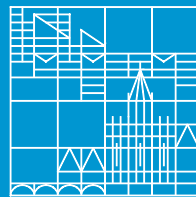


Doktorarbeit

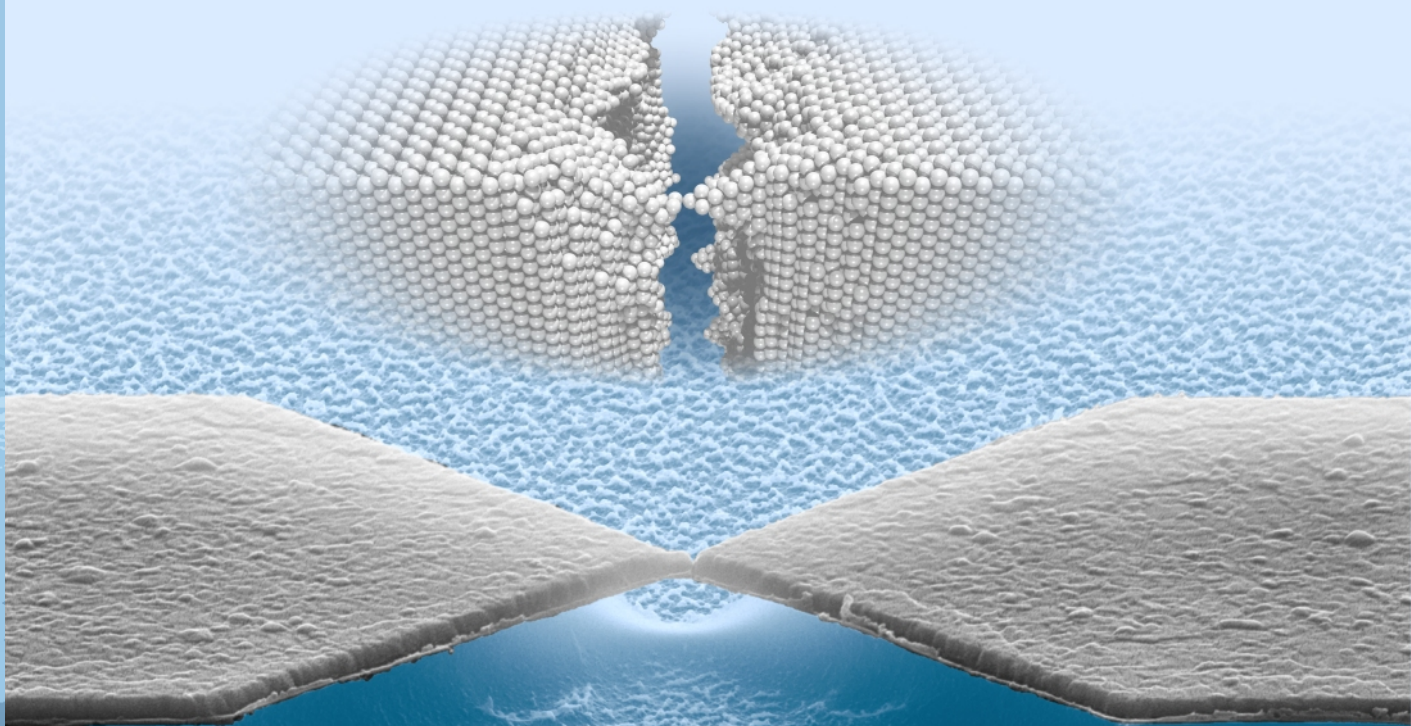


Universität
Konstanz



Christian Schirm

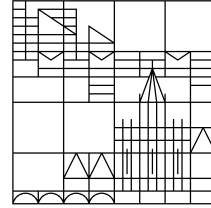
Einfluss hoher Ströme auf atomare Kontakte



Fachgebiet Physik

Doktorarbeit

Universität
Konstanz



Christian Schirm

Einfluss hoher Ströme auf atomare Kontakte

Dissertation zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
an der Universität Konstanz, Fachbereich Physik,
vorgelegt von Christian Schirm.

Tag der mündlichen Prüfung : 08.10.2009

Referenten: Prof. Dr. Elke Scheer

Prof. Dr. Juan Carlos Cuevas

Inhaltsverzeichnis

1. Einleitung	9
Übersicht über die Kapitel	10
2. Grundlagen	11
2.1. Mesoskopischer Transport	11
2.1.1. Mesoskopische Systeme	11
2.1.2. Kurze historische Einordnung der Theorie	11
2.1.3. Theorie des mesoskopischen Transports	12
Mehrere Zuleitungen	14
2.1.4. Leitwertquantisierung	14
2.1.5. Atomare Punktkontakte	15
2.2. Supraleitende SNS-Kontakte	17
2.2.1. Multiple Andreev-Reflexionen	17
2.2.2. IV-Kennlinien	18
2.3. Leitwerthistogramme an Bruchkontakten	20
2.3.1. Leitwertquantisierung	20
2.3.2. Schaleneffekte	21
2.3.3. Theoretische Histogramme	22
2.4. Spannungsabhängiger Leitwert atomarer Kontakte	23
2.4.1. Leitwertfluktuationen	23
2.4.2. Anregung von Vibrationsmoden oder Phononen	24
2.4.3. Gitterverzerrung	24
2.4.4. Einfluss der Bandstruktur	25
2.4.5. Temperaturabhängigkeit	25
2.5. Elektromigration	25
3. Numerische Methoden	27
3.1. Numerische Kanalanalyse	27
3.1.1. Problemstellung	27
3.1.2. Symmetrie-Betrachtungen	28
Symmetrie in Φ	28
Entartung des globalen Minimums	30
Konsequenz	30
3.1.3. Anzahl der Theoriekurven	30
3.1.4. Bandlücke	31

3.1.5. Fit-Algorithmen	34
Systematisches Abrastern	34
GAUSS'sche Regression	35
Simuliertes Ausheilen	36
Automatische Bestimmung der Zahl der Kanäle	38
3.1.6. Effiziente Berechnung des Fitfehlers Φ	40
3.2. Glättalgorithmus	41
3.2.1. Bekannte Glättalgorithmen	42
Echte Glättung	42
Vergleich	44
3.2.2. Neue Glättmethode	45
Funktionsweise des Algorithmus	45
Schnelle Berechnung der Gewichtung	46
Herleitung der Iteration	48
Variante mit DIRAC'scher Deltafunktionen	49
Verbesserte GAUSS-Ähnlichkeit	49
4. Experimentelle Techniken	51
4.1. Überblick	51
4.1.1. Bekannte und neue Techniken	51
4.1.2. Anforderungen	52
4.1.3. Ablauf der Experimente	52
4.2. Gesamtaufbau	53
4.3. Bruchmechanik	54
4.4. Kryostat	55
4.5. Elektrische Verkabelung des Kryostats	55
4.5.1. Schaltbare Verkabelung	55
Lösungsidee	56
Fehlgeschlagene Varianten	57
Beste Variante	58
Beispielmessungen für beide Varianten	58
Details zum Relais	59
Details zu den Kabeln	60
4.5.2. Kupferpulverfilter	60
Filter A	61
Filter B	62
4.5.3. Erdung und Abschirmung	62
4.6. Geräte und Steuerung	63
4.6.1. Messelektronik	63
4.6.2. Computersteuerung	65
4.6.3. Skriptsprache Python	66
Anwendungsbeispiele für Python	66

4.7. Probenherstellung	68
4.7.1. Mechanische Optimierung des Substrats	68
4.7.2. Optimierung des Probendesigns	71
5. Messergebnisse	73
5.1. Atomare Stabilität bei hohen Strömen	73
5.1.1. Messung sprunghafter atomarer Umordnungen	73
5.1.2. Statistik über viele Messungen	74
Steueralgorithmus und Auswertung des Experiments	75
5.1.3. Ergebnis und Interpretation	76
Histogramm über Sprunghöhen	77
Erklärung für die Rückkehr zum mechanischen Histogramm	78
Entstehungsort der Bevorzugten Leitwerte	78
Stabilität	80
5.1.4. Lokale Temperatur	81
Berechnung	82
Fazit	85
5.2. Bistabiles Schalten	86
Vergleich vieler Hysteresen	87
Trainingsphase	90
Atomares Schalten in anderen Arbeiten	90
Eignung als Speicher?	91
5.3. Materialgüte und Supraleitung	92
Zuleitungswiderstand und RRR	92
Energiespektrum	93
Kritische Größen I_c , B_c , T_c	93
5.4. Kanalanalyse	95
5.5. Leitwertkanalhistogramme	101
5.6. Hysterese-Formen	104
Ursache der Formen	104
Ähnlichkeiten der Form	106
Fazit	106
5.6.1. Einzelbeobachtungen	106
5.7. Übersicht über alle Messungen	109
A. Grafische Veranschaulichung	111
A.1. Simulation zur Veranschaulichung eines Bruchkontakts	111
Funktionsweise	111
Ergebnis	113
Literaturverzeichnis	114
Zusammenfassung	125

INHALTSVERZEICHNIS

*»Wo damals die Grenzen
der Wissenschaft waren,
da ist jetzt die Mitte.«*

— Georg Christoph
Lichtenberg¹

1

KAPITEL 1.

Einleitung

Kein physikalisches Phänomen lässt sich technisch so universell einsetzen wie der elektrische Strom. Er ermöglicht den gebündelten und unkomplizierten Transport von Energie und Information, zwei grundlegende Elemente heutiger Technik. Die Wandlungsfähigkeit zu anderen Energie- und Informationsformen und die Einsatzmöglichkeiten wachsen weiter mit dem Wissen über seine grundlegenden physikalischen Gesetzmäßigkeiten. In Zukunft lässt besonders der technologische Vorstoß in den atomaren Bereich auf neue Anwendungen des elektrischen Stroms in der Datenverarbeitung hoffen.

In der vorliegenden Arbeit wurden physikalische Zusammenhänge zwischen dem elektrischen Transportverhalten und der atomaren Struktur wenige Atome umfassender Leiter untersucht. Dazu wurde die Technik der mechanisch kontrollierbaren Bruchkontakte [RAP⁺96] verwendet, um atomare Punktkontakte herzustellen, die in ihrer Geometrie auf mehr oder weniger kontrollierbare Weise verändert werden können. Da kein mikroskopisches Verfahren bekannt ist, um die genaue atomare Geometrie der Kontakte zu bestimmen, kann nur auf gewisse Eigenschaften der Geometrie durch Beobachtung des Leitwerts rückgeschlossen werden. Die möglichst gezielte Beeinflussung der Geometrie und zum Vergleich die Messung der dadurch veränderten Leitwerte standen im Vordergrund dieser Arbeit.

Mit Leitwerthistogrammen, der Kanalspektrenanalyse mittels Supraleitung und der Messung von Strom-Spannungs-Kennlinien wurden die atomaren Kontakte untersucht. Mit diesen Untersuchungsmethoden wurden in den letzten Jahren die meisten wichtigen experimentellen Ergebnisse zu grundlegenden Fragestellungen der atomaren Leitfähigkeit erzielt. So konnten Effekte der Leitwertquantisierung und auch der geometrischen Bevorzugung bestimmter atomarer Konfigurationen (was sich zum Teil als scheinbare Leitwertquantisierung bemerkbar macht) gefunden werden. Außerdem konnten am mikroskopischen Limit eines einzelnen Atoms die von R. LANDAUER vorhergesagten quantenmechanischen Leitwertkanäle [Lan57] einzelnen Orbitalen des zentralen Kontaktatoms zugeordnet werden [SJE⁺97].

Eine gewisse Klassifizierung vieler Eigenschaften zeigt sich in der Unterscheidung zwischen monovalenten und multivalenten Metallatomen. Die Unterschiede erkennt man in

¹1742–1799, erster deutscher Professor für Experimentalphysik, Zitat aus „Sudelbuch“ H, entstanden 1784–1788.

den Kanalspektren von Ein-Atom-Kontakten [SAC⁺98], im Verhalten von Leitwertfluktuationen [Lud99] und (eher in Ausnahmefällen) am Auftreten von Leitwertquantisierungseffekten [KRF⁺95, YR97, XNOS04, THNA05]. In dieser Arbeit wurde primär das trivalente Metall Aluminium näher untersucht.

Übersicht über die Kapitel

Kapitel 2 vermittelt die theoretischen Grundlagen und die wichtigen Ergebnisse bisheriger Experimente zur atomaren Leitfähigkeit. Dabei wird auf die LANDAUER-Theorie der Leitwertkanäle eingegangen, auf Supraleitungseffekte an atomaren Kontakten (multiple Andreev-Reflexionen) und auf einige theoretische Rechnungen zur Leitfähigkeit und zur atomaren Struktur. Ferner werden Quanteninterferenzen (Leitwertfluktuationen), geometrisch bevorzugte Schaleneffekte, Schwingungsanregungen und Elektromigrationskräfte behandelt.

Kapitel 3 zeigt numerische Methoden, die für die Auswertung der Messdaten verwendet und weiterentwickelt wurden. Der erste Teil behandelt verbesserte Methoden für die Kanalanalyse zum Fit von theoretischen Kennlinien supraleitender Kontakte an experimentelle Kennlinien. Ziel der Weiterentwicklung war die Absicherung gegenüber numerischen Artefakten und Fitfehlern. Nebenbei wurde eine deutliche Steigerung der Rechengeschwindigkeit erreicht. Im zweiten Teil wird ein Glättalgorithmus vorgestellt, der ursprünglich der Untergrundbehandlung von Leitwertfluktuationen dienen sollte, jedoch auch in dieser Arbeit mehrfach angewendet wurde.

Kapitel 4 beschreibt den experimentellen Aufbau. Ein paar Verbesserungen gegenüber bisher bekannten Aufbauten werden vorgestellt. Insbesondere wird auf eine spezielle Verkabelung des Kryostats eingegangen, die für die hier durchgeführten Experimente entwickelt werden musste.

Kapitel 5 zeigt die Messergebnisse. Zunächst werden Messungen zu stromgetriebenen Umordnungen von atomaren Kontakten beschrieben und die Ergebnisse und statistischen Analysen diskutiert. Um die Änderungen der atomaren Beeinflussung stärker einzugrenzen, wird durch wechselnde Stromrichtungen ein reversibles lokales Schaltverhalten erzwungen. Die bistabilen Schaltzustände werden mittels Kanalanalyse untersucht. Histogramme werden über Leitwerte und Kanaltransmissionen gebildet und miteinander verglichen. Auch auf die Formen von IV-Kennlinien wird näher eingegangen. Dabei wird der Frage nachgegangen, wo die Ursachen der Effekte liegen, ob lokal am atomaren Kontakt, oder in der näheren Umgebung. Zum Schluss werden Einzelbeobachtungen von bistabilen Schalthysteresen gezeigt und interpretiert.

»Ich will diese elektrischen
Teilchen im Anschluss an
neuere Bezeichnungen
Elektronen [...] nennen.«

— Paul Drude, 1900¹

2 KAPITEL 2. Grundlagen

2.1. Mesoskopischer Transport

2.1.1. Mesoskopische Systeme

Für viele physikalische Systeme genügt ein mikroskopisches, atomares Modell und die Kenntnis der makroskopischen Geometrie für eine vollständige theoretische Beschreibung der physikalischen Vorgänge. Im elektrischen Transport finden dagegen einige Phänomene auf einer mittleren Längenskala statt, die das Gesamtverhalten stark beeinflussen können. Für Systeme in dieser Größenordnung wurde der Name *mesoskopische Systeme* (griechisch: *meso* = mittel) eingeführt. Verantwortlich für die Phänomene dieser Größe sind diverse Längenskalen, wie z. B. die Fermiwellenlänge der Elektronen ($\lambda_f \sim 5 \text{ \AA}$), die elastische mittlere freie Weglänge ($l_e \sim 100 \text{ nm}$) und die Phasenkohärenzlänge ($l_\phi \sim 1 \text{ }\mu\text{m}$). In Abbildung 2.1 sind drei Klassen solcher mesoskopischer Leiter gezeigt, in denen diese Längenskalen entscheidend für das Verhalten sind.

2.1.2. Kurze historische Einordnung der Theorie

Die Entwicklung der Theorien über den elektronischen Strom begann 1900, drei Jahre nach dem Nachweis des Elektrons [Fal97], mit dem von PAUL DRUDE vorgeschlagenen Teilchenmodell [Dru00]. Der zweite Meilenstein folgte eben so kurz nach Beginn der modernen Quantenmechanik (1926, Schrödingergleichung) und der Entdeckung der Wellennatur des Elektrons (1927, [DG27]), als FELIX BLOCH 1928 die neuen Erkenntnisse in einer neuen Theorie der freien Metallelektronen anwenden konnte [Blo29], wodurch auch Einflüsse der zugrundeliegenden Kristallstruktur erklärbar wurden. 1957 wurde schließlich durch ROLF LANDAUER eine mikroskopische Theorie entworfen, die sich der inzwischen ausgereiften Methoden der Atomphysik und Quantentheorie bediente, wie der Streutheorie und der quantenmechanischen Beschreibung elektronischer Eigenzustände. Die Theorie war erstmals in der Lage, elektrische

¹Zitat aus: *Zur Elektronentheorie der Metalle* [Dru00].

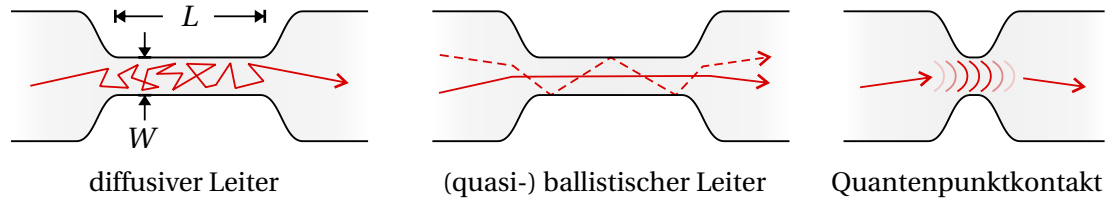


Abbildung 2.1.: Verschiedene Grenzfälle bei mesoskopischen Leitern. Diffusiv: $l_\Phi > W, L \gg l_e$, ballistisch: $L, W < l_e$, Quantenpunktkontakt: $W, L \lesssim \lambda_f$

Leiter mit stark eingeschränkter Geometrie korrekt zu beschreiben. Sie war zunächst auf effektive Zweipunkt-Kontakte begrenzt und wurde 1986 durch MARKUS BÜTTIKER auf echte Vierpunkt-Kontakte erweitert [Bü86]. Die dabei hergeleitete LANDAUER-BÜTTIKER-Formel kann zur Beschreibung von mesoskopischen Systemen als Erweiterung oder Ersatz für das ohmsche Gesetz verwendet werden. Sie soll im Folgenden in groben Zügen kurz erläutert werden.

2.1.3. Theorie des mesoskopischen Transports

Abbildung 2.2 zeigt ein Modell des elektronischen Transports im mesoskopischen Leiter bei einem Ungleichgewicht des chemischen Potentials um eU , entsprechend einer angelegten äußeren Spannung U . Die Bandkante der Elektronen in den Zuleitungen ist links und rechts durch die Fermifunktion dargestellt, die bei endlicher Temperatur um etwa $2k_B T$ verschmiert ist. Die besetzten und unbesetzten Zustände liegen oft auf parabelförmigen Dispersionsrelationen, hier abgebildet für beide k -Richtungen parallel zum Leiter. Im mesoskopischen Leiter selbst stehen nur wenige diskrete Zustände als Transportmoden zur Verfügung.

Ein Transport von Elektronen durch den Leiter kann nur von einem besetzten Zustand des linken zu einem unbesetzten Zustand des rechten Reservoirs stattfinden. Im Leiter selbst müssen freie Zustände, genannt Kanäle, im dazwischenliegenden Energieintervall zur Verfügung stehen.

Zur Berechnung des gesamten Stroms durch den Leiter betrachten wir zunächst einen solchen einzelnen Kanal. Sein Beitrag zum Gesamtleitwert hängt hauptsächlich vom Überlapp der Wellenfunktionen der Zustände in den Elektroden mit dem Kanal ab. Es zeigt sich jedoch, dass selbst im günstigsten Fall für einen einzelnen Kanal eine obere Grenze des Leitwerts existiert, die nicht überschritten werden kann und der eine universelle Konstante darstellt. Dieses Ergebnis liefert folgende etwas grobe Abschätzung. Dafür wird der Leitwert

$$G = \frac{I}{U} = \frac{\Delta Q}{\Delta t} \frac{1}{U} \quad (2.1)$$

zunächst für den Transport einzelner Elektronen mit $\Delta Q = e$ und $\Delta E = eU$ als

$$G = \frac{e^2}{\Delta E \Delta t} \quad (2.2)$$

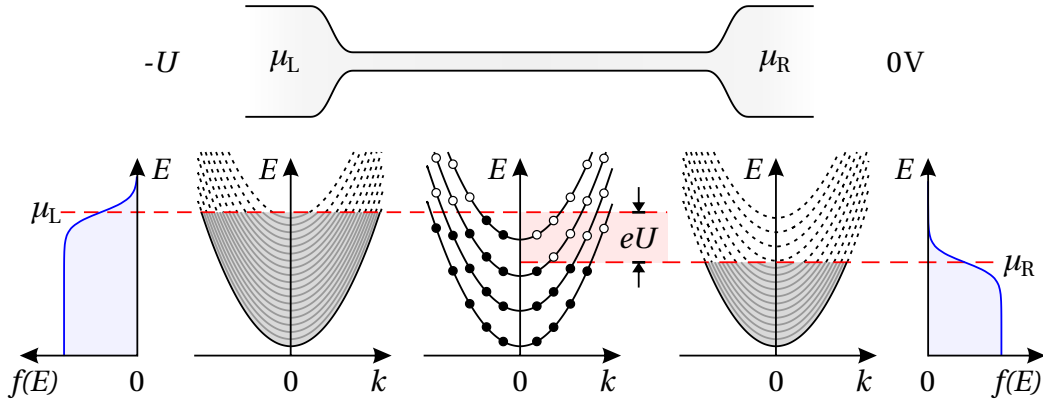


Abbildung 2.2.: Modell des elektronischen Transports durch einen ballistischen mesoskopischen Leiter bei einer angelegten Spannung U und endlicher Temperatur.

geschrieben. In [Bat02] wird nun mit der Energie-Zeit-Unschärfe $\Delta E \Delta t \geq h$ argumentiert, dass der im Energieintervall $\Delta E = \mu_2 - \mu_1$ stattfindende Transportvorgang eine Mindestverweildauer $\Delta t \geq h/\Delta E$ für das Elektron zur Folge hat. Für nachfolgende Elektronen ist wegen des Pauli-Verbots der Kanal die entsprechende Zeit lang blockiert. Dadurch wird der Leitwert

$$G \leq \frac{e^2}{h} \quad (2.3)$$

eines einzelnen Kanals nach oben begrenzt. Das Zweifache dieses maximalen Leitwerts

$$G_0 = \frac{2e^2}{h} \quad (2.4)$$

für den Transport durch beide Spinzustände wird *Leitwertquantum* genannt und mit G_0 abgekürzt. Die Summe über die Leitwerte aller zur Verfügung stehender Kanäle

$$G = \frac{2e^2}{h} \sum_i \tau_i \quad (2.5)$$

wird als LANDAUER-Formel bezeichnet. Die τ_i sind dabei Werte zwischen 0 und 1, welche die Kopplungsstärken der äußeren mit dem Kanal überlappenden Moden wiedergeben. Sie werden als *Transmissionen* bezeichnet und geben die Wahrscheinlichkeit wieder, mit der ein Elektron den Kanal durchläuft, ohne dabei zurückreflektiert zu werden.

Die Herleitung der Formel wurde in der originalen [Lan57] und in späteren Arbeiten und Lehrbüchern detaillierter und sorgfältiger durchgeführt. Diese hier gezeigte kurze Herleitung oder Abschätzung, die „auf den ersten Blick eher zufällig zum exakten Ergebnis zu führen

scheint“², vermittelt jedoch eine plausible Idee dafür, wie selbst bei perfekten ballistischen Leitern dem Leitwert einzelner Kanäle eine universelle Grenze gesetzt sein kann.

Mehrere Zuleitungen

Wenn die Zuleitungen der Anregung und Spannungsmessung nahe am mesoskopischen Leiter innerhalb des ballistischen Bereichs zusammengeführt werden, liefert die Leitwertmessung andere Werte. BÜTTIKER stellte eine entsprechende Formel auf, die von vier (oder einer beliebigen Zahl) Zuleitungen ausgeht. Dabei werden alle Zuleitungen gleich behandelt, unabhängig von ihrer Funktion. Das Ergebnis ist die LANDAUER-BÜTTIKER-Formel

$$I_p = \sum_q G_{pq} [U_p - U_q], \quad (2.6)$$

die den Strom I_p durch die Zuleitung mit Nummer p angibt. U_q ist die Spannung an einer anderen Zuleitung q und

$$G_{pq} \equiv \frac{2e^2}{h} \tau_{p \leftarrow q} = \frac{2e^2}{h} \sum_{m=1}^{N_p} \sum_{n=1}^{N_q} |S_{mp,nq}|^2 \quad (2.7)$$

sind die Koeffizienten der Leitwertmatrix \mathbf{G} . Die Streumatrix \mathbf{S} enthält die Übergangsamplituden des n -ten Zustands der Leitung q zum m -ten Zustand der Leitung p .

Die beiden Gleichungen ersetzen das altbekannte ohmsche Gesetz

$$I = GU \quad \text{mit} \quad G = \sigma \frac{A}{L}, \quad (2.8)$$

welches letztenendes als Grenzfall darin enthalten ist.³

2.1.4. Leitwertquantisierung

1988 gelang VAN WEES ET AL. die Herstellung eines mesoskopischen Leiters mit ausschließlich Kanälen der Transmission 1. In einem 2D-Elektronengassystem eines Halbleiters wurden mit zwei Gate-Elektroden mit variabler Spannung Verarmungszonen räumlich derart variiert, dass ein dazwischen verlaufender leitender Kanal nach Belieben eingeschnürt werden konnte. Die Ränder dieses Leiters schienen genügend glatt zu sein, so dass bei kontinuierlicher Änderung der Breite die Zahl der darin Platz findenden Wellenleitermoden regelmäßig zunahm, wie in Abb. 2.3 zu sehen ist. Die Temperatur war hier klein genug (0.6 K), und die Fermikante

²Zitat aus der Veröffentlichung [Bat02] selbst. Hier findet sich auch eine alternative Herleitung über die Ort-Impuls-Unschärfe.

³Für tiefere Details der mesoskopischen Transporttheorie wird auf das Lehrbuch von S. DATTA [Dat97] verwiesen.

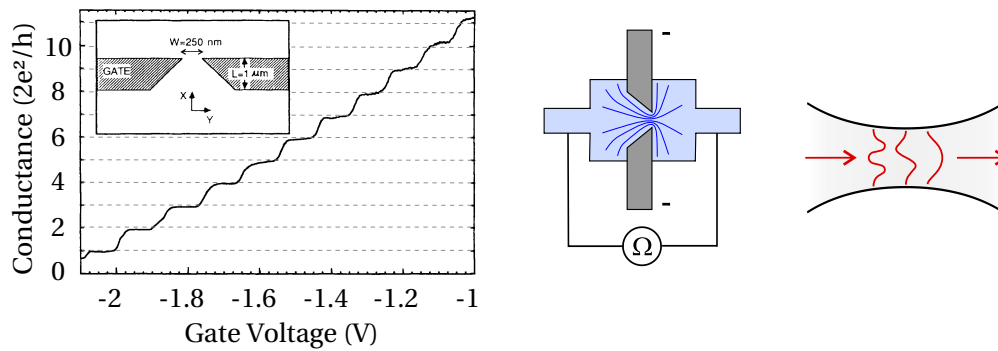


Abbildung 2.3.: Leitwertquantisierung bei perfekt ballistischem Leiter mit variabler Breite [WHB⁺88].

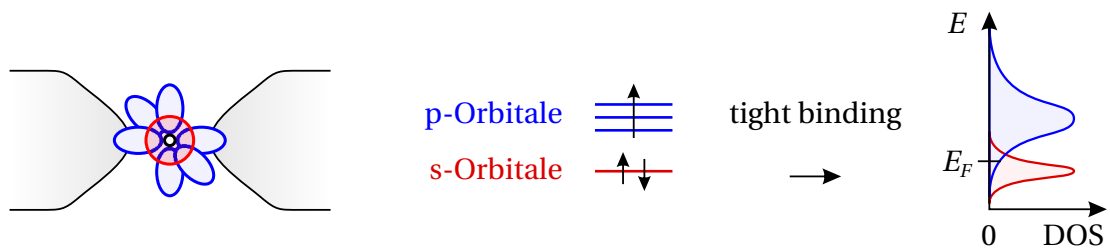


Abbildung 2.4.: Atomarer Punktkontakt aus Aluminium

entsprechend scharf, so dass die Moden jeweils ab einer bestimmten Schwelle in Erscheinung traten.

Der Leitwert, der pro Mode um $1G_0$ ansteigt, ist hier effektiv quantisiert. Die Leitwertquantisierung darf jedoch nicht im strengen Sinn verstanden werden, wie z. B. die Quantisierung der elektrischen Ladung durch Ladungsquanten, den Elektronen, die Grundsätzlich nicht teilbar sind. Dass Leitwerte prinzipiell in beliebigen, nicht ganzzahligen G_0 auftreten können, insbesondere bei Kontakten mit nur einem Kanal, zeigt insbesondere das Beispiel im nächsten Abschnitt.

2.1.5. Atomare Punktkontakte

Transportmoden, wie sie in Abb. 2.3 auftreten, setzen ein entsprechendes Wellenleiterpotenzial voraus. Im kugelsymmetrischen Kernpotenzial eines einzelnen Atoms sind solche Moden nicht möglich. In atomaren Punktkontakten, wie in Abb. 2.4 links schematisch dargestellt, werden die Elektronen durch das Potenzial eines einzelnen Atoms gezwängt. Hier liegt die Vermutung nahe, dass die Elektronenwellenfunktionen beim Durchlaufen anstelle ebenwelliger Transportmoden vornehmlich die Form der atomaren Orbitale annehmen.

Abbildung 2.4 zeigt in der Mitte das Niveauschema eines freien Aluminium-Atoms. Die Valenzelektronen treten in keulenförmigen p- und kugelsymmetrischen s-Orbitalen auf. Angekoppelt an die Kontakte werden die Energieeigenfunktionen der Orbitale verändert und

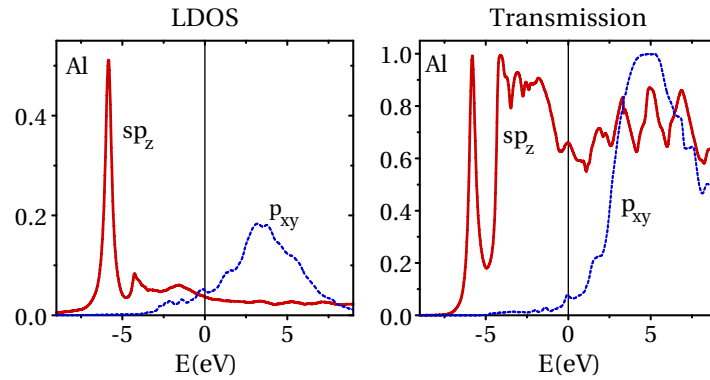


Abbildung 2.5.: Theoretische Rechnung der lokalen Zustandsdichte und der Transmission eines atomaren Punktkontakts von Aluminium wie in Abb. 2.4 schematisch gezeigt, jeweils gegen die Position der Fermi-Energie aufgetragen (aus [Cue99]). Nach neueren Rechnungen können an Ein-Atom-Kontakten bereits Leitwerte bis $2 G_0$ auftreten [PVH⁺08].

dadurch energetisch in der Zustandsdichte aufgeweitet (Abbildung links). Die ursprünglich diskreten Niveaus können dadurch überlappen, und insbesondere mit der Fermienergie E_F in Berührung kommen. Dies erlaubt es, dass die Elektronen der Zuleitung die Orbitale kurzzeitig besetzen können, um auf die andere Seite zu gelangen. Die Besetzungswahrscheinlichkeit ist dabei proportional zur Zustandsdichte an der Fermienergie. So sind manche Orbitale entsprechend stärker und andere schwächer leitend. Eine Leitwertquantisierung zu ganzzahligen G_0 ist daher bei atomaren Kontakten nicht zu erwarten.

Abbildung 2.5 zeigt eine theoretische Rechnung der in Abb. 2.4 schematisch dargestellten aufgeweiteten Niveaus [Cue99]. Außerdem wurde hier die Transmission berechnet. Das s-Orbital im Aluminium hybridisiert mit einem der drei p-Orbitale, welches in Transportrichtung liegt. Der antibindende Zustand des sp_z ist nicht leitend, so bleiben ein sp_z - und zwei p-Orbitale als Kanäle für den Transport übrig. In der Rechnung wurde mit idealisierter Geometrie gerechnet (pyramidenförmige Spitzen in (111)-Wachstumsrichtung). Das sp_z -Orbital leitet wesentlich besser als die beiden gleichen p-Orbitale.

Zur Bestätigung dieser Signatur eines einzelnen Aluminium-Atoms werden im Kapitel 5 Messungen gezeigt. Erstmals wurden solche Messungen in [SJE⁺97] veröffentlicht. Darin konnten unter Verwendung von Supraleitungseffekten die Transmissionswerte der einzelnen Kanäle bestimmt werden, insbesondere bei mehreren gleichzeitig am Stromtransport beteiligten Kanälen.⁴ Abb. 2.6 zeigt eine Öffnungskurve eines atomaren Bruchkontakts, mit den gleichzeitig bestimmten Transmissionen der individuellen Kanäle. Die Messungen bestätigen das atomare Modell recht gut, da an den niedrigsten Plateaus die drei gemessenen Transmissionswerte den individuellen atomaren Niveaus des Aluminium-Atoms überzeugend zugeordnet werden können. Sichtbar sind nämlich zwei ähnliche Kanäle und ein sich anders verhaltender mit höherer Transmission, was mit der gerechneten lokalen Zustandsdichte übereinstimmt. Auch das Ansteigen des sp -Kanals mit wachsendem Abstand ist durch theoretische

⁴Die Methode dazu wird in den Abschnitten 2.2 und 3.1 näher erläutert.

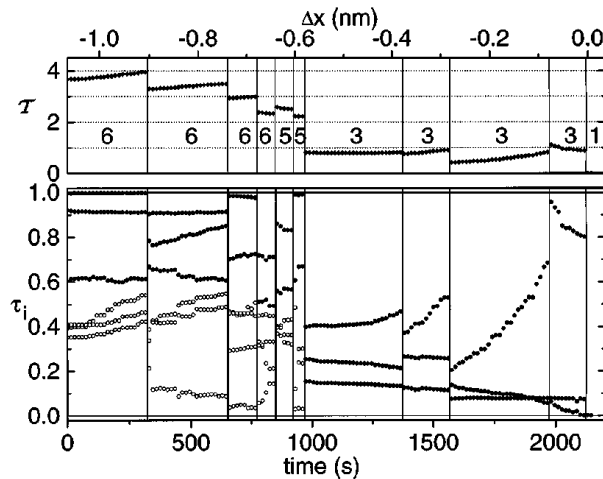


Abbildung 2.6.: Öffnungskurve eines Aluminium-Bruchkontakts mit Bestimmung der Transmissionen der individuellen Leitwertkanäle (aus [SJE⁺97]).

Rechnungen bestätigt worden.

Neben Aluminium wurden auch andere Metalle untersucht [SAC⁺98]. Es konnte jeweils gezeigt werden, dass am letzten Plateau des atomaren Kontakts die Zahl der Kanäle kleiner oder gleich der Zahl der freien Valenzelektronen des Metallatoms ist.

Bei Punktkontakten, die durch mehr als ein Atom in der Engstelle verbunden sind, ist in Anbetracht dieses und des vorausgehenden Abschnitts unklar, ob hier eine Quantisierung zu erwarten ist, oder die lokalen atomaren Zustände den Leitwert bestimmen. Tatsächlich ist meist nur eine Quantisierungstendenz zu erkennen, wie weiter unten gezeigt wird.

2.2. Supraleitende SNS-Kontakte

2.2.1. Multiple Andreev-Reflexionen

Supraleitung⁵ tritt auch in mesoskopischen Leitern auf, wobei an sehr schmalen Kontakten der kritische Strom schnell überschritten wird und bei atomaren Punktkontakten kein kristallines Gitter mehr vorhanden ist, welches eine Voraussetzung für Supraleitung ist. Im Supraleiter-Normalleiter-Supraleiter-Kontakt (kurz SNS-Kontakt) kann es jedoch bei genügend kleiner räumlicher Ausdehnung des N-Bereichs zu einem Überlapp der Cooperpaar-Wellenfunktionen der beiden supraleitenden Seiten kommen (JOSEPHSON-Kontakt). Hierbei können Cooperpaare und normalleitende Elektronen je nach Spannungsdifferenz die Barriere, mehr oder weniger direkt, in bestimmten Vielteilchen-Prozessen passieren [And64]. Abb. 2.7 zeigt oben die elektronischen Zustände der linken und rechten supraleitenden Zuleitung. Der dazwischenliegende Punktkontakt ist durch die gestrichelte Linie angedeutet. Bei einer

⁵Auf die Grundlagen der Supraleitung wird hier nicht näher eingegangen. Eine geeignete Übersicht liefert die von SCHRIEFFER und TINKHAM auf 4 Seiten zusammengefasste Geschichte der Supraleitung [ST99].

Potenzialdifferenz von mindestens $eU \geq 2\Delta$ können normalleitende Elektronen direkt aus dem besetzten in den unbesetzten Bandbereich der anderen Seite oberhalb der Bandlücke gelangen. Bei geringeren Spannungen ist dieser Übergang nicht möglich, jedoch könnten sich zwei solche Elektronen zu einem Cooperpaar zusammenschließen, was eine Spannung von mindestens $eU \geq \Delta$ erfordert. Unterhalb dieser Spannung ist zumindest ein 3-Teilchen-Prozess mit einem virtuellen Zwischenniveau möglich, wie grafisch veranschaulicht wird. Es zeigt sich, dass n -Teilchen-Prozesse ab einer Spannung von $eU \geq 2\Delta/n$ auftreten. Alternativ wurde im weiter unteren Teil der Abbildung die abfallende Potenzialdifferenz nicht eingezeichnet, entsprechend bewegen sich die Elektronen um eU nach oben. Beim n -Teilchen-Prozess kann der Teilchenübergang abwechselnd als Elektron und als in die andere Richtung laufendes Loch eingezeichnet werden, bis am Ende in der Summe einmal die Bandlücke überwunden wurde. Aus je zwei virtuellen Niveaus, symmetrisch um das Cooperpaarniveau, kann sich am Ende ein Cooperpaar bilden, oder bei ungeraden n kann außerdem ein normalleitendes Elektron in das unbesetzte Band gehoben werden. In einer gewissen Hilfsvorstellung werden hier mehrfach Elektronen als Löcher und Löcher wieder als Elektronen zurückreflektiert, woraus die Bezeichnung *multiple Andreev-Reflexionen* (MAR) entstanden ist.

2.2.2. IV-Kennlinien

Obwohl die Vielteilchenprozesse im Allgemeinen bei großer Anzahl n mehr Ladung transportieren, ist der Strombeitrag geringer, da der Prozess seltener stattfindet. Die Transmissionswahrscheinlichkeit eines Elektrons ist im normalleitenden Zustand durch τ gegeben, daher müssen in n -Teilchen-Prozessen die Einzelwahrscheinlichkeiten beim gleichzeitigen Auftreten multipliziert werden, wodurch eine Wahrscheinlichkeit $P \propto \tau^n$ maßgeblich für den Strombeitrag ist. Der aus den Übergängen resultierende Strom nimmt daher mit wachsendem n ab. Abb. 2.7 zeigt unten den Beitrag jedes Übergangs zum Strom-Spannungsdiagramm, der unterhalb einer Schwelle von $eU = 2\Delta/n$ auf Null abfällt. Die Summe aller Beiträge ergibt hier eine stufenförmige Kurve. Durch die unterschiedlichen Potenzen von τ verschieben sich die Verhältnisse der Einzelbeiträge und damit die relativen Stufenhöhen, so dass unterschiedliche Kurvenformen für unterschiedliche τ zu erwarten sind. Dies erweist sich als Glücksfall, weil die damit verbundene lineare Unabhängigkeit zur Analyse der Messdaten ausgenutzt werden kann.

Die tatsächlichen Kurvenformen weichen von dem idealisierten Modell (Abb. 2.7 unten) deutlich ab. Die genauen Kennlinien können jedoch numerisch berechnet werden. Das Ergebnis einer solchen Rechnung zeigt Abb. 2.8. Die Schwellen der Vielteilchen-Prozesse sind hier nicht als Stufen, sondern mehr als kleine Knicke erkennbar. Die Übereinstimmung dieser Rechnungen mit Experimenten erweist sich jedoch als nahezu perfekt. In Abb. 2.8 wurden keine absoluten SI-Einheiten verwendet (A und V), sondern Einheiten der Energielücke. Auf diese Weise wird das Ergebnis universell für unterschiedliche Metalle gültig, da dies der einzige und außerdem linear eingehende Materialparameter ist. Unterschiedliche Transportkanäle, die parallel verlaufen (nicht im räumlichen Sinn, sondern elektronisch als effektive Parallelschaltung), führen zu einer Summe solcher IV-Kennlinien. Dabei gehen außer den Transmissionswerten

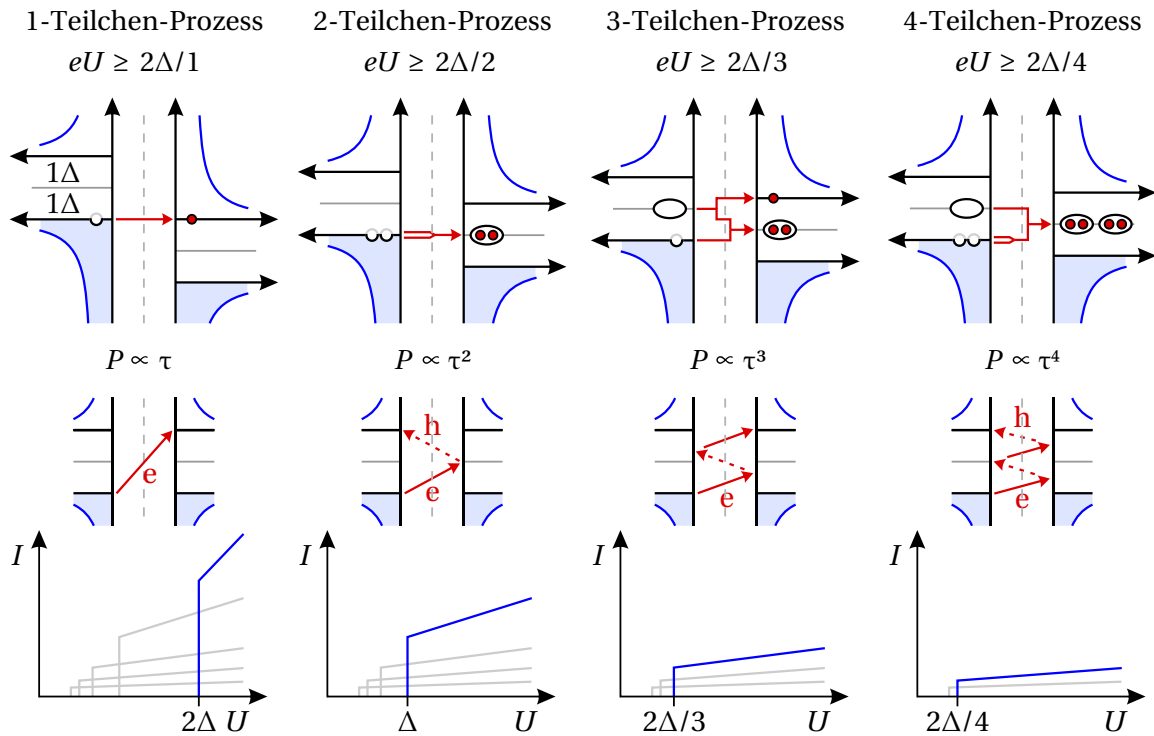


Abbildung 2.7.: Oben und Mitte: Mechanismen des Ladungstransports im SNS-Kontakt. Ab jedem Überschreiten einer Potenzialdifferenz von $eU = 2\Delta/n$ wird zusätzlich zu den bisherigen Übergängen mit $n + 1, n + 2, \dots$ Teilchen ein weiterer Übergang mit nur n Teilchen ermöglicht. Man nennt die Übergänge *Quasiteilchen-Übergang* ($n=1$), *Andreev-Reflexion* ($n=2$) und *multiple Andreev-Reflexionen* ($n \geq 3$). Unten: Einfache Vorstellung zur Erklärung der Strom-Spannung-Kennlinien.

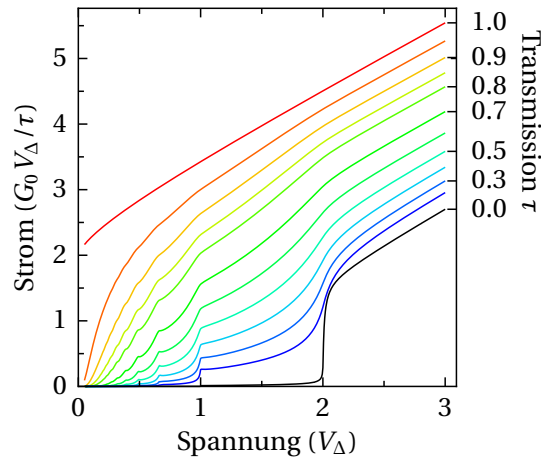


Abbildung 2.8.: Von C. CUEVAS berechnete IV-Kurven [CMRY96], zur Verdeutlichung der Form mit τ normiert (unnormierte Kurven siehe Abb. 3.1, Seite 27). Für jeden Transmissionswert hat die Kennlinie eine individuelle Form, die linear unabhängig von den anderen ist.

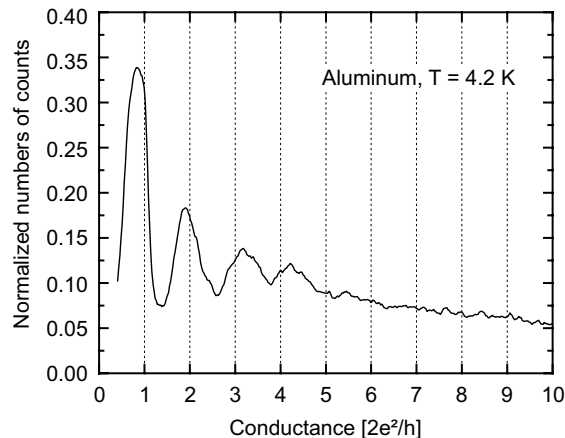


Abbildung 2.9.: Histogramm aus 30 000 Öffnungsvorgängen eines Aluminium-Bruchkontakts bei 10 mV (aus [YR97, AYR03]).

τ_i und der Bandlücke keine weiteren Parameter ein, um die IV-Kennlinie eines bestimmten mehrkanaligen Kontakts nachzubilden. Durch eine geeignete Fitprozedur (siehe Abschnitt 3.1) kann die im Experiment gewonnene IV-Kennlinie auch in die einzelnen IV-Kennlinien zerlegt werden, um so die einzelnen Transmissionen τ_i zu bestimmen.

2.3. Leitwerthistogramme an Bruchkontakten

2.3.1. Leitwertquantisierung

Mechanisch kontrollierbare Bruchkontakte, die auch in dieser Arbeit untersucht wurden, können durch Öffnen und Schließen in ständig neue Geometrien überführt werden, als würde die Probe permanent ausgetauscht werden. Durch diese ergodische Annahme können in einem Experiment Statistiken über Millionen von Proben, d. h. atomare Kontakte, gebildet werden. Besonders eignen sich zur Auswertung Leitwerthistogramme, also Häufigkeitsverteilungen der vorkommenden Leitwerte. In [KRF⁺95] wurde auf diese Weise eine Tendenz zur Leitwertquantisierung in Natrium- und Kupfer-Kontakten gezeigt, die mit zylindrischen Moden erklärbar ist. In [YR97] (siehe Abb. 2.9) wurde auch für das hier untersuchte Material Aluminium ein Histogramm aus 30 000 Öffnungsvorgängen gebildet. Obwohl, wie im vorigen Abschnitt dargestellt wurde, die individuellen Leitwertkanäle keine Andeutung einer Leitwertquantisierung zeigen, sind auf den ersten Blick in der Summe der Kanäle (also im Gesamtleitwert) bevorzugte Werte erkennbar. Diese liegen sogar in der Nähe der Leitwertquanten. In [YR97] wird jedoch hervorgehoben, dass die Maxima zum Teil auch etwas oberhalb der ganzzahligen Vielfachen des Leitwertquantums liegen. Insgesamt deutet dies an, dass keine reine Quantisierung der elektrischen Moden als Ursache in Frage kommt. Vielmehr könnte eine geometrische Bevorzugung die Form des Histogramms prägen. Dies wurde durch spätere Messungen zu Schaleneffekten auch deutlich gezeigt. Ebenso musste durch die Experimente

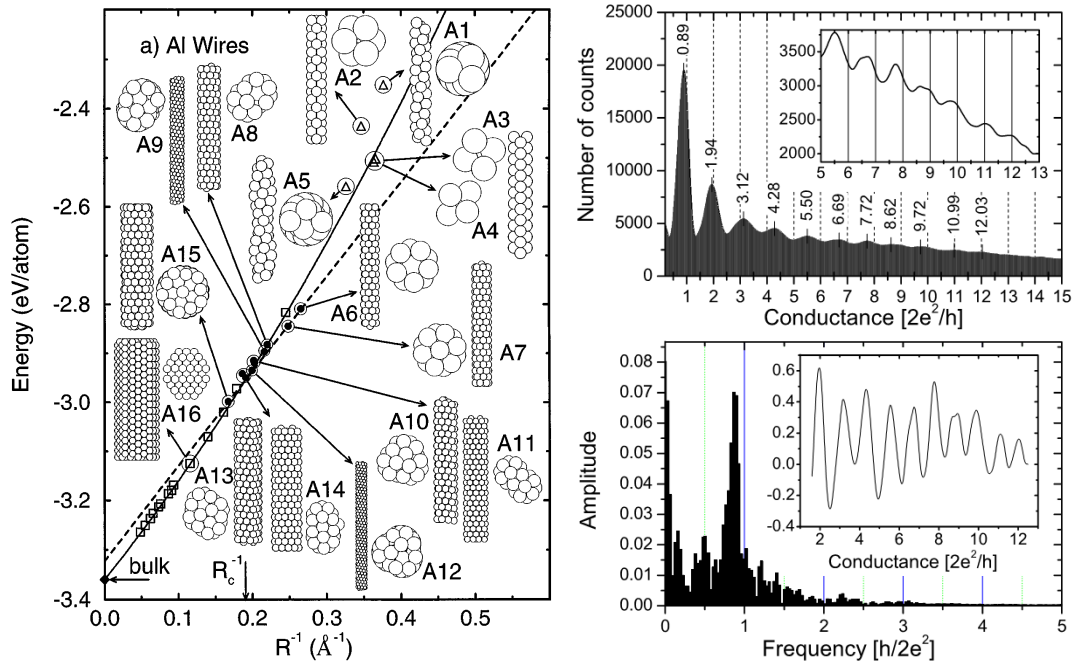


Abbildung 2.10.: Links: Theoretische Vorhersagen von Schalenstrukturen bei Aluminium (aus [GET98]). Rechts: Gemessene Histogramme an gehärtetem Aluminium mit sehr regelmäßigen Abständen mit $\Delta G > 1$ (aus [YSRS08]).

die ergodische Annahme wieder zum Teil verworfen werden, da die Vorgeschichte des Metalls (die kristalline Struktur) die Leitwerthistogramme reproduzierbar beeinflussen kann.

2.3.2. Schaleneffekte

Experimente die eine reine Leitwertquantisierung zeigen, können durch Eigenzustände der Elektronenwellen erklärt werden [KRF⁺95]. Treten jedoch bei der Erzeugung des Kontakts bevorzugte geometrische Konfigurationen auf, wird dies ebenfalls als Häufung der damit verbundenen Leitwerte im Histogramm sichtbar. Die Häufungen liegen zum Teil in regelmäßigen Abständen in einer Größenordnung um G_0 , was die Unterscheidung von der Leitwertquantisierung erschwert. Die bevorzugten Geometrien weisen eine Ähnlichkeit zu metallischen Clustern auf. Diese bilden auf der Oberfläche der rundlichen Partikel Schalenstrukturen aus, zum Teil geometrischer Art mit monoatomaren Schichten, und zum Teil elektronischer Art, ähnlich den Schalen im Atommodell. Cluster mit abgeschlossenen Schalen werden *magisch* genannt und sind besonders stabil und entsprechend häufig. Geometrische und elektronische Schalen sind meist nicht bei der selben Teilchenzahl magisch. Falls doch, spricht man von *doppelt magisch*.

Bei Nanodrähten treten ähnliche Schalen im Querschnitt des Drahtes auf [YYR99, Mar06]. Abb. 2.10 links zeigt einige theoretische Vorhersagen solcher bevorzugter Geometrien bei Aluminiumdrähten [GET98]. Experimentell können die Shell-Strukturen über Bruchkontakt-

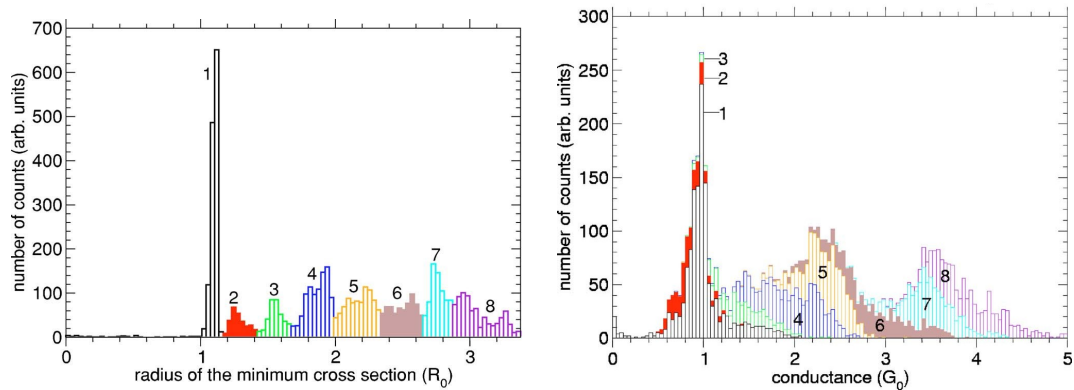


Abbildung 2.11.: Molekulardynamische Simulation mehrerer Öffnungskurven eines Gold-Bruchkontakts von M. DREHER [DPH⁺05]. Links: Histogramm der Querschnittsradien. Rechts: Leitwerthistogramm mit farbig hervorgehobenen Zuordnungen zu den Querschnittsradien.

experimente mit Spannungen $\gg 10$ mV und mit Leitwerthistogrammen nachgewiesen werden [YYR99]. Speziell bei Aluminium können unter bestimmten Voraussetzungen elektronische oder geometrische Schalen dominieren [Mar06, MUB⁺07]. Oder wie in Abb. 2.10 rechts können an gehärtetem⁶ Aluminium sehr regelmäßige Maxima im Abstand $1.15 G_0$ auftreten [YSRS08]. Eine reine Leitwertquantisierung [KRF⁺95], ist hier durch die geometrische Quantisierung völlig unterdrückt.

2.3.3. Theoretische Histogramme

Zur Klärung dieser und anderer Fragen wurde eine umfangreiche Simulation von M. DREHER durchgeführt [DPH⁺05, Dre08]. Dabei wurden mit molekulardynamischen Simulationen vollständige Öffnungsvorgänge von Bruchkontakten einschließlich der Leitwerte berechnet. In der Auswertung wurden auch die jeweiligen Querschnittsradien des Kontakts bestimmt, die im Experiment nicht zugänglich sind. Obwohl deutliche Peaks im Histogramm der Querschnittsradien erkennbar sind (Abb. 2.11 links), sind dieselben Querschnitte im Leitwerthistogramm über größere Bereiche verteilt. Sie sind insbesondere nicht mit Vielfachen des Leitwertquantums korreliert. Die Abweichung der im Experiment beobachteten Quantisierung an Gold wird in [Dre08] den Einflüssen der relativ schmalen Anfangsgeometrie zugeschrieben. Aluminium konnte hier aufgrund bestimmter komplizierter Eigenschaften des Metalls leider nicht in der selben Fülle und Genauigkeit wie Gold analysiert werden.

⁶Der Unterschied zwischen weichem und kalt gehärtetem Material ist bei Gold um einiges deutlicher sichtbar [YSC⁺05, HWHS08].

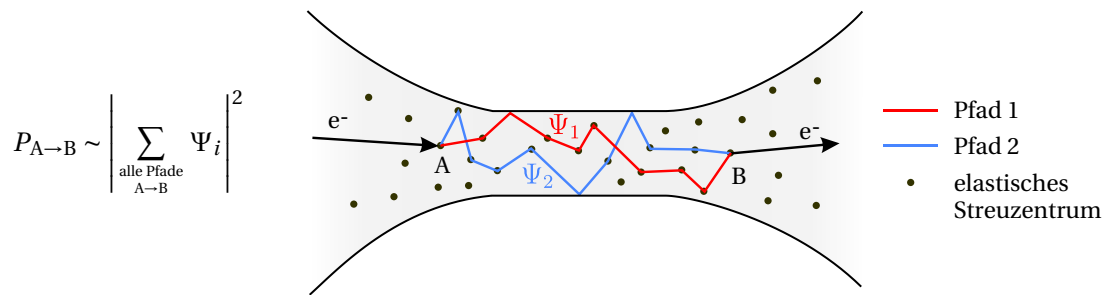


Abbildung 2.12.: Leitwertfluktuationen im diffusiven mesoskopischen Leiter

2.4. Spannungsabhängiger Leitwert atomarer Kontakte

2.4.1. Leitwertfluktuationen

In mesoskopischen Leitern ist im Gegensatz zu ohmschen Widerständen der Leitwert meist spannungsabhängig. Im diffusiven Regime können Interferenzen zwischen unterschiedlichen Elektronenpfaden auftreten (siehe Schema in Abb. 2.12). Betrachtet man die Elektronenwellen als FEYNMANSche Pfade, die sich beim Eintritt in den Leiter aufteilen und am Ende und innerhalb der Phasenkohärenzlänge wieder zusammenlaufen, so kann die dortige Größe der quantenmechanischen Amplitude mit der Summe über alle Pfadintegrale entlang der Wege berechnet werden. Als Wahrscheinlichkeit, für ein Elektron von einem Ort A zu einem Ort B zu gelangen, ergibt sich rechnerisch

$$P_{A \rightarrow B} \sim \left| \sum_{\substack{\text{alle Pfade} \\ A \rightarrow B}} \Psi_i \right|^2 \quad \text{mit} \quad \Psi_i = \Psi_{A,i} \exp(i\varphi_i) \quad \text{und} \quad \varphi_i = \frac{1}{\hbar} \int_{t_A}^{t_B} L(\vec{x}, \dot{\vec{x}}, t) dt. \quad (2.9)$$

Die als $L = T - V$ angesetzte Lagrange-Funktion wächst mit steigendem Spannungsgefälle. Dadurch sind zwischen Paaren von Pfaden mit unterschiedlicher Weglänge sinusperiodische Schwankungen der Amplitude als Funktion der Spannung zu erwarten. Der Gesamtleitwert setzt sich aus einer Summe unterschiedlicher Frequenzen zusammen, die jeweils Aufschluss über die Weglängenunterschiede und indirekt über die Weglängen geben können.

Im atomaren Punktkontakt sind ebenfalls Leitwertfluktuationen zu erwarten. Abb. 2.13 zeigt schematisch die Voraussetzung dafür. Im ballistischen Regime müssen Streuzentren vorhanden sein, die Teilen der Elektronenwelle ermöglichen, den Kontakt z. B. ein zweites Mal phasenverschoben zu durchlaufen. Rechts sind Messungen von Leitwertfluktuationen gezeigt mit unterschiedlich starker Amplitude. Nahe beim Leitwert $1 G_0$ ist die Fluktuationsamplitude sehr klein, da es sich meist um einen vollen Kanal handelt, bei dem der reflektierte Anteil klein ist gegenüber der transmittierten Amplitude.

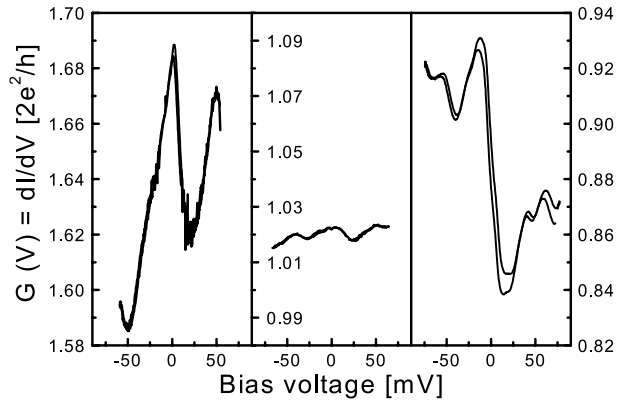
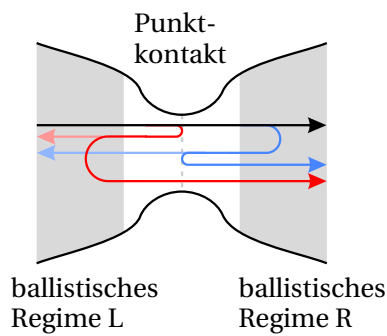


Abbildung 2.13.: Links: Schema von Leitwertfluktuationen an atomaren Punktkontakten. Die ballistische Streuung sorgt für zweifaches Durchlaufen des Kontakts mit unterschiedlichen Teilen der Wellenfunktion des Elektrons. Rechts: Messung der Leitwertfluktuationen bei verschiedenen Leitwerten. Aus [Lud99, LDE⁺99].

2.4.2. Anregung von Vibrationsmoden oder Phononen

Weitere Spannungsabhängigkeiten des Leitwerts können auftreten, wenn ab bestimmten Schwellenspannungen Vibrationsmoden einzelner Atome oder Phononen in den Zuleitungen angeregt werden. Die Anregungen treten symmetrisch in positiver und negativer Stromrichtung in Erscheinung. Oberhalb der Schwelle führt die Anregung zu einer Änderung des Leitwerts.

Vibrationsmoden konnten bei einem einzelnen H_2 -Molekül zwischen zwei atomaren Platin-Elektroden (und bei massereicheren Isotopenmolekülen) sehr deutlich nachgewiesen werden [SNU⁺02]. Doch auch an einem rein metallischen Aluminium-Kontakt ließen sich (transversale) Vibrationsmoden identifizieren [BES09, Bö08]. Der Effekt kann teilweise nicht von Leitwertfluktuationen unterschieden werden. Merkmal von Vibrationsanregungen und Phononen sind ein Peak bei einer positiven und ein Dip bei der selben negativen Spannung in der zweiten Ableitung von $I(V)$ nach V . Das entspricht einem (meist schlecht sichtbarem) Knick in der Kennlinie $I(V)$.

2.4.3. Gitterverzerrung

Auch die spannungs- bzw. stromabhängigen Verzerrungen der lokalen atomaren Geometrie können zu deutlichen Leitwertänderungen in IV-Kennlinien führen. Hinweise darauf liefern die Messungen dieser Arbeit, wie später noch gezeigt wird. Physikalische Gründe für die dafür verantwortlichen Kräfte auf einzelne Atome werden im Abschnitt 2.5 zur Elektromigration genannt. Die mit der Spannung steigenden Kräfte lenken einzelne Atome aus ihrer Ruhelage immer weiter aus, was wiederum einen spannungsabhängigen Leitwert verursacht.

2.4.4. Einfluss der Bandstruktur

Wie an der theoretischen lokalen Zustandsdichte in Abb. 2.5, S. 16 erkennbar ist, kann die Zustandsdichte und damit die Transmission vom chemischen Potenzial der Elektronen abhängen. Bei Variation der Spannung kann sich das chemische Potenzial lokal ändern, und damit also auch der Leitwert. Dieser Effekt ist jedoch bei den experimentell zugänglichen Spannungen verschwindend gering.

2.4.5. Temperaturabhängigkeit

Die Temperatur kann am Ein-Atom-Kontakt mit wachsendem Strom ansteigen. Die dort dissipierte Wärme kann bereits mit dem Thermometer nachgewiesen werden, obwohl die Wärme, wenn sie am Thermometer ankommt, bereits sehr ausgedünnt ist. Am Entstehungsort um den Bruchkontakt herum ist sie dagegen stark konzentriert. Die lokale Temperatur kann über Phononen oder Vibrationen einen Einfluss auf den Leitwert ausüben.

2.5. Elektromigration

Etwas gegenteilig zu den bisherigen Themen zeigt dieser Abschnitt, wie der Strom den Leiter. Die bisherigen Themen des ersten Kapitels zeigten Einflüsse des Leiters auf das Transportverhalten der Elektronen. Was durch den Titel dieser Arbeit angekündigt wurde, der Atomaren behandelten den Einfluss der von dem dar, was bisher Der Titel dieser Arbeit bezieht sich auf das Gegenteil von dem, was bisher in diesem Kapitel beschrieben wird. Die Fortbewegung einzelner Metallatome durch elektrischen Strom wird durch zweierlei Kräfte verursacht, die in gegensätzliche Richtungen zeigen. In [Lod05] wurden die historischen zunächst getrennt prognostizierten Kräfte zu einer vereinheitlichten Theorie zusammengefasst, die das Zusammenspiel beider Kräfte beschreibt. Benannt werden die beiden Kräfte mit *direkter Kraft* und *Windkraft*. Für ein Atom ergibt sich im Metall die Kraft

$$\mathbf{F} = \mathbf{F}_{\text{direct}} + \mathbf{F}_{\text{wind}} = (Z_{\text{direct}} + Z_{\text{wind}}) e\mathbf{E} = (Z_i + Z^{\text{scr}} + Z_{\text{wind}}) e\mathbf{E} \quad (2.10)$$

in einem elektrischen Feld \mathbf{E} . Die direkte Kraft ist dabei die Kraft des Feldes auf den positiven Atomkern mit der effektiven Ladung Z_{direct} , die sich aus der Ionenwertigkeit Z_i (positiv) und einer effektiven Abschirmung der Ionenladung durch die Valenzelektronen Z^{scr} (negativ) zusammensetzt. Die Windkraft ist dagegen eine indirekte Kraft auf den Kern in die entgegengesetzte Richtung. Die negative Elektronenhülle spürt die abstoßende Kraft der ankommenden Leitungselektronen, die wiederum durch repulsive Pauli-Kräfte einen Abstand zum Kern halten. Die Windkraft ist also eine Kraft aufgrund des Drucks des fließenden elektrischen Stroms, während die direkte Kraft durch die elektrische Spannung hervorgerufen wird.

Es ist letztendlich entscheidend, welche der Kräfte überwiegt, d. h. wie stark die Abschirmung ist. Abb. 2.14 zeigt eine Rechnung, wie die Abschirmung vom Metall und von der genauen Potenzialform abhängen kann, in der sich das Teilchen befindet.

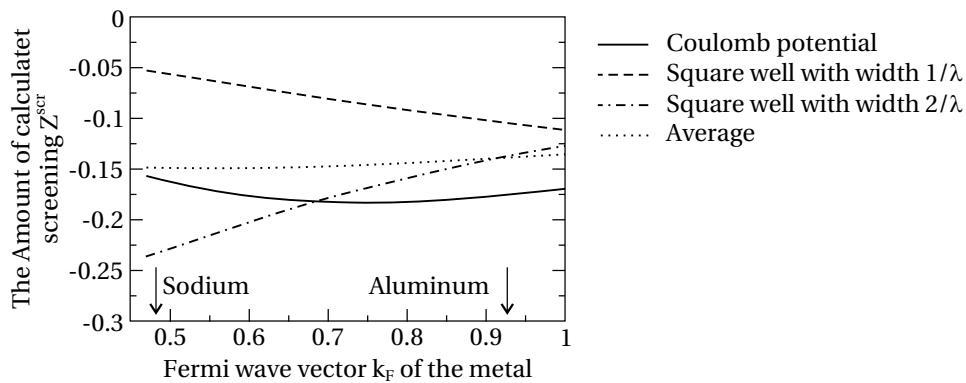


Abbildung 2.14.: Theoretische Rechnung der Abschirmung, abhängig vom Fermi-Wellenvektor und von der Art des umgebenden Potentials (aus [Lod05]).

An experimentellen Arbeiten sind besonders die von STAHLMECKE ET AL. zu erwähnen [SHC⁺06, SCHED06, SD07a, SD07b, Sta08]. Hier wurden in-situ-Beobachtungen von Elektromigrationsvorgängen mit einem Rasterelektronenmikroskop durchgeführt. Unter anderem konnte gezeigt werden, wie Material durch die Elektromigration von einer zur anderen Seite und zurück transportiert werden kann, je nach Stromrichtung. Bei unterschiedlichen Leitern (z. B. polykristallines Gold und einkristallines Silber oder bei unterschiedlich verunreinigten Oberflächen) wurden sogar gegensätzliche Elektromigrationsrichtungen beobachtet [Sta08], was bedeutet, dass im einen Fall die direkte Kraft, im anderen die Windkraft stärkeren Einfluss auf die Atome ausübt.

»Die Rechenautomaten haben etwas von den Zauberern im Märchen. Sie geben einem wohl, was man sich wünscht, doch sagen sie einem nicht, was man sich wünschen soll.«

— Norbert Wiener¹

3 KAPITEL 3.

Numerische Methoden

In diesem Kapitel werden einige in dieser Arbeit verwendeten und weiterentwickelten numerischen Methoden zur Datenauswertung beschrieben.

3.1. Numerische Kanalanalyse

3.1.1. Problemstellung

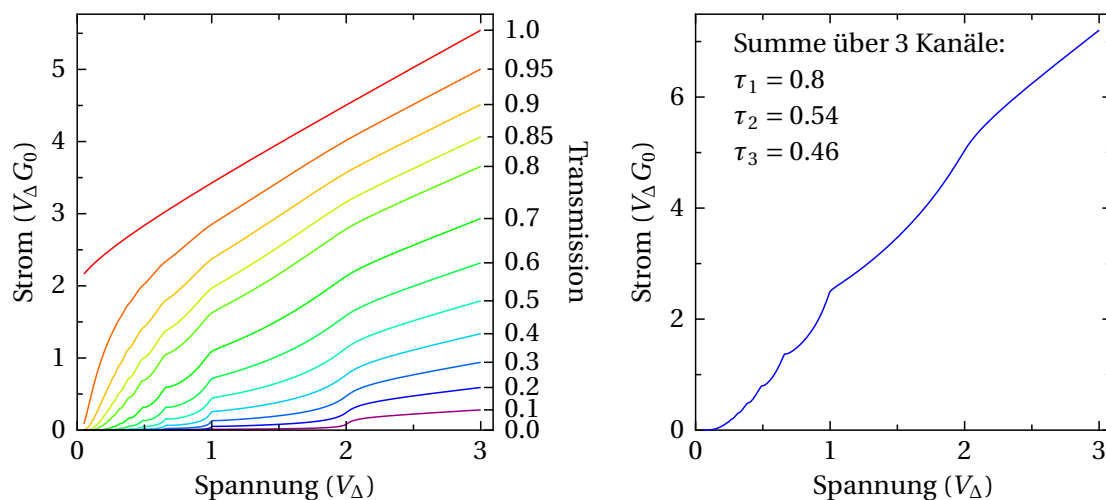


Abbildung 3.1.: MAR-Theoriekurven, berechnet von C. CUEVAS und Beispiel einer Summe aus drei bestimmten Kanälen mit den Transmissionen 0.8, 0.54 und 0.46.

Wie im Abschnitt 2.2 bereits erwähnt, können die Transportkanäle eines atomaren Punktkontakts experimentell bestimmt werden, indem die gemessene Strom-Spannungs-Kennlinie numerisch in die Beiträge einzelner Kanäle zerlegt wird. Die von C. CUEVAS berechneten

¹US-amerikanischer Mathematiker, 1894–1964, (entwickelte d. WIENER-Filter.). Quelle unklar, ev. nicht wörtlich.

Kennlinien einzelner solcher Kanäle sind in Abb. 3.1 gezeigt und standen zur Analyse der Messdaten dieser Arbeit zur Verfügung. Jede dieser Theoriekurven ist eindeutig mit einer Transmission² eines Kanals verknüpft. Durch eine Zerlegung in die Einzel-Kanal-Kurven wird also der Transmissionswert jedes beitragenden Kanals bestimmt. Dadurch liegt auch die Anzahl der im Experiment vorhandenen Kanäle fest, indem Kanäle mit Transmission nahe 0 als nicht vorhandene Kanäle betrachtet werden.

Zur Bestimmung der Kanäle werden n Theoriekurven mit verschiedenen Transmissionen $(\tau_1, \tau_2, \dots, \tau_n)$ so kombiniert, dass der mittlere Fehler

$$\Phi(\tau_1, \tau_2, \tau_3, \dots, \tau_n) = \sqrt{\int \left(\sum_{i=1}^n I_{\text{Theo}}(\tau_i, U) - I_{\text{Mess}}(U) \right)^2 dU} \quad (3.1)$$

minimal wird. n muss größer oder gleich der tatsächlichen Zahl von Kanälen sein. Die verschiedenen Kurven erweisen sich als linear unabhängig, so dass genau ein Minimum existiert und die Zerlegung folglich eindeutig und korrekt möglich sein sollte. Praktisch geht jedoch bei zunehmendem Messrauschen und großer Zahl von Kanälen die enthaltene Information über die Kanäle immer mehr verloren. Im Folgenden wird das Auftreten eines Artefaktes erörtert, das ähnliche Kanäle zu scheinbar gleichen Kanälen verschmelzen lässt.

3.1.2. Symmetrie-Betrachtungen

Symmetrie in Φ

Der Fehler Φ bildet eine N -dimensionale Potenziellandschaft. Aus der Vertauschbarkeit der Summanden folgt für Φ eine Symmetrie bezüglich paarweiser Vertauschung zweier beliebiger Koordinaten τ_i und τ_j . Für jeden Wert in Φ mit ungleichen Koordinaten existieren folglich $N!$ identische Kopien. Die Potenziellandschaft zerfällt daher in $N!$ identische, zum Teil gespiegelte Teilbereiche. Die inneren Grenzflächen

$$S_{p,q} := \{(\tau_1, \tau_2, \dots, \tau_p, \dots, \tau_q, \dots, \tau_N) \mid \tau_i \in [0, 1], \tau_p = \tau_q\} \quad (3.2)$$

sind $\frac{N!}{2!}$ lineare Unterräume der Dimension $N - 1$ für alle Paare p und q mit $p \neq q$. Sie bilden Spiegelebenen, da für jeden Punkt im Abstand $\sqrt{2}d$ zu einer solchen Ebene $S_{p,q}$ das Potenzial

$$\Phi(\tau_1, \tau_2, \dots, \tau_p - d, \dots, \tau_q + d, \dots, \tau_N) \quad (3.3)$$

wegen $\tau_p = \tau_q$ und der Symmetrievoraussetzung zu

$$\Phi(\tau_1, \tau_2, \dots, \tau_p + d, \dots, \tau_q - d, \dots, \tau_N) \quad (3.4)$$

identisch ist. Die Abstandsvektoren $\mathbf{d}^+ = (0, \dots, d, \dots, -d, \dots, 0)$ und $\mathbf{d}^- = (0, \dots, -d, \dots, d, \dots, 0)$ zum Mittel- und Symmetriepunkt stehen dabei orthogonal zur $N - 1$ -dimensionalen Spiegele-

²Die Transmission (die Wahrscheinlichkeit für ein Elektron den Kanal zu durchlaufen) kann jeweils durch Multiplikation mit $G_0 = \frac{2e^2}{h} \approx \frac{1}{12906\Omega}$ direkt in den Leitwert des Kanal umgerechnet werden.

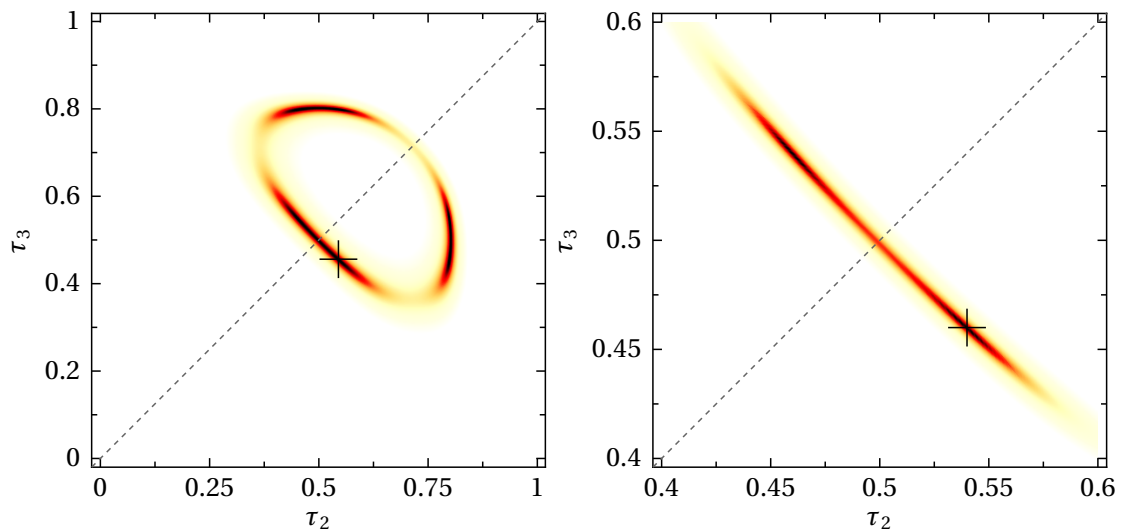


Abbildung 3.2.: Links: Fehler beim Fitten mit drei Parametern, aufgetragen gegen zwei der Parameter in unterschiedlichen Ausschnitten. Dargestellt ist der jeweilige minimale Fehler bezüglich des dritten Parameters in Farben von dunkel (klein) bis hell (groß) in stark verzerrter Farbskala. Gefittet wird hier eine theoretische Kurve mit den Kanälen 0.8, 0.54 und 0.46. Das gefundene Minimum (schwarzes Kreuz) liegt wie zu erwarten bei einem der $3! = 6$ Spiegelbilder mit $\tau_1 = 0.8000$, $\tau_2 = 0.5400$ und $\tau_3 = 0.4600$. Rechts: Ausschnittsvergrößerung.

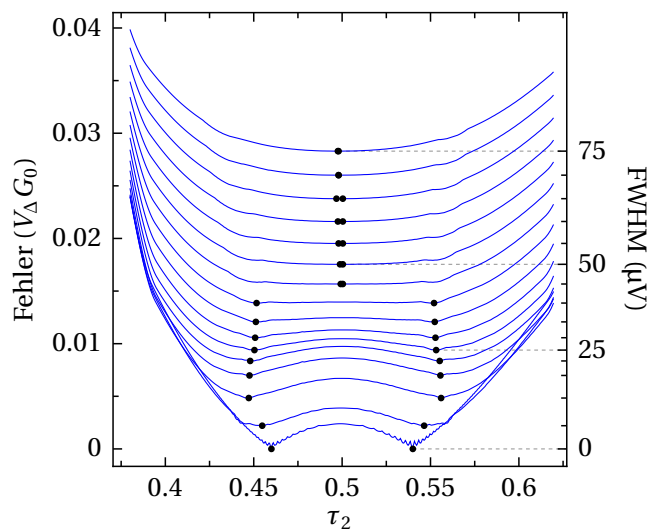


Abbildung 3.3.: Verlauf des Minimums entlang der τ_2 -Richtung mit den globalen Minima, dargestellt durch schwarze Punkte. Von unten nach oben wurde die zu fittende Kurve zunehmend verrundet. Die rechte Skala ordnet den einzelnen Kurven die entsprechende Gaußbreite der numerischen Verrundung zu. Ab einer Verrundung von $45 \mu\text{V}$ fallen in diesem Beispiel die Minima fälschlicherweise zusammen.

bene $S_{p,q}$ wie das verschwindende Skalarprodukt

$$\langle \mathbf{d}^+, \mathbf{s}_2 - \mathbf{s}_1 \rangle = d(\tau_{p,2} - \tau_{p,1}) - d(\tau_{q,2} - \tau_{q,1}) = 0 \quad \text{mit } s_1, s_2 \in S_{p,q} \quad (3.5)$$

zeigt. Abbildung 3.2 zeigt ein Beispiel eines Potenzials mit einer sichtbaren Symmetrieebene bzw. -linie und sechs identischen Minima, sowie eine Ausschnittsvergrößerung.

Entartung des globalen Minimums

Eine weitere Symmetrieeigenschaft des Problems zeigt sich in der speziellen Form der gegebenen Kurven. Hier hat der lineare Bereich im höheren Spannungsbereich ein besonders starkes Gewicht. In rein linearen Kurven würde die fehlende lineare Unabhängigkeit dazu führen, dass in Φ kein eindeutiges globales Minimum mehr existiert. Vielmehr tritt eine Entartung der eigentlichen Minima \mathbf{w} in die Richtungen gleicher Summe $\sum \tau_i$ auf. Die identischen Fehler liegen auf einer $N - 1$ -dimensionalen Ebene

$$K_{\mathbf{w}} := \{ \mathbf{w} + (\delta_1, \delta_2, \dots, \delta_N) \mid \delta_i \in [0, 1], \sum \delta_i = 0 \}, \quad (3.6)$$

die aufgrund von

$$\langle (\delta_1, \delta_2, \dots, \delta_N), (1, 1, \dots, 1) \rangle = \sum \delta_i = 0 \quad (3.7)$$

senkrecht zum Ursprungsvektor $(1, 1, \dots, 1)$ und zu der in allen Spiegelebenen enthaltenen Raumdiagonalen steht. Damit schneiden sich alle Spiegelebenen $S_{p,q}$ mit den Ebenen $K_{\mathbf{w}}$ mit identischem Fehler. Da es nicht identische, sondern nur ähnliche Eigenschaften sind, ergeben sich zwar dennoch absolute Minima, die jedoch in länglich-flachen Potenzialmulden liegen, wie auch in Abb. 3.2 zu sehen ist.

Konsequenz

Befinden sich im Experiment zwei oder mehr Kanäle bei ähnlichen Transmissionen, so ist der Abstand zur Spiegelebene und auch zu den Spiegelbildern gering. Verringert sich aufgrund von Messungenauigkeiten die Tiefe der Potenzialmulde um das globale Minimum, besteht die Gefahr, dass sich die Mulde und ihr Spiegelbild gegen die Konkavform des umgebenden Potenzials nicht mehr durchsetzen können. Dadurch verschmelzen beide spiegelbildliche Mulden miteinander und das globale Minimum rutscht genau auf die Spiegelebene. Als Konsequenz werden zwei oder mehr ähnliche Kanäle scheinbar identisch. Abb. 3.3 zeigt im Querschnitt eines Potenzials, wie ab einem bestimmten Rauschen der ungewollte Fall zweier scheinbar gleicher Kanäle auftritt.

3.1.3. Anzahl der Theoriekurven

Die Simulation und Berechnung einer Theoriekurve mit dem Programm von C. CUEVAS benötigt einige Rechenzeit, so dass es sinnvoll ist, die Kurven einmalig in genügend kleinem Raster zu berechnen und zu speichern. Für 5000 Kurven wurden mehrere Tage auf verteilten

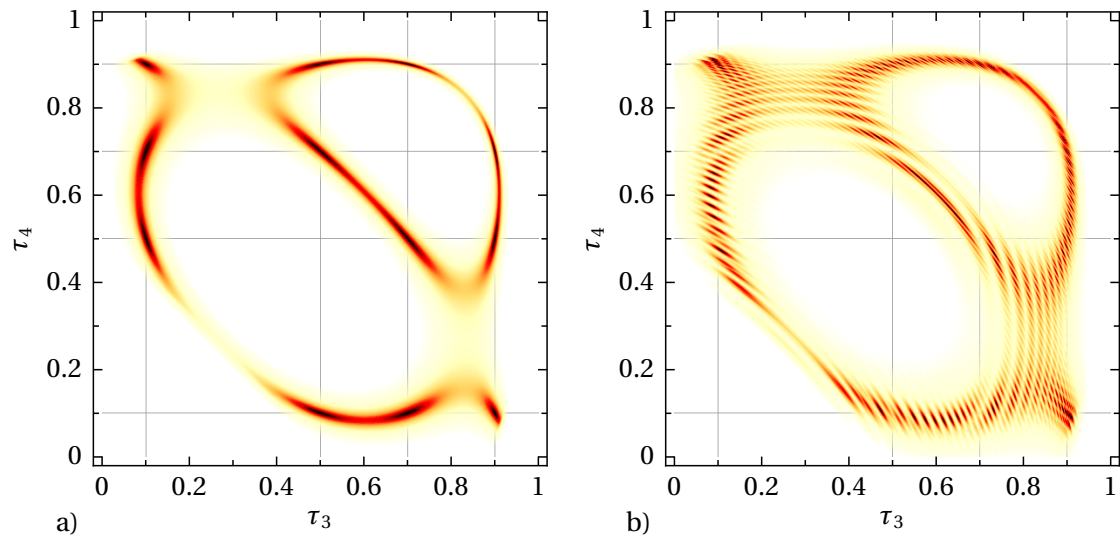


Abbildung 3.4.: Fit-Fehler als Farbskala bei einer Testkurve mit den Kanälen 0.1, 0.5, 0.7 und 0.9, aufgetragen gegen zwei der insgesamt vier Fitparameter τ_3 und τ_4 in einem Raster von je 500 Schritten. Die beiden anderen Kanäle τ_1 und τ_2 wurden in a) 1000 Schritten und b) 50 Schritten gerastert. Die Einschränkung der Genauigkeit der Kanäle τ_1 und τ_2 auf ein Raster von 0.02 wirkt sich in der Ebene von τ_3 und τ_4 sehr viel stärker aus bis zu einem um ca. 0.1 falsch positionierten Minimum.

Rechnern beansprucht. 5000 Abstufungen entsprechen bei einer Transmission zwischen 0 und 1 einem Raster im Abstand 0.0002. Das scheint zunächst übertrieben genau, da sich im Fit eine absolute Genauigkeit der einzelnen Kanäle besser als 0.05 kaum begründen lässt. Wie sich aber zeigt, wirkt sich eine Ungenauigkeit eines Kanals sehr viel stärker auf die Genauigkeit anderer Kanäle aus. Insbesondere gilt das für Abweichungen bei Kanälen mit hohen Transmissionen, die sich sehr stark auf kleine Kanäle auswirken. Abb. 3.4 zeigt den Fehler als Farbskala aufgetragen gegen zwei der insgesamt vier Kanäle. Während in a) die beiden nicht aufgetragenen Kanäle τ_1 und τ_2 mit 1000 Zwischenwerten gerastert wurden, zeigt b) das Ergebnis für nur 50 Zwischenwerte. Die Quantisierungsartefakte führen zu lokalen Minima an falschen Positionen. Die Rauigkeit der Potenziellandschaft erschwert zudem die Konvergenz des Monte-Carlo-Optimierungsverfahrens, das weiter unten beschrieben wird. 5000 scheint eine ausreichende Zahl von Kurven und Transmissionswerten zu sein. Liegen die Kurven in größerem Raster vor, könnten Zwischenwerte linear interpoliert werden. Hier haben jedoch Versuche mit Kanalhistogrammen gezeigt, dass etwa 100 Stützstellen noch grenzwertig sind. Außerdem vergrößert die Interpolation die Rechenzeit pro Iteration um den Faktor 2 bis 3.

3.1.4. Bandlücke

Die theoretisch berechneten IV-Kurven (s. Abb. 3.1, S. 27) gelten universell für eine große Klasse leitender Nanostrukturen, unabhängig von Materialparametern mit Ausnahme der Bandlücke der Supraleitung V_{Δ} . Diese Größe geht jedoch in die IV-Kennlinien nur als lineare Skalierung

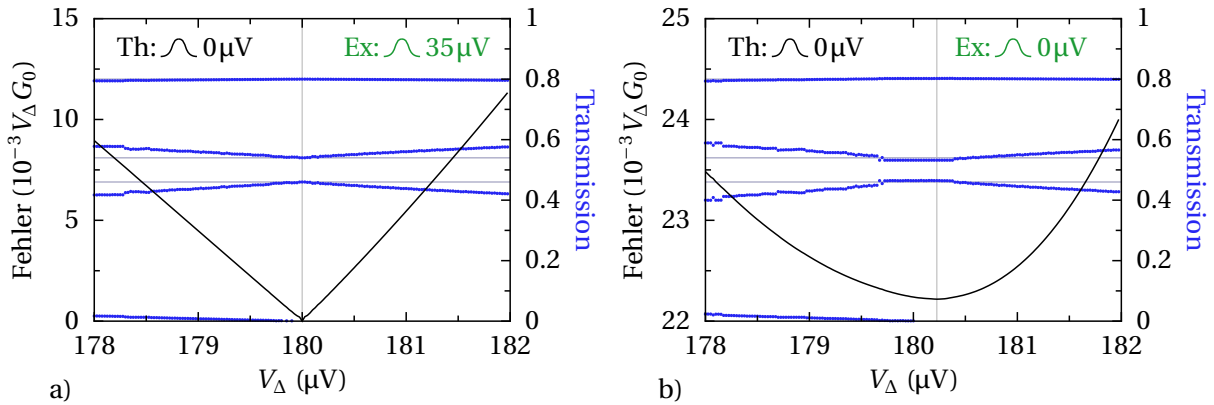


Abbildung 3.5.: Fit einer Testkurve mit drei Kanälen (0.8, 0.54, 0.46) in Abhängigkeit von V_{Δ} . Minimaler Fehler (schwarze Linie) mit den jeweils resultierenden Kanälen (blaue Punkte) bei a) nicht verrundeter und b) verrundeter (35 μV) Testkurve. Die vertikale graue Linie liegt an der Position des minimalen Fitfehlers. Die horizontale grauen Linien entsprechen den wahren Kanalwerten.

ein, so dass bei geeigneter Wahl der Einheiten die Materialabhängigkeit eliminiert wird. Beim Vergleich mit dem Experiment spielt der Wert jedoch eine Rolle, da die Einheiten ineinander umgerechnet werden müssen. V_{Δ} hat beim hier verwendeten Material Aluminium einen Wert von ca. 180 μV . Neben dem Material ist auch ein geringfügiger Einfluss der Geometrie des Kontakts auf V_{Δ} von maximal 1 μV denkbar, was die Analysen andeuten. In wie weit sich ein nicht genau bekannter Wert von V_{Δ} negativ auf die Bestimmbarkeit der Kanäle auswirkt, soll nun anhand von Beispielen untersucht werden.

Zunächst werden aus Theoriekurven zusammengesetzte Testkurven untersucht. Konstruiert wurden die Kurven mit einer Bandlücke von $V_{\Delta} = 180 \mu\text{V}$. Analysiert wird mit Werten zwischen 178 μV und 182 μV . Abbildung 3.5 a) zeigt Untersuchungen von Testkurven mit Kanälen identisch zu den Beispielen des vorigen Kapitels. Als schwarze durchgezogene Linie ist der minimale Fehler des Fits eingezeichnet. Bei der tatsächlichen Bandlücke von 180 μV ist dieser Fehler Null. Abweichende Werte führen zu höheren, nach beiden Seiten etwa linear ansteigenden Fehlern. Die zugehörigen Kanäle des jeweiligen Fits sind durch blaue Punkte dargestellt. Sie weichen zunehmend von den wahren Werten (graue horizontale Linien) ab. Bei realistischer Verrundung der konstruierten Kennlinien mit 35 μV (Abb. 3.5 b)) liegt das Minimum des Fehlers zwar weiterhin bei 180 μV , die Kanäle zeigen bei diesem Wert jedoch kleine Abweichungen.

Sehr viel empfindlicher ist das Ergebnis mit einem zusätzlichen Kanal bei 0.25 (siehe Abb. 3.6). Hier werden im Fall einer unverrundeten Testkurve (a), wie zu erwarten, die richtigen Kanäle ermittelt. Bei der verrundeten Kurve jedoch (b) liegt erstens das Minimum leicht tiefer bei 179.8 μV und zweitens die Kanäle bei sowohl diesem Wert als auch bei genau 180 μV weit daneben. Bei einem bestimmten größeren Wert $V_{\Delta} \approx 181.3 \mu\text{V}$ werden die Kanäle richtig wiedergegeben. Ein solcher Wert findet sich häufig in vielen anderen Berechnungsbeispielen. Gezielt einen etwas größeren Wert von V_{Δ} zu wählen, eignet sich jedoch nicht als sichere Stra-

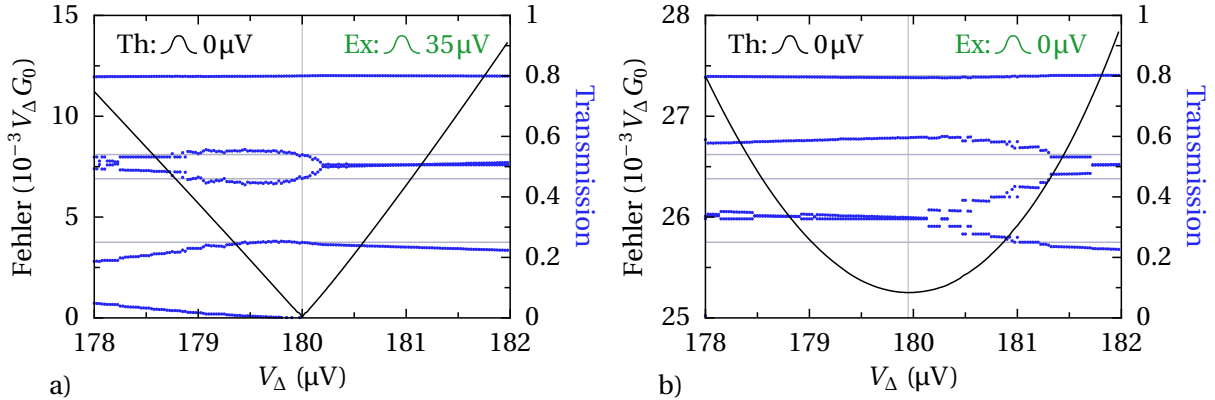


Abbildung 3.6.: Fit-Ergebnisse wie in Abb. 3.5, nur mit vier Kanälen (gegenüber Abb. 3.5 ein weiterer Kanal bei 0.25).

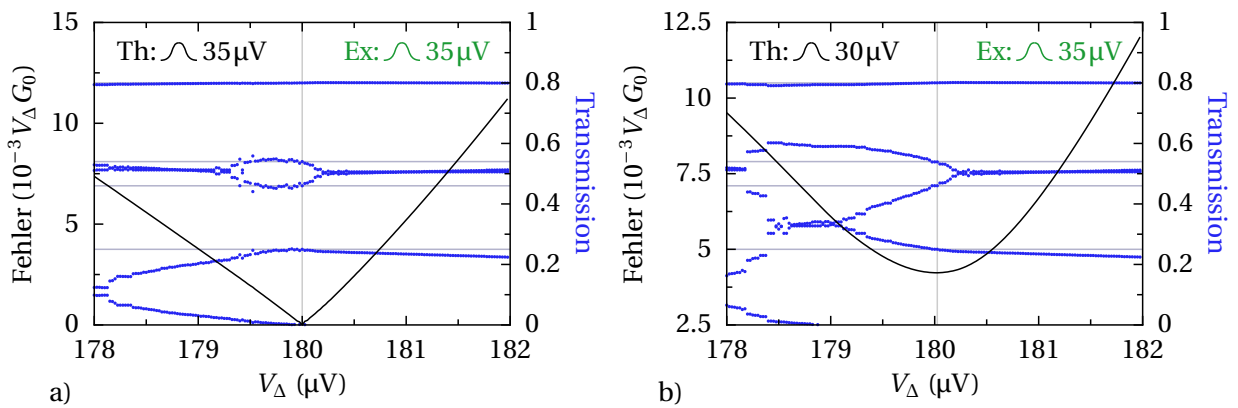


Abbildung 3.7.: Verrundete Testkurve ($35\mu\text{V}$) wie in Abb. 3.5 b) mit verrundeten Theoriekurven mit a) $35\mu\text{V}$ und b) $30\mu\text{V}$. Wird also die Art der Verrundung nicht genau modelliert, kann dadurch das Ergebnis des Fits stark beeinflusst werden.

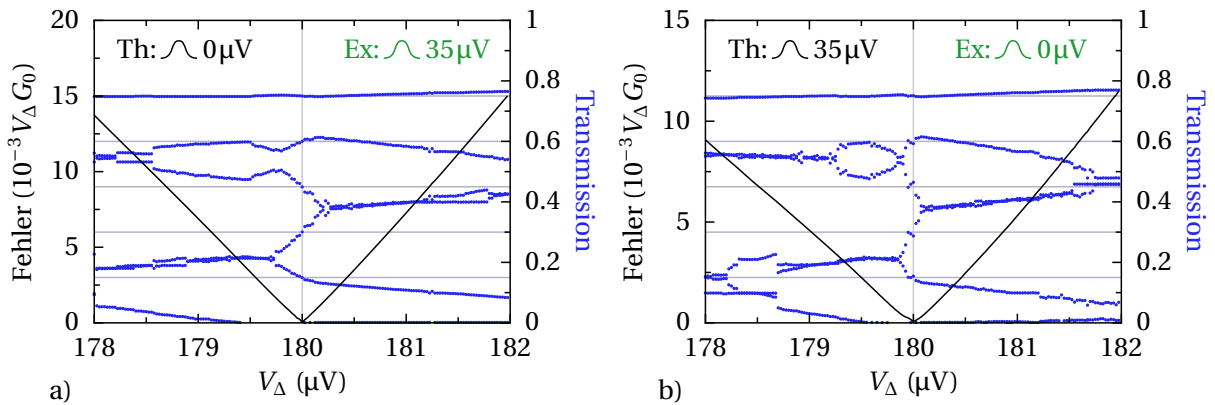


Abbildung 3.8.: Schwierig zu fittender Satz von Kanälen mit den Werten 0.15, 0.30, 0.45, 0.60 und 0.75 mit a) keiner und b) $35\mu\text{V}$ starker Verrundung. Die Bandlücke muss hier relativ genau bekannt sein. Im Normalfall sind die Kanalabstände viel weniger ausgeglichen, was die Gefahr der Verschmelzung zweier Kanäle noch begünstigt.

ategie zur Ermittlung der richtigen Kanäle, da der Abstand zu $180\ \mu\text{V}$ oder zum Minimalfehler zu sehr variiert. Die Abweichungen sind jedoch oberhalb von $180\ \mu\text{V}$ im Allgemeinen etwas geringer.

Als alternative Strategie wird nun versucht, die Theoriekurven, die zum Fitten kombiniert werden, dem Experiment besser anzugleichen, indem diese möglichst realistisch verrundet werden. Dazu wird eine Faltung mit einer GAUSS'schen Glockenfunktion durchgeführt als Modellierung der im Experiment vorhandenen Rauschverteilung. Das rein elektrische Rauschen wird dabei als Hauptquelle der Verrundung gesehen und die endliche Temperatur als effektiv ähnlicher Effekt eingestuft. Das Ergebnis zeigt Abb. 3.7. In Bild a) wurde dieselbe Verrundung von $35\ \mu\text{V}$ sowohl auf die Theoriekurven als auch auf die zu fittende Kurve angewendet. Wie zu erwarten, verbessert dies nicht die Stabilität des Fits gegenüber unverrundeten Kurven. Bei korrekter Wahl von V_{Δ} liefert der Fit jedoch die richtigen Kanäle. In Bild b) wurde dieselbe mit $35\ \mu\text{V}$ verrundete Messkurve untersucht, jedoch schwächer verrundete Theoriekurven ($30\ \mu\text{V}$) zum Fitten verwendet. Dies führt immer noch zu einem Gewinn an Genauigkeit gegenüber dem Fit mit unverrundeten Theoriekurven (siehe Abb. 3.6 b)) bei sowohl $180\ \mu\text{V}$, als auch an der Position des Minimalfehlers. Das Ergebnis zeigt jedoch auch, dass die genaue Art der Verrundung einen starken Einfluss hat. Die etwa gaußförmige Rauschverteilung konnte im Experiment zwar direkt gemessen werden, jedoch war unklar, welcher Anteil des Rauschens an der Probe anliegt und welcher z. B. vom Verstärker im Nachhinein aufgeprägt wird. So stand keine genaue gemessene Verteilung für die Faltung zur Verfügung. Eine GAUSS-Kurve konnte jedoch durch systematische Fit-Versuche in der Breite recht gut angepasst werden. Eine gewisse Bestätigung lieferte auch der Vergleich des Fitfehlers zum Symmetriefehler. Die Analyse der echten experimentellen Daten wird im Kapitel 5 gezeigt.

3.1.5. Fit-Algorithmen

Wie im vorigen Abschnitt gezeigt, reduziert sich das Problem der Kanalanalyse numerisch auf die Bestimmung eines globalen Minimums einer mehrdimensionalen Potenziellandschaft, gebildet aus der quadratischen Abweichung des Fits vom Experiment. Hierfür kommen verschiedene Strategien bzw. Algorithmen in Frage, von denen sich das *simulierte Ausheilen* als am vorteilhaftesten herausgestellt hat. Die GAUSS'sche Regression erwies sich als instabil und das systematische Abrastern aller Werte als zu rechenaufwendig.

Systematisches Abrastern

Das vollständige Abrastern aller möglichen Parameter erforderte eine Zahl von Rechenschritten proportional zur Ordnung $O(N^n)$, wobei n die Zahl der zu erwarteten Kanäle ist und N die Zahl der zur Verfügung stehenden Theoriekurven (hier ist $N = 5000$). Abbildung 3.2 zeigt das Ergebnis eines solchen Scans bei drei Kanälen, was einige Minuten Rechenzeit benötigt. Bei vier Kanälen müsste die Rechenzeit bereits in Tagen bemessen werden, ab fünf Kanälen in Jahren. Die Berücksichtigung der Symmetrie beschleunigt die Berechnung um Faktor $n!$. Eine Einschränkung auf kleine Bereiche um das Minimum reduziert ebenfalls die Rechenzeit, erfordert jedoch eine manuelle Auswahl oder eine vorausgegangene gröbere Rasterung.

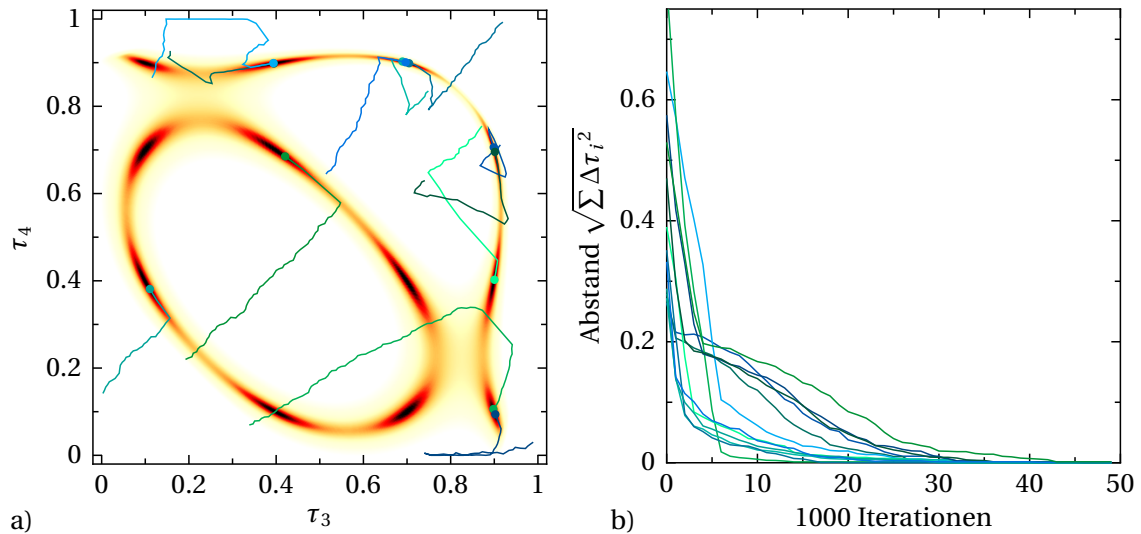


Abbildung 3.9.: Beispiel von 12 Fits mit simuliertem Ausheilen einer aus den Werten 0.1, 0.4, 0.7 und 0.9 zusammengesetzten Kurve. a) Pfade der simulierten Teilchentrajektorien ausgehend von einem zufälligen Startpunkt entlang der ersten 50 000 Iterationsschritte bis zu einem der 24 Minima (hier sind nur $4!/2! = 12$ sichtbar, da in dieser Projektion je zwei aufeinander liegen). b) Kleinster Abstand zu einem der Minima aufgetragen gegen die ersten 50 000 Iterationsschritte. 1 Million Iterationsschritte benötigen etwa eine Sekunde Rechenzeit.

GAUSS'sche Regression

Natürlich wäre es attraktiv, das globale Minimum direkt analytisch bestimmen zu können. Dies ist durch die GAUSS'sche Regressionsmethode (siehe Abschnitt 3.2.2) möglich, sofern die Fitparameter linear in die Fitfunktion eingehen. Das ist hier zwar nicht der Fall, jedoch könnte man alle einzelnen zur Verfügung stehenden Theoriekurven mit einem linearen Faktor als freien Fitparameter versehen. Auf diese Weise würde das Problem linear werden, jedoch mit zusätzlichen Freiheitsgraden. Zu hoffen wäre nun, dass die zusätzlichen Freiheitsgrade nicht besetzt werden und dass fast alle Faktoren den Wert Null bekommen, mit Ausnahme von einigen wenigen, die auf ganze Zahlen gerundet auftreten. Der ganzen Zahl entspräche dann die Anzahl der Kanäle zur zugehörigen Transmission. Versuche zeigten jedoch, dass dies nur mit exakten (künstlich erstellten) Messkurven möglich ist. Bereits bei kaum sichtbarem Rauschen erhöht sich die Zahl der Kanäle drastisch und die Faktoren nehmen verschiedenste nicht ganzzahlige und auch negative Werte an. Der verbleibende Fehler wird dadurch zwar verschwindend klein, und selbst verrauschte Messkurven werden perfekt approximiert, jedoch unter Verwendung von physikalisch unsinnigen Parametern, wie negativen Transmissionen. Das Verfahren ist also leider ungeeignet. Immerhin fand sich in den GAUSS'schen Gleichungen eine sehr effiziente Methode zur Berechnung von Φ (siehe Kapitel 3.1.6).

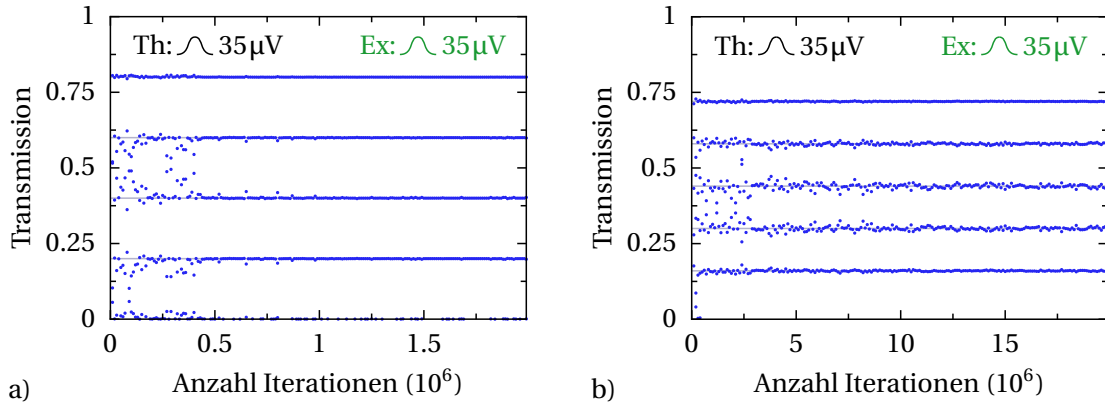


Abbildung 3.10.: Unterschiedliche Fits in Abhängigkeit von der Anzahl der Iterationsschritte. Alle Fits sind unabhängig voneinander und beginnen mit Zufallswerten. a) Fit von vier weit voneinander entfernten Kanälen (0.8, 0.6, 0.4, 0.2) mit fünf freien Parametern. b) Fünf etwas dichter liegende Kanäle (0.72, 0.68, 0.44, 0.3, 0.16) gefittet mit sechs freien Parametern. Hier wird für ein stabiles Ergebnis mehr als die zehnfache Zahl an Iterationsschritten benötigt.

Simuliertes Ausheilen

Als besonders geeignet stellte sich die Methode des *simulated annealing* (oder *simuliertes Ausheilen*) heraus³, ein Monte-Carlo-Optimierungsverfahren, das mit hoher Wahrscheinlichkeit schnell ein genaues Ergebnis liefert. Der Algorithmus erlaubt es, ein globales Minimum in einer Potenziallandschaft zu finden, falls die lokalen Minima weder sehr tief sind, noch isoliert auftreten. Diese beiden Anforderungen erfüllt das untersuchte Problem, wie gezeigt werden kann. Das Verfahren startet zunächst an einer zufälligen Anfangsposition in der Potenziallandschaft und ändert die Koordinaten um kleine Schritte in zufällige Richtungen. Der Schritt wird verworfen, falls es bergauf geht und gleichzeitig die Bedingung $\exp(-\Phi_{\text{neu}}/\Phi_{\text{Ref}}) > r$ erfüllt ist, mit dem Fehlerquadrat der neuen Position Φ_{neu} , einem Parameter Φ_{Ref} und einer Zufallszahl $r \in [0, 1]$. Dies entspricht der Bewegung eines Teilchens in einer Potenziallandschaft bei einer Temperatur Φ_{Ref} und der Besetzungswahrscheinlichkeit eines höher liegenden Potentials gemäß der BOLTZMANN-Verteilung. Das Teilchen bewegt sich also mit sehr großer Wahrscheinlichkeit bergab. Dennoch kann es aus einem flachen lokalen Minimum mit einer gewissen Wahrscheinlichkeit entkommen. Die Temperatur kann je nach Tiefe der lokalen Minima an das Problem angepasst werden. Auch während des Ablaufs kann die Temperatur langsam gesenkt werden. Das bringt in diesem Fall jedoch keine Vorteile, da die lokalen Minima hier nur auf kleiner Skala auftreten. Auf der gesamten Trajektorie wird die Position des minimalen Potentials gespeichert. Hat der Algorithmus die Mulde des globalen Minimums näherungsweise erreicht, verweilt er dort für alle kommenden Iterationsschritte und tastet die unmittelbare Umgebung aufgrund der endlichen Temperatur gründlich ab. Um die Aufenthaltszeit in kleinen Mulden zusätzlich zu begrenzen, gibt ein weiterer Parameter (Kickoff) eine Wahrscheinlichkeit für einen Bewegungsschritt an, der unabhängig vom Temperaturkriterium ausgeführt wird. Ta-

³Die Idee zur Verwendung der Methode stammt aus der Gruppe von GABINO RUBIO-BOLLINGER aus Madrid.

Parameter	guter Wert	Beschreibung
Strength	0.001	Maximale Stärke der Zufallsänderung in jeder Iteration. Da die Theoriekurven im Raster 0.0002 vorliegen, dürfte diese maximale Schrittweite ausreichen. Höhere Werte führen zu langsamerer Konvergenz.
ErrRef	10^{-7}	Temperatur für die Besetzungswahrscheinlichkeit höherer Potentiale. Dieser Wert ist so gewählt, dass bei großen Fehlern praktisch nur Bewegungen nach unten möglich sind und bei kleinen Fehlern nur ab und zu Verschlechterungen stattfinden.
KickOff	0.0003	Wahrscheinlichkeit für Änderung in jedem Fall. Dieser Parameter zeigt seine Wirkung in seltenen Fällen an Sattelpunkten oder bei kleinen tiefen Mulden.

Tabelle 3.1.: Bedeutung der Parameter beim Fit der IV-Kennlinien mit simuliertem Ausheilen mit der Angabe gut geeigneter Werte.

belle 3.1 zeigt eine Übersicht aller Parameter und gibt Werte an, die als geeignet befunden wurden.

Zur Demonstration der Methode zeigt Abb. 3.9 den Verlauf eines Fits mit der Methode der simulierten Ausheilung. In a) wurde der bezüglich zweier Parameter minimale Fehler farbkodiert gegen die beiden anderen der insgesamt vier freien Parameter aufgetragen. Die schwarzen Stellen zeigen die spiegelbildlichen globalen Minima. Als Startpunkt werden nun vier Zufallswerte gewählt und die Trajektorie während der Iterationsschritte eingezeichnet. Ein Punkt markiert die Stelle, die nach 500 000 Iterationen erreicht wurde. b) zeigt den Abstand zum nächstgelegenen Minimum, aufgetragen gegen die Iterationsschritte. Nach 50 000 Iterationen wurde hier in den 12 Beispielen das Minimum sehr gut erreicht. Die Trajektorie bewegt sich erst sehr schnell, dann immer langsamer auf ein Minimum zu. Fern von den Minima scheint das Potenzial in der Regel sehr steil zu sein, was auch Abb. 3.3 zeigt. In seltenen Fällen bleibt die Trajektorie kurzzeitig an einem Sattelpunkt hängen.

Ein weiteres Beispiel soll ein Gefühl dafür vermitteln, wie viele Iterationsschritte sinnvoll und ausreichend sind. Leichte Fitprobleme bei 4 Kanälen wie in 3.9 b) liefern schon nach 50 000 Iterationen ein recht gutes Ergebnis. Abb. 3.10 a) zeigt Fits von 5 freien Parametern bei 4 gegebenen Kanälen im Abstand 0.2. Die Fits sind voneinander unabhängig berechnet worden mit jeweils zufälliger Startposition. Mindestens 500 000 Iterationen sollten hier abgewartet werden. In Abbildung b) beträgt der Abstand der Kanäle nur 0.14. Hier werden 5 Millionen Iterationen benötigt. Die unterschiedlichen Kanäle lassen sich unterschiedlich gut fitten. Besonders gut konvergieren im Allgemeinen die obersten Kanäle. Schlecht konvergieren die Kanäle mit benachbarten Kanälen und solche mit sehr ähnlichen IV-Kennlinien, wie etwa zwischen 0.2 und 0.5.

Insgesamt zeigen die Beispiele, dass aufgrund der asymptotischen Konvergenz in der Regel völlig falsche Ergebnisse nach einigen Iterationen selten sind, sofern das globale Minimum an der richtigen Position liegt. Letzteres dürfte in der Regel die größere Fehlerquelle darstellen.

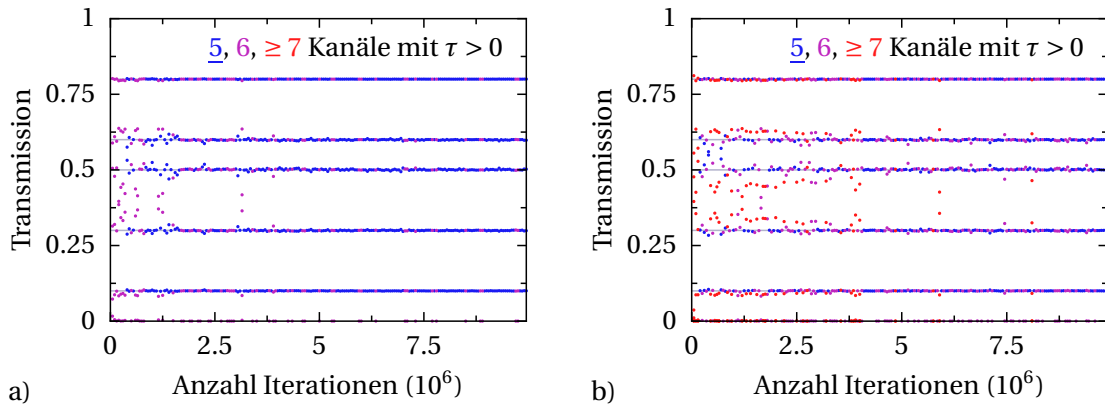


Abbildung 3.11.: Abhängigkeit eines Fits von der Anzahl der Iterationsschritte wie in Abb. 3.10 bei fünf vorgegebenen Kanälen 0.8, 0.6, 0.5, 0.3, 0.1 wenn a) maximal 6 freie Parameter und b) maximal 20 freie Parameter zugelassen sind. Trotz der 20 freien Parameter ist die Konvergenz nicht wesentlich langsamer. Nur in wenigen Fällen bleiben von den anfänglichen 20 Kanälen am Ende mehr als 6 offen, wie der Farbcode zeigt. Häufig ist selbst der Reservekanal exakt Null.

Automatische Bestimmung der Zahl der Kanäle

Die Anzahl der Kanäle ist vor dem Fit nicht bekannt. Daher muss mit einer genügend großen Zahl an freien Parametern begonnen werden. Jeder Freiheitsgrad verlangsamt jedoch die Konvergenz des Fits. Bei der simulierten Ausheilung macht dies jedoch nur etwa einen Faktor 2 aus, wenn der zusätzliche Freiheitsgrad nicht besetzt wird. Das liegt daran, dass die Kanäle, die sich in der Nähe von Null aufhalten, bei der Zufallsänderung jedes zweite Mal auf einen negativen Wert fallen und negative Werte auf Null gesetzt werden. So ist jeder zweite Iterationsschritt identisch zu einem Fit mit einem auf Null fixierten Kanal bzw. einer um eins kleineren Zahl von Freiheitsgraden. Bei zwei zusätzlichen Kanälen ist das im Schnitt jedes vierte Mal der Fall und so weiter. n zusätzliche freie Fitparameter verlangsamen die Konvergenz also nur um 2^n , nicht um $(N + n)!/N!$. Dennoch lohnt es sich, mit einer Erweiterung des Algorithmus die Konvergenz zu beschleunigen, indem die Zahl der Kanäle automatisch angepasst wird.

Um das zu erreichen, wird, sobald mehr als zwei Kanäle den Wert Null erreicht haben, ein Freiheitsgrad entfernt. Wenn kein Kanal mehr den Wert Null hat, wird dagegen wieder ein Freiheitsgrad hinzugefügt, bis zu einer vorgegebenen Maximalzahl von Freiheitsgraden. Es wird also immer genau ein Reservekanal offen gelassen, der den momentanen Wert Null hat. Änderungen der Anzahl der Freiheitsgrade sollten allerdings nicht in jedem Iterationsschritt erfolgen, sondern (zur Verringerung der Rechenzeit) immer nur dann, wenn ein neuer Bestwert gefunden wird.

Abbildung 3.11 demonstriert das Verhalten bei unterschiedlich vielen Freiheitsgraden. In beiden Fällen wurden Fits mit unterschiedlicher Anzahl von Iterationen gestartet. Dabei wurden zunächst alle Werte auf Zufallszahlen gesetzt. Nach einer gewissen Zeit werden in der Regel alle überflüssigen Freiheitsgrade, bis auf einen, geschlossen und die Konvergenz wird so insgesamt nicht verlangsamt.

Im Folgenden wird der Quelltext des Algorithmus in vereinfachter Notation gezeigt (Pseudocode). Er wurde in C++ kompiliert und als dynamische Bibliothek (DLL) in Python als Funktion implementiert. Grundlage diese Codes war ein Algorithmus von R. GABINO-BOLLINGER. Die Verbesserungen bestehen im Verzicht auf die Interpolation, in der effizienteren Berechnung des Fehlers und in der Erweiterung, die die Kanäle um Null automatisch sperrt oder wieder öffnet. Alle drei Verbesserungen dienen der Geschwindigkeitssteigerung der Konvergenz um mindestens Faktor 20.

```

1  for count = 0..maxiter:                                     //Iterationen zählen
2  //----- Berechnung des neuen Fehlers -----
3  ErrNew = c                                                //Φ = c + ...
4  for i = 0..channels-1:
5      ErrNew += b[TNew[i]]                                  //+bi
6      for j = 0..i-1:
7          ErrNew += A[TNew[i],TNew[j]]                    //+Aij
8  //----- Verbesserung eingetreten? -----
9  if ErrNew < ErrNow:
10     ErrNow = ErrNew                                       //→ übernehmen
11     for i = 0..channels-1: TNow[i] = TNew[i]             //→ übernehmen
12 //----- Verschlechterung eingetreten? -----
13 else:
14     if ((exp(-ErrNew/ErrRef)) > (random(0..1))):         //Temperaturkriterium
15         ErrNow = ErrNew                                   //→ übernehmen
16         for i = 0..channels-1: TNow[i] = TNew[i]         //→ übernehmen
17     else:
18         if random(0..1) < KickOff:                       //Kick-Off-Kriterium
19             ErrNow = ErrNew                               //→ übernehmen
20             for i = 0..channels-1: TNow[i] = TNew[i]     //→ übernehmen
21 //----- neuer Bestwert gefunden? -----
22 if ErrNow < ErrBest:
23     ErrBest = ErrNow                                     //→ speichern
24     zerochannels = 0
25     for i = 0..channels-1:
26         TBest[i] = TNow[i]                               //→ speichern
27         if TNow[i] = 0: zerochannels += 1                 //zähle Nullkanäle
28 //----- Kanäle öffnen oder schließen? -----
29     if zerochannels ≠ 1:
30         if zerochannels > 1:                               //Kanal schließen?
31             i = 0; while TBest[i] ≠ 0: i += 1             //erster Nullkanal?
32             TNow[i] = TBest[i] = TBest[channels-1]       //leeren Kanal...
33             TNow[channels-1] = TBest[channels-1] = 0    //...ans Ende setzen
34             if channels > 1: channels -= 1               //→ Kanal schließen
35         else:
36             if channels < channelsMax: channels += 1     //→ Kanal öffnen
37 for i = 0..channels-1:                                    //Zufallsänderung
38     TNew[i] = min(TMax,max(0,TNow[i]+int (Strength*(random(-1..1))))
39 //----- Ende -----

```

Listing 3.1: Simuliertes Ausheilen mit automatischer Kanalreduktion (Pseudocode)

3.1.6. Effiziente Berechnung des Fitfehlers Φ

Die Effizienz der Berechnung des Fitfehlers

$$\Phi(\tau_1, \tau_2, \tau_3, \dots, \tau_n) = \sqrt{\int \left(\sum_{i=1}^n I_{\text{Theo}}(\tau_i, U) - I_{\text{Mess}}(U) \right)^2 dU} \quad (3.8)$$

pro Iterationsschritt scheint zunächst durch die Integration begrenzt zu sein. Sie entspricht bei diskreten Messwerten der Summation über die Anzahl der Stützstellen. Bei üblicherweise etwa 300 Messpunkten sind $300(n+2)$ Rechenoperationen bei n Kanälen notwendig. Diese Zahl kann jedoch durch geeignete Vorbereitung auf $\frac{1}{2}n(n+1)$ Additionen verringert werden. Dazu wird ausgenutzt, dass nur endlich viele (hier $N = 5000$) theoretische Kurven φ_j bei diskreten Transmissionswerten j/N vorliegen. Mit

$$f \equiv I_{\text{Mess}}(U), \quad g \equiv \sum_{i=1}^n I_{\text{Theo}}(\tau_i, U) \quad \text{und} \quad \varphi_j \equiv I_{\text{Theo}}(j/N, U) \quad (3.9)$$

und mit ganzzahligen Koeffizienten $\alpha_j \in \mathbb{N}$ (≥ 2 bei mehreren gleichen Kanälen), so dass gilt

$$g = \sum_{i=1}^n \varphi_{(\tau_i, N)} = \sum_{j=1}^N \alpha_j \varphi_j, \quad (3.10)$$

kann der Fehler umformuliert werden zu

$$\Phi^2(\tau_1, \tau_2, \dots, \tau_n) = \|g - f\|_2^2 = \langle g - f, g - f \rangle \quad (3.11a)$$

$$= \left\langle \sum_{j=1}^N \alpha_j \varphi_j - f, \sum_{k=1}^N \alpha_k \varphi_k - f \right\rangle \quad (3.11b)$$

$$= \langle f, f \rangle - 2 \sum_{j=1}^N \alpha_j \langle f, \varphi_j \rangle + \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \langle \varphi_j, \varphi_k \rangle \quad (3.11c)$$

$$= \langle f, f \rangle + \sum_{j=1}^N \alpha_j \left(-2 \langle f, \varphi_j \rangle + \langle \varphi_j, \varphi_j \rangle + 2 \sum_{k=1}^{j-1} \alpha_k \langle \varphi_j, \varphi_k \rangle \right) \quad (3.11d)$$

$$= c + \sum_{j=1}^N \alpha_j \left(b_j + \sum_{k=1}^{j-1} \alpha_k A_{jk} \right) \quad (3.11e)$$

$$= c + \sum_{j=1}^n \left(b_{(\tau_j, N)} + \sum_{k=1}^{j-1} A_{(\tau_j, N)(\tau_k, N)} \right). \quad (3.11f)$$

In 3.11e wurde $A_{jk} := 2 \langle \varphi_j, \varphi_k \rangle$, $b_j := \frac{1}{2} A_{jj} - 2 \langle f, \varphi_j \rangle$ und $c := \langle f, f \rangle$ definiert. Zuletzt werden die eingeführten ganzzahligen α_j wieder eliminiert, indem die Indizes einfach oder mehrfach gezählt oder ausgelassen werden. Die Skalarprodukte, d. h. die rechenaufwendigen Integrale, können nun vor der Fitprozedur einmalig berechnet und gespeichert werden. Während b_j

und c für jede Messkurve vor dem Fit einmal berechnet werden muss (ca. 50 ms Rechenzeit auf einem normalen Bürorechner), kann die Matrix A_{jk} einmalig für alle Messkurven berechnet (ca. 5 Minuten Rechenzeit) und für zukünftige Fits gespeichert werden (200 MByte-Datei bei 5000×5000 Werten). Damit ist die Rechenzeit für Φ während des Fits jetzt sehr gering und unabhängig von der Zahl der Stützstellen der Messkurve.

Bei z. B. $n = 4$ Kanälen ist der Rechenaufwand für die Berechnung von Φ pro Iterationsschritt von ursprünglich $300(n + 2) = 1800$ Rechenoperationen (zum Teil Multiplikationen) auf nur noch $\frac{1}{2}n(n + 1) = 10$ Additionen gesunken, also um mindestens Faktor 180 schneller geworden. Für eine Abrasterung benötigt ein aktueller Prozessor (2.8 GHz Quadcore) pro Kern etwa 30 ns Rechenzeit für die Berechnung von Φ , d. h. man erhält 33 Millionen Werte pro Sekunde. Ein Iterationsschritt der simulierten Ausheilung benötigt $0.7 \mu\text{s}$. Die Berechnung von Φ nimmt hier nur noch einen sehr kleinen Teil der Zeit in Anspruch. Pro Sekunde kommt man immer noch auf 1.4 Millionen Iterationsschritte.

Ein Verlust an Rechengenauigkeit macht sich bei der Methode nur in den letzten 4–5 von insgesamt 15–16 Dezimalstellen des Double-Typs bemerkbar, was vernachlässigbar ist.

3.2. Glättalgorithmus

In diesem Abschnitt wird ein Algorithmus zur Kurvenglättung beschrieben, der auf der *lokalen Regression*⁴ basiert und in dieser Arbeit verbessert (auf $O(n)$ beschleunigt und für eine numerisch stabile Interpolation erweitert) und mehrfach angewendet wurde. Er sollte der Untergrundbehandlung von Leitwertfluktuationen dienen, erwies sich jedoch auch als ein nützliches Werkzeug zur Behandlung von verschiedenen Messdaten und ist dabei anderen Glättmethoden überlegen.

Anforderungen

Glättalgorithmen werden oft nur zur besseren grafischen Darstellung von Daten verwendet, um bestimmte Tendenzen für den Betrachter leichter erkennbar zu machen. Hängen jedoch Rechenergebnisse empfindlich von der Art und Weise der Glättung ab, sind die einfachen bekannten Algorithmen unzureichend. Das ist beispielsweise bei der Untergrundbehandlung von Leitwertfluktuationen der Fall, wenn daraus Autokorrelationsfunktionen berechnet werden sollen, oder bei den im Experiment vorliegenden Rauschartefakten, die in theoretischen Stom-Spannungs-Kennlinien (siehe Abschnitt 3.1.4) modelliert werden sollen.

Glättalgorithmen sollten insbesondere zwei Eigenschaften besitzen: Sie sollten die Kurve möglichst glatt darstellen und gleichzeitig im Mittel möglichst wenig von der ursprünglichen Kurve abweichen. Diese beiden kombinierten Anforderungen definieren einerseits qualitativ das Glätten und beschreiben andererseits recht genau die in den Messdaten verborgene Information, die bei den genannten Anwendungsbeispielen herausgelesen werden soll. Das

⁴In der Literatur findet sich z. B. eine ausführliche Behandlung in [Loa99]. Dort wird erwähnt, dass die Methode bereits 1866 vom italienischen Astronom SCHIAPARELLI („Entdecker“ der Marskanäle) verwendet wurde.

Filtern bestimmter Frequenzanteile wäre hier eine unpassende Beschreibungssprache des Problems. Oft werden schon die Begriffe Filtern und Glätten kaum sprachlich unterschieden. Ebenso wenig gezielt werden die verschiedenen Algorithmen eingesetzt, die nur von weitem betrachtet ähnliche Wirkung zeigen. Während beim *Filtern* bestimmte Bestandteile abgetrennt werden sollen, meist Frequenzanteile, drückt *Glätten* die Eigenschaft des Ergebnisses aus, in der Krümmung begrenzt zu sein. Ferner sollte ein Glättalgorithmus auch durch verständliche und dem Problem angepasste, physikalisch motivierbare Kontrollparameter präzise gesteuert werden können.

3.2.1. Bekannte Glättalgorithmen

In der Literatur sind eine Vielzahl von Glättalgorithmen zu finden. Verbreitet in der Sekundärliteratur und in Numerikpaketen sind jedoch ausgerechnet diejenigen, die nicht immer zuverlässig glätten, wie man es qualitativ erwarten würde oder auch von Hand zeichnen könnte. Die bekannten Algorithmen zeigen sogar fast alle am Rand einer schrägen Geraden eine Zunahme der Krümmung.

Sehr verbreitet und leicht zu programmieren ist der gleitende Durchschnitt, bei dem über eine feste Zahl von benachbarten Messpunkten gemittelt wird. In Nachschlagewerken zur Numerik (z. B. *Numerical Recipes* [Pre02]) oder in numerischen Softwarepaketen (z. B. NumPy/Scipy, Origin, NAG-library) finden sich jedoch auch fortgeschrittenere Algorithmen, wie z. B. eine *Spline-Approximation* [Die95], eine *FFT-Filterung*, die *GAUSS-Faltung*, die *optimale (WIENER-) Filterung* [Wie64, Pre02] oder die *SAVITZKY-GOLAY-Glättung* [SG64, Pre02]. In wissenschaftlichen Publikationen findet sich dagegen eine unüberschaubar große Zahl von Filter- und Glättalgorithmen, häufig für Zeitreihen (d. h. asymmetrisch) oder für die digitale Bildbearbeitung optimiert.

Echte Glättung

Am strengsten erfüllt die weniger verbreitete WHITTAKER-HENDERSON-Methode [Whi23]⁵ die Anforderungen des Glättens. Hier wird das mathematische Optimum zwischen den Forderungen nach Glattheit (kleine mittlere quadratische n -te Ableitung) und Genauigkeit (kleine mittlere quadratische Abweichung vom Original) berechnet. Der Kontrollparameter ist das Verhältnis beider Größen. Der Parameter ist jedoch etwas unhandlich und eine Längenskala muss erst umständlich ermittelt werden (siehe dazu [Eil03]). Die x -Werte müssen außerdem in einem äquidistanten Raster vorliegen, was die Methode einschränkt.

Ähnlich gute Resultate liefert die hier verwendete lokale Regression mit glockenförmigem Glättkern. Die geglättete Kurve setzt sich hierbei aus punktuell ausgewerteten Regressionspolynomen zusammen, die durch lokale Gewichtung der jeweiligen näheren Umgebung gewonnen wurden. Als Kontrollparameter kann direkt die Breite der Glockenkurve als Längenskala der Glättung angegeben werden, was sich gut physikalisch motivieren lässt. Bei sehr schwacher

⁵Die WHITTAKER-HENDERSON-Methode [Whi23], in der Ökonomie auch unter HODRICK-PRESCOTT-Filter bekannt [HP97], geht nach [HP97] ebenfalls auf den italienischen Astronom SCHIAPARELLI zurück (1867).

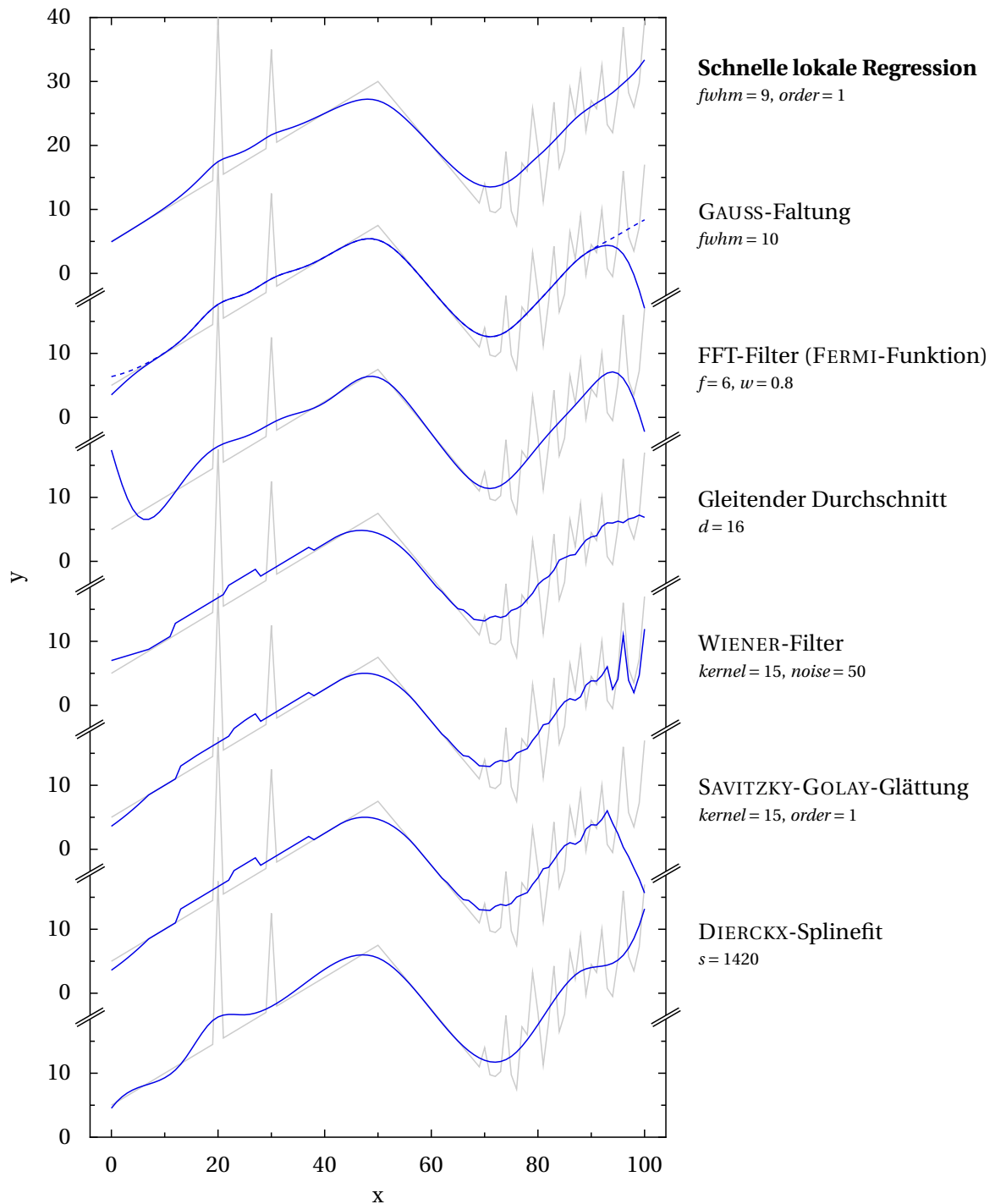


Abbildung 3.12.: Vergleich der lokalen Regression (mit iterativer Glockenfaltung) mit verschiedenen gängigen Algorithmen. Die meisten davon weisen Schwächen an den Rändern auf. Viele reagieren auf einzelne Ausreißer mit sich ungünstig überlappenden rechteckigen Unstetigkeiten. Neben den hier sichtbaren Mängeln stößt man bei der Anwendung einiger Filter auf weitere Schwierigkeiten, wie z. B. eine schlechte Kontrollierbarkeit des Verhaltens mit den zur Verfügung stehenden Parametern.

Glättung ähnelt das Ergebnis der Faltung mit der Glockenkurve, bei sehr starker Glättung geht das Ergebnis gegen das Regressionspolynom. Weitere Vorteile sind: Die x -Werte müssen nicht äquidistant sein und Interpolation ist möglich. Die lokale Regression wurde daher als Grundlage für den hier beschriebenen Algorithmus verwendet.

Vergleich

Abb. 3.12 zeigt einen qualitativen Vergleich der verschiedenen Glättmethoden mit dem hier vorgestellten Algorithmus. Die Testkurve wurde aus einer zweifach geknickten Geraden aus 101 Punkten konstruiert. Links wurden Peaks an zwei einzelnen Punkten hinzugefügt, rechts Rauschen bis zum linken Ende.⁶ Die in der Abbildung offensichtlichen und weitere Nachteile werden im Folgenden zusammengefasst:

GAUSS-Faltung: Ungewollte Krümmung der Kurve an den Rändern. Erweitert man die Gauß-Faltung derart, dass die Fläche der Gauß-Kurve genau im Überlappbereich immer auf eins normiert bleibt, verbessern sich die Rand-Effekte (gestrichelte Kurve), jedoch streben dann die Ränder immer in Richtung des Mittelwerts der Kurve und auch nicht gegen die Gerade.

FFT-Filter (FERMI): Die Abweichung am linken Rand hängt ungünstigerweise sogar vom Kurvenverlauf am rechten Rand ab. Dies liegt an der Periodizität der FOURIER-Approximation. Als Abschneidefunktion wurde die FERMI-Funktion gewählt.

Gleitender Durchschnitt: Unstetigkeiten bleiben immer auch in der geglätteten Kurve unstetig, was nicht den Erwartungen des Glättens entspricht.

WIENER-Filter: Die sogenannte optimale Filterung erzeugt in der besten Einstellung Artefakte ähnlich denen des gleitenden Durchschnitts.

SAVITZKY-GOLAY-Glättung: Zeigt ebenfalls Unstetigkeiten, selbst wenn eine höhere Polynomordnung in den Parametern gewählt wird. Außerdem ist der Kontrollparameter zur Einstellung der Stärke der Glättung nicht kontinuierlich.

DIERCKX-Splinefit: Splines erscheinen zwar glatt und zeigen keine Randartefakte, sie sind jedoch oft unvorhersehbar: Manche Details werden berücksichtigt, andere ignoriert (siehe die beiden Peaks im Vergleich). Außerdem ändern sich Splines sprunghaft bei kontinuierlicher Änderung des Kontrollparameters und bei kontinuierlicher Änderung von Funktionswerten der Kurve. Dies liegt daran, dass die Kurve durch eine endliche ganze Zahl von Splineknoten beschrieben wird. Bei besonders starker Glättung sind nur sehr wenige Knoten vorhanden und dadurch die Quantisierungsartefakte besonders auffällig und störend.

⁶Werte der Testkurve für Vergleiche: $x = 0 \dots 100$, $y = 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 11.5, 12, 12.5, 13, 13.5, 14, 14.5, 40, 15.5, 16, 16.5, 17, 17.5, 18, 18.5, 19, 19.5, 35, 20.5, 21, 21.5, 22, 22.5, 23, 23.5, 24, 24.5, 25, 25.5, 26, 26.5, 27, 27.5, 28, 28.5, 29, 29.5, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 14, 9.75, 9.5, 10.25, 19, 9.75, 7.5, 17.25, 16, 25.75, 19.5, 11.25, 18, 26.75, 16.5, 19.25, 29, 24.75, 31.5, 22.25, 27, 25.75, 32.5, 23.25, 22, 27.75, 38.5, 28.25, 26, 29.75, 39.5$.

3.2.2. Neue Glättmethode

Die neue Glättmethode beruht auf der lokalen Regression mit glockenförmigem Glättkern, die im Folgenden beschrieben wird. Anstelle einer Gaußkurve wird jedoch eine dazu ähnliche Glockenkurve verwendet, die eine schnellere Berechnung zulässt, so dass der Rechenaufwand des Algorithmus insgesamt mit $O(n)$, also linear mit der Anzahl der gegebenen Messpunkte n skaliert und insbesondere unabhängig ist von der Stärke der Glättung oder der Punktezahl $m = n$ der Faltungsfunktion.

In der Literatur fanden sich auch nach intensiver Recherche keine solchen gaußartigen Filter mit dieser seltenen „constant-time“-Eigenschaft, wie sie z. B. bei einem Median-Filter (Weichzeichner) gefunden wurde [PH07]. Faltungen, die anstelle von $O(n^2)$ mit $O(n)$ skalieren (schneller wie die FFT-Faltung mit $O(n \log n)$), sind dagegen in [Wer03] für eine größere Klasse von Funktionen beschrieben, jedoch ist keine Glockenkurve ersichtlich darin enthalten. In diesem Abschnitt wird die hier gefundene $O(n)$ iterative Glockenfaltung vorgestellt. Zunächst soll jedoch die bereits bekannte lokale Regression erläutert werden.

Funktionsweise des Algorithmus

Zur Bestimmung eines geglätteten Werts einer Messkurve werden die Koeffizienten zur Berechnung eines Regressionspolynoms um die nähere Umgebung glockenförmig gewichtet. Das durch diese lokale Gewichtung gewonnene Regressionspolynom wird an der jeweiligen Position ausgewertet. Das selbe Spiel wird für jeden zu berechnenden Wert durchgeführt und die Glockenkurve dazu jeweils mit ihrem Maximum zu diesem Wert geschoben. Aus diesen Werten bildet sich eine geglättete Kurve, die lokal einem Regressionspolynom entspricht, jedoch einem ähnlichen Einfluss der Umgebung unterliegt wie bei einer Faltung mit der Glockenkurve. Die Faltung und die Berechnung eines Regressionspolynoms werden in gewisser Weise zu einem einzigen Algorithmus vereint.

Das Regressionspolynom kann sehr schnell berechnet werden. Auf CARL FRIEDRICH GAUSS und ADRIEN-MARIE LEGENDRE geht die „*Methode der kleinsten Quadrate*“ zurück. Diese erlaubt es, für eine gemessene Funktion f mit den Messwerten $(x_1, y_1) \dots (x_n, y_n)$ ein Fit-Polynom k -ten Grades mit minimalem Fehlerquadrat, oder auch eine andere lineare Modellfunktion, zusammengesetzt aus beliebigen linear unabhängigen Basisfunktionen φ_i , direkt analytisch zu berechnen.

Die Koeffizienten c_i vor den Basisfunktionen bzw. die Polynomkoeffizienten erhält man direkt als Lösung des linearen Gleichungssystems

$$\begin{pmatrix} \langle \varphi_0, \varphi_0 \rangle & \langle \varphi_0, \varphi_1 \rangle & \cdots & \langle \varphi_0, \varphi_k \rangle \\ \langle \varphi_1, \varphi_0 \rangle & \langle \varphi_1, \varphi_1 \rangle & \cdots & \langle \varphi_1, \varphi_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \varphi_k, \varphi_0 \rangle & \langle \varphi_k, \varphi_1 \rangle & \cdots & \langle \varphi_k, \varphi_k \rangle \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_0 \rangle \\ \langle f, \varphi_1 \rangle \\ \vdots \\ \langle f, \varphi_k \rangle \end{pmatrix}, \quad (3.12)$$

das aus der Forderung nach der kleinsten Fehlerquadratsumme hergeleitet wird.⁷ Als Basisfunktionen φ_p wählt man die Basis der Polynome, also die x -Potenzen $\varphi_p = x^p$. Die Skalarprodukte definiert man bei der *lokalen* Regression abhängig von der zu berechnenden x -Position x_i mittels glockenförmiger Gewichtung durch

$$\langle f, \varphi_p \rangle^{(i)} := \sum_{j=1}^n f(x_j) \varphi_p(x_j) g(x_j - x_i) \quad (3.15)$$

mit einer Glockenfunktion $g(x)$, z. B.

$$g_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right). \quad (3.16)$$

Aus den $c_p^{(i)}$ erhält man nun durch Auswertung des Polynoms an der Stelle x_i den gesuchten geglätteten i -ten Wert

$$s_i = \sum_{p=0}^k c_p^{(i)} x_i^p. \quad (3.17)$$

Die beschriebene Prozedur wird für alle Positionen x_i durchgeführt, wobei sich das Maximum der Glockenfunktion jeweils an der selben Position x_i befindet. Die x_i in 3.15 und 3.17 müssen dabei nicht notwendigerweise mit den x_j der Messung übereinstimmen. Dadurch kann bei der Glättung auch eine Interpolation und Neuverteilung der Stützstellen erreicht werden.

Schnelle Berechnung der Gewichtung

Im beschriebenen Algorithmus müssen für jede Position x_i die Skalarprodukte berechnet werden, die eine Faltung mit einer Glockenkurve beinhalten. Die Zahl der Rechenoperationen für die Glättung wächst dann insgesamt mit der Ordnung $O(n^2)$ oder über die FFT-Faltung mit

⁷Die Herleitung des Gleichungssystem erfolgt ähnlich zur Gleichung 3.11, S. 40, mit der Messwertfunktion f , der Fitfunktion g , den Basisfunktionen φ_i , den Koeffizienten c_i und durch die Umformung der quadratischen Abweichung

$$\Phi^2(c_1, c_2, \dots, c_k) = \|g - f\|_2^2 = \langle g - f, g - f \rangle \quad (3.13a)$$

$$= \left\langle \sum_{i=1}^k c_i \varphi_i - f, \sum_{j=1}^k c_j \varphi_j - f \right\rangle \quad (3.13b)$$

$$= \langle f, f \rangle - 2 \sum_{i=1}^k c_i \langle f, \varphi_i \rangle + \sum_{i=1}^k \sum_{j=1}^k c_i c_j \langle \varphi_i, \varphi_j \rangle \quad (3.13c)$$

und durch die Extremalforderung

$$\frac{1}{2} \frac{\partial \Phi}{\partial c_i} = -\langle f, \varphi_i \rangle + \sum_{j=1}^k c_j \langle \varphi_i, \varphi_j \rangle = 0, \quad \text{für } i = 0, 1, \dots, k. \quad (3.14)$$

Ferner kann gezeigt werden, dass die zweite Ableitung (Hessematrix) positiv definit ist, so dass es sich beim gefundenen Extremum um ein Minimum handelt [SK09, Gro69].

$O(n \log n)$. Die Rechenzeit auf $O(n)$ zu reduzieren gelang, indem die GAUSS'sche Glockenkurve 3.16 durch die im Folgenden *iterative Glocke* genannte Kurve

$$\tilde{g}_\sigma(x) = \frac{1}{\sigma} \exp\left(-\frac{|x|}{\sigma}\right) - \frac{1}{\gamma\sigma} \exp\left(-\gamma\frac{|x|}{\sigma}\right) \quad \text{mit} \quad \gamma = \frac{21}{20} \quad (3.18)$$

ersetzt wurde. Diese Funktion ist zweimal stetig differenzierbar und, wie in Abb. 3.13 erkennbar, relativ glockenförmig. Die Ähnlichkeit zur GAUSS-Form kann noch gesteigert werden, wie später gezeigt wird.

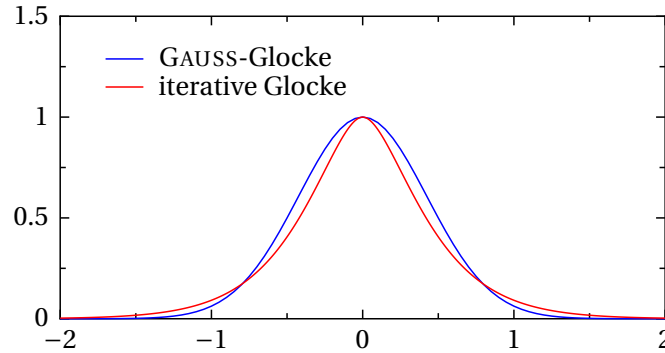


Abbildung 3.13.: Vergleich der glockenförmigen GAUSS'schen Normalverteilung (3.16) mit der iterativen Glocke (3.18). Beide Funktionen sind auf die Fläche 1 und das Maximum 1 normiert.

Der Wert γ sollte wegen Rundungsfehlern zwar nicht zu nahe bei eins liegen, damit aber die Form nicht zu spitzig wird, auch nahe genug bei eins. Der Effekt der iterativen Glocken-Faltung wird nun erkennbar, wenn die Funktion (3.18) durch

$$\tilde{g}_\sigma(x) = h_\sigma(x) - \frac{1}{\gamma^2} h_{\sigma/\gamma}(x) - h_\sigma(-x) + \frac{1}{\gamma^2} h_{\sigma/\gamma}(-x) \quad (3.19)$$

in vier halbe e^{-x} -Funktionen

$$h_\sigma(x) = \frac{1}{\sigma} \exp\left(-\frac{x}{\sigma}\right) \Theta(x) \quad \text{mit} \quad \Theta(x) = \begin{cases} 0: & x < 0 \\ 1: & x \geq 0 \end{cases} \quad (3.20)$$

zerlegt wird. Jede der in negativer Richtung abklingenden $h_\sigma(x)$ -Funktionen (mit $x \geq 0$) kann nun mit der aufsteigenden Iterationsformel

$$\tilde{y}_i = \tilde{y}_{j(i)} r_i + y(x_i) (1 - r_i) \quad \text{mit} \quad r_i = \exp\left(-\frac{x_i - \bar{x}_{j(i)}}{\sigma}\right) \quad (3.21)$$

in endlichen Schritten schnell mit $O(n)$ gefaltet werden. Analoges gilt für die in positiver Richtung abklingenden Funktionen mit einer absteigenden Iteration. Als Startwert der Iteration wird $\tilde{y}_1 = 0$ gesetzt. Die \bar{x}_j sind die Mittelpunkte der beiden benachbarten gemessenen x -Werte und markieren die sprunghaften Übergänge zwischen zwei diskret vorliegenden

y -Werten. $j(i)$ gibt den Index vom vorhergehenden Grenzwert an, bei dem gegebenen neuen (eventuell interpolierenden) Positionsindex i . Ohne Interpolation ist $j(i) = i$. y steht stellvertretend für alle zu faltenden Größen x^p und yx^p . Die \tilde{y}_i enthalten nach der Iteration die mit $h_\sigma(x)$ gefalteten Werte. Nach dem Zusammensetzen von vier Faltungen gemäß Gleichung 3.19 erhält man die Skalarprodukte, die in Gleichung 3.12 eingesetzt werden.

Herleitung der Iteration

Zu zeigen ist, dass die Faltung einer Messkurve $y(x)$ mit der Funktion $h_\sigma(x)$ durch die Iteration 3.21 gegeben ist. Die Faltung in kontinuierlicher Schreibweise

$$\tilde{y}(x) = \frac{1}{\sigma} \int_{x_1}^{x_n} y(\xi) \exp\left(-\frac{x-\xi}{\sigma}\right) \Theta(x-\xi) d\xi \quad (3.22)$$

ist wegen der HEAVISIDE-Funktion Θ identisch zu

$$\tilde{y}(x) = \frac{1}{\sigma} \int_{x_1}^x y(\xi) \exp\left(-\frac{x-\xi}{\sigma}\right) d\xi. \quad (3.23)$$

Das Integral dieser Faltung bis x_i kann an der letzten Übergangsstelle $\bar{x}_{j(i)}$ des y -Werts in

$$\tilde{y}(x_i) = \frac{1}{\sigma} \int_{x_1}^{\bar{x}_{j(i)}} y(\xi) \exp\left(-\frac{x_i-\xi}{\sigma}\right) d\xi + \frac{1}{\sigma} \int_{\bar{x}_{j(i)}}^{x_i} y(\xi) \exp\left(-\frac{x_i-\xi}{\sigma}\right) d\xi \quad (3.24)$$

aufgespaltet werden und lässt sich mit einem vorausgegangenen Iterationsschritt $\tilde{y}(\bar{x}_{j(i)})$ durch

$$\tilde{y}(x_i) = \tilde{y}(\bar{x}_{j(i)}) \exp\left(-\frac{x_i-\bar{x}_{j(i)}}{\sigma}\right) + \frac{1}{\sigma} \int_{\bar{x}_{j(i)}}^{x_i} y(\xi) \exp\left(-\frac{x_i-\xi}{\sigma}\right) d\xi \quad (3.25)$$

ausdrücken. Nimmt man nun in $y(\xi)$ zwischen $\bar{x}_{j(i)}$ und x_i den konstanten Wert $y(x_i)$ an⁸, kann das verbleibende Integral gelöst werden zu

$$\tilde{y}(x_i) = \tilde{y}(\bar{x}_{j(i)}) \exp\left(\frac{\bar{x}_{j(i)}-x_i}{\sigma}\right) + y(x_i) \left(1 - \exp\left(-\frac{x_i-\bar{x}_{j(i)}}{\sigma}\right)\right). \quad (3.26)$$

Vereinfacht ergibt sich daraus mit $\tilde{y}_i \equiv \tilde{y}(x_i)$ und $\tilde{y}_{j(i)} \equiv \tilde{y}(\bar{x}_{j(i)})$ die Iterationsformel

$$\tilde{y}_i = \tilde{y}_{j(i)} r_i + y(x_i) (1 - r_i) \quad \text{mit} \quad r_i = \exp\left(-\frac{x_i-\bar{x}_{j(i)}}{\sigma}\right). \quad (3.27)$$

⁸Damit werden die Wert an den diskreten Stützstellen über ein Intervall zwischen den Stützstellen verteilt. Bei einer schwachen Glättung niedriger Ordnung kommt die Verteilung als Treppenfunktion zum Vorschein. Alternativ können die Stützstellen als DIRAC'sche Deltafunktionen verteilt werden, wie im nächsten Abschnitt gezeigt wird.

Für alle Werte, die x -Potenzen enthalten, wird ebenfalls ein im Intervall $[\bar{x}_j, \bar{x}_{j+1}]$ konstanter Wert bereitgestellt. Damit selbst bei sehr schmaler Glockenkurve die Matrix von Gleichung 3.12 positiv definit bleibt und eine eindeutige Lösung existiert, müssen genügend Werte zum Skalarprodukt beitragen. Konkret muss jedes Skalarprodukt, das x -Potenzen enthält, aus einer Summe über mindestens $k - 1$ verschiedene Werte gebildet werden. Eine numerisch sichere Lösung ist, für jedes Intervall $[\bar{x}_j, \bar{x}_{j+1}]$ den integralen Mittelwert der x -Potenzen

$$\frac{1}{\bar{x}_j - \bar{x}_{j-1}} \int_{\bar{x}_{j-1}}^{\bar{x}_j} x^p dx = \frac{\bar{x}_j^{p+1} - \bar{x}_{j-1}^{p+1}}{(p+1)(\bar{x}_j - \bar{x}_{j-1})} \quad (3.28)$$

zu berechnen und diesen dann als konstanten Wert über das ganze Intervall zu verteilen.

Variante mit DIRAC'scher Deltafunktionen

Alternativ zur Gleichverteilung des Stützstellenwertes über das Intervall zwischen Messwert und Interpolationspunkt kann der Stützstellenwert auch punktförmig konzentriert bleiben. Anstelle einer Treppenfunktion wären die Werte der Stützstellen dann in Form von DIRAC'schen Deltafunktionen verteilt.

Dieser Ansatz vermeidet zwar Treppenstufen-Artefakte, führt jedoch bei zu schwacher Glättung dazu, dass nicht mehr sichergestellt ist, dass ein ausreichender Anteil der benachbarten Stützstellenwerte im Bereich zwischen den Stützstellen gemischt wird. Dadurch treten dort Null-durch-Null-Divisionen oder nicht mehr lösbare Gleichungssysteme auf.

Als Iterationsformel ergäbe sich dann

$$\tilde{y}_i = \tilde{y}_{i-1} r_i + \frac{1}{2\sigma} (y_{i-1} r_i + y_i) \quad \text{mit} \quad r_i = \exp\left(\frac{x_{i-1} - x_i}{\sigma}\right). \quad (3.29)$$

Dabei werden die Deltafunktionen durch die x_i -Werte in der Mitte halbiert. Als Startwert setzt man $\tilde{y}_1 = \frac{1}{2} y_1 / (1 + r_2)$.

Verbesserte GAUSS-Ähnlichkeit

Die Form der *iterativen Glocke*, wie in Abb. 3.13, ist gegenüber der GAUSSschen Normalverteilung etwas spitzer und unten breiter. Um näher an die physikalisch leichter motivierbare GAUSS-Form heranzukommen, kann die iterative Glockenfaltung mehrfach durchgeführt werden mit jeweils einer um einen geeigneten Faktor (bei doppelter Faltung ≈ 0.6) reduzierten Breite.

Abbildung 3.14 zeigt das Ergebnis einer doppelten Faltung im Vergleich zur GAUSSschen Glocke. Da die Faltung dem Gesetz der Assoziativität folgt, wird durch das mehrfache Falten effektiv mit der abgebildeten Funktion gefaltet.

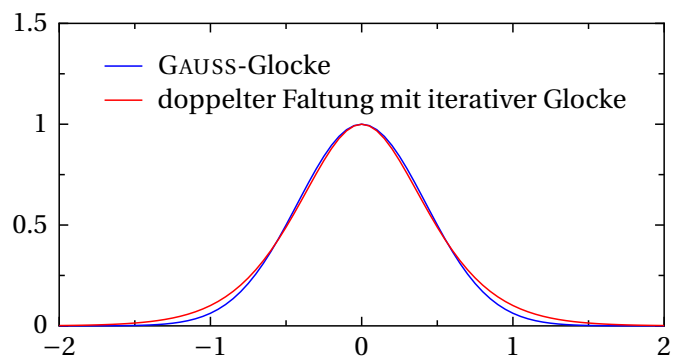


Abbildung 3.14.: Vergleich der glockenförmigen GAUSS'schen Normalverteilung (3.16) mit der Glockenfunktion der zweifach auf eine Deltafunktion angewandten iterativen Glockenfaltung. Beide Funktionen sind wie in Abb. 3.13 auf die Fläche 1 und das Maximum 1 normiert. Sie sind sich nun wesentlich ähnlicher.

»Jede hinreichend fortgeschrittene
Technologie ist von Magie nicht
mehr zu unterscheiden.«

— Arthur C. Clarke¹

4 KAPITEL 4. Experimentelle Techniken

4.1. Überblick

Bei den Experimenten zu dieser Arbeit wurden elektrische Messungen an mechanisch kontrollierbaren atomaren Bruchkontakten durchgeführt. Dabei sind schwache Supraleitungseffekte gemessen und Einflüsse durch hohe Ströme untersucht worden. Dieses Kapitel zeigt die verschiedenen dafür verwendeten und weiterentwickelten experimentellen Techniken.

Abbildung 4.1 zeigt den zentralen Gegenstand der Experimente, hier eine Goldprobe in einer elektronenmikroskopischen Aufnahme und die grafische Veranschaulichung eines atomaren Kontakts, dargestellt mittels Computergrafik².

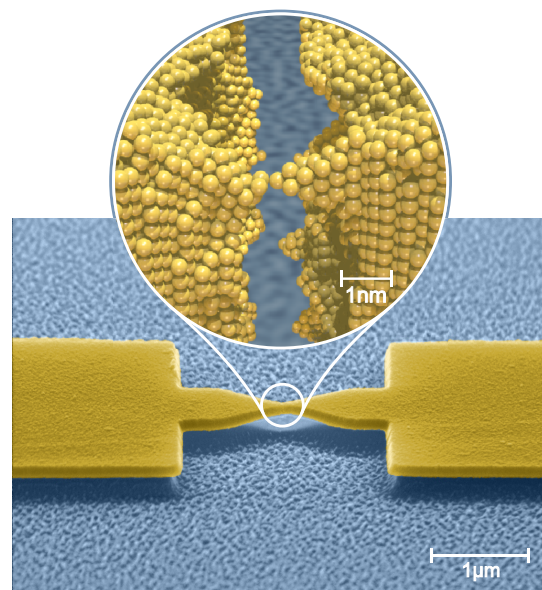


Abbildung 4.1.: REM-Bild und Modell

4.1.1. Bekannte und neue Techniken

Beschrieben wird in diesem Kapitel ein Bruchkontaktexperiment mit lithografisch erzeugten Proben, wie es erstmals in [RAP⁺96] verwendet und später in der Arbeit [SJE⁺97] erweitert wurde, um individuelle Leitwertkanäle bestimmen zu können. Für die hier vorliegende Arbeit wurde eine Erweiterung der Messelektronik entwickelt, die zusätzlich Elektromigrationsexperimente erlaubt.

Obwohl die meisten in diesem Kapitel beschriebenen experimentellen Techniken nicht prinzipiell neu sind und bereits veröffentlicht wurden, wird hier dennoch ein vollständiger

¹Britischer Science-Fiction-Schriftsteller, 1917–2008, Zitat: *drittes Clarke'sches Gesetzes*.

²Die Methode zur Erzeugung dieser „künstlerischen Darstellung“ wird im Anhang A.1 beschrieben.

Überblick über die wesentlichen Details wiedergegeben, die zum Nachbau des Experiments wichtig sind. Insgesamt stellt der Aufbau ein für die beschriebenen Experimente optimales System dar, d. h. einerseits wurden weder beim Kryostat noch bei der Messelektronik unnötig komplizierte und kostspielige Geräte verwendet, andererseits ist vom Einsatz minderwertiger Geräte deutlich abzuraten. Grenzen der Qualität der Messdaten lagen vermutlich nicht an den Geräten, sondern an Details der Abschirmung, Erdung oder Wärmeankopplung.

Die wichtigsten *neu* entwickelten Techniken sind in den folgenden drei Punkten zusammengefasst:

- Umschaltbare Verkabelung im Kryostat mittels Relais bei tiefen Temperaturen für zwei Messungen, die unterschiedliche Anforderungen an die Kabel stellen → Abschnitt 4.5.1.
- Einkerbung des Substrats zur Optimierung der mechanischen Eigenschaft der Probe beim Brechen des Bruchkontakts → Abschnitt 4.7.1.
- Python als alternative Messsoftware und guter Kompromiss zwischen interaktiven Programmen (z. B. Labview) und klassischen Sprachen (C oder Pascal) → Abschnitt 4.6.3.

4.1.2. Anforderungen

Eine erfolgreiche elektrische Messung an extrem kleinen Strukturen wie einzelnen Atomen am Bruchkontakt ist unweigerlich an das Erreichen ebenso extremer experimenteller Bedingungen gekoppelt. Die Erzeugung eines stabilen atomaren Bruchkontakts erfordert ein entsprechend kleines mechanische Übersetzungsverhältnis. Dafür ist wiederum eine äußerst schmale (100 nm breite) und flach aufliegende Sollbruchstelle notwendig. Um Fremdmoleküle (z. B. Sauerstoff, Wasser) aus der Atmosphäre von den zu messenden Atomen fernzuhalten ist ein sehr gutes, kryogenes Vakuum erforderlich, insbesondere bei leicht oxidierenden Proben aus Aluminium. Auch zur mechanischen Stabilisierung auf atomarer Skala und zur Erzeugung von Supraleitung sind die tiefen Temperaturen eines Kryostats zwingend notwendig. Und schließlich müssen extrem störungsarme und kleine elektrische Ströme gemessen werden. Das erfordert gute elektrische Abschirmungen und durchdachte Erdungen der Messleitungen. Um an statistischer Aussagekraft zu gewinnen, war es nötig, die Experimente automatisch wiederholen zu lassen. Die mehreren tausend Einzelexperimente wären anders nicht durchführbar gewesen. Die vollständige Computersteuerung erforderte einige kleine Erweiterungen des Aufbaus. Softwareseitig musste aufgrund der Komplexität der Vorgänge ein neues Konzept gegenüber Standardlösungen wie LabView oder Compilersprachen entwickelt werden.

4.1.3. Ablauf der Experimente

Zunächst werden Proben wie in Abb. 4.1 mit einer 100 nm engen brückenartigen Sollbruchstelle lithografisch hergestellt. Tabelle 4.3, S. 69 beschreibt den Herstellungsprozess. Als Substrat wird das bei tiefen Temperaturen elastische Material Bronze verwendet. Die Probe wird durch zwei mit Steckern versehenen Kupferdrähten kontaktiert und an die Messleitungen angeschlossen. Danach wird sie unter geringer mechanischer Spannung in die Bruchmechanik eingesetzt.

Anschließend wird die Vakuumschleuse des Kryostats geschlossen und die Kaltfahrprozedur durchgeführt. Nach einigen Vorversuchen zum Testen der Probe und der Verkabelung kann das eigentliche Experiment beginnen. Mittels eines Motors mit starkem Untersetzungsgetriebe und mit der im Kryostat befindlichen Bruchmechanik wird die Probe langsam gebogen. Währenddessen wird der Leitwert gemessen und beobachtet. Ist die Engstelle gebrochen und ein atomarer Kontakt hergestellt, können die Versuche mit steigender Stromstärke durchgeführt und die atomaren Umordnungen beobachtet werden. Befindet sich der Kryostat auf Basistemperatur (250 mK), kann die Verkabelung umgeschaltet werden und die Messung von multiplen Andreev-Reflexionen zur Kanalanalyse erfolgen. Nach einigen anfänglichen Versuchen wurde zur besseren Statistik das Experiment computergesteuert vielfach wiederholt. Bei reinen Elektromigrationsexperimenten genühten etwa zwei Wochen bei einer Probe. Die zeitaufwendigeren Experimente mit der Erzeugung bistabiler Zustände und gleichzeitigen Messungen der Kanalspektren liefen drei Monate.

4.2. Gesamtaufbau

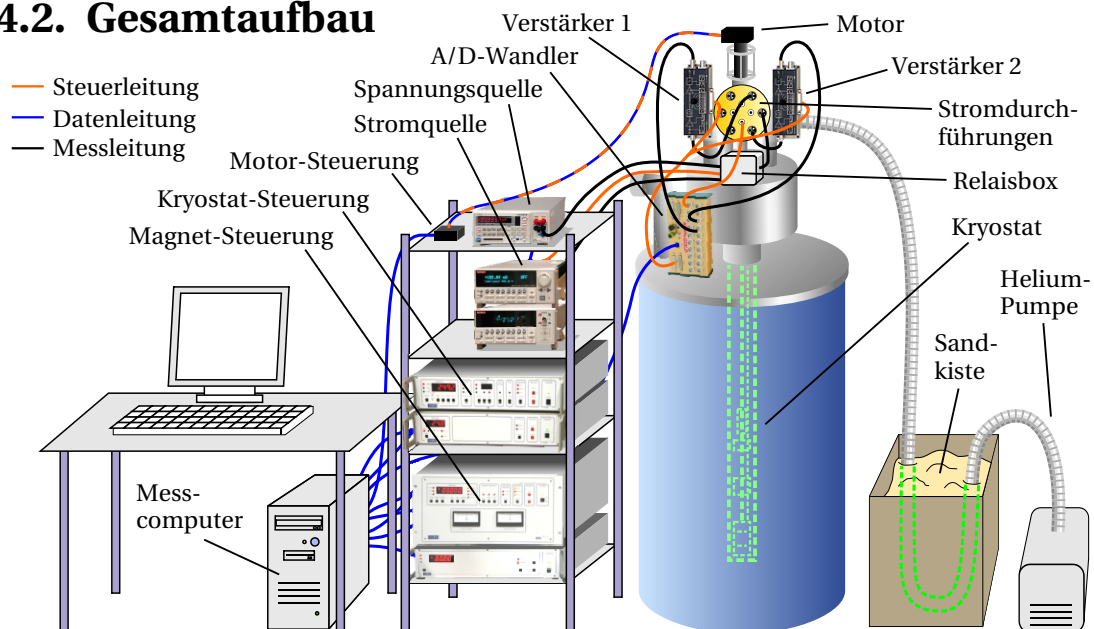


Abbildung 4.2.: Aufbau des Experiments im Überblick

Der experimentelle Aufbau besteht aus einem ^3He -Verdampfungskryostat mit mehreren koaxial geschirmten elektrischen Durchführungen und einer mechanischen Drehdurchführung. Mit einem computergesteuerten Motor kann die Achse zum Öffnen und Schließen des Bruchkontakts bewegt werden. Zur elektrischen Anregung der Probe wird eine Strom- und eine Spannungsquelle verwendet und zur Messung des Widerstands rauscharme hochohmige Vorverstärker sowie ein Datenerfassungssystem mit A/D- und D/A-Wandler. Im Dewar des Kryostats befindet sich eine supraleitende Spule für ein externes Magnetfeld zur Unterdrückung der Supraleitung in der Probe. Am Kryostat ist eine für den Betrieb notwendige externe Drehschieberpumpe angeschlossen. Die Vibrationen der Pumpe werden gedämpft, indem der

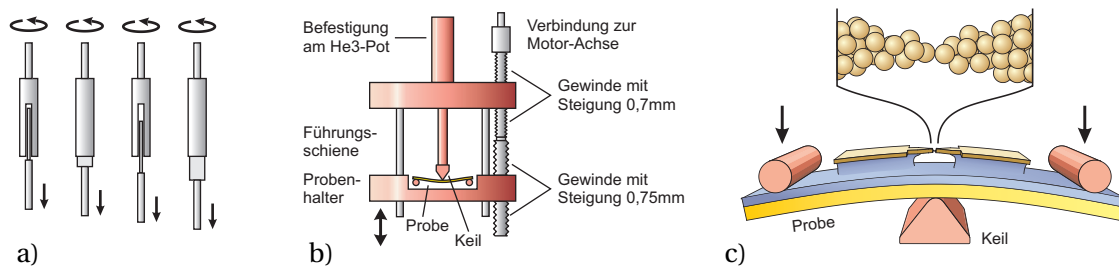


Abbildung 4.3.: Prinzip der Bruchmechanik. a) Die von der Drehdurchführung kommende Achse ist durch ein Zwischenstück in der Länge variabel. b) Schraubt sich die Achse durch zwei Gewindebuchsen mit unterschiedlicher Steigung, bewegen sich die Gewindebuchsen relativ zueinander um $50\ \mu\text{m}$ pro Umdrehung. c) Dadurch wird die Probe von zwei Gegenlagern auf einen Keil gedrückt und gebogen. Die Sollbruchstelle reißt dabei sehr langsam ein.

Wellschlauch zwischen Pumpe und Kryostat durch eine mit Sand gefüllte Kiste geführt wird. Der Dewar mit Kryostat hängt in einem Gestell, ist aber sonst nicht zusätzlich vom Boden mechanisch entkoppelt. Elektrisch ist er vollständig isoliert und nur über die Strom- oder Spannungsquelle mit einer speziellen Messerde verbunden.

4.3. Bruchmechanik

Bestandteil der Bruchmechanik ist eine mechanische Vakuum-Drehdurchführung. Die Drehbewegung wird über ein in der Länge variables Zwischenstück (Abb. 4.3 links) und über eine Schraube mit Differenzgewinde (Abb. 4.3 mitte) in eine vertikale Bewegung des Probenhalters übersetzt. Dieser drückt die von zwei Gegenlagern gehaltene Probe gegen einen Keil und biegt sie bis zu $2.5\ \text{mm}$ durch (Abb. 4.3 rechts). Das $300\ \mu\text{m}$ dicke Substrat wird dabei an der konkaven Seite gestaucht und auf der konvexen Oberfläche horizontal gestreckt, was zum langsamen Einreißen des $100\ \text{nm}$ dicken Metallfilms an der Sollbruchstelle führt. Das Untersetzungsverhältnis liegt dabei in der Größenordnung $1 : 10^4$. Von außen wird die Achse mit einem geregelten Präzisionsmotor gesteuert.

Der hier ungünstig große Abstand von der Schraube zur Probe, welcher durch die geometrischen Gegebenheiten des Kryostats vorgegeben ist, erzeugt eine starke Hebelkraft, die auf das Gewinde wirkt. Das führt zu einem hohen mechanischen Verschleiß und auf Dauer zum Klemmen der Achse. Je stärker die Achse klemmt, um so höher ist die Energie, die im $250\ \text{mK}$ -Bereich pro Umdrehung in Wärme umgesetzt wird. Dadurch steigt beim Drehen die Temperatur an und das flüssige ^3He wird schneller verbraucht. Aus diesem



Abbildung 4.4.: unteres Ende des Kryostats mit Bruchmechanik, Filterarray und eingesteckter Probe.

Grund musste ein Weg gefunden werden, wie die Probe mit weniger Kraft gebogen werden kann. Die Lösung ist im Kapitel 4.7.1 beschrieben.

4.4. Kryostat

Der verwendete Kryostat ist ein ^3He -Verdampfungskryostat vom Typ Heliox VL der Firma Oxford Instruments mit einer Basistemperatur von 250 mK. Diese Temperatur war für die Messung ausreichend. Verrundungen des Tunnelspektrums der supraleitenden Bandlücke waren deutlich höher, als durch Temperatureffekte zu erwarten. Begrenzend für die Auflösung waren hauptsächlich eingekoppelte Störsignale oder möglicherweise noch nicht ganz ausgereizte Temperatur- und Entstörfilter.

Die Wirkungsweise des Kryostats beruht auf der Verdampfungskühlung von gepumptem flüssigem ^3He . Das ^3He befindet sich dabei in einem geschlossenen System des Kryostats und kann abwechselnd verflüssigt und wieder verdampft werden. Der Kondensationsvorgang dauert etwa eine halbe Stunde. Anschließend kühlt das wieder verdampfende ^3He das System für etwa zwei Tage, wenn nicht durch die Reibung der Bruchmechanik oder auch durch Elektromigrationsexperimente mit hohen Strömen Wärme eingebracht und die Standzeit verringert wird.

Das Prinzip des Kryostats ist in Abb. 4.5 dargestellt. Der Wechsel von der ^3He -Verdampfung zur erneuten ^3He -Kondensation wird dabei lediglich durch eine elektrische Heizung in der Sorbtionspumpe vollzogen, ohne dass die Stellung des Nadelventils verändert werden muss. Dadurch vereinfacht sich, hinsichtlich einer vollautomatischen Computersteuerung des Experimentes, die Kontrollierbarkeit des Kryostats deutlich.

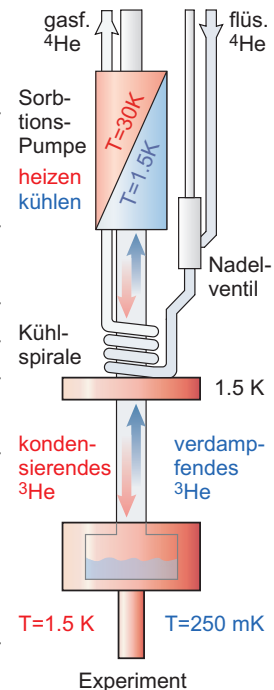


Abbildung 4.5.: Prinzip des ^3He -Verdampfungskryostats.

4.5. Elektrische Verkabelung des Kryostats

In diesem Abschnitt wird eine für diese Experimente spezialisierte Kabelführung im Kryostat vorgestellt, die eine Kanalanalyse bei Elektromigrationsexperimenten erlaubt. Die hier entwickelte Technik wurde in *Review of Scientific Instruments* veröffentlicht [SPS09].

4.5.1. Schaltbare Verkabelung

Die Verkabelung stellt in Tieftemperatur-Experimenten der Nanoelektronik oft den schwierigsten Teil des Aufbaus dar. Hier kann in der Regel auf keine kommerziellen Lösungen auf technisch neuem Stand zurückgegriffen werden, wie es z. B. beim Kryostat und bei den

elektrischen Messgeräten der Fall ist. Typische, sich zum Teil widersprechende Anforderungen sind hier:

- Minimale Störsignale und minimales Rauschen
- Effektive Filterung der Elektronentemperatur
- Geringe Wärmeübertragung entlang der Kabel
- Niedrige Kabelwiderstände für geringen Spannungsabfall
- Niedrige Kabelkapazitäten für schnelle Signalübertragung

Die Erfahrung auf diesem Gebiet der Forschung wächst jedoch. So konnte z. B. in den letzten Jahren die Widersprüchlichkeit zwischen einer effektiven thermischen Filterung und gleichzeitig geringen Kabelkapazitäten mithilfe der neu entwickelten Kupferpulverfilter zufriedenstellend gelöst werden (siehe Abschnitt 4.5.2). Einen weiteren Widerspruch galt es für die Experimente dieser Arbeit zu lösen: Experimente mit sehr rauscharmen Messungen im Mikrovolt-Regime schienen unvereinbar mit Experimenten, bei denen hohe Ströme bis 1 mA über vergleichsweise kleine Zuleitungswiderstände zur Probe geführt werden sollen. In Zahlen sind die hier geforderten Voraussetzungen eine Spannungsauflösung bei IV-Kurven von $< 10 \mu\text{V}$ bei einer Stromauflösung von 100 pA. Gleichzeitig sollen Ströme bis mindestens $200 \mu\text{A}$ möglich sein bei Leitungen mit Widerständen deutlich kleiner als der Probenwiderstand.³

Lösungsidee

Eine universelle Kabelführung konnte nicht gefunden werden. Als einziger Weg blieben umschaltbare Kabel, die zwischen zwei für die jeweilige Anforderung optimierten Kabelführungen hin- und herschalten. Abb. 4.6 a) zeigt beide Kabelführungen. Die linke Variante ist optimiert für hohe Ströme, die für Experimente zur Elektromigration erforderlich sind. Die Spannung fällt bei Proben mit $R > 900 \Omega$ zum größeren Teil über der Probe ab, so dass bei abrupten Leitwertänderungen kein Spannungsabfall der Leitungen sprungartig auf die Probe übergehen kann. Ebenso wenig heizen unnötige Verluste in den Leitungen den Kryostat zu sehr auf. Die rechte Variante zeigt im Fall der hochauflösenden Messungen, wie durch die Anregungsströme eingebrachte Störsignale durch einen Spannungsteiler in der Amplitude auf $1/37$ reduziert werden können. Der Spannungsteiler ist im Bereich tiefer Temperaturen platziert, um kein thermisches Rauschen zu erzeugen. Er verhindert allerdings, dass der Strom durch die Probe in verlässlicher Genauigkeit gemessen werden kann, so dass ein zusätzlicher Referenzwiderstand bekannter Größe zur Bestimmung des Stroms nötig ist. Er wird in Serie zur Probe angebracht

³Mit der Auflösung ist übrigens nicht die Rauscharmut der Messsignale gemeint. Sie wird vielmehr über die kleinstmöglichen Details definiert, die in nichtlinearen Strom-Spannungs-Kennlinien gemessen werden können. Zeitliches Rauschen des Messsignals kann im Gegensatz dazu bei genügend langsamem Messen oder bei vielen Datenpunkten im Nachhinein, bei der Auswertung im Computer, durch Mitteln minimiert werden. An der Probe anliegendes Rauschen, das zu verrundeten Strom-Spannungs-Kennlinien führt, zerstört dagegen unwiderruflich Information. Effektiv wird die zu messende Kennlinie mit der Rauschverteilungsfunktion gefaltet.

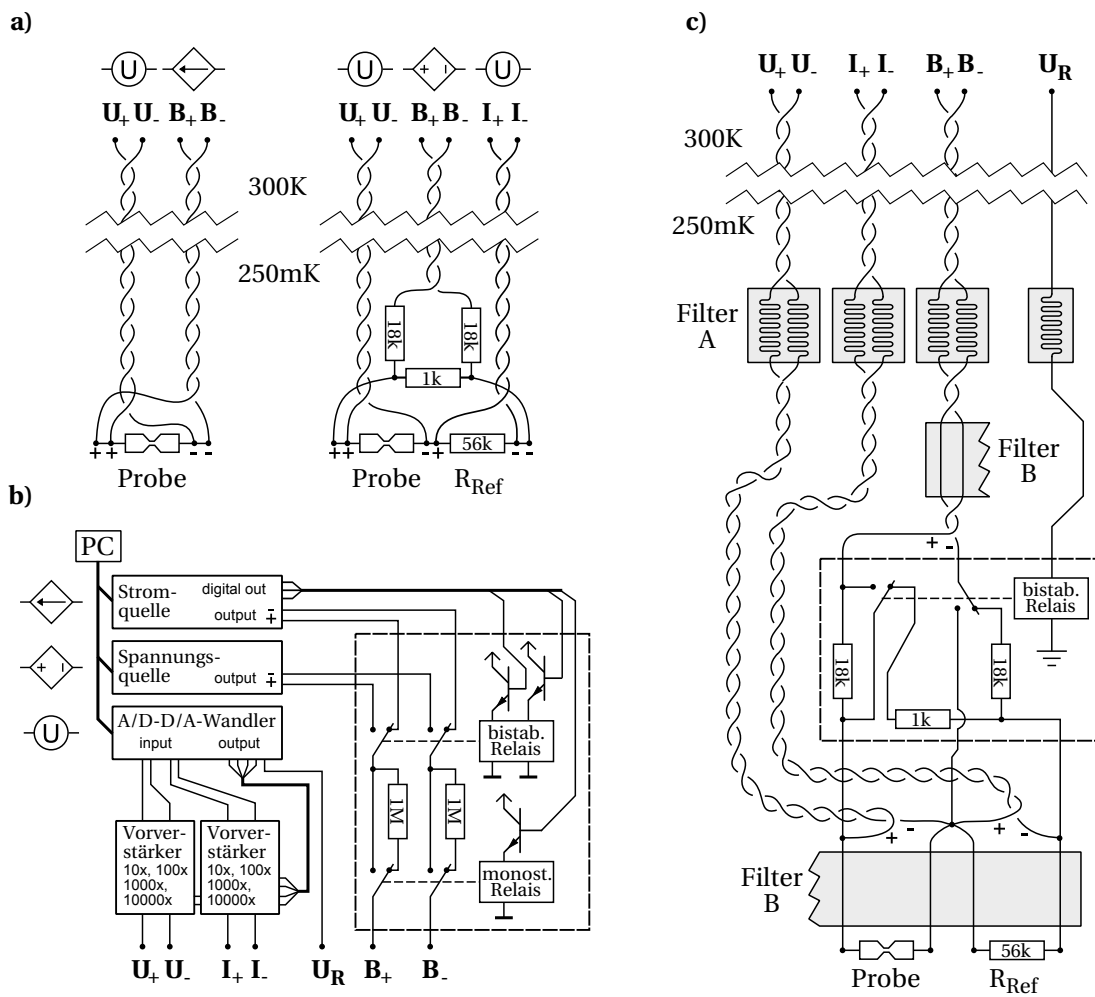


Abbildung 4.6.: Umschaltbare Verkabelung: a) Effektive Verkabelungen, zwischen denen hin- und hergeschaltet werden kann. b) Verkabelung außerhalb des Kryostats mit Relais zum Umschalten zwischen Strom- und Spannungsquelle. c) Verkabelung innerhalb des Kryostats mit Relais zum Aktivieren und Deaktivieren eines Spannungsteilers und eines seriellen Referenzwiderstands.

und verkleinert die Anregungsspannung und die eingebrachten Störsignale zusätzlich um den Faktor $R_{Probe}/(R_{Probe} + R_{Ref})$.

Fehlgeschlagene Varianten

Um die Vorteile beider Verkabelungen zu vereinen, wurden verschiedene Versuche durchgeführt. Zum Beispiel wurde die Probe direkt mit einer Stromquelle am Spannungsteiler vorbei angeregt, indem Leitungen außerhalb des Kryostats umfunktioniert worden sind. Pro-

blematisch war dabei jedoch der parasitäre Strom durch den zur Probe parallel liegenden $1\text{ k}\Omega$ -Widerstand des Spannungsteilers und durch den Referenzwiderstand. Dieser führte zur Verfälschung des Messwerts, zu einer zusätzlichen Masseschleife und einer damit verbundenen Störanfälligkeit. Außerdem rief die äußere Relais-Schaltbox zum Wechseln der Kabelfunktion und die darin nicht koaxial geschirmten Unterbrechungen ebenfalls Störsignalen hervor. Die Alternative, stattdessen weitere Kabel mit jeweils fest bleibender Funktion und je eigenem Messgerät zu ergänzen, würde weitere Störungen verursachen. Grund dafür sind unvermeidbare Masseschleifen durch die beiden ständig verbundenen Quellen (Strom- und Spannungsquelle). Beim völligen Verzicht auf den Spannungsteiler konnte die erforderliche Rauscharmut und Messauflösung nicht erreicht werden. Eine weitere Alternative wäre, die Widerstände außerhalb des Kryostats in flüssigem Stickstoff zu lagern. Dies wurde teilweise in der Arbeit von P. KONRAD umgesetzt [Kon03, KBS⁺05], wo mit einem externen Spannungsteiler eine ausreichende Messauflösung erreicht werden konnte. Darauf aufbauend ließe sich der externe Spannungsteiler über Schalter deaktivieren. Der interne Serienwiderstand von $100\text{ k}\Omega$ könnte dann durch Umfunktionieren der Kabel umgangen werden. Hierzu wurden jedoch keine weiteren Versuche unternommen, da die Reproduzierbarkeit dieser als Notlösung entworfenen Variante sehr unsicher erschien.

Beste Variante

Abbildung 4.6 b) und c) zeigen die in dieser Arbeit realisierte Lösung des Problems. Kernstück ist ein bei tiefen Temperaturen platziertes elektromechanisches Relais, das den Spannungsteiler und Referenzwiderstand zu- und abschalten kann. Damit wird effektiv zwischen den beiden in Abb. 4.6 a) gezeigten Varianten hin- und hergeschaltet. Die Messleitungen **U** behalten dabei die feste Funktion der Spannungsmessung über der Probe. Leitungen **I** werden nur bei Verwendung des Spannungsteilers benötigt. Leitungen **U** und **I** können also beide unterbrechungsfrei mit dem äußeren Messgerät verbunden bleiben. Einzig die Anregungsleitungen **B** müssen beim Umschalten verändert werden. Diese sind aber weniger problematisch, da eingebrachte Störungen bei der hochauflösenden Messung durch den Spannungsteiler minimiert werden. Die Messleitungen **U** und **I** kommen ohne derartige Schutzwiderstände aus. Sie sind mit sehr hochohmigen ($1\text{ G}\Omega$) Vorverstärkern verbunden. Durch sie fließt also kaum Strom, wodurch wenige äußere Störsignale eingebracht werden. Die mittels Antennenwirkung eingekoppelten Störungen werden durch Verdrillung der Leitungspaare (Twisted-Pair-Kabel) minimiert.

Abbildung 4.6 b) zeigt die Verkabelung außerhalb des Kryostats. Mithilfe der äußeren Umschaltbox kann die Stromquelle (Keithley 6221a) gegen die Spannungsquelle (Yokogawa) ausgetauscht werden. Während des Umschaltvorgangs werden $1\text{ M}\Omega$ -Widerstände zwischengeschaltet, die bei leichten Potenzialunterschieden die Probe vor kurzen Stromspitzen schützen sollen.

Beispielmessungen für beide Varianten

In Abbildung 4.7 wird zur Demonstration der Eigenschaften beider Schaltzustände je eine Beispielmessung gezeigt. a) zeigt die Kennlinie des kritischen Stroms mit hohen Strömen in der

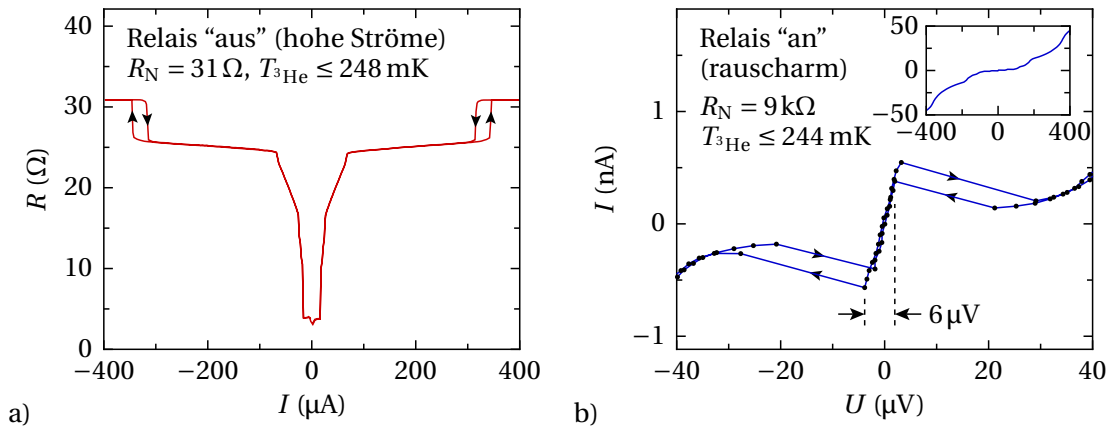


Abbildung 4.7.: Demonstration der Leistungsfähigkeit der Verkabelung in den beiden Relaisstellungen:
 a) Experiment zur Messung des kritischen Stroms, zum Teil bis 1 mA b) Hochauflösende Messung einer multiplen Andreev-Reflexion, Detail um Null und größerer Ausschnitt im Inset.

Variante ohne Spannungsteiler. Bei 0.4 mA tritt noch kein nennenswerter Temperaturanstieg am He^3 -Thermometer auf. Maximal sind etwa 50 mA möglich. In b) wird die höchste erreichte Auflösung bei aktiviertem Spannungsteiler gezeigt. Details der Kennlinie von etwa $6 \mu\text{V}$ sind noch deutlich erkennbar.

Details zum Relais

Bei dem im Kryostat verwendeten Relais vom Typ RAL-5 D W-K handelt es sich um ein bistabiles Zweifachrelais. Das bedeutet, dass die Magnetspule nur für kurze Zeit mit Strom versorgt werden muss (für $\geq 6 \text{ ms}$). Der geschaltete Zustand wird dann über einen Permanentmagneten dauerhaft gehalten. Zurückgeschaltet wird mit umgekehrter Polarität oder wahlweise über eine zweite umgekehrt gewickelte Spule. Um ein zusätzliches Kabel zu vermeiden, wurde erstere Option gewählt. Nur mit einem bistabilen Relais ist es möglich, bei tiefen Temperaturen, wie hier bei 250 mK, zu schalten, ohne den Kryostat dabei aufzuheizen. Ein Relais dieser Art hat bei 5 V eine Leistungsaufnahme von 150 mW. Bei einem monostabilen Relais wäre der Kryostat mit seiner Kühlleistung von $40 \mu\text{W}$ bei 300 mK vielfach überlastet. Die kurze Schaltzeit von 15 ms zeigte aber pro Schaltvorgang bei 250 mK einen sehr geringen Temperaturanstieg im ^3He von nur 2 mK für wenige Minuten. Das Kunststoff-Gehäuse des Relais wurde entfernt, so dass beim Abpumpen keine Luft im Gehäuse bleibt, die auf den Kontakten gefrieren und diese verschlechtern könnte. Stattdessen wurde das Relais in einem aus Kupferblech geformten Gehäuse untergebracht, das für elektrische Abschirmung und einen ausreichenden Wärmekontakt zum Kühlsystem sorgt, und über ein gebogenes Röhrchen evakuiert werden kann. In Abb. 4.6 c) ist das Kupfergehäuse mit einer gestrichelten Linie dargestellt. Die Widerstände des Spannungsteilers befinden sich im selben Gehäuse. Als Widerstände wurden SMD-Bauteile verwendet, die soweit möglich mit Kupferfolie umwickelt und thermisch angekoppelt wurden, damit die entstehende Wärme abfließen kann.

Eigenschaften der Kabelpaare U, I und B	Wert
Kabelmaterial	Manganin
Kabeldurchmesser	50 μm
Kabellänge	1.5 m
Abschirmungsmaterial	Edelstahl
Äußerer Durchmesser der Abschirmung	400 μm
Innerer Durchmesser der Abschirmung	200 μm
Gesamtwiderstand	490 Ω
Spezifischer Widerstand (Herstellerangabe)	224.9 Ω/m
Spezifische Kapazität (bei 5 kHz)	20 pF/m
Spezifische Induktivität	140 nH/m
Errechnete Dämpfung bei 5 GHz	-58 dB/m
Errechnete Abschneidefrequenz	14 MHz

Tabelle 4.1.: Eigenschaften der verwendeten Kabel.

Als Nachteil der Relais-Schaltung sei erwähnt, dass keine größeren Magnetfelder ($\gtrsim 50$ mT) beim Experiment verwendet werden können. Das Relais würde sonst ungewollt schalten, oder bei stärkeren Feldern mechanische Schäden verursachen.

Details zu den Kabeln

Die Kabelpaare **U, I** und **B** bestehen aus jeweils verdrehten Manganin-Drähten der Länge 1.5 m und des Durchmessers 50 μm in einer Edelstahlkapillare. Im Zwischenbereich wurde die Kapillare mit verdünntem Silikon ausgefüllt (Lösungsmittel Toluol). Das Silikon hält die Drähte mechanisch stabil in der Abschirmung und sorgt als Elektrolyt für eine gewisse Kapazität. Der für ein Metall recht hohe Widerstand des Manganins, die Kapazität und Induktivität bilden zusammen einen elektrischen Tiefpass mit einer Abschneidefrequenz von 14 MHz und einer Dämpfung von -58 dB bei 5 GHz (gerechnet für 1 m Kabellänge). 5 GHz entspricht der mittleren thermischen Energie der Anregungen im verwendeten Kryostat mit $T = 250$ mK umgerechnet in die Photonenfrequenz $\nu = k_B T/h$. Tabelle 4.1 fasst einige Daten zu den Kabeln zusammen.⁴

4.5.2. Kupferpulverfilter

Neben der schon erwähnten Filterwirkung der Koaxialkabel wurden weitere Filter zur Senkung der elektronischen Temperatur und zum Abschneiden hoher Frequenzen ab 100 MHz verwendet. Die Weitergabe der Wärme durch den Leiter erfolgt hauptsächlich durch elektrische Felder (der sich mit fast Lichtgeschwindigkeit ausbreitende „Schall“ des Elektronengases) und

⁴Herstellung und Vermessung der Kabel erfolgte durch H.-F. PERNAU

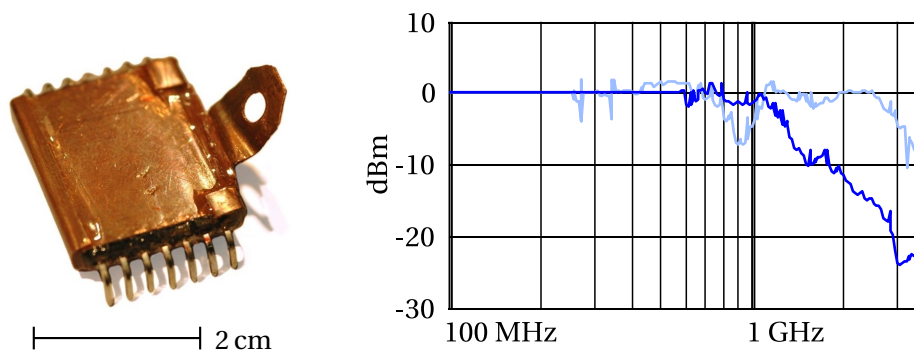


Abbildung 4.8.: Filter B. Links: Abbildung des Filterarrays im kompakten Kupfergehäuse mit Öse zur Befestigung und thermischen Ankopplung. Rechts: Filterkurven mit der transmittierten Leistung durch zwei der sieben Kanäle (Messung: Wissenschaftliche Werkstätten).

durch Diffusion. Die Elektronen bewegen sich selbst zwar auch sehr schnell ($v_F \approx 1600$ km/s in Cu), jedoch sind die mittleren Driftgeschwindigkeiten bei den verwendeten Stromstärken vernachlässigbar klein (ca. 10 nm/s bis 1 μ m/s)⁵. Das Durchlaufen des nur 5 cm langen Filterdrahtes benötigt für ein Elektron Tage bis Monate. Die Aufgabe des Filters ist, die Elektronen zu kühlen und die Frequenzen des thermischen Rauschens im elektrischen Signal zu dämpfen. Hier wurden ausschließlich die neu entwickelten Kupferpulverfilter [MDC87, BGH⁺03] anstelle herkömmlicher RC-Glieder eingesetzt. Die Vorteile sind schärfere Abschneidefrequenzen und stärkere Dämpfung bei hohen Frequenzen. Jede der von Raumtemperatur kommenden Leitungen durchläuft, wie in Abb. 4.6, S. 57 gezeigt, einen Kupferpulverfilter. Zwei Arten dieser Filter wurden eingebaut:

Filter A

Als Filter A werden Filter mit isolierten Drähten bezeichnet. Die Länge des Drahtes muss für eine ausreichende Filterwirkung im Bereich von Metern liegen (hier 2.4 m). Aufgrund des Skin-Effekts findet der elektrische Transport hoher Frequenzen nahe an der Oberfläche des Leiters statt. Als evaneszente Wellen überlappen diese durch die Isolierung hindurch mit dem Kupferpulver und geben dorthin ihre Leistung ab. Die Kupferpartikel sorgen aufgrund ihrer elektrisch isolierenden Oxid-Oberfläche für eine sehr selektive Weiterleitung und allmähliche Dämpfung hochfrequenter Signale, außerdem durch die gute Wärmeleitfähigkeit für ausreichende Kühlung der entstandenen Wärme. Der 2.4 m lange Draht muss in einem ca. 3 cm langen Filtergehäuse untergebracht werden. Dies geschieht mit einer speziellen Wicklung mit verschwindender Induktivität, damit magnetische Störfelder möglichst keine Ströme induzieren können. Die Dämpfung dieser Filter liegt bei einer Frequenz von 600 MHz bei -70 dB. Damit wird kein nennenswertes thermisches Rauschen, das Temperaturen von 250 mK und höher entspricht, zur Probe hin durchgelassen.

⁵Bei einer MAR-Messung am Ein-Atom-Kontakt ist z.B. $I = 0 \dots 100$ nA und $v_{\text{Drift}} = I / (ne\pi r^2)$, mit $r = 50$ μ m, $n = 8.45 \times 10^{22}$ cm⁻³ [Kit06] und $e = 1.6 \times 10^{-19}$ C.

Eigenschaften	Filter A	Filter B
Korngröße des Kupferpulvers	< 40 μm	< 40 μm
Länge des Kupferdrahtes	2.4 m	5 cm
Durchmesser des Kupferdrahtes	50 μm	100 μm
Isolierung	Kapton ca. 5 μm	keine
Abschneidefrequenz (–3 dB)	100 MHz	1 bis 3 GHz

Tabelle 4.2.: Eigenschaften der verwendeten Filter.

Filter B

Filter B bezeichnet einen Kupferpulverfilter mit einem nicht isolierten Draht, der direkten Kontakt zum Kupferpulver hat. Sehr feinkörniges Kupferpulver, wie es hier verwendet wurde, ist aufgrund der Oxid-Oberfläche bezüglich der Leitfähigkeit bei niedrigen Frequenzen und auf großen Distanzen nicht perkolierend. Der direkte Kontakt zum Pulver verursacht daher keinen Kurzschluss zur Masse, verbessert jedoch die Filterwirkung beträchtlich, so dass wesentlich kürzere Drähte im Filter zur gleichen Filterung führen. Filtergehäuse können so wesentlich kompakter gebaut werden, da keine Wicklung notwendig ist. In ca. 5 cm Abstand von der Probe konnte hier ein solch kompaktes (25 mm \times 21 mm \times 5 mm) siebenfaches Filterarray in der innersten Abschirmhülle gut untergebracht werden. Nachteile sind ab und zu auftretende leitende Pfade durch das Kupferpulver zur Gehäusemasse oder zu benachbarten Drähten. Dies muss vor jeder Messung geprüft werden. Liegt ein solcher Fall vor, kann das Kupferpulver durch leichte Stöße umgeordnet oder gelockert werden, bis keine Kurzschlüsse mehr gemessen werden. Abbildung 4.8 zeigt die Messung zweier Kanäle des Filters. Beide Kurven sind sehr unterschiedlich und hängen von der Anordnung der Pulverpartikel ab. Die Filterkurven wurden unter geänderten Bedingungen außerhalb des Kryostats bei Raumtemperatur gemessen. Sie zeigen jedoch, dass trotz der vereinfachten Bauweise ohne ein durchgehend koaxial abgeschirmtes Gehäuse eine Filterwirkung erkennbar ist. Der Filter dient als zweite Filterstufe. Ein zusätzlicher Vorteil ist die Kühlung der weiter zur Probe führenden nur 5 cm langen Drähte. Abgesehen von den Metallen im Stecker und in den Lötverbindungen wird dadurch die Probe über einen Kupferdraht direkt an den vom ^3He -Pot kommenden Kupferstab angekoppelt. Die 1.5 μm dicke Polyimidschicht zwischen der 100 nm dicken Aluminiumstruktur und dem Bronzesubstrat isoliert die Probe thermisch, daher spielt die Kühlung durch die Zuleitungen eine sehr wichtige Rolle. Tabelle 4.2 fasst die wichtigsten Daten der beiden Filtertypen zusammen.

4.5.3. Erdung und Abschirmung

Zur Verringerung der Einkopplung von Störsignalen ist eine gute Abschirmung und eine durchdachte Erdung notwendig. Letzteres bedeutet insbesondere, dass die Erdung sternförmig von der speziellen Messerde ausgeht, damit keine Masseschleife und kein Kontakt zur Netzerde gebildet wird. Variationen dieser Grundstrategie mussten jedoch durch Ausprobieren

optimiert werden. Als erfolgreich erwies sich, den kompletten Kryostat zunächst von jeder Masse abzukoppeln. Auch die Abschirmung jeglicher Messleitungen wurden nur einseitig verbunden. Der Anschluss zur Masse war dann über eine einzige Abschirmung der Anregungsleitung gegeben. Eine masselose Messung bezüglich der Messleitungen selbst erwies sich als ungünstig. So wurde der Minuspol der jeweiligen Quelle auf Masse gelegt. Damit lag die Probe selbst nicht mehr auf dem Massepotenzial der Abschirmung, was sich jedoch aufgrund des Zuleitungswiderstands bis zur Probe und des Spannungsteilers nicht vermeiden ließ.

4.6. Geräte und Steuerung

4.6.1. Messelektronik

Bei der Wahl der Messgeräte wurden die Aufgaben auf jeweils spezialisierte Geräte verteilt. Im Einzelnen sind das eine Gleichstromquelle, Gleichspannungsquelle, zwei Messsignal-Verstärker, ein D/A-Wandler und, nicht zu vergessen, das Gerät, welches auf die möglichst alleinige „Denkarbeit“ spezialisiert sein sollte, der Computer. So wurde für jede Aufgabe das meiner Meinung nach beste erhältliche Gerät eingesetzt und dabei eine hohe Flexibilität bewahrt. Bei der Arbeit mit Universalgeräten (SMUs (source and measure units) oder Digitaloszilloskope) ist man oft auf den mehr oder weniger durchdachten Funktionsumfang eingeschränkt und steht zum Teil vor unüberwindlichen Grenzen bei bestimmten, experimentell zwingenden Vorgaben.⁶

Anforderungen

Aufgrund der temperaturabhängigen Widerstände der Zuleitungen von den Messgeräten zur Probe musste eine Vierpunktmessung durchgeführt werden, was separate Quellen und Messgeräte erfordert. Die Messung der multiplen Andreev-Reflexionen benötigt eine extrem stabile Gleichspannungsquelle und einen guten Verstärker, um ein Signal von wenigen μV rauschfrei zu erfassen. Da sich die Kabelkapazitäten je nach Probenwiderstand als zeitliche Verzögerungen bemerkbar machen können und die Verstärker eine eingeschränkte Bandbreite besitzen, wurde auf Wechsellspannung verzichtet und entsprechend langsam, etwa im 0.5 bis 1 s-Takt gemessen. Für manche Messungen, z. B. zur Beurteilung des Rauschens, sollte dennoch die Möglichkeit bestehen, am selben Aufbau schnelle oszilloskopische Messungen durchführen zu können.

Wahl der Geräte

Als Quelle fiel die Wahl auf die sehr rauscharme Gleichstromquelle Keithley 6221 bzw. die ebenso rauscharme Gleichspannungsquelle Yokogawa 7651. Messseitig wurde das Signal kurz

⁶Besonders tückisch bei Quellen ist der vom Gerät gut gemeinte automatische Wechsel des Verstärkungsbereiches, der jedoch durch Spannungspulse den atomaren Kontakt zerstören oder verändern kann. Vorbildlich passiv und durchschaubar ist z. B. die Spannungsquelle Yokogawa 7651.

nach den Vakuumdurchführungen mit den sehr kompakt gebauten, rausch- und driftarmen FET-Spannungsverstärkern DLPVA-100-F-D der Marke Femto verstärkt. Das hinterher unempfindliche Signal zwischen -10 V und 10 V wird mittels Datenerfassungssystem ADwin digital erfasst. Als ein weiteres Gerät für vorbereitende Messungen wurde teilweise auch anstatt Verstärker und Datenerfassung das Nanovoltmeter Keithley 2182A verwendet. Alle Geräte wurden über GPIB bzw. über Ethernet vom Computer gesteuert und synchronisiert. Die Geräte im Einzelnen:

Stromquelle Keithley 6221: Primär eine Gleichstromquelle mit einigen Wechselstromfunktionen. Eine Spannungsbegrenzung ist in relativ kleinen Abstufungen möglich, was zum Schutz der Probe wichtig ist. Weitere verwendete Funktion: digitaler Ausgang, womit die äußere Relaisbox geschaltet werden konnte. Ein tückischer Nachteil soll noch erwähnt werden: Befindet sich die Erdung im Stromkreis zu weit vom Minuspol entfernt, sendet die Quelle extreme Störsignale aus. Davor wird im Handbuch nicht explizit gewarnt. Das schließt jedoch z. B. die Möglichkeit aus, die Erdung des Stromkreises direkt an die Probe zu legen, da dieser Punkt bereits durch die Zuleitungswiderstände zu weit vom Minuspol entfernt liegt.

Spannungsquelle Yokogawa 7651: Die Gleichspannungsquelle ist äußerst stabil und rauscharm, und nicht mit unnötig komplizierter Software ausgestattet. Sie bietet die Möglichkeit, mit einer zusätzlichen Feedbackleitung die Spannung weiter vom Ausgang entfernt zu überwachen und so die auf dem Weg entstandenen Störsignale zu kompensieren. Von letzterer Funktion wurde jedoch kein Gebrauch gemacht, da keine Verbesserung erreicht werden konnte.

Spannungsverstärker Femto DLPVA-100-F-D: Zwei dieser Verstärker wurden messseitig mit nur 20 cm kurzen Koaxialleitungen sehr nahe an der Stromdurchführung platziert. Das sehr kleine Gehäuse konnte mit Paketklebeband leicht am Kryostat befestigt werden. Der Verstärker ist ausreichend rauscharm und hat Einstellmöglichkeiten für die zwei Bandbreiten 1 kHz und 100 kHz sowie vier Spannungs-Verstärkungsbereiche 10^1 , 10^2 , 10^3 und 10^4 . Die Verstärkungsbereiche können über einen optisch entkoppelten digitalen Eingang vom Computer aus eingestellt werden.

Datenerfassungssystem ADwin: Dieses handliche Gerät besitzt zwei Arten von A/D-Wandlern: 16 bit bei 200 kHz oder 14 bit bei 2 MHz. Darüber hinaus bietet es 16 bit D/A-Ausgänge und einige digitale Ausgänge. Anstelle von umfangreichen Funktionen stellt das Gerät einen einfach gehaltenen Basic-Compiler zur Verfügung, um beliebige Funktionen mit einem internen 40 MHz-Prozessor abarbeiten zu lassen. So können z. B. Daten bereits als Mittelwerte an den Computer weitergeleitet werden, oder als Histogramm oder gepuffertes und getriggertes Oszilloskopbild. Für unsere Experimente wurde ein Mittelwert gebildet und bei Stromrampen zusätzlich eine bestimmte Einschwingzeit nach jeder Stromänderung abgewartet. Jeder Messpunkt ist ein numerischer Mittelwert über einen genau festgelegtes Zeitintervall mit etwa 100 000 Einzelmesspunkten. Die Belastung des eigentlichen Messcomputers und der Datenleitungen bleibt dabei sehr

gering. Die Ausgänge des Gerätes wurden zum Umschalten der Verstärker und zum Schalten des Relais im Kryostat verwendet.

Nanovoltmeter Keithley 2182A: Dieses Gerät wurde ursprünglich anstelle der Femto-Verstärker und des Datenerfassungssystem ADwin verwendet. Vorteil ist ein gepulster Modus für differenzielle Messungen, um einen Widerstand beim Kaltfahren des Kryostats driftfrei ohne thermischen Effekte auslesen zu können. Dabei wird ein Synchronisations- und Datenkabel zur Stromquelle Keithley 6221 verwendet. Der Widerstand wird dann am Display der Stromquelle angezeigt. Nachteil des Gerätes ist, dass keine schnellen Oszilloskop-Messungen möglich sind.

4.6.2. Computersteuerung

Es erwies sich als äußerst vorteilhaft, dass alle Geräte des Experiments vollständig für den Fernsteuerbetrieb per Computer ausgelegt waren. Die fünf elektronischen Messgeräte, die Geräte zur Steuerung des Kryostats und des Magnetsystems, sowie der Controller für den Motor der Bruchmechanik konnten alle über digitale Datenleitungen (*GPiB*, *Ethernet*, *RS232*) vom Computer aus gesteuert bzw. ausgelesen werden. Die Relais zum Umschalten der Kabel waren ohne Schwierigkeit über eines der anderen Geräte schaltbar. Nur so war es überhaupt möglich, die Experimente in der hohen Quantität durchzuführen. Durch die Computersteuerung konnten für die vielen Einzelvorgänge und Entscheidungen Abläufe und Regeln in Form von Python-Skripten einmalig formuliert und vielfach durchgeführt werden. Der folgende kurze Auszug aus der Logdatei

```
08-02-24 02:29:45 Strom steigt
08-02-24 02:30:19 Sprung von 0.904 nach 0.538
08-02-24 02:30:19 Strom sinkt
08-02-24 02:30:23 Sprung von 0.563 nach 0.904
08-02-24 02:30:24 Strom sinkt
08-02-24 02:31:32 Strom steigt
08-02-24 02:31:33 Feld ausschalten...
08-02-24 02:32:29 Feld ist auf 0mT
08-02-24 02:32:29 Andreev-Reflexionen
08-02-24 02:32:38 He3-Pot: warten auf T<280mK
08-02-24 02:42:41 Kondensieren...
08-02-24 03:05:07 Kryo: warten auf T<280mK
08-02-24 03:22:21 Kryo: weiter kühlen...
08-02-24 03:22:43 Andreev-Reflexionen messen
08-02-24 03:36:53 html-Datei (He-Level: 97.9%)
08-02-24 03:36:54 Feld anschalten...
08-02-24 03:37:54 Feld ist auf 18mT
08-02-24 03:38:03 Strom steigt
08-02-24 03:39:10 Sprung von 0.844 nach 1.499
08-02-24 03:39:11 Strom sinkt
```

zeigt 20 von insgesamt 150 000 Zeilen mit Vorgängen, die vom Messprogramm während der Experimente an Probe Nummer cs_b37 in einem Zeitraum von über 4 Monaten durchgeführt und mitgeloggt wurden. Hinter jedem mitgeloggtten Vorgang verbergen sich meist viele weitere Einzelschritte, wie das Anpassen von Messbereich, Verstärkungsfaktoren, Offset und Schrittweite der Rampe, das Umschalten von einigen Relais, das Kontrollieren des Zustandes der Probe und der Temperatur, das Speichern von Messdaten und einiges mehr.

All diese Vorgänge konnten nach und nach als Funktionen programmiert und später als Bausteine zu einer komplett automatisierten Ablaufsteuerung zusammengesetzt werden. So war

es möglich, die Experimente ununterbrochen 24 Stunden am Tag über Monate durchlaufen zu lassen. Während der Messung ließen sich parallel die Messergebnisse auswerten und bestimmte Parameter feinjustieren. Über das Internet wurden in einer html-Datei die aktuellen Messkurven und wichtige Informationen wie z. B. der Heliumfüllstand des Kryostats angezeigt, und konnten von jedem Ort aus ständig überwacht werden. Per VNC (*Virtual Network Computing* [VNC]) war es auch leicht möglich, das Experiment per Internet fernzusteuern, etwa vom Büroarbeitsplatz aus oder sogar von unterwegs während einer Physiktagung. Abgesehen von der komfortableren Bedienung wäre die manuelle Durchführung derartiger Messungen – mit vermutlich Millionen von einzelnen Handgriffen – ohne vollständige Computersteuerung in dieser Form praktisch unmöglich.

4.6.3. Skriptsprache Python

Als Messsoftware wurde die objektorientierte interaktive Skriptsprache Python [Ros, VRD09] anderen zur Verfügung stehenden Softwarelösungen, wie z. B. LabView oder der Compilersprache Delphi (Pascal), vorgezogen. In einem Editor, geeignet zum Ausführen einzelner Python-Programmzeilen, kann ein Experiment in Einzelschritten durchgeführt und schnell ausprobiert werden. Komplexere Vorgänge sind in Form von Funktionen programmierbar. Gespeichert als Textdatei können die wichtigen Funktionen eines Experimentes, ohne Nachteile gegenüber Buttons und Eingabefeldern wie in LabView, in beliebiger Reihenfolge ausgeführt werden. Auch für numerische Berechnungen, die Analyse und Auswertung der Messdaten bis zur Erzeugung fertiger Graphen wurde mit Python in einheitlicher Infrastruktur gearbeitet.

Als Philosophie hinter Python steht die Idee eines rein auf Text basierten Universalprogramms mit komfortabler Syntax ohne überflüssigen Ballast. Während der Programmtext und die Einzelkommandos selbst aus Text bestehen, verfolgt Python auch bei der Handhabung von Daten die Philosophie einer Bijektivität zwischen den Objekten im Speicher und einer menschenlesbaren Repräsentation als Klartext. So wird vom Texteditor aus die vollständige und einfache Kontrollierbarkeit des Systems erreicht. Für Python existieren zahlreiche Erweiterungspakete, auch zur Erzeugung von Fenstern (zum Plotten von Funktionen), oder zur Steuerung von GPIB-Schnittstellen und anderer Hardware. Insbesondere enthält Python eine Funktion um beliebige dll-Bibliotheken anzusprechen, was einen Zugriff auf die meisten mitgelieferten Treiber von wissenschaftlichen Geräten ermöglicht.

Anwendungsbeispiele für Python

Sobald in Python die wichtigsten Funktionen zur Steuerung der Geräte in Textform und als handliche Objekte verfasst sind, gelingt das Experimentieren sehr einfach und flexibel. Im Folgenden wird anhand kurzer Beispiele ein Eindruck vermittelt, wie mit Python in dieser Arbeit experimentiert wurde. Nach dem Start von Python wird mit der Zeile

```
import Messgeraete
```

ein vorher angelegtes Pythonskript importiert, in dem eine Sammlung von Messgeräten, Auslesewerten und die GPIB-Schnittstelle definiert wurden. Anschließend lassen sich sofort

Messungen durchführen. Zum Beispiel nach Eingabe von

```
ISoll(500)
UProbe()/URef()*RRef
```

wird die Stromquelle auf 500 nA gesetzt, zwei Spannungen gemessen und mit der Variablen RRef der Probenwiderstand berechnet. Das Ergebnis des Ausdrucks wird im Ausgabefenster angezeigt. Durch erneutes Betätigen der Eingabetaste in dieser Zeile kann die Messung wiederholt werden. Durch komplexere Befehle wie

```
kondensieren(T=30, t=25)
motorpos(U=2, UPM=500)
relais(quelle='U')
```

wird das Kondensieren von ^3He für 25 Minuten bei 30 K gestartet, die Achse der Bruchmechanik bis zur zweiten absoluten Umdrehung gedreht oder alle Relais für den Betrieb mit Spannungsteiler, Referenzwiderstand und Spannungsquelle umgeschaltet. Ein weiteres Beispiel zeigt, wie in einer Zeile ein neues Messgerät integriert werden kann. Sobald das Gerät, z. B. die Spannungsquelle Yokogawa, angeschlossen ist, gibt man ein:

```
yokogawa=GPIB(3)
yokogawa.write('S+15.0000E-3E')
```

Die erste Zeile legt ein neues Messgerät an der GPIB Schnittstelle mit Adresse 3 an und speichert es als Objekt in der Variable yokogawa. In der zweiten Zeile wird das Gerät Yokogawa bereits auf eine Ausgangsspannung von 15 mV gesetzt (das Übergabeformat muss im Handbuch nachgelesen werden). Diese Zeile kann auch leicht über eine Funktion auf `USOLLY(15)` verkürzt werden. Über solche Objekte und Funktionen, wie sie in der Datei `Messgeraete.py` gesammelt wurden, kann auf alle vorhandenen Messgeräte und Messwerte permanent (und vor allem auf einheitliche Weise) zugegriffen werden. In einem weiteren Beispiel kann in einer einzelnen Zeile

```
mm.rampe([0, -20, 20, 0], dy=1, dt=1, set=BFeld, get=[zeit, UProbe, GProbe, THE3],
         filename='magnetsweep1.txt')
```

eine Messung gestartet werden, wobei das Magnetfeld von 0 mT bis -20 mT, $+20$ mT und wieder bis 0 mT durchgefahren wird. Dabei werden die Zeit, die Probenspannung, der Leitwert und die Temperatur im Sekundentakt ausgelesen und alles in einer Datei `magnetsweep1.txt` gespeichert. Nach Beginn dieser Messung erzeugt die Zeile

```
mm.plot(UProbe, THE3)
```

ein Fenster in dem die Temperatur gegen die Probenspannung geplottet und permanent aktualisiert wird. Die letzten beiden gezeigten Funktionen sind Teil der selbstgeschriebenen Funktionssammlung `xyMess.py`.

All diese häufig benötigten Anwendungen verbergen sich hinter einzelnen Zeilen aus handlichem Text. Dieser ist über gewöhnliche Editierfunktionen wie copy & paste sehr flexibel handhabbar und die dahinter verborgene Funktion kann mit nur einer Taste (Enter) ausgeführt werden. Im Gegensatz zu LabView oder Compilersprachen kann ein Experiment somit mit einem minimalen Aufwand Zeile für Zeile ausprobiert und leicht variiert und verbessert werden.

Eine größere Ansammlung von Python-Befehlen lässt sich zu einer Funktion zusammensetzen, womit komplexere Abläufe wiederum über einzelne Befehlszeilen ausgeführt werden können. Eine der besonderen Stärken der Syntax von Python sind flexible Funktionsparameter mit optionalen Schlüsselwörtern, wie z. B. bei der Funktion `kondensieren(ramp=0.5)`, wo nur der Parameter `ramp` einen speziellen Wert bekommen soll und alle anderen Parameter auf voreingestellten Standardwerten bleiben. Diese Syntax ist z. B. in der Mathematiksoftware Maple üblich. Schleifen, `if`-Anweisungen und andere wichtige Hochsprachen-Elemente sind in Python ebenfalls enthalten. So stellt Python eine mächtige Infrastruktur bereit, die den hier umgesetzten Ansatz der Experimentsteuerung sowohl bei der erreichbaren Komplexität der Abläufe, als auch in der Entwicklungszeit anderen Programmen wie LabView überlegen macht. Voraussetzung zum Experimentieren ist allerdings eine entsprechende kleine Sammlung an grundlegenden Funktionen. Im Internet existiert bisher leider noch kein umfassendes Paket für Python als vollständige Messsoftware. Zwar werden zur wissenschaftlichen Datenauswertung zusammengestellte Komplettpakete angeboten⁷, jedoch müssen zur Gerätesteuerung weitere Pakete und einige kleine selbst geschriebenen Funktionen ergänzt werden.

4.7. Probenherstellung

In diesem Abschnitt werden zwei Verbesserungen der Proben bezüglich ihrer mechanischen Eigenschaften beschrieben. Zur Elektronenstrahlolithografie mit Lift-off-Technik soll dagegen nicht viel gesagt werden. Das Verfahren ist in der Literatur ausreichend bekannt und vielfach beschrieben worden⁸, so dass auf eine nähere Diskussion der Technik nicht eingegangen wird. Die Prozessschritte, Materialdaten und Parameter zur Herstellung der Proben sind kompakt in Tabelle 4.3 zusammengestellt. Die Seite kann so auch einzeln als handliche Rezeptur für das Labor ausgedruckt werden. Eine REM-Aufnahme einer fertig prozessierten Probe zeigt Abb. 4.11, S. 72.

4.7.1. Mechanische Optimierung des Substrats

Wie schon im Abschnitt 4.3 erwähnt, musste die Wahl des Substrats hinsichtlich der mechanischen Eigenschaften optimiert werden, um die Biegekraft zu minimieren. Daher wurden verschiedene Substrate durchprobiert:

Bronze (300 μm): Die Kraft beim Biegen beträgt hier 6 N bei 1 mm und 7 N ab 1.5 mm (Raumtemperatur). Das führt zum Teil bereits zum Klemmen der Mechanik.

Bronze (250 μm): Die Proben brachen erst bei höheren Krümmungen, und entsprechend stärker gekrümmt mussten sie eingespannt werden. Das stellt ein Risiko dar und übt zu Beginn bereits eine hohe Kraft auf die Mechanik aus.

⁷Zum Beispiel die *Enthought Python Distribution* oder *Python(x,y)*.

⁸Eine Beschreibung findet sich z. B. in meiner Diplomarbeit [Sch02]. Darin ist auch der speziell für freitragende Bruchkontakte notwendige Schritt mit einer Opferschicht aus Polyimid beschrieben.

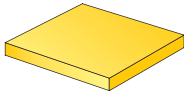
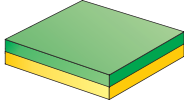
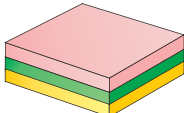
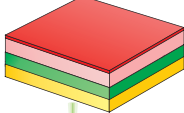
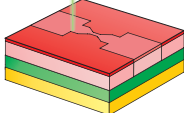
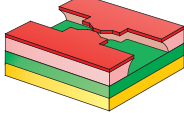
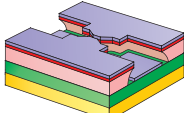
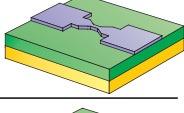
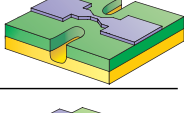
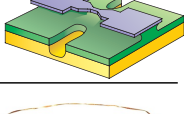
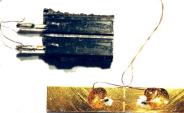
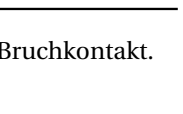
Prozessschritt	Materialien, Geräte, Parameter und Vorgehensweise	
Waferherstellung	Unmagnetisches ■ 300 μm Bronzeblech \rightarrow Metalldrehbank: Aufkleben und mit Drehstahl runde Scheiben (\varnothing 5.5 cm) ausschneiden	
Wafer polieren	1000er Schleifpapier, Bohrmaschine mit Polieraufsatz	
Wafer reinigen	mit Lösungsmittel Aceton und Isopropanol (IPA)	
Trocknen	Heizplatte: 1 min bei 100 $^{\circ}\text{C}$, abblasen mit Stickstoff	
Opferschicht	Wafer bestreichen mit zähflüssiger Polyamidsäure „Durimide 115 A“	
Schleudern	Spincoater: 30 s bei 300 U/min, Rampe in 3 s 90 s bei 5000 U/min, Rampe in 3 s	
Vorbacken	Labor-Ofen: 5 min bei 130 $^{\circ}\text{C}$	
Ausbacken	Vakuump-Heizplatte 30 min bei 430 $^{\circ}\text{C}$ (mit Rampe \nearrow und \searrow in jeweils 1:30 h), Imidisierung \rightarrow ■ 1.5 μm Polyimid-Schicht	
Trocknen	Heizplatte 1 min bei 100 $^{\circ}\text{C}$, abblasen mit Stickstoff	
Pufferschicht	Copolymer MMA-MAA gelöst zu 11% in Ethylacetat	
Schleudern	Spincoater: 4.5 s bei 400 U/min, Rampe in 3 s 90 s bei 2500 U/min, Rampe in 3 s	
Zwischenbacken	Heizplatte: 5 min bei 100 $^{\circ}\text{C}$ \rightarrow ■ 540 nm Copolymer-Puffer	
Elektronenlack	PMMA, Polymer mit 950 kg/mol, gelöst zu 4% in Anisol	
Schleudern	Spincoater: 4.5 s bei 400 U/min, Rampe in 3 s 60 s bei 5000 U/min, Rampe in 3 s	
Ausbacken	Labor-Ofen: 30 min bei 170 $^{\circ}\text{C}$ \rightarrow ■ 160 nm PMMA-Lackschicht	
Wafer zerteilen	spezielle Schneidevorrichtung mit fast parallelem Messer, Probengröße 3 mm \times 18 mm	
Belichten	Schreiben der Struktur mit Raster-Elektronen-Mikroskop 30 kV, Belichtungsdosis ca. 280 $\mu\text{A}/\text{cm}^2$ (variiert), große Flächen mit 80%, 100 nm-Engstelle mit 130% der Dosis.	
Entwickeln	In 1:3 verdünntem MIBK:IPA 20 s entwickeln, dann 2 min in IPA. Entwicklung stoppen durch Trockenblasen mit Stickstoff.	
Aufdampfen	PVD-Anlage mit thermischem Verdampfungstiegel: Aufdampfen von z. B. ■ 110 nm Aluminium (99.999% Reinheit) bei 5 $\text{\AA}/\text{s}$ während eines Kammerdrucks von ca. 10^{-8} bis 10^{-7} mbar.	
Liftoff	Unbelichtetes PMMA in Aceton auf 50 $^{\circ}\text{C}$ heißer Heizplatte auflösen. Mit Pinzette restliche Aluminiumschicht am Rand der Probe vorsichtig abstreifen.	
Einkerben	Optional, zur Reduktion der Biegekraft: Unter dem Binokular wird an beiden Seiten der Sollbruchstelle das Bronze substrat mit einem Skalpell bis auf die Hälfte der Dicke eingedrückt.	
Trockenätzen	Reaktive-Ionen-Ätzanlage: Plasma aus 1 mbar reinem O_2 , 10 min bei 50 W Mikrowellenleistung und 40 min bei 15 W \rightarrow entfernt 400 bis 500 nm der Opferschicht.	
Kontaktieren	Unter dem Binokular: Mit dünnem Lötendraht Klebetrophen aus UHU-Sofortfest aufbringen zur Fixierung und mechanischer Entlastung des Drahtes (90 μm Kupferdraht, lackisoliert). Lötendraht, hölzerner Zahnstocher: Leitsilberkleckse zur großflächigen Kontaktierung mit dem Metall der Kontaktflächen und zur Fixierung.	

Tabelle 4.3.: Alle wichtigen Schritte und Parameter zur Herstellung einer Probe mit Bruchkontakt.

Polyimidfolie (180 μm): (hier Apical® besser bekannt als Kapton®) kann ohne nennenswerte Kraft gebogen werden ($< 0.1\text{ N}$). Auch der zeitaufwendige Schritt der Polyimidbeschichtung schien zunächst überflüssig. Die Ätzparameter waren jedoch zu unterschiedlich, so dass die Schicht dennoch aufgebracht werden musste. Desweiteren benötigen nichtleitende Substrate eine zusätzliche 10 nm bis 15 nm dicke Aluminium-Schicht, die temporär aufgedampft und später wieder mit NaOH entfernt wird, um die Elektronenstrahlithografie überhaupt zu ermöglichen. Als größten Nachteil der Polyimidfolie stellte sich jedoch heraus, dass gar keine plastische Verformung auftritt wie beim Metall, wodurch auch kein Knick an der Stelle des Keils die Krümmung lokal verstärkt. Damit wurde die Untersetzung so stark erhöht, dass die Probe selbst am oberen Anschlag der Bruchmechanik noch nicht brach. Zu starkes Vorspannen kann die Probe sehr leicht zerstören. Alle Versuche endeten mit einem der beiden unerwünschten Fälle. Das Einschneiden der Folie links und rechts der Sollbruchstelle mit einer Schere (zur lokalen Erhöhung der Krümmung) machte die Proben viel zu empfindlich.

Polyimidfolie mit gekerbtem Draht: Bei der *notched wire*-Technik wird ein dünner Draht an einer Stelle eingekerbt und mit zwei Klebepunkten auf dem Substrat fixiert. Diese Technik findet häufige Verwendung, erfordert jedoch eine Bruchmechanik mit Piezo, da die Untersetzung wesentlich geringer ist als bei lithografischen Proben. Um diese Technik dennoch für die hier vorhandene Bruchmechanik nutzbar zu machen, wurde der mit einer Rasierklinge ringsherum eingekerbte Aluminiumdraht (25 μm) auf ein vollständig mit Klebstoff (UHU-Sofortfest) bestrichenes Polyimid-Substrat geklebt. Aufgrund der Oberflächenspannung des Klebstoffs bleibt genügend viel Luft in dem Schlitz, so dass eine freitragende Brücke bestehen bleibt. Das zu erwartende Untersetzungsverhältnis sollte nur geringfügig niedriger sein, als bei einer lithografischen Brücke, was insgesamt mit dem dünnen Substrat ein brauchbares Bruchverhalten erwarten lässt. Praktisch brach die Probe jedoch viel zu schnell. Zum einen löste sich der Draht in der Nähe der Kerbe vermutlich etwas vom Klebstoff. An der Kerbe bildete der Klebstoff außerdem kleine Risse, da der Klebstoff selbst an dieser Stelle geschwächt ist. Eine bisher noch nicht getestete Idee, welche die Probleme lösen könnte, ist, den Draht auch von oben mit Klebstoff zu überziehen. Dadurch sollte sich das Einreißen noch weiter verlangsamen, und das erwünschte Untersetzungsverhältnis erreicht werden. Mit der Dicke der Klebstoffschicht kann sogar die Untersetzung feinjustiert werden. Die Dicke kann beim Glattstreichen mit einer Rasierklinge durch unterschiedlich starken Druck variiert werden.

Kupfer (300 μm): hat einen wesentlich kleineren elastischen Bereich von etwa 0.5 mm (bei R.T.). Das begrenzt die Kraft auf 4 N und konzentriert die Krümmung früher auf den Knick. Der elastische Bereich sollte ausreichen, um den Bruchkontakt beliebig oft wieder zu schließen. Die Verformung geschieht jedoch bei tiefen Temperaturen vermutlich ruckartig in kleinen Stößen. Alle Proben brachen abrupt und konnten nicht wieder geschlossen werden.

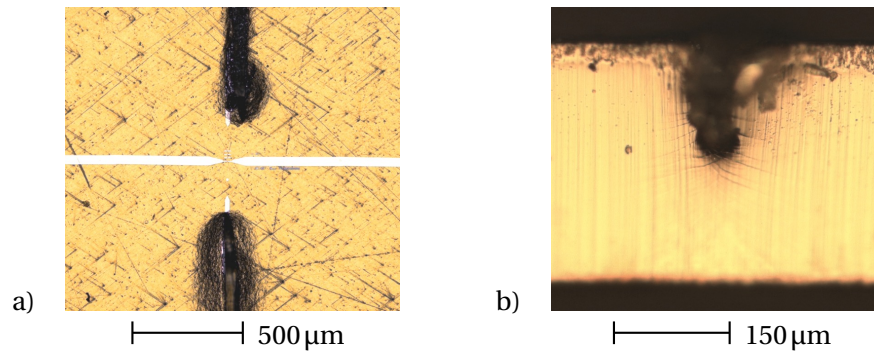


Abbildung 4.9.: Die Einkerbungen sollten unter dem Binokular wie in a) gezeigt möglichst nahe an den Bruchkontakt gesetzt werden und b) mit dem Skalpell bis in die Mitte des Substrats eingedrückt werden.

Bronze mit Einkerbung: Den Erfolg brachte schließlich wieder das ursprüngliche Material Bronze (300 µm), diesmal mit zwei Einkerbungen links und rechts neben der Sollbruchstelle. Diese ließen sich nach dem Aufdampfen des Probenmaterials und vor dem Unterätzen der Brücke gefahrlos mit einem Skalpell in das Bronzematerial eindrücken. Abb. 4.9 zeigt eine solche eingekerbte Probe von oben und der Seite. Das Einkerben hat zwei Vorteile. Erstens wird die Kraft reduziert auf 2.2 N bei 1 mm und 2.5 N bei 1.5 mm Biegeweg. Zweitens wird die Oberfläche um die Bruchstelle herum stärker gestreckt, da die Grenzfläche zwischen Stauch- und Streckbereich in das untere Viertel der Substratdicke verlagert wird. Diese Herstellungsmethode wurde von nun an ausschließlich verwendet.

4.7.2. Optimierung des Probendesigns

Als weiterer Weg, das Brechen früher eintreten zu lassen und damit die Biegekraft klein zu halten, wurde der freitragende Bereich der Brücke vergrößert, da dieser linear in das Unteretzungsverhältnis eingeht. Um überhaupt einen freitragenden Bereich der Sollbruchstelle zu erhalten, musste die Umgebung jedoch genügend schmal sein, damit sie unterätzt werden kann. Es stellte sich heraus, dass diese Strategie den gegenteiligen Effekt hatte und die dabei höhere Instabilität und Elastizität das Brechverhalten insgesamt verschlechterte. So wurde wieder mit einem kleineren freitragenden Bereich gearbeitet, dafür wurden massivere Zuleitungen dreiecksförmig bis zur 100 nm-Engstelle geführt, wie in Abb. 4.10 zu sehen ist.

4. EXPERIMENTELLE TECHNIKEN

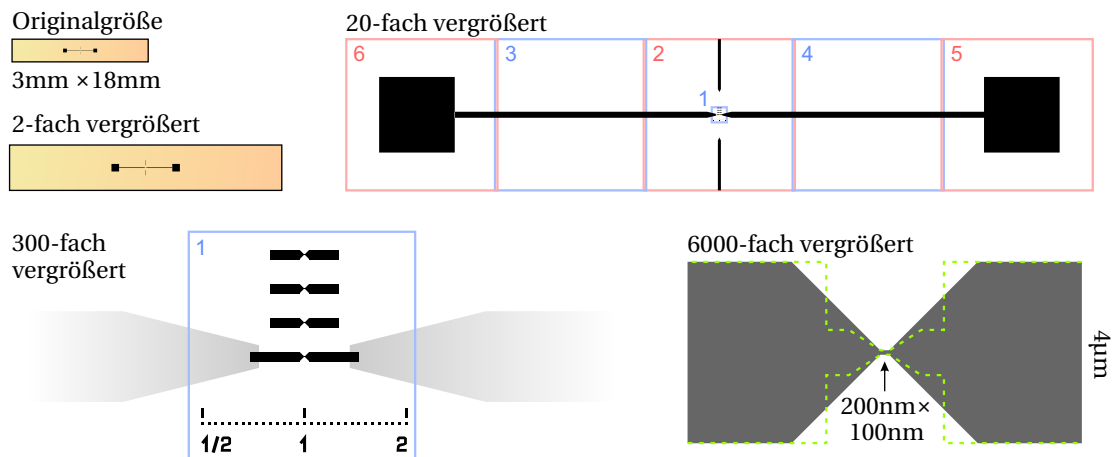


Abbildung 4.10.: Probedesign für die Belichtung mit Elektronenstrahlen. Markierung der Schreibfelder einzelner Belichtungsschritte. Links unten: Inneres Schreibfeld mit Belichtungsskala. Rechts unten: Sollbruchstelle in neuer (dunkelgrau) und alter Form (grün gestrichelt).

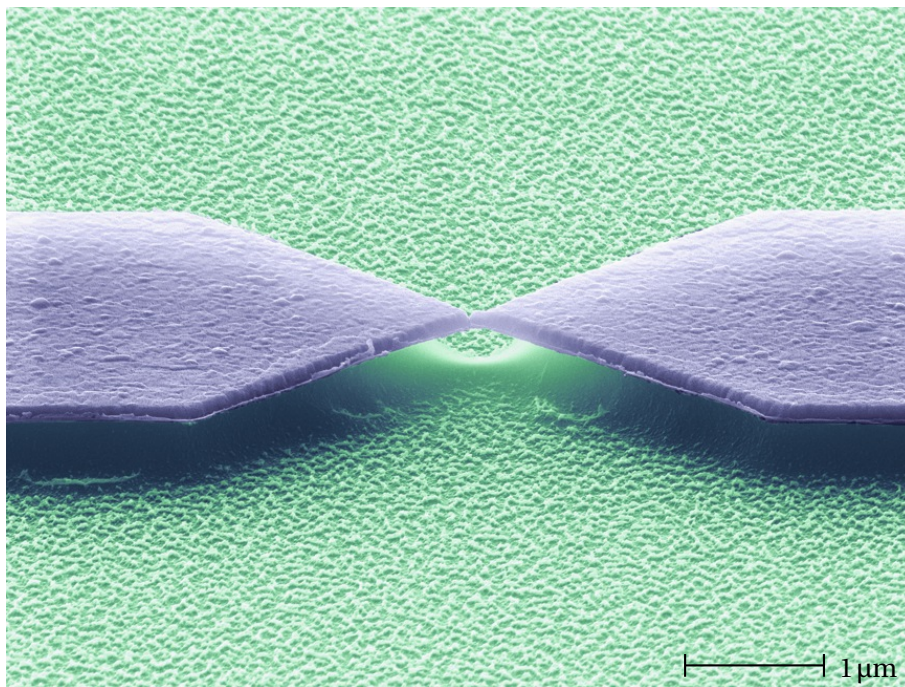


Abbildung 4.11.: Rasterelektronenmikroskopische Aufnahme einer fertiggestellten Nanobrücke aus Aluminium (in Falschfarben). Der rauhe Untergrund (grün eingefärbt) zeigt die teilweise weggeätzte Opferschicht aus Polyimid.

»Durch bloßes logisches Denken vermögen wir keinerlei Wissen über die Erfahrungswelt zu erlangen; alles Wissen über die Wirklichkeit geht von der Erfahrung aus und mündet in ihr.«

— Albert Einstein¹

5 KAPITEL 5. Messergebnisse

Dieses Kapitel zeigt Messergebnisse zur Untersuchung von Gold- und Aluminium-Bruchkontakten unter dem Einfluss hoher Ströme. Analysiert wurden die Kontakte mittels Strom-Spannungs-Kennlinien und durch die Bestimmung der einzelnen quantenmechanischen Leitwertkanäle.

Ziel der Untersuchungen war, verschiedene Zusammenhänge zwischen Eigenschaften des Leitwerts und der atomaren Geometrie zu ermitteln. Dabei stand die lokale Manipulation der Geometrie mit stromgetriebenen Umordnungen im Vordergrund.

5.1. Atomare Stabilität bei hohen Strömen

5.1.1. Messung sprunghafter atomarer Umordnungen

Wird mittels Bruchmechanik ein atomarer Kontakt hergestellt und der Biegevorgang gestoppt, bleibt die atomare Struktur bei tiefen Temperaturen ohne äußere Einflüsse über mehrere Tage stabil. Bei steigendem Strom, der über den Kontakt fließt, ändert sich dagegen der Leitwert ab etwa 50 mV sprunghaft und irreversibel. Dies kann nur als geometrische Änderung der atomaren Struktur gedeutet werden. Abb. 5.1 zeigt ein Beispiel einer solchen Leitwertänderung bei steigendem Strom. Mehrere Umordnungen führen hier zu kleinen Änderungen des Leitwerts. Wird der Strom konstant gehalten (was die Abbildung nicht zeigt), treten keine Umordnungen auf und der Leitwert bleibt konstant. Bei etwa 26 μA wird im Beispiel der Strom langsam wieder auf Null heruntergefahren, was ebenfalls zu keinen weiteren Sprüngen führt. Lediglich eine leichte kontinuierliche Änderung ist sichtbar, die als Leitwertfluktuation gedeutet werden kann. Eine Erhöhung des Stroms in negativer Richtung führt zunächst ebenfalls nur zu kontinuierlichen Änderungen. Bei $-24 \mu\text{A}$ tritt in diesem Beispiel eine weitere kleine Umordnung auf. In dieser nun erreichten atomaren Geometrie bleibt der Kontakt gegenüber allen Strömen zwischen $-26 \mu\text{A}$ und $26 \mu\text{A}$ stabil, was ein weiterer kompletter Durchlauf der Stromrampe bestätigt (blaue Kurve).

¹1879–1955, Zitat aus *Mein Weltbild* [ES05]

Die Messung kann durch eine atomare Umordnung gedeutet werden, die zu einer stabileren Geometrie führt. Eine gegenüber positiven Strömen stabile Geometrie ist dabei meistens auch gegenüber negativen Strömen stabil. Die Ausnahme dieser Regel, d. h. eine Asymmetrie der Stabilität, wurde in einer späteren Messung speziell untersucht und zum reversiblen atomaren Schalten ausgenutzt.

Auch sinkende Ströme können Umordnungen auslösen, jedoch nur in sehr seltenen Fällen. Diese könnten als eine eigentlich instabile Geometrie gedeutet werden, die von einer Gegenkraft durch den permanenten elektrischen Stroms oder die abfallende Spannung stabilisiert wird.

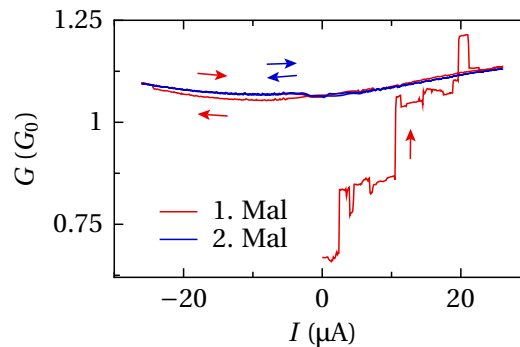


Abbildung 5.1.: Sprunghafte irreversible Änderung des Leitwerts bei steigendem Strom. Im zweiten kompletten Durchlauf treten keine Sprünge mehr auf.

5.1.2. Statistik über viele Messungen

Eine Einzelmessung wie in Abb. 5.1 mit einer bestimmten Anfangsgeometrie ist nicht reproduzierbar. Einige Eigenschaften konnten dennoch herausgelesen werden, andere verbergen sich hinter der Zufälligkeit der Formierung des Kontakts und kommen nur zum Vorschein, wenn sehr viele Einzelmessungen durchgeführt und statistisch analysiert werden. Um dieser Strategie nachzugehen, wurde in den folgenden Messungen jeweils der Strom so lange erhöht, bis der Kontakt ganz unterbrochen war. Anschließend wurde der Kontakt geschlossen und ein neuer zufälliger atomarer Kontakt eingestellt. Der Startleitwert ist nicht völlig zufällig, sondern kann beeinflusst werden, indem der Öffnungsvorgang bei einem bestimmten Leitwert beendet wird. Für eine aussagekräftige Statistik muss die Verteilung dieser Anfangswerte auf eine definierte und reproduzierbare Form gebracht werden.

Abbildungen 5.2 b) und 5.3 b) zeigen die Verteilungen, die beim gleichmäßigen mechanischen Öffnen des Kontakts auftreten. Es ist jedoch unklar, welche physikalischen Eigenschaften des Kontakts zu diesen Verteilungen führen, wie dem Abschnitt 2.3, S. 20 zu entnehmen ist. Einerseits können bevorzugte oder besonders stabile Geometrien zu dieser ungleichmäßigen Verteilung des Leitwerts führen, andererseits können die unterschiedlichen Geometrien dazu neigen, bestimmte gemeinsame Leitwerte häufiger anzunehmen. Im letzten Fall wäre eine physikalische Ursache zu suchen, die unabhängig von der konkreten atomaren Geometrie

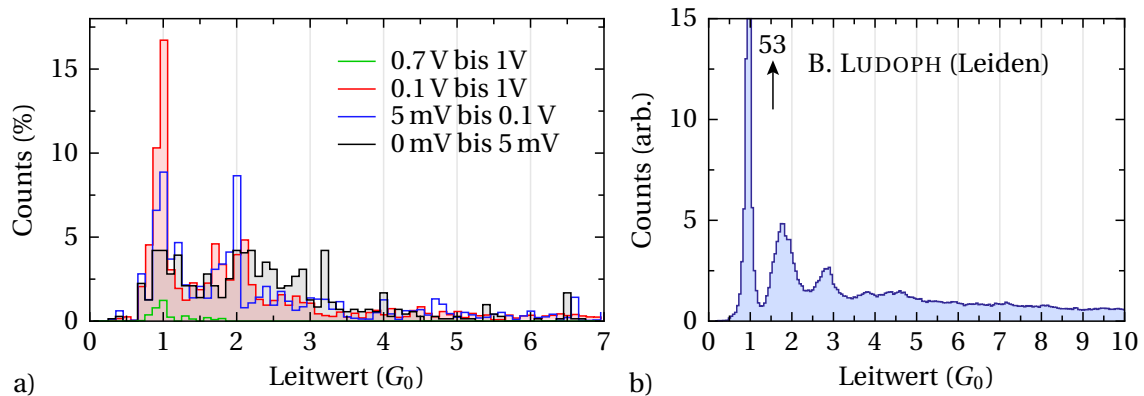


Abbildung 5.2.: Histogramme bei Gold. a) 108 stromgetriebene Öffnungsvorgänge bei nicht ganz gleichmäßiger Anfangsverteilung wie in Abb. 5.3, jedoch mit gleicher Aussage. b) 3500 rein mechanische Öffnungsvorgänge [Lud].

wirkt, wie z. B. Modeneffekte, die allein durch den Öffnungswinkel der Kontaktspitze erklärbar sind.

In diesem Experiment wurde versucht, die Anfangsverteilung der Leitwerte ausgeglichen zu präparieren, bevor die Leitwerte durch stromgetriebene Umordnung zu einer veränderten Verteilung überführt werden. Um diese Anfangsverteilung zu präparieren, wurden viele der Messungen mit entsprechend häufigen Leitwerten selektiert. Das geschah teilweise während des Experiments, noch bevor das Umordnungsexperiment gestartet wurde (im manchen Fällen auch durch gezieltes Hin- und Herdrehen der Mechanik), und zum Teil später bei der Auswertung. Die Histogramme so gewonnener Anfangsverteilungen und die Histogramme nach der Behandlung mit hohen Spannungen sind jeweils in den Abbildungen 5.2 a) für Gold und 5.3 a) für Aluminium zu sehen. Bei den Goldproben (5.2 a)) konnte mit 161 Messkurven und einer anschließenden Selektion eine nur grobe Gleichverteilung im Bereich 1 bis 3 G_0 erreicht werden. Mit der höheren Zahl von 1354 Messungen an einer Aluminium-Probe gelang die Verteilung über einen Bereich von 0.5 bis 5.5 G_0 nahezu gleichmäßig (5.3 a)).

Ausgehend von allen Kontakten des schwarzen Histogramms wurde jeweils ein Umordnungsexperiment gestartet. Nach langsamem Steigern des Stroms ändert sich die atomare Geometrie sprunghaft und neue Leitwerte kommen zustande. Sortiert man die gemessenen Leitwerte in unterschiedliche Histogramme, je nach anliegender Spannung, ergeben sich jeweils neue Histogramme, die ebenfalls in den Abbildungen zu sehen sind.

Steueralgorithmus und Auswertung des Experiments

Um eine gleichmäßige Anfangsverteilung zu erhalten, muss jedes Leitwertintervall eines Histogrammbalkens mit gleicher Häufigkeit als Startwert vorhanden sein, um von dort ausgehend das Umordnungsexperiment zu starten. Gezielt die gewünschten Intervalle zu treffen, gelingt nur in Einzelfällen, da die kleinen Intervalle bei den sprunghaften Änderungen meist übersprungen werden. So wurden etwas größere Intervalle zugelassen. Während des Experiments

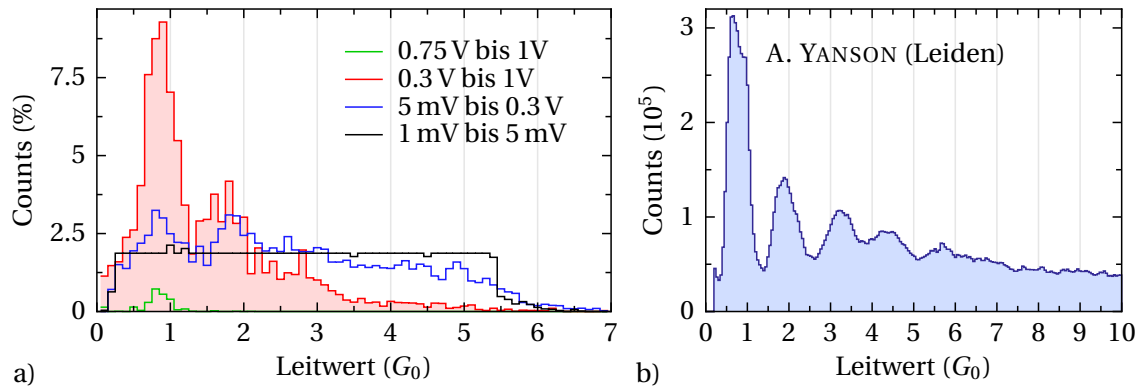


Abbildung 5.3.: Histogramme bei Aluminium. a) 1354 stromgetriebene Öffnungsvorgänge: Erzwingene Anfangsverteilung bis zur stromgetriebenen veränderten Verteilungen in unterschiedlichen Spannungsintervallen. Der höchste Spannungsbereich wurde bezüglich des zweithöchsten normiert. b) 6200 Öffnungsvorgänge: Histogramm bei gleichmäßigen mechanischen Öffnungsvorgängen des Kontakts [Yan].

wurde das Histogramm permanent berechnet und die Lücken häufiger als Zielwert gewählt. Die so erreichte Verteilung war nur grob ausgeglichen. Um in den wochenlang laufenden Messungen eine hohe Ausbeute zu bekommen, wurden die Intervalle dennoch stärker verkleinert und die Ziel-Leitwerte stattdessen erzwungen, indem der Motor der Bruchmechanik die Richtung automatisch änderte, wenn das gewünschte Intervall übersprungen wurde, solange bis ein Zwischenwert erreicht wurde. Damit ergab sich jedoch, wie noch gezeigt wird, eine gewisse Einschränkung der Aussagekraft. Werte mit geänderter Motor-Drehrichtung mussten für eine weitere wichtige Auswertung wieder aussortiert werden.

Selbst mit diesen Techniken war das am Ende berechnete Histogramm noch etwas unausgewogen. Daher wurden bei der Auswertung alle Messungen mit etwas zu häufigen Leitwerten mit einem Gewichtungsfaktor versehen und schwächer gewichtet. Das hat statistisch die gleiche Wirkung wie eine Selektion einzelner Messungen, ist jedoch stufenlos justierbar.²

5.1.3. Ergebnis und Interpretation

In den Abbildungen 5.3 und 5.2 ist zu erkennen, dass die künstlich unterdrückten Bevorzugungseffekte (schwarze Histogramme) nach einigen stromgetriebenen Umordnungen wieder zum Vorschein kommen (rote Histogramme). Dabei liegt die Breite der Peaks bei den

²Als kleines Detail soll noch erwähnt werden, dass auch die Schrittweite bei den stromgetriebenen Umordnungen als eine Art Gewichtung in das Histogramm eingeht. Da bei unterschiedlichen Leitwerten unterschiedliche Spannungsschrittweiten verwendet wurden, musste jeder einzelne Messwert der stromgetriebenen Umordnung eine Gewichtungskorrektur erhalten. Die Gewichtung wurde so gewählt, dass die Histogramme bezüglich gleicher Spannungsschrittweiten einheitlich gewichtet wurden. Die Spannung scheint eher für die Stabilität relevant zu sein (siehe Abb. 5.7, S. 80), so dass z. B. eine Gewichtung bezüglich gleicher Stromschrittweiten oder gleicher Leistungsschrittweiten weniger geeignet wäre. Auch nicht sinnvoll scheinen Histogramme, in denen jedes Plateau nur einmal gezählt wird. Die Stabilität eines Plateaus würde sonst nicht berücksichtigt werden und die Vergleichbarkeit zu den vorliegenden mechanischen Histogrammen wäre weniger gut möglich.

bevorzugten Leitwerten in derselben Größenordnung wie beim gleichmäßigen mechanischen Öffnen. Zu erkennen sind jedoch nur zwei Peaks. Das liegt daran, dass Kontakte mit höherem Leitwert weniger spannungsstabil sind und die Zahl der vorhandenen stabilen Kontakte hier sehr schnell abnimmt. Das bestätigt der jeweils höchste Spannungsbereich ab 0.7 V bzw. 0.75 V, der hier nicht auf die Fläche 1 normiert, sondern in identischer Einheit zur jeweils roten Kurve gewählt wurde. Man kann durch diese Auftragung erkennen, in wie wenigen Fällen der Kontakt bei diesen Spannungen überhaupt noch stabil war. Die Leitwerte dieser Situationen sind in beiden Messungen um den ersten Peak des mechanischen Histogrammes herum konzentriert. Bei diesen Leitwerten liegt also die höchste Spannungsstabilität. Weitere Diskussion zur maximalen Stabilität siehe Seite 80.

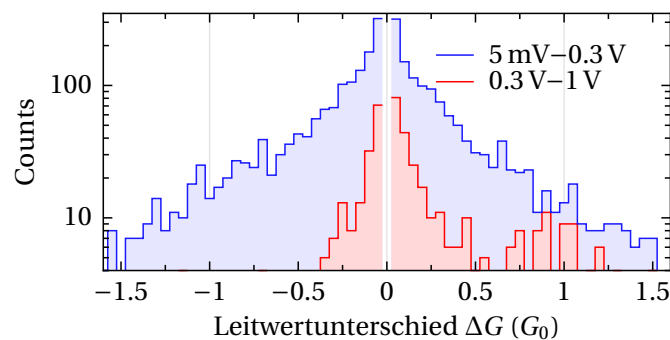


Abbildung 5.4.: Histogramme (logarithmisch) über positive und negative Leitwertsprunghöhen beim Al-Experiment, für zwei verschiedene Spannungsbereiche der Spannung vor dem Sprung. Eine Bevorzugung ist im niedrigen Spannungsbereich nicht erkennbar, bis auf kleine symmetrische Peaks. Der hohe Spannungsbereich zeigt ein stärkeres Abklingen und einen Peak um $+1 G_0$.

Histogramm über Sprunghöhen

Änderungen, die durch die stromgetriebenen Umordnungen zustande kommen, werden als zufällig und lokal angenommen. Die Zufälligkeit wird dadurch bestätigt, dass die Sprunghöhen des Leitwerts, wie in Abb. 5.4 zu sehen ist, fast keine Bevorzugungen zeigen. Sprunghöhnhistogramme können weitere Informationen über Bevorzugungen von Kontakten aufdecken, welche im absoluten Leitwerthistogramm nicht herausgelesen werden können (wie z. B. in einem Sprunghöhnhistogramm über Öffnungs- und Schließkurven [SKB⁺06]).

In Abb. 5.4 ist für den niedrigen Spannungsbereich lediglich ein kleiner Peak bei $+0.65 G_0$ und $-0.7 G_0$ erkennbar. Auch bei $0.8 G_0$ und $1.05 G_0$ sind spiegelbildliche Strukturen vorhanden. Da bei allen Experimenten der Strom ausschließlich erhöht wird, kann die Spiegelsymmetrie nur durch eine echte Bevorzugung bestimmter Sprunghöhen erklärt werden, kaum durch statistische Artefakte. Eine Rückkehr zu einer früheren Geometrie ist bei einer nur steigenden Stromrampe nicht zu beobachten. Nur zwischen einem Experiment und dem darauffolgenden wird die Stromrichtung gewechselt. Die negativen Sprunghöhen haben allerdings nichts mit

der negativen Stromrichtung zu tun. Der Leitwert ist immer positiv. Negative Leitwertänderungen sind atomare Umordnungen zu kleineren Leitwerten und wahrscheinlich kleineren Querschnitten des Kontakts. Eine bevorzugte Sprunghöhe könnte einen Übergang zwischen zwei bevorzugten Leitwerten (oder Geometrien) bedeuten, von Klein nach Groß oder genauso oft umgekehrt. Im höheren Spannungsbereich zeigt sich eine Tendenz des Leitwerts, sprunghaft um etwa $1 G_0$ zuzunehmen. Dass hier kein negatives Spiegelbild vorhanden ist, kann daran liegen, dass mit einer Stromquelle gearbeitet wurde. Verringert sich der Leitwert, steigt die Spannung der Stromquelle entsprechend an und würde bei einem großen Sprung um $-1 G_0$ den Kontakt zerstören.

Die Steigungen der Histogramme zeigen im logarithmischen Plot für die beiden Spannungsbereiche ein unterschiedlich starkes, exponentielles Abklingverhalten. Im hohen Spannungsbereich scheinen große Änderungen sehr viel seltener zu sein, als im niedrigen Spannungsbereich. Das kann daran liegen, dass die Umordnungen bei hohen Spannungen auch weiter weg von der Engstelle stattfinden können. Gleichzeitig sind die lokaleren Freiheitsgrade bereits gehärtet. Die Änderungen, die weiter von der Engstelle entfernt stattfinden, zeigen kleine Leitwertänderungen und sind auch häufiger, da hier mehr Freiheitsgrade zu erwarten sind.

Abgesehen von den Details scheinen die beiden Histogramme etwa exponentiell abzufallen. Nimmt man an, dass kleine Änderungen sowohl statistisch häufiger als auch bei kleineren Strömen auftreten, ist diese Tendenz zu erwarten. Die Messung zeigt, dass die durch den Strom angeregten Änderungen näherungsweise zufällig und nicht auf nur wenige Änderungsarten beschränkt sind.³

Erklärung für die Rückkehr zum mechanischen Histogramm

Das Wiederauftreten der zuvor künstlich unterdrückten Leitwertanhäufungen kann durch eine ungünstige Vorbereitung des Ausgangshistogrammes verursacht worden sein. Häufiger vorkommende Leitwerte wurden nicht durch reine Selektion reduziert, sondern durch mehrfache Drehrichtungsänderungen des Motors, so dass schwer zu erreichende Leitwerte erzwungen wurden. Dadurch könnte sich ein lokales Ungleichgewicht eingestellt haben, welches durch die stromgetriebene Umordnung wieder gelöst wurden. So bleibt die Möglichkeit bestehen, dass die eigentlich wahre Ursache der Bevorzugung weiter von der Engstelle entfernt liegt.

Entstehungsort der Bevorzugten Leitwerte

Um eine strikte Trennung von lokalen und nicht lokalen Eigenschaften erreichen zu können, müssen, wie in Abb. 5.6 schematisch gezeigt, zwei Arten von Änderungen miteinander vergli-

³Bei der Berechnung dieses Histogramms wurden nicht die Differenzen zwischen benachbarten Datenpunkten verwendet. Bei einem Sprung taucht mit gewisser Wahrscheinlichkeit ein falscher Zwischenwert auf, der durch die endliche Mittelungszeit eines Datenpunktes zustandekommt. Der eigentliche, sehr schnelle Sprung kann im ungünstigen Fall genau in der Mitte des zeitlichen Mittelungsintervalls eines Messwerts liegen und einen scheinbaren Datenpunkt auf halber Höhe generieren. Der Sprung taucht dann zweifach beim halbierten Wert im Histogramm auf. Solche Artefakte wurden herausgefiltert, indem die Sprünge nur zwischen den echten Niveaus berechnet werden. Diese wurden dadurch ermittelt, indem zwei benachbarte Messwerte (jeweils vor und nach dem Übergang) mit etwa gleichem Wert als Referenz verwendet wurden.

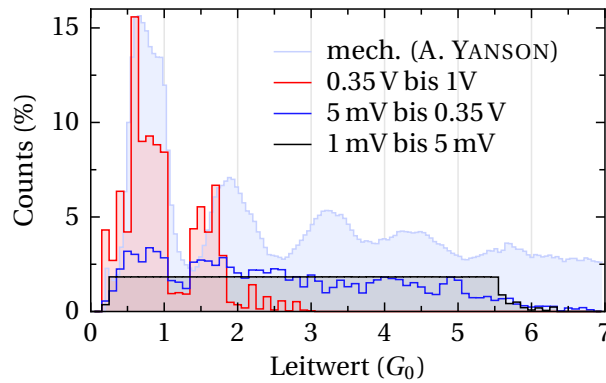


Abbildung 5.5.: Histogramm über lokale Umordnungen, nachdem die nicht-lokalen Transporteigenschaften als Quelle der Bevorzugung durch Selektion ausgeschlossen wurden.

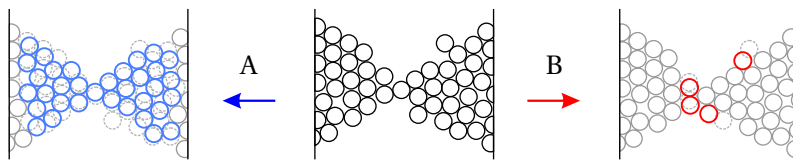


Abbildung 5.6.: Schema eines atomaren Kontakts mit zwei Arten von Änderungen der atomaren Positionen. Änderung A: Nachdem der Kontakt mechanisch geschlossen und wieder geöffnet wurde (Unterschiede in blau). Änderung B: Nach einer stromgetriebenen lokalen Umordnung (Unterschiede in rot).

chen werden. Änderungen über große Bereiche (A) können durch Schließen des Bruchkontakts bis $50 G_0 \approx 260 \Omega$ vorgenommen werden. Wenn im klassischen Widerstandsregime gerechnet wird, entspräche dies einem Durchmesser von ca. 20 % des ungebrochenen Drahtes. Der Kontakt wird also verhältnismäßig weit geschlossen und über weite Bereiche verformt. Lokale Änderungen (B) ergeben sich durch stromgetriebene Umordnungen, die an der Engstelle durch die lokale Konzentration der Stromdichte herbeigeführt werden.

Um zu bestimmen, ob die Bevorzugungseffekte lokalen oder nichtlokalen Ursprungs sind, müssen beide Änderungen miteinander verglichen werden. Bei jedem Experiment muss zunächst der Bruchkontakt gründlich neu formiert werden, indem der Kontakt geschlossen (bis $50 G_0$) und wieder bis zu einem bestimmten Wert geöffnet wird. Anschließend wird die lokale stromgetriebene Umordnung durchgeführt. Wie beim vorigen Experiment muss jeder Leitwert mit gleicher Häufigkeit als Startwert gewählt werden, um ein ausgeglichenes Ausgangshistogramm zu erzeugen. Dabei darf ein seltener Leitwert nicht durch ein lokales Ungleichgewicht ausgeglichen werden, indem ein Leitwert durch Drehrichtungsänderung der Bruchmechanik erzwungen wird. Für diese Auswertung wurden daher nur solche Werte zugelassen, die direkt von $50 G_0$ mit einer reinen Öffnungsbewegung auf einen gewünschten Anfangsleitwert gebracht wurden. Für Gold standen leider nicht genügend solcher Messungen zu Verfügung.

Das Ergebnis für Aluminium ist in Abb. 5.5 zu sehen. Hier wird zum Vergleich nochmals das

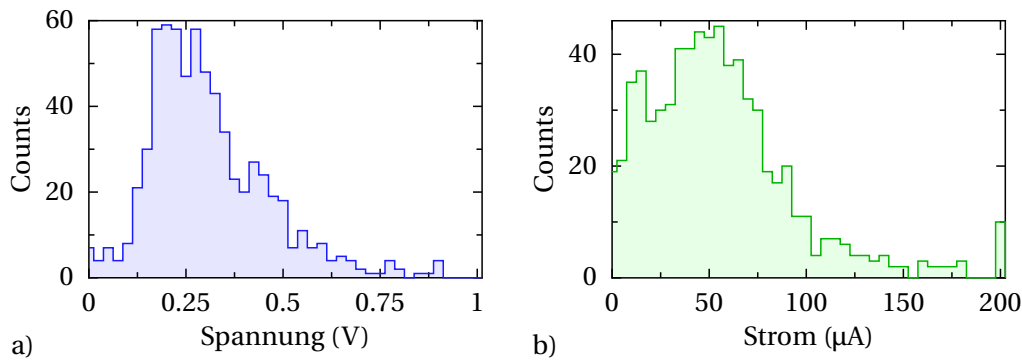


Abbildung 5.7.: Histogramme über die a) Spannungen und b) Ströme, die zur Unterbrechung des Kontakts führten.

mechanisch erzeugte Histogramm gezeigt. Das selektierte bzw. gewichtete Histogramm ist vor der Strombehandlung sehr eben bis $5.5 G_0$. Nach der lokalen Umordnung bilden sich auch hier wieder zwei Peaks mit einem deutlichen Minimum aus. Während der erste Peak in der Position und Breite mit dem mechanischen Histogramm übereinstimmt, ist der zweite allerdings in Richtung kleinerer Werte verschoben und schmaler.⁴ Auch der erste Peak ist leicht zu kleinen Werten hin verschoben. Hierfür könnte der hohe Strom verantwortlich sein, der bei gegebener Spannung selektiv jeweils nur die größeren Leitwerte eines stabilen Peaks instabil macht. Der zweite Peak ragt allerdings etwas ins Minimum des mechanischen Histogramms. Das kann nur bedeuten, dass hier die statistisch unterdrückten nicht-lokalen Eigenschaften des Kontakts tatsächlich durch ihr Fehlen eine Änderung bewirken.

Stabilität

Abbildungen 5.3 und 5.2 zeigen jeweils Histogramme sehr hoher Spannungen ab 0.7 V bzw. 0.75 V, die reduziert auf einen einzelnen schmalen Peak um $0.8 G_0$ (bei Al) und $1 G_0$ (bei Au) auftreten. Das bedeutet, dass die Ein-Atom-Kontakte die höchsten Spannungen aushalten. Bei höheren Werten tritt sehr viel häufiger der Fall ein, dass der Kontakt ganz unterbrochen wird. Das kann durch den höheren Strom erklärt werden, der bei höherem Leitwert und gleicher Spannung fließt. Mit zunehmendem Durchmesser des Kontaktes wächst außerdem die Zustandsdichte der Phononen, die zur Wärmeentwicklung und damit zur Zerstörung des Kontaktes führen kann.

Die geringere Spannungsstabilität bei massiveren Leitern setzt sich übrigens bis in den makroskopischen Bereich fort. Ein Leiterstück aus reinem Aluminium von 2 cm Länge und 1 cm Durchmesser würde bei 1 V einen Strom von 30 000 A leiten. Es würden somit 30 kW Wärmeleistung entstehen, die das Aluminium sofort schmelzen ließe. Das liegt an der kleineren maximalen Stromdichte, die ein massiver Leiter gegenüber einem atomaren Leiter

⁴Hier könnte die berechtigte Vermutung entstehen, dass eine zu kleine Anzahl an Messungen den zweiten Peak verfälscht. Eine genaue Analyse des Histogrammes zeigt aber, dass zum zweiten Peak immerhin 37 verschiedene Einzelmesskurven mehr oder weniger stark beitragen. Beim ersten Peak sind es 50 Messkurven.

aushält. Der bei 1 V durch einen Aluminium-Ein-Atom-Kontakt fließende Strom von $65 \mu\text{A}$ entspräche bei der Querschnittsfläche eines Aluminium-Atoms von $\pi(0.125 \text{ nm})^2$ einer Stromdichte wie $1.3 \times 10^9 \text{ A}$ durch ein Kabel mit 1 mm^2 Querschnitt. Die dabei entstehende Wärme geht in die Bewegung der Atome über und kann nicht schnell genug abgeführt werden. Im atomaren, ballistischen Kontakt wird die Wärme dagegen nicht lokal am größten Widerstand, dem Zentralatom, in atomare Bewegungsenergie umgewandelt. Hier nehmen vielmehr die Elektronen die Wärme auf [DSDV06] und transportieren diese in die nähere Umgebung, wo sie verdünnt an die Atome der Zuleitung abgegeben und abgeführt wird. Die lokale Temperatur kann dadurch niedrig bleiben.

Zur Frage, ob der Strom oder eher die Spannung einen Kontakt zerstört, wurden in Abb. 5.7 Histogramme über die Strom- und Spannungswerte gebildet, die zuletzt vor dem Abreißen des Kontakts gemessen wurden und noch stabil waren. Das Histogramm der *Abrissspannungen* zeigt einen Peak um 0.25 V, ab dem die meisten Kontakte zerstört werden. Im Histogramm des *Abrissstroms* tritt die Zerstörung gleichmäßiger ein. Das ist jedoch nicht verwunderlich, da aufgrund der verschiedenen Leitwerte unterschiedliche Stromdichten zugrunde liegen können. Im klassischen ohmschen Leiter ist die Stromdichte J proportional zur Querschnittsfläche A und diese proportional zum Leitwert G . Klassisch wäre die Stromdichte daher $J = I/A \sim I/G = U$, also proportional zur Spannung. In atomaren Kontakten ist diese proportionale Abhängigkeit nicht gegeben (siehe z. B. die Rechnung in Abb. 2.11, S. 22).

Die Stromdichte wäre eine geeignetere Größe um die Stromstabilität zu beurteilen, als der Strom selbst, der je nach Querschnitt in unterschiedlicher Dichte auftreten kann. Im Stromdichte-Histogramm wäre eher eine Häufung wie im Spannungshistogramm zu erwarten.

5.1.4. Lokale Temperatur

An mehreren Stellen in dieser Arbeit wurde bei der Interpretation der beobachteten Phänomene über thermische Effekte gesprochen. Tatsächlich kann selbst unter kryogenen Umgebungsbedingungen die Temperatur in stromdurchflossenen Nanodrähten lokal weit über Raumtemperatur ansteigen. Besonders in Nanokontakten konzentriert sich die Leistungsdisipation auf ein sehr kleines Gebiet und führt dort zu einer sehr hohen lokalen Temperatur, die nach außen schnell abfällt. Es lohnt sich daher, mit einigen Überlegungen diese lokale Temperatur abzuschätzen.

Im Fall von Gold wurden spezielle Experimente durchgeführt, um die lokale Temperatur zu messen, die während der Elektromigration vorliegt [TVDMVW06, Tro09]. Dazu machten sich die Autoren den thermoresistiven Effekt des Leiters zu Nutze. Der Vergleich der Widerstandswerte bei angeschaltetem und (fast) abgeschaltetem Strom mit unterschiedlichen Umgebungstemperaturen zeigte, dass eine bestimmte kritische Temperatur maßgeblich ist, ab der die Elektromigration einsetzt. Diese Temperatur lag bei etwa 400 K und erwies sich als unabhängig von der Umgebungstemperatur. Die Strukturen hatten zu Beginn eine Engstelle von 100 nm Breite und 15 nm Schichtdicke.

Bei Bruchkontakten aus wenigen Atomen könnten die Instabilitäten möglicherweise bei ganz anderen Temperaturen und durch andere Mechanismen eintreten. Da keine geson-

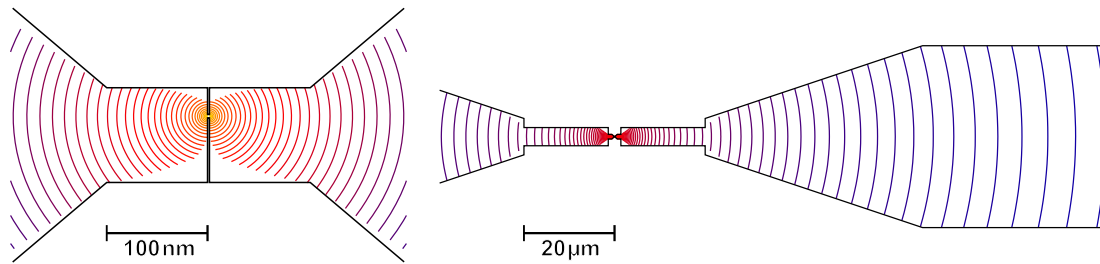


Abbildung 5.8.: Mittlere Schnittebene einer 3D-Berechnung von Schnittflächen zwischen Kugelschalen um den atomaren Kontakt und der Zuleitung. Aus den Flächen wird anschließend das Temperaturgefälle entlang des Leiters berechnet. Für realistischere Werte wird nur eine bestimmte endliche Winkeländerungsrate der Flächenränder zugelassen.

derten Temperaturmessungen durchgeführt wurden, soll im Folgenden eine rechnerische Abschätzung der Temperatur am Beispiel von Aluminium gezeigt werden.

Berechnung

Zur Berechnung der Temperatur muss die elektrische Leistung bekannt sein, die am Kontakt in Wärme umgewandelt wird, sowie die Umgebungstemperatur, die Wärmeleitfähigkeit und die Geometrie des Leiters. In der Literatur zeigt sich, dass die Wärmeleitfähigkeit sehr stark temperaturabhängig [Woo05b, Woo05a] und bei Nanodrähten auch geometrieabhängig [LSC02] ist. Gold zeigt qualitativ ein ähnliches Verhalten [Whi53].

Die Probengeometrie wurde bei der lithografischen Herstellung festgelegt und ist daher genau bekannt (siehe auch Abbildung 4.10, S. 72). Strömt nun im Gleichgewicht die thermische Leistung P durch jedes Leiterelement, so berechnet sich der resultierende Temperaturanstieg $\Delta T(x)$ aus der Länge Δx des Leiterelements und dessen Querschnittsfläche $A(x)$ und aus der Wärmeleitfähigkeit λ über die Wärmeleitgleichung

$$\Delta T(x) = \frac{P \Delta x}{A \lambda}. \quad (5.1)$$

Die Wärmeleitfähigkeit $\lambda(T, b, h)$ hängt von der unbekanntem Temperatur und ebenfalls von den Leiterabmessungen (Breite b und Höhe h) ab. T und b sind wiederum ortsabhängige Größen. Die Wärmeleitfähigkeit ist außerdem stark vom Material und der Verunreinigung abhängig, welches jedoch konstante Größen sind. Die Berechnung des Temperaturverlaufs beginnt an der Kontaktierungsfläche der lithographischen Struktur, wo als Temperatur die Umgebungstemperatur $T_{\text{Kryo}} \approx 1 \text{ K}$ des Kryostaten angenommen wird. Die Temperatur $T(x)$ berechnet sich über die iterative Lösung der Differenzialgleichung

$$T(x) = T_{\text{Kryo}} + \int_0^l \frac{P}{A(x)\lambda(T(x), b(x), h(x))} dx. \quad (5.2)$$

Da die Temperatur am Ende des Leiters $x = l$ bekannt ist, kann von dort beginnend schrittweise entlang des Leiters integriert werden, bis der Bruchkontakt $x = 0$ erreicht wird.

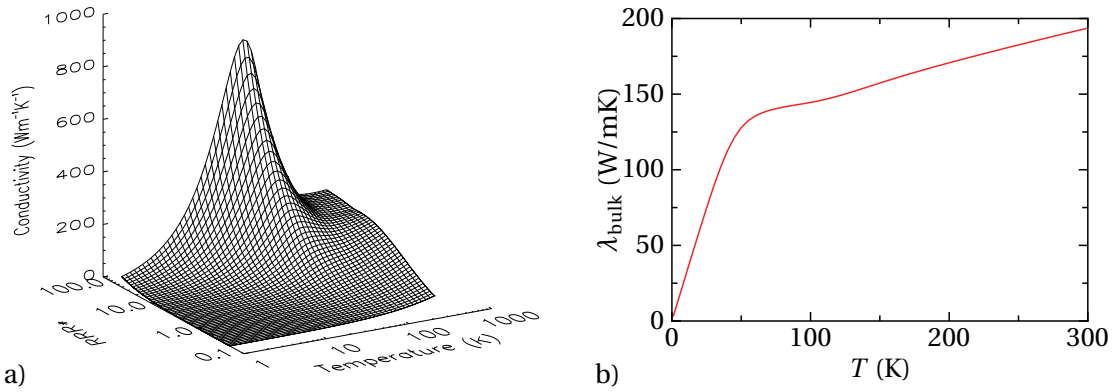


Abbildung 5.9.: a) Wärmeleitfähigkeit in Abhängigkeit des Restwiderstandverhältnisses (RRR) und der Temperatur (Grafik aus [Woo05b]). b) Wärmeleitfähigkeit beim Restwiderstandverhältnis von $RRR = 4$. Quelle: im Internet veröffentlichte Tabellen der Daten von [Woo05b].

Zunächst muss die Querschnittsfläche $A(x)$ berechnet werden. Dazu wurde ein 3D-Modell des Leiters erstellt (siehe Abbildung 5.8). Um den Bruchkontakt herum wurden Kugelschalen mit wachsendem Radius gelegt und die Schnittflächen $A(x)$ mit dem Leiter berechnet. Diese Näherung führt bei Unstetigkeiten der Leitergeometrie zu Unstetigkeiten der Schnittflächen $A(x)$. Um physikalisch unmögliche Knicke in der Gleichgewichtswärmestromdichte zu vermeiden, wurden für die Flächen nur endliche Änderungsraten erlaubt. Dies führt die Näherung näher an die Geometrie realistischer Äquipotenzialflächen des Wärmestroms heran. Als zweite geometrische Größe wird $b(x)$ für die geometrieabhängige Wärmeleitfähigkeit berechnet.

Die Wärmeleitfähigkeit $\lambda(T, RRR)$ für Aluminium kann der Literatur entnommen werden. Für die von A. WOODCRAFT zusammengetragenen Messungen [Woo05b, Woo05a] wurden Tabellen veröffentlicht, die als 3D-Plot in Abbildung 5.9 gezeigt sind. Hier wurden die Wärmeleitfähigkeiten für den gemessenen konstanten Wert $RRR = 4$ verwendet. Die Abhängigkeit von $\lambda(b)$ von der Breite b des Nanodrahtes findet sich in [LSC02, SBMH09], wo ein Korrekturfaktor

$$\frac{\sigma_{\text{Nanodraht}}}{\sigma_{\text{bulk}}} = \left[1 + \frac{3}{8} \left(\frac{\lambda_e(T)}{b} + \frac{\lambda_e(T)}{h} \right) + \frac{36}{25} \frac{\lambda_e}{D} \frac{R}{1-R} \right]^{-1} \quad (5.3)$$

in Abhängigkeit der Leiterabmessungen angegeben wird. Über das Wiedemann-Franz-Gesetz lässt sich mit diesem Verhältnis der elektrischen Leitfähigkeiten die thermische Leitfähigkeit abschätzen. Hier ist D die Korngröße des Aluminiums. Für die in dieser Arbeit hergestellten Aluminiumfilme wurde mit Röntgendiffraktometrie eine Korngröße von $D = 40 \pm 10$ nm abgeschätzt. Der Wert für den Korngrenzenreflexionskoeffizient $R = 0.59$ wurde aus der genannten Literatur übernommen. Die Schichtdicke des Leiters beträgt $h = 100$ nm. Als variable Größen verbleiben die Breite $b(x)$ und die mittlere freie Weglänge der Elektronen $\lambda_e(T)$. Letztere Größe kann aus der umgeformten Drudeformel

$$\lambda_e(T) = \frac{m^* v_F}{e^2 n \rho(T)} \quad (5.4)$$

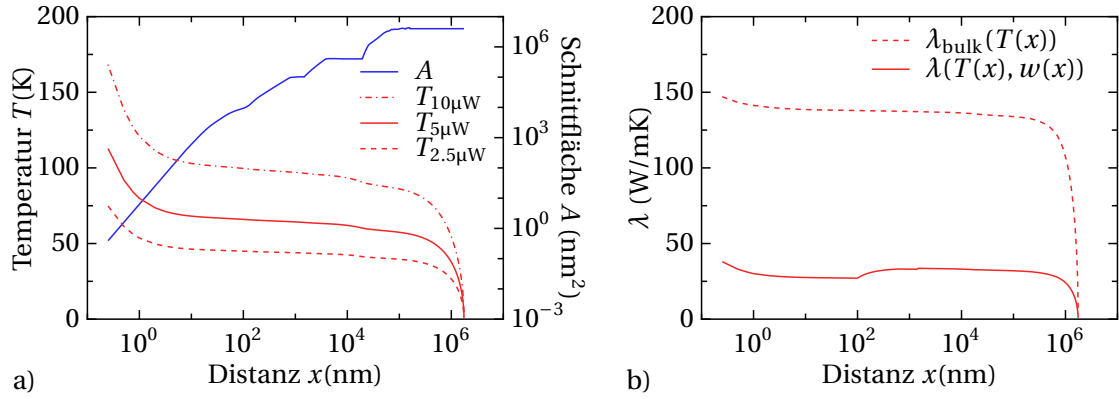


Abbildung 5.10.: a) Berechnete Querschnittsfläche des Wärmetransports entlang des Leiters (blau) und die berechnete Temperatur (rot) im Aluminium-Leiter (für $2.5 \mu\text{W}$, $5 \mu\text{W}$ und $10 \mu\text{W}$). Begonnen wurde die Berechnung an der Kontaktierungsfläche der Zuleitung, wo die Temperatur bei ungefähr 1 K liegt. Bei der Berechnung des Temperaturverlaufs entlang des Leiters wurde die sich ändernde temperaturabhängige Wärmeleitfähigkeit berücksichtigt. b) Berechnete Wärmeleitfähigkeit entlang des Leiters. Zum Vergleich ist die Wärmeleitfähigkeit gezeigt, wie sie ohne Berücksichtigung des Größeneffekts im Bulk-Aluminium auftreten würde. Der Raumtemperatur-Bulk-Wert liegt bei 273 W/m K .

mit der Fermigeschwindigkeit v_F , der Elektronendichte n und der effektiven Elektronenmasse m^* berechnet werden. Die Zahlenwerte sind in Tabelle 5.1 zusammengefasst. Aus dem Restwiderstandsverhältnis RRR und der Raumtemperatur-Elektron-Phonon-Streulänge $\rho_{\text{el-phon}}$ erhält man $\rho_0 = \rho(T = 0)$ über die Definition

$$\text{RRR} = \frac{\rho_0 + \rho_{\text{el-phon}}}{\rho_0} \quad (5.5)$$

und die Temperaturabhängigkeit von $\rho(T)$ über das BLOCH-GRÜNEISEN-Gesetz⁵

$$\rho(T) = \rho_0 + \alpha \left(\frac{T}{\theta_D} \right)^5 \int_0^{T/\theta_D} \frac{x^5}{(e^x - 1)(1 - e^{-x})} dx. \quad (5.6)$$

Hier ist θ_D die Debye-Temperatur und α eine durch die Anpassung an ρ_0 und $\rho_0 + \rho_{\text{el-phon}}$ gewonnene Konstante. Für das untere Temperaturlimit ergibt sich für die mittlere Freie Weglänge $\lambda_e(T = 0) = 63 \text{ nm}$.

Sind alle Abhängigkeiten der Größen bestimmt, kann ausgehend von den Zuleitungen die Temperatur ΔT stückweise aufsummiert werden. Das Ergebnis ist in Abbildung 5.10 links für drei verschiedene typische Leistungen dargestellt. Die rechte Grafik zeigt, wie sich die thermische Leitfähigkeit entlang des Leiters verändert. Die Rechnung ist bestenfalls eine Abschätzung mit einer Unsicherheit von ca. 50%. Das Ergebnis reagiert sehr empfindlich auf die Korngröße des Aluminiumfilms, die nur näherungsweise gemessen werden konnte

⁵Quellen und tabellierte Werte der BLOCH-GRÜNEISEN-Formel siehe Landolt-Börnstein-Datenbank [Bas]

und am Kontakt selbst ihre Gültigkeit als mittlere Größe verliert. Die umgebungsabhängige Wärmeleitfähigkeit wirkt sich andererseits im Vergleich zur äußeren Umgebungstemperatur stabilisierend auf die Temperatur am Bruchkontakt aus, so dass die Randbedingung der äußeren Temperatur keinen großen Einfluss auf die berechnete Temperatur hat.

Konstante / Parameter	Symbol	Wert	Quelle
Restwiderstandsverhältnis	RRR =	4	Messergebnis
Korngröße	$D =$	40 ± 10 nm	Messergebnis ⁶
Korngrenzenreflexionskoeffizient	$R =$	0.59	[SBMH09]
Fermigeschwindigkeit	$v_F =$	2.02×10^8 cm/s	[Kit06, Tab. 6.1]
Elektronendichte	$n =$	18.06×10^{22} /cm ³	[Kit06, Tab. 6.1]
Schichtdicke	$h =$	100 nm	Probenparameter
Effektive Elektronenmasse	$m^* =$	$1.45 m_e$	[Pyt67]
Spezif. Widerstand von Al (297 K)	$\rho_{\text{el-phon}} =$	2.74×10^{-6} Ω cm	[Kit06, Tab. 6.3]
Spezif. Widerstand von Al (0 K)	$\rho_0 =$	9.13×10^{-6} Ω cm	Berechnet
Debye-Temperatur von Aluminium	$\theta_D =$	396 K	[WGW02]
Konstante in BLOCH-GR.-Formel	$\alpha =$	5.97	Fitparameter

Tabelle 5.1.: Material-Konstanten und Parameter der hier untersuchten Aluminium-Probe, als Grundlage für die gezeigte Temperaturberechnung.

Fazit

Die Temperatur steigt gemäß dieser Rechnung am Ein-Atom-Kontakt deutlich an bis etwa 100 K. Die gezeigte lokale Konzentration der Temperatur Abbildung 5.10 in der Größenordnung einzelner Atome basiert jedoch auf einer idealisierten Rechnung. Einerseits wird die Temperatur dadurch überschätzt. So würde man bei einem einzelnen Kanal der Transmission Eins erwarten, dass die elektrische Leistung nicht vollständig in die thermische atomare Bewegung der Zentralatome übergeht, sondern dass die Elektronen der näheren Umgebung diese Energie größtenteils aufnehmen [DSDV06], was das lokale Temperaturmaximum entsprechend niedriger ausfallen ließe. Andererseits gehen Materialparameter wie die Korngröße oder Streulänge in die Rechnung ein, die unterhalb der eigenen Größenordnung nicht mehr als mittlere homogene Eigenschaften interpretiert werden können, insbesondere, da der Ein-Atom-Kontakt selbst eine Korngrenze darstellt. Dies könnte die lokale Temperatur um einiges höher oder niedriger ausfallen lassen, als in der idealisierten Rechnung eines Kontinuums.

Wenn die gesamte Leistung tatsächlich in einem ein oder zwei Atome umfassenden Steuerelement als Wärme konzentriert an das Gitter abgegeben wird, zeigt die Rechnung, dass das Aluminium beim Schalten immer noch weit von seiner Schmelztemperatur (933 K, [Lid04]) entfernt wäre. Für eine Erklärung als thermisches Phänomen wäre die abgeschätzte Temperatur auch zu unterschiedlich. Auch bei den Messungen an Gold [TVDMVW06, Tro09] liegt die Temperatur mit 400 K noch weit von der Schmelztemperatur (1337 K) entfernt.

⁶Durchführung und Interpretation der Röntgenmessung erfolgten durch HANS BALLOT und ANTON PLECH.

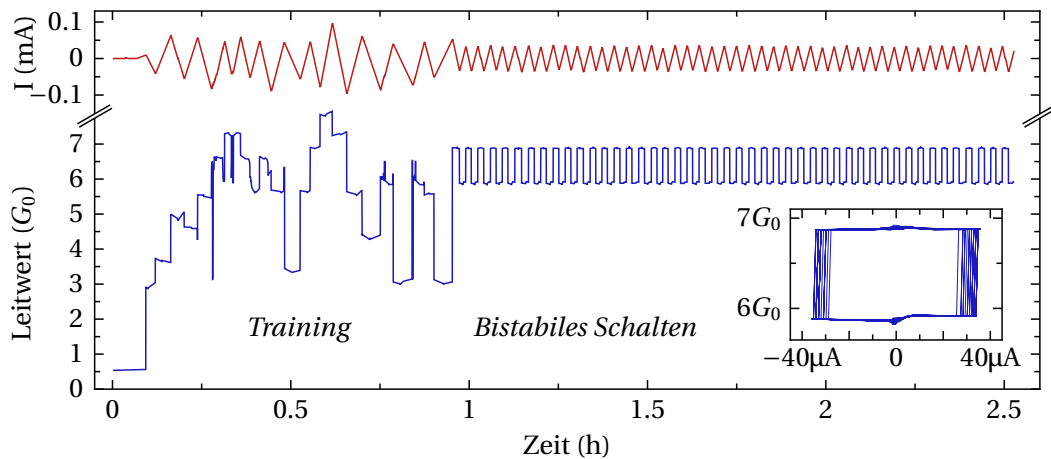


Abbildung 5.11.: Bistabiles reversibles Schalten zwischen zwei Leitwerten, nach einigen anfänglichen Umordnungen.

5.2. Bistabiles Schalten

Bei einzelnen strominduzierten atomaren Umordnungsvorgängen ist unklar, wie viele Atome sich bei einzelnen Leitwertsprüngen ändern. Durch die mechanische Spannung, die sich beim Erzeugen des Kontakts aufbaut, könnten sich mehrere Atome in einer energetisch ungünstigen Situation befinden. Sobald ein Atom durch Elektromigration umgeordnet wird, könnten dadurch weitere instabile Atome zur Umordnung angeregt werden und eine stabilere Position einnehmen. Die Elektromigrationsereignisse könnten daher eine lokale Härtung oder Ausheilung eines Verbunds von mehreren Atomen darstellen. Die Experimente des vorigen Abschnitts (Abb. 5.1) deuten eine solche Härtung an. Für den Fall, dass sich viele Atome pro Umordnung bewegen, könnte ein spezieller vorausgehender Härtungsvorgang die Zahl der pro Umordnung bewegten Atome immer weiter reduzieren.

Abbildung 5.11 zeigt ein Experiment, bei dem eine solche Härtung durchgeführt wurde. Eine Stromrampe wird dabei so lange hochgefahren, bis die erste Umordnung stattfindet. Der Kontakt ist dadurch etwas stabiler geworden und der Strom müsste weiter erhöht werden, um eine weitere Umordnung auszulösen. Das erhöht die Gefahr, den Kontakt ganz zu zerstören. Daher wird die Richtung des Stroms umgekehrt und der Strom in negativer Richtung hochgefahren bis zur nächsten Umordnung, und so weiter. Die jeweils instabilste Stelle, die vom Strom in Bewegung versetzt wird, kann so gefestigt werden. Durch das Wechseln der Stromrichtung wird allerdings dafür gesorgt, dass möglichst eine andere als die zuletzt gefestigte Stelle mechanisch beansprucht wird. Dadurch treten insgesamt mehr Umordnungen auf, bevor der Strom zu hoch und der Kontakt ganz unterbrochen wird. Außerdem treibt die Elektromigration Material nicht immer in die selbe Richtung, was zur Ausdünnung der Engstelle führen könnte. In einem von zehn Fällen tritt ein Sonderfall auf, der in Abb. 5.11 ab einer bestimmten Zeit zu sehen ist. Offenbar ist der Kontakt größtenteils gehärtet. Es wird nur noch zwischen zwei energetisch ähnlichen Situationen hin und her geschaltet. Dieser sehr eingeschränkte Spielraum kann eine

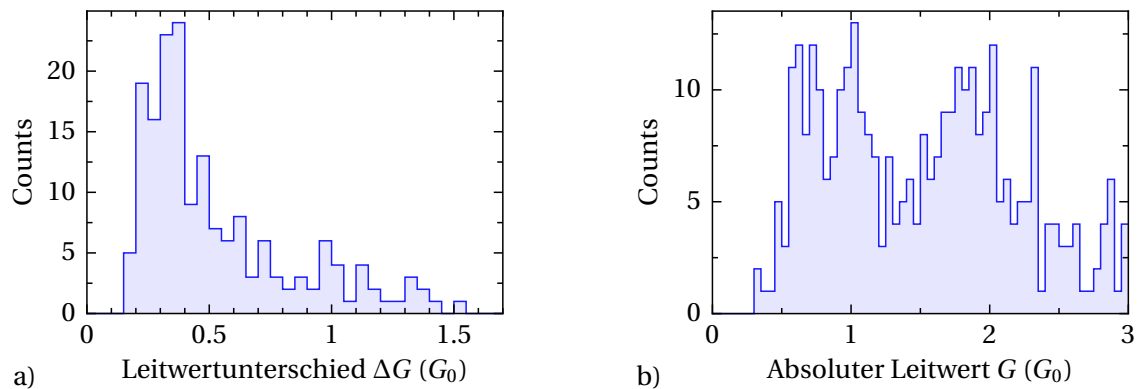


Abbildung 5.12.: a) Histogramm über die Sprunghöhen zwischen den beiden bistabilen Zuständen. b) Die selben Daten als Histogramm über die absoluten Leitwerte.

vermutlich starke räumliche Einschränkung des Schaltvorgangs bedeuten. Möglicherweise ist hier nur noch ein einzelnes Atom im Spiel, das zwischen zwei Positionen hin und her springt. Das Inset zeigt die Hystereseschleife, bei der der Leitwert gegen den Strom beim Schaltvorgang aufgetragen wurde.

Durch die bistabilen Schaltzustände werden aus den stromgetriebenen Umordnungen solche ausgewählt, die sehr kleinen und lokalen Änderungen des Kontakts entsprechen. Damit lassen sich zwei Kontakte mit minimalem lokalem Unterschied in ihren Transporteigenschaften vergleichen, um differenzierter der Frage nach dem Einfluss der atomaren Geometrie auf den elektronischen Transport nachgehen zu können.

Vergleich vieler Hysteresen

Abbildung 5.13 zeigt viele solcher Hystereseschleifen von bistabilen Schaltsituationen. Alle unterscheiden sich im Leitwert, in der Leitwertdifferenz und der Stromschwelle, das heißt in der Breite der Hysterese. Außerdem sind die Formen der IV-Kennlinien sehr unterschiedlich. Bistabile Schaltzustände können also nicht bestimmten atomaren Kontaktgeometrien zugeordnet werden, sondern treten generell in Erscheinung. Der Leitwertbereich wurde hier eingeschränkt auf Leitwerte bis etwa $3 G_0$. Wie jedoch das Beispiel in Abb. 5.11 zeigt, sind bistabile Kontakte auch bei höheren Leitwerten zu finden. Die Leitwertunterschiede der Hysteresen häufen sich, wie in Abb. 5.12 a) zu sehen ist, um $0,3 G_0$ und werden zu höheren Leitwerten seltener. Die genaue Form wird allerdings stark von den Vorgaben des Experiments beeinflusst. Hier wurden Intervalle vorgegeben, innerhalb denen sich der Leitwert bewegen darf und als Sprung wurden nur Umordnungen mit $\Delta G > 0,125 G_0$ gewertet. Um einen Unterschied von $1 G_0$ herum scheint dennoch eine leichte Häufung zu liegen, die nicht durch die Vorgaben beeinflusst wurde.

Das Leitwerthistogramm über die zugehörigen Absolutwerte, in Abb. 5.12 b), zeigt die gewohnten Maxima des Öffnungskurven-Histogramms. Auf das kleine Minimum beim typischen Ein-Atom-Kontakt um $0,8 G_0$ wird später noch eingegangen (Abschnitt 5.5, S. 101). Auch hier sind Einflüsse der experimentellen Parameter für die genaue Form verantwortlich. Die jeweili-

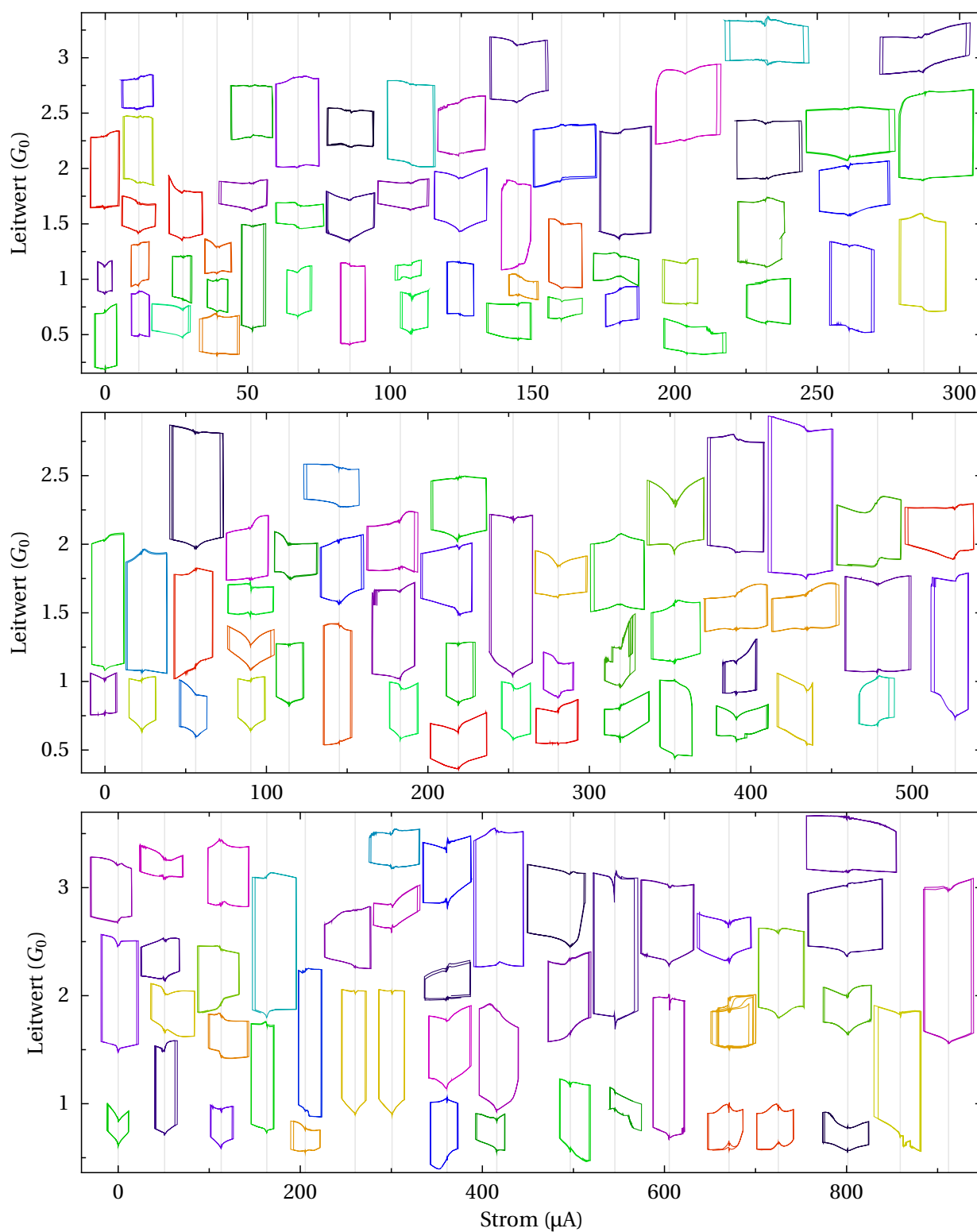


Abbildung 5.13.: Hystereseschleifen beim Schalten bistabiler Kontakte. Die Messungen wurden zur Übersicht mit einem Offset im Strom dargestellt. Der Leitwert ist unverändert. Die Sortierung folgt grob der Breite und die Farben deuten eine zeitliche Reihenfolge an.

gen Startwerte lagen im Bereich $0.5 G_0$ bis $2.5 G_0$ (zum Teil bis $3 G_0$), möglichst gleichmäßig verteilt mit einer Toleranz bei der Selektion von $\pm 0.4 G_0$.

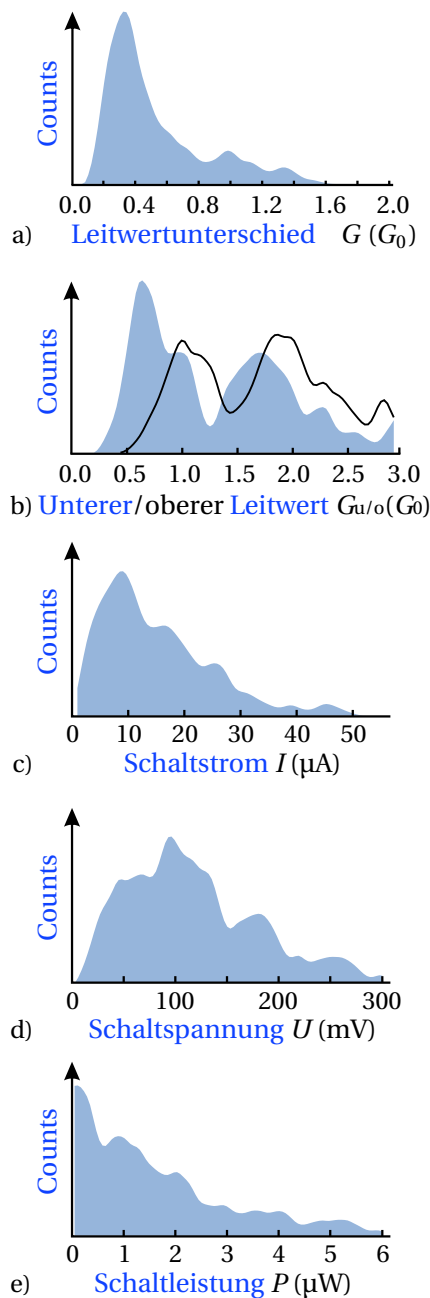


Abbildung 5.14.: Histogramme über diverse Größen

Abb. 5.14 zeigt Histogramme weiterer Werte. Zur besseren Übersicht sind diese Histogramme geglättet dargestellt. Das Histogramm in Abb. 5.14 a) zeigt nochmals die selbe Kurve von Abb. 5.12 a) zur Demonstration der Glättung. Im Histogramm b) wurden zwei separate Histogramme über absolute Leitwerte gebildet und in die jeweils oberen und unteren Leitwerte der Hysterese aufgeteilt. Wie zu erkennen ist, ergibt sich ein ähnliches jedoch leicht versetztes Histogramm mit zwei Maxima. Der Offset entspricht in etwa dem Maximum im Leitwertunterschiedshistogramm a). Der Schwellstrom und die Schwellspannung, die zum Schalten führen, sind in den Histogrammen c) und d) gezeigt. Berechnet man aus diesen Größen die Leistung, die kurz vor dem Schalten am Kontakt dissipiert wurde, ergibt sich das Histogramm e). Die Verteilung war hier sehr breit und lag zwischen 8 nW und $15 \mu\text{W}$, wobei die kleineren Werte häufiger auftraten. In der Hälfte der Fälle war die Leistung kleiner als $2.5 \mu\text{W}$. Mit der Abschätzung der Temperatur in Abschnitt 5.1.4, S. 81 wurde gezeigt, dass die Temperatur bei diesen Heizleistungen in der Größenordnung 70 K und weit unter der Schmelztemperatur von Aluminium (933 K) liegt. Die sehr variable Leistung bei sehr ähnlichen Kontakten lässt auch keine andere kritische Temperatur erwarten, die für das Eintreten der Umordnungen verantwortlich sein könnte.

Viele der bistabilen Kontakte ändern sich nach einigen Durchläufen. Maximal wurden 500 identische Durchläufe gemessen, meist wird die Hysterese schon früher verlassen. Reproduzierbare Durchläufe sind auch im Fluktuationmuster und im Kanalspektrum (s. Abschnitt 5.4) identisch. Leichte Variationen treten häufig im Schaltstrom auf. Der Strom besitzt selbst eine sehr hohe Stabilität, und kann nicht für das unterschiedliche Schalten verantwortlich sein. Stattdessen kommt gleichermaßen der Tunneleffekt und die endliche Temperatur in Frage, was den Schaltstrom variieren lässt. Beide Effekte könnten auch dafür verantwortlich sein, dass der bistabile Zustand sich meist nach einigen Durchläufen verändert oder ganz verlassen wird.

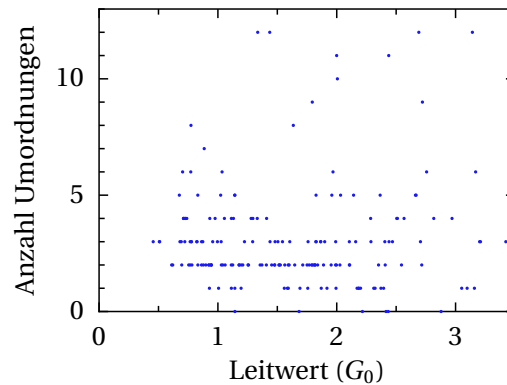


Abbildung 5.15.: Anzahl der atomaren Umordnungen, bis ein bistabiler Zustand eingenommen wurde (als *Training* bezeichneter Abschnitt in Abb. 5.11, S. 86).

Trainingsphase

Der in Abb. 5.11 als *Training* bezeichnete Abschnitt, in dem kein bistabiles Schalten auftritt, kann bei kleinen Leitwerten sehr kurz werden. Abbildung 5.15 zeigt für die Messungen in Abb. 5.13 die Länge der vorausgegangenen Trainingsphase als Anzahl der Umordnungen. In einigen Fällen war der Kontakt nach der mechanischen Neuformierung von vornherein bistabil. Obwohl zunächst vermutet wurde, dass bistabile Kontakte höherer Leitwerte längeres Training benötigen (wie z. B. Abb. 5.11 vermuten lässt), zeigt sich im Bereich bis $3 G_0$ noch keine deutliche Zunahme der benötigten Umordnungen bis zum bistabilen Kontakt. Insgesamt ist die Zahl der Umordnungen bei dieser Probe relativ klein im Gegensatz zu anfänglichen Messungen. Das kann aber daran liegen, dass beim Schließen und Neuformieren des Kontaktes die bereits gehärtete Struktur nicht wieder völlig neutralisiert wurde. Das Gedächtnis für die Vorgeschichte wurde durch die Messungen der Shell-Effekte bereits gezeigt (Abschnitt 2.3.2, S. 21). So könnte die Probe auch insgesamt „trainiert“ worden sein, bistabiles Schaltverhalten zu zeigen.

Atomares Schalten in anderen Arbeiten

Das mechanische Schalten eines rein metallischen atomaren Kontakts, angeregt durch eine elektrische Spannung, wurde auch in anderen Arbeiten gezeigt. In der Arbeit von XIE ET AL. [XNOS04] wurde ein Silberdraht in einer Flüssigkeit elektrochemisch bis auf einen atomaren Kontakt verjüngt. Durch das elektrochemische Potenzial konnte in dem gezeigten Fall der Leitwert zwischen 0 und $1 G_0$ reproduzierbar hin und her geschaltet werden. Interessanterweise konnten auch die Leitwerte $2 G_0$ und $3 G_0$ gezielt eingestellt werden, wie behauptet wurde.

Ein weiterer atomarer Schalter wurde von THERABE ET AL. [THNA05] vorgestellt. Hier ist zwar kein einzelnes Silber-Atom in Bewegung, sondern mehrere. Sie lassen sich jedoch ebenfalls allein durch eine anliegende Spannung in Bewegung versetzen. Die Spannung treibt die Silberatome als Ionen durch ein Schwefelgitter und aus dem Schwefel heraus, worauf sie einen Kontakt zu einem Platindraht herstellen können. Das Schalten kann ebenfalls zwischen

einem Leitwert 0 und $1 G_0$ reproduzierbar wechseln. Und auch hier lassen sich Leitwerte von ganzzahligen G_0 bis zu $6 G_0$ direkt anspringen.

In beiden Fällen scheint sehr fraglich zu sein, in wie weit diese deutliche Leitwertquantisierung tatsächlich bevorzugten Leitwert zuzuschreiben ist, oder in wie weit sie durch eine spezielle Wahl des Schaltimpulses aufgeprägt wird, da reine Silber-Leitwerthistogramme nur bei $1 G_0$ einen scharfen Peak zeigen. In beiden Veröffentlichungen wurde diese Leitwertquantisierung nicht detailliert untersucht und beschrieben. In verunreinigten atomare Kontakten können auch kontinuierliche Leitwertänderungen mit der Öffnungsbewegung erzielt werden, wo sich jeder beliebige Wert, möglicherweise reproduzierbar, einstellen lässt. Die beiden sehr speziellen Fälle von einerseits einer Flüssigkeit und andererseits eines Kontaktes zwischen möglicherweise unterschiedlichen Metallen, könnte aber auch zu neuen Effekte führen, die in reinen Silberkontakten nicht auftreten.

Im Vergleich zu diesen Arbeiten konnten die hier gezeigten Schaltvorgänge sehr viel genauer untersucht werden. Insbesondere kann durch die Öffnung des Kontaktes bei tiefen Temperaturen eine Verunreinigung ausgeschlossen werden. Die Bestätigung liefert die Kanalanalyse (Abschnitt 5.4), die sehr deutlich die zentralen Atome als Aluminium identifiziert.

Eignung als Speicher?

Zahlreiche Veröffentlichungen der letzten Jahre zeigen vielfältige Ideen, wie bistabile oder hysteretische Systeme jeglicher physikalischer Art für Computer-Speicher genutzt werden könnten.⁷ Die hier gezeigte Informationsspeicherung stellt die kleinste denkbare räumliche Konzentration eines „mechanischen“ Speicher-Bits dar, das außerdem zum Beschreiben das Minimum von nur zwei elektrischen Zuleitungen benötigt. Das Schalten geht prinzipiell sehr schnell, auch wenn die ersten Versuche mit 100 kHz bisher zu keiner Hysterese geführt haben. Schnelle Pulse, anstelle von Rampen zeigten jedoch den selben Schalteffekt oberhalb einer bestimmten Spannungsschwelle, wie in Abbildung 5.16 zu sehen ist. Die beiden Leitwertniveaus liegen jedoch von einem zum anderen erzeugten Kontakt leider bei sehr verschiedenen Werten. Für eine Verwendung des Effektes als Speicher müssten atomar ähnliche Kontakte vorliegen, dass die Hysteresen vergleichbare Größen annehmen. Die Leitwertunterschiede von 10 bis 100 % und Schaltspannungen von 50 bis 300 mV würden zumindest technisch durchaus brauchbare Größen für crossed-wire-Speicher darstellen [Mus06]. Allerdings sollte die Richtung definiert vorliegen, ansonsten müsste bei jedem Schreib- oder Lesevorgang zunächst die Orientierung bestimmt werden oder in einem zweiten Bit müsste die Information über die Orientierung des ersten in einem anfänglichen Formatierungsvorgang gespeichert werden. Die Stabilität bei Raumtemperatur dürfte auch ein großes Problem darstellen, da selbst bei tiefen Temperaturen die Hysteresen, wie oben erwähnt, bereits thermische Instabilitäten zeigen.

Direkt lässt sich also kein Speicher damit realisieren, auch wenn einige der vielen Vorausset-

⁷Zum Beispiel [GML09, YZNT08, YPX⁺08, LKDK08, JWB⁺07, SSBW06, THNA05, COL⁺03, NNWA01] und einige weitere. Drei Systeme mit resistivem Schaltverhalten scheinen sich besonders als Favoriten für den Massenmarkt für nichtflüchtige Speicher hervorzutun [Mei08].

zungen dafür erfüllt sind. Ein erfolgreiches Beispiel eines Speichers mit atomaren Kontakten zeigt dagegen z. B. die bereits erwähnte Arbeit von TERABE ET AL. [THNA05]. Hier wird eine etwas größere Anzahl Silberatome bzw. Silberionen bewegt, die sich mittels Spannung durch ein Schwefelgitter hindurchbewegen lassen. Sobald das Silber über den Schwefel hinausragt, kann es zu einem Platindraht einen Ein-Atom-Kontakt bilden und so den Speicherzustand "1" repräsentieren. Durch weitere Miniaturisierung könnte der schaltbare Speicher prinzipiell auf sehr wenige Atome reduziert werden.

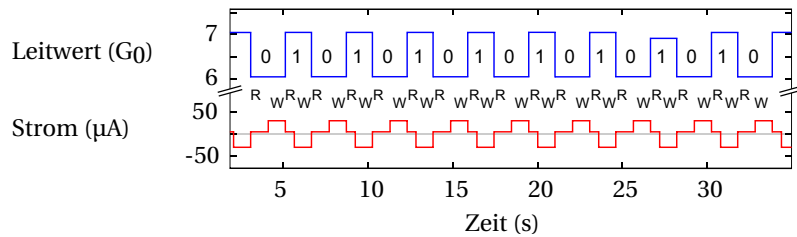


Abbildung 5.16.: Schreib-/Lesezyklen eines atomaren 1-Bit-Speichers

5.3. Materialgüte und Supraleitung

Zur Kontrolle der Probenqualität und der Kontakte der Zuleitungen konnten Temperatur- und Supraleitungseffekte zu Nutze gemacht werden. Bei 250 mK und reinem Aluminium-Ausgangsmaterial (99.999%) mit einer nominellen Sprungtemperatur um 1.175 K [Lid04] sind die Voraussetzungen für Supraleitung erfüllt.

Zuleitungswiderstand und RRR

Vor den eigentlichen Messungen wurde zunächst der Widerstand der Probe bei drei charakteristischen Temperaturen verglichen. Für die Probe der Nummer cs_b37 fiel der Widerstand von 38.2Ω bei Raumtemperatur auf 9.4Ω bei tiefen Temperaturen (1.5 K). Im supraleitenden Zustand (250 mK) verblieben $< 100 \text{ m}\Omega$, was für eine gute Kontaktierung der Probe spricht. Das Restwiderstandsverhältnis RRR (residual resistance ratio, als $\text{RRR} = R(300 \text{ K})/R(4 \text{ K})$ definiert) von $38.2/9.4 \approx 4$ stellt einen guten Wert für Proben dieser Art dar. Messungen an gekerbten $25 \mu\text{m}$ -Aluminium-Drähten eines Wire-Bonders ergaben höhere Werte (> 10). Hohe Werte bedeuten, dass der Widerstand bei Raumtemperatur hauptsächlich durch Phononenstreuung dominiert ist. Bei tiefen Temperaturen sind die Phononen ausgefroren und der Widerstand entsprechend geringer. Bei kleinem RRR findet diese Verringerung nur in geringen Maßen statt. Ursachen sind temperaturunabhängige Störstellen, wie Korngrenzen oder Fremdatome, oder reduzierte Phononenzustandsdichten bei eingeschränkter Geometrie. Für reines Bulk-Aluminium ist $\text{RRR} \approx 20000$ [Lid04].

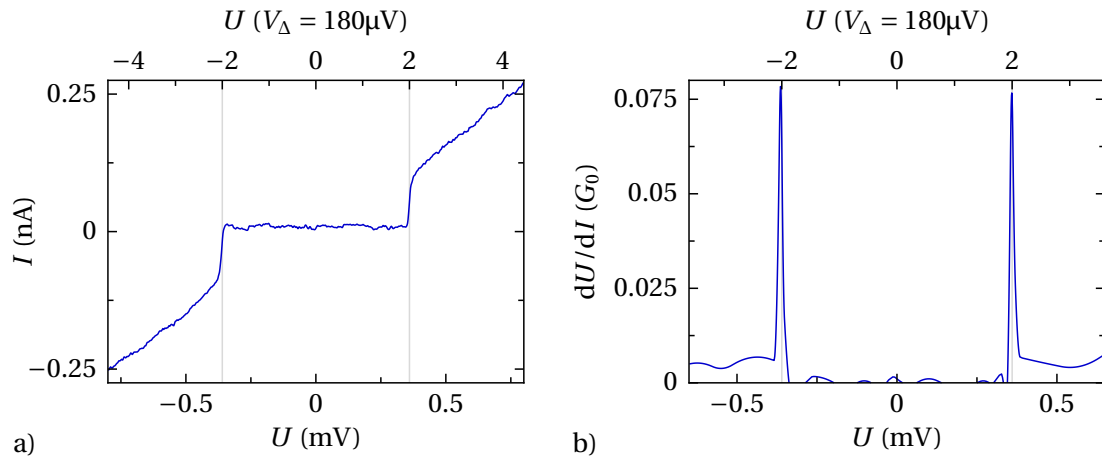


Abbildung 5.17.: Bandlücke der Supraleitung. a) gemessene IV-Kennlinie, b) numerische Ableitung mit Hilfe von glättenden Splines.

Energiespektrum

Wird die Probe gebrochen und auf einen Tunnel-Widerstand von etwa $3 \text{ M}\Omega$ eingestellt, kann durch Variation der Spannung die elektronische Zustandsdichte mithilfe des differentiellen Widerstands abgebildet werden. Im Bereich bis 1 mV kann die Bandlücke der Supraleitung beobachtet werden. Abb. 5.17 zeigt in a) die gemessene IV-Kennlinie und in b) eine numerische Ableitung mittels analytisch abgeleiteter Spline-Funktionen. Das Messrauschen schränkt diese numerische Behandlung etwas ein, jedoch sind die scharfen Bandkanten bei $\pm 2\Delta$ deutlich zu sehen. In der oberen Skala sind die Einheiten dieser Bandlücken aufgetragen, wobei hier $\Delta = 180 \mu\text{eV}$ angenommen wurde. Die Peaks sollten bei diesen Temperaturen von 250 mK noch viel schärfer und höher ausfallen. Die Ursache der Verrundung liegt allein im Rauschen der Anregungsspannung.

Kritische Größen I_c , B_c , T_c

Zusätzliche Aussagen über das Material liefern die kritischen Größen der Supraleitung. Alle drei Größen wurden erst am Ende der Messungen bestimmt, als die Probe cs_b37 bereits über 4 Monate geöffnet war. Obwohl, wie oben gezeigt, der Widerstand anfangs $< 100 \text{ m}\Omega$ war, konnte die Probe hinterher nur noch bis zu einem Wert von $5\text{--}10 \Omega$ geschlossen werden. Dieser Widerstand entsteht vermutlich durch die sich mit der Zeit veränderte Form der Bruchstelle (durch ständige mechanische Bearbeitung). Abbildung 4.11, S. 72 zeigt eine REM-Aufnahme der Probe mit Nummer cs_b37. Hier ist die Bruchstelle deutlich zu erkennen (die selbe Aufnahme ist auf dem Umschlag dieser Arbeit etwas größer gezeigt, Seite 1). Die beiden Seiten scheinen abgerundet zu sein, was häufig (aber nicht immer) an bereits gebrochenen Kontakten beobachtet werden kann. Der eigentliche Kontakt ist daher sehr klein, was die $5\text{--}10 \Omega$ ausmachen könnte. Das bestätigt Abb. 5.18, wo der kritische Strom an der inneren Struktur sehr schnell überschritten wird. Abb. 5.19 zeigt im Vergleich eine Messung an einer Probe, die

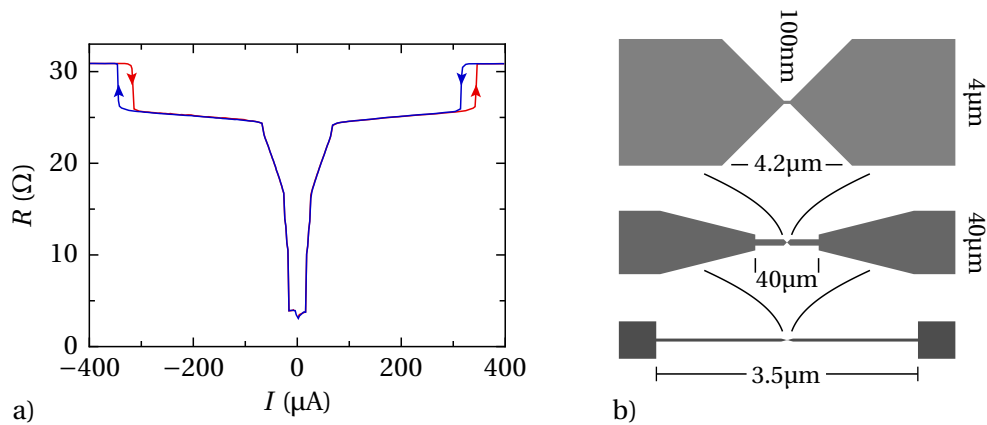


Abbildung 5.18.: Kritischer Strom. Die Messung erfolgte nach wieder geschlossener Brücke.

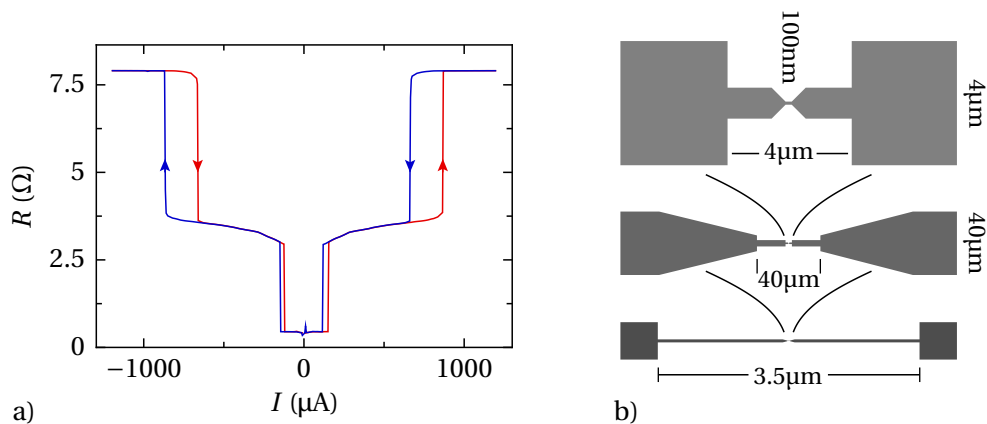


Abbildung 5.19.: Kritischer Strom einer anderen, ungebrochenen Probe.

vorher nicht gebrochen wurde. Die Geometrie war hier zwar etwas anders gewählt, das sollte jedoch keinen Einfluss auf die Kennlinie haben. Der innere Abschnitt ist hier etwas breiter.

Die Stufenform kann durch die unterschiedlich breiten Bereiche des Leiters erklärt und über die Widerstände der Normalleitung identifiziert werden. Die 3,5 mm lange Zuleitung macht in beiden Fällen den Sprung beim höchsten Strom aus. Einiges früher springen die schmalere Bereiche.

Die Hysteresen zeigen einen Effekt, der durch die Bandlücke der Supraleitung zustande kommt. Bei steigendem Strom, unterhalb des kritischen Stromes, fällt zunächst die komplette Spannung über dem Zuleitungswiderstand ab. Wird der kritische Strom überschritten und dadurch ein Teil des Leiters normalleitend, muss die Stromquelle, um den Strom zu halten, zusätzliche Spannung auf das System legen. Beim Senken des Stromes befindet sich im normalleitenden Bereich lokal die Spannung zunächst außerhalb der (doppelten) Bandlücke, die in Abb. 5.17 gezeigt ist. Berührt die lokale Spannung den Peak der Bandkante, steigt der Widerstand stark an. Die Stromquelle muss die Spannung entsprechend hochregeln. Damit

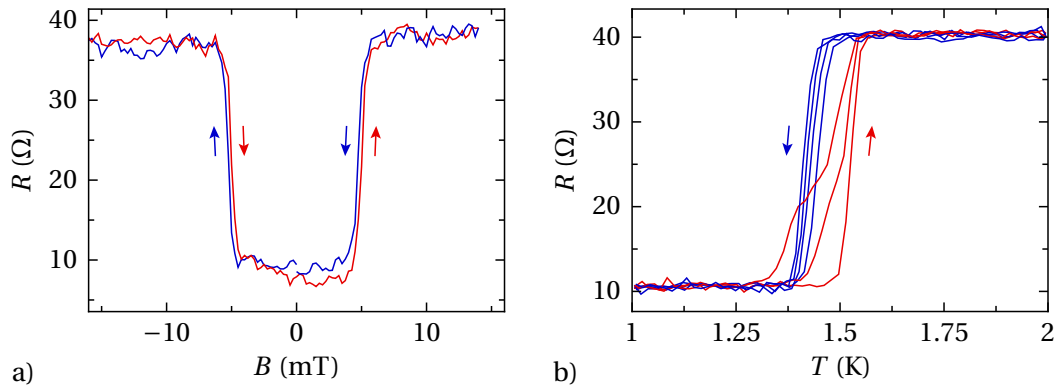


Abbildung 5.20.: a) Kritisches B-Feld und b) kritische Temperatur.

steigt die lokale Spannung wieder und die lokale Spannung bleibt am Peak hängen. Dies verzögert den Eintritt in den supraleitenden Bereich, wodurch eine Hysterese in der IV-Kennlinie sichtbar wird.

Auch bei einem kritische B-Feld und einer kritische Temperatur konnte der Zusammenbruch der Supraleitung beobachtet werden (Abb. 5.20). Da diese Messungen ebenfalls nach Beendigung der Bruchkontakt-Experimente durchgeführt wurden und der verbleibende Widerstand bereits auf $10\ \Omega$ gestiegen war, entsprechen die Übergänge möglicherweise nicht mehr denen des reinen Aluminiums. Die kritische Temperatur (gemessen mit dem probenfernen ^3He -Thermometer) lag bei deutlich zu hohen Werten.

5.4. Kanalanalyse

Die bisherigen Ergebnisse basieren alle auf Messungen des Gesamtleitwerts des atomaren Kontakts. Mikroskopisch durchlaufen die Elektronen den Kontakt jedoch in unterschiedlichen quantenmechanischen Moden, den Leitwertkanälen, die im Abschnitt 2.1.5, S. 15 beschrieben wurden und die mit Messungen der Multiplen Andreev-Reflexionen (MAR) (Abschnitt 2.2, S. 17) und durch eine numerischen Analyse (Abschnitt 3.1, S. 27) experimentell zugänglich sind. Die Messung erfolgt, indem das zur Unterdrückung der Supraleitung angelegte Magnetfeld heruntergefahren wird und wie in Kapitel 4 beschrieben, die hochauflösende IV-Kennlinie bis $1\ \text{mV}$ gemessen wird.

Abbildung 5.21 zeigt den Ablauf einer Messung mit bistabilem Schalten und Messung der MAR. a) zeigt links eine Öffnungskurve zur Erzeugung eines atomaren Kontakts als Funktion der Strecke, über die das Substrat durchgebogen wurde. Weiter rechts wird als Funktion der Zeit das bistabile Schalten im Leitwert und die zugehörige Stromrampe gezeigt. In diesem Beispiel ist (bis auf eine kaum sichtbare kleine Umordnung) der Kontakt fast von Anfang an in einem bistabilen Zustand, ohne eine längere Trainingsphase. Nach der Detektion der Bistabilität, wird jeweils im Nulldurchgang der Stromrampe die Messung unterbrochen und eine MAR-Messung durchgeführt. Die dafür benötigte Zeit (jeweils ca. 15 Min.) wurde aus der

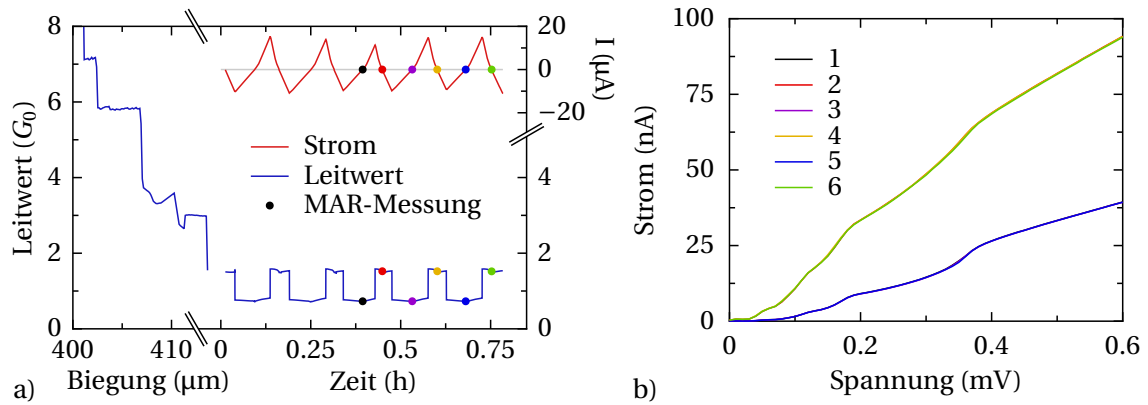


Abbildung 5.21.: a) Ablauf einer Messung mit bistabilem Schalten. b) Sechs gemessene MAR-Kennlinien. Dabei überdecken sich sowohl die drei geraden als auch die drei ungeraden vollständig.

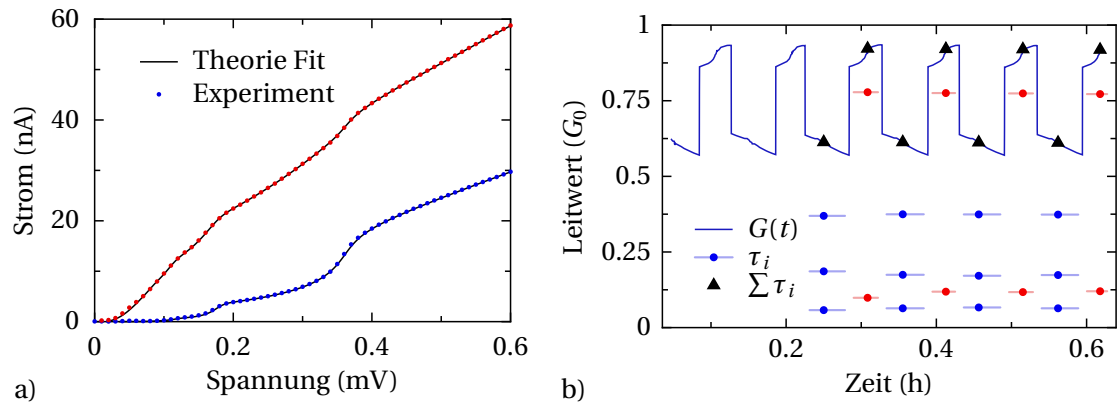


Abbildung 5.22.: a) Gemessene MAR-Kennlinie und Fit zweier Zustände. b) Zeitlicher Leitwertverlauf mit den eingetragenen Leitwerten der ermittelten einzelnen Kanäle.

Zeitachse herausgeschnitten. b) zeigt alle sechs gemessenen MAR-Kennlinien in einem Graph. Jeweils drei Kennlinien liegen dabei exakt übereinander, was bestätigt, dass tatsächlich nur zwei verschiedene Zustände eingenommen werden.

Abbildung 5.22 zeigt das Ergebnis der im Kapitel 3.1 beschriebenen numerischen Kanalanalyse, angewendet auf eine Reihe von Messungen. a) zeigt die beiden ersten MAR-Kennlinien aus b). Die Messpunkte wurden hier in reduzierter Zahl dargestellt, um die dahinter liegende durchgehende gefittete Theoriekurve sichtbar werden zu lassen. Die mittlere Abweichung zwischen Messung und Fit liegt bei etwa 40 pA. Die gefundenen Fitparameter geben nun die Transmissionswahrscheinlichkeiten der einzelnen Kanäle wieder und können direkt als Vielfache des Leitwertquantums G_0 zusammen mit dem Gesamtleitwert in b) geplottet werden. Die Summen der Leitwertkanäle wurden ebenfalls eingetragen als schwarze Dreiecke. Sie liegen sehr genau auf dem direkt gemessenen Leitwert und bestätigen die MAR-Messungen, die mit einer ganz anderen Messelektronik (außerdem im supraleitenden Zustand) durchge-

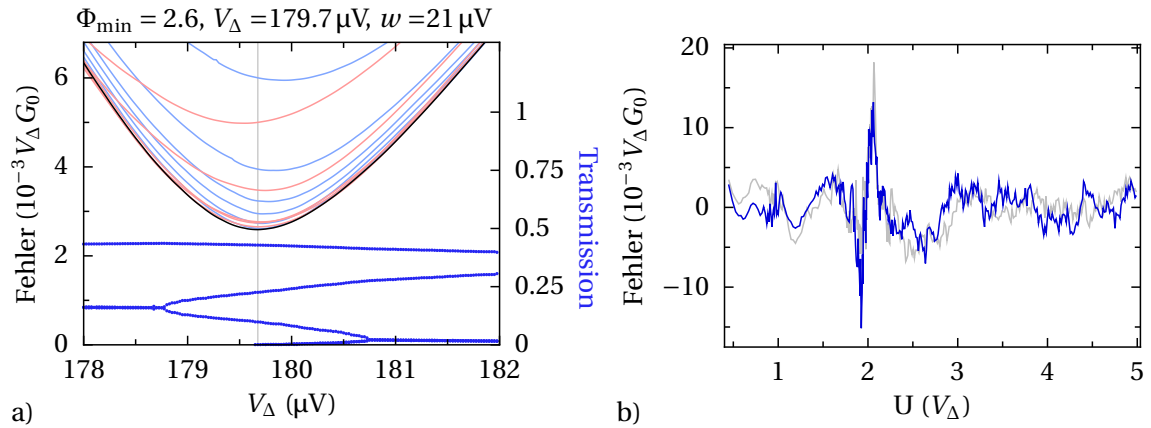


Abbildung 5.23.: Fitfehler. a) Abhängigkeit des Fitfehlers von der angenommenen Bandlücke V_{Δ} und zusätzliche von der Variation der Gaußbreite w der Verrundung. Schwarz: Fitfehler beim besten w , hellrot/hellblau: größere/kleinere Werte von w , blau: gefundene Kanäle passend zum besten w . b) Verbleibende Differenz zur Theoriekurve, positive und negative Flanke der IV-Kennlinie. Die Einheit $10^{-3} V_{\Delta} G_0$ der Fehlerachse entspricht etwa 14 pA.

führt wurden. Hier hätte der direkt gemessene Leitwert einen ganz anderen und außerdem spannungsabhängigen Wert. Dadurch ist sowohl die Messung, die MAR-Theorie, als auch die numerische Analyse bestätigt.

Gültig sind diese Kanalspektren nur für den Nulldurchgang der Stromrampen. Der mit dem Strom sich kontinuierlich ändernde Leitwert muss sich auch in den einzelnen Kanälen bemerkbar machen. Sie sind jedoch nur um Null herum messbar, da die Bestimmung selbst auf einer Spannungsrampe um Null herum basiert.

Der Theoretische Fit enthält als weitere Parameter die Breite w (Halbwertsbreite FWHM) des Rauschens und der Temperatur (die die MAR-Kennlinie verrunden) und den genauen Wert der Bandlücke V_{Δ} der Supraleitung. Diese Werte konnte nicht direkt gemessen werden und mussten zunächst ermittelt werden. Dazu wurden Fits mit unterschiedlichen Werten w und V_{Δ} durchgeführt und das Minimum des Fehlers bestimmt. Abb. 5.23 a) zeigt eine solche Variation der verschiedenen Parameter. Die durchgezogenen Linien geben den minimalen Fehler als Funktion der Bandlücke an. In schwarz ist die Kurve gezeigt, die bezüglich des Rauschens w den kleinsten Fehler liefert. In rot sind die entsprechenden Kurven für höhere Werte von w zu sehen und in blau für kleinere Werte. In allen Untersuchungen fanden sich die kleinsten Fehler bei einem Rauschen um $w=20 \mu\text{V}$ und einer Bandlücke um $V_{\Delta}=180.0 \mu\text{V}$. Dieses Wertepaar wurde für alle folgenden Fitprozeduren verwendet. Abb. 5.23 b) zeigt zum gefundenen Minimum die Differenz zwischen der Messung und dem Fit für die positive und negative Flanke der MAR-Kennlinie. Während in diesem Beispiel der Fit bezüglich Variationen relativ stabil ist, zeigt Abb. 5.23 zwei Messungen, bei denen die gefundenen Transmissionen (blau) sehr viel stärker, bzw. sehr viel schwächer von der angenommenen Bandlücke abhängen.

Die verbleibenden mittleren Fehler des Fits liegen meist um $5 \cdot 10^{-3} V_{\Delta} G_0 \approx 70 \text{ pA}$ und sind

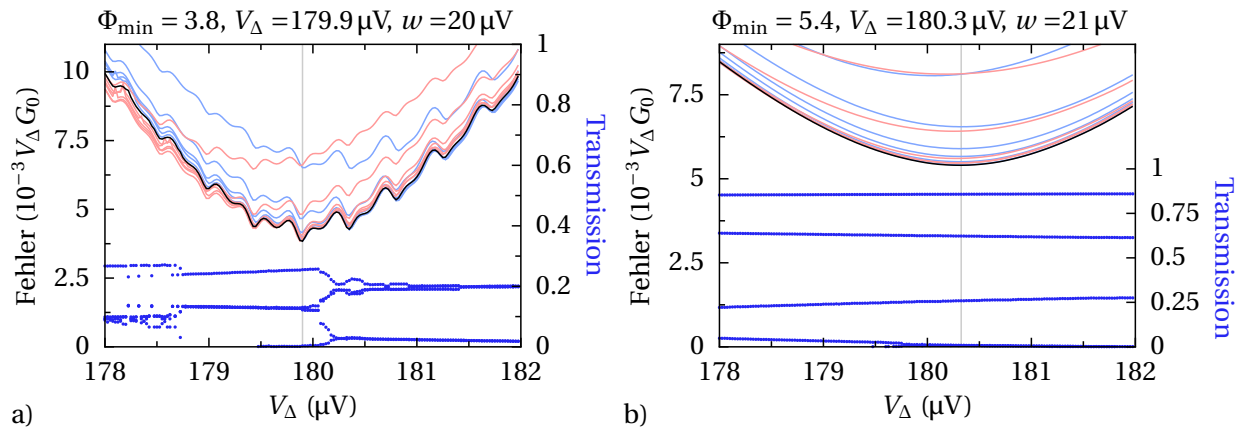


Abbildung 5.24.: a) Extrem empfindlicher Fit bei vier sehr kleinen ähnlichen Kanälen. b) Sehr robustes Ergebnis bei drei weit auseinander liegenden nicht zu kleinen Kanälen.

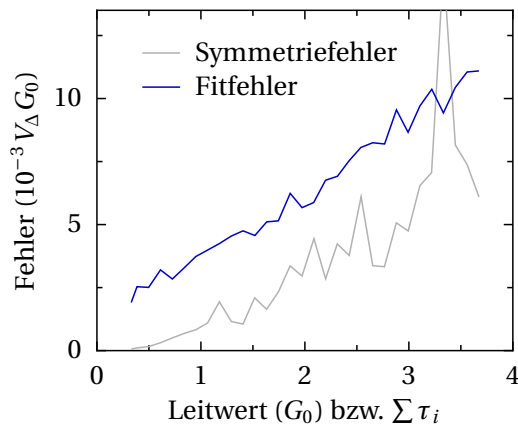
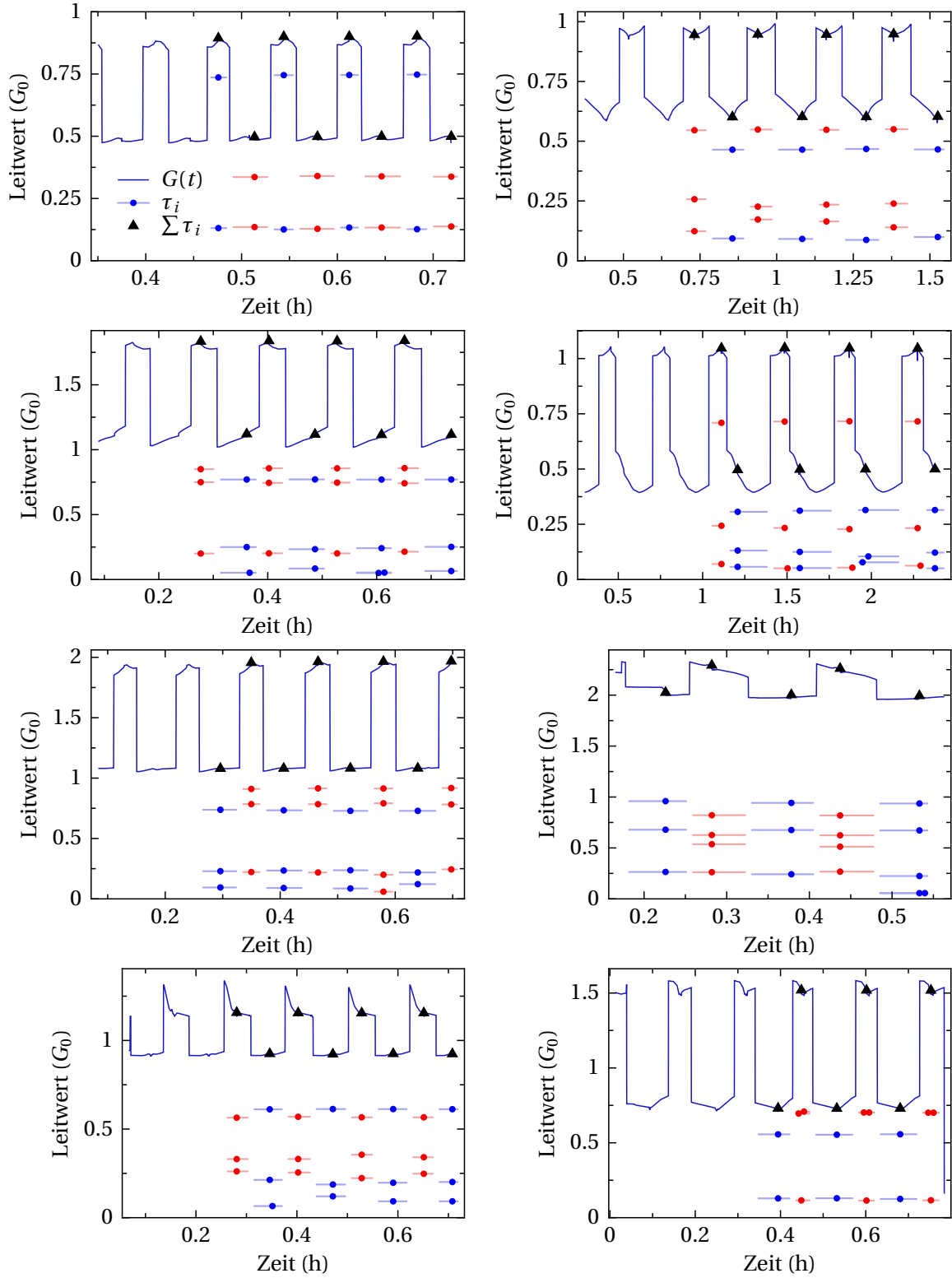


Abbildung 5.25.: Symmetriefehler (mittlere Abweichung zur punktgespiegelten IV-Kennlinie) und Fitfehler (mittlere Abweichung zum Theoriefit) in Abhängigkeit des Gesamtleitwerts. Dargestellt sind gemittelte Fit-Ergebnisse von 1580 gemessenen MAR-Kennlinien.

kaum sichtbar. Dennoch können (aufgrund der hohen Zahl von Freiheitsgraden) Abweichungen dieser Größenordnung die Ergebnisse stark beeinflussen. Vor jedem Fit wird zunächst der genaue Symmetriepunkt zwischen positiver und negativer Flanke ermittelt, um den meist geringen Offset der Quelle und der Verstärker zu kompensieren. Die Differenz der gespiegelten Flanken liegt im Mittel etwas niedriger bei etwa $3 \cdot 10^{-3} V_{\Delta} G_0$. Abb. 5.25 zeigt die Abhängigkeit des Symmetrie- und Fitfehlers vom Leitwert des Kontakts. Der Symmetriefehler wächst ungefähr linear mit dem Leitwert, was als Stromrauschen der Quelle (bzw. als Spannungsrauschen über dem Referenzwiderstand) oder als Spannungsrauschen des Verstärkers gedeutet werden kann. Der Fitfehler hängt auch vom selben Rauschen wie der Symmetriefehler ab, und zusätzlich von der nicht genau modellierbaren Rauschverteilungsfunktion. Er ist daher jeweils größer.

Abbildung 5.26 (Seiten 99 u. 100) zeigt 14 weitere Beispiele von insgesamt 190 ähnlichen



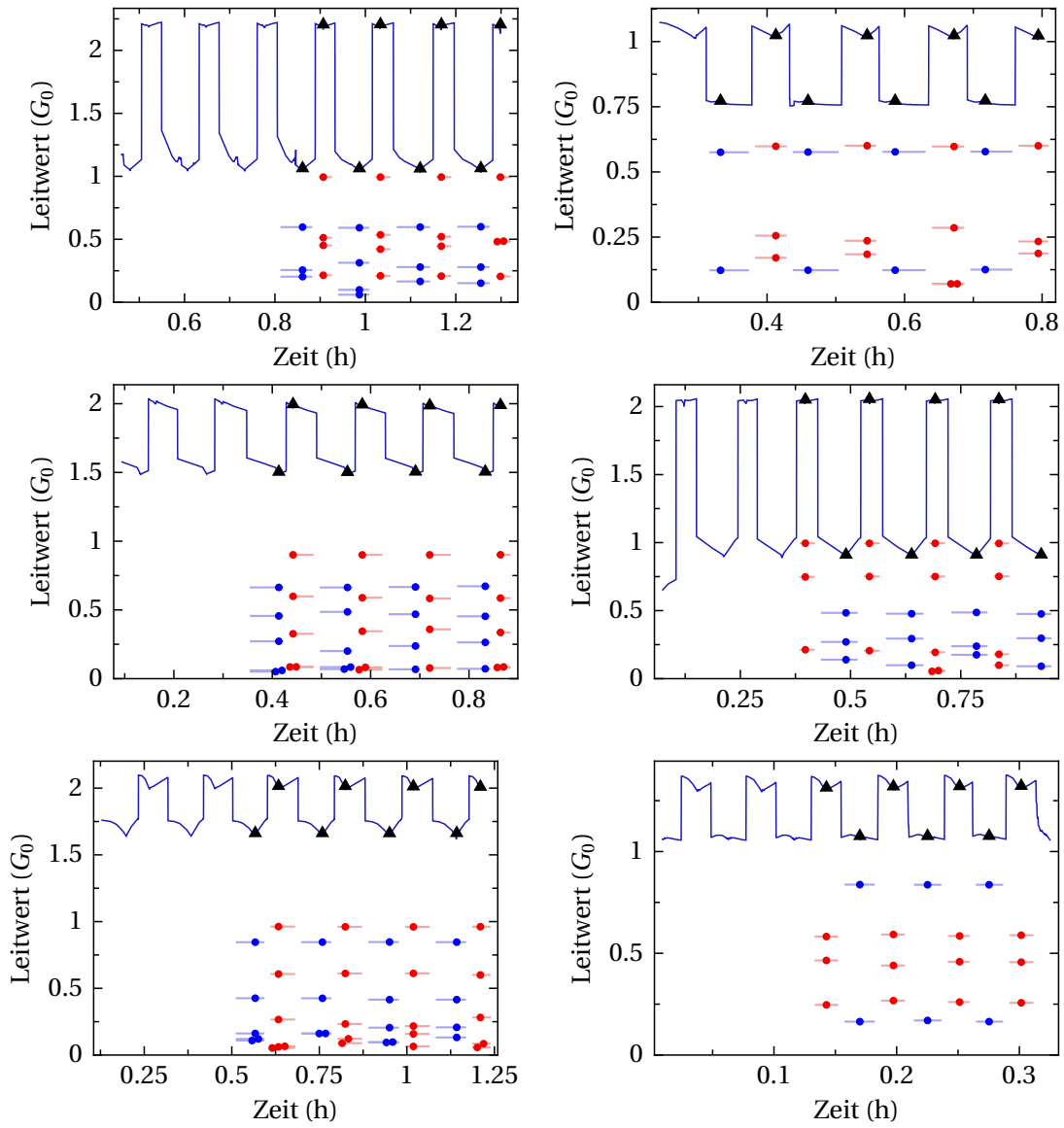


Abbildung 5.26.: Eine Auswahl von 14 weiteren der insgesamt 190 ähnlichen Messungen. Geplottet wurden nur Kanäle $\tau > 0.05$. Beim Fit wurden alle Kanäle unterhalb der Transmission $\tau < 0.01$ bis auf einen Reservekanal geschlossen.

Messungen. Die Auswahl zeigt Fälle mit besonders deutlicher Reproduzierbarkeit in allen Details. Die gewonnenen Kanalspektren zeigen ebenfalls die selben Kanäle. Die meisten hier nicht gezeigten Messungen weisen eine etwas schlechtere Reproduzierbarkeit der numerisch gewonnenen Kanalspektren auf, obwohl die zugehörigen MAR-Kennlinien, mit blosem Auge beurteilt, exakt gleich erscheinen. In einer früheren Auswertung ohne die Modellierung der Rauschkurve war die Reproduzierbarkeit der Spektren unter den zusammengehörigen Zuständen einiges besser. Gleichzeitig tauchten jedoch auch Kanalgruppen auf, die (wie im Kapitel 3.1 gezeigt wurde) auf schlechte Fits hindeuten und die mit der scheinbar sehr hohen Stabilität des Fits zusammenhängen. Den jetzigen Werten kann mehr Verlässlichkeit zugetraut werden, obwohl sie stärker schwanken. Fehlerbalken wurden hier weggelassen, da eine verlässliche quantitative Größe für einzelne Messungen nur mit großem Rechenaufwand zu bekommen wäre. Als Richtwert kann in den typischen Zusammensetzungen bei hohen Transmissionen ein Fehler von etwa ± 0.03 angenommen werden, bei kleinen Transmissionen etwa ± 0.1 , wobei diese Werte stark von der genauen Zahl und Konstellation der Kanäle beeinflusst werden.

5.5. Leitwertkanalhistogramme

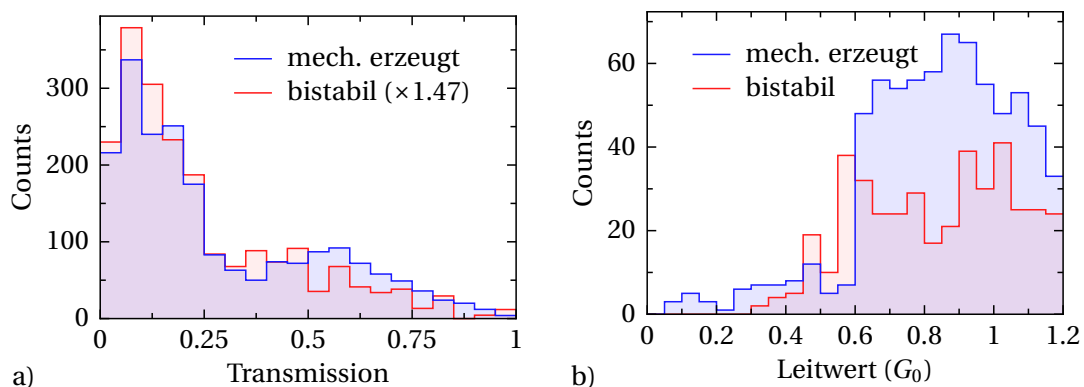


Abbildung 5.27.: a) Histogramm über 2018 (mech.) und 1367 (bistab.) Kanäle aus 713 und 409 Kanalspektren bis zum Leitwert $1.2 G_0$. b) Histogramm über die jeweils zugehörigen Summen der Kanäle bzw. über die Leitwerte.

Um die Zusammensetzung der Ein-Atom-Kontakte aus einzelnen Kanälen zu untersuchen, wurden von allen 190 gemessenen bistabilen Schaltzuständen diejenigen Kanalspektren mit Gesamtleitwert $< 1.2 G_0$ ausgewählt (409 Spektren) und in einem Histogramm (Abb. 5.27 a) dargestellt. Zum Vergleich wurde ein Experiment durchgeführt, bei dem Spektren nach einer rein mechanischen Formierung der Kontakte gemessen wurden (713 Spektren). Nach jeweils 5–10 Messungen an nur leicht veränderten Kontakten wurde der Kontakt bis zu einem Leitwert von $50 G_0$ geschlossen und neu geformt, so dass die gute statistische Durchmischung zu einer reproduzierbaren Aussage führt.

Es ist kein deutlicher Unterschied zwischen bistabilen und mechanisch erzeugten Kontakten

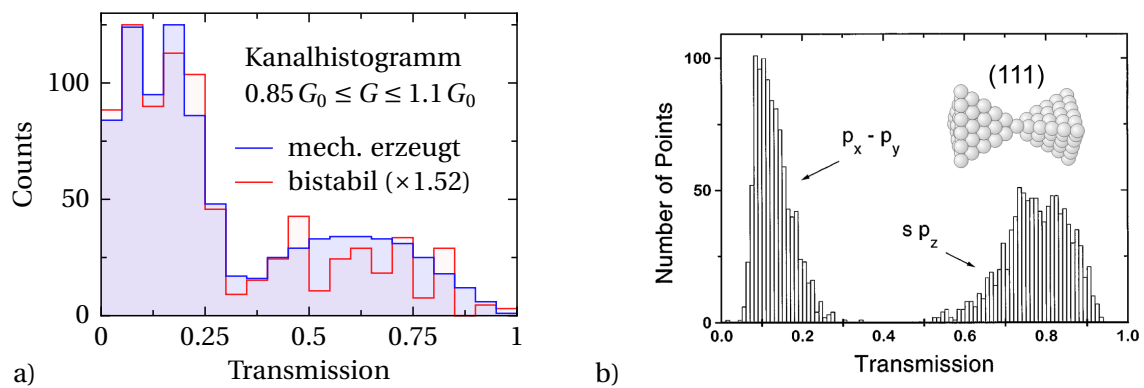


Abbildung 5.28.: a) Kanalhistogramm über 288 bzw. 156 Kontakte mit $0.85 G_0 \leq G \leq 1.1 G_0$. b) Theoretische Rechnung des Histogramms von C. CUEVAS [CYMR98].

erkennbar. Nur der Peak oberhalb von 0.5 scheint bei bistabilen Zuständen in Abb. 5.27 a) abgeschwächt zu sein. Dieser Effekt kann jedoch erklärt werden.

Betrachtet man das Histogramm der Summen in Abb. 5.27 b), erkennt man auch im ersten Aluminium-Peak bei 0.85 eine Abflachung oder sogar ein Minimum bei bistabilen Kontakten⁸. Möglicherweise sind Kontakte um 0.85 so stabil, dass hier sehr selten bistabiles Schalten beobachtet werden kann. Wie im Abschnitt 5.1.3 gezeigt wurde, halten diese Werte besonders hohe Spannungen aus. Der Partnerzustand dieses bistabilen Kontakts müsste diese hohe Spannung jedoch ebenfalls aushalten können, um nicht direkt nach dem Schalten zerstört zu werden. Solche Partnerzustände bei anderen Leitwerten existieren jedoch nicht oder nur selten, so dass nur wenige der bistabilen Paare einen Wert um 0.85 enthalten. Der abgeschwächte 0.85-Peak macht sich natürlich auch im Kanalhistogramm bemerkbar und betrifft offenbar besonders die höheren Kanäle, vermutlich die sp_z -Orbitale. Die quer zur Transportrichtung liegenden keulenförmigen p_x - und p_y -Orbitale, die jeweils etwa 80 % der ankommenden Elektronen zurückreflektieren, stellen sich ankommenden Elektronen eher in den Weg (in bildlicher Sprechweise) und sind leichter schaltbar. So verwundert es nicht, diese etwas häufiger in bistabilen Kontakten zu finden. Kanäle mit hoher Transmission wirken dagegen eher als „Blitzableiter“ und sind stabiler.

Die Zuordnung der beiden Peaks des Kanalhistogramms zu den sp_z , p_x - und p_y -Orbitalen liefert ein Vergleich zur Theorie. Abb. 5.28 b) zeigt eine Rechnung eines Kanalhistogramms mit Zuordnung der Orbitale. In dieser Rechnung wurden pyramidenförmigen (111)-Kontakten angenommen und über eine gewisse Unordnung der atomaren Positionen gemittelt. Die Peaks fallen dabei schärfer aus als im Experiment. Abb. 5.28 a) zeigt im Vergleich ein Kanalhistogramm von rein mechanisch erzeugten Kontakten, das (eingeschränkt auf $0.85 G_0 \leq G \leq 1.1 G_0$) eine gute Übereinstimmung aufweist. In neueren Rechnungen wurden in Kontakten dieser Art Gesamtleitwerte bis $2 G_0$ berechnet [PVH⁺08]. Weitere Rechnungen dieser Methode, durchge-

⁸Die steile Flanke beim Histogramm der mechanisch erzeugten Kontakte ist durch die Vorgabe der Steuerparameter für die Bruchmechanik im Experiment bedingt und ist keine Gesetzmäßigkeit.

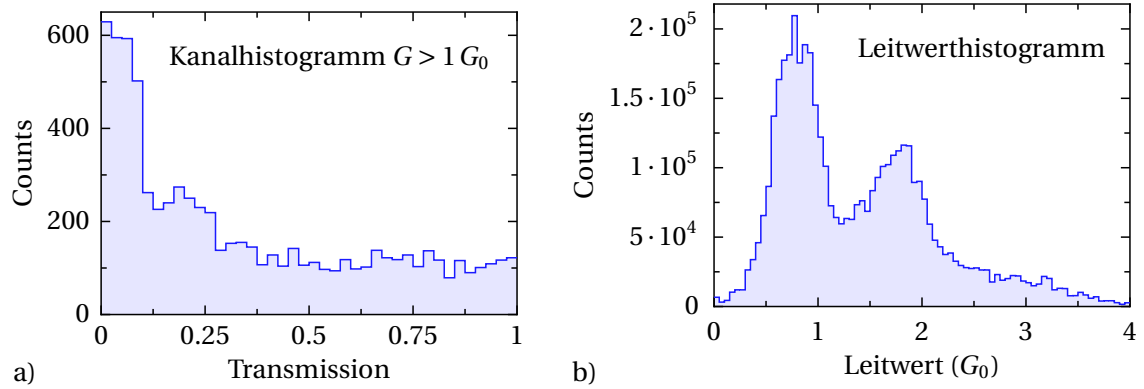


Abbildung 5.29.: a) Kanalhistogramm über 7426 Kanäle aus 1380 Kanalspektren mit Leitwerten $> 1 G_0$.
b) Histogramm über alle bistabilen und nicht bistabilen stromgetriebenen Umordnungsexperimente von Probe cs_b37, zur Überprüfung der Reinheit.

führt von M. HÄFNER und in [Sch08] veröffentlicht, zeigen, wie sich die Kanäle durch Winkel oder Abstände zwischen dem Dimerkontakt stark verändern können. Demnach zeigt diese Häufung keine idealen, sondern hauptsächlich verspannte Kontakte.

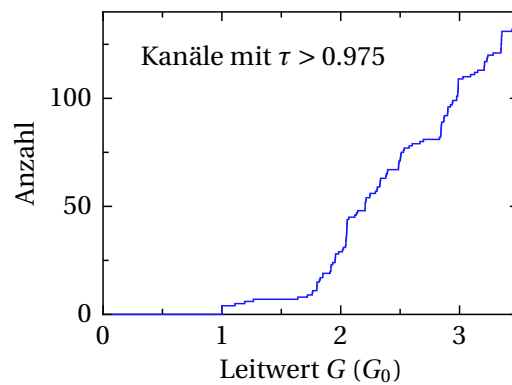


Abbildung 5.30.: Anzahl der Kanäle mit Transmission nahe $1 G_0$, aufsummiert über alle Kontakte deren Leitwert G nicht übersteigt, aufgetragen gegen G . Untersucht wurden hier etwa 1000 bistabile und 1000 mechanisch erzeugte Kontakte.

Die Auswertung aller Kanalspektren mit Summen > 1 (Abb. 5.29 a)) liefert eine sehr ausgeglichene Verteilung hoher Transmissionen, aber immer noch eine Bevorzugung kleiner Werte. Bei den vielen Kanälen mit hohen Leitwerten schwindet jedoch die Aussagekraft der numerischen Auswertung sehr kleiner Kanäle, so dass Höhe und Form der Peaks um 0.05 und 0.2 nicht gesichert sind. Abbildung 5.30 zeigt, ab welchem Gesamtleitwert der in Abb. 5.29 höchste Histogrammbalken anfängt anzusteigen. Ein paar wenige Kontakte bestehen aus einem einzelnen voll transmittierenden Kanal, wie in den Rechnungen von M. HÄFNER in [Sch08] bei Dimer-Kontakten, ohne verspannung vorhergesagt wird. Ansonsten beginnen erst ab einem Gesamtleitwert von etwa $1.7 G_0$ die Zahl der vollen Kanäle linear mit der Anzahl der hinzugekommenen Kontakte anzusteigen.

Abb. 5.29 b) zeigt ein Leitwerthistogramm über alle Umordnungsexperimente. Dieses setzt sich größtenteils aus nicht bistabilen stromgetriebenen Leitwertänderungen zusammen, bei jedoch wechselnder Stromrampe nach jeder Umordnung (d. h. die fehlgeschlagenen Versuche, stabile Kontakte zu erzeugen). Beim Experiment wurde ein Intervall von 0 bis maximal $3G_0$ zugelassen, was das Histogramm nach oben hin begrenzt. Die beiden Peaks haben eine typische Breite und sitzen an gewohnten Stellen, was die Unversehrtheit gegenüber Verschmutzung der über fast fünf Monate geöffneten Probe bestätigt. Die Temperatur war zwar durchgehend unter 1.5 K, jedoch könnten Fremdatome oder Moleküle beim häufigen Öffnen und wieder Schließen in den atomaren Kontakt wandern. Das könnte das Histogramm, z. B. durch einen zusätzlichen Peak, auffällig verändern.

5.6. Hysterese-Formen

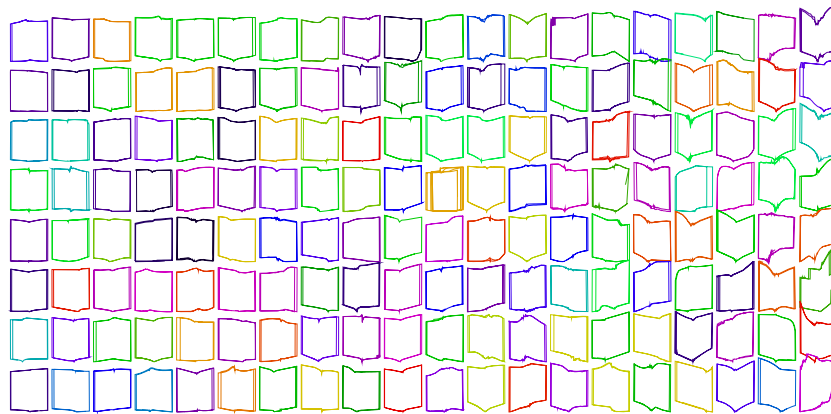


Abbildung 5.31.: Hysteresen bistabiler Kontakte. Zur Normierung wurden die IV-Kennlinien der oberen und unteren Seite jeder Hysterese zueinander vertikal so weit verschoben, dass sich ein Quadrat ergibt. Zur weiteren Normierung wurden die Quadrate in x - und y -Richtung gleichmäßig skaliert. Durch dieses Vorgehen behalten die Steigungen untereinander vergleichbare Einheiten. Die Sortierung folgt aufsteigend nach unten bzw. nach rechts der Standardabweichung.

Ursache der Formen

in Abb. 5.31 sind nochmals die meisten der bistabilen Schalthysteresen gezeigt. Die Hysteresen wurden in der Höhe und Breite normiert und nach der Standardabweichung der Kennlinien sortiert. Auffällig unter den vielfältigen Formen sind viele mit Spitzen am Nulldurchgang. Dies widerspricht der Annahme, dass die Variation des Leitwerts allein mit Leitwertfluktuationen, d. h. Quanteninterferenzen zwischen verschiedenen Elektronenpfaden zur erklären ist. Diese wurden z. B. in [Lud99, LDE⁺99] an Gold meistens ganz ohne Knick am Nulldurchgang beobachtet (siehe z. B. Abb. 2.13, S. 24). Weitere Effekte (vgl. auch Abschnitt 2.4, S. 23) wie

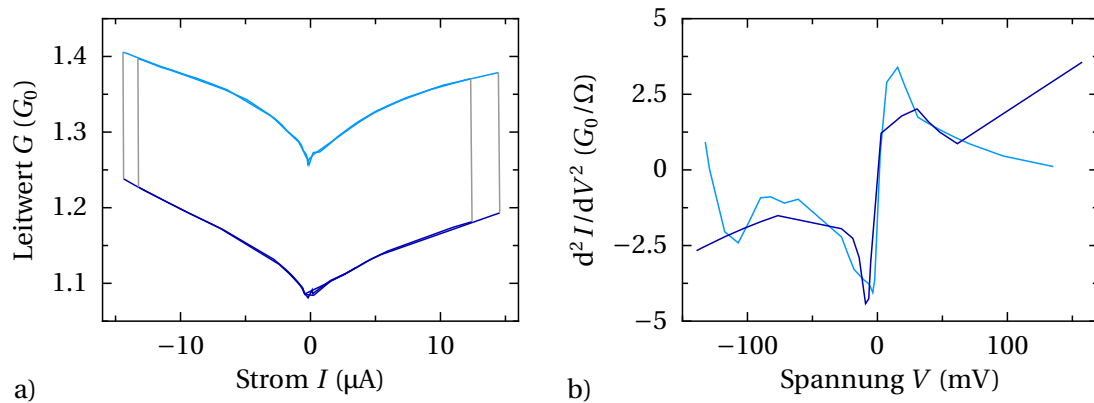


Abbildung 5.32.: Hysterese mit sehr deutlichem Knick. a) Absoluter Leitwert gegen den Strom. b) Zweifach differenzieller Leitwert gegen die Spannung (numerische Ableitung).

Vibrationsanregungen lassen ebenfalls keinen Knick bei Null erwarten. In Frage kommen nur Elektromigrationskräfte, die einzelne Teilchen kontinuierlich aus der Ruhelage auslenken, bis der Rand der Potenzialmulde überschritten wird und der Sprung stattfindet. Diese Auslenkung kann, je nach Situation, nur vom Betrag des Stroms abhängen und unabhängig sein von der Richtung, so dass eine spitze Form am Nulldurchgang erklärbar ist. Verschiedene denkbare Zero-Bias-Effekte könnten zwar auch eine Spitze sichtbar werden lassen, jedoch nur in ein und derselben Richtung. Die Spitzen kommen aber in beiden Richtungen vor, was für die Elektromigrationskräfte spricht.

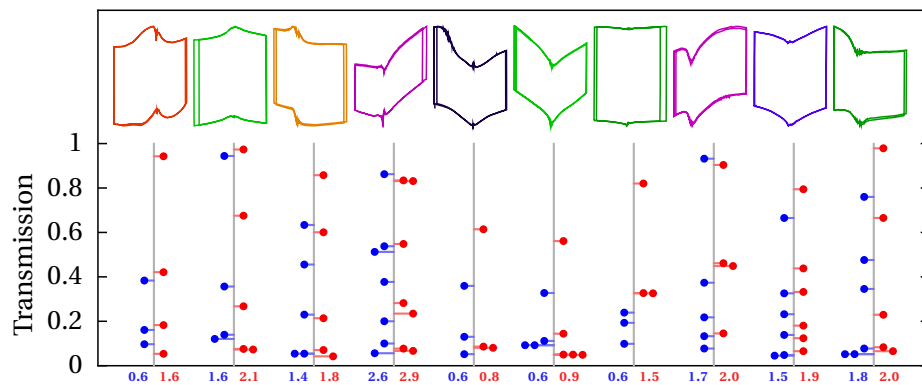


Abbildung 5.33.: Hysteresen mit starken Ähnlichkeiten zwischen oberer und unterer Kennlinie mit den zugehörigen Kanälen. Die auf Einheitsgröße normierten Hysteresen liegen bei unterschiedlichen absoluten Leitwerten, wie die unten angegebenen Zahlen in G_0 erkennen lassen.

Ähnlichkeiten der Form

Die Standardabweichung der Hysteresen in Abb. 5.31 ist meist oben und unten ähnlich. Zu beobachten sind auch einige Hysteresen, bei denen sich sogar die Form der oberen und unteren Seite ähnelt. Diese Übereinstimmung trotz unterschiedlichen Leitwerts zeigt, dass die Ursache der Form nicht lokal an einzelnen Atomen zu suchen ist, sondern durch Eigenschaften der entfernteren Umgebung bestimmt wird. Da für jede der Hysteresen das Kanalspektrum gemessen wurde, können diese Fälle genauer untersucht werden. Abbildung 5.33 zeigt eine solche Auswahl sich ähnelnder Leitwertänderungen. Unter den Hysteresen sind die ermittelten Kanäle eingezeichnet für das jeweils obere und untere Niveau. Kanalspektren mit ähnlichen Kanälen wurden nach rechts sortiert. Bei den linken beiden scheint ein zusätzlicher Kanal im oberen Zustand hinzuzukommen, während die anderen Kanäle sich weiterhin ähneln. In den nächsten paar Fällen verschwindet je ein Kanal und ein anders liegender kommt hinzu. In weiteren Fällen ist keine Ähnlichkeit zu erkennen. Die Fälle mit Ähnlichkeit unterscheiden sich jedoch nicht signifikant von solchen mit zufälliger Ähnlichkeit. Genauso oft lassen sich auch Ähnlichkeiten von einer zu einer anderen Hysterese finden.

Fazit

Durch diese Ergebnisse und weitere Betrachtungen anderer Hysteresen kann festgestellt werden, dass die minimalen lokalen atomaren Umordnungen der bistabilen Kontakte die Kanalverteilung stark verändern. In vielen Fällen erscheinen jedoch die Formen der Hysteresen durch weiter entfernte Geometrien bedingt zu sein.

5.6.1. Einzelbeobachtungen

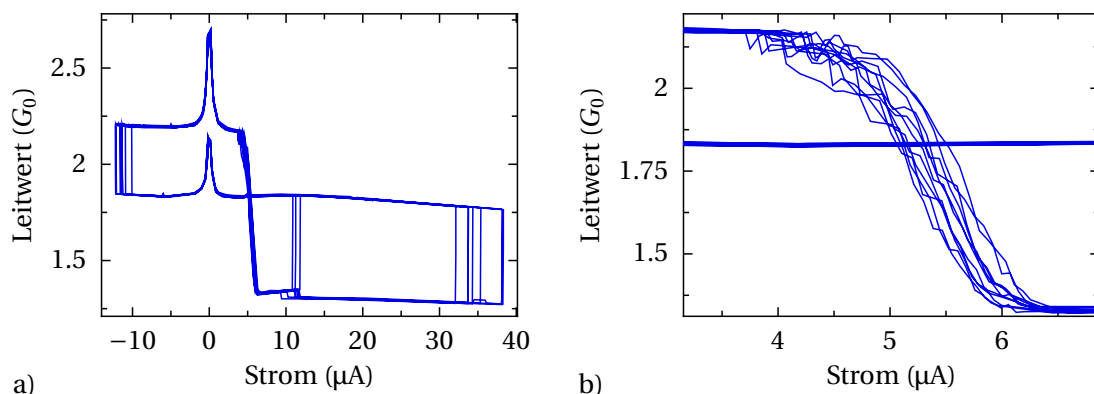


Abbildung 5.34.: Hysterese mit scheinbar steilem kontinuierlichem Übergang, der nur den Mittelwert einer schnellen zeitlichen Zwei-Niveau-Fluktuation zeigt.

Im Folgenden werden einige ausgesuchte Hysteresen betrachtet, die gewisse Besonderheiten aufweisen. Abb. 5.34 zeigt eine Hysterese mit einem scheinbar kontinuierlichen Übergang. Im Zoom b) erkennt man an dieser Stelle ein erhöhtes Rauschen. Interpretiert wird dies als ein

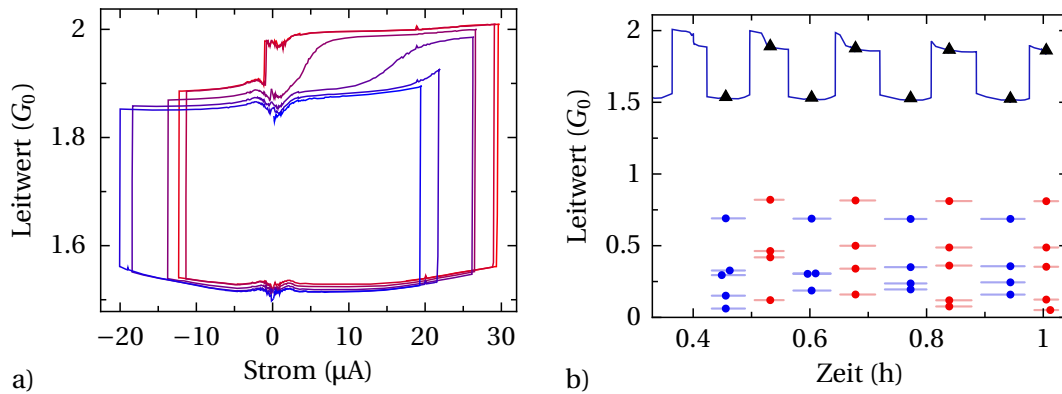


Abbildung 5.35.: a) Ungeklärter Einzelfall einer langsamen stetigen Änderung des Kontaktes. Die Einfärbung von Rot nach Blau gibt den zeitlichen Verlauf wieder. b) Zeitlich aufgetragenes Schaltverhalten mit den gemessenen Kanalspektren.

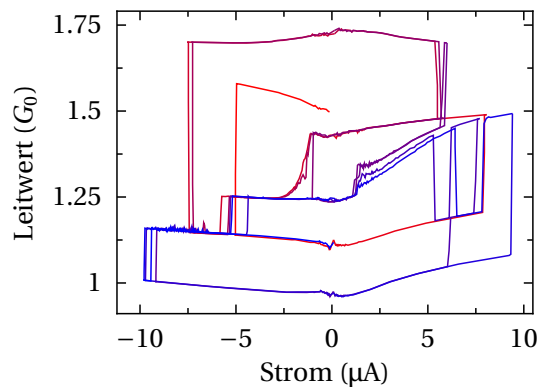


Abbildung 5.36.: Beispiel für die eingeschränkten Freiheitsgrade der Umordnungen. Die zeitliche Reihenfolge wird durch die Farben Rot bis Blau gezeigt. Zu Beginn und ganz am Ende scheint der Kontakt in der selben Konfiguration zu sein.

zeitlich fluktuierendes Zwei-Niveau-System. Mit kontinuierlicher Änderung des Stroms verlagert sich die Aufenthaltswahrscheinlichkeit zunehmend von der einen zur anderen Position. Jeder über 500 ms gemittelte Messwert ergibt daher einen scheinbar dazwischen liegenden Leitwert, der langsam vom einen zum anderen Leitwert übergeht. Derartige Beobachtungen wurden mindestens zweimal gemacht.

Einen bisher nicht geklärten Einzelfall zeigt Abb. 5.35 a). Hier tritt eine schrittweise stetige Änderung der Hysterese auf, die sich deutlich in der oberen Form der Leitwertvariation äußert. Die Farbgebung entspricht der zeitlichen Reihenfolge (von rot nach blau). Da sich die Messung über einen Zeitraum von ca. 40 Minuten abspielt, ist eine Relaxation des Metalls oder Substrats unwahrscheinlich. Die Temperatur am ^3He -Thermometer sinkt in dieser Zeit von 0.3 K auf 0.249 K. Möglicherweise ist dieser Temperaturverlauf an der Probe sehr viel stärker und beeinflusst ebenfalls eine zeitliche Fluktuation zwischen zwei Positionen, die hier sehr schnell ist, und kein verstärktes Rauschen zeigt. Die Kanalspektren in b) sind zwar bei den kleinen

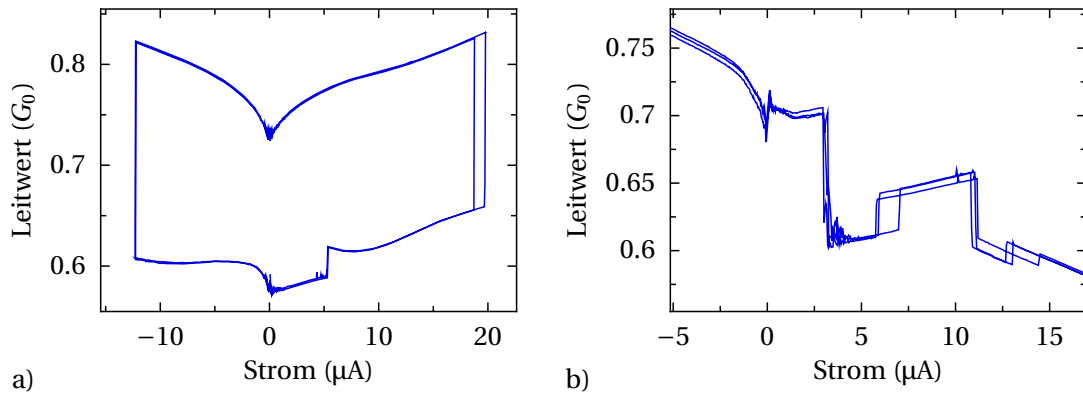


Abbildung 5.37.: Kleine Umordnungen, die in der selben Reihenfolge mehrfach durchlaufen werden.

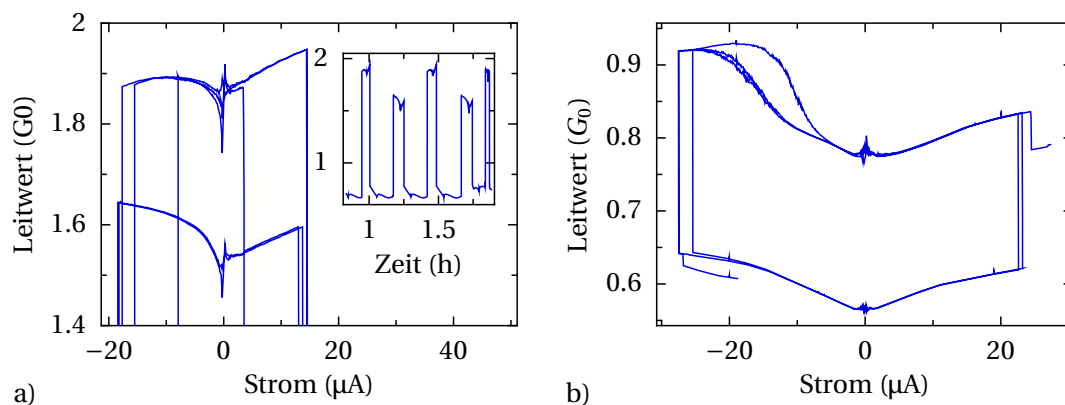


Abbildung 5.38.: Zufällige kleine Änderungen, die zu sehr ähnlichen Verläufen des Leitwerts mit dem Strom führen.

Kanälen wenig reproduzierbar, es zeigt sich jedoch auch keine Tendenz einer systematischen Veränderung.

Abb. 5.36 zeigt, dass die Geometrieänderungen auf kleinem Raum konzentriert sind. An der zeitlichen Reihenfolge (von rot nach blau) sieht man, dass viele geometrische Zustände immer wieder eingenommen werden. Insbesondere wird das zweitunterste Niveau weit am Anfang (rot) und weit am Ende (blau) der Messung eingenommen. Diese Einschränkung auf immer wiederkehrende Zustände kann nur bedeuten, dass ein räumlich stark eingeschränkter Bereich für die Änderungen zuständig ist, in dem nur wenige mögliche Geometrien eingenommen werden können.

Kleinere reproduzierbare Mehrfach-Umordnungen erkennt man auch in den beiden Hysteresen in Abb. 5.37. In beiden Fällen zeigt die jeweils untere Seite der Hysterese zusätzliche Leitwertsprünge bei steigendem Strom, was durch eine Umordnung mit sehr vielen beteiligten Atomen schwer erklärbar wäre. In beiden Beispielen kann zwischen mehr als zwei atomaren Geometrien reproduzierbar umgeschaltet werden. Solche Situationen können durch spezielle Experimente gezielt gesucht werden, indem man die Stromrampe z. B. erst nach der zweiten

Leitwertänderung umkehrt. In diesem Fall war die Leitwertänderung zu klein, um von der Experimentsteuerung als Sprung detektiert zu werden.

In Abb. 5.38 bewirken die lokal kleinen (eigentlich ungewollten) Änderungen der Leitwertvariation, die den groben Verlauf unbeeinflusst lassen. Dies zeigt erneut, dass solche Formen nicht allein durch die lokale Geometrie bestimmt sind oder chaotisches Verhalten dominiert.

Dass kleine lokale Änderungen der Geometrie auch nur kleine Änderungen der Form der Leitwertvariation verursachen, zeigen die Beispiele in Abb. 5.38. Zwar wurde bereits in [Lud99] gezeigt, wie sich die Leitwertfluktuationen an Gold kontinuierlich mit dem Öffnen des Kontakts verändern, dabei ist jedoch eine komplette Hälfte aller Atome des Kontakts gegenüber der anderen Hälfte in Bewegung. Auch wenn die mechanischen Änderungen bei dieser Technik beliebig klein gewählt werden können, könnten Interferenzeffekte durch das Auseinanderziehen des Kontaktes eventuell eine deutliche Änderung bewirken (wie z. B. in einem Fabry-Perot-Interferometer). Die lokalen Änderung durch bistabile Umordnungen sind klein in einem Sinn, der die räumliche Ausdehnung des geänderten Bereichs betrifft.

5.7. Übersicht über alle Messungen

In der folgenden Tabelle 5.2 sind alle wichtigen in dieser Arbeit durchgeführten Messungen aufgelistet und den abgebildeten Ergebnissen zugeordnet.

Zeitraum	Probe	Art	Anzahl Messungen und Referenz zum Ergebnis
16.03.2003	cs_61	Al	einzelne Versuche zu stromgetriebenen Umordnungen → Abb. 5.1
12.04.2006–16.04.2006	cs_a35	Au	46 stromgetriebene Umordnungen → Abb. 5.2
22.04.2006–30.04.2006	cs_a38	Au	123 stromgetriebene Umordnungen → Abb. 5.2
24.06.2006–03.07.2006	cs_b2	Al	479 stromgetriebene Umordn. → Abb. 5.3a, 5.4, 5.5, 5.7
19.03.2007–10.04.2007	cs_b14	Al	906 stromgetriebene Umordn. → Abb. 5.3a, 5.4, 5.5, 5.7
15.03.2007–23.03.2007	cs_b14	Al	90 Versuche für bistabiles Schalten → Abb. 5.11
05.12.2007–14.12.2007	cs_b33	Al	67 stromgetriebene Umordnungen mit einer MAR-Messung pro Umordnung → bisher unausgewertet
14.12.2007–18.12.2007	cs_b33	Al	70 Versuche für bistabiles Schalten mit MAR-Messungen → Abb. 5.16
01.02.2008–23.04.2008	cs_b37	Al	2449 Versuche für bistabiles Schalten, davon 190 bistabil mit 1583 MAR-Messungen → alle Abbildungen von 5.21 bis 5.37 (außer 5.28)
08.02.2008–03.06.2008	cs_b37	Al	662 Kontakte um $0.85 G_0$ mit MAR-Messungen und stromgetriebener Umordnung → Abb. 5.27, 5.5, 5.28
04.06.2008–13.06.2008	cs_b37	Al	440 MAR-Messungen an mechanisch erzeugten Kontakten → Abb. 5.27, 5.28, 5.29

Tabelle 5.2.: Übersicht aller wichtigen Messungen. In der Nomenklatur der Proben bedeutet a = Apical®-Substrat (Polyimid), b = Bronze-Substrat. „Stromgetriebene Umordnung“ bezeichnet hier Experimente mit einer Steigerung des Stroms in nur eine Stromrichtung bis zur Unterbrechung des Kontakts.

*» Ich male die Dinge,
wie ich sie denke,
nicht wie ich sie sehe. «*

— Pablo Picasso

A ANHANG A.

Grafische Veranschaulichung

A.1. Simulation zur Veranschaulichung eines Bruchkontakts

Die Simulation, die im Folgenden gezeigt wird, verfolgt das Ziel, eine realistisch wirkende grafische Veranschaulichung des Bruchkontakts als Kugelmodell mit einem Raytracing-Programm zu erzeugen. Dabei geht es nicht um ein physikalisch korrektes Modell, sondern nur um ein Bild für didaktische Zwecke. Das manuelle Positionieren der Kugeln im Computer, was zunächst naheliegend erschien, entpuppte sich als viel zu zeitaufwendig und aus der theoretischen Physik war nur Bildmaterial mit sehr wenigen Atomen zu bekommen. Diese Simulationen generiert dagegen Bilder mit über einer Million Teilchen. Nach dem Einbau einiger heuristischer Terme und Parameter konnten in der Simulation sogar einige Effekte beobachtet werden, die nicht explizit eingebaut wurden, aber dennoch plausibel erscheinen. Zunächst folgt ein kurzer Überblick der Funktionsweise.

Funktionsweise

Als Potenzial der Atome wird das LENNARD-JONES-Potenzial

$$V \sim \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \quad (\text{A.1})$$

verwendet. Die Dynamik des Systems wird jedoch nicht durch eine Bewegungsgleichung zweiter, sondern erster Ordnung beschrieben. Das heißt, die Teilchen erfahren keine Beschleunigung durch das Potenzial, sie erhalten stattdessen direkt die Geschwindigkeit aus dem Gradient des Potenzials. Dazu wird die Ableitung

$$V' \equiv \frac{\partial V}{\partial r} \sim - \left(2 \left(\frac{\sigma}{r}\right)^{13} - \left(\frac{\sigma}{r}\right)^7 \right) \quad (\text{A.2})$$

vom Potenzial V gebildet und dieser Wert für jedes Nachbaratom pro Iterationsschritt ausgewertet und zur Koordinate addiert. Als Grundgleichung der Simulation ergibt sich mit

A. GRAFISCHE VERANSCHAULICHUNG

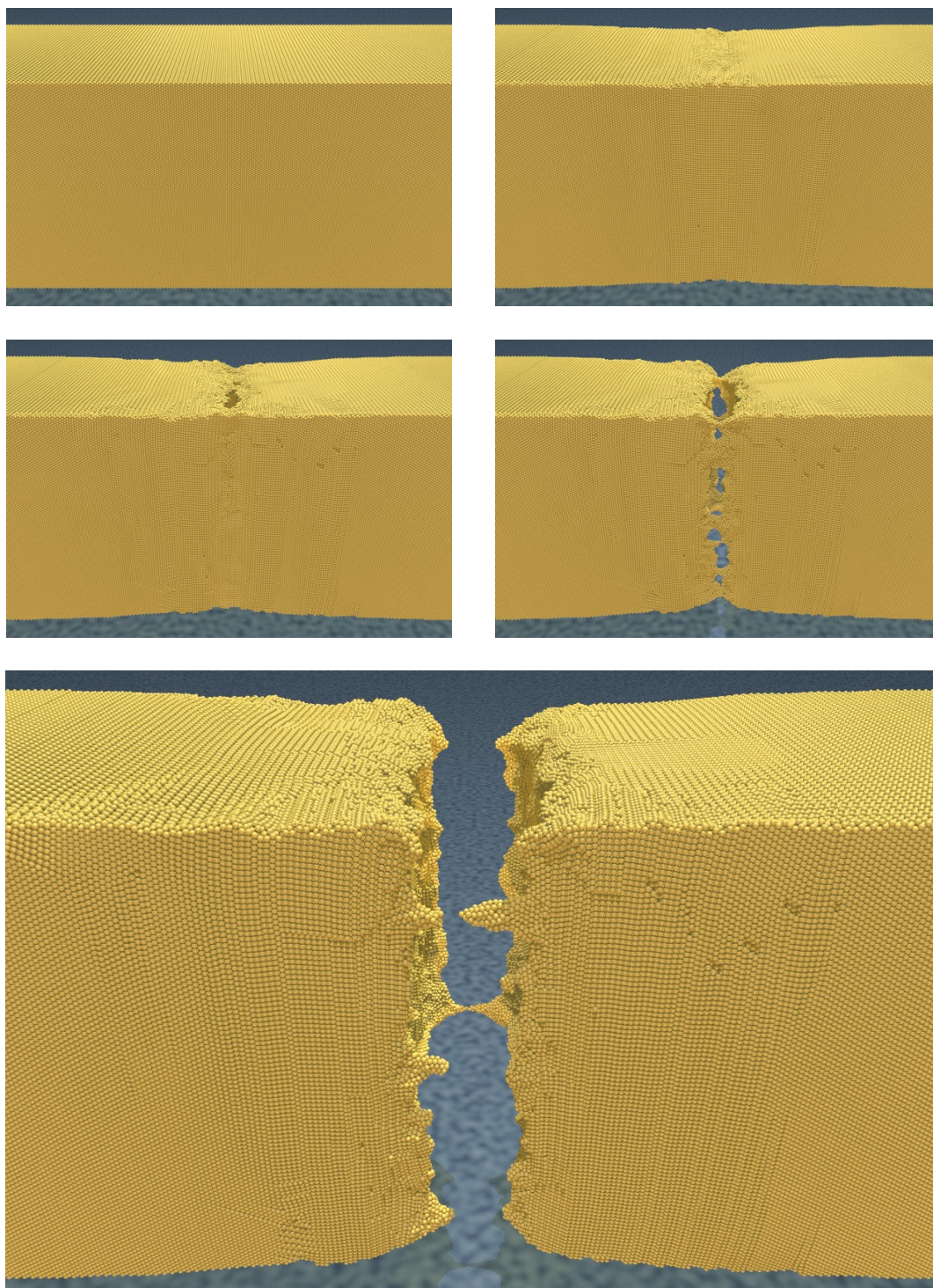


Abbildung A.1.: Simulation des Öffnungsvorganges eines Bruchkontakts mit einem $100 \times 100 \times 200$ Atome umfassenden Draht, wobei 1.2 Millionen Atome echt beweglich sind.

den Abständen zu den Nachbaratomen $\mathbf{d}_{ij} = (\mathbf{x}_j - \mathbf{x}_i)$ und $d_{ij} = |\mathbf{d}_{ij}|$ für das i -te Atom die Iterationsformel

$$\mathbf{x}_i = \mathbf{x}_i + \sum_j \frac{\mathbf{d}_{ij}}{d_{ij}} V'(d_{ij}). \quad (\text{A.3})$$

Zur Steigerung der Rechengeschwindigkeit wird die Wechselwirkung nicht mit jedem Atom berechnet. Stattdessen erhält jedes Atom eine begrenzte Zahl von nächsten Nachbarn (hier 14) mit denen eine Wechselwirkung stattfindet. Da sich mit der Zeit die Atome durchmischen, wird in regelmäßigen Abständen aus der Menge der Nachbarn, sowie der Nachbarn der Nachbarn, die 14 Atome mit der kleinsten Distanz als neue Nachbarn gespeichert. Auf diese Weise skaliert die Rechenzeit nur mit $\sim O(N)$, also linear mit der Teilchenzahl N . Eine weitere Beschleunigung wurde erreicht, indem $V'(d_{ij})/d_{ij}$ in einem Initialisierungsschritt für 5000 Werte berechnet und in einem Array gespeichert wird. Um die Zahl der Iterationsschritte klein halten zu können, wird die Gleichung A.3 sofort, und nicht erst am Ende des Iterationsschrittes, auf die Koordinate angewendet. Dadurch können die nachfolgenden Teilchen mit noch aktuelleren Positionen der Nachbarn rechnen. Auf diese Weise kann bereits innerhalb eines Iterationsschrittes ein gewisser Impuls weit durch das System getragen werden. Die Reihenfolge der Teilchen im Array muss dafür passend zu der anfänglichen Kristallordnung gewählt sein und jeder zweite Iterationsschritt muss in umgekehrter Reihenfolge ausgeführt werden. Durch diese Strategie wird besonders die spontane Umkristallisation am Anfang auch bei wenigen Iterationsschritten ermöglicht. Insgesamt kann mit diesem Algorithmus eine Simulationen mit über einer Million Teilchen in einer Nacht auf einem gängigen Bürocomputer berechnet werden.

Neben der Grundgleichung wurden ein paar heuristische Zusatzterme eingebaut, wie eine Verlangsamung von Teilchen, die sich ungefähr im Bindungsabstand zu Nachbaratomen befinden, oder ein Parameter für eine zusätzliche Oberflächenspannung. Die Ziehbewegung wird von einigen nicht beweglichen Teilchen am linken und rechten konkav nach innen gewölbten Rand erzwungen. Außerdem können die beweglichen Atome entweder der linken oder der rechten Seite zugehören und bewegen sich entsprechend durch einen Zusatzterm mit derselben Geschwindigkeit wie der zugehörige Rand. Die Zugehörigkeit ist dabei ein stufenloser Parameter zwischen -1 und 1 , der sich mit der Zeit ändern kann, abhängig von der Zugehörigkeit der Atome in der Nachbarschaft. Auf diese Weise kann ein Verbund von Atomen an einer der beiden gebrochenen Hälften haften bleiben und der konstanten Bewegung folgen, ähnlich wie bei einer Bewegungsgleichung zweiter Ordnung. Temperatureffekte enthält das System keine, aufgrund der Bewegungsgleichung erster Ordnung.

Ergebnis

Abbildung A.1 zeigt das Ergebnis einer Simulation eines $100 \times 100 \times 200$ Atome umfassenden Drahtes mit 1.2 Millionen echt beweglichen Atomen, dargestellt mit einem Raytracing-Programm¹. Da es sich um ein unspezifisches Kugelmodell handelt, das nicht durch quantenmechanische Vorgänge motiviert ist, ist dadurch kein bestimmtes Material beschrieben. Dennoch unterliegt es erkennbar einigen grundlegenden Materialeigenschaften auf größerer Skala, die gemeinsam für verschiedene Systeme aus Atomen, Kolloiden, Schaum oder

¹Hier wurde das Freeware-Programm Povray verwendet [pov].

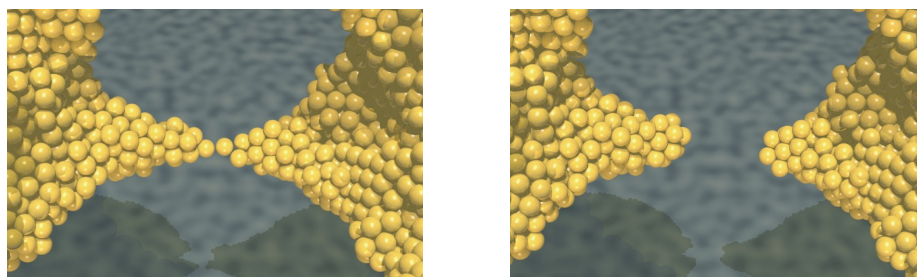


Abbildung A.2.: Atomare Spitze die sich selbst gebildet hat. Der Einfluss der vorgegebenen Oberflächenspannung trägt zu einer spitzeren Form bei, bewirkt aber nach dem Zerreißen wiederum eine leichte Verrundung.

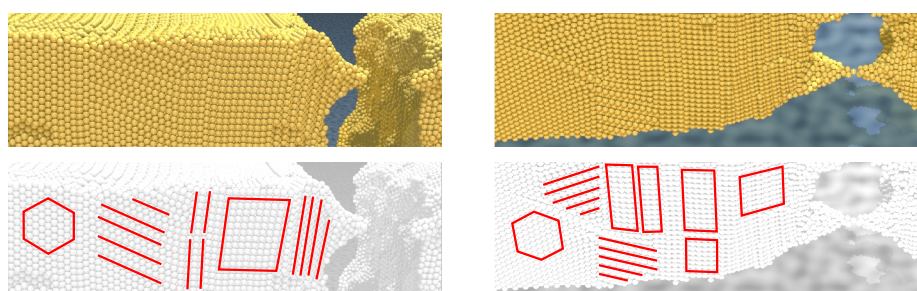


Abbildung A.3.: Korngrenzen zwischen unterschiedlichen Kristallregionen, die durch die Zugspannung lokal zur Umkristallisation angeregt wurden.

ähnlichen adhäsiven Objekten gelten.

Nach einer anfänglichen leichten Verjüngung fangen im Beispiel einige Bereiche an, unter Zugspannung ihre Kristallorientierung zu ändern. An einer Stelle löst sich die Kristallstruktur lokal auf. Dort bilden sich erste Hohlräume und bald ein Spalt. Bei der hier mittels Parameter eingestellten Oberflächenspannung (z. B. in Anlehnung an Gold) bilden sich zum Teil fadenartige Brücken, die sich kurz vor dem Abreißen zu Spitzen verformen. Die letzte Berührung zwischen den beiden Hälften bildet im gezeigten Beispiel ein Dimer-Kontakt.

Abbildung A.2 zeigt eine solche Spitze, die sich von selbst gebildet hat. Die Oberflächenspannung, die bei der Formierung hilft, lässt die beiden Spitzen nach dem Zerreißen wieder etwas verrunden. Spitzen wie diese finden sich auch in den physikalisch korrekten atomaren Simulationen der theoretischen Physik, wie in [Dre08, DPH⁺05].

In Abb. A.3 sind die nach dem Brechen entstandenen Kristallbereiche zwischen den Korngrenzen durch Linien hervorgehoben. Die Kristalle können dabei sehr unterschiedliche Winkel zu einander annehmen. Die ursprünglich gewählte Kristallrichtung, auf der Sechseckformen zu sehen sind, scheint unter der Zugspannung am wenigsten stabil zu sein und ist um den Bruch herum nie zu finden. Auch diese Effekte sind in Wirklichkeit denkbar.

Insgesamt erzeugt die Simulation trotz ihrer sehr einfachen Annahmen realistisch wirkende Strukturen. Sie erfüllt daher ihren Zweck, anschauliches Bildmaterial zur Anregung der Vorstellungskraft und der Fantasie zu liefern, während die vielfältigen Vorgänge in der „Black Box“ des Experiments nur über eine einzige skalare Größe, den Leitwert, „sichtbar“ sind.

Literaturverzeichnis

- [And64] ANDREEV, A. F.: The Thermal Conductivity of the Intermediate State in Superconductors. In: *Soviet physics JETP-USSR* 19 (1964), Nr. 5, S. 1228–1231 [17](#)
- [AYR03] AGRAÏT, N. ; YEYATI, Alfredo L. ; RUITENBEEK, Jan M.: Quantum properties of atomic-sized conductors. In: *Physics Reports* 377 (2003), April, Nr. 2-3, S. 81–279. [http://dx.doi.org/10.1016/S0370-1573\(02\)00633-6](http://dx.doi.org/10.1016/S0370-1573(02)00633-6). – DOI 10.1016/S0370-1573(02)00633-6. – ISSN 03701573 [20](#)
- [Bü86] BÜTTIKER, M.: Four-Terminal Phase-Coherent Conductance. In: *Phys. Rev. Lett.* 57 (1986), Oct, Nr. 14, S. 1761–1764. <http://dx.doi.org/10.1103/PhysRevLett.57.1761>. – DOI 10.1103/PhysRevLett.57.1761 [12](#)
- [Bö08] BÖHLER, Tobias: *Elektronischer Transport durch einzelne C60-Moleküle*. Universitätsstr. 10, 78457 Konstanz, Universität Konstanz, Diss., 2008. <http://kops.ub.uni-konstanz.de/volltexte/2008/5829> [24](#)
- [Bas] BASS, J.: 1.5 Bloch-Grüneisen resistivity function. http://dx.doi.org/10.1007/10307022_30. In: *Landolt-Börnstein - Group III Condensed Matter* Bd. 15a. Hellwege, K.-H., Olsen, J. L. (ed.). SpringerMaterials - The Landolt-Börnstein Database. – DOI 10.1007/10307022_30 [84](#)
- [Bat02] BATRA, Inder P.: From uncertainty to certainty in quantum conductance of nanowires. *124* (2002), Nr. 12, S. 463–467. [http://dx.doi.org/10.1016/S0038-1098\(02\)00550-1](http://dx.doi.org/10.1016/S0038-1098(02)00550-1). – DOI 10.1016/S0038-1098(02)00550-1 [13](#), [14](#)
- [BES09] BÖHLER, T. ; EDTBAUER, A. ; SCHEER, E.: Point-contact spectroscopy on aluminium atomic-size contacts: longitudinal and transverse vibronic excitations. In: *New Journal of Physics* 11 (2009), Nr. 013036, S. 013036. <http://dx.doi.org/10.1088/1367-2630/11/1/013036>. – DOI 10.1088/1367-2630/11/1/013036 [24](#)
- [BGH⁺03] BLADH, K. ; GUNNARSSON, D. ; HÜRFELD, E. ; DEVI, S. ; KRISTOFFERSSON, C. ; LANDER, B. S. ; PEHRSON, S. ; CLAESON, T. ; DELSING, P. ; TASLAKOV, M.: Comparison of cryogenic filters for use in single electronics experiments. In: *Review of Scientific Instruments* 74 (2003), Nr. 3, S. 1323–1327. <http://dx.doi.org/10.1063/1.1540721>. – DOI 10.1063/1.1540721 [61](#)

- [Blo29] BLOCH, Felix: Über die Quantenmechanik der Elektronen in Kristallgittern. In: *Zeitschrift für Physik A Hadrons and Nuclei* 52 (1929), Nr. 7-8, S. 555–600. <http://dx.doi.org/10.1007/BF01339455>. – DOI 10.1007/BF01339455 11
- [CMRY96] CUEVAS ; MARTÍN-RODERO ; YEYATI: Hamiltonian approach to the transport properties of superconducting quantum point contacts. In: *Phys Rev B Condens Matter* 54 (1996), Sep, Nr. 10, S. 7366–7379. <http://dx.doi.org/10.1103/PhysRevB.54.7366>. – DOI 10.1103/PhysRevB.54.7366 19
- [COL⁺03] CHEN, Yong ; OHLBERG, Douglas A. A. ; LI, Xuema ; STEWART, Duncan R. ; WILLIAMS, R. S. ; JEPPESEN, Jan O. ; NIELSEN, Kent A. ; STODDART, J. F. ; OLYNICK, Deirdre L. ; ANDERSON, Erik: Nanoscale molecular-switch devices fabricated by imprint lithography. In: *Applied Physics Letters* 82 (2003), Nr. 10, S. 1610–1612. <http://dx.doi.org/10.1063/1.1559439>. – DOI 10.1063/1.1559439 91
- [Cue99] CUEVAS, Juan C.: *Electronic Transport in Normal and Superconducting Nanocontacts*, Universidad Autónoma de Madrid, Diss., 1999 16
- [CYMR98] CUEVAS, J. C. ; YEYATI, A. L. ; MARTÍN-RODERO, A.: Microscopic Origin of Conducting Channels in Metallic Atomic-Size Contacts. In: *Phys. Rev. Lett.* 80 (1998), Feb, Nr. 5, S. 1066–1069. <http://dx.doi.org/10.1103/PhysRevLett.80.1066>. – DOI 10.1103/PhysRevLett.80.1066 102
- [Dat97] DATTA, Supriyo: *Electronic Transport in Mesoscopic Systems*. Cambridge : Cambridge University Press, 1997. – ISBN 9780521599436 14
- [DG27] DAVISSON, C. ; GERMER, L. H.: The Scattering of Electrons by a Single Crystal of Nickel. In: *Nature* 119 (1927), S. 558–560. <http://dx.doi.org/10.1038/119558a0>. – DOI 10.1038/119558a0 11
- [Die95] DIERCKX, Paul: *Curve and Surface Fitting with Splines*. (1995). ISBN 9780198534402 42
- [DPH⁺05] DREHER, Markus ; PAULY, Fabian ; HEURICH, Jan ; CUEVAS, Juan C. ; SCHEER, Elke ; NIELABA, Peter: Structure and conductance histogram of atomic-sized Au contacts. In: *First publ. in: Physical Review B* 72 (2005), 7, Article 075435 (2005). <http://dx.doi.org/10.1103/PhysRevB.72.075435>. – DOI 10.1103/PhysRevB.72.075435 22, 114
- [Dre08] DREHER, Markus: *Untersuchung elektronischer Eigenschaften komplexer Materialien mittels Computer-Simulationen*. Universitätsstr. 10, 78457 Konstanz, Universität Konstanz, Diss., 2008. <http://kops.ub.uni-konstanz.de/volltexte/2008/6094> 22, 114
- [Dru00] DRUDE, Paul: Zur Elektronentheorie der Metalle. In: *Annalen der Physik* 306 (1900), Nr. 3, S. 566 – 613. <http://dx.doi.org/10.1002/andp.19003060312>. – DOI 10.1002/andp.19003060312 11

- [DSDV06] D'AGOSTA, Roberto ; SAI, Na ; DI VENTRA, Massimiliano: Local Electron Heating in Nanoscale Conductors. In: *Nano Letters* 6 (2006), dec, Nr. 12, S. 2935–2938. <http://dx.doi.org/10.1021/nl062316w>. – DOI 10.1021/nl062316w. – ISSN 1530–6984 81, 85
- [Eil03] EILERS, Paul H C.: A perfect smoother. In: *Anal Chem* 75 (2003), Jul, Nr. 14, S. 3631–3636. <http://dx.doi.org/10.1021/ac034173t>. – DOI 10.1021/ac034173t 42
- [ES05] EINSTEIN, A. ; SEELIG, C.: *Mein Weltbild*. Bd. 30. Ullstein, 2005 73
- [Fal97] FALCONER, Isobel: J J Thomson and the discovery of the electron. In: *Physics Education* 32 (1997), Nr. 4, S. 226–231. <http://dx.doi.org/10.1088/0031-9120/32/4/015>. – DOI 10.1088/0031–9120/32/4/015 11
- [GET98] GÜLSEREN, Oguz ; ERCOLESSI, Furio ; TOSATTI, Erio: Noncrystalline Structures of Ultrathin Unsupported Nanowires. In: *Phys. Rev. Lett.* 80 (1998), Apr, Nr. 17, S. 3775–3778. <http://dx.doi.org/10.1103/PhysRevLett.80.3775>. – DOI 10.1103/PhysRevLett.80.3775 21
- [GML09] GOMEZ-MARLASCA, F. ; LEVY, P.: Resistance switching in silver – manganite contacts. In: *Journal of Physics: Conference Series* 167 (2009), 012036 (6pp). <http://dx.doi.org/10.1088/1742-6596/167/1/012036>. – DOI 10.1088/1742–6596/167/1/012036 91
- [Gro69] GROSSMANN, Walter: *Grundzüge der Ausgleichsrechnung nach der Methode der kleinsten Quadrate nebst Anwendung in der Geodäsie*. 3., erw. A. Springer-Verlag GmbH, 1969. – ISBN 9783540044956 46
- [HP97] HODRICK, R. J. ; PRESCOTT, E.C.: Postwar US Business Cycles: An Empirical Investigation. In: *Journal of Money, Credit & Banking* 29 (1997), Feb, Nr. 1, S. 1–16 42
- [HWHS08] HOFFMANN, R. ; WEISSENBERGER, D. ; HAWECKER, J. ; STÖFFLER, D.: Conductance of gold nanojunctions thinned by electromigration. In: *Applied Physics Letters* 93 (2008), Nr. 4, S. 043118. <http://dx.doi.org/10.1063/1.2965121>. – DOI 10.1063/1.2965121 22
- [JWB⁺07] JOOSS, C. ; WU, L. ; BEETZ, T. ; KLIE, RF ; BELEGGIA, M. ; SCHOFIELD, MA ; SCHRAMM, S. ; HOFFMANN, J. ; ZHU, Y.: Polaron melting and ordering as key mechanisms for colossal resistance effects in manganites. In: *Proceedings of the National Academy of Sciences* 104 (2007), Nr. 34, S. 13597. <http://dx.doi.org/10.1073/pnas.0702748104>. – DOI 10.1073/pnas.0702748104 91
- [KBS⁺05] KONRAD, Patrick ; BACCA, Cecile ; SCHEER, Elke ; BRENNER, Patrice ; MAYER-GINDNER, Andreas ; LÖHNEYSSEN, Hilbert: Stable single-atom contacts of zinc

- whiskers. In: *First publ. in: Applied Physics Letters* 86 (2005), 213115. <http://dx.doi.org/10.1063/1.1926408>. – DOI 10.1063/1.1926408 58
- [Kit06] KITTEL, Charles: *Einführung in die Festkörperphysik*. 14. Auflage. Oldenbourg Wissenschaftsverlag, 2006. – ISBN 3–486–57723–9. – Daten aus den Tabellen 6.1 und 6.3. 61, 85
- [Kon03] KONRAD, Patrick: *Herstellung und elektronischer Transport von Kontakten atomarer Größe aus Zink*. Universitätsstr. 10, 78457 Konstanz, Universität Konstanz, Diplomarbeit, 2003 58
- [KRF⁺95] KRANS, J. M. ; RUITENBEEK, J. M. ; FISUN, V. V. ; YANSON, I. K. ; JONGH, L. J.: The signature of conductance quantization in metallic point contacts. In: *Nature* 375 (1995), S. 767 – 769. <http://dx.doi.org/10.1038/375767a0>. – DOI 10.1038/375767a0 10, 20, 21, 22
- [Lan57] LANDAUER, Rolf: Spatial variation of currents and fields due to localized scatterers in metallic conduction. In: *IBM Journal of Research and Development* 1 (1957), Nr. 3, 223-231. <http://researchweb.watson.ibm.com/journal/rd/013/ibmrd0103D.pdf> 9, 13
- [LDE⁺99] LUDOPH, B. ; DEVORET, M. H. ; ESTEVE, D. ; URBINA, C. ; RUITENBEEK, J. M.: Evidence for Saturation of Channel Transmission from Conductance Fluctuations in Atomic-Size Point Contacts. In: *Phys. Rev. Lett.* 82 (1999), Feb, Nr. 7, S. 1530–1533. <http://dx.doi.org/10.1103/PhysRevLett.82.1530>. – DOI 10.1103/PhysRevLett.82.1530 24, 104
- [Lid04] LIDE, David: *Crc Handbook of Chemistry and Physics: a Ready Reference Book of Chemical and Physical Data: Special Student Ed.* City : CRC Press Inc, 2004. – ISBN 0849305977 85, 92
- [LKDK08] LI, Fushan ; KIM, Tae W. ; DONG, Wenguo ; KIM, Young-Ho: Formation and electrical bistability properties of ZnO nanoparticles embedded in polyimide nanocomposites sandwiched between two C₆₀ layers. In: *Applied Physics Letters* 92 (2008), Nr. 1, S. 011906. <http://dx.doi.org/10.1063/1.2830617>. – DOI 10.1063/1.2830617 91
- [Loa99] LOADER, Clive: *Local Regression and Likelihood*. Berlin : Springer, 1999. – ISBN 9780387987750 41
- [Lod05] LODDER, A.: Electromigration theory unified. In: *Europhysics Letters* 72 (2005), Nr. 5, S. 774–780. <http://dx.doi.org/10.1209/epl/i2005-10306-9>. – DOI 10.1209/epl/i2005-10306-9 25, 26
- [LSC02] LÜ, X. ; SHEN, WZ ; CHU, JH: Size effect on the thermal conductivity of nanowires. In: *Journal of Applied Physics* 91 (2002), S. 1542. <http://dx.doi.org/10.1063/1.1427134>. – DOI 10.1063/1.1427134 82, 83

- [Lud] LUDOPH, Bas: *Private Kommunikation mit E. Scheer* 75
- [Lud99] LUDOPH, Bas: *Quantum conductance properties of atomic-size contacts*, Universiteit Leiden, Diss., 1999 10, 24, 104, 109
- [Mar06] MARES, Ancuta I.: *Search for shell effects in metallic nanowires*, Universiteit Leiden, Diss., 2006 21, 22
- [MDC87] MARTINIS, John M. ; DEVORET, Michel H. ; CLARKE, John: Experimental tests for the quantum behavior of a macroscopic degree of freedom: The phase difference across a Josephson junction. In: *Phys. Rev. B* 35 (1987), Apr, Nr. 10, S. 4682–4698. <http://dx.doi.org/10.1103/PhysRevB.35.4682>. – DOI 10.1103/PhysRevB.35.4682 61
- [Mei08] MEIJER, G. I.: MATERIALS SCIENCE: Who Wins the Nonvolatile Memory Race? In: *Science* 319 (2008), Nr. 5870, 1625-1626. <http://dx.doi.org/10.1126/science.1153909>. – DOI 10.1126/science.1153909 91
- [MUB⁺07] MARES, A I. ; URBAN, D F. ; BURKI, J ; GRABERT, H ; STAFFORD, C A. ; RUITENBEEK, J M.: Electronic and atomic shell structure in aluminium nanowires. In: *Nanotechnology* 18 (2007), Nr. 26, S. 265403 (9pp). <http://dx.doi.org/10.1088/0957-4484/18/26/265403>. – DOI 10.1088/0957-4484/18/26/265403 22
- [Mus06] MUSTAFA, Jakob: *Design and Analysis of Future Memories Based on Switchable Resistive Elements*, Rheinisch-Westfälische Technische Hochschule Aachen, Diss., 2006. <http://darwin.bth.rwth-aachen.de/opus3/volltexte/2006/1605/> 91
- [NNWA01] NAKAMURA, Jun ; NAKAYAMA, Tomonobu ; WATANABE, Satoshi ; AONO, Masakazu: Structural and Cohesive Properties of a C₆₀ Monolayer. In: *Phys. Rev. Lett.* 87 (2001), Jul, Nr. 4, S. 048301. <http://dx.doi.org/10.1103/PhysRevLett.87.048301>. – DOI 10.1103/PhysRevLett.87.048301 91
- [PH07] PERREAULT, Simon ; HÉBERT, Patrick: Median Filtering in Constant Time. In: *IEEE Transactions on Image Processing* 16 (2007), Sept., Nr. 9, S. 2389–2394. <http://dx.doi.org/10.1109/TIP.2007.902329>. – DOI 10.1109/TIP.2007.902329 45
- [pov] *Povray*. Freeware, Vision Raytracer Pty. Ltd. <http://www.povray.org/>. Version: 3.7.0.beta33, Freeware 113
- [Pre02] PRESS, William: *Numerical Recipes in C++*. Cambridge : Cambridge University Press, 2002 <http://www.nr.com/>. – ISBN 0521750334 42
- [PVH⁺08] PAULY, F ; VILJAS, J K. ; HUNIAR, U ; HAFNER, M ; WOHLTHAT, S ; BURKLE, M ; CUEVAS, J C. ; SCHON, G: Cluster-based density-functional approach to quantum transport through molecular and atomic contacts. In: *New Journal of*

- Physics* 10 (2008), Nr. 12, 125019 (28pp). <http://stacks.iop.org/1367-2630/10/125019> 16, 102
- [Pyt67] PYTTE, E.: Contribution of the electron-phonon interaction to the effective mass, superconducting transition temperature, and the resistivity in aluminum. In: *Journal of Physics and Chemistry of Solids* 28 (1967), Nr. 1, S. 93–103. – ISSN 0022–3697 85
- [QS07] QUADBECK-SEEGER, H.J.: *Der Wechsel allein ist das Beständige*. Wiley-VCH, 2007. – ISBN 978–3–527–50343–8
- [RAP⁺96] RUITENBEEK, J. M. ; ALVAREZ, A. ; PINEYRO, I. ; GRAHMANN, C. ; JOYEZ, P. ; DEVORET, M. H. ; ESTEVE, D. ; URBINA, C.: Adjustable nanofabricated atomic size contacts. In: *Review of Scientific Instruments* 67 (1996), Nr. 1, S. 108–111. <http://dx.doi.org/10.1063/1.1146558>. – DOI 10.1063/1.1146558 9, 51
- [Ros] ROSSUM, Guido van: *Python*. Open Source. <http://www.python.org/>. Version: 2.5 66
- [SAC⁺98] SCHEER, Elke ; AGRAIT, Nicolas ; CUEVAS, Juan C. ; YEYATI, Alfredo L. ; LUDOPH, Bas ; MARTIN-RODERO, Alvaro ; BOLLINGER, Gabino R. ; RUITENBEEK, Jan M. ; URBINA, Cristian: The signature of chemical valence in the electrical conduction through a single-atom contact. In: *Nature* 394 (1998), Nr. 6689, S. 154–157. <http://dx.doi.org/10.1038/28112>. – DOI 10.1038/28112. – ISSN 0028–0836 10, 17
- [SBMH09] STOJANOVIC, N. ; BERG, JM ; MAITHRIPALA, DHS ; HOLTZ, M.: Direct measurement of thermal conductivity of aluminum nanowires. In: *Applied Physics Letters* 95 (2009), Nr. 9, S. 091905. <http://dx.doi.org/10.1063/1.3216035>. – DOI 10.1063/1.3216035. – ISSN 0003–6951 83, 85
- [Sch02] SCHIRM, Christian: *Leitwertfluktuationen von Ein-Atom-Kontakten*. Universitätsstr. 10, 78457 Konstanz, Universität Konstanz, Diplomarbeit, 2002. <http://kops.ub.uni-konstanz.de/volltexte/2003/940> 68
- [Sch08] SCHECKER, Olivier: *Von Punktkontakten zu Nano-Elektro-Mechanischen Systemen – NEMS*, Universität Konstanz, Diss., 2008 103
- [SCHD06] STAHLMECKE, B. ; CHELARU, L. I. ; HERINGDORF, F.-J. M. ; DUMPICH, G.: Electromigration in Gold and Single Crystalline Silver Nanowires. In: *AIP Conference Proceedings* 817 (2006), Nr. 1, 65–70. <http://dx.doi.org/10.1063/1.2173533>. – DOI 10.1063/1.2173533 26
- [SD07a] STAHLMECKE, B. ; DUMPICH, G.: Influence of the electron beam on electromigration measurements within a scanning electron microscope. In: *Applied Physics Letters* 90 (2007), Nr. 4, 043517. <http://dx.doi.org/10.1063/1.2432304>. – DOI 10.1063/1.2432304 26

- [SD07b] STAHLMECKE, B. ; DUMPICH, G.: Resistance behaviour and morphological changes during electromigration in gold wires. In: *Journal of Physics: Condensed Matter* 19 (2007), Nr. 4, 046210 (11pp). <http://dx.doi.org/10.1088/0953-8984/19/4/046210>. – DOI 10.1088/0953-8984/19/4/046210 26
- [SG64] SAVITZKY, A. ; GOLAY, M. J. E.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures. In: *Analytical Chemistry* 36 (1964), Nr. 8, S. 1627–1639 42
- [SHC⁺06] STAHLMECKE, B. ; HERINGDORF, F.-J. M. ; CHELARU, L. I. ; HOEGEN, M. H. ; DUMPICH, G. ; ROOS, K. R.: Electromigration in self-organized single-crystalline silver nanowires. In: *Applied Physics Letters* 88 (2006), Nr. 5, 053122. <http://dx.doi.org/10.1063/1.2172012>. – DOI 10.1063/1.2172012 26
- [SJE⁺97] SCHEER, E. ; JOYEZ, P. ; ESTEVE, D. ; URBINA, C. ; DEVORET, M. H.: Conduction Channel Transmissions of Atomic-Size Aluminum Contacts. In: *Physical Review Letters* 78 (1997), Nr. 18, S. 3535–3538. <http://dx.doi.org/10.1103/PhysRevLett.78.3535>. – DOI 10.1103/PhysRevLett.78.3535 9, 16, 17, 51
- [SK09] SCHWARZ, Hans R. ; KÖCKLER, Norbert: *Numerische Mathematik*. Teuber, 2009 (7., überarb. Aufl.). <http://dx.doi.org/10.1007/978-3-8348-9282-9>. <http://dx.doi.org/10.1007/978-3-8348-9282-9>. – ISBN 978-3-8348-9282-9 46
- [SKB⁺06] SCHEER, Elke ; KONRAD, Patrick ; BACCA, Cecile ; MAYER-GINDNER, Andreas ; LÖHNEYSSEN, Hilbert ; HÄFNER, Michael ; CUEVAS, Juan C.: Correlation between transport properties and atomic configuration of atomic contacts of zinc by low-temperature measurements. In: *First publ. in: Physical Review B* 74 (2006), 205430 (2006). <http://dx.doi.org/10.1103/PhysRevB.74.205430>. – DOI 10.1103/PhysRevB.74.205430 77
- [SNU⁺02] SMIT, R. H. M. ; NOAT, Y. ; UNTIEDT, C. ; LANG, N. D. ; HEMERT, M. C. ; RUITENBEEK, J. M.: Measurement of the conductance of a hydrogen molecule. In: *Nature* 419 (2002), oct, Nr. 6910, 906–909. <http://dx.doi.org/10.1038/nature01103>. – DOI 10.1038/nature01103. – ISSN 0028-0836 24
- [SPS09] SCHIRM, C. ; PERNAU, H.-F. ; SCHEER, E.: Switchable wiring for high-resolution electronic measurements at very low temperatures. In: *Review of Scientific Instruments* 80 (2009), Nr. 2, 024704. <http://dx.doi.org/10.1063/1.3073962>. – DOI 10.1063/1.3073962. – weitere Quelle: <http://kops.ub.uni-konstanz.de/volltexte/2009/7521> 55, 125
- [SSBW06] SZOT, Krzysztof ; SPEIER, Wolfgang ; BIHLMAYER, Gustav ; WASER, Rainer: Switching the electrical resistance of individual dislocations in single-crystalline SrTiO₃. In: *Nat Mater* 5 (2006), apr, Nr. 4, 312–320. <http://dx.doi.org/10.1038/nmat1614>. – DOI 10.1038/nmat1614. – ISSN 1476-1122 91

- [ST99] SCHRIEFFER, J. R. ; TINKHAM, M.: Superconductivity. In: *Rev. Mod. Phys.* 71 (1999), Mar, Nr. 2, S. S313–S317. <http://dx.doi.org/10.1103/RevModPhys.71.S313>. – DOI 10.1103/RevModPhys.71.S313 17
- [Sta08] STAHLMECKE, Burkhard: *Elektromigration in Gold und Silber Nanostrukturen*, Universität Duisburg-Essen, Diss., 2008. <http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:hbz:464-20080130-095740-3> 26
- [THNA05] TERABE, K. ; HASEGAWA, T. ; NAKAYAMA, T. ; AONO, M.: Quantized conductance atomic switch. In: *Nature* 433 (2005), Jan, Nr. 7021, 47–50. <http://dx.doi.org/10.1038/nature03190>. – DOI 10.1038/nature03190 10, 90, 91, 92
- [Tro09] TROUWBORST, M. L.: *Electron Transport Through Single Gold Atoms an Hydrogen Molecules*, University of Groningen, Diss., 2009 81, 85
- [TVDMVW06] TROUWBORST, ML ; VAN DER MOLEN, SJ ; VAN WEES, BJ: The role of Joule heating in the formation of nanogaps by electromigration. In: *Journal of Applied Physics* 99 (2006), S. 114316. <http://dx.doi.org/10.1063/1.2203410>. – DOI 10.1063/1.2203410. – weitere Quelle: <http://irs.ub.rug.nl/dbi/44f55f0d1fc1c> 81, 85
- [VNC] *TightVNC*. Open Source. <http://www.tightvnc.com/>. – VNC = *Virtual Network Computing*, Protokoll zur passwortgeschützten Übertragung von Bildschirminhalt und Tastatur- und Mausereignissen über das Internet 66
- [VRD09] VAN ROSSUM, G. ; DRAKE, F.L.: Python language reference. (2009) 66
- [Wer03] WERMAN, Michael: Fast Convolution. In: *Journal of WSCG* Bd. 11, 2003. – <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.64.6978> 45
- [WGW02] WITCZAK, Zbigniew ; GONCHAROVA, Valeria A. ; WITCZAK, Przemyslaw: Elastic properties of a polycrystalline sample of the L12 Al₅CrTi₂ intermetallic compound under hydrostatic pressure up to 1 GPa at room temperature. In: *Journal of Alloys and Compounds* 337 (2002), Nr. 1-2, 58 - 63. [http://dx.doi.org/10.1016/S0925-8388\(01\)01957-0](http://dx.doi.org/10.1016/S0925-8388(01)01957-0). – DOI 10.1016/S0925-8388(01)01957-0. – ISSN 0925-8388. – Zitiert nach http://www.springermaterials.com/docs/VSP/datasheet/lpf-p/01103000/LPFP_1103887.html 85
- [WHB⁺88] WEES, B. J. ; HOUTEN, H. van ; BEENAKKER, C. W. J. ; WILLIAMSON, J. G. ; KOUWENHOVEN, L. P. ; MAREL, D. van d. ; FOXON, C. T.: Quantized conductance of point contacts in a two-dimensional electron gas. In: *Phys. Rev. Lett.* 60 (1988), Feb, Nr. 9, S. 848–850. <http://dx.doi.org/10.1103/PhysRevLett.60.848>. – DOI 10.1103/PhysRevLett.60.848 15
- [Whi23] WHITTAKER, E.T.: On a new method of graduation. In: *Proceedings of the Edinburgh Mathematical Society* 41 (1923), S. 63–75 42

- [Whi53] WHITE, G K.: The Thermal Conductivity of Gold at Low Temperatures. In: *Proceedings of the Physical Society. Section A* 66 (1953), Nr. 6, 559. <http://dx.doi.org/10.1088/0370-1298/66/6/307>. – DOI 10.1088/0370-1298/66/6/307 82
- [Wie64] WIENER, Norbert: *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge : The MIT Press, 1964. – ISBN 0262730057 42
- [Woo05a] WOODCRAFT, Adam L.: Predicting the thermal conductivity of aluminium alloys in the cryogenic to room temperature range. In: *Cryogenics* 45 (2005), Nr. 6, 421 - 431. <http://dx.doi.org/10.1016/j.cryogenics.2005.02.003>. – DOI 10.1016/j.cryogenics.2005.02.003. – ISSN 0011-2275 82, 83
- [Woo05b] WOODCRAFT, A.L.: Recommended values for the thermal conductivity of aluminium of different purities in the cryogenic to room temperature range, and a comparison with copper. In: *Cryogenics* 45 (2005), Nr. 9, S. 626-636. <http://dx.doi.org/10.1016/j.cryogenics.2005.06.008>. – DOI 10.1016/j.cryogenics.2005.06.008 82, 83
- [XNOS04] XIE, F-Q. ; NITTLER, L. ; OBERMAIR, Ch ; SCHIMMEL, Th: Gate-controlled atomic quantum switch. In: *Phys Rev Lett* 93 (2004), Sep, Nr. 12, S. 128303. <http://dx.doi.org/10.1103/PhysRevLett.93.128303>. – DOI 10.1103/PhysRevLett.93.128303 10, 90
- [Yan] YANSON, A.: *Private Kommunikation mit E. Scheer* 76
- [YPX⁺08] YANG, J. J. ; PICKETT, Matthew D. ; XUEMA, Li ; OHLBERG, Douglas A. A. ; STEWART, Duncan R. ; WILLIAMS, R. S.: Memristive switching mechanism for metal/oxide/metal nanodevices. In: *Nat Nano* 3 (2008), jul, Nr. 7, 429-433. <http://dx.doi.org/10.1038/nnano.2008.160>. – DOI 10.1038/nnano.2008.160. – ISSN 1748-3387 91
- [YR97] YANSON, A. I. ; RUITENBEEK, J. M.: Do Histograms Constitute a Proof for Conductance Quantization? In: *Physical Review Letters* 79 (1997), Nr. 11, 2157-2157. <http://dx.doi.org/10.1103/PhysRevLett.79.2157>. – DOI 10.1103/PhysRevLett.79.2157 10, 20
- [YSC⁺05] YANSON, I. K. ; SHKLYAREVSKII, O. I. ; CSONKA, Sz. ; KEMPEN, H. van ; SPELLER, S. ; YANSON, A. I. ; RUITENBEEK, J. M.: Atomic-Size Oscillations in Conductance Histograms for Gold Nanowires and the Influence of Work Hardening. In: *Phys. Rev. Lett.* 95 (2005), Dec, Nr. 25, S. 256806. <http://dx.doi.org/10.1103/PhysRevLett.95.256806>. – DOI 10.1103/PhysRevLett.95.256806 22
- [YSRS08] YANSON, I. K. ; SHKLYAREVSKII, O. I. ; RUITENBEEK, J. M. ; SPELLER, S.: Aluminium nanowires: Influence of work hardening on conductance histograms.

- In: *Physical Review B (Condensed Matter and Materials Physics)* 77 (2008), Nr. 3, 033411. <http://dx.doi.org/10.1103/PhysRevB.77.033411>. – DOI 10.1103/PhysRevB.77.033411 21, 22
- [YYR99] YANSON, A. I. ; YANSON, I. K. ; RUITENBEEK, J. M.: Observation of shell structure in sodium nanowires. In: *Nature* 400 (1999), jul, Nr. 6740, S. 144–146. <http://dx.doi.org/10.1038/22074>. – DOI 10.1038/22074. – ISSN 0028–0836 21, 22
- [YZNT08] YAO, Jun ; ZHONG, Lin ; NATELSON, Douglas ; TOUR, James M.: Etching-dependent reproducible memory switching in vertical SiO₂ structures. In: *Applied Physics Letters* 93 (2008), Nr. 25, 253101. <http://dx.doi.org/10.1063/1.3045951>. – DOI 10.1063/1.3045951 91

»Sie wissen, dass ich langsam schreibe,
[...] weil ich mir nie anders gefallen
kann, als wenn in kleinem Raum
möglichst viel ist, und kurz zu schreiben
viel mehr Zeit kostet als lang.«

— Carl Friedrich Gauß²

Zusammenfassung

Techniken

Im Rahmen dieser Doktorarbeit wurden Leitwertmessungen an mechanisch kontrollierbaren Ein-Atom-Kontakten bei tiefen Temperaturen durchgeführt und deren Verhalten bei kleinen lokalen Änderungen der atomaren Konfiguration studiert. Zur möglichst lokalen Manipulation wurden mit hohen Strömen einzelne Elektromigrationsereignisse angeregt und im Leitwert die entsprechenden Sprünge beobachtet. Für die so erzeugten und veränderten Kontakte konnten zur Analyse mithilfe von supraleitenden Strom-Spannungs-Kennlinien, den sogenannten multiplen Andreev-Reflexionen (MAR), die quantenmechanischen Leitwertkanäle ermittelt werden.

Vorbereitend mussten experimentelle Techniken entwickelt werden, um die hohen Anforderungen der Rauscharmut für die MAR-Messung zu gewährleisten und gleichzeitig genügend niederohmige Leitungen für die Elektromigrationsanregung zu erhalten. Die spezielle Kabelführung im Kryostat wurde in *Review of Scientific Instruments* veröffentlicht [SPS09]. Weitere experimentelle Details, die zum Reproduzieren der Messungen wichtig sind, wurden beschrieben.

Ein zur Analyse der MAR-Messungen und zur Bestimmung der Kanalspektren wichtiges numerisches Verfahren (ein spezieller Fitalgorithmus), musste weiter verfeinert werden, da das Rauschen im System (eventuell auch die endliche Temperatur von 250 mK) zu leicht verrundeten MAR-Kennlinien geführt hat. Da sehr viele Messungen vorlagen, wurde außerdem die Effizienz des numerisch aufwendigen Optimierungsproblems deutlich verbessert, wodurch mehrere Wochen Rechenzeit auf einem Bürorechner auf wenige Stunden reduziert wurden. Zur Modellierung des Rauschens und für andere Behandlungen von Messdaten, wurde außerdem ein vielseitig verwendbarer Glättalgorithmus entwickelt, basierend auf der *lokalen Regression*.

Messergebnisse

Anhand von Einzelbeispielen konnte gezeigt werden, dass mit hohen Strömen atomare Umordnungen in der Nähe eines Ein-Atom-Kontaktes ausgelöst werden können, in einer Weise, die zu stabileren Konfigurationen führt. Ein lokales Schmelzen kommt nicht als Ursache der atomaren Änderungen in Frage, wie eine Abschätzung der Temperatur zeigt. Mit einer statistischen Analyse vieler solcher Messungen an Gold- und Aluminium-Bruchkontakten konnten bevorzugte Leitwerte beobachtet und durch stromgetriebene Umordnungen manipuliert werden. Werden lokale Änderungen von globaleren Änderungen isoliert untersucht, zeigen sich

²1833, in einem Briefwechsel mit H. C. Schumacher

ebenfalls bevorzugte Werte, jedoch bei verschobenen Leitwerten. Ein Zusammenspiel von lokalen und nichtlokalen Eigenschaften scheint also für Form der Gesamtleitwert-Histogramme verantwortlich zu sein. Das Histogramm über die Höhe der Leitwertsprünge deutet an, dass bei höheren Spannungen die Änderungen in größerer Entfernung vom atomaren Kontakt stattfinden.

In der Analyse der Stabilität zeigt sich, dass Ein-Atom-Kontakte die höchste Spannung aushalten. Das trifft auf Gold mit einem Leitwert von $1 G_0$, aber auch auf Aluminium zu, wo der Ein-Atom-Kontakt bei einem etwas kleineren Leitwert liegt.

Durch eine spezielle Stromrampe konnten reversibel schaltbare bistabile Kontaktgeometrien erzeugt werden, die als Ein-Atom-Schalter interpretiert werden. Die Schalthysteresen lagen dabei bei zufällig verteilten Werten und unterschieden sich in Form und Größe. Eine Analyse vieler Hysteresen zeigte ein vermindertes Auftreten von Leitwerten, die einem Ein-Atom-Kontakt entsprechen und in anderen Experimenten als besonders stabil beurteilt wurden. Die Erklärung ist, dass eine zu stabile Konfiguration keinen zweiten Partner-Zustand findet, der die Voraussetzung erfüllt, in einer Stromrichtung stabiler zu sein, als in der anderen.

Die Analyse der Kanalspektren der bistabilen Hysteresen zeigt einen ähnlichen Effekt, dass Kanäle mit hohen Transmissionen seltener auftreten, als in rein mechanisch erzeugten Kontakten. Bei hohen Transmissionen können aufgrund der geringeren Rückstreuwahrscheinlichkeiten die Elektromigrationskräfte nur sehr geringen Druck auf das Atom ausüben, um es zur Auslenkung zubringen. Dadurch ist die Voraussetzung für bistabiles Schalten weniger gut erfüllt.

Ein Kanalhistogramm von mechanisch erzeugten Aluminium-Kontakten um den Ein-Atom-Kontakt zeigt eine gute Übereinstimmung im Vergleich mit einer theoretischen Rechnung eines pyramidenförmigen Ein-Atom-Kontakts in (111)-Kristallrichtung. Für Kontakte mit höheren Leitwerten findet im Kanalhistogramm zunehmend eine Gleichverteilung statt. Volle Kanäle mit Transmission 1 treten erst ab einem Gesamtleitwert von etwa $1.7 G_0$ in Erscheinung. Einzelne Aluminium-Atome hingegen haben mehrere deutlich unter $1 G_0$ liegende Kanäle, wobei einer der Kanäle meist gegenüber den anderen dominiert.

Anhand der Strom-Spannungs-Kennlinien der bistabilen Hysteresen können zwei Beobachtungen gemacht werden. Zum einen können kleine atomare Änderungen des Kontaktes die Kanalverteilung deutlich verändern, während in der Regel die Form der Kennlinien ähnlich bleibt, manchmal sogar identisch. Dies zeigt, dass das Schalten ein sehr lokales Phänomen ist, mit wenig beteiligten Atomen. Zum anderen deutet eine um Null meist spitz zulaufende symmetrische Form der Kennlinien an, dass bei steigender Spannung möglicherweise eine stetige geometrische Verzerrung stattfindet, die durch Elektromigrationskräfte hervorgerufen wird und in beide Stromrichtungen einen ähnlichen Effekt hervorruft.

Danksagung

Auf dieser Seite möchte ich noch allen danken, die diese Arbeit direkt oder indirekt unterstützt haben:

Meiner Doktormutter Elke Scheer für die langjährige Unterstützung, den Vorarbeitern am Experiment, hier besonders Hansi Pernau, allen Kollegen, die immer sehr zuverlässig für das Funktionieren bestimmter Anlagen gesorgt haben (z. B. Vojko Kunej für das REM und Hansi Pernau, Michael Wolz, Stephan Egle für die „Dampfe“), Ansgar Fischer für seine präzisen Metallarbeiten, Sabine Lukas für ihre ständige Hilfsbereitschaft, den Mitarbeitern der wissenschaftlichen Werkstätten, meinem Vater und meinem Bruder fürs Korrekturlesen der Arbeit und für das externe Zweitgutachten Carlos Cuevas von der Universität von Madrid.

Indirekte Hilfe erhielt ich von Herrn Froböse, der mich in die Allzweck-Programmiersprache Python eingeführt hat. An dieser Stelle sollte man eigentlich auch mal manchen unsichtbaren Helfer aus dem Internet dankbar erwähnen, wie die Open-Source-Entwickler von Python (Guido van Rossum et al.), den Numerik-Paketen Numpy/Scipy (T. Oliphant et al.), den Satzprogrammen \LaTeX (Knuth/Lamport et al.) und KOMA-Script (Markus Kohm et al.), dem Plotprogramm PyX (A. Wobst, S. Lehmann, M. Schindler) und bei den vielen Namenlosen, die sich hinter all den Open-Source-Projekten verbergen.

