex variables  $z$ . Here, however, we will only focus on real, negative values  $x$  [cf. Eq. (6.32) on page 165 in Chapter 6] for which we have plotted the dilogarithm in Fig. B.1.

For  $|x| < 1$ , the dilogarithm can be written as a series

$$\text{Li}_2(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}, \quad (\text{B.2})$$

from which we see that

$$\text{Li}_2(x) \simeq x \quad \text{for } x \ll 1. \quad (\text{B.3})$$

If the first-order approximation for  $\text{Li}_2(x)$  (shown as the dashed line in Fig. B.1) is used in the expression for the zero temperature mean transmission in the parabolic barrier model [Eq. (6.32)], we obtain the WKB expression for the transmission as shown in Chapter 6. For large negative values of  $x$ , the dilogarithm

---

<sup>1</sup>In the literature many alternative definitions and notations for the dilogarithm exist. Here we adopt the notation used by Lewin [1].

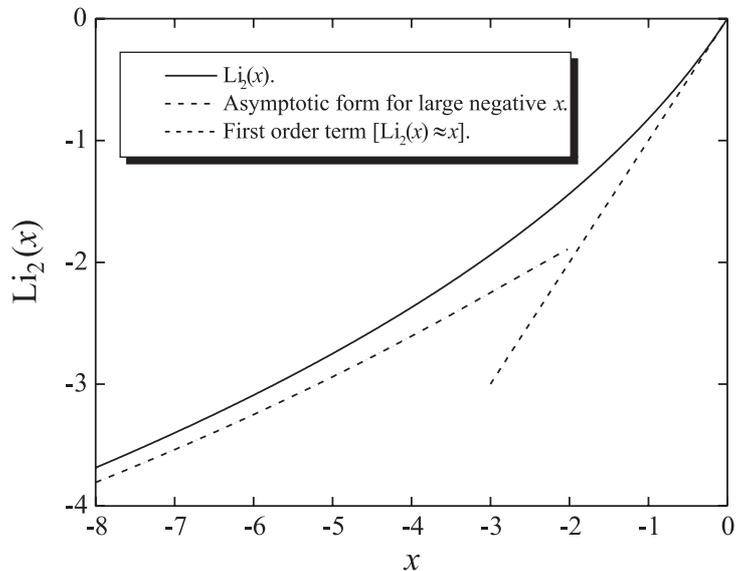


Figure B.1: The dilogarithm for real, negative values of the argument. For comparison, the first-order term  $\text{Li}_2(x) \simeq x$  (the WKB limit in the parabolic barrier model) is shown (dashed line) as well as the asymptotic form for large negative values (dashed-dotted line).

has the asymptotic form

$$\text{Li}_2(x) = -\frac{\pi^2}{6} - \frac{1}{2}[\ln(-x)]^2 \quad \text{for } x \rightarrow -\infty, \quad (\text{B.4})$$

which is shown for comparison as the dashed-dotted line in Fig. B.1.

We have found that the dilogarithm can be evaluated efficiently using the following recipe:

1. If  $x < -1$ , the relation

$$\text{Li}_2(x) = -\frac{\pi^2}{6} - \frac{1}{2}[\ln(-x)]^2 - \text{Li}_2(1/x)$$

is first used to invert the argument and map it into  $[0; 1]$ .

2. If  $x \ll 1$  (say, smaller than 0.01), the series in Eq. (B.2) is used directly for evaluating  $\text{Li}_2(x)$ .
3. Otherwise,  $\text{Li}_2(x)$  is found by Romberg integration [2] of Eq. (B.1).

## References

- [1] L. Lewin, *Polylogarithms and associated functions* (North Holland, New York, 1981).
- [2] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal* (Cambridge University Press, Cambridge, 1989).



## Appendix C

# Stratton's Result

In 1962 Stratton developed an arbitrary barrier approximation in the WKB limit, which could be used for finding the current density  $j(V)$  through a thin insulating film [1]. For arbitrary symmetric barriers, he found the following general formula [Eq. (23) in Ref. 1]

$$j(V) \simeq j_0 \frac{\pi c_{10} k_B T}{\sin(\pi c_{10} k_B T)} \exp[-b_{12}(eV)^2] 2 \sinh(c_{10} eV/2),$$

where

$$j_0 = \frac{4\pi m e \exp(-b_{10})}{h^3 c_{10}^2},$$

and  $c_{10}$ ,  $b_{10}$  and  $b_{12}$  are constants which can be calculated from the barrier shape at equilibrium. We may now write Stratton's corresponding mean transmission  $\mathcal{T}^{\text{Str}}$  by using  $I(V) = j(V)A = G_s \mathcal{T}(V)V$ , whereby we find the transmission valid for arbitrary symmetric (and opaque) barriers

$$\begin{aligned} \mathcal{T}^{\text{Str}}(V) &= \frac{\pi c_{10} k_B T}{\sin(\pi c_{10} k_B T)} \frac{\exp(-b_{10})}{c_{10} E_F} \\ &\times \exp[-b_{12}(eV)^2] \frac{2 \sinh(c_{10} eV/2)}{c_{10} eV}. \end{aligned} \quad (\text{C.1})$$

The parameters  $b_{10}$  and  $b_{12}$  can be found by expanding the logarithm of the WKB transmission given by Eq. (6.25) to second order in the voltage [1]

$$-\ln [T_{1D}^{\text{WKB}}(E_F, V)] \equiv b_1 = b_{10} - b_{11}V + b_{12}V^2 + \dots \quad (\text{C.2})$$

The last parameter  $c_{10}$  is found from the following WKB-like integral over the barrier at equilibrium [1]

$$c_{10} = \frac{\sqrt{2m}}{\hbar} \int_{z_1}^{z_2} \frac{dz}{\sqrt{\phi(0; z) - E_F}}, \quad (\text{C.3})$$

where  $z_1$  and  $z_2$  are the classical turning points at the incoming electron energy  $E_z = E_F$ .

For the parabolic barrier, we find  $b_1^P = \gamma(\phi_V - E_F)$  directly by comparing Eq. (C.2) with Eq. (6.26). Using the expression for  $\phi_V$  given by Eq. (6.24), we then find  $b_{10}^P = \gamma(\phi_0 - E_F)$  and  $b_{12}^P = \gamma/(16\phi_0)$ . When the equilibrium barrier  $\phi(0; z)$  given by Eq. (6.21) is inserted in Eq. (C.3), we get an elementary integral. Upon insertion of the expression for the curvature given by Eq. (6.22), we find  $c_{10}^P = \gamma$ . Inserting the values for  $b_{10}^P$ ,  $b_{12}^P$  and  $c_{10}^P$  in Eq. (C.1), we get

$$\begin{aligned} \mathcal{T}^{\text{Str,P}}(V) &= \frac{\pi\gamma k_B T}{\sin(\pi\gamma k_B T)} \frac{\exp[-\gamma(\phi_0 - E_F)]}{\gamma E_F} \\ &\times \exp\left[-\frac{\gamma}{16\phi_0}(eV)^2\right] \frac{2 \sinh(\gamma eV/2)}{\gamma eV} \end{aligned} \quad (\text{C.4})$$

Finally, we expand the first term

$$\frac{\pi\gamma k_B T}{\sin(\pi\gamma k_B T)} = 1 + \frac{\pi^2}{6} \gamma^2 (k_B T)^2 + \dots, \quad (\text{C.5})$$

and we see that to second order in  $\gamma k_B T$ , the Stratton result is identical to the opaque barrier limit of the parabolic model [compare Eqs. (6.15), (6.35), (6.36), (C.4) and (C.5)].

## References

- [1] R. Stratton, "Volt-current characteristics for tunneling through insulating films," *J. Phys. Chem. Solids* **23**, 1177–1190 (1962).

## Appendix D

# The Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (KS) test is a statistical test, which can be used for testing whether two measured cumulated probability functions,  $S_{N_1}(x)$  and  $S_{N_2}(x)$ , with  $N_1$  and  $N_2$  data points are drawn from *different* underlying distributions.<sup>1</sup> (See Ref. 1 and references therein). To be more specific, the KS test provides an answer to the following question:

Assuming that  $S_{N_1}(x)$  and  $S_{N_2}(x)$  are drawn from the same underlying distribution, what is the probability of finding a worse agreement between the functions if the experiments are repeated?

A few clarifying remarks are in order here. First we have to specify what is meant by an *agreement* between two probability functions. In the KS test, the agreement, or rather the disagreement, is quantified by a particularly simple measure called the Kolmogorov-Smirnov  $D$ . This quantity is defined as the maximum value of the absolute difference between the two probability functions, or, in mathematical terms,

$$D = \max_{-\infty < x < \infty} |S_{N_1}(x) - S_{N_2}(x)| . \quad (\text{D.1})$$

---

<sup>1</sup>In contrast, it is *impossible* to test whether two measured cumulated probability functions are drawn from the *same* underlying distribution. This is because the difference between two different underlying distributions can be infinitely small such that an infinite number of data points are needed to map out the true distributions.

Using  $D$  alone is not enough since we have not yet taken into account the number of data points  $N_1$  and  $N_2$ . Clearly, if  $N_1$  and  $N_2$  are small,  $D$  can be large even when the distributions are drawn from the same underlying distribution. We therefore have to normalize  $D$  with  $N_1$  and  $N_2$  to get a measure  $\lambda$ , which is independent of counting statistics. If  $N_1, N_2 \gtrsim 20$  (as is easily fulfilled for our data),  $\lambda$  can be defined as [1]

$$\lambda = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} D.$$

If  $S_{N_1}$  and  $S_{N_2}$  are drawn from the same underlying distribution, the distribution of  $\lambda$  *itself* can be calculated, at least to a useful approximation, thus giving the significance of any observed nonzero value of  $\lambda$ . The probability that a new measurement would result in a larger  $\lambda$  (worse agreement) can be calculated from the function  $Q_{KS}$ , which is defined in terms of the following infinite sum

$$Q_{KS}(\lambda) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2j^2 \lambda^2), \quad (\text{D.2})$$

which a monotonic function with limiting values

$$Q_{KS}(0) = 1 \quad \text{and} \quad Q_{KS}(\infty) = 0. \quad (\text{D.3})$$

For instance, if  $S_{N_1}(x)$  and  $S_{N_2}(x)$  are identical,  $D = 0$  and thus  $Q_{KS} = 1$ , indicating that there is a 100% probability of getting a worse agreement between the functions if the measurements are repeated. The alternating sum in Eq. (D.2) can be calculated straightforwardly on a computer [1].

## References

- [1] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal* (Cambridge University Press, Cambridge, 1989).