

Transport in Nanostructures

DAVID K. FERRY
STEPHEN M. GOODNICK

CAMBRIDGE STUDIES IN
SEMICONDUCTOR PHYSICS AND
MICROELECTRONIC ENGINEERING



TRANSPORT IN NANOSTRUCTURES

DAVID K. FERRY

*Regents' Professor of Electrical Engineering
Arizona State University*

and

STEPHEN M. GOODNICK

*Professor and Chair of Electrical Engineering
Arizona State University*



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, United Kingdom
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1997

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 1997

Printed in the United States of America

Typeset in Times Roman

Library of Congress Cataloging-in-Publication Data

Ferry, David K.

Transport in nanostructures / David K. Ferry, Stephen Marshall Goodnick.
p. cm. – (Cambridge studies in semiconductor physics and
microelectronic engineering ; 6)

ISBN 0-521-46141-3

I. Nanostructures. 2. Mesoscopic phenomena (Physics). 3. Solid
state electronics. I. Goodnick, Stephen Marshall, 1955–

II. Title III. Series.

QC176.8.N35F47

537.6'22 – dc21 1997

96-36825

CIP

A catalog record for this book is available from the British Library

ISBN 0 521 46141 3 hardback

The advent of semiconductor structures whose characteristic dimensions are smaller than the mean free path of carriers has led to the development of novel devices and given rise to many advances in our theoretical understanding of these mesoscopic systems or nanostructures. This book reviews the results of experimental research into mesoscopic devices and develops a detailed theoretical framework for understanding their behavior.

The authors begin by discussing the key observable phenomena in nanostructures, including phase interference and weak localization. They then describe quantum confined systems, transmission in nanostructures, quantum dots, and single electron phenomena. Separate chapters are devoted to interference in diffusive transport and temperature decay of fluctuations, and the book concludes with a chapter on non-equilibrium transport and nanodevices. Throughout, the authors interweave experimental results with the appropriate theoretical formalism.

The book will be of great interest to graduate students taking courses in mesoscopic physics or nanoelectronics, as well as to anyone working on semiconductor nanostructures or the development of new ultrasmall devices.

**Cambridge Studies in Semiconductor Physics
and Microelectronic Engineering: 6**

EDITED BY

Haroon Ahmed

Cavendish Laboratory, University of Cambridge

Michael Pepper

Cavendish Laboratory, University of Cambridge

Alec Broers

Department of Engineering, University of Cambridge

TRANSPORT IN NANOSTRUCTURES

TITLES IN THIS SERIES

- 1 Doping in III-V Semiconductors
E. F. Schubert
- 2 The Physics and Applications of Resonant Tunnelling Diodes
H. Mizuta and T. Tanoue
- 3 Electronic Transport in Mesoscopic Systems
S. Datta
- 4 Photoinduced Defects in Semiconductors
D. Redfield and R. H. Bube
- 5 Electrons and Phonons in Semiconductor Multilayers
B. K. Ridley
- 6 Transport in Nanostructures
D. K. Ferry and S. M. Goodnick

Preface

This book has grown out of our somewhat disorganized attempts to teach the physics and electronics of mesoscopic devices over the past decade. Fortunately, these have evolved into a more consistent approach, and the book tries to balance experiments and theory in the current understanding of mesoscopic physics. Whenever possible, we attempt to first introduce the important experimental results in this field followed by the relevant theoretical approaches. The focus of the book is on electronic transport in nanostructure systems, and therefore by necessity we have omitted many important aspects of nanostructures such as their optical properties, or details of nanostructure fabrication. Due to length considerations, many germane topics related to transport itself have not received full coverage, or have been referred to by reference. Also, due to the enormity of the literature related to this field, we have not included an exhaustive bibliography of nanostructure transport. Rather, we have tried to refer the interested reader to comprehensive review articles and book chapters when possible.

The Introduction of Chapter 1 gives a general overview of the important effects that are observable in small systems that retain a degree of phase coherence. These are also compared to the needs that one foresees in future small electron devices. Chapter 2 provides a general introduction to quantum confined systems, and the nature of quasi-two-, quasi-one- and quasi-zero-dimensional systems including their dielectric response and behavior in the presence of an external magnetic field. It concludes with an overview of semi-classical transport in quantum wells and quantum wires including the relevant scattering mechanisms in quantum confined systems.

Chapter 3 begins with the general principle of quantum mechanical tunneling, and its historical evolution towards present day resonant tunneling diodes. The concepts of quantum mechanical flux, reflection, and transmission are introduced and applied towards understanding the characteristics of tunneling diodes. These concepts are then generalized to more complicated quantum wave guide systems, which introduces the tunneling/transmission connection upon which the Landauer formula is based. The concept of quantized conductance is introduced, and its connection to the experimentally observed conductance quantization in quantum point contacts. This is then followed by an elaboration of simulation techniques used for modeling wave guide structures and multi-terminal structures.

In Chapter 4, we focus exclusively on quantum dot structures beginning with their electronic structure, and then the experimental results and theoretical formalism related to single electron effects in such structures such as Coulomb Blockade. This is followed by discussion of more complicated systems of multiple quantum dots, and transport through such structures.

Chapter 5 begins to discuss the effects of weak localization and universal conduction fluctuations, which are direct results of inhomogeneities and phase coherence in small structures. First, the experiments and simple understanding are presented; then the more formal treatment by Green's functions provides the detailed understanding that is necessary. This chapter is closed by discussions of open quantum dot systems and the reinterpretation of "universality" that is necessary for the theory.

Chapter 6 extends the above treatments to real temperatures, and begins the discussion of how the phase breaking process, important for loss of coherence, occurs in these systems. The temperature, or Matsubara, Green's functions are introduced in order to handle the underlying physics for this process.

Finally, Chapter 7 discusses nonequilibrium transport in nanostructure under high bias potentials. A review of the important experimental observations under nonequilibrium conditions is given, followed by the introduction of the nonequilibrium, or real time, Green's functions, which provide the formal theoretical basis for treating transport in such systems.

Currently, we are teaching a two semester graduate sequence on the material contained in the book. In the first course, which is suitable for first-year graduate students, the experiments and simpler theory, such as that for tunneling, edge states, and the Landauer-Büttiker method, are introduced. This covers parts of each of the chapters, but does not delve into the topic of Green's functions. Rather, the much more difficult treatment of Green's functions is left to the second course, which is intended for more serious-minded doctoral students. Even here, the developments of the zero-temperature Green's functions in Chapter 6, followed by the Matsubara Green's functions in Chapter 7 and the nonequilibrium (real-time) Green's functions in Chapter 8, are all coupled closely to the experiments in mesoscopic devices.

In spite of the desire to consistently increase the level of difficulty and understanding as one moves through the book, there remain some anomalies. We have chosen, for example, to put the treatment of the lattice expansion and recursive Green's functions in the chapter with wave guide modal expansions, since these two quantities are closely coupled. Nevertheless, the reader would be well served to go through Chapter 5 and its introduction of the Green's functions prior to undertaking an in depth study of the recursive Green's function. This, of course, signals that topics have been grouped together in the chapters in a manner that lies on their connection to one another in physics, rather than in a manner that would be optimally chosen for a textbook. Nevertheless, we are convinced that one can use this book in graduate coursework.

Acknowledgements

The authors are indebted to a great number of people, who have read (and suggested changes) in all or parts of the manuscript during its presentation. In addition, there are the students who suffered through the courses mentioned above without an adequate textbook. In particular, we wish to thank (in alphabetical order) Richard Akis, John Barker, Jon Bird, Paolo Bordone, Kevin Connolly, Manfred Dür, Hal Grubin, Allen Gunther, Anu Krishnaswamy, David Pivin, Dragica Vasileska, and Christoph Wasshuber. In particular we would like to thank Andreas Ecker for substantial contributions to the material on recursive Green's functions in Chapter 3 as part of his Diplom Thesis at the University of Karlsruhe.

The authors are also indebted to several groups and institutes who supported the writing of this manuscript. These include Tom Zipperian at Sandia Laboratories, Peter Vogl at the Technical University of Munich, and Chihiro Hamaguchi at Osaka University. Finally, the authors would like to thank Larry Cooper, at the Office of Naval Research, for his support in the publication of this book.

Contents

<i>Preface</i>	<i>page</i> ix
<i>Acknowledgements</i>	xi
1 Introduction	1
1.1 Nanostructures: The impact	2
1.1.1 Progressing technology	2
1.1.2 Some physical considerations	4
1.2 Mesoscopic observables in nanostructures	8
1.2.1 Ballistic transport	8
1.2.2 Phase interference	11
1.2.3 Universal conductance fluctuations	13
1.2.4 Weak localization	14
1.2.5 Carrier heating in nanostructures	16
1.3 Space and time scales	18
1.4 An introduction to the subsequent chapters	19
1.5 What is omitted	21
2 Quantum confined systems	23
2.1 Nanostructure materials	24
2.2 Quantization in heterojunction system	26
2.2.1 Quantum wells and quasi-two-dimensional systems	27
2.2.2 Coupled wells and superlattices	31
2.2.3 Doped heterojunction systems and self-consistent solutions	34
2.3 Lateral confinement: Quantum wires and quantum dots	39
2.3.1 Nanolithography	39
2.3.2 Quantum wire and quantum dot structures	41
2.4 Electronic states in quantum wires and quantum dots	44
2.5 Magnetic field effects in quantum confined systems	46
2.5.1 Magnetic field in a 2DEG	47
2.5.2 Magnetic field and 1D waveguides: Edge states	51
2.6 Screening and collective excitations in low-dimensional systems	54
2.6.1 Dielectric function in quasi-2D systems	56
2.6.2 Dielectric function in quasi-1D systems	62
2.7 Homogeneous transport in low-dimensional systems	63
2.7.1 Semiclassical transport	64
2.7.2 Relaxation time approximations	66

2.7.3	Elastic scattering mechanisms	68
2.7.4	Lattice scattering	75
2.7.5	Experimental mobility in 2DEG heterostructures	81
2.7.6	Magnetotransport in quantum confined structures	84
3	Transmission in nanostructures	91
3.1	Tunneling in planar barrier structures	92
3.2	Wavefunction treatment of tunneling	96
3.2.1	Single rectangular barrier	97
3.2.2	The double barrier case	104
3.2.3	Tunneling time	111
3.3	Current in resonant tunneling diodes	113
3.3.1	Coherent tunneling	114
3.3.2	Incoherent or sequential tunneling	118
3.3.3	Space charge effects and self-consistent solutions	121
3.4	Landauer formula	124
3.5	The multi-channel case	128
3.6	Quantized conductance in nanostructures	131
3.6.1	Experimental results in quantum point contacts	131
3.6.2	Adiabatic transport model	134
3.6.3	Temperature effects	137
3.6.4	Inhomogeneous effects	138
3.6.5	Nonlinear transport	141
3.7	Transport in quantum waveguide structures	145
3.7.1	Mode-matching analysis	146
3.7.2	Transport through bends	151
3.7.3	Lateral resonant tunneling	153
3.7.4	Coupled waveguides	155
3.8	Lattice Green's function method	156
3.8.1	Single-particle Green's functions	159
3.8.2	Tight-binding Hamiltonian	161
3.8.3	Lattice Green's functions	162
3.8.4	Analytic forms of lattice Green's functions	166
3.8.5	Relation between Green's functions and S-matrix	173
3.8.6	Recursive Green's function method	174
3.8.7	Application to specific geometries	178
3.9	Multi-probe formula	181
3.9.1	Specific examples	185
3.9.2	Experimental multi-probe measurements ($B = 0$)	191
3.10	Magnetic fields and quantum waveguides	194
3.10.1	Quantized conductance in a perpendicular field	195
3.10.2	Edge states and the quantum Hall effect	198
3.10.3	Selective population of edge states	201
4	Quantum dots and single electron phenomena	209
4.1	Electronic states in quantum dot structures	209
4.1.1	Noninteracting electrons in a parabolic potential	209
4.1.2	Dot states in a magnetic field	213

4.1.3	Multi-electron quantum dots	215
4.1.4	Quantum dot statistics	219
4.1.5	Spectroscopy of quantum dots	220
4.2	Single electron tunneling and Coulomb blockade	226
4.2.1	Introduction to Coulomb blockade and experimental studies	226
4.2.2	Orthodox theory of single electron tunneling	249
4.2.3	Co-tunneling of electrons	262
4.2.4	Coulomb blockade in semiconductor quantum dots	264
4.3	Coupled dots and quantum molecules	264
4.4	Transport in anti-dot systems	267
4.4.1	The low-magnetic field regime	268
4.4.2	The high-magnetic field regime	272
5	Interference in diffusive transport	280
5.1	Weak localization	282
5.1.1	Semiclassical treatment of the conductance	283
5.1.2	Effect of a magnetic field	286
5.1.3	Size effects in quantum wires	291
5.1.4	The magnetic decay "time"	294
5.1.5	Extension to short wires	297
5.2	Universal conductance fluctuations	299
5.3	The Green's function in transport	305
5.3.1	Interaction and self-energies	308
5.3.2	Impurity scattering	309
5.3.3	Beyond the Drude result	314
5.4	Weak-localization correction to the conductance	317
5.4.1	The cooperon correction	319
5.4.2	Role of a magnetic field	323
5.4.3	Periodic eigenvalues for the magnetic effects	325
5.5	Quantum treatment of the fluctuations	328
5.5.1	The correlation function in energy	330
5.5.2	Correlation function in a magnetic field	334
5.6	Summary of universality	336
5.6.1	The width dependence of the fluctuations	336
5.6.2	Size variation of the correlation magnetic field	338
5.6.3	Breakdown of the universality in a magnetic field	340
5.7	Fluctuations in quantum dots	346
6	Temperature decay of fluctuations	361
6.1	Temperature decay of coherence	362
6.1.1	Decay of the coherence length	363
6.1.2	Decay of the coherence time	368
6.1.3	Summary	371
6.2	The role of temperature on the fluctuations	372
6.2.1	Fluctuation amplitudes	375
6.2.2	Dimensional crossover	376
6.2.3	Correlation ranges	377

6.3	Electron-electron interaction effects	379
6.3.1	Electron energy loss in scattering	381
6.3.2	Screening and plasmons	382
6.3.3	Temperature Green's functions	386
6.3.4	One-particle density of states	391
6.3.5	The effective interaction potential	392
6.3.6	Electron-electron interactions in disordered systems—The self-energy	403
6.4	Conductivity	408
6.5	The phase-breaking time	414
6.5.1	Interactions coupled to background fields	415
6.5.2	Modifications of the self-energy	419
7	Nonequilibrium transport and nanodevices	423
7.1	Nonequilibrium transport in mesoscopic devices	425
7.1.1	Nonequilibrium effects in tunnel barriers	426
7.1.2	Ballistic transport in vertical and planar structures	429
7.1.3	Thermopower in nanostructures	431
7.1.4	Measuring the hot electron temperature	436
7.1.5	Hot carriers in quantum dots	437
7.1.6	Breakdown of the Landauer-Büttiker formula	442
7.2	Real-time Green's functions	445
7.2.1	Equations of motion for the Green's functions	447
7.2.2	The Langreth theorem	449
7.2.3	The Green-Kubo formula	450
7.3	Transport in an inversion layer	453
7.3.1	Coulomb scattering	454
7.3.2	Surface-roughness scattering	455
7.3.3	The retarded function	458
7.3.4	The "less-than" function	463
7.4	Considerations of mesoscopic devices	465
7.4.1	A model device	465
7.4.2	Proportional coupling in the leads	469
7.4.3	A noninteracting resonant-level model	471
7.4.4	Another approach to the phonon-assisted tunneling	473
7.5	Nonequilibrium transport in high electric fields	476
7.5.1	The retarded function	477
7.5.2	The "less-than" function	482
7.5.3	Gauge-invariant formulations	484
7.6	Screening with the Airy-transformed Green's function	487
7.7	Other approaches to quantum transport in nonequilibrium systems	490
7.7.1	The density matrix	492
7.7.2	The Wigner distribution	494
<i>Index</i>		507

Introduction

It is often said that *nanostructures* have become the system of choice for studying transport over the past few years. What does this simple statement mean?

First, consider transport in large, macroscopic systems. Quite simply, for the past fourscore years, emphasis in studies of transport has been on the Boltzmann transport equation and its application to devices of one sort or another. The assumptions that are usually made for studies are the following: (i) scattering processes are local and occur at a single point in space; (ii) the scattering is instantaneous (local) in time; (iii) the scattering is very weak and the fields are low, such that these two quantities form separate perturbations on the equilibrium system; (iv) the time scale is such that only events that are slow compared to the mean free time between collisions are of interest. In short, one is dealing with structures in which the potentials vary slowly on both the spatial scale of the electron thermal wavelength (to be defined below) and the temporal scale of the scattering processes.

In contrast to the above situation, it has become possible in the last decade or so to make structures (and devices) in which characteristic dimensions are actually smaller than the appropriate mean free paths of interest. In GaAs/GaAlAs semiconductor heterostructures, it is possible at low temperature to reach mobilities of $10^6 \text{ cm}^2/\text{Vs}$, which leads to a (mobility) mean free path on the order of $10 \mu\text{m}$ and an inelastic (or phase-breaking) mean free path even longer. (By “phase-breaking” we mean decay of the energy or phase of the “wave function” representing the carrier.) This means that transport in a regime in which the Boltzmann equation is clearly invalid becomes easily accessible. Each of the assumptions detailed above provides a factor that is neglected in the usual Boltzmann transport picture. Structures (and devices) can readily be built with dimensions that are much smaller than these dimensions, so new physical processes become important in the overall transport. These devices have come to be called *nanostructures*, *nanodevices*, or *mesoscopic devices*, depending upon the author. Perhaps the best description is that of a mesoscopic device, where the prefix “meso-” is used to indicate structures that are large compared to the microscopic (atomic) scale but small compared to the macroscopic scale upon which normal Boltzmann transport theory has come to be applied.

A simple consideration illustrates some of the problems. If the basic semiconductor material is doped to 10^{18} cm^{-3} , then the mean distance between impurity atoms is 10 nm , so that any discrete device size, say $0.1 \mu\text{m}$, spans a countably small number of impurity atoms. That is, a cubic volume of $0.1 \mu\text{m}$ on a side contains only 1000 atoms. These atoms are not uniformly distributed in the material; instead they are randomly distributed with large fluctuations in the actual concentration on this size scale. Again, the variance in the actual number N in any volume (that is, the difference from one such volume to another) is

roughly $\sqrt{N}$, which in this example is about 32 atoms (or 3.3% of the doping). Since these atoms often compose the main scattering center at low temperatures, the material is better described as a highly conducting but disordered material, since the material is certainly not uniform on the spatial scale of interest here. As the current lines distort to avoid locally high densities of impurities, the current density becomes non-uniform spatially within the material; this can be expected to lead to new effects. Since the dimensions can be smaller than characteristic scattering lengths, transport can be ballistic and highly sensitive to boundary conditions (contacts, surfaces, and interfaces). To complicate the problem, many new effects that can be observed depend upon the complicated many-body system itself, and simple one-electron theory no longer describes these new effects. Finally, the size can be small compared to the phase-breaking length, which nominally describes the distance over which the electron wave's phase is destroyed by some process. In this case, the phase of the particle becomes important, and many phase-interference effects begin to appear in the characteristic conductance of the material.

Our purpose in this book is twofold. First, we will attempt to review the observed experimental effects that are seen in mesoscopic devices. Second, we want to develop the theoretical understanding necessary to describe these experimentally observed phenomena. But in the remainder of this chapter, the goal is simply to give an introduction into the type of effects that are seen, and to discuss why these effects will be important to future technology, as well as for their interesting physics.

1.1 Nanostructures: The impact

1.1.1 Progressing technology

Since the introduction of the integrated circuit in the late 1950s, the number of individual transistors that can be placed upon a single integrated circuit chip (often just called the *chip*) has approximately quadrupled every three years. The fact that more functionality can be put on a chip when there are more transistors, coupled to the fact that the basic cost of the chip (in terms of \$/cm²) changes very little from one generation to the next, leads to the conclusion that greater integration leads to a reduction in the basic cost per function for high-level computation as more functions are placed on the chip. It is this simple functionality argument that drives the chip progress. In 1980, Hewlett-Packard produced a single-chip microprocessor containing approximately 0.5 M devices in its 1 cm² area [1]. This chip was produced with transistors having a nominal 1.25 μm gate length and was considered a remarkable step forward. Today, the dynamic random-access memory chip (DRAM) is the technology driver; the 64 Mbit is currently in production, with the 256 Mbit expected in 1998. The former chip obviously contains on the order of 64 million transistors. With this progress, one can expect to see 10^9 devices on a single chip just after the turn of the century. In general, this rapid progress in chip density has followed a complicated scaling relationship [2]. The reduction in critical feature size, such as the gate length, is actually a factor of only 0.7 each generation, and this produces only a doubling of the device density. (Other factors are an increase in the actual chip area and changes in the circuit implementation, such as the introduction of trench capacitors.) Still, this leads to some remarkable projections. The 64 Mbit chip uses nominally 0.35- μm gate length transistors. Following the scaling relationships will lead to gate lengths of only 0.1 μm in just over a decade (for the 16 Gbit chip, which scaling suggests will arrive in full production in 2007).

From this discussion, one can reasonably ask just how far the size of an individual electron device can be reduced, and if we understand the physical principles that will govern the behavior of devices as we approach this limiting size. In 1972, Hoeneisen and Mead [3] discussed the minimal size expected for a simple MOS gate (as well as for bipolar devices). Effects such as oxide breakdown, source-drain punch-through, impact ionization in the channel, and so on were major candidates for processes to limit downscaling. Years later, Mead [4] reconsidered this limit in terms of the newer technologies that have appeared since the earlier work, concluding that one could easily downsize the transistor to a gate length of 30 nm *if macroscopic transport theory continued to hold*.

The above considerations tell us that the industry is pushing the critical dimensions downward at a very rapid rate. In contrast to this, research has led to the fabrication of really small individual transistors that operate (at room temperature) in a more-or-less normal fashion. For example, Schottky-gate FETs and high-electron-mobility FETs in GaAs have been made with gate lengths down to 20 nm [5]–[7], and MOSFETs in Si have been made with fabricated gate lengths down to 40 nm [8]. In the latter, the effective gate length was as short as 27 nm. While these devices appear to be normal, there is already evidence that the transport is changing, with tunneling through the gate depletion region becoming more important [9]. Perhaps the more important attribute is the variation that can be expected as one moves from device to device across a chip containing several million transistors. If there is a significant fluctuation in the number of impurities (and/or the number of electrons/holes), then the performance of the devices varies significantly across the chip. This is a major reliability problem, which translates into a dramatic reduction in the effective *noise margin* (the range over which a voltage level can vary without changing the state of a logic gate) of the devices in the chip. This in turn translates into reduced performance of the chip.

Granted that the technological momentum is pushing to ever smaller devices, and that the technology is there to prepare really small devices, it becomes obvious that we must now ask whether our physical understanding of devices and their operation can be extrapolated down to very small space and time scales without upsetting the basic macroscopic transport physics – or do the underlying quantum electronic principles prevent a down-scaling of the essential semi-classical concepts upon which this macroscopic understanding is based? Preliminary considerations of this question were presented more than a decade ago [10]. Suffice it to say, though, that experiment has progressed steadily as well, and ballistic (and therefore coherent and unscattered) transport has been seen in the base region of a GaAs/AlGaAs hot electron transistor [11]. From this, it is estimated that the inelastic mean free path for electrons in GaAs may be as much as $0.12\ \mu\text{m}$ *at room temperature*. Moreover, there are simulations that suggest that it is less than a factor of two smaller in Si [12]. The inelastic mean free path is on the order of (and usually equal to) the energy relaxation length $l_e = v\tau_e$, where τ_e is the energy relaxation time and v is a characteristic velocity (which is often the Fermi velocity in a degenerate system). (There is some ambiguity here because the energy relaxation time is usually defined as the effective inverse decay rate for the mean electron energy, or temperature. The definition here talks about a mean-free path for energy relaxation, which is not quite the same thing. This is complicated by the fact that, in mesoscopic systems, one really talks about a phase-breaking time, which is meant to refer to the average time for relaxation of the coherent single particle phase of a charge carrier. Again, this is a slightly different definition. This ambiguity exists throughout the literature, and although we will probably succumb to it in later chapters, the reader should recognize

these subtle differences.) This tells us that, even in Si, the electron inelastic mean free path may be 50–100 nm and comparable to gate lengths of the devices that are expected to be made in the not-too-distant future. Since the phase will likely remain coherent over these distances, it is quite natural to expect phase interference effects to appear in the transport, and to expect most of the assumptions inherent in the Boltzmann picture to be violated. The small device will then reflect the intimate details of the impurity distribution in the particular device, and macroscopic variations can then be expected from one device to another. These effects are, of course, well known in the world of mesoscopic devices. Thus, the study of mesoscopic devices, even at quite low temperatures, provides significant insight into effects that may well be expected to occur in future devices.

Consider, as an example, a simple MOSFET with a gate length of 50 nm and a gate width of 100 nm. If the number of carriers in the channel is $2 \times 10^{12} \text{ cm}^{-2}$, there are only about 100 electrons on average in the open channel. If there is a fluctuation of a single impurity, the change in the conductance will not be 1%, but will be governed by the manner in which the phase interference of the carriers is affected by this fluctuation. This effect is traditionally taken to be of order e^2/h , which leads to a fluctuation in conductance of about $40 \mu\text{S}$. If our device were to exhibit conductance of 1 S/mm (of gate width), the absolute conductance would only be $100 \mu\text{S}$, so that the fluctuation is on the order of 40% of the actual conductance. This is a very significant fluctuation, arising from the lack of ensemble averaging in the limited number of carriers in the device. In fact, this may well be a limiting mechanism for the down-scaling of individual transistor sizes.

1.1.2 Some physical considerations

In macroscopic conductors, the resistance that is found to exist between two contacts is related to the bulk conductivity and to the dimensions of the conductor. In short, this relationship is expressed by

$$R = \frac{L}{\sigma A}, \quad (1.1)$$

where σ is the conductivity and L and A are the length and cross-sectional area of the conductor, respectively. If the conductor is a two-dimensional conductor, such as a thin sheet of metal, then the conductivity is the conductance per square, and the cross-sectional area is just the width W . This changes the basic formula (1.1) only slightly, but the argument can be extended to any number of dimensions. Thus, for a d -dimensional conductor, the cross-sectional area has the dimension $A = L^{d-1}$, where here L must be interpreted as a “characteristic length.” Then, we may rewrite (1.1) as

$$R = \frac{L^{2-d}}{\sigma_d}. \quad (1.2)$$

Here, σ_d is the d -dimensional conductivity. Whereas one normally thinks of the conductivity, in simple terms, as $\sigma = ne\mu$, the d -dimensional term depends upon the d -dimensional density that is used in this definition. Thus, in three dimensions, σ_3 is defined from the density per unit volume, while in two dimensions σ_2 is defined as the conductivity per unit square and the density is the sheet density of carriers. The conductivity (in any dimension) is not expected to vary much with the characteristic dimension, so we may take the logarithm