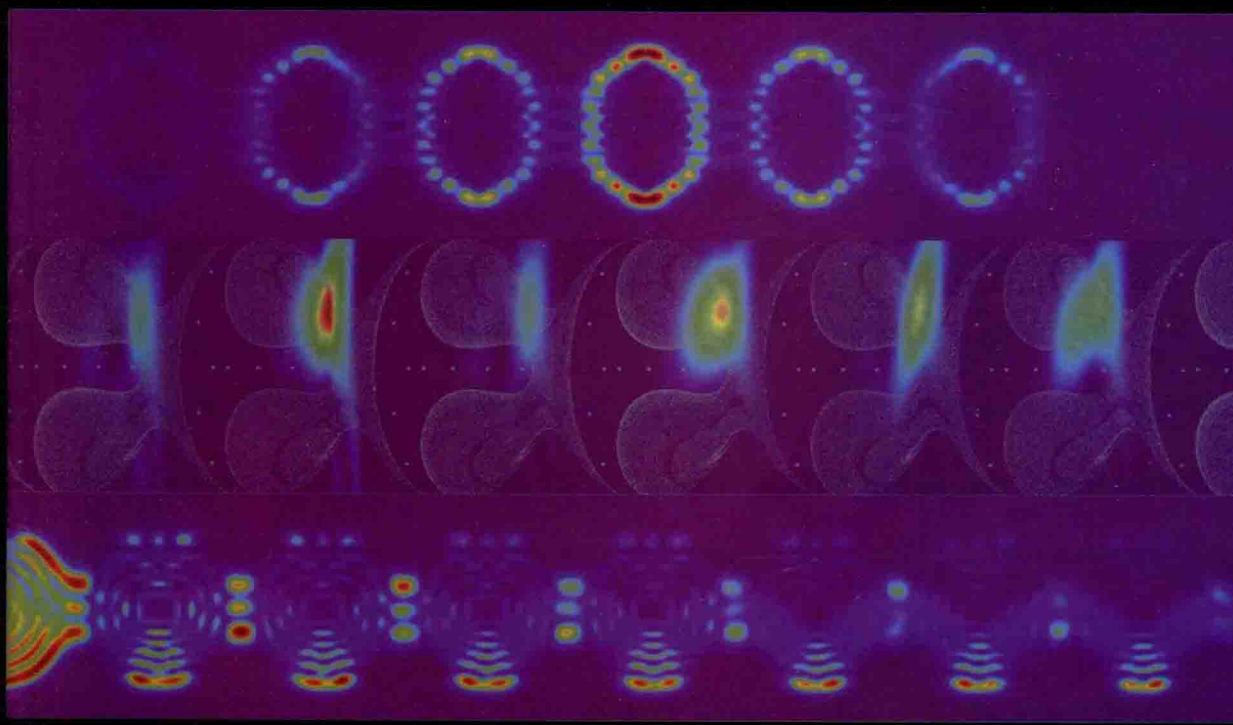
David K. Ferry,
Stephen M. Goodnick,
Jonathan Bird

Transport in Nanostructures

Second Edition

纳米结构中的输运 第2版

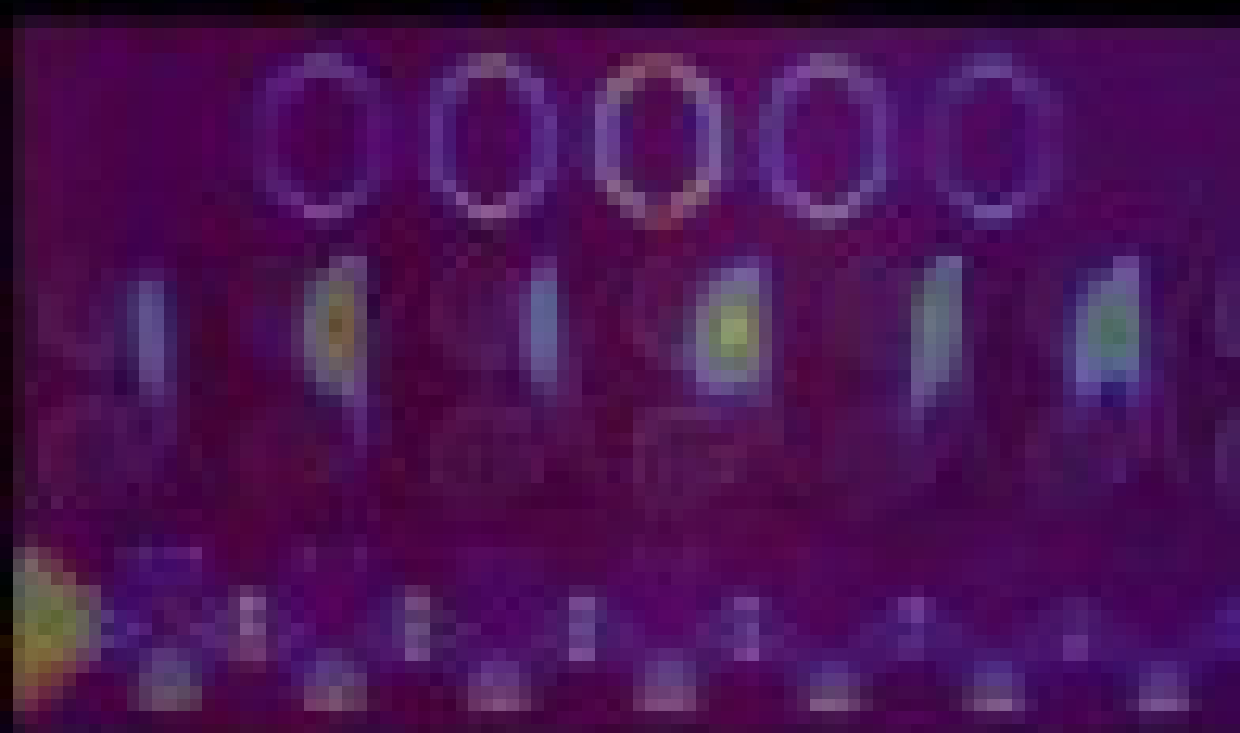


CAMBRIDGE

世界图书出版公司
www.wpcbj.com.cn

Transport in Nanostructures

纳米结构中的输运



Transport in Nanostructures

Second Edition

David K. Ferry
Arizona State University
Stephen M. Goodnick
Arizona State University
Jonathan Bird
University at Buffalo

CAMBRIDGE
UNIVERSITY PRESS

图书在版编目（CIP）数据

纳米结构中的输运：第2版 = Transport in Nanostructures 2nd ed. ：英文/（美）费里(Ferry,D. K.) 著. —影印本. —北京：世界图书出版公司北京公司，2014. 10
ISBN 978 - 7 - 5100 - 8776 - 9

Ⅰ. ①纳… Ⅱ. ①费… Ⅲ. ①纳米材料—薄膜—输运过程—研究—英文 Ⅳ. ① 0484. 3

中国版本图书馆 CIP 数据核字（2014）第 241765 号

书　　名：	Transport in Nanostructures Second Edition
作　　者：	David K. Ferry, Stephen M. Goodnick, Jonathan Bird
中 译 名：	纳米结构中的输运 第2版
责任编辑：	高蓉　刘慧
出 版 者：	世界图书出版公司北京公司
印 刷 者：	三河市国英印务有限公司
发　　行：	世界图书出版公司北京公司（北京朝内大街 137 号 100010）
联系电话：	010 - 64021602, 010 - 64015659
电子信箱：	kjb@ wpcbj. com. cn
开　　本：	16 开
印　　张：	42
版　　次：	2015 年 3 月
版权登记：	图字：01 - 2014 - 2643
书　　号：	978 - 7 - 5100 - 8776 - 9　　定　　价：　149. 00 元

Transport in Nanostructures

Second Edition

Providing a much-needed update on the latest experimental research, this new edition has been thoroughly revised and develops a detailed theoretical framework for understanding the behavior of mesoscopic devices.

The second edition now contains greater coverage of the quantum Hall effect, in particular, the fractional quantum Hall effect; one-dimensional structures, following the growth of research in self-assembled nanowires and nanotubes; nanoscale electronic devices, due to the evolution of device scaling to nanometer dimensions in the semiconductor industry; and quantum dots.

The authors combine reviews of the relevant experimental literature with theoretical understanding and interpretation of phenomena at the nanoscale. This second edition will be of great interest to graduate students taking courses in mesoscopic physics or nanoelectronics, and researchers working on semiconductor nanostructures.

DAVID K. FERRY is Regents' Professor in the Department of Electrical Engineering at Arizona State University. His areas of research include nanoelectronic devices, quantum transport, and nonequilibrium transport. He is a Fellow of the American Physical Society, the Institute of Electrical and Electronics Engineers, and of the Institute of Physics (UK).

STEPHEN M. GOODNICK is Director of the Arizona Institute for Nanoelectronics and Professor of Electrical Engineering at Arizona State University, where he researches transport in semiconductor devices, computational electronics, quantum and nanostructured devices, and device technology.

JONATHAN BIRD is a Professor in the Department of Electrical Engineering at the University of Buffalo and a Visiting Professor in the Graduate School of Advanced Integration Science, Chiba University. His research interests lie in the area of nanoelectronics. He is the co-author of more than two hundred peer-reviewed publications, and of undergraduate and graduate textbooks.

Preface

The original edition of this book grew out of our somewhat disorganized attempts to teach the physics and electronics of mesoscopic devices over the past decade. Fortunately, these evolved into a more consistent approach, and the book tried to balance experiments and theory in the current, at that time, understanding of mesoscopic physics. Whenever possible, we attempted to first introduce the important experimental results in this field followed by the relevant theoretical approaches. The focus of the book was on electronic transport in nanostructure systems, and therefore by necessity we 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 did not receive full coverage, or were referred to only by reference. Also, due to the enormity of the literature related to this field, we did not include an exhaustive bibliography of nanostructure transport. Rather, we tried to refer the interested reader to comprehensive review articles and book chapters when possible.

The decision to do a second edition of this book was reached only after long and hard consideration and discussion among the authors. While the first edition was very successful, the world has changed significantly since its publication. The second edition would have to be revised extensively and considerable new material added. A decision to go ahead was made only after welcoming Jon Bird to the author's team. Once this was done, we then carefully discussed the revison and its required redistribution of material among several new chapters. Even so, the inclusion of this considerable material has meant that a lot of material has been left out of this second edition, in order to bring it down to a tractable size. This even included a considerable amount of material that was in the first edition, but no longer appears. We hope that the reader is not put off by this; as the first edition was so successful, we anticipate many of the readers will already have that tome. But, it was essential to include more up-to-date material and topics while maintaining a rational size for the book. Thus, the decision was in principle already made for us.

Currently, we still are teaching a two semester graduate sequence (at ASU) 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, quantum Hall effect, quantum dots, 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 now in Chapter 3, followed by the Matsubara Green's functions in Chapter 8, and the nonequilibrium (real-time) Green's functions in Chapter 9, 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 recursive Green's functions in the chapter with waveguide transport and the recursive solutions to the Lippmann–Schwinger equation, since these two treatments of quantum transport are closely coupled. Nevertheless, the reader would be well served to go through the 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 relies 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, as is clear from our own courses.

Chapter 1 is, of course, an introduction to the material in the entire book, but new material on nanodevices has been added, as progress in silicon technology has brought the normal metal-oxide-semiconductor field-effect transistor (MOSFET) into the mesoscopic world. Additionally, a new introduction to nanowires and carbon nanotubes is given. Chapter 2 remains a discussion of quantum confined systems, but now focuses more on the one-dimensional structures rather than the two-dimensional ones. New material here includes a discussion of numerical solutions to the Schrödinger equation and Poisson's equation as well as non-self-consistent Born approximations to scattering in quasi-one-dimensional systems. Chapter 3 remains focused upon very low temperature transport, but now includes the introduction to the different approaches to (equilibrium) quantum transport. Chapter 4 is a completely new chapter focused upon the quantum Hall effect and the fractional quantum Hall effect.

Chapter 5 is also a new chapter which focuses upon quantum wires, and includes some of the modern investigations into various effects in these wires. Chapter 6 focuses upon quantum dots with new material on the role of spin. The focus upon single electron tunneling remains, but considerable new material on coupled dots has been added. Then, Chapter 7 discusses weakly disordered systems. New material on (strong) localization is presented so that weak localization can be placed in its context in relation to this material.

Chapter 8 is mostly carried over from the first edition and discusses the role of temperature. Here, the Matsubara Green's functions are introduced in addition

to the semiclassical approach. Finally, Chapter 9 discusses nonequilibrium transport and nanodevices. Here, considerable new material on semiconductor nanodevices has been added. Device simulation via the scattering matrix implementation based upon the Lippmann–Schwinger equation appears as well as the treatment with nonequilibrium Green's functions.

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 preparation, and particularly since the first edition appeared. In addition, there are the students who suffered through the courses mentioned above without an adequate text-book. In particular, we wish to thank (in alphabetical order) Richard Akis, Joy Barker, Paolo Bordone, Adam Burke, Kevin Connolly, Neil Deutscher, Manfred Dür, Matthew Gilbert, Allen Gunther, Irena Knezevic, Anu Krishnaswamy, David Pivin, Chetan Prasad, Jo Rack, Steve Ramey, Lucian Shifren, Dragica Vasileska, and Christoph Wasshuber. In particular, a considerable amount of the material on quantum transport has come from their experiments and theory, as well as from the Purdue group.

The authors are also indebted to several groups and institutes who supported the original 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, formerly at the Office of Naval Research, for his support in the publication of the original version of this book.

Contents

1
Introduction

Nanostructures are generally regarded as ideal systems for the study of electronic transport. What does this simple statement mean?

First, consider transport in large, macroscopic systems. In bulk materials and devices, transport has been well described via the Boltzmann transport equation or similar kinetic equation approaches. The validity of this approach is based on the following set of assumptions: (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.

Since the late 1960s and early 1970s, researchers have observed quantum effects due to confinement of carriers at surfaces and interfaces, for example along the Si/SiO_2 interface, or in heterostructure systems formed between lattice-matched semiconductors. In such systems, it is still possible to separate the motion of carriers parallel to the surface or interface, from the quantized motion perpendicular, and describe motion semiclassically in the unconstrained directions. Since the 1980s, however, it has been possible to pattern structures (and devices) in which characteristic dimensions are actually smaller than the appropriate mean free paths of interest. In GaAs/AlGaAs semiconductor heterostructures, it is possible at low temperature to reach mobilities in excess of 10^7 cm^2/Vs, which leads to a (mobility) mean free path on the order of 100 μm 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, where the latter term 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 μm, spans a countably small number of impurity atoms. That is, a cubic volume of 0.1 μm 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. 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 comprise the main scattering centers 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 nanoscale and mesoscopic devices. Second, we want to develop the theoretical understanding necessary to describe these experimentally observed phenomena. 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 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 $\$/cm^2$) has changed very little from one generation to the next (until recently), 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 has driven device feature reduction according to a complicated scaling relationship [1]. In 1980, Hewlett-Packard produced a single-chip microprocessor containing approximately 0.5 M devices in its 1 cm^2 area [2]. This chip was produced with transistors having a nominal 1.25 μm gate length and was considered a remarkable step forward. In contrast, by 2007, the functionality of the dynamic random access memory (DRAM) is on the order of 2 Gbit, a number which is expected to double by 2010 [3]. The printed gate length of production microprocessor transistors in this same year was 48 nm and the physical gate length closer to 25 nm. Research devices have been demonstrated down to 10 nm gate length or less. Clearly, current integrated circuit manufacturing is truly a nanoscale technology.

For a 25 nm gate length Si device, the number of atoms spanning the channel is on the order of a 100 or less. Hence, 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 [4] 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 [5] 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 current ITRS roadmap (2007) now predicts scaling solutions down to 10 nm gate length before a serious "brick wall" is encountered, and 15 nm gate lengths are scheduled for production by 2010. Laboratory MOSFET devices with gate lengths down to 15 nm have been reported by Intel [6] and AMD [7] which exhibit excellent *I–V* characteristics, and 6 nm gate length *p*-channel transistors have been reported by IBM [8].

Given this rapid scaling of device technology towards 1 nm feature sizes, 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 semiclassical concepts upon which this macroscopic understanding is based? Preliminary considerations of this question were presented more than two decades ago [9]. If transport is ballistic, meaning carriers suffer few or no scattering as they traverse the channel, quantum effects are expected to play a major role. Ballistic (and therefore coherent and unscattered) transport was

already observed in the base region of a GaAs/AlGaAs hot electron transistor [10]. From this, it is estimated that the inelastic mean free path for electrons in GaAs may be as much as 0.12 μm at room temperature. Simulation results in Si indicate that at room temperature, the ballistic mean free path may be much smaller, only a few nanometers [11], which may partly explain the success in scaling of Si MOSFETs discussed above. 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).[1] 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. A 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 2×10^{12} 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 μS. If our device were to exhibit conductance of 1 S/mm (of gate width), the absolute conductance would only be 100 μS, 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, when trying to realize circuit architectures involving 100s of millions of transistors that have to perform within a relatively narrow range of tolerance, necessitating entirely new types of fault tolerant designs to accommodate such fluctuations.

[1] 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.

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

$$R = \frac{L}{\sigma A}, \tag{1.1}$$

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

$$R = \frac{L^{2-d}}{\sigma_d}. \tag{1.2}$$

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 of the last equation. Then, taking the derivative with respect to $\ln(L)$ leads to

$$\frac{\partial \ln(R)}{\partial \ln(L)} = 2 - d. \tag{1.3}$$

This result is expected for macroscopic conducting systems, where resistance is related to the conductivity through Eq. (1.2). We may think of this limit as the *bulk* limit, in which any characteristic length is large compared to any characteristic transport length.

In mesoscopic conductors, the above is not necessarily the case, since we must begin to consider the effects of ballistic transport through the conductor. (For ballistic transport we generally adopt the view that the carrier moves through the structure with very little or no scattering, so that it follows normal phase space trajectories.) However, let us first consider a simpler situation. We have assumed that the conductivity is independent of the length, or that σ_d is a constant. However, if there is surface scattering, which can dominate the mean free path, then one could expect that the latter is $l \sim L$. Since $l = v_F\tau$, where v_F is the Fermi velocity in a degenerate semiconductor and τ is the mean free time, this leads to