VOLUME TWENTY

FRONTIERS OF NANOSCIENCE

SEMICONDUCTOR NANODEVICES

Physics, Technology and Applications

Edited by

DAVID A. RITCHIE

Professor of Experimental Physics,
Cavendish Laboratory,
University of Cambridge,
Cambridge, United Kingdom

Professor of Semiconductor Science and Technology,
Swansea University,
Swansea, United Kingdom



ELSEVIER

Frontiers of Nanoscience

Series Editors

Richard E. Palmer
Nanomaterials Laboratory, College of Engineering, Swansea University, UK

Vol. 1 **Nanostructured Materials** edited by
Gerhard Wilde

Vol. 2 **Atomic and Molecular Manipulation** edited by
Andrew J. Mayne and Gérald Dujardin

Vol. 3 **Metal Nanoparticles and Nanoalloys** edited by
Roy L. Johnston and J.P. Wilcoxon

Vol. 4 **Nanobiotechnology** edited by
Jesus M. de la Fuente and V. Grazu

Vol. 5 **Nanomedicine** edited by
Huw Summers

Vol. 6 **Nanomagnetism: Fundamentals and Applications** edited by
Chris Binns

Vol. 7 **Nanoscience and the Environment** edited by
Jamie R. Lead and Eugenia Valsami-Jones

Vol. 8 **Characterization of Nanomaterials in Complex Environmental and Biological Media** edited by
Mohammed Baalousha and Jamie R. Lead

Vol. 9 **Protected Metal Clusters: From Fundamentals to Applications** edited by
Tatsuya Tsukuda and Hannu Häkkinen

Vol. 10 **Structure and Properties of Nanoalloys** edited by
Riccardo Ferrando

Vol. 11 **Materials and Processes for Next Generation Lithography** edited by
Alex P. G. Robinson and Richard A. Lawson

Vol. 12 **Computational Modelling of Nanoparticles** edited by
Stefan T. Bromley and Scott M. Woodley

Vol. 13 **Self-Assembly of Nano- and Micro-structured Materials using Colloidal Engineering** edited by
Dwaipayan Chakrabarti and Stefano Sacanna

Vol. 14 **Nanostructured Thin Films** edited by
Maria Benelmekki and Andreas Erbe

Vol. 15 **Cluster Beam Deposition of Functional Nanomaterials and Devices** edited by
Paolo Milani and Mukhles Sowwan

Vol. 16 **Colloids for Nanobiotechnology: Synthesis, Characterization and Potential Applications** edited by
Wolfgang J. Parak and Neus Feliu

Vol. 17 **Computational Modelling of Nanomaterials** edited by
Panagiotis Grammatikopoulos

Vol. 18 **Nanoscale Electrochemistry** edited by
Andrew J. Wain and Edmund J.F. Dickinson

Vol. 19 **Nanomaterials for Electrochemical Energy Storage** edited by
Rinaldo Raccichini and Ulderico Ulissi

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States

Copyright © 2021 Elsevier Ltd. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-822083-2
ISSN: 1876-2778

For information on all Elsevier publications visit our website at
https://www.elsevier.com/books-and-journals

Publisher: Matthew Deans
Acquisitions Editor: Edward Payne
Editorial Project Manager: Charlotte Rowley
Production Project Manager: Nirmala Arumugam
Cover Designer: Mark Rogers

Typeset by TNQ Technologies

ELSEVIER Book Aid International Working together to grow libraries in developing countries
www.elsevier.com • www.bookaid.org

Cover Credits: Harvey Beere provided the photograph of the University of Cambridge molecular beam epitaxy system, Pedro Vianez and Christopher Ford provided the device images.

Contributors

A.J. Bennett
School of Engineering, Queen's Buildings, Cardiff University, Cardiff, Wales

Siming Chen
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom

Yong P. Chen
Department of Physics and Astronomy, Purdue University, West Lafayette, IN, United States; Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, United States

Glattli D. Christian
SPEC, CEA Saclay Nanoelectronics Group, Service de Physique de l'Etat Condense, Universite Paris-Saclay, CEA Saclay, Gif-sur-Yvette, France

Christopher Ford
Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom

M. Fernando Gonzalez-Zalba
Quantum Motion Technologies, London, United Kingdom

Mark J. Holmes
Institute for Nano Quantum Information Electronics, The University of Tokyo, Meguro-ku, Tokyo, Japan

Gulzat Jaliel
The London Centre for Nanotechnology, University College London, London, United Kingdom

Masaya Kataoka
National Physical Laboratory, Teddington, United Kingdom

Morteza Kayyalha
Department of Electrical Engineering, Pennsylvania State University, University Park, PA, United States

S. Kumar
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom

Huiyun Liu
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom

Yoshitaka Okada
Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Tokyo, Japan; Department of Advanced Interdisciplinary Studies, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

Rachel A. Oliver
Department of Materials Science and Metallurgy, The University of Cambridge, Cambridge, United Kingdom

Jae-Seong Park
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom

Patrick Parkinson
Department of Physics and Astronomy and the Photon Science Institute, University of Manchester, Manchester, United Kingdom

M. Pepper
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom

David A. Ritchie
Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom; Department of Physics, Singleton Campus, Swansea University, United Kingdom

Leonid P. Rokhinson
Department of Physics and Astronomy, Purdue University, West Lafayette, IN, United States; Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, United States

Simon Schaal
London Centre for Nanotechnology, University College London, London, United Kingdom; Intel Components Research, Intel Corporation, Hillsboro, OR, United States

Yasushi Shoji
Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Tokyo, Japan; Global Zero Emission Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

Ryo Tamaki
Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Tokyo, Japan

Mingchu Tang
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom

Oleksandr Tsyplyatyev
Institut für Theoretische Physik, Universität Frankfurt, Frankfurt, Germany

Pedro Vianez
Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom

Katsuhisa Yoshida
Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Tokyo, Japan; Institute of Applied Physics, University of Tsukuba, Tsukuba, Japan

CHAPTER ONE

Introduction, background and contents

David A. Ritchie[1,2,*]
[1]Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
[2]Department of Physics, Singleton Campus, Swansea University, United Kingdom
*Corresponding author: dar11@cam.ac.uk

Nanodevices are all around us. Laptops, tablets, televisions, mobile phones, cameras, cars, and many other everyday items contain billions of silicon devices with active region dimensions as small as 5 nm (around 20 atoms across), decreasing by an order of magnitude over the past decade. These devices are made with remarkable silicon fabrication technology using deep ultraviolet lithography for patterning. Recent developments such as three-dimensional integrated circuit and device architecture, the use of high-k dielectrics and high-mobility materials, have enabled an ongoing improvement in device and integrated circuit performance. In these developments, the device designers are fighting against quantum effects due to the small number of electrons contained in each device, along with problems due to quantum tunnelling through barriers and so on. In this book, we describe research which embraces, explains, and tries to exploit many of these quantum effects, discussing work carried out using a wide range of nanodevices.

One can imagine that in the reasonably near future, conventional computers will incorporate quantum processors, which will be able to solve many currently intractable problems. These computers will be linked together by optical devices, such as sources, modulators, repeaters, and detectors based on nanostructures and nanodevices. These devices may be powered by solar cell technology, fabricated with advanced materials based on nanostructures.

The development of nano devices started in the 1960s and 1970s with the fabrication of silicon transistors and integrated circuits. The availability of this technology enabled the use of silicon devices for the study of semiconductor physics, for example, the two-dimensional electron gas in the inversion layer of MOSFETs (Metal-oxide-semiconductor field effect transistors) was used for the study of electron localisation effects at low

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00008-3
© 2021 Elsevier Ltd.
All rights reserved.

temperatures, which led in 1980 to the discovery of the quantum Hall effect and a value of resistance of great accuracy. This is an effect where electrons are localised to one-dimensional edge states by a large magnetic field and which still has important implications for current work, including device studies of topologically protected quantum states.

In parallel with the use of silicon devices, was the development of III–V semiconductors, particularly the GaAs/AlGaAs material system where the invention and development of the crystal growth technology; molecular beam epitaxy (MBE) led to the fabrication of devices containing very high purity GaAs giving rise, for example, to ultra-high-mobility two-dimensional electron gases with electron mean free paths above 100 microns.

The technological developments in MBE included the addition of load-locks and buffer chambers with sample outgassing, the use of low-outgassing materials in thermal sources and the substrate heaters, such as tantalum, molybdenum and boron nitride, and the use of liquid nitrogen-cooled cryopanels, along with large helium-cooled cryopumps. In-situ diagnostics such as reflection high energy electron diffraction and optical measurement of band gaps enabled the measurement of epilayer surface structure and temperature during growth. In addition, significant improvements were made in the purity of source materials and the availability of low defect substrates. As well as these system improvements, the development of the epilayer structure design and growth protocols were important, in particular, the concept of modulation doping was crucial for the growth of high-mobility two-dimensional electron gases. Developments in MBE technology often relied on good interactions between research scientists and equipment manufacturers, which widened the availability of these techniques and generally pushed the field forward.

These advances in MBE technology largely took place during the 1980s and 1990s, and their effect is illustrated by the rapid development in the features observed in the fractional quantum Hall effect. This ranged from the first report in 1982 where the main observation was at filling factor 1/3 though to 1987 when features at around 20 different filling factors were observed; during this time, the electron mobility had increased by a factor of around 18. It was to increase by a further factor of 20 over the next decade or so.

The growth of self-assembled InAs quantum dots, important for work described in several chapters of this book, started as curiosity-driven research in the mid-1990s, trying to understand the science of dot formation. Optical measurements however quickly showed that these quantum dots had

remarkable optical properties, with very narrow optical emission line widths and nanosecond lifetimes; consequently, self-assembled dots were quickly adopted for a wide range of quantum optical experiments.

At a similar time, device patterning techniques were being developed with the aim of reducing dimensionality; confining two-dimensional electrons to one-dimensional wires and zero-dimensional quantum dots. The use of electron beam lithography with metal gate deposition and wet or dry etching was important for fabrication of the finer features of bespoke nanodevices. The use of the metal split-gate technique developed in 1985 was crucial and has been exploited in many developments ever since. Electron beam lithography has moved from the use of converted scanning electron microscopes with a home-built pattern generator and capable of producing fairly well-defined 100 nm lines of metal forming single layer patterns on small chips of semiconductor wafers, to multimillion pound commercial systems which can produce well-defined lines of 10 nm in width and which can accurately align several different layers, doing all of this on 8-in. wafers.

In addition to this, scientific discovery required the development of high-quality experimental measurement techniques. For example, in the study of electron transport, the development of ultra-low temperature cryostats, operating down to millikelvin temperatures with a high magnetic field in the sample space, and with quick sample exchange was important. In more recent times, access to the sample space for high frequency and optical experiments has been required. Once again, the close interaction of scientists with equipment manufacturers has led to a range of improvements in this area. In addition to this, the development of sensitive optical measurement techniques has been important. This has been driven by the availability of pulsed laser systems, fast and sensitive detectors based on avalanche photodiodes and superconducting nanowires as well as the remarkable developments in optical communications technology leading to plug and play optical components for quantum optics experiments.

In this book, of which this is the first chapter, we discuss a wide range of devices where the operation of the devices is dominated by an element with nanoscale dimensions. These devices range from those intended to study the fundamental physics of nanoscale semiconductors at low temperatures, to the development of devices for engineering applications such as lasers and solar cells. The book is divided into two areas of work, focussing first on electron nanodevices (Chapters 2–8) and second on optical nanodevices (Chapters 9–13).

The second chapter describes the fundamentals of electron transport in one-dimension, a field of research which accelerated rapidly with the discovery of the quantisation of electrical conductance in quantum point contacts, first observed in 1988, but which is an area of research that is still producing unexpected results; for example, in this chapter, the recent observation of the fractional quantisation of conductance in one-dimension and zero magnetic is described. This leads nicely in to the third chapter which describes measurements of strong electron correlation effects in nanodevices; one area covered in detail describes how magneto-tunnelling spectroscopy of electron tunnelling between parallel one-dimensional wires are interpreted by a model of spin-charge separation in Tomonaga–Luttinger liquid theory. The fourth chapter describes the observation of thermal transport and other effects in quantum dots and the development of nanoscale energy harvesters. The fifth chapter discusses the development of electron pumping devices with potential for applications to metrology such as use as a current standard. Chapter 6 is an extensive study of noise measurements in semiconductor nanodevices from both a theoretical and experimental perspective. This area is becoming increasingly important, particularly when considering high sensitivity measurements of devices containing small numbers of electrons. Chapter 7 describes an area of increasing interest; transport studies using topological insulators and in particular topological insulator nanoribbons. Chapter 8 describes recent developments in the areas of silicon quantum dot devices with applications to quantum computing, currently a field of great scientific and technological interest.

The second part of the book (Chapters 9–13) describes optoelectronic nanodevices. Much of this work is reliant on the formation, during the semiconductor growth process, of self-assembled nanostructures, typically quantum dots formed from InAs or InGaAs. These strained quantum dots can be grown defect free by both MBE and metal organic vapour phase deposition; consequently, they can have excellent optical properties with linewidths as low as a few μeV in an emission energy of around 1 eV, which has led to their use as 'artificial atoms' in many sophisticated optical experiments.

Chapter 9 describes the use of this type of quantum dot in devices which emit single-photons or pairs of entangled photons. The emphasis in this work is on the high-speed electrical control of these devices, using applied voltages to control the nature of the light emitted as well as the time at which it is emitted. Chapter 10 describes the applications of high densities of self-assembled quantum dots in solar cells, where the quantum dot energy levels produce an intermediate band, increasing the spectrum of light that

can be absorbed. Chapter 11 concerns the use of InAs quantum dots in III–V semiconductor lasers grown on silicon substrates. The advantage of using the quantum dots in lasers being the modification of the density of states available for the lasing action, the availability of high-quality growth on silicon substrates enables easier incorporation into silicon photonic circuits than is currently possible. Chapter 12 describes the use of self-assembled nanowires in lasers, the nanowire acting as an integrated waveguide, cavity and gain material, with potential for novel applications. Finally, Chapter 13 describes the use of the III–N material system for the fabrication of devices such as single photon sources operating in the blue region of the spectrum. A major advantage here is being the high temperature of operation due to the relatively large bandgaps in this material system.

The study of semiconductor nanodevices has been of great interest for around half a century, often driven along by close interaction between industry and academia. As we move into an era where quantum effects in devices become increasingly important, the development of new materials and technology to fabricate increasingly complex devices and measure their physical properties often under extreme conditions means that this field of research is likely to continue well into the future.

CHAPTER TWO

Advances in interaction effects in the quasi one-dimensional electron gas

S. Kumar* **and M. Pepper**
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom
*Corresponding author: sanjeev.kumar@ucl.ac.uk

1. Introduction

Experimental efforts in confining electrons below two dimensions and manipulating their wavefunctions in semiconductor nanostructures have been one of the major research areas in condensed matter physics [1–5]. One of the breakthroughs in semiconductor nanostructure fabrication which further accelerated interest in the field was achieved in 1986 when Thornton and colleagues demonstrated the use of split-gate technology to confine the two-dimensional electron gas (2DEG) formed at the interface of a GaAs/AlGaAs heterostructure to a narrow, one-dimensional (1D) channel [6]. The split-gate technology subsequently led to the first observations of quantised conductance through 1D quantum wires [7] and quantum point contacts (QPCs) [8]. The conductance measured, as a function of gate voltage, of a typical 1D constriction in the GaAs heterostructure consists of quantised plateaus in units of $N2e^2/h$, where N is the number of allowed 1D subbands and the factor 2 is due to spin degeneracy. This surprising result provided new insight in that controllable energy quantisation could be achieved through an electric field without utilising a magnetic field, unlike the quantum Hall effect (QHE) [4]. Since then, the 1D research field has generated enormous interest in fundamental quantum physics as well as for future technological applications [9–13]. During that time when many researchers were intrigued by the recently discovered fractional quantum Hall effect (FQHE) [5,14], several theorists, starting with Haldane, were looking into the possibility of whether fractional conductance could be seen in low-dimensional systems without Landau levels [15–19]. Despite theoretical and experimental progress, there were very few many-body effects observed in 1D systems [20] until almost a decade later when an anomalous effect was observed; a new plateau

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00006-X
© 2021 Elsevier Ltd.
All rights reserved.

at a conductance of 0.7 ($2e^2/h$) [21]. The 0.7 structure, or 0.7 ($2e^2/h$), on application of an in-plane magnetic field became fully spin polarised with a conductance of e^2/h [21–23]. Later using transverse electron focussing (TEF), it was shown that the 0.7 structure appeared due to intrinsic spin polarisation [24–27] via the exchange interaction [28,29] in the channel. The literature is divided on its origin, and accordingly, it has been a widely investigated quantum phenomena in the 1D system. The 0.7 structure has also been reported in systems as due to the Kondo effect [30], as well as systems where both the intrinsic spin polarised 0.7 structure and the Kondo induced 0.7 have been reported [31–35].

Despite a typical 1D system being very simple, just comprising a pair of split gates, its many-body physics is not very well understood. Many new effects awaited discovery but could not be seen due to the challenges in introducing an additional control parameter into the system. An additional gate over the split gates was introduced to control the carrier concentration within the 1D channel [36], allowing the formation of a very low-density 1D system suitable for the investigation of many-body effects [36–41]. Controlling carrier density and confinement potential allowed observation of a new effect where the first plateau at $2e^2/h$ disappeared, creating a new ground state at $4e^2/h$, which was understood to be the incipient Wigner lattice in the 1D system, based on theoretical work by Meyer and Matveev [42–44]. On reducing the carrier density, the Coulomb interaction between electrons overcomes the confinement potential to reconfigure a line of 1D electrons into a zigzag and then into a two-row spatial arrangement. This new state was considered to be rich in spin and charge phases depending on neighbour and nearest next neighbour exchange interactions [44]. Recently, a weakly confined 1D quantum wire, in the zigzag phase, resulted in the observation of fractional conductance plateaus in units of e^2/h at: (a) 1/2 for the spin degenerate case and 1/4 in the spin polarised case in a Ge-based hole gas [45] and (b) 1/2, 1/4, 1/6, 2/5 and many other fractional states for electrons in GaAs in the absence of a quantising magnetic field [38,41]. These recent results show that the 1D system is rich in new physics, the advent of new means of realising complex gate geometries, as well as the availability of new, high-quality material systems, may result in the emergence of many exciting results including entangled states [46], electron pairing [47–49], etc. In this chapter, we present some recent work on electron transport in 1D quantum wires, primarily investigating many-body effects which lead to the incipient Wigner lattice, TEF of spin polarised electrons and zero-magnetic field fractional quantisation of conductance.

2. Experimental

The devices used in various sections are discussed here. A typical 1D device is fabricated from GaAs/AlGaAs heterostructures grown using molecular beam epitaxy. Two kinds of wafers were used, the first is bulk modulation–doped and the second is delta modulation–doped. For studies leading to Wigner crystallisation and fractional conductance, delta-doped GaAs/AlGaAs semiconductor heterostructures in which the 2DEG is formed up to 290 nm beneath the surface were utilised. At 1.5 K, the 2DEG had a low-temperature mobility of 1.6×10^6 cm^2/Vs and an electron density of 9.0×10^{10} cm^{-2}. For TEF measurement, the devices were fabricated from a modulated-doped GaAs/AlGaAs heterostructure with the 2DEG formed at 90 nm from the surface with electron density (mobility) of 1.80×10^{11} cm^{-2} (2.17×10^6 cm^2/Vs). A pair of split gates of length (width) of 0.4 μm (0.7 μm), and a top gate of length 1 μm separated by a 300 nm thick insulating layer of cross-linked polymethyl methacrylate were patterned by a standard electron beam lithographic technique (see Fig. 2.1a and b). The samples were cooled to a lattice temperature of 25 mK in a cryofree dilution refrigerator [38,40]. The two-terminal differential conductance (G) measurements were performed by sweeping the split gate voltage in the presence of an excitation voltage of 10 μV at 73 Hz. The electron focusing experiments were performed using a four-terminal conductance measurement technique [25,26].

3. Quantum transport properties of 1D devices

3.1 Cooldown

A 1D quantum wire based on a GaAs/AlGaAs heterostructure and its conductance characteristics are shown in Fig. 2.1. The device schematic is shown in (a) and (b) and explained previously in the experimental section, with S and D forming the source and drain for the two-terminal conductance measurements, respectively. The device was mounted in a dilution refrigerator and conductance measurements were performed while sweeping the split gate voltage, with a value V_{sg}. In this measurement, the split gate voltage was continuously swept as the cryostat was cooling down, i.e. as a function of decreasing temperature. The resulting plot is shown in the inset (c) where traces are overlapped. For clarity, a horizontal offset is

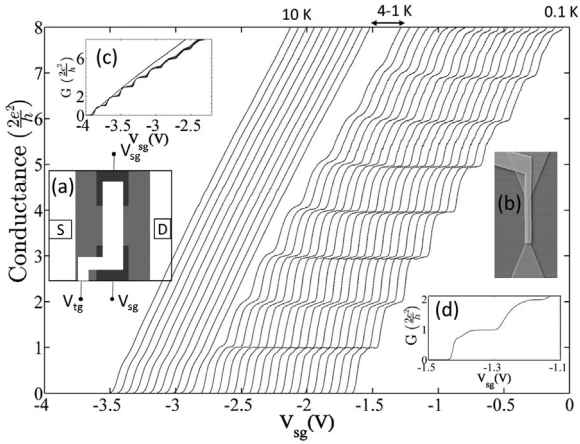


Figure 2.1 Cooldown characteristic and device layout of a typical one-dimensional (1D) quantum wire. (a) A schematic of the device consisting of a pair of split gates (in black), and over it, a cross-linked polymethyl methacrylate (PMMA) in grey acts as a dielectric for the top gate in white. S and D form the source and drain for the two terminal conductance measurements. (b) A magnified scanning electron micrograph of a top-gated, split gate device. The top gate spans an area of 1 μm, and the split gate is 400 nm long and 700 nm wide. (c) A conductance plot of a 1D device as the sample was cooling down in a dilution refrigerator. The starting temperature was 10 K and continuous split gate voltage sweeps were performed until the base temperature was 100 mK. The main plot shows the data in (c) replotted by setting a horizontal offset between the consecutive traces. On the top, a temperature range is indicated which is not true to scale. (d) A well-defined 0.7 conductance anomaly is observed at 100 mK in a different cooldown.

set between the consecutive traces, and the result is shown in the main plot. The conductance traces are shown from 10 K until 100 mK (not to scale, on moving from left to right). At 10 K, the thermal energy, kT, where k is the Boltzmann constant and T is the operating temperature, is more than the 1D subband spacing, therefore, no plateaus are resolved until around 1K where 1D subbands start to become depopulated when the split gate was swept negative. A negatively swept split gate would raise the conduction band bottom with respect to the Fermi energy which will increase the energy spacing between the 1D energy levels due to increased transverse confinement. On further lowering the temperature, the plateaus became more pronounced and the $0.7(2e^2/h)$ structure strengthens. Fig. 2.1d shows a magnified view of the $0.7(2e^2/h)$ structure from a different experimental cooldown at 100 mK.

3.2 Transverse electron focussing

Within a 1D system, the spin degeneracy can be partially lifted by tuning the exchange interaction between neighbouring electrons, thus giving rise to a spin energy gap in the system [28,29]. This spin gap, which gives rise to spin polarisation, can be measured through TEF where the height of each focussing peak is proportional to the population of detected 1D electrons [25–27]. Fig. 2.2 shows the experimental setup along with a typical focussing spectrum obtained using a 1D-based TEF device [25]. The injector and detector quantum wires are at right angles to each other forming an orthogonal focussing geometry. In the presence of a small positive transverse magnetic field (perpendicular to wafer surface), B, electrons are focused from injector to detector leading to peaks periodic in B using the relation, $B = \sqrt{2} \cdot \hbar \cdot k_F / eL$, where e is the electronic charge, ħ is the reduced Planck constant, k_F the Fermi wavevector, L is the separation between injector and detector and the prefactor $\sqrt{2}$ is due to the orthogonal focussing geometry. The experimentally observed periodicity of 60 mT is in agreement with the calculated value. Apart from the well-resolved focussing peaks, it may be noticed that the first and the third peaks (with increasing B) are split into two sub-peaks,

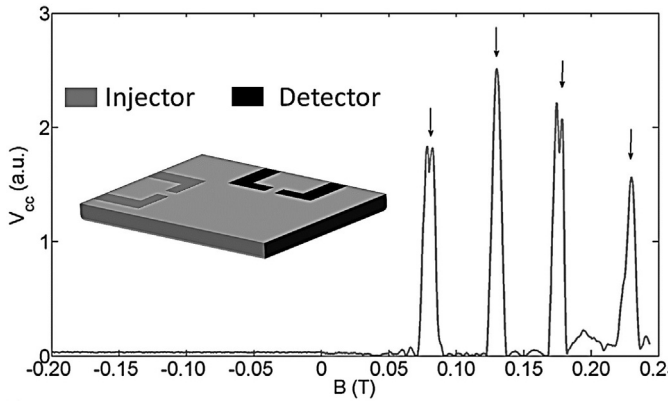


Figure 2.2 A representative result of transverse electron focussing as a function of applied transverse magnetic field. Periodic focussing peaks are well defined and the positions are in good agreement with calculated values as highlighted by the arrows. Inset shows an illustration of the diagonal focussing geometry. Lithographically defined separation between the injector and detector is 1.5 μm (along the diagonal direction). The width (confinement direction) of the quantum wire is 500 nm and the length (transport direction) is 400 nm. *Adapted from the Open Access [25].*

the splitting for the first peak is around 6 mT, while the second and fourth peaks remain single. Similar observations are also made for p-type GaAs [50,51], n-type InSb [52] and n-type $In_xGa_{1-x}As$ [53]. In this work, the shape dependence of the injector quantum wire was investigated and it was realised that when the injector is a rectangular shape, rather than the pointed gate, as in a QPC, the odd-focussing peaks show splitting [25]. Also, for the rectangular-shaped injector, the splitting of the odd-focussing peaks was pronounced when the injector conductance was increased from $0.7(2e^2/h)$ to $2(2e^2/h)$. The splitting of odd-numbered peaks indicates that the effect arises from the spin-orbit interaction (SOI). When the partially spin polarised 1D electrons are injected into the 2D region, the presence of the SOI will cause the Fermi surface to split into two, due to the different spin orientations. Therefore, the cyclotron radii will be different for the two spin states, and a split in odd-focussing peaks is observed. The absence of splitting in even-focussing peaks is due to reflection at the boundary which changes the momentum from k to $-k$ while the spin orientation remains conserved [54]. This will result in hopping of spin down (up) states between inner (outer) and outer (inner) Fermi surfaces and the reunion of the two spin states to create a single peak which is measured using the detector. Although the odd-peak splitting is a manifestation of the SOI in 2D, the asymmetry of the two sub-peaks reflects the spin polarisation of the injected 1D electrons. A contribution from dynamic spin polarisation [55,56] cannot be ruled out in enhancing spin splitting or odd peak splitting. It appears that the tuning of the SOI in the 2D regime by spatial control of spin states of injected 1D electrons could open a new regime of spin-engineering for spintronics as well as quantum information schemes.

3.3 Spin repulsion via transverse electron focussing

In these experiments, exchange-driven spin repulsion has been investigated via TEF. The quantum wire used for the injector and detector has a width of 500 nm and a length of 800 nm. The TEF measurement in this device showed spitting in odd-focussing peaks and even peaks remained as single peaks. In the presence of an in-plane magnetic field (magnetic field parallel to wafer surface), in addition to the transverse magnetic field, the splitting of the first peak is enhanced from 5.5 to 8.3 mT, whereas the second peak started to showing a tendency to split due to the Zeeman effect, thus confirming the effect to be spin-related [25–27,54,57].

Fig. 2.3a shows the first focussing peak when the detector is set at $G_0 = 2e^2/h$, and the injector conductance is varied from 0.4 to $3G_0$ [26]. In the lowest injector conductance, G_i, in a regime ($0.4\ G_0 < G_i < 0.6\ G_0$), a single highly asymmetric peak occurs around 0.044 T; however, by widening the injector further, a pronounced peak splitting is observed, resulting in subpeaks I and II, which survive up to $2G_0$. We note that the asymmetric single peak in the low injector conductance regime aligns with peak I rather than the central dip in the sub-peaks, suggesting that sub-peak I represents a spin-state. The split in the focussing peak persists up to $2G_0$ which is consistent with the experimental observation of $1.7G_0$ in the conductance measurement which was attributed to a spontaneous spin polarisation in the 1D system [58]. The two spin states become degenerate at high injector conductance, resulting in a single broad peak for $3G_0$. A significant feature of the results is the alternation of the height of the spin-split peaks when the injector is widened to vary the injector conductance from

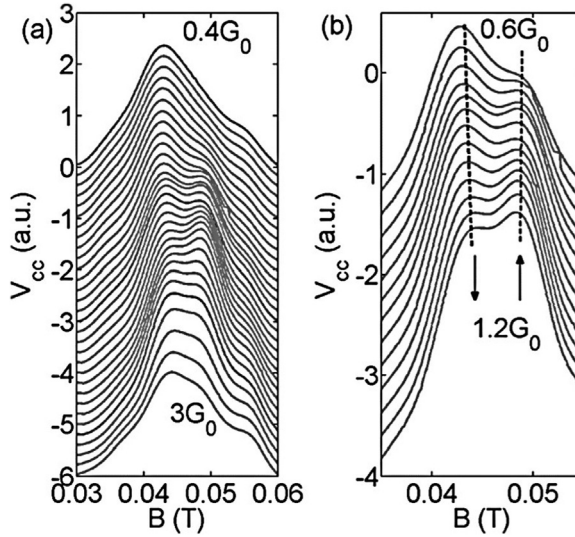


Figure 2.3 Transverse electron focussing as a function of injector conductance. (a) Injector conductance was increased from $0.4G_0$ (top trace) to $3G_0$ (bottom trace). On widening the injector to $0.6G_0$, two sub-peaks start to be resolved, and then merge to form a broad peak at $3G_0$. (b) A magnified view of the data in (a) for $0.6G_0 < G_i < 1.2G_0$. The dotted lines are a guide to the eye, reflecting the emergent alteration or flip-flop of the two spin states. Data in (a) and (b) have been offset vertically for clarity. *Adapted from the Open Access [26].*

$0.6G_0$ to $1.2G_0$ (see Fig. 2.3b). Due to the exchange repulsive interaction between 1D electrons, a spin gap is created. In the first level, subpeak I, representing the population of spin-down states, has a higher magnitude than subpeak II, spin-up states, when the injector conductance is below G_0. On entering the second level, the first spin states to populate will be spin-up, so the spin-up electron population will rise compared to the number of electrons in spin-down states resulting an alternation in the magnitude of sub-peak heights as the channel is widened and the levels progressively become filled. We suggest that the interaction between the spin levels depends on the density of states at the Fermi energy. One of the silent features of this result is that the level with the smallest number of electrons and lowest Fermi energy also has the highest density of states at the Fermi energy; hence, it is this effect which is responsible for the alternation in the spin levels. The spin repulsion thus obtained may have implications for basic quantum physics and spintronics devices.

3.4 Many-body effects and the Incipient Wigner Crystal

In a low-density 1D system, when the confinement is sufficiently weak, the enhanced Coulomb interaction can overcome the kinetic energy causing the electrons to rearrange into a zigzag, and with a further increase in interaction a two-row situation arises, known as the incipient Wigner crystal [36,37,40]. It was suggested that the zigzag would occur when $r_s \cong r_0$, where $r_s \sim 1/2n_{1D}$ is the Wigner Seitz radius with n_{1D} the density of 1D electrons, and $r_0 \sim (2e^2/\varepsilon m^*\omega^2)^{3/2}$, is the characteristic length scale of the transverse displacement at which the Coulomb interaction between neighbouring electrons becomes comparable to the harmonic confinement, and $a_B << 1/n_{1D}$, where a_B is the Bohr radius, ε is the dielectric constant and ω is the angular frequency of the harmonic oscillator [44]. A Wigner crystal would exist if $a_B d^{-2} << n_{1D} << a_B^{-1}$, provided $d >> a_B$, where d is the distance of the metal gates from the 2DEG. In experimental devices, the 1D constriction is formed electrostatically by metal gates which could screen the Coulomb interactions between the carriers therefore it is preferred to utilise a system with a deep 2DEG (around 300 nm from the surface).

Fig. 2.4 shows the conductance characteristics of a top gated, split gate device (see, Fig. 2.1a and b for a typical top gated, split gate device). The purpose of the top gate is to alter the carrier concentration within the 1D channel. The top gate is varied from −5 V (left) to −9.5 V (right) in steps of $\Delta V_{tg} = -50$ mV. Strong (weak) confinement is on the left (right) of

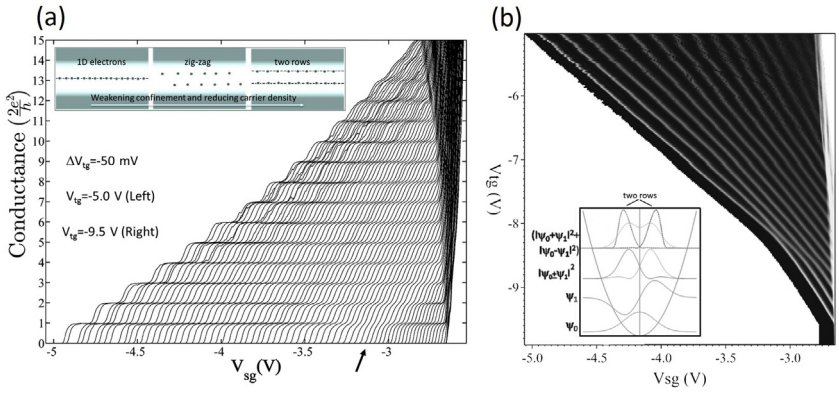


Figure 2.4 (a) Conductance characteristics of a quasi-one-dimensional channel as a function of top gate voltage, V_{tg} from −5.0 (left) to −9.5 V (right). On reducing the carrier concentration by making the top gate more negative, the confinement weakens, as does the plateau corresponding to $2e^2/h$, and on further reduction of carrier density the plateau reappears. Inset shows a cartoon of the mechanism taking place as a function of weakening confinement potential and reducing carrier density. (b) Grayscale plot of dG/dV_{sg} for the data in (a); here the white regimes are the risers in conductance and the dark regions are the conductance plateaus. Inset shows a cartoon of the hybridisation of the ground state and the first excited state wavefunction in a double-row system. The simple wavefunctions are shown, whereas in reality, they are heavily distorted by the electron-electron interaction. *The inset is adapted from the Open Access Ref.* [40].

the plot. In the strong confinement on the left, the usual quantised conductance plateaus in units of $2e^2/h$ are observed. Weakening the confinement results in the weakening of the first plateau at $2e^2/h$ plateau and $4e^2/h$ becomes the ground state (shown by an arrow), eventually $2e^2/h$ reappears when the confinement is further weakened (Fig. 2.4a). This can be seen at $V_{sg} \sim -3.25$ V ($V_{tg} \sim -8.2$ V). The greyscale of the transconductance (dI/dV) plot in Fig. 2.4b shows how the energy levels behave as the confinement is weakened and the interaction becomes dominant; a clear anticrossing of the initial ground and first excited states is seen resulting in the formation of two rows in the system. A cartoon in the inset of Fig. 2.4a shows the evolution of two rows as a function of decreasing carrier concentration. The measured conductance depends on the number of occupied subbands below the Fermi level. The first plateau representing the ground state, the 0 state, corresponds to a transverse wave function which is half-sine like. The second plateau at $4e^2/h$, resulting from the occupation of the second energy level, the 1-state, is represented by the transverse wave function sine like. As the confinement is weakened, the 0 and 1 states are

no longer the ideal single-electron wave functions but are modified by the electron–electron interaction. The rate of change of the 0 state is faster than the 1 state which moves roughly parallel with the higher energy levels. In the 1-state, the electrons can describe a highly correlated motion in alternate lobes of the state. The 1- and 0-states both drop in energy as the channel width increases, which also reduces the electron–electron repulsion. The 0 state has the electrons confined near the middle of the channel and does not experience as large reduction in electron repulsion as the 1 state where the electrons are mainly in two parallel rows (Inset, Fig. 2.4b). When the energy levels overlap, the original single electron wavefunctions will be highly distorted forming a hybridised wave function with a greater magnitude located at the edges, corresponding to the formation of two rows and producing a conductance of $4e^2/h$. When the confinement potential is further weakened, the conductance plateau $2e^2/h$ is observed indicating that past the anti-crossing the ground wavefunction comprises a single highly distorted row [40]. The numerical simulation of the charge density distribution in the ground state has been shown to divide into two for a two-electron system in a wide square potential well, which is similar to the case presented here [59]. The closure of the conductance traces on the right side in Fig. 2.4a indicates that the coupling between the split gates and the electron system has increased as would be expected if the electron distribution broadened across the channel. In addition, the density distribution corresponding to each row in the two row situation or the incipient Wigner lattice was imaged using TEF, which showed the first focussing peak splitting into two sub-peaks when the 1D channel entered in the zigzag regime otherwise remained a single peak [60]. A further confirmation of the formation of two distinct rows can be gained by dc source-drain bias measurements.

3.5 Source drain bias across two rows

The linear transport in a 1D system is characterised by small differences $\mu_S - \mu_D$ with respect to the subband spacing, i.e., $\mu_S - \mu_D = eV_{SD}$, where, μ_S, μ_D and V_{SD} are the source chemical potential, drain chemical potential and the applied source drain bias around the Fermi energy, respectively. If V_{SD} is increased so that it becomes comparable to the subband spacing, the transport becomes non-linear, which results in the appearance of additional half integer plateaus [61].

Fig. 2.5 shows conductance measurements performed on a different 1D sample with an applied dc source-drain bias, at a point where the ground and

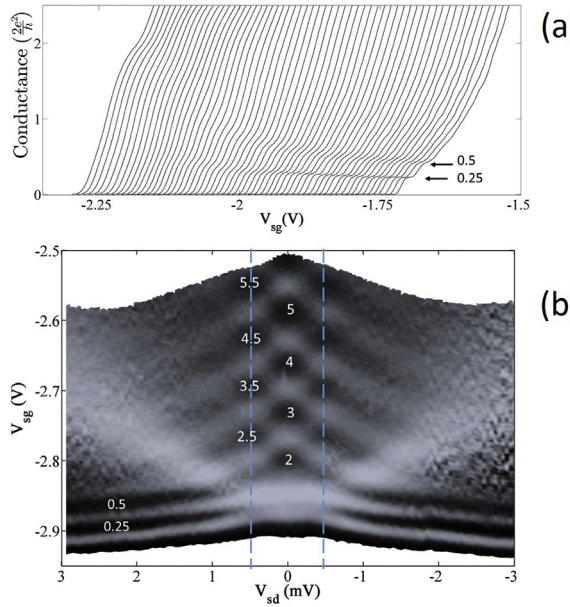


Figure 2.5 Effect of dc source-drain bias on a two row situation in a quasi-one-dimensional channel. (a) Conductance plot for a source-drain bias of 0 to −3 mV with increments of −0.05 mV from left to right. (b) The grayscale plot of dG/dV_{sg} for the data in (a); the white regions are conductance risers and the black regions are conductance plateaus. The plateaus indexing is in units of $2e^2/h$.

first excited states showed an anticrossing. The dc bias was applied across the 1D device between source and drain and was varied from 0 to −3 mV with an increment of −0.05 mV, the split gates were swept for a particular dc bias resulting in a series of conductance plots as shown in Fig. 2.5a. It is observed that there is the usual $0.25(2e^2/h)$ due to the lifting of both momentum and spin degeneracy by the strong dc bias [62,63], which indicates that either the two rows move together initially in the same state or else the source-drain voltage assists one to form before the other [40]. A structure at $0.5(2e^2/h)$ is also observed which is indicative of the addition of the conductance from two separate rows [37,40]. Fig. 2.5b is the grayscale plot of dG/dV_{sg} of the data in (a) as a function of source-drain bias (horizontal axis) and V_{sg} (vertical axis). The $0.25(2e^2/h)$ and $0.5(2e^2/h)$ structures can be traced by looking at the dark regions which extend almost parallel to each other as the source-drain bias is increased. In addition, there are half-integer plateaus indicated by dotted lines which are symmetric about zero source-drain voltage. The plot shows the absence of impurity effects and mesoscopic fluctuations as the system is symmetric about zero source-drain bias.

3.6 Effect of an in-plane magnetic field on two rows

The effect of an in-plane magnetic field of 12 T (along the transport direction) on the conductance characteristics is shown in Fig. 2.6a. Due to the in-plane magnetic field, the spin degeneracy is lifted due to the Zeeman splitting and plateaus appear at integer multiples of Ne^2/h, where $N = 1,2,3 \ldots$ as shown on the left for the strong confinement. The missing plateaus at 6 and 8 arise from the overlap of spin down and spin up levels from different subbands. Assuming the spin down ↓ state to be the lowest spin energy, and adopting the indexation of states as 1↓ (0.5), 1↑(1), 2↓ (1.5), 2↑(2), 3↓(2.5), 3↑(3) …. a merger of 2↑ and 3↓ leads to the missing plateau $4e^2/h$. On weakening the confinement and reducing the carrier density by making the top gate more negative a second merger of 1↑ and

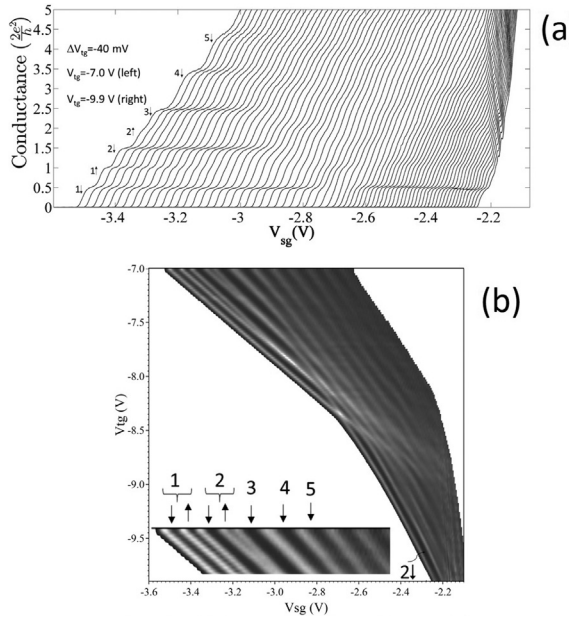


Figure 2.6 (a) Conductance characteristics of a quasi-1D channel from a different cooldown as a function of top gate voltage, V_{tg} from −7.0 (left) to −9.9 V (right) with increment of −40 mV in the presence of an in-plane magnetic field of 12 T. Conductance characteristics showing overlapping of energy levels due to Zeeman splitting, and crossing of the 1↓ state with the higher subbands as the confinement weakens. (b) A greyscale plot of dG/dV_{sg} of the data in (a) where white regions represent the conductance risers and the dark regions are the conductance plateaus. Inset shows a magnified view of the indexed spin states at $V_{tg} = -7.0$ V in the main plot.

$2\downarrow$ is seen at around $V_{sg} \sim -3.1$ V, which is more clearly visible in the greyscale plot in Fig. 2.6b and its inset. As the channel is further widened this merged state remains and is stable. Further weakening of confinement results in a complex overlapping of energy levels. Interestingly, the $1\downarrow$ disappears with the appearance of $2e^2/h$, which in the absence of an in-plane magnetic field would be $4e^2/h$, a spin degenerate two-row state. The result indicates that the two rows are not spin neutral. Fig. 2.6b further reveals that $1\downarrow$ changes at a much slower rate than the higher levels, so cutting across them and giving rise to the pattern of crossing events. Comparing this with the results in the absence of a magnetic field (Fig. 2.4b), shows the two processes are different due to level crossings imposed by the Zeeman splitting [40]. The increasing mixing as levels move together also explains why a new plateau emerges from one plateau then drops in value before settling at the next quantised value. A significant feature is the disappearance of the $2e^2/h$ plateau and the reintroduction of e^2/h as the ground state; this arises from the $2\downarrow$ dropping through many states to become the new ground state, as shown in Fig. 2.6b. This is a surprising manifestation of the modification of the energy levels by the electron interaction. The $1\downarrow$ level passes through several higher levels during its trajectory producing a series of crossing/anti-crossings with higher levels [40].

3.7 Non-magnetic fractional states

Haldane first suggested using the tight-binding model on a honeycomb lattice that Landau levels were not necessary for the observation of a quantised Hall conductance [15]. He showed that the existence of a quantum Hall state does not necessarily depend upon the presence of external magnetic field, but on the symmetries of the system and its topological phases, and since then there have been various discoveries including the anomalous Hall effect [64], spin Hall effect [65,66], etc., in various material systems including topological insulators [67]. Recently, it was shown that in the absence of a magnetic field, when a quasi-1D system is allowed to reconfigure into a relaxed state such that the 1D electrons have access to the second dimension, a variety of fractional conductance states emerge as a function of increasing asymmetry in the confinement potential [38,41]. It is now well understood that the IQHE/FQHE are due to states which are topologically protected from impurity scattering [15,18], unlike 1D systems which are very sensitive to impurity effects [68,69]. Therefore, materials of highest quality would be required if fractional states were to be seen in 1D quantum

wires. In this section, experimental results showing the recently discovered fractional conductance quantisation in 1D GaAs channels are presented.

The zigzag phase, as discussed previously, occurs in a weakly confined regime and is very sensitive to the confinement potential and changes in the kinetic energy of the relaxed electrons, which may result in non-trivial characteristics if any of the parameters are perturbed. One possible way to perturb the confinement is to alter the symmetry of the confinement potential by applying an offset to one of the split gates [41]. Fig. 2.7a shows the results obtained as a function of increasing asymmetry in the confinement potential, which is achieved by increasing an offset voltage, ΔV_{sg} between the split gates. The top gate voltage, V_{tg}, was kept constant at -0.53 V, and the initial 2D carrier concentration was 2×10^{10} cm^{-2}. For comparison, experiments were also performed in the presence of an in-plane magnetic field, B_{II} of 10 T applied in the plane of the 2D electron gas and normal to the current direction. In Fig. 2.7a, the asymmetry in confinement potential is gradually increased until the total offset across the device is -245 mV; the first trace on the right is for a symmetric confinement potential where there was no structure apart from 3/2, i.e., $0.7(2e^2/h)$. As the asymmetry and width of the 1D channel

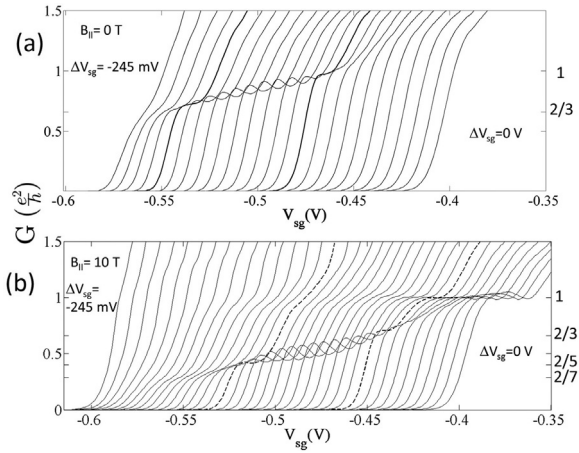


Figure 2.7 Effect of asymmetry in the confinement potential on the conductance characteristics. (a) At $B_{II} = 0$ T, plateaus at 1 and 2/3 were observed as a function of increasing asymmetry, ΔV_{sg} (0 V, right, symmetric potential) to -245 mV (left) in steps of -10 mV; $V_{tg} = -0.53$ V. (b) At $B_{II} = 10$ T, the system is fully spin polarised, the previous 1 now becomes the 2/3, and the previous 2/3 [observed in (a)] is now the 2/5, with a further hint of 2/7 when the channel is fully widened. *Adapted from [41]. Reproduced with permission from AIP.*

increase, the system becomes spin polarised with a plateau forming close to 1 (trace shown in thick black). On further increasing asymmetry, the plateau at 1 becomes unstable and drops down before settling at the 2/3. When the confinement potential has almost reached an asymmetry of −245 mV, the previously settled 2/3 plateau now weakens and turns into a featureless transition to 2D conduction. In the presence of an in-plane magnetic field of 10 T, the system is already spin polarised, which is evident due to the presence of a plateau at 1 on the right in (b). On moving towards the left, the feature at 1 divides into two, one feature stays at 1 and the other drops to form a flat plateau at 2/3, which was seen in the absence of B_{II} in (a) indicating 2/3 is a spin-polarised fractional state. On further increasing the asymmetry by increasing the offset voltage, the feature at 2/3 undergoes an instability as before and then stabilises at 2/5. On further increasing the asymmetry, the strength of the feature weakens and in a very weak confinement regime a shoulder appears near 2/7 before the result becomes featureless indicating a transition to 2D conduction.

The zigzag is the precursor for the onset of fractional states in a 1D system which could allow a 'ring exchange' which has been suggested previously as an explanation for the FQHE [49,70,71]. In the zigzag arrangement, which is essentially a two-row situation, the transverse localisation of the electrons contributes to the energy of the system. The energy of the system can be reduced by the increase in spatial freedom given by a cyclic rotation which appears to create the conductance fractionalisation [38]. A systematic increase in asymmetry in confinement potential could result in the formation of different rings, each of which forms a different fractional state. We think the ring formation could also allow tunnelling between different phases which may be a possible reason for the instability in the 2/3 fractional state as the confinement was altered. It was recently suggested that the appearance of the 2/5 fractional state arises from coherent backscattering in the strongly interacting 1D channel [72].

3.8 Effect of the variation of carrier density

Fig. 2.8 illustrates the effect of changing carrier density within the 1D channel at a fixed in-plane magnetic field of 10 T normal to the current direction. A fixed offset on the split gate voltage produces the asymmetric, wide confinement necessary for the observation of fractional states. The top gate voltage is changed from −0.39 V (left) to −0.62 V (right) in steps of −10 mV. A weak structure on the left at around 3/5 drops in value until it

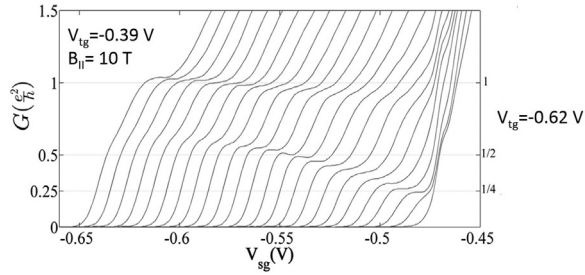


Figure 2.8 Effect of variation of carrier concentration on the fractional conductance in the presence of $B_{||} = 10$ T, V_{tg} was varied from -0.45 (left) to -0.62 V (right) in steps of -10 mV; $\Delta V_{sg} = -180$ mV. The fractional states at 1/2 and 1/4 are observed as the carrier density is further reduced. *Adapted from the Open Access [38].*

flattens on either side of a conductance value of 1/2. The traces on either side of the indexed 1/2 exhibit flat plateau in the range 6/11–5/11, suggesting the system is attempting to stabilise. On further reducing the carrier concentration, the structure drops in value and flattens at around 1/4 before smearing out indicating a transition to 2D behaviour. The appearance of flat plateaus on either side of the 1/4 between 3/10 and 6/25 indicates that there is a band of stable states in the vicinity of the 1/4 in a similar way to the states formed on either side of the 1/2. In the FQHE, the 5/2 state is of particular interest due to its potential for non-Abelian statistics, and the effective electronic charge associated with it has been estimated to be 1/4 [73,74]. The similarity in the behaviour of the structures near the 1/2 and 1/4 values found here indicates that they may be linked to the proposed non-Abelian nature of the 1/4, in that the differing details of the exchange path provide distinct states that are degenerate. The differences due to the small changes in the confinement provide a different exchange path; the change in confinement energy alters the level energy to provide the band around the 1/4. Similar considerations apply to the formation of the band around the 1/2, particularly if this arises from pairs of 1/4 states [38,73].

3.9 Effect of an in-plane magnetic field

The nature of the fractional plateau at 1/6 changes as a function of the in-plane magnetic field. This is shown in Fig. 2.9, where both the 2D density and asymmetry in confinement are held constant. The structure on the left at 1/6 at zero magnetic field drops in value as the magnetic field is gradually increased. During its downward trajectory, the original 1/6 fractional state

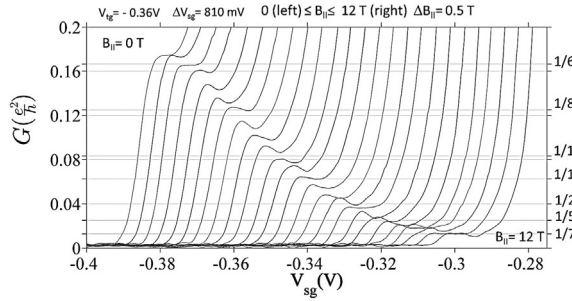


Figure 2.9 Effect of variation of in-plane magnetic field on the conductance characteristics in the fractional regime. $B_{||}$ was varied from 0 (left) to 12 T (right) in steps of 0.5 T; $V_{tg} = -0.36$ V; $\Delta V_{sg} = 800$ mV.

passes through different fractional conductance regimes and on several occasions, it acquires a flat plateau, in other places, it has a structure with a gradient. The fractional states which are flat and with minimal gradient occur in the vicinity of 1/6, 1/8, 1/12, 1/16 and 1/25. It was noted that when the applied B was more than 8 T, that much smaller in value but wider conductance plateaus, are observed as if the square of a fractional state (i.e., 1/25, 1/40, 1/50 and 1/78) is involved [47,72,75]. The continuous movement of the structure with in-plane magnetic field implies that it may not be related to spin but, rather, is dependent on flux quanta like the composite fermions [38,76].

3.10 Non-linear transport with fractional states

As discussed in Section 3.5, the non-linear transport gives rise to the half integer plateaus in a typical non-interacting 1D system when the applied source-drain bias is comparable to the subband spacing [61]. The situation is expected to be different in the case of fractional states observed in a quasi-1D system given this is a new system based on many body effect within the 1D channel. Fig. 2.10 shows dc source-drain bias measurements performed on the fractional states, 2/5 and 1/4 [38]. Fig. 2.10a shows a conductance plateau at 2/5 at $V_{sd} = 0$ mV (right); on increasing the source-drain bias, the structure eventually flattens and settles at 1/5. This result, where the plateau value is halved, seems similar to the normal integer case [61] and indicating that fractional band conduction is occurring as opposed to a fractional transmission coefficient. Also, as the differential conductance decreased towards the 1/5 with increasing source-drain bias, a plateau emerged at 1/4 perhaps indicating that another state was attempting to

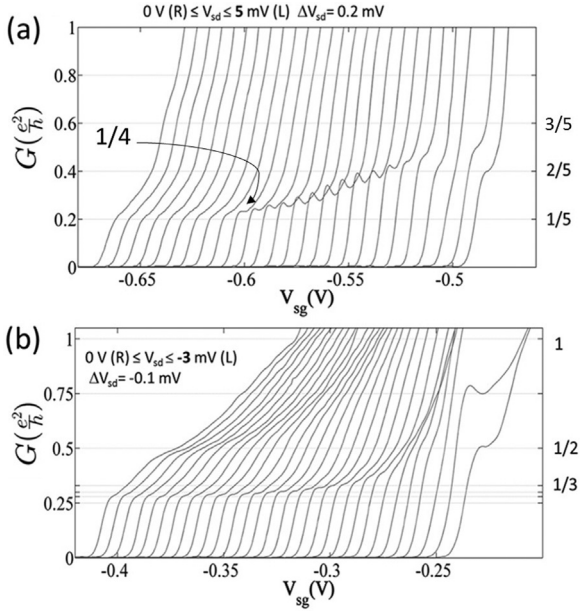


Figure 2.10 (a) Effect of dc source-drain bias on the fractional conductance at 2/5, V_{sd} was incremented from 0 (right) to 5 mV (left) in steps of 0.2 mV; $\Delta V_{sg} = 955$ mV, $V_{tg} = -0.36$ V. (b) Investigation of the effect of dc bias of the 1/2 structure; V_{sd} was incremented from 0 (right) to −3 mV (left) in steps of −0.1 mV; $\Delta V_{sg} = 0$ mV, $V_{tg} = -0.62$ V. *Adapted from the Open Access [38].*

form as the conducting subband was pulled further down below the source chemical potential and the carrier concentration increased. It may be noted that a structure at 3/5 is also forming at large source-drain bias in addition to the 1/5 state. For the 2/5, on application of large source-drain bias, it is seen that two new fractional states at 3/5 and 1/5 are formed. Being a highly asymmetrically defined 1D system, it is possible that the voltage would not drop symmetrically across the saddle point as generally is the case with adiabatic transport so the expected 1/2 drop of the source-drain bias would not occur. It does appear though a voltage drop of 1/5 is probably causing the conducting band to split into 3/5 (2/5 + 1/5) and 1/5(2/5−1/5) when the source-drain bias is increased. Fig. 2.10b shows the differential conductance of a fractional plateau at 1/2 in a symmetrical confinement channel. The differential conductance drops, as expected, with increasing source drain bias but attempts to flatten and form a plateau at 1/3 followed by a flattening at 3/10 with intervening structure. The structure always remains slightly above the value of 1/4 produced by the lifting of the momentum

degeneracy [38]. On application of large source-drain bias, a plateau close to 1/4 is seen along with 1/2. This could arise from the fact that the 1/2 state has been suggested as a pairing of two 1/4 states, one of which is formed as a separate conducting band on applying source-drain bias, so resulting in both 1/2 and 1/4 states at large source drain bias. It may be argued that the 1/2 state at large source drain bias could be the usual $0.25(2e^2/h)$ state seen generally in a typical 1D system due to the lifting of spin and momentum degeneracy, however, the absence of such feature in (a) possibly rules out this option.

In conclusions, the fractional quantum states in a 1D system described here have implications for furthering fundamental physics of low-dimensional electron systems and the emerging quantum technologies. The results show the emergence of a wide variety of non-magnetic fractional states in a weakly confined, wider 1D channel which will enable further development of new experimental platforms as well as investigations on topological properties of the system. The recent developments in the field presented here would be of interest to multidisciplinary research community, and those involved in developing schemes for future quantum computation and spintronics.

Acknowledgements

We acknowledge the Engineering and Physical Sciences Research Council (EPSRC) and a United Kingdom Research and Innovation (UKRI) Future Leaders Fellowship (Reference: MR/S015728/1) for financial support. We thank our collaborators Prof David A. Ritchie, Prof S Bose, Prof G Gumbs, the late Prof K Berggren, Dr I Farrer, Dr P See, Dr C Yan, Prof K J Thomas, Dr H Montagu and Dr T-M Chen for many fruitful discussions.

References

[1] Berggren K, Pepper M. New directions with fewer dimensions. Phys World 2002; 15(10):37.
[2] Thouless D. Maximum metallic resistance in thin wires. Phys Rev Lett 1977;39(18): 1167.
[3] Dean C, Pepper M. The transition from two-to one-dimensional electronic transport in narrow silicon accumulation layers. J Phys C Solid State Phys 1982;15(36):L1287.
[4] Klitzing Kv, Dorda G, Pepper M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. Phys Rev Lett 1980; 45(6):494.
[5] Tsui DC, Stormer HL, Gossard AC. Two-dimensional magnetotransport in the extreme quantum limit. Phys Rev Lett 1982;48(22):1559–62.
[6] Thornton TJ, Pepper M, Ahmed H, Andrews D, Davies GJ. One-dimensional conduction in the 2D electron gas of a GaAs-AlGaAs heterojunction. Phys Rev Lett 1986;56(11):1198.

[7] Wharam D, Thornton T, Newbury R, Pepper M, Ahmed H, Frost J, Hasko D, Peacock D, Ritchie D, Jones G. One-dimensional transport and the quantisation of the ballistic resistance. J Phys C Solid State Phys 1988;21(8):L209.
[8] Van Wees B, Van Houten H, Beenakker C, Williamson JG, Kouwenhoven L, Van der Marel D, Foxon C. Quantized conductance of point contacts in a two-dimensional electron gas. Phys Rev Lett 1988;60(9):848.
[9] Smith C, Pepper M, Ahmed H, Frost J, Hasko D, Peacock D, Ritchie D, Jones G. The transition from one-to zero-dimensional ballistic transport. J Phys C Solid State Phys 1988;21(24):L893.
[10] de-Picciotto R, Reznikov M, Heiblum M, Umansky V, Bunin G, Mahalu D. Direct observation of a fractional charge. Nature 1997;389(6647):162–4.
[11] Field M, Smith CG, Pepper M, Ritchie DA, Frost JEF, Jones GAC, Hasko DG. Measurements of Coulomb blockade with a noninvasive voltage probe. Phys Rev Lett 1993;70(9):1311–4.
[12] Wharam DA, Pepper M, Newbury R, Ahmed H, Hasko DG, Peacock DC, Frost JEF, Ritchie DA, Jones GAC. Observation of Aharonov-Bohm oscillations in a narrow two-dimensional electron gas. J Phys Condens Matter 1989;1(21):3369–73.
[13] Meirav U, Kastner MA, Wind SJ. Single-electron charging and periodic conductance resonances in GaAs nanostructures. Phys Rev Lett 1990;65(6):771–4.
[14] Laughlin RB. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. Phys Rev Lett 1983;50(18):1395.
[15] Haldane FDM. Model for a quantum Hall effect without Landau levels: condensed-matter realization of the "parity anomaly". Phys Rev Lett 1988;61(18):2015–8.
[16] Roy R, Sondhi SL. Fractional quantum Hall effect without Landau levels. Physics 2011;4:46.
[17] Trugman S. Localization, percolation, and the quantum Hall effect. Phys Rev B 1983; 27(12):7539.
[18] Neupert T, Santos L, Chamon C, Mudry C. Fractional quantum Hall states at zero magnetic field. Phys Rev Lett 2011;106(23):236804.
[19] Girvin SM, MacDonald AH, Platzman PM. Collective-excitation gap in the fractional quantum Hall effect. Phys Rev Lett 1985;54(6):581–3.
[20] Tarucha S, Honda T, Saku T. Reduction of quantized conductance at low temperatures observed in 2 to 10 μm-long quantum wires. Solid State Commun 1995;94(6): 413–8.
[21] Thomas K, Nicholls J, Simmons M, Pepper M, Mace D, Ritchie D. Possible spin polarization in a one-dimensional electron gas. Phys Rev Lett 1996;77(1):135.
[22] Thomas KJ, Nicholls JT, Appleyard NJ, Simmons MY, Pepper M, Mace DR, Tribe WR, Ritchie DA. Interaction effects in a one-dimensional constriction. Phys Rev B 1998;58(8):4846–52.
[23] Thomas KJ, Nicholls JT, Pepper M, Tribe WR, Simmons MY, Ritchie DA. Spin properties of low-density one-dimensional wires. Phys Rev B 2000;61(20):R13365–8.
[24] Yan C, Kumar S, Pepper M, See P, Farrer I, Ritchie D, Griffiths J, Jones G. Temperature dependence of spin-split peaks in transverse electron focusing. Nanoscale Res Lett 2017;12(1):553.
[25] Yan C, Kumar S, Thomas K, See P, Farrer I, Ritchie D, Griffiths J, Jones G, Pepper M. Engineering the spin polarization of one-dimensional electrons. J Phys Condens Matter 2018;30(8):08LT01.
[26] Yan C, Kumar S, Thomas KJ, Pepper M, See P, Farrer I, Ritchie D, Griffiths J, Jones G. Direct observation of exchange-driven spin interactions in one-dimensional system. Appl Phys Lett 2017;111:042107.

[27] Yan C, Kumar S, Thomas K, See P, Farrer I, Ritchie D, Griffiths J, Jones G, Pepper M. Coherent spin amplification using a beam splitter. Phys Rev Lett 2018;120(13): 137701.
[28] Jaksch P, Yakimenko I, Berggren K-F. From quantum point contacts to quantum wires: density-functional calculations with exchange and correlation effects. Phys Rev B 2006;74(23):235320.
[29] Berggren K-F, Yakimenko I. Effects of exchange and electron correlation on conductance and nanomagnetism in ballistic semiconductor quantum point contacts. Phys Rev B 2002;66(8):085323.
[30] Meir Y, Hirose K, Wingreen NS. Kondo model for the "0.7 anomaly" in transport through a quantum point contact. Phys Rev Lett 2002;89(19):196802.
[31] Chen T-M, Graham A, Pepper M, Farrer I, Ritchie D. Non-Kondo zero-bias anomaly in quantum wires. Phys Rev B 2009;79(15):153303.
[32] Sfigakis F, Ford CJB, Pepper M, Kataoka M, Ritchie DA, Simmons MY. Kondo effect from a tunable bound state within a quantum wire. Phys Rev Lett 2008;100(2): 026807.
[33] Sarkozy S, Sfigakis F, Das Gupta K, Farrer I, Ritchie DA, Jones GAC, Pepper M. Zero-bias anomaly in quantum wires. Phys Rev B 2009;79(16):161307.
[34] Aryanpour K, Han JE. Ferromagnetic spin coupling as the origin of 0.7 anomaly in quantum point contacts. Phys Rev Lett 2009;102(5):056805.
[35] Bauer F, Heyder J, Schubert E, Borowsky D, Taubert D, Bruognolo B, Schuh D, Wegscheider W, von Delft J, Ludwig S. Microscopic origin of the '0.7-anomaly' in quantum point contacts. Nature 2013;501(7465):73–8.
[36] Hew W, Thomas K, Pepper M, Farrer I, Anderson D, Jones G, Ritchie D. Incipient formation of an electron lattice in a weakly confined quantum wire. Phys Rev Lett 2009;102(5):056804.
[37] Smith L, Hew W, Thomas K, Pepper M, Farrer I, Anderson D, Jones G, Ritchie D. Row coupling in an interacting quasi-one-dimensional quantum wire investigated using transport measurements. Phys Rev B 2009;80(4):041306.
[38] Kumar S, Pepper M, Holmes S, Montagu H, Gul Y, Ritchie D, Farrer I. Zero-magnetic field fractional quantum states. Phys Rev Lett 2019;122(8):086803.
[39] Kumar S, Thomas K, Smith L, Pepper M, Farrer I, Ritchie D, Jones G, Griffiths J. Effect of low transverse magnetic field on the confinement strength in a quasi-1D wire. AIP Conf Proc 2013;1566:245.
[40] Kumar S, Thomas KJ, Smith LW, Pepper M, Creeth GL, Farrer I, Ritchie D, Jones G, Griffiths J. Many-body effects in a quasi-one-dimensional electron gas. Phys Rev B 2014;90(20):201304.
[41] Kumar S, Pepper M, Ritchie D, Farrer I, Montagu H. Formation of a non-magnetic, odd-denominator fractional quantized conductance in a quasi-one-dimensional electron system. Appl Phys Lett 2019;115(12):123104.
[42] Klironomos A, Meyer JS, Hikihara T, Matveev K. Spin coupling in zigzag Wigner crystals. Phys Rev B 2007;76(7):075302.
[43] Mehta AC, Umrigar CJ, Meyer JS, Baranger HU. Zigzag phase transition in quantum wires. Phys Rev Lett 2013;110(24):246802.
[44] Meyer JS, Matveev K. Wigner crystal physics in quantum wires. J Phys Condens Matter 2009;21(2):023203.
[45] Gul Y, Holmes SN, Myronov M, Kumar S, Pepper M. Self-organised fractional quantisation in a hole quantum wire. J Phys Condens Matter 2018;30(9):09LT01.
[46] Gumbs G, Huang D, Hon J, Pepper M, Kumar S. Tunneling of hybridized pairs of electrons through a one-dimensional channel. Adv Phys X 2017;2(3):545–68.
[47] Shavit G, Oreg Y. Electron pairing induced by repulsive interactions in tunable one-dimensional platforms. Phys Rev Res 2020;2:043283.

[48] Greiter M, Wen X-G, Wilczek F. Paired hall states. Nucl Phys B 1992;374(3):567–614.
[49] Kane CL, Stern A, Halperin BI. Pairing in luttinger liquids and quantum Hall states. Phys Rev X 2017;7(3):031009.
[50] Rokhinson LP, Larkina V, Lyanda-Geller YB, Pfeiffer LN, West KW. Phys Rev Lett 2004;93:146601.
[51] Rokhinson LP, Pfeiffer LN, West KW. Phys Rev Lett 2006;96:156602.
[52] Dedigama AR, Deen D, Murphy SQ, Goel N, Keay JC, Santos MB, Suzuki K, Miyashita S, Hirayama Y. Current focusing in InSb heterostructures. Phys E Low-dimens Syst Nanostruct 2006;34(1):647–50.
[53] Lo S-T, Chen C-H, Fan J-C, Smith LW, Creeth GL, Chang C-W, Pepper M, Griffiths JP, Farrer I, Beere HE, Jones GAC, Ritchie DA, Chen T-M. Controlled spatial separation of spins and coherent dynamics in spin-orbit-coupled nanostructures. Nat Commun 2017;8(1):15997.
[54] Reynoso A, Usaj G, Sánchez MJ, Balseiro CA. Phys Rev B 2004;70:235344.
[55] Kawamura M, Ono K, Stano P, Kono K, Aono T. Electronic magnetization of a quantum point contact measured by nuclear magnetic resonance. Phys Rev Lett 2015; 115(3):036601.
[56] Córcoles A, Ford CJB, Pepper M, Jones GAC, Beere HE, Ritchie DA. Nuclear spin coherence in a quantum wire. Phys Rev B 2009;80(11):115326.
[57] Reynoso A, Usajand G, Balseiro CA. Phys Rev B 2007;75:085321.
[58] Graham A, Thomas K, Pepper M, Cooper N, Simmons M, Ritchie D. Interaction effects at crossings of spin-polarized one-dimensional subbands. Phys Rev Lett 2003; 91(13):136404.
[59] Bryant GW. Electronic structure of ultrasmall quantum-well boxes. Phys Rev Lett 1987;59(10):1140–3.
[60] Ho S-C, Chang H-J, Chang C-H, Lo S-T, Creeth G, Kumar S, Farrer I, Ritchie D, Griffiths J, Jones G, Pepper M, Chen T-M. Imaging the zigzag wigner crystal in confinement-tunable quantum wires. Phys Rev Lett 2018;121(10):106801.
[61] Patel NK, Martin-Moreno L, Pepper M, Newbury R, Frost JEF, Ritchie DA, Jones GAC, Janssen JTMB, Singleton J, Perenboom JAAJ. Ballistic transport in one dimension: additional quantisation produced by an electric field. J Phys Condens Matter 1990;2(34):7247–54.
[62] Chen T-M, Graham A, Pepper M, Farrer I, Anderson D, Jones G, Ritchie D. Direct observation of nonequilibrium spin population in quasi-one-dimensional nanostructures. Nano Lett 2010;10(7):2330–4.
[63] Chen T-M, Graham A, Pepper M, Farrer I, Ritchie D. Bias-controlled spin polarization in quantum wires. Appl Phys Lett 2008;93(3):032102.
[64] Chang C-Z, Zhang J, Feng X, Shen J, Zhang Z, Guo M, Li K, Ou Y, Wei P, Wang L-L, Ji Z-Q, Feng Y, Ji S, Chen X, Jia J, Dai X, Fang Z, Zhang S-C, He K, Wang Y, Lu L, Ma X-C, Xue Q-K. Experimental observation of the quantum anomalous Hall effect in a magnetic topological insulator. Science 2013;340(6129):167–70.
[65] Bernevig BA, Zhang S-C. Quantum spin Hall effect. Phys Rev Lett 2006;96(10): 106802.
[66] Kato Y, Myers R, Gossard A, Awschalom D. Science 2004;306:1910.
[67] Fu L, Kane CL. Topological insulators with inversion symmetry. Phys Rev B 2007; 76(4):045302.
[68] Smith LW, Thomas KJ, Pepper M, Ritchie DA, Farrer I, Griffiths JP, Jones GAC. Disorder and interaction effects in quantum wires. J Phys Conf Ser 2012;376:012018.
[69] Williamson JG, Timmering CE, Harmans CJPM, Harris JJ, Foxon CT. Quantum point contact as a local probe of the electrostatic potential contours. Phys Rev B 1990;42(12):7675–8.

[70] Kivelson S, Kallin C, Arovas DP, Schrieffer JR. Cooperative ring exchange and the fractional quantum Hall effect. Phys Rev B 1987;36(3):1620–46.
[71] Thouless DJ, Li Q. Ring exchange and the fractional quantum Hall effect. Phys Rev B 1987;36(8):4581–3.
[72] Shavit G, Oreg Y. Fractional conductance in strongly interacting 1D systems. Phys Rev Lett 2019;123(3):036803.
[73] Willett RL. The quantum Hall effect at 5/2 filling factor. Rep Prog Phys 2013;76(7): 076501.
[74] Radu IP, Miller JB, Marcus CM, Kastner MA, Pfeiffer LN, West KW. Quasi-particle properties from tunneling in the 5/2 fractional quantum Hall state. Science 2008; 320(5878):899–902.
[75] Oreg Y, Sela E, Stern A. Fractional helical liquids in quantum wires. Phys Rev B 2014; 89(11):115402.
[76] Jain JK. Microscopic theory of the fractional quantum Hall effect. Adv Phys 1992; 41(2):105–46.

CHAPTER THREE

Semiconductor nanodevices as a probe of strong electron correlations

Pedro Vianez[1], Oleksandr Tsyplyatyev[2] and Christopher Ford[1,*]
[1]Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
[2]Institut für Theoretische Physik, Universität Frankfurt, Frankfurt, Germany
*Corresponding author: cjbf@cam.ac.uk

1. The failure of Fermi liquid theory

Many-body interactions of electrons in three or two dimensions can, generally speaking, be treated using Fermi liquid theory, proposed by Lev Landau in 1956 [1,2]. Landau showed that, despite the enormous strength of electron–electron interactions, these systems could be described as collective excitations that behave in almost every way very similarly to free electrons, but have different effective masses. In Fermi liquid theory, electrons can therefore be modelled as if being part of a noninteracting system.

Let us consider a so-called 'normal' Fermi liquid. Here, the system is in a ground state when all interactions are switched off, the states being filled according to Fermi–Dirac statistics. This forms a well-defined Fermi surface at low temperatures. In the noninteracting regime, an excited state can then be represented by a filled Fermi circle or sphere, corresponding to an equilibrium state of zero momentum, together with an additional particle of momentum p. As interactions are switched on, conservation of the total momentum enforces conservation of the momentum p of the same excited state, even as the additional particle moves across the system causing a perturbation. Landau proved that this separate entity – which he called a quasiparticle, composite of contributions from many interacting particles – could move around and only very weakly interact with its surroundings, essentially as if interactions were absent.

Non-Fermi liquids mark the departure from this view. For an in-depth review of the field, see Ref. [3]. In general, however, there are two scenarios where Fermi liquid theory usually fails:

1. In higher dimensions, such as two- (2D) or three-dimensional (3D) systems, if the electron liquid is highly correlated, for instance, when

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00007-1
© 2021 Elsevier Ltd.
All rights reserved.

the energy of the Coulomb interaction greatly exceeds the kinetic energy;
2. In lower dimensions, like one-dimensional (1D) wires, where spatial confinement alone dictates strong correlations regardless of the strength of the interactions in question.

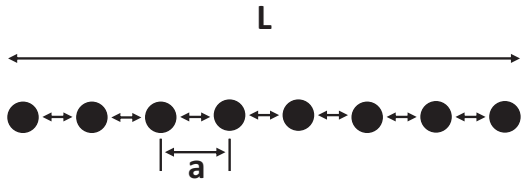


Figure 3.1 Schematic representation of a one-dimensional electron system with density per unit length $N/L = 1/a$, where L is the length of the system and a is the interparticle spacing.

In this chapter, we focus on the latter and on how tunnelling spectroscopy can be used to experimentally probe such correlations. When confined to a 1D geometry (see Fig. 3.1), electrons cannot be modelled as Fermi quasi-particles, since each one is now strongly affected by its two neighbours. In other words, spatial confinement alone dictates that they are unable to go over, under or around each other, the only option left being to try to go through, were this not prevented by the diverging Coulomb repulsion at short distances. Hence, strong correlations arise.

One-dimensional wires mark a drastic departure from their higher-dimensional counterparts as, unlike the latter, there is no longer such a thing as locality of interactions. Instead, the motion of one electron causes a collective response of the whole system. This is why the single-particle properties, on which Fermi liquid theory rests, break down completely. What emerges instead are the hydrodynamic modes of the Tomonaga–Luttinger liquid (TLL), with bosonic instead of fermionic statistics. We start this chapter by introducing the main ideas behind the TLL model, which replaces Fermi liquid theory for electron liquids in one dimension, together with its main predictions.

1.1 The Tomonaga–Luttinger liquid model

Many attempts have been made to describe strongly correlated systems, especially in one and two dimensions. Whilst at first sight, perturbation theory may seem like the obvious approach, it is important to note that any such theory would only work under the condition of weak interactions.

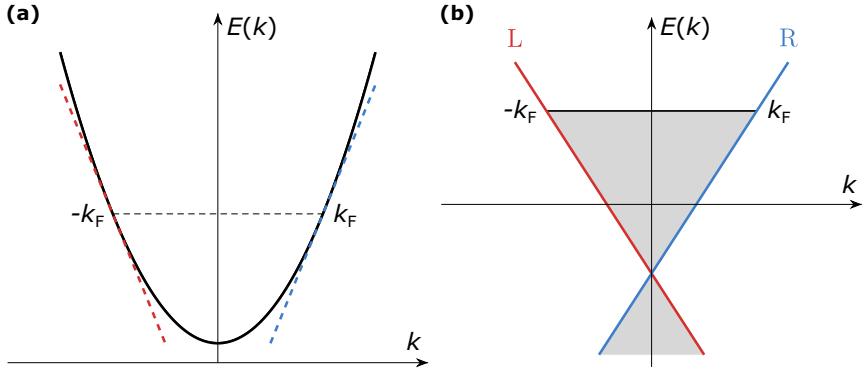


Figure 3.2 The Tomonaga–Luttinger liquid (TLL) model. (a) The original energy dispersion relation is shown in black. For electrons in free space, this is parabolic. According to the TLL picture, this can be approximated by a linear dispersion close to the Fermi points (red [grey in print version] and blue [grey in print version] *dashed lines*). (b) The linear dispersion used in the TLL model. Red [grey in print version] and blue [grey in print version] curves correspond, respectively, to left- and right-moving electrons, whilst the shaded area marks the filled Dirac sea below the Fermi energy. *From [10], reprinted with permission.*

Proposed by Tomonaga in 1950 [4] and later refined by Luttinger in 1963 [5], the TLL model was introduced as an attempt to describe interacting 1D electron systems at arbitrary interaction strength. The correct solution of the model was first obtained by Mattis and Lieb in 1965 [6] and later improved by Luther and Peschel in 1974 [7], before being finalised by Haldane in 1981 [8]. The idea was that by using Bloch's theory of sound waves, electron excitations could be described using a Bose field, effectively laying the groundwork for what today is known as the method of bosonisation.

The Tomonaga–Luttinger liquid model is used to describe highly correlated interacting 1D systems. It was initially developed as a simple spinless theory but has since been extended to include spin [9]. Although exactly solvable, it starts from two key assumptions, that the dispersion is linear near the Fermi points and that the 1D system is infinitely long, see Fig. 3.2. We know, however, that the dispersion of free electrons, given by $E_k = \hbar^2 k^2/(2m^*)$, is parabolic. Nevertheless, when low-energy excitations occur, electrons can only move from just below to just above the Fermi energy and so, under these conditions, even though the true dispersion is in fact slightly curved (i.e., nonlinear), one can argue approximate linearity. Similarly, the assumption of infinitely long wires can be justified as long as end effects are not expected to play a significant role. A final assumption is that the wire is infinitely narrow, i.e., that the subband spacing is infinite, so that there is no mixing with higher 1D subbands. This makes it hard to be sure whether the

theory is applicable to any particular experimental observation (where subband spacing is never huge compared with the interaction energy).

In three or two dimensions, low-energy excitations with arbitrarily low momenta can occur [9]. This is because in both 3D and 2D the Fermi surface in k-space can be modelled as a sphere or a circle, respectively, with radius k_F, and electrons are able to occupy any state within it. In 1D systems, however, this is no longer the case, as the Fermi surface is now not only discontinuous but also consists of only two specific points, at k_F and $-k_F$. Within the TLL assumptions, one can write the Hamiltonian of a 1D interacting electron system as

$$\widehat{H} = \widehat{H}_0 + \widehat{H}_{\text{int}}. \tag{3.1}$$

Here, $\widehat{H}_0$ represents the kinetic energy, whilst $\widehat{H}_{\text{int}}$ refers to the interaction energy of the system. They can be written, respectively, as

$$\widehat{H}_0 = \sum_{k,\alpha=\pm 1} v_F(\alpha k - k_F)\widehat{a}^{\dagger}_{k,\alpha}\widehat{a}_{k,\alpha}, \tag{3.2}$$

where $v_F = \hbar k_F/m$ is the Fermi velocity, $\widehat{a}^{\dagger}_{k,\alpha}$ and $\widehat{a}_{k,\alpha}$ are the creation and annihilation operators at the right $(\alpha = 1)$ or left $(\alpha = -1)$ branches of the dispersion, and

$$\widehat{H}_{\text{int}} = \frac{1}{2L}\sum_{q\neq 0} v_q \widehat{\rho}_{-q}\widehat{\rho}_q, \tag{3.3}$$

where $\widehat{\rho}_q$ is the electron density fluctuation operator and v_q is the Fourier transform of the interaction potential. Together, $\widehat{H}$ accounts for particles moving in either of the right (R) and left (L) branches, as well as for inter- and intra-branch interactions [11].

Generally speaking, it is $\widehat{H}_{\text{int}}$ that characterises the type of interactions. For a 1D electron system with spacing a, these will be Coulomb in nature and of two types (see Fig. 3.3a): if an electron is excited within the same branch (left or right) from a state just below the Fermi energy, then only a small change in momentum is required, $q \sim 0$, with the direction of motion remaining unchanged. This is called forward scattering. On the other hand, if there is an exchange involving both branches, then a larger change in momentum will be necessary, $q \sim 2k_F$, with the direction of motion now being reversed. This constitutes backward scattering. All of these processes appear naturally from the density fluctuation operator when the particles in Fig. 3.3a are split into the left and the right moving subbands around the $\pm k_F$ points,

$$\widehat{\rho}_q \simeq \begin{cases} \sum_k (\widehat{a}^\dagger_{k-q,\mathrm{R}} \widehat{a}_{k,\mathrm{R}} + \widehat{a}^\dagger_{k-q,\mathrm{L}} \widehat{a}_{k,\mathrm{L}}) & \text{for } |q| \sim 0, \\ \sum_k \widehat{a}^\dagger_{k-q,\mathrm{L}} \widehat{a}_{k,\mathrm{R}} & \text{for } q \sim 2k_\mathrm{F}, \\ \sum_k \widehat{a}^\dagger_{k-q,\mathrm{R}} \widehat{a}_{k,\mathrm{L}} & \text{for } q \sim -2k_\mathrm{F}. \end{cases} \tag{3.4}$$

By substituting Eq. (3.4) into Eq. (3.3), the interaction term in the Hamiltonian can be divided into two terms depending on the direction of motion of the interacting electrons. We then take the long-wavelength limit, i.e., by considering a system of infinite length or, in other words, by assuming arbitrarily high excitation energies, finding that the total TLL Hamiltonian is given by

$$\widehat{H}_\mathrm{TLL} = \widehat{H}_0 + \frac{1}{2L}\sum_{q\neq 0} V_1(q)(\widehat{\rho}_{-q,\mathrm{R}}\,\widehat{\rho}_{q,\mathrm{R}} + \widehat{\rho}_{-q,\mathrm{L}}\,\widehat{\rho}_{q,\mathrm{L}}) + \frac{1}{2L}\sum_{q\neq 0} V_2(q)(\widehat{\rho}_{-q,\mathrm{R}}\,\widehat{\rho}_{q,\mathrm{L}} + \widehat{\rho}_{-q,\mathrm{L}}\,\widehat{\rho}_{q,\mathrm{R}}), \tag{3.5}$$

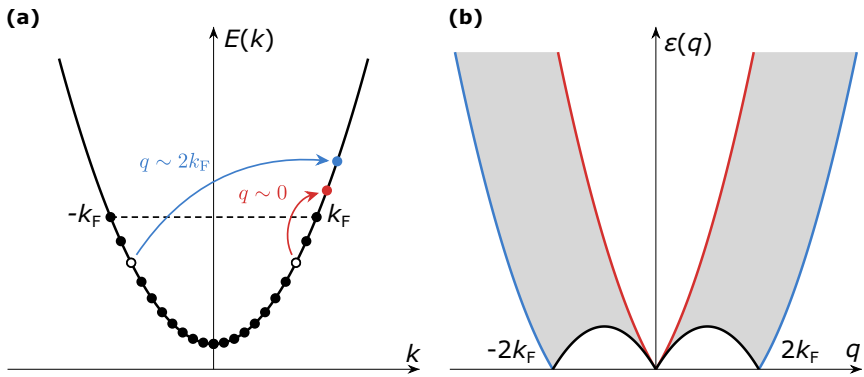


Figure 3.3 Excitation spectrum for 1D fermions. (a) Single-particle excitations constituting forward (red [grey in print version]) and backward (blue [grey in print version]) scattering for small and large momentum change across the Fermi points, respectively. Occupied states are represented by filled dots and holes by empty circles. (b) Electron–hole pair spectrum. Only shaded areas are energetically accessible. The forbidden region indicated by the gap is unique to the 1D geometry and is not present at higher dimensions. The position of the spectral edge, separating the accessible and forbidden regions, is marked by the black curve. It can be observed directly, for instance, by neutron scattering in some magnetic materials [12–14]. *From [10], reprinted with permission.*

where $V_1(q) = v_q$ and $V_2(q) = v_q - v_{2k_F}$. Note that the choice of V_1 and V_2 is, in general, completely arbitrary and highly dependent on the interaction process in question. Attempts to verify their form have been made by Dash et al. [15] assuming a Gaussian dependence, by Creffield et al. [16,17] using Monte Carlo simulations, and by Haüsler et al. [18] and Matveev et al. [19] analytically.

It can also be shown that the TLL Hamiltonian is diagonalisable if written in a bosonic basis. This means that every operator, including the fermionic creation and annihilation operators, can be represented in terms of boson operators together with operators that raise or lower the *total* number of particles. This is in stark contrast with the representation used so far, which assumed action on individual electrons and not on the system as a whole. This is, however, not surprising. As argued before, in 1D systems, collective response replaces individual behaviour.

A full review of the bosonisation method can be found in Apostol [20] or von Delft and Schoelle [21] as well as in several textbooks [2,9]. The main result is that a 1D system, following the TLL assumptions, is equivalent to a system of independent massless bosons, where the dispersion is given by $\omega_q = v|q|$, see Fig. 3.3b, with the velocity

$$v = \lim_{q\to 0} \sqrt{\left| v_F + \frac{V_1(q)}{2\pi\hbar} \right|^2 - \left| \frac{V_2(q)}{2\pi\hbar} \right|^2}. \tag{3.6}$$

Rewriting $v = v_F/K$ in Eq. (3.6) allows one to extract the TLL parameter K as

$$K = \lim_{q\to 0} \frac{1}{\sqrt{\left| 1 + \frac{V_1(q)}{2\pi\hbar v_F} \right|^2 - \left| \frac{V_2(q)}{2\pi\hbar v_F} \right|^2}}. \tag{3.7}$$

Here, K encapsulates both the interaction potentials $V_1(q)$ and $V_2(q)$ as well as the Fermi velocity v_F. Note that all of these parameters are independent of temperature.

1.2 Spinful Tomonaga–Luttinger liquids

Any model attempting to describe electron behaviour must account for spin. This plays an important role when considering the interaction process given by forward and backward scattering, as was shown by Luther et al. [7,22], since both parallel and anti-parallel configurations are now possible. One

can generalise both the kinetic $\widehat{H}_0$ and the interaction $\widehat{H}_{\rm int}$ terms to include spin, see Section 9.12 of [11]. The Hamiltonian for the spinful Tomonaga–Luttinger liquid then becomes

$$\widehat{H}_{\rm sTLL} = \widehat{H}_0 + \frac{1}{2L}\sum_{q\neq 0} V_1(q)(\widehat{\rho}_{-q,\rm R}\,\widehat{\rho}_{q,\rm R} + \widehat{\rho}_{-q,\rm L}\,\widehat{\rho}_{q,\rm L}) + \frac{1}{2L}\sum_{q\neq 0,\sigma}[V_2(q) - V_1(q)](\widehat{\rho}_{-q\sigma,\rm R}\,\widehat{\rho}_{q\sigma,\rm L} + \widehat{\rho}_{-q\sigma,\rm L}\,\widehat{\rho}_{q\sigma,\rm R}), \tag{3.8}$$

where $\widehat{\rho}_{q,\alpha} \equiv \widehat{\rho}_{q\uparrow,\alpha} + \widehat{\rho}_{q\downarrow,\alpha}$ is the electron density operator and $\sigma = \uparrow, \downarrow$ is the spin index.

Note the similarities between the spinless (Eq. 3.5) and the spinful (Eq. 3.8) Hamiltonians, the only difference being the potential prefactor of the last term. Therefore, the bosonisation method remains applicable as long as both bosonic operators and the (spin-) density fluctuation operator are also modified to include spin.

One of the most striking features of the spinful TLL is what came to be known as spin–charge separation. Once $\widehat{H}_{\rm sTLL}$ is rewritten in terms of the spinful bosonic operators, one sees that it decouples into two independent Hamiltonians,

$$\widehat{H}_{\rm sTLL} = \widehat{H}_{\rm C} + \widehat{H}_{\rm S}, \tag{3.9}$$

which can be associated with charge and spin modes, respectively. This result is analogous to the one discussed in the spinless case. Now, however, one has two independent collective excitations given by the characteristic frequencies $\omega_q^{\rm C} = |q|v^{\rm C}$ and $\omega_q^{\rm S} = |q|v^{\rm S}$, where

$$v^{\rm C} = \lim_{q\to 0}\sqrt{\left(v_{\rm F} + \frac{V_1(q)}{\pi\hbar}\right)^2 - \left(\frac{V_1(q) + V_2(q)}{2\pi\hbar}\right)^2} \tag{3.10}$$

is the velocity of the charge-density wave (CDW) and

$$v^{\rm S} = \lim_{q\to 0}\sqrt{v_{\rm F}^2 - \left(\frac{V_1(q) - V_2(q)}{2\pi\hbar}\right)^2} \tag{3.11}$$

is the velocity of the spin-density wave (SDW). These are, generally speaking, different.

Similarly to the spinless scenario, one can also define two interaction parameters for the spinful TLL, $K_{\rho,\sigma}$ given by $v_\rho \equiv v^{\rm C} = v_{\rm F}/K_\rho$ and $v_\sigma \equiv v^{\rm S} = v_{\rm F}/K_\sigma$ for the charge and spin modes, respectively. A spinful TLL is

therefore fully determined by four parameters, v_ρ, v_σ, K_ρ and K_σ, compared to just two, v and K, in the spinless scenario.

1.3 Spectral functions and power-law behaviour

A rigorous way of describing single-particle properties in a many-body system can be given via the spectral function $A(k,\omega)$. Formally, it represents the probability density of increasing or decreasing the energy of an N-particle system by adding or removing a single particle from a state characterised by a momentum k and energy ω, see an example in Fig. 3.4a for a spinful TLL. For comparison, in the absence of interactions, $A(k,\omega)$ can be shown [11] to take the form of a δ-function in frequency, $A(k,\omega) \propto \delta(\omega - E_k)$, i.e., adding or removing a particle to/from a plane-wave state always results in an exact eigenstate of the system. Here, $E_k = \hbar^2(k_x^2 + k_y^2)/(2m^*)$ for a 2D system.

In order to account for broadening arising from disorder, the spectral function can be convolved with a Lorentzian function. Following the notation used by Kardynal et al. [25], we have

$$A_{1(2)}\left(k_x, k_y, k_x', n; \mu\right) = \frac{\Gamma/\pi}{\left[\Gamma^2 + \left(\mu - \xi_{1(2)}\right)^2\right]}, \tag{3.12}$$

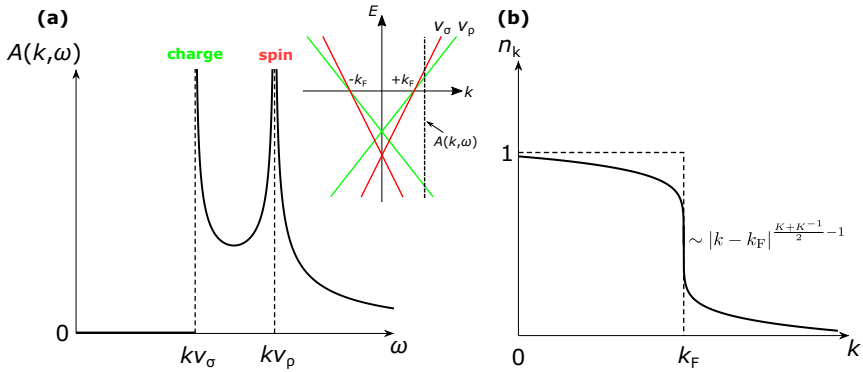


Figure 3.4 Spectral function of the spinful TLL and occupation numbers. (a) Spectral function $A(k,\omega)$ of a spinful TLL, evaluated at fixed k, showing charge and spin power law singularities at different energies [23,24]. Inset: Elementary excitations of a spinful TLL with $v_\rho \neq v_\sigma$. Green (grey in print version) and red (dark grey in print version) lines mark charge- and spin-type modes, respectively. (b) Occupation numbers n_k within the TLL theory without spin. *(a) From [3], reprinted with permission. (a) inset and (b) From [10], reprinted with permission.*

where Γ is the spectral line width associated with the single-particle disorder broadening, μ is the chemical potential and $\xi_{1(2)}$ are the energy spectra of the system, given by

$$\xi_1 = E_n + \frac{\hbar^2 k_x'^2}{2m^*} + E_{0,1}, \tag{3.13}$$

$$\xi_2 = \frac{\hbar^2\left(k_x^2 + k_y^2\right)}{2m^*} + E_{0,2} \tag{3.14}$$

with $E_n = (n+1/2)\hbar\omega_0$ being the energy of the nth subband and $E_{0,1(2)}$ denoting the bottom of the 2D subband as the electron gas is confined into a 1D (1) or 2D (2) geometry.

Experimentally, however, most 1D systems are not *perfect* 1D systems, i.e., infinitely narrow, having instead a finite transverse extension. This can be accounted for by treating the confinement potential as being parabolic, with electron wave functions in the transverse direction given by the eigenstates of the quantum harmonic oscillator, as is the case in semiconducting quantum wires. Consequently, $A_1(\mathbf{k}, n; \mu)$ must be modified to account for a harmonic dependence along the y-direction, resulting in energy levels known as 1D subbands and given by

$$A_{1,\mathrm{non-int}}\left(k_x', k_y, E, \Gamma\right) = \sum_n \frac{\Gamma}{\pi} \frac{H_n\left(k_y a\right) e^{-\left(k_y a\right)^2}}{\Gamma^2 + \left(\mu - \xi_1\right)^2}, \tag{3.15}$$

where $a = m_{1\mathrm{D}}\omega/\hbar$ is the finite width of the wire and $H_n(x)$ are the Hermite polynomials.

One can also find the momentum distribution function n_k of the system in a similar fashion, as it is related to the spectral function via the real-space Green's function, see Fig. 3.4b. A full analysis with the Green's function formalism is outside the scope of the review but can be found in Refs. [11,26]. Most importantly, however, at $T = 0$ K and using Eq. (3.15), $n(k)$ takes the following form:

$$n(k) \propto |k - k_\mathrm{F}|^{\frac{K+K^{-1}}{2}-1}. \tag{3.16}$$

Note that, at $k = k_\mathrm{F}$, $n(k)$ exhibits power-law behaviour with exponent $\alpha = (K + K^{-1})/2 - 1$. There is then a natural connection between the interaction parameter K, derived previously in the context of a spinless TLL, and correlation functions like $n(k)$ [27]. As we will see in the next sections, early experimental studies of TLL-type behaviour focussed exactly on

this property, by trying to establish the existence of an interaction parameter K by looking for power-law behaviour in the density of states around the Fermi points.

2. Early results on Tomonaga–Luttinger liquid behaviour

The basic principle behind any spectroscopy technique is the measurement of the spectral function $A(k, \omega)$ which, for a 1D system, combines all the relevant information about the energy and momentum of the elementary excitations. This probes both the hole sector, in which a particle is removed, and the particle sector, in which it is added, revealing various correlation effects emerging in many-body systems in addition to observing just single-particle dispersions of non-interacting electrons. From the solutions of the spinful TLL model, we know that the spectral function of an interacting 1D electronic system is double peaked, leading to spin–charge separation, and that the height of those peaks follows a power law in both temperature T and source-drain bias V_{dc}.

Early experimental work on verifying the TLL model focussed on establishing this power-law nature as well as observing the separation between the spin and charge modes. It can be grouped into three categories following the experimental technique used and includes photoemission spectroscopy, momentum-resolved tunnelling spectroscopy and transport measurements. The range of systems utilised is large and their claims to be a 1D system vary significantly. Nevertheless, important progress has been made using GaAs quantum wires [28,29], carbon nanotubes [30–32], the high T_c superconductor $SrCuO_2$ [33], the 1D metal $Li_{0.9}Mo_6O_{17}$ [34], 1D organic conductors [35,36] and quantum-Hall edge states [37,38] to name but a few.

2.1 Photoemission spectroscopy

The most direct way of measuring a system's spectral function is by photoemission spectroscopy. This is a difficult experiment to do with high resolution and as such had limited applicability when it came to measuring if 1D systems truly behaved as given by the TLL model.

The principle of operation is, however, relatively simple, see Fig. 3.5a. An incident photon $E_\gamma = \hbar\omega$ of known energy causes the transition of an electron from an initial band state at energy E_i to a final, plane wave state at energy E_f above the vacuum energy, where $E_f = \hbar^2 k_f^2/(2m) = E_i + \hbar\omega - \phi$ and ϕ is the work function of the material. This photo-emitted electron leaves the crystal and can be collected by a detector. This gives

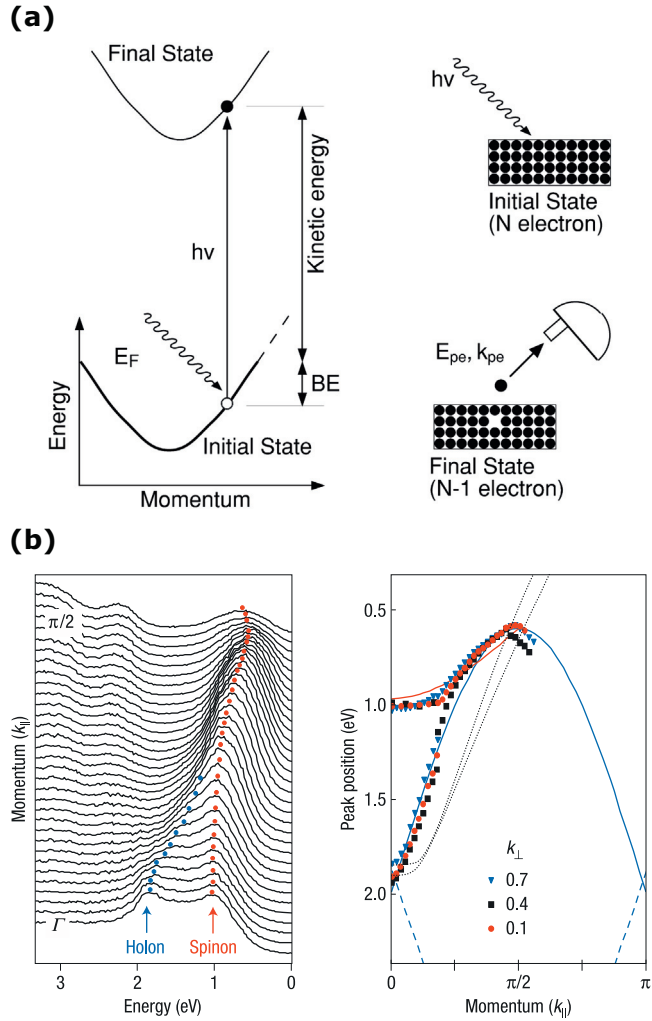


Figure 3.5 (a) Schematic representation of the direct photoemission spectroscopy, used in mapping the band structure of a system. (b) Left: angle-resolved photoemission spectroscopy (ARPES) data showing the many-body spectrum of $SrCuO_2$, where a two-peak spinon-holon structure can be seen, as predicted by the spinful Tomonaga–Luttinger model. Right: Comparison between experimental (symbols) and theoretical (solid and dashed lines) dispersions. *(a) From [40], reprinted with permission. (b) From [41], reprinted with permission.*

the emission intensity $I(k,\omega)$, which can be related to the spectral function $A(k,\omega)$ by

$$I(k,\omega) = M(k)A(k,\omega). \tag{3.17}$$

Here, $M(k)$ is the photoemission matrix element involving the initial and final states of the emitted photoelectrons. By mapping the system's response of photo-emitted electron energies as a function of analyser angle, information about the band structure can be obtained. For a review of angle-resolved photoemission spectroscopy experiments, see Ref. [39].

However, this technique is limited on two fronts. First, it assumes the 1D system is relatively close to the surface on which the light is incident. Anisotropic layered materials are a particularly good example of this, where there is little dispersion of the electronic bands perpendicular to the surface, making the data also easier to interpret. However, only certain classes of materials, such as specific high-temperature superconductors, fall into this category. Second, by its very nature, only occupied states can be probed. It is in principle possible to map unoccupied bands by inserting an electron of known energy into an empty band state and measuring the outgoing, ejected photon (this is known as *inverse* photoemission spectroscopy), but little work has been carried out in the 1D field using this technique.

Photoemission spectroscopy was first applied by both Kim et al. [33,40,41] and Fujisawa et al. [42,43] in studying the high-temperature superconductor $SrCuO_2$. Here, a good agreement to the experimental data was obtained by solving for the band structure using the Hubbard model, with the broadness of the observed peak being interpreted as resulting from the separation of the spinon and holon modes, see Fig. 3.5b. This work provided early evidence of spin–charge separation in a 1D Mott insulator which, although not metallic, still behaved as a 1D system as described by the TLL model.

Another experiment using photoemission spectroscopy was carried out by Gweon et al. [34] using the quasi-1D metal $Li_{0.9}Mo_6O_{17}$. Here, results similar to those shown above were also obtained, i.e., spectral separation between the spinon and holon modes in the observed dispersion. From the spectral functions, the velocities of each mode were also estimated, giving $v_\rho/v_\sigma = 2$. Their experiment also yielded good theoretical agreement, with a power-law exponent of about $\alpha = 0.9$, characteristic of a system with strong backscattering.

2.2 Transport measurements

Experimental studies on the TLL power-law behaviour were also performed by Bockrath et al. [30] and Yao et al. [44] using carbon nanotubes. Their experiments consisted of measuring electrical transport along single-walled

carbon nanotubes (SWNTs) with different contact geometries, either bulk or end. This method relied on approximating the Luttinger liquid parameter as

$$K_\rho = \left[1 + \frac{2U}{\Delta}\right]^{-1/2}, \tag{3.18}$$

where U is the charging energy of the nanotube and Δ is the single-particle level spacing. Bockrath et al. showed that, in a 1D system, the conductance along the system is suppressed at zero bias, vanishing as a power-law dependence with temperature T for small biases $|eV| \ll k_B T$,

$$G(T) \propto T^\alpha, \tag{3.19}$$

whilst with source-drain bias V_{dc} at larger biases $|eV| \gg k_B T$,

$$dI/dV \propto V^\alpha, \tag{3.20}$$

as predicted by the TLL model. Here, α depends on the number of 1D channels and on whether the tunnelling is happening into the bulk of the system or whether it is affected by the ends. This can be found from the TLL theory by calculating the momentum distribution function once the spectral function is known. In SWNTs, depending on the contact geometry, the power-law exponent α can then be related to K_ρ as either

$$\alpha_{end} = \left(K_\rho^{-1} - 1\right)/4 \tag{3.21}$$

or

$$\alpha_{bulk} = \left(K_\rho^{-1} + K_\rho - 2\right)/8 \tag{3.22}$$

as different contact geometries slightly change the prominence of the long-range Coulomb interactions.

In the experiment, sometimes individual nanotubes were placed on top of two metallic contacts, perturbing the nanotube only weakly and therefore yielding a long 1D system and hence 'bulk' tunnelling. Alternatively, contact metal was deposited on top of the nanotube, yielding a short 1D system in which any tunnelling excitation would propagate to the ends and be affected by the boundary condition there, producing 'end' tunnelling. This is determined by comparing the energy of the tunnelling electron to $\Delta E = 2\hbar v_{F1D}/(K_\rho L)$, which is related to the inverse time scale for the TLL quasi-particles to travel to the end of a wire of length L [45,46]. For $k_B T$, $eV_{dc} \gg \Delta E$, the ends are unimportant.

Conductance measurements as a function of temperature T and source-drain bias V_{dc} were performed at zero magnetic field, resulting in $\alpha_{end} \approx 0.6$ and $\alpha_{bulk} \approx 0.3$, in good agreement with the theoretical predictions for SWNTs, $\alpha_{end}(\text{theory}) = 0.65$ and $\alpha_{bulk}(\text{theory}) = 0.24$ [47,48], see insets in Fig. 3.6.

Whilst different contact geometries translated into different transport mechanisms, both systems were observed to still follow a power law in temperature once the data were re-scaled using the universal relation

$$\frac{dI}{dV} = AT^{\alpha}\cosh\left(\gamma\frac{eV}{2k_BT}\right)\left|\Gamma\left(\frac{1+\alpha}{2}+\gamma\frac{ieV}{2\pi k_BT}\right)\right|^2, \tag{3.23}$$

where $\Gamma(x)$ is the gamma function, γ takes into account the voltage drops at the two tunnel junctions and A is an arbitrary constant (see Fig. 3.6). Note that this result assumed that the leads were at $T = 0$ K. From here, the bulk-contacted nanotubes were observed to follow $\gamma_{bulk} \approx 0.46$, whilst end-contacted nanotubes gave $\gamma_{end} \approx 0.63$, in good agreement with the

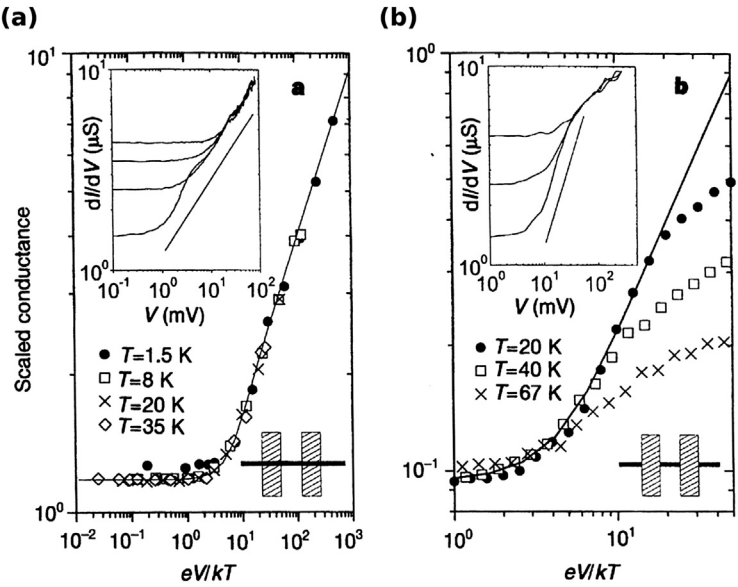


Figure 3.6 Differential conductance dI/dV taken at various temperatures for both (a) bulk- and (b) end-contacted carbon nanotubes. Insets show dI/dV on a log–log plot where power-law behaviour, $dI/dV \propto T^{\alpha}$, can be seen in both geometries (see straight line). Main panels show the same data collapsed onto a single curve after using the scaling relation described in Eq. (3.23). The extracted values of γ were (a) $\gamma = 0.46$ and (b) $\gamma = 0.63$. *From [30], reprinted with permission.*

theoretical predictions of Fisher and Kane [49], where $\gamma_{\mathrm{bulk}}(\mathrm{theory}) = 0.50 \pm 0.1$ and $\gamma_{\mathrm{end}}(\mathrm{theory}) = 0.60 \pm 0.1$. This was the first experimental evidence that electrons in a metallic nanotube behaved as a TLL.

2.3 Magnetotunnelling spectroscopy

Following Haldane's suggestion in 1981 and after initial measurements done on the 1D spectral function by Kardynal et al. [50], Barnes proposed in 1999 a novel method for probing TLL behaviour [51].

The idea was to measure the tunnelling conductance between a quantum wire and a parallel 2D electron system (2DES) as a function of both the potential difference between them, V, and an in-plane magnetic field, B, see Fig. 3.7a. The dependence on wave vector k and frequency ω of the spectral function $A(k, \omega)$ could then be determined by analysing the differential tunnelling conductance $\mathrm{d}I/\mathrm{d}V$. In particular, the authors argued that presence of spin–charge separation should manifest itself as emerging singularities in the $I - V$ characteristics, in a markedly different fashion from what would be expected in a noninteracting system. This idea was later further refined by Grigera et al. [52], who extended it to arbitrary values of

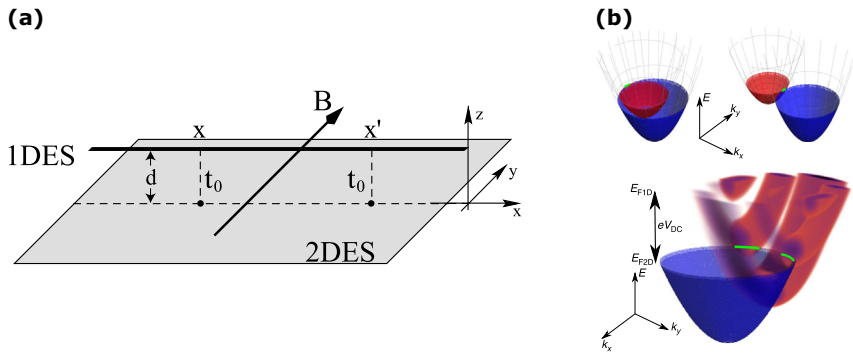


Figure 3.7 (a) Device configuration for mapping the dispersion of a 1D system by measuring tunnelling of electrons from it to a nearby 2DES. (b) Magnetotunnelling spectroscopy. Top panels show the overlap of the spectral functions of two 2D systems. Magnetic field B displaces the blue (dark grey in print version) paraboloid to the right, making the Fermi circles touch from the inside (top left) or from the outside (top right). This allows the states of the red (grey in print version) paraboloid near $-k_{\mathrm{F}}$ and k_{F} to be probed at the Fermi energy E_{F}. Bottom panel shows tunnelling between a 2D (blue [dark grey in print version]) and a 1D (red [grey in print version]) system at finite bias V_{dc}. Multiple 1D subbands can be probed by varying both B and V_{dc}. *(a) From [51], reprinted with permission. (b) From [53], reprinted with permission.*

the interaction parameter K_ρ whilst accounting for the effects of Zeeman splitting in both the TLL and the 2DES.

When an electron tunnels quantum mechanically through a barrier between two regions, it must conserve energy and momentum transverse to the barrier (modulo a reciprocal-lattice vector, in a crystal). It must also obey the Pauli exclusion principle, so it cannot tunnel into an occupied state. In a solid at low temperature, nearly all the states below the Fermi energy E_F are occupied, so this provides a reference energy below which tunnelling is forbidden. We will focus on semiconductor heterostructures where a planar barrier separates two parallel quantum wells containing electrons, which are free to move in two (or fewer) dimensions, so that in those directions their wave functions are described by plane-wave Bloch states with wave vector $\mathbf{k}$.

The tunnelling probability then depends on the overlap of the wave functions on either side of the barrier, integrated perpendicular to the barrier, at a rate given by Fermi's golden rule [54–56]. This picks out pairs of initial and final states with matching $\mathbf{k}$, embodying the conservation of momentum, with resonances in I being observed each time the 2D band and the nth subband dispersions are aligned. This, however, is unlikely to happen under equilibrium conditions.

By applying a magnetic field B parallel to the layers, a Lorentz force is produced as the electrons tunnel, boosting the momentum by an amount $\mathbf{\Delta k} = eBd/\hbar$, where d is the tunnelling distance (the distance between the centres of the wave functions). This allows one to map out the electron dispersion relations on either side of the barrier using the magnetic field B and the voltage V_dc applied across the barrier, which provides extra energy eV_dc. Putting it all together, we obtain

$$I \propto \int \mathrm{d}\mathbf{k}\mathrm{d}E[\, f_T(E - E_\mathrm{F1D} - eV_\mathrm{dc}) - f_T(E - E_\mathrm{F2D})] \times A_1(\mathbf{k}, E) A_2(\mathbf{k} + ed(\mathbf{n} \times \mathbf{B})\,/\,\hbar, E - eV_\mathrm{dc}). \tag{3.24}$$

Here, e is the electronic charge, $f_T(E)$ is the Fermi–Dirac distribution, d is the separation between the wells, $\mathbf{n}$ is a unit normal to the plane, $\mathbf{B} = -B\widehat{\mathbf{y}}$ is the (in-plane) magnetic-field vector and $A_1(\mathbf{k}, E)$ and $A_2(\mathbf{k}, E)$ are the corresponding spectral functions of the 1D and 2D systems, together with their Fermi energies, E_F1D and E_F2D, respectively (see Fig. 3.7b).

Both the 1999 proposal of Barnes (Altland et al. [51]) and the subsequent work of Grigera et al. [52] in 2004 concerned the idea of detecting spin–charge separation in a quantum wire. Neither, however, made any attempt at modelling the anomalous suppression of tunnelling current at zero bias, as

predicted by the TLL model. This feature, commonly referred to as the zero-bias anomaly (ZBA), is also expected to depend strongly on the strength of the interactions and to follow a power-law behaviour in both bias V and electron temperature T.

The first experimental realisation of tunnelling from a 1D wire that was in the single-subband regime was achieved by Auslaender et al. [28,57], with later theoretical analysis by Carpentier et al. [58]. Here, the authors used a pair of coupled quantum wires grown via cleaved-edge overgrowth (CEO), meaning that the tunnelling process took place between two 1D systems instead of between a wire and a 2DES, as discussed before, see Fig. 3.8a and b. The CEO technique allowed very strong confinement to be obtained, equal to roughly the width of the well itself, and as such in principle better suited for observing interaction effects. Unfortunately, the probe layer was now also of the same nature as the system being probed, meaning that interpretation and disentanglement of each wire's response became significantly more complicated. These experiments, nevertheless, showed early evidence in support of the TLL picture by observing an enhancement of the excitation velocity relative to the bare electron velocity v_F, see Fig. 3.8c and d. Previous work by the same authors had also observed power-law behaviour of the tunnelling density of states around the zero-bias suppression region [45,59].

In 2009, Jompol et al. reported the first clear observation of spin–charge separation together with power-law suppression of the tunnelling current at zero bias. In contrast to the CEO samples, the authors fabricated an electrostatically gated 1D system and measured tunnelling between it and a nearby 2DES. Thus, the probe layer used was now a well-understood system, making interpretation much more straightforward.

Fig. 3.9 shows the tunnelling conductance G as a function of both interlayer bias V_{dc} and in-plane magnetic field B at lattice temperatures of 1 K and $\sim$40 mK, when only one 1D subband was occupied. The dashed and solid black curves mark the expected 1D and 2D parabolic dispersion, respectively, arising from single-particle tunnelling in the absence of interactions. A parasitic 2D–2D tunnelling signal originating from the ungated injection region is also shown, as green (grey in print version) dash-dotted lines. From the crossing points at $V_{dc} = 0$, B^- and B^+, the 1D Fermi wave vector $k_{F1D} = ed(B^+ - B^-)/2\hbar$ together with the 1D electron density $n_{1D} = 2k_{F1D}/\pi$ can be extracted. For the data shown, the approximate electron density in the wires was $n_{1D} \cong 40\ \mu m^{-1}$.

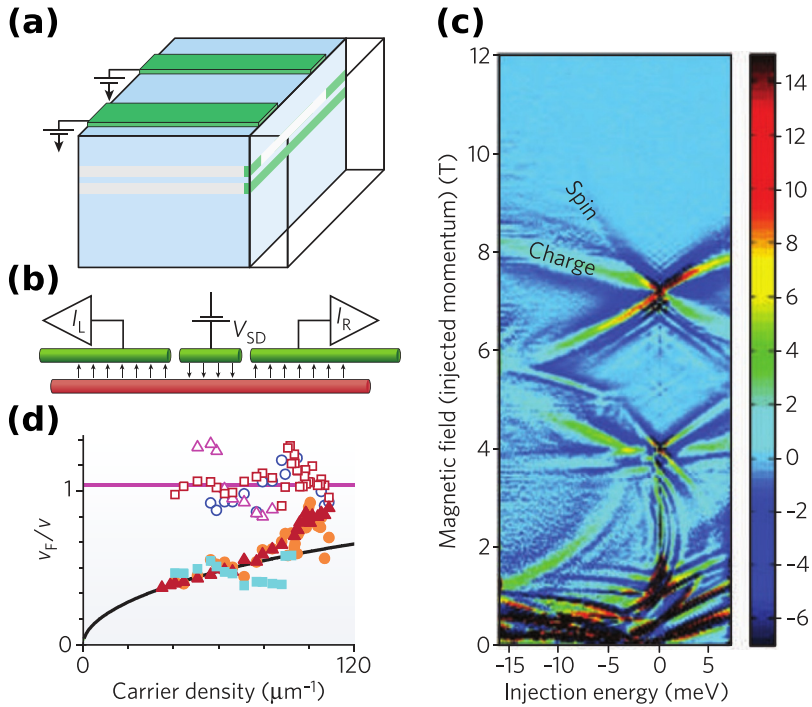


Figure 3.8 Probing spin–charge separation via momentum-resolved magnetotunnelling spectroscopy between two interacting 1D wires. (a) Schematic representation of double-wire structure made via cleaved-edge overgrowth (CEO). The system resides at the edge of a GaAs/AlGaAs double-well heterostructure. (b) Schematic of the circuit used in the 1D–1D tunnelling process. Here, I_L is the left-moving current, I_R the right-moving current and V_{SD} the source-drain bias voltage. (c) Differential conductance $G = dI/dV$ plotted as a function of both V_{SD} ($\propto$ energy) and B ($\propto$ momentum). Two features, labelled charge and spin, can be seen branching off at around zero energy and a magnetic field of 7 T. (d) Measured spin velocities (open symbols) and charge velocities (filled symbols), plotted as a function of carrier density and normalised with respect to the Fermi velocity v_F. At lower densities, repulsion between the electrons is stronger and hence the charge excitation velocity larger, unlike the spin mode which remains largely unaffected. Solid curves show the theoretical fits. *From [60], reprinted with permission.*

In both Fig. 3.9a and b, an additional region of high conductance at positive biases near B^+ can be observed. This can be more clearly seen in Fig. 3.9c by looking at the differential dG/dB. In both cases, a feature moving diagonally up and to the left of B^+ is visible. Fig. 3.10b shows detailed measurements taken around the same region, for a different device, also in the single-subband regime. For comparison, Fig. 3.10a and d show the theoretical

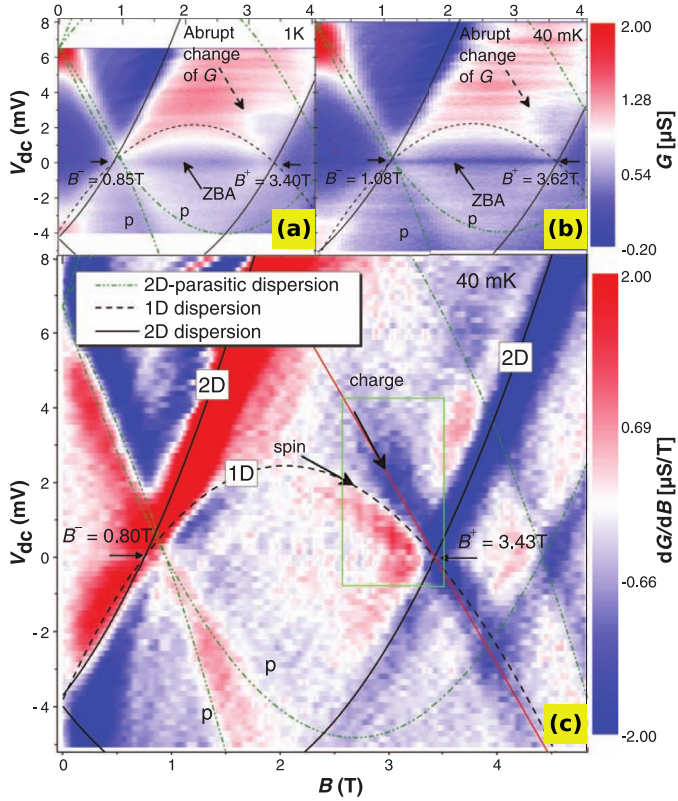


Figure 3.9 Tunnelling conductance G versus dc-bias V_{dc} and magnetic field B, at lattice temperatures T of (a) 1 K and (b) 40 mK. Tunnelling resonances between a system of 1D wires and a 2DES as predicted by the noninteracting model are marked by solid and dashed black lines, whilst green (grey in print version) dash-dotted curves indicate the location of unavoidable parasitic 2D–2D tunnelling 'p'. In addition, signs of interactions can be seen in the suppression of G around zero bias (ZBA) and its abrupt drop at positive biases around B^+. (c) dG/dB differential of the data shown in (b). The noninteracting parabolae shown in (a) and (b) are labelled 1D, 2D or p, depending on which system's dispersion is being probed. The straight red (grey in print version) line, marking the position of the abrupt drop in G, is $\sim$1.4 times steeper on this plot than the 1D parabola at $V_{dc} = 0$. According to the TLL model, we identify the first as a charge-type mode (holon), whilst the second as a spin-type excitation (spinon). *From [29], reprinted with permission.*

prediction for a noninteracting system. Here, a strong dark blue (dark grey in print version) feature of negative dG/dB can be seen; however, it follows the 1D parabola unlike what is observed experimentally which moves away from it (dashed red [grey in print version] line). It follows therefore that the 1D parabola and the dashed red (grey in print version) line track the dispersion

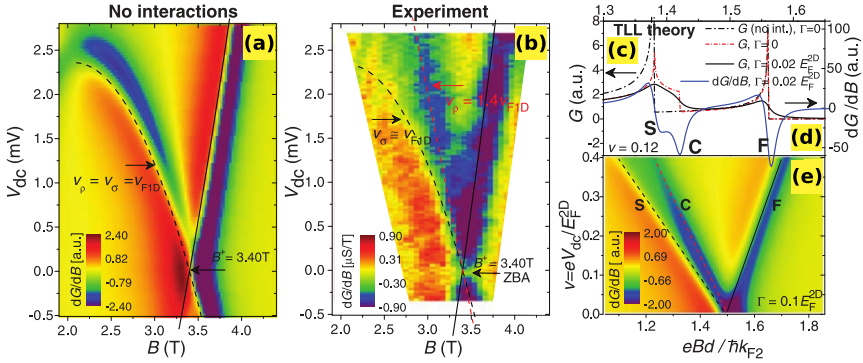


Figure 3.10 Spin–charge separation around the $+k_F$ point: comparison between experiment and theory. (a) dG/dB for a system of noninteracting electrons with disorder broadening $\Gamma = 0.6$ meV. (b) High-resolution mapping of dG/dB. The red (grey in print version) dashed line marks a feature that does not follow the noninteracting parabola (black dashed) and is absent in (a). The spinon velocity is given by $v_\sigma \equiv v_{F1D}$. The extracted holon velocity was $v_\rho = 1.4v_{F1D}$. (c) and (d) show, respectively, the calculation of G and dG/dI for both a noninteracting and a TLL system of electrons. Spin and charge excitations are labeled by S and C, respectively, whilst F labels the noninteracting 2D dispersion curve. The cuts were taken at a bias of $v = eV_{dc}/E_F^{2D} = 0.12$ (given here in dimensionless units). (e) dG/dB as a function of B and v according to the TLL model. Note how, unlike (a), this model predicts a charge feature C as seen in (b). *From [29], reprinted with permission.*

of two independent features which, according to the TLL picture, can be associated with spin- and charge-type excitations, respectively. The velocity of the charge mode was estimated to be $v_\rho = 1.4v_{F1D}$, resulting, under the linear approximation, in an interaction parameter $K_\rho \approx 0.71$. For completeness, Fig. 3.10c and e show the theoretical result as expected for 1D interacting electrons tunnelling to and from a 2D system according to the TLL model, where a similar feature to that observed can also be seen.

Another feature that cannot be explained by the noninteracting model is the ZBA, visible as a dark blue (dark grey in print version) horizontal line in Fig. 3.9a and b. As discussed before, this anomalous suppression in the tunnelling current results from interaction effects and is likely to be associated with the extra energy cost for an electron to tunnel in or out of the wire, at zero bias, as it inevitably disturbs the remaining electrons already present. The tunnelling conductance G as a function of V_{dc} for different temperatures T can be seen in Fig. 3.11a, at a field B approximately midway between B^- and B^+. Similarly, Fig. 3.11b and c show $G(V_{dc} = 0, T)$ and $G(|V_{dc}|, T < 70\text{mK})$, respectively, on a log–log plot. Here, clear variation as a power

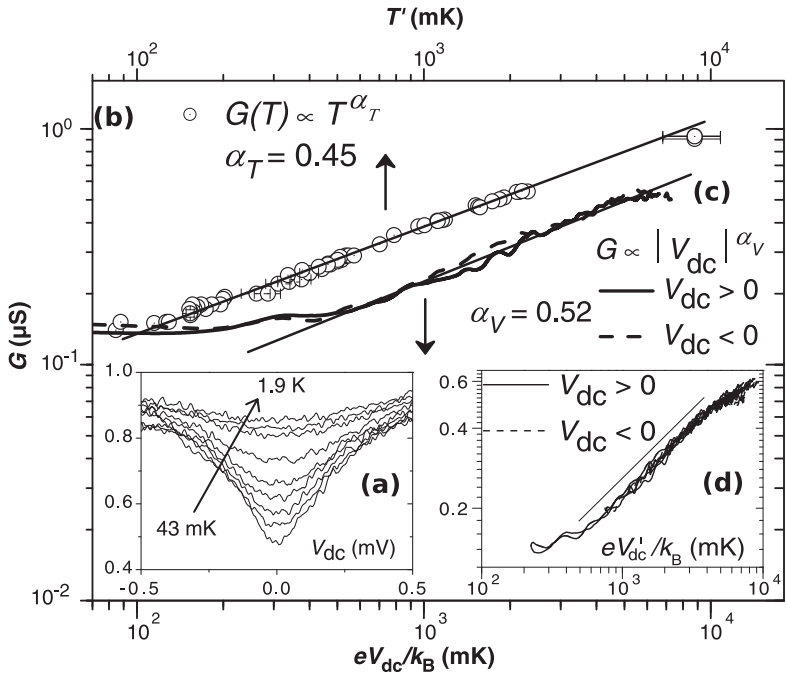


Figure 3.11 The zero-bias anomaly. (a) Tunnelling conductance G as a function of bias V_{dc}, taken at a field $B = 2.33$ T, for different temperatures T ranging from 43 mK to 1.9 K. (b) Log–log plot of $G(V_{dc} = 0)$ versus T, showing power-law behaviour $\propto T^{\alpha}$ with $\alpha_T \approx 0.45$. (c) Same as in (b) but showing $G(|V_{dc}|)$ versus V_{dc} with $\alpha_V \approx 0.52$. (d) $G(|V_{dc}|)$ versus $V'_{dc} = \sqrt{|V_{dc}|^2 + (3k_BT/e)^2 + V_{ac}^2}$, obtained by adding different sources of energy smearing as noise in quadrature. The value of $3k_BT$ was chosen so that the curves in (b) and (c) and all the curves in (d) would superpose. For comparison, using the scaling relation in Eq. (3.23) results in a universal curve with $\alpha \approx 0.51$. *From [29], reprinted with permission.*

law in both V_{dc} and T over a range of orders of magnitude is observed. Both V_{dc} and T play a similar role in smearing the energy, as illustrated in Fig. 3.11d.

The interaction parameter K_ρ was obtained from the extracted power-law exponents, $\alpha_T \approx 0.45 \pm 0.04$ and $\alpha_V \approx 0.52 \pm 0.04$, depending on whether excitations from electrons tunnelling into or out of the wire generally propagated as far as the ends ('end tunnelling') or not ('bulk tunnelling'), see section 2.2. For the latter, the measured exponents gave $K_\rho \approx 0.28$ and 0.26, respectively, significantly smaller than the value extracted from the holon branch in Fig. 3.10b. By considering 'end' tunnelling instead, the values obtained were of $K_\rho = 0.53$ for α_T and $K_\rho = 0.49$ for

α_V, considerably closer to that obtained in Fig. 3.10. However, this implies that the wire length is shorter than the thermal length $L_T = 2\hbar v_{F1D}/(K_\rho k_B T)$ at all temperatures measured (100 mK–10 K), or else the power-law exponent would change at some temperature, as observed in 1D–1D tunnelling by Tserkovnyak et al. [45].

Overall, the work carried out by Jompol et al. showed clear evidence in support of the TLL model. It was obvious, though, that the 'charge' line observed persisted far beyond the energy range over which the dispersions could be approximated to be linear, making the spin–charge separation a phenomenon more robust than that expected from the range of validity of the TLL model. This raised the question, therefore, of what happens to both spin and charge modes in the nonlinear region, that is, far from the Fermi points, where the linear approximation breaks down. The role of interactions in 1D wires at arbitrary energy and momentum has only recently started to be understood. Such 'nonlinear Tomonaga–Luttinger liquids' are the topic of the next section.

3. Beyond the linear Tomonaga–Luttinger liquid approximation

TLL theory works remarkably well in describing the low-energy behaviour of 1D systems. As we have shown, these exhibit a number of unusual properties, including long-lived plasmonic excitations, the separation of the spin and charge degrees of freedom and interaction-driven power-law behaviour [8]. In general, however, the 1D spectrum is curved, i.e., nonlinear, meaning that deviations from linearity are expected to occur away from the Fermi points. We will now focus our attention on two particular nonlinear theories.

3.1 Mobile impurity model of nonlinear Tomonaga-Luttinger liquids

Excitations away from the Fermi energy are described in a 1D system by nonlinear hydrodynamics, in what came to be known as the mobile impurity model, see Fig. 3.12 [61,62]. The name itself is reminiscent of another problem, X-ray scattering in metals, where a hole is created far below the Fermi level upon absorption of a high-energy X-ray photon. As in the TLL model, this system is also known to have power-law singularities around E_F. However, in contrast to that model, perturbation analysis on the resulting interaction between the deep hole and the low-energy quasi-particles is divergent.

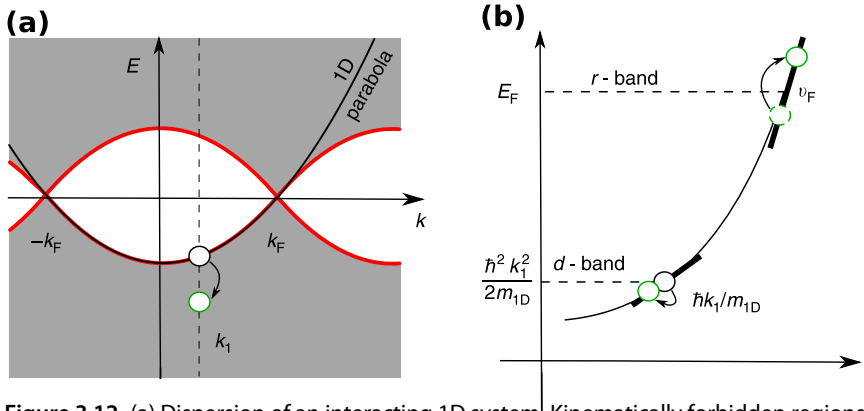


Figure 3.12 (a) Dispersion of an interacting 1D system. Kinematically forbidden regions are shown in white, whilst the grey shaded area represents the many-body continuum of excitations. The spectral edge marks the border between these two regions and is shown in red (dark grey in print version). A higher-energy excitation (green [grey in print version] circle) is obtained by changing the state of a heavy hole (black circle) deep below the Fermi energy and simultaneously creating some linear TLL excitations around the Fermi energy. This moves the total energy of the many-body state away from the spectral edge as described by the mobile impurity model – see text. (b) Splitting of the fermionic dispersion into two subbands, one for a heavy hole deep in the Fermi sea, with characteristic velocity given by $\hbar k_1/m_{1D}$, and one for excitations close to the Fermi point k_F, with velocity v_F. *From [53], reprinted with permission.*

Nozières and De Dominicis [63] solved this problem by introducing the so-called heavy impurity model. This construction, resulting from Fermi liquid quasi-particles interacting with a hole state deep below the Fermi level, can be diagonalised exactly, in a framework analogous to that of a TLL system. It also predicts power-law behaviour away from the Fermi energy, in the form of a threshold exponent. Most interestingly though, the 1D spectral function, as derived from the mobile impurity model, is

$$A_1(k_x, E) \propto \frac{1}{|E + \mu - \xi_1|^{\alpha(k_x)}}, \tag{3.25}$$

implying that there is now a momentum dependence in the exponent $\alpha(k_x)$ of the interacting 1D spectral function [62]. The explicit expression for spinful fermions that is applicable for electronic systems is found to be [64].

$$\alpha(k_x) = 1 - \frac{(1 - C(k_x))^2}{4K_\rho} - \frac{K_\rho(1 - D(k_x))^2}{4}, \tag{3.26}$$

where $C(k) = (k^2 - k_F^2)/(k^2/K_\sigma - k_F^2 K_\sigma/K_\rho^2)$ and $D(k) = (k - k_F)(k_F/K_\rho^2 + k/K_\sigma^2)/(k^2/K_\sigma^2 - k_F^2/K_\rho^2)$.

Having now extended the universality of a TLL from the low- to the finite-energy regime, one finds that a momentum-dependent power law emerges due to the finite curvature of the spectral edge in energy and momentum space. This is a unique feature of nonlinear hydrodynamics in 1D and a hallmark of nonlinear Tomonaga-Luttinger liquid (NLL) behaviour. Another implication of this theory is the emergence of 'replicas' of the principal dispersion, i.e., all red (dark grey in print version) lines in Fig. 3.12a are predicted to be pronounced in the spectral function instead of just the one coinciding with the principal 1D parabola.

Yet another prediction from the NLL model is the drastic reduction of the plasmon's lifetime. This can be understood by analogy with the linear case. If the spectrum is linear, each electron–hole pair can be identifiable by an energy linear to its momentum. If, instead, the spectrum is curved, the possibility exists of having different electron–hole pairs with different energies yet the same momentum. These can morph into a plasmon with wave number q once the interactions are turned on, with a finite spectral width, approximately equal to the spread in energies, $\delta\omega(q) = \omega_+(q) - \omega_-(q)$, and hence a reduced lifetime $\tau \propto 1/\delta\omega$.

3.2 A mode hierarchy away from the Fermi points

The mobile impurity model, described in the previous section, has released the TLL theory from the restriction of the linear approximation. Nevertheless, it still does not provide a description for arbitrary nonlinear excitations since it is applicable only in the proximity of the spectral threshold. This is because, being constructed by analogy with the edge singularity in the X-ray scattering problem, the nonlinear hydrodynamics become less well defined away from the spectral edge separating the forbidden zone from the continuum of many-body excitations in 1D, where such a phenomenological approach becomes inconsistent.

A more systematic understanding of the general picture for nonlinear excitations came from a microscopic analysis via the Bethe ansatz. It was found [65,66] that an exponentially large number of excitations are separated into levels of a mode hierarchy according to their spectral strength, which is proportional to integer powers of a small parameter R^2/L^2, where R is the radius of the two-body interaction potential and L is the length of system. A detailed analysis of these amplitudes revealed formation of a

parabola-like dispersion, see Fig. 3.13a, similar to that of original noninteracting fermions but with some re-normalisation due to interactions, by the strongest excitations with the zeroth power of R^2/L^2. All other many-body modes have powers of R^2/L^2 greater than zero with the general trend of increasing the integer by one with each discrete step of $2k_F$ away from the principal parabola.

As an example, the hole part of the principal parabola, between the $\pm k_F$ points, has the largest amplitude but its mirror in the particle sector – a 'replica' in the shape of a dome marked as *p0b* in Fig. 3.13a – has a parametrically smaller amplitude proportional to the first power of R^2/L^2. In the spectral function, the strength of this replica is predicted to be

$$A_1(k_x, E) \propto \frac{R^2}{L^2} \frac{k_F^2 k_x^2}{\left(k^2 - k_F^2\right)^2} \delta(E - \mu + \xi_1). \tag{3.27}$$

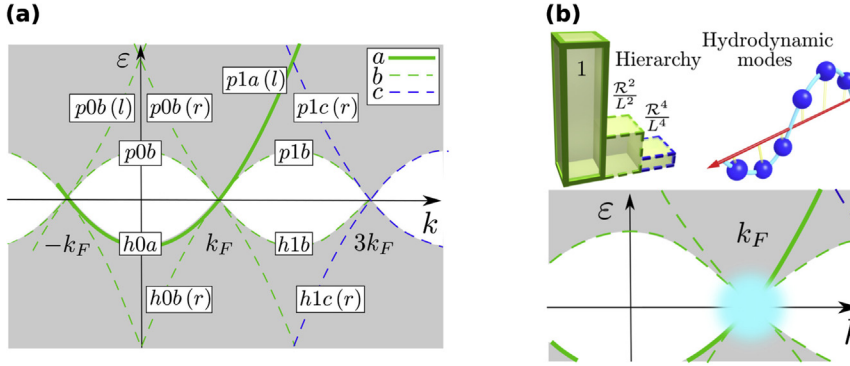


Figure 3.13 (a) Spectral function of spinless fermions according to the mode hierarchy picture. First (second)-level modes are shown in the region $-k_F < k < k_F$ ($k_F < k < 3k_F$) and are labelled by 0 (1), where k_F is the Fermi wave vector. Accessible and forbidden regions are marked by grey and white, respectively. Particle (hole) sectors are marked by *p* (*h*) for positive (negative) energies, *a*, *b* and *c* correspond, respectively, to the level in the mode hierarchy in powers of 0, 1 and 2 of R^2/L^2 and (r, l) correspond to the origin in the range, right or left. (b) Bottom, regions of validity for different theories on the energy–momentum plane. Top right, TLL hydrodynamic modes, where the deviations from the red (grey in print version) line mark a variation in density and which dominate at low energies (see cyan region in the bottom panel). Top left, mode hierarchy up to third order, where the bars denote qualitatively the amplitude of different many-body modes and which is relevant for the rest of the plane. *(a) From [65], reprinted with permission. (b) From [66], reprinted with permission.*

Therefore, for almost all momenta, this mode will be unobservable in the thermodynamic limit. The only exceptions are the regions around the $\pm k_F$ points, where the singularity in the denominator starts to compete with the parametric smallness, resulting in a large amplitude overall. On the other hand, the measurement of the whole mode requires using smaller systems, in which the R^2/L^2 parameter still leaves the amplitude of the whole mode above the background from the other processes.

Close to the spectral threshold the mode hierarchy reproduces the predictions of the mobile impurity model. The small parameter R^2/L^2 provides a path for accounting for the principal amplitudes of the exponential continuum of many-body excitations, which leads to the microscopic calculation of the spectral function. The resulting threshold exponents match those predicted by the phenomenologically introduced nonlinear hydrodynamic model.

Whilst the mode hierarchy emerges away from the Fermi points, close to them (where the spectrum is almost linear) it transitions into the usual linear TLL, see Fig. 3.13b. The hydrodynamic modes of the latter consist of a huge number of many-body modes, all of which are of similar spectral strengths, making the two regimes distinct already on the microscopic level. The change from one into the other can be traced quite easily using a macroscopic quantity, the density of states. It can be calculated exactly by numerical means using Bethe ansatz, exhibiting the power-law suppression around the Fermi energy E_F predicted by the TLL model and a crossover into a finite density $\propto 1/\sqrt{E}$ predicted by the mode hierarchy away from the linear region, where the nonlinearity of the single-particle dispersion already destroys the hydrodynamic modes of the TLL.

4. Recent work on nonlinear effects

4.1 1D 'replica' modes

In 2015, the first observation of structure resembling higher-order excitations, as predicted by the mode hierarchy model, was reported in Tsyplyatyev et al. [65] in devices similar to that used by Jompol et al. [29] but with air bridges [67], see Fig. 3.14. The experiment consisted of measuring momentum-resolved tunnelling to and from an array of 1D wires to a 2DES in a GaAs/AlGaAs heterostructure, for a variety of different wire lengths. The authors observed structure resembling second-level spinon excitations, near the $+k_F$ point, branching away from the 1D mode and dying away rapidly at high momentum, in line with the theoretical predictions.

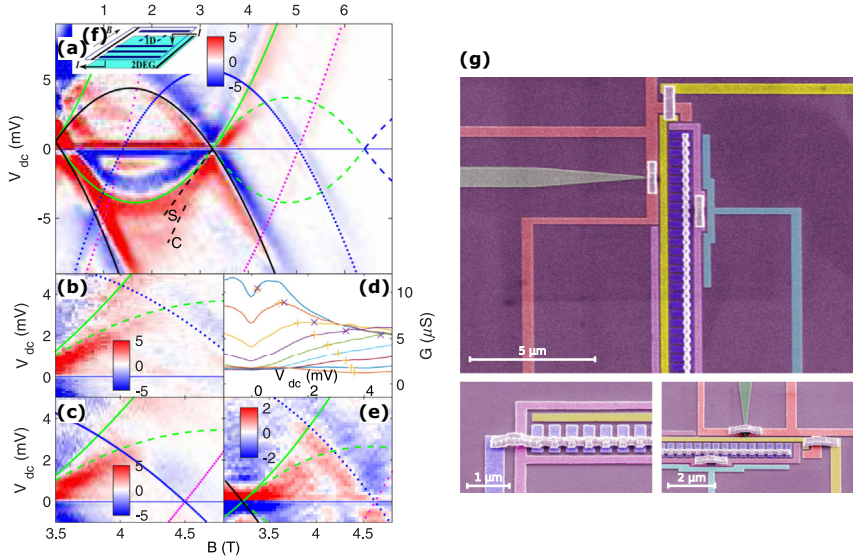


Figure 3.14 Tunnelling differential conductance $G = \mathrm{d}I/\mathrm{d}V$ for two samples, consisting of a set of identical wires of length $L = 10$ μm (a)–(d) and $L = 18$ μm (e). A schematic representation of the device measured can be seen in inset (f). (a) $\mathrm{d}G/\mathrm{d}V_{\mathrm{dc}}$ map obtained at a lattice temperature of $T = 300$ mK. The 1D signal is marked by solid green (light grey in print version) curves for the *a* modes, dashed green (light grey in print version) for the *b* modes and dashed blue (dark grey in print version) for the *c* modes, see Fig. 3.13 for details. Magenta (grey in print version) and blue (dark grey in print version) curves indicate the parasitic 2D–2D signal, whilst the black curve follows the dispersion of the 2D system as probed by the 1D wires. Spin and charge modes are labelled by the S and C lines, respectively. (b) and (c) show an enlargement of the replica region to the right of $+k_{\mathrm{F}}$, as seen in (a), for $v_{\mathrm{PG}} > 0$ and $v_{\mathrm{PG}} = 0$, respectively, where PG is a gate running over most of the parasitic 'p' region, see blue (dark grey in print version) curves. (d) G vs V_{dc} at various fields B, around the replica region. The symbols $+$ and $\times$ on each curve mark, respectively, the dashed and solid green (light grey in print version) curves in (a) and (b). (e) $\mathrm{d}G/\mathrm{d}V_{\mathrm{dc}}$ map for a second device obtained at $T < 100$ mK, showing a similar feature to that observed in (a)–(c). (g) Scanning electron micrographs of a tunnelling device with air bridges, showing how air bridges are used not only to connect gates together across other gates, but also to link together all the finger gates defining the 1D wires, which avoids variation of the gate potential along each of these very short wires, as happens when the gates are joined together at one end [67,68]. *From [65] reprinted with permission.*

Most interestingly, however, this feature was observed to emerge as a function of system length, its spectral weight increasing for shorter channels. It was also noted to be visible even when more subbands started being occupied. A comprehensive review on the mode hierarchy picture, highlighting both theoretical and experimental progress, can be found in Tsyplyatyev et al. [66].

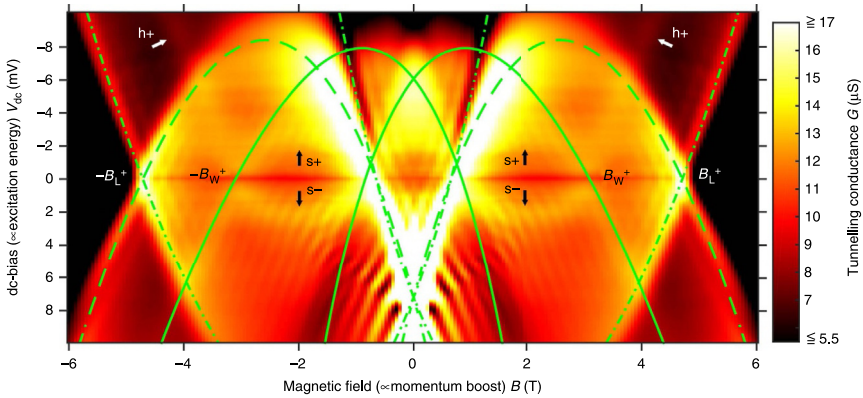


Figure 3.15 Tunnelling conductance G plotted as a function of dc-bias V_{dc} and in-plane magnetic field B, for the regime of a single 1D subband occupied. The electron density in the wires is $n_{1D} \approx 35\ \mu m^{-1}$. The solid green (grey in print version) curves mark the dispersion of the 2D system as mapped by the 1D wires, whilst the green (grey in print version) dashed and dash-dotted lines indicate the resonances arising from 2D–2D parasitic tunnelling. Spinon bands can be seen in the hole (s–) and particle (s+) sectors, whilst a holon band is present in the particle sector (h+). The labels $\pm\ B_{W,L}^{+}$ indicate specific magnetic fields at which each resonance crosses the $V_{dc} = 0$ axis. *From [69], reprinted with permission.*

In 2016, Moreno et al. [69], also using tunnelling spectroscopy, found evidence for the existence of another higher-order mode, this time a first-level inverted spinon band (see Fig. 3.15). The tunnelling device used consisted of an array of 6000 GaAs wires, 1 μm in length, considerably shorter than previously used in Tsyplyatyev et al. [65,66], which were 10 μm and 18 μm in length. The mode reported by Moreno et al. could not be observed in the longer-length wires, further establishing the mode hierarchy picture as one of the main nonlinear TLL theories. More recently, a full experimental study of the length dependence of replicas has been carried out by Vianez et al. [68], confirming the prediction in the mode hierarchy picture that the replicas get stronger as the length decreases [65,66].

4.2 Momentum-dependent power law

As was discussed in section 3.1, the mobile impurity model, developed in the context of 1D quantum fluids at arbitrary energy and momentum, makes a remarkable prediction regarding the power-law behaviour of the observed density of states near the spectral edge. Here, due to the finite curvature of the 1D dispersion, the exponent becomes momentum dependent, in clear

contrast with the linear model, where no such dependence is expected. For electrons, i.e., spin-1/2 fermions, the observation of a momentum dependence of the threshold exponent away from the Fermi points would constitute a hallmark of 1D nonlinear hydrodynamics.

In 2019, Jin et al. [53] reported the first observation of this new type of power law using a system of interacting electrons in a quantum wire, probed via tunnelling spectroscopy. Here, a finite length scale was present – given by the particle's wavelength – unlike in the usual physical picture of phase transitions where length scales become infinite, making it the first power law observed in the absence of a scale invariance.

The authors observed that an enhancement of the tunnelling conductance below the bottom of the 1D dispersion (see Fig. 3.16) could not be explained solely by considering either a noninteracting or linear Luttinger framework (i.e., a momentum-independent power law). Instead, the best fits to the data were obtained by considering a momentum-dependent

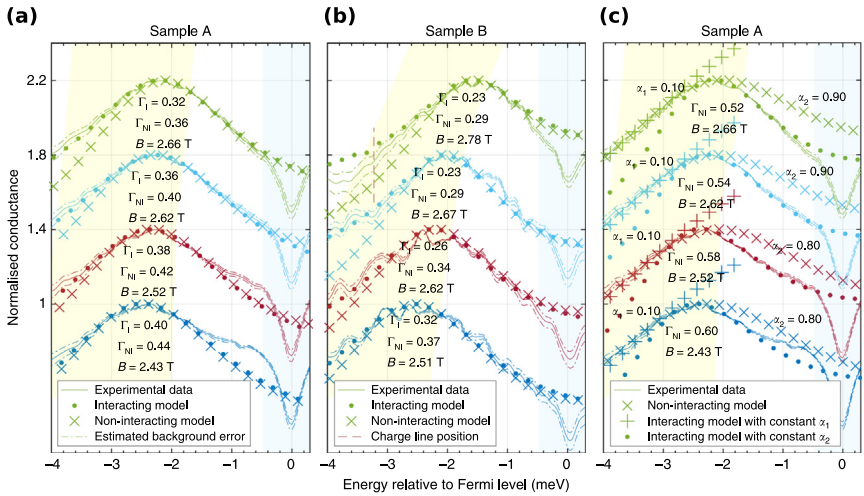


Figure 3.16 Momentum-dependent power law. (a) and (b) Fits to the tunnelling conductance of two different samples, around the bottom of the 1D subband, for a variety of different fields, normalised to their peak value and shifted vertically for clarity. Two different models, noninteracting ($\times$) and interacting with a momentum-dependent exponent (•), were used when fitting the data. The effects of disorder-induced broadening were also considered. Sample A, consisting of 18 μm long wires, was measured at 50 mK in a He^3/He^4 dilution refrigerator, whilst sample B, with 10 μm, was measured at 330 mK in a He^3 cryostat. (c) Fittings obtained using a momentum independent power law (• and $+$) and a noninteracting model ($\times$), matched to specific parts of the data. *From [53], reprinted with permission.*

power law as predicted by the mobile impurity model. In addition to the interaction effects, the effects of disorder-induced broadening were also considered. This result was observed in multiple devices, with different wire lengths and measured at different temperatures.

The extracted Luttinger parameter from the nonlinear regime was $K_\rho = 0.70 \pm 0.03$. This is in good agreement with the values obtained from other 1D interaction effects, most notably the ZBA ($K_\rho = 0.59 \pm 0.13$ assuming end tunnelling) and the spin–charge separation ($K_\rho = 0.76 \pm 0.07$), both also present in this experiment. Note that throughout this analysis the authors considered $K_\sigma = 1$, in accordance to re-normalisation group arguments [9].

4.3 Lifetime of spinon and holons at high energies

One of the early studies of nonlinear TLL dynamics was done by Barack et al. [70]. Here, the authors selectively injected holes and particles into a quantum wire in order to study their relaxation properties via momentum-resolved tunnelling spectroscopy. Their measurements concluded that whilst energetic particles underwent a rapid thermal relaxation, holes remained largely unchanged, in clear contrast with the linear TLL theory, in which energy relaxation is largely forbidden. This sets the limits on the relaxation times to $\tau < 10^{-11}$ s for particles and $\tau \gg 10^{-11}$ s for holes.

It is important to highlight that these results rest on the assumption that the interactions are weak. Previous work done by the same authors on similar samples obtained a Luttinger parameter for the electrons of $K_\rho \approx 0.55$, comfortably sitting in the strongly interacting regime. This apparent contradiction however was explained by suggesting that relaxation dynamics are largely independent of the strength of the interactions. The model derived showed good agreement with the data. Its phenomenological nature however illustrated the clear need for a microscopic description of kinetics and energy relaxation in the strongly interacting limit.

4.4 Nonlinear carbon nanotubes

Recent work by Wang et al. [71] has provided further experimental evidence of nonlinear TLL dynamics beyond the limitations observed in Barak et al. By growing long and clean semiconducting carbon nanotubes (CNTs), the authors were able to utilise low-noise near-field scanning optical measurements to map the plasmonic excitations in a nanotube, even in the

low-density limit. Semiconducting CNTs have a bandgap and as such the linear approximation is expected to break as one approaches the bottom of the conduction band. By coupling the nanotube to a gate electrode, the authors were able to tune the position of the Fermi level and use it to probe the effect of the nonlinear dispersion on the plasmons. Similar measurements were also carried out on metallic CNTs, where the band structure is gapless and so captured well by the linear model.

One of the key predictions of the NLL theory is a drastic reduction of the plasmon lifetime. By injecting a plasmon wave into the nanotube, Wang and collaborators were able to observe the interference between the incident and reflected excitations at one end of the nanotube and from there extracted the quality factor and the wavelength of the mode in the system. This allowed them to measure the increase of the quality factor (and consequently, of the plasmon lifetime) as a function of the electron density, see Fig. 3.17a. By contrast, the linear model predicts infinitely lived excitations, with no intrinsic relaxation mechanism and therefore gate independent. Optical and transport measurements on similar samples also confirmed that the decay was indeed associated with nonlinear dynamics rather than scattering of impurities.

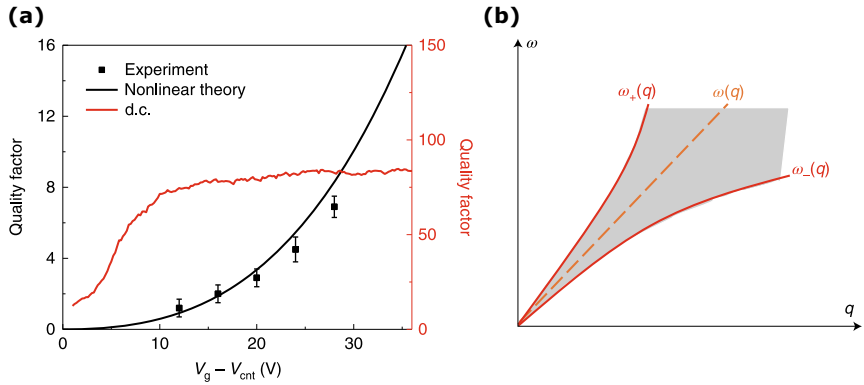


Figure 3.17 Nonlinear Luttinger plasmons in carbon nanotubes. (a) Plasmon quality factor as a function of gate voltage. The experimental results match the NLL well (black) and cannot be captured by alternative mechanisms like impurity scattering (red [grey in print version]). (b) DSF diagram characteristic of the NLL model. In a nonlinear TLL, the plasmon mode is not an exact eigenstate of $\omega(q) = v_p q$ (dashed) but instead is broadened with upper and lower bounds given by $\omega_\pm(q)$, respectively. The spectral width, $\delta\omega(q) = \omega_+(q) - \omega_-(q)$, therefore increases with q, indicating a reduction in the plasmon lifetime, as observed. *From [71], reprinted with permission.*

Their results can be understood from the dynamic structure factor (DSF) of an NLL, see Fig. 3.17b. For a linear TLL, such as in a metallic nanotube, the DSF takes the form $S(q,\omega) = 2K|q|\delta(\omega - v_p q)$. This results in a linear dispersion for the plasmon mode, given by $\omega(q) = v_p q$. Here, K is the Luttinger parameter, v_p is the plasmon velocity (charge mode) and q is its wave vector.

In contrast, for a nonlinear TLL, like in a semiconducting CNT, the DSF is given by

$$S(q,\omega) = 2\frac{\widetilde{m}K}{|q|}\theta\left(\frac{q^2}{2\widetilde{m}} - |\omega - v_p q|\right), \tag{3.28}$$

where $\widetilde{m}$ is an interaction-dependent effective mass and θ is the Heaviside function.

This means that the plasmon mode is no longer an exact eigenstate of the system but instead is broadened, with upper and lower limits given by $\omega_\pm = v_p q \pm q^2/(2m)$. Consequently, for a given frequency, i.e., energy, there will now be multiple available plasmons with different momenta. A high-energy plasmon can then easily decay into multiple low-energy plasmons. The damping and lifetime associated with this process is determined by the width of the broadening $\delta\omega(q) = q^2/\widetilde{m}$.

5. Other work on 1D interaction effects

5.1 Coulomb drag

Free carriers in semiconducting devices screen each other weakly and so the Coulomb interaction between charges is screened over a finite distance, which is however often rather long. This means that when two parallel conducting wires come together, separated only by a thin insulating barrier, a current through one of them can induce a net charge displacement in the other. This is called Coulomb drag. The drag resistivity $R_D = -V_{drag}/I_{drive}$ is determined by the drive current in the inducing wire, I_{drive}, and the measured drag voltage in the induced wire, V_{drag}.

Initially, it was thought that this phenomenon was due to electrons in the drive wire transferring their momentum to their counterparts in the other wire, via the Coulomb interaction, keeping the original direction of motion unchanged. However, recent work by Laroche et al. [72,73] strongly challenged this picture. They found that the drag resistance could be both positive and negative, whilst the momentum transfer model can only account

for the former. Negative Coulomb drag, however, has been predicted to exist in models where the drag is induced by charge fluctuations, therefore introducing a new transport paradigm in addition to the momentum-transfer picture.

Later work by the same group [74] also observed an upturn in drag resistance at a temperature $T^{\star}$ when in the single-subband regime. This confirmed a long-held prediction of the TLL model, wherein a crossover between drag dominated by forward (for $T > T^{\star}$) or backward scattering (for $T < T^{\star}$) should occur, thus establishing the 1D–1D drag dependence on temperature. For comparison, the conductance along a 1D channel in the ballistic regime has only a very weak dependence on T.

5.2 Helical current

Conductance along a ballistic, 1D, quantum wire is quantised in units of $2e^2/h$, where the factor of 2 arises from the spin degeneracy of the electrons. This is not affected by interactions in a clean system, as it is solely determined by the contact resistance to the Fermi leads. In the presence of disorder, however, it is reduced following a TLL power law.

Recent work carried out in the ultra-low-temperature regime ($T \sim 10$ mK) in CEO wires revealed promising new evidence for helical nuclear magnetism in the TLL regime. Tracking the conductance of the first mode as a function of temperature, Scheller et al. [75] observed the expected value of $2e^2/h$ at $T \gtrsim 10$ K drop to e^2/h as the temperature was lowered, becoming independent of T at $T \lesssim 0.1$ K. This behaviour was seen to be independent of both density and magnetic field up to at least 3 T.

After ruling out several other potential interpretations, the authors concluded that the most likely explanation is the appearance of a nuclear spin helix – an ordered state of nuclear spins hosted by atoms of the same crystal – thus explaining the drop in conductance by a factor of 2 at lower temperatures, since the wire now only transmits spin-down right and spin-up left movers. This is consistent with a recent theory proposed by Braunecker et al. [76–78].

5.3 Cold atoms

For completeness, we conclude by commenting on a different platform, outside solid state, for studying the physics of strongly correlated many-body systems. In recent years, experiments have emerged reporting the formation of clean, disorder-free 1D chains by trapping ultra-cold atoms in

optical lattices [79–81]. Here, the confinement potential is given by the laser field itself, therefore allowing the simulation of a very controlled Fermi–Hubbard model not otherwise accessible to material systems. Recent work using quantum gas microscopy, that is, probing the atomic chain with single-atom/single-lattice-site resolution, has revealed signatures of spin-charge separation [82,83], with real-space tracking of the de-confinement and evolution of each separate plasmon having also been reported [84]. These systems are particularly attractive because of the relative ease of varying the strength of the interactions within them, and because they also offer the prospect for tracking crossover from 1D to 2D physics.

6. Conclusion

Tunnelling spectroscopy has proven itself over the past years as a powerful technique capable of mapping the dynamics of many-body systems across the whole spectrum of momentum and energy. This has been shown to be extremely important, for not only has it allowed the confirmation of many of the predictions made by the original, linear TLL model, but also, more recently, it has provided much-needed evidence in support of its non-linear counterparts. Improvements in wafer quality, nanofabrication and tunnelling resolution have also recently allowed for a full mapping to occur at high energies, with early results suggesting signs of physics beyond the Tomonaga–Luttinger paradigm. We anticipate that, provided that two conductors lie close together with a well-controlled tunnel barrier in between, this technique can be applied to study different classes of materials, such as topological insulators, therefore offering itself as a powerful tool for probing exotic single-particle and many-body physics in new systems.

References

[1] Landau LD. J Exp Theor Phys 1957;3:920.
[2] Altland A, Simons BD. Condensed matter field theory. Cambridge University Press; 2010.
[3] Schofield AJ. Contemp Phys 1999;40:95.
[4] Tomonaga S-i. Prog Theor Phys 1950;5:544.
[5] Luttinger JM. J Math Phys 1963;4:1154.
[6] Mattis DC, Lieb EH. J Math Phys 1965;6:304.
[7] Luther A, Peschel I. Phys Rev B 1974;9:2911.
[8] Haldane FDM. J Phys C Solid State Phys 1981;14:2585.
[9] Giamarchi T. Quantum physics in one dimension, international series of monographs on physics. Oxford, New York: Oxford University Press; 2003.
[10] Jin Y. Measurement of electron-electron interactions in one dimension with tunnelling spectroscopy. University of Cambridge; 2020 [Ph.D. thesis].

[11] Giuliani G, Vignale G. Quantum theory of the electron liquid. Cambridge: Cambridge University Press; 2005.
[12] Lake B, Tennant DA, Frost CD, Nagler SE. Nat Mater 2005;4:329.
[13] Mourigal M, et al. Nat Phys 2013;9:435.
[14] Lake B, et al. Phys Rev Lett 2013;111:137205.
[15] Dash LK, Fisher AJ. J Phys Condens Matter 2001;13:5035.
[16] Creffield CE, Klepfish EG, Pike ER, Sarkar S. Phys Rev Lett 1995;75:517.
[17] Creffield CE, Häusler W, MacDonald AH. EPL 2001;53:221.
[18] Häusler W, Kecke L, MacDonald AH. Phys Rev B 2002;65:085104.
[19] Matveev KA. Phys Rev B 2004;70:245319.
[20] Apostol M. J Phys C Solid State Phys 1983;16:5937.
[21] Von Delft J, Schoeller H. Ann Phys 1998;7:225.
[22] Luther A, Emery VJ. Phys Rev Lett 1974;33:589.
[23] Voit J. Phys Rev B 1993;47:6740.
[24] Meden V, Schönhammer K. Phys Rev B 1992;46:15753.
[25] Kardynał B, et al. Phys Rev B 1997;55:R1966.
[26] Abrikosov AA, Gorkov LP, Dzyaloshinski IE. Methods of quantum field theory in statistical physics. New York: Dover Publications Inc.; 1976.
[27] Voit J. J Phys Condens Matter 1993;5:8305.
[28] Auslaender OM, et al. Science 2005;308:88.
[29] Jompol Y, et al. Science 2009;325:597.
[30] Bockrath M, et al. Nature 1999;397:598.
[31] Ishii H, et al. Nature 2003;426:540.
[32] Kim NY, et al. Phys Rev Lett 2007;99:036802.
[33] Kim C, et al. Phys Rev Lett 1996;77:4054.
[34] Gweon G-H, Allen JW, Denlinger JD. Phys Rev B 2003;68:195117.
[35] Zwick F, et al. Phys Rev Lett 1998;81:2974.
[36] Pouget JP, et al. Phys Rev Lett 1976;37:437.
[37] Chang AM. Rev Mod Phys 2003;75:1449.
[38] Grayson M, Tsui DC, Pfeiffer LN, West KW, Chang AM. Phys Rev Lett 1998;80:1062.
[39] Shen ZX, Dessau DS. Phys Rep 1995;253:1.
[40] Kim C. J Electron Spectrosc Relat Phenom 2001;117–118:503.
[41] Kim BJ, et al. Nat Phys 2006;2:397.
[42] Fujisawa H, et al. J Electron Spectrosc Relat Phenom 1998;88–91:461.
[43] Fujisawa H, et al. Solid State Commun 1998;106:543.
[44] Yao Z, Postma HWC, Balents L, Dekker C. Nature 1999;402:273.
[45] Tserkovnyak Y, Halperin BI, Auslaender OM, Yacoby A. Phys Rev B 2003;68:125312.
[46] Kane CL, Fisher MPA. Phys Rev B 1992;46:15233.
[47] Kane C, Balents L, Fisher MPA. Phys Rev Lett 1997;79:5086.
[48] Egger R, Gogolin AO. Phys Rev Lett 1997;79:5082.
[49] Kane CL, Fisher MPA. Phys Rev Lett 1992;68:1220.
[50] Kardynał B, et al. Phys Rev Lett 1996;76:3802.
[51] Altland A, Barnes CHW, Hekking FWJ, Schofield AJ. Phys Rev Lett 1999;83:1203.
[52] Grigera SA, Schofield AJ, Rabello S, Si Q. Phys Rev B 2004;69:245109.
[53] Jin Y, et al. Nat Commun 2019;10:2821.
[54] Bardeen J. Phys Rev Lett 1961;6:57.
[55] Schrieffer JR, Scalapino DJ, Wilkins JW. Phys Rev Lett 1963;10:336.
[56] Smoliner J. Semicond Sci Technol 1996;11:1.
[57] Auslaender OM, et al. Science 2002;295:825.
[58] Carpentier D, Peça C, Balents L. Phys Rev B 2002;66:153304.

[59] Tserkovnyak Y, Halperin BI, Auslaender OM, Yacoby A. Phys Rev Lett 2002;89: 136805.
[60] Deshpande VV, Bockrath M, Glazman LI, Yacoby A. Nature 2010;464:209.
[61] Imambekov A, Glazman LI. Science 2009;323:228.
[62] Imambekov A, Schmidt TL, Glazman LI. Rev Mod Phys 2012;84:1253.
[63] Noziéres P, De Dominicis CT. Phys Rev 1969;178:1097.
[64] Tsyplyatyev O, Schofield AJ. Phys Rev B 2014;90:014309.
[65] Tsyplyatyev O, et al. Phys Rev Lett 2015;114:196401.
[66] Tsyplyatyev O, et al. Phys Rev B 2016;93:075147.
[67] Jin Y, et al. Appl Phys Lett 2021;118:162108.
[68] Vianez PMT, et al. arXiv 2021. 2102.05584.
[69] Moreno M, et al. Nat Commun 2016;7:12784.
[70] Barak G, et al. Nat Phys 2010;6:489.
[71] Wang S, et al. Nat Mater 2020;19:986.
[72] Laroche D, Bielejec ES, Reno JL, Gervais G, Lilly MP. Phys E Low-dimens Syst Nanostruct 2008;40:1569.
[73] Laroche D, Gervais G, Lilly MP, Reno JL. Nat Nanotechnol 2011;6:793.
[74] Laroche D, Gervais G, Lilly MP, Reno JL. Science 2014;343:631.
[75] Scheller CP, et al. Phys Rev Lett 2014;112:066801.
[76] Braunecker B, Simon P, Loss D. Phys Rev B 2009;80:165119.
[77] Braunecker B, Simon P, Loss D. Phys Rev Lett 2009;102:116403.
[78] Braunecker B, Japaridze GI, Klinovaja J, Loss D. Phys Rev B 2010;82:045127.
[79] Pagano G, et al. Nat Phys 2014;10:198.
[80] Boll M, et al. Science 2016;353:1257.
[81] Yang T, et al. Phys Rev Lett 2018;121:103001.
[82] Hilker TA, et al. Science 2017;357:484.
[83] Salomon G, et al. Nature 2019;565:56.
[84] Vijayan J, et al. Science 2020;367:186.

CHAPTER FOUR

Thermoelectric properties of a quantum dot

Gulzat Jaliel*
The London Centre for Nanotechnology, University College London, London, United Kingdom
*Corresponding author: g.jaliel@ucl.ac.uk

Quantum dots (QDs) are model systems for studying quantum thermoelectric behaviour because of their ability to control and measure the effects of electron-energy filtering and quantum confinement on thermoelectric properties. In this chapter, we will discuss the thermoelectric properties of QDs. It starts with a discussion of thermoelectric coefficients in quantum limits in the linear regime, followed by the discussions on the Wiedemann–Franz law, figure of merit and nonlinear responses. At the end, we will briefly look into the applications for using thermoelectric properties of QDs.

1. Landauer–Büttiker formalism of thermoelectricity

The Landauer–Büttiker formalism [1,2] relates the transport properties of a conductor to the transmission probabilities between reservoirs that are in local equilibrium. Let us assume that only two such reservoirs are present. In equilibrium, the reservoirs are at Fermi energy E_F and temperature T. In the regime of linear response, the current I and heat flow J are related to the chemical potential difference $\Delta\mu$ and the temperature difference ΔT by the constitutive equations [3]:

$$\begin{pmatrix} I \\ J \end{pmatrix} = \begin{pmatrix} G & L \\ M & K \end{pmatrix} \begin{pmatrix} \Delta\mu/e \\ \Delta T \end{pmatrix}. \tag{4.1}$$

The thermo-electric coefficients L and M are related by an *Onsager relation,* which in the absence of a magnetic field is

$$M = -LT. \tag{4.2}$$

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00001-0
© 2021 Elsevier Ltd.
All rights reserved.

L is also known as Lorenz ratio, which will be discussed in detail in Section 7. Eq. (4.1) is often re-expressed with the current I rather than the electrochemical potential $\Delta\mu$ as an independent variable [3]:

$$\begin{pmatrix} \Delta\mu/e \\ J \end{pmatrix} = \begin{pmatrix} R & S \\ \Pi & -\kappa \end{pmatrix} \begin{pmatrix} I \\ \Delta T \end{pmatrix}. \tag{4.3}$$

The resistance R is the reciprocal of the isothermal conductance G. The thermopower S is defined as

$$S \equiv \left(\frac{\Delta\mu/e}{\Delta T}\right)_{I=0} = -L/G. \tag{4.4}$$

The Peltier coefficient Π, defined as

$$\Pi \equiv \left(\frac{J}{T}\right)_{\Delta T=0} = M/G = ST \tag{4.5}$$

is proportional to the thermopower S in view of the Onsager relation 1.2. Finally, the thermal conductance κ is defined as

$$\kappa \equiv -\left(\frac{J}{\Delta T}\right)_{I=0} = -K\left(1 + \frac{S^2 GT}{K}\right). \tag{4.6}$$

The thermo-electric coefficients are given in the Landauer–Büttiker formalism by the following [4,5]

$$G = -\frac{2e^2}{h}\int_0^\infty dE \frac{\partial f}{\partial E} t(E), \tag{4.7}$$

$$L = -\frac{2e^2}{h}\frac{kB}{e}\int_0^\infty dE \frac{\partial f}{\partial E} t(E)(E - E_F)/k_B T, \tag{4.8}$$

$$\frac{K}{T} = \frac{2e^2}{h}\left(\frac{k_B}{e}\right)^2 \int_0^\infty dE \frac{\partial f}{\partial E} t(E)[(E - E_F)/k_B T]^2. \tag{4.9}$$

These integrals are convolutions of the transmission probability $t(E)$, which characterises the conductor, and a kernel of the form $\varepsilon^m df/d\varepsilon$, $m = 0,1,2$, with $\varepsilon \equiv (E - E_F)/k_B T$, and f the Fermi function: $f(\varepsilon) = [\exp(\varepsilon) + 1]^{-1}$ [3].

The Landauer–Büttiker formalism has been useful for the interpretation of experiments on low-dimensional systems relating to universal fluctuations [6,7], Aharonov–Bohm oscillations [8], ballistic transport [9,10], the integer quantum Hall effect [11,12] and its quenching at low magnetic fields [13].

2. The quantum dot model

Following the discussions by X. Zianni [14], we will next focus on the calculations of the electron thermal conductance of a QD weakly coupled to two electrode leads in the sequential-tunnelling regime within linear response. Two major simplifications are applied in this model. First, virtual tunnelling processes are neglected; second, the electrostatic energy is described by the classical charging energy: $(Ne)^2/2C$, where N is the number of electrons in the dot and C is the capacitance of the surroundings [15].

We consider a QD that is consisted of a double-barrier tunnel junction, and is weakly coupled to two electron reservoirs via tunnel barriers. Fig. 4.1a shows an equivalent circuit for a QD, and Panel (b) presents a schematic diagram of its energies. Each reservoir is assumed to be in thermal equilibrium and there are a voltage difference V and a temperature difference ΔT between the two reservoirs. A continuum of electron states is assumed in the reservoirs that are occupied according to the Fermi–Dirac distribution:

$$f(E - E_F) = [1 + \exp[(E - E_F)/k_B T]]^{-1}, \tag{4.10}$$

where the Fermi energy, E_F, in the reservoirs is measured relative to the local conduction band bottom.

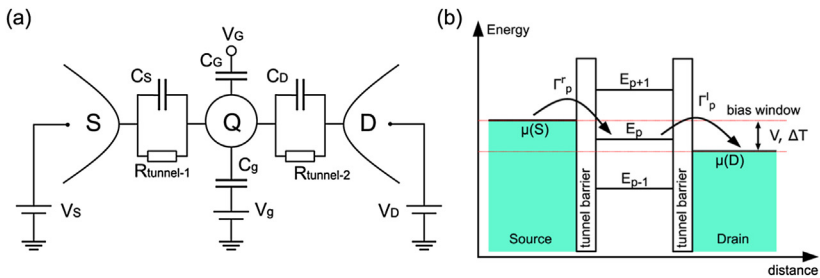


Figure 4.1 (a) Equivalent circuit for a typical quantum dot. The dot has a charge Q, and it is connected by two tunnel barriers to Source and Drain reservoirs, which are at voltages of V_S and V_D, respectively. The electrochemical potential of the dot can be controlled by the voltage of a capacitively coupled gate electrode (V_g). The charging energy is determined by the total capacitance: $C_\Sigma = C_S + C_D + C_g + C_G$. (b) Schematic diagram of the energies of a quantum dot system as a function of distance. Located between the tunnel barriers is the quantum dot, represented as a ladder of states. All the states below the source and drain chemical potentials are filled. Each state, separated by the single-particle energy E_{0D}, can hold an up and a down spin electron. The difference between source and drain voltages creates the so-called bias window: if a state falls within this window, current can flow.

The QD is characterised by discrete energy levels E_p ($p = 1, 2, \ldots$) that are measured from the bottom of the potential well. Degeneracies can be included by multiple counting of the levels. Each level is occupied by N electrons. It is assumed that the energy spectrum is not changed by the number of electrons in the dot. The electrostatic energy $U(N)$ of the dot with charge $-Ne$ is

$$U(N) = (Ne)^2/2C - N\phi_{ext}, \tag{4.11}$$

where ϕ_{ext} is the contribution of external charges. The tunnelling rates through the left and right barriers from the level p to the left and right reservoirs are denoted by Γ_p^l and Γ_p^r, respectively. It is assumed that energy relaxation rates for the electrons are fast enough with respect to the tunnelling rates that we can characterise the state of the dot by a set of occupation numbers, one for each energy level. The transport through the dot can be described by rate equations. It is also assumed that inelastic scattering takes place exclusively in the reservoirs not in the dot.

Due to the voltage difference V and the temperature difference ΔT between the two reservoirs, electric and thermal currents pass through the dot. The stationary current I and the heat flux J through the left barrier are given by the following equations [14]:

$$I = \frac{e}{k_B T}\sum_{p=1}^{\infty}\sum_{N=1}^{\infty}\frac{\Gamma_p^l\Gamma_p^r}{\Gamma_p^r+\Gamma_p^l}P_{eq}(N)F_{eq}\left(E_p/N\right) \times\left[1-f\left(\varepsilon_p-E_F\right)\right]\left[eV-\frac{\Delta T}{T}\left(\varepsilon_p-E_F\right)\right], \tag{4.12}$$

$$J = -\frac{1}{k_B T}\sum_{p=1}^{\infty}\sum_{N=1}^{\infty}\frac{\Gamma_p^l\Gamma_p^r}{\Gamma_p^r+\Gamma_p^l}P_{eq}(N)F_{eq}\left(E_p/N\right)\left[1-f\left(\varepsilon_p-E_F\right)\right] \times\left(\varepsilon_p-E_F\right)\left[eV-\frac{\Delta T}{T}\left(\varepsilon_p-E_F\right)\right], \tag{4.13}$$

where $\varepsilon_p \equiv E_p + U(N) - U(N-1)$, $P_{eq}(N)$ is the probability that the QD contains N electrons in equilibrium and $F_{eq}(E_p/N)$ is the conditional probability in equilibrium that level p is occupied given that the QD contains N electrons. The above equilibrium probabilities are defined as follows:

$$P_{eq}(N) = \sum_{\{n_i\}} P_{eq}(\{n_i\})\delta_{N,\Sigma_i n_i}, \tag{4.14}$$

$$F_{eq}\left(E_p/N\right) = \frac{1}{P_{eq}(N)}\sum_{n_i} P_{eq}(n_i)\delta_{n_{p,1}}\delta_{N,\Sigma_i n_i}. \tag{4.15}$$

$P_{eq}(\{n_i\})$ is the Gibbs distribution in the grand canonical ensemble

$$P_{eq}(\{n_i\}) = Z^{-1} exp\left[-\frac{1}{k_B T}\left(\sum_{i=1}^{\infty} E_i n_i + U(N) - NE_F \right)\right], \tag{4.16}$$

where $N \equiv \Sigma_i n_i$s and Z is the partition function

$$Z = \sum_{\{n_i\}} exp\left[-\frac{1}{k_B T}\left(\sum_{i=1}^{\infty} E_i n_i + U(N) - NE_F \right).\right] \tag{4.17}$$

By comparison of the above definitions of the transport coefficients in the previous section and the linearised expressions for I and J, the following expressions are extracted for coefficients [14–16]:

$$G = \frac{e^2}{k_B T} \sum_{p=1}^{\infty} \sum_{N=1}^{\infty} \gamma_p P_{eq}(N) F_{eq}\left(E_p/N\right)\left\{1 - f\left[E_p + U(N) - U(N-1) - E_F\right]\right\}, \tag{4.18}$$

$$S = -\frac{e}{k_B T^2 G} \sum_{p=1}^{\infty} \sum_{N=1}^{\infty} \gamma_p \left[E_p + U(N) - U(N-1) - E_F\right] P_{eq}(N) F_{eq}\left(E_p/N\right)\left\{1 - f\left[E_p + U(N) - U(N-1) - E_F\right]\right\}, \tag{4.19}$$

$$K = -\frac{1}{k_B T^2} \sum_{p=1}^{\infty} \sum_{N=1}^{\infty} \gamma_p \left[E_p + U(N) - U(N-1) - E_F\right] P_{eq}(N) F_{eq}\left(E_p/N\right)\left\{1 - f\left[E_p + U(N) - U(N-1) - E_F\right]\right\}. \tag{4.20}$$

where $\gamma_p \equiv \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^r + \Gamma_p^l}$.

3. Quantum limit

In the quantum limit, where $\Delta E = E_p - E_{p-1} \gg k_B T$, the discreteness of the energy spectrum of the QD plays a predominant role. In this limit, the term with $N = N_{min}$ gives the dominant contribution to the sums over N in Eqs. (4.12) and (4.13), where N_{min} is the integer that minimises the absolute value of

$$\Delta(N) = E_N + U(N) - U(N-1) - E_F. \tag{4.21}$$

Then, we define $\Delta \equiv \Delta(N_{min})$ and $\Delta_p \equiv E_p - E_{N_{min}}$.

In the quantum limit, the distribution functions for nondegenerate energy levels can be approximated by the following expressions:

$$P_{eq}(N_{min}) = \frac{1}{1 + e^{\Delta/k_B T}}, \tag{4.22}$$

$$F_{eq}(E_p/N_{min}) = \begin{cases} 1, & \text{for } p \leq N_{min}, \\ e^{-\Delta_p/k_B T}, & \text{for } p > N_{min}, \end{cases} \tag{4.23}$$

and

$$1 - f(\Delta_p + \Delta) = \begin{cases} 1, & \text{for } p > N_{min}, \\ e^{(\Delta_p + \Delta)/k_B T}, & \text{for } p > N_{min}, \\ \dfrac{e^{\Delta/k_B T}}{1 + e^{\Delta/k_B T}} & \text{for } p = N_{min}. \end{cases} \tag{4.24}$$

Using the above approximations in Eqs. (4.18) and (4.20) plus the definition in Eq. (4.6), we obtain for G and κ for an equidistant energy-level spectrum $(E_p = p\Delta E)$ and level-independent tunnelling rates, i.e., $\Gamma_p^{l,r} = \Gamma^{l,r}$:

$$G^{QL} = \frac{e^2}{k_B T} \gamma \frac{1}{4 \cosh^2(\Delta/2k_B T)}, \tag{4.25}$$

$$S^{QL} = -\frac{1}{eT}\left[-\frac{\Delta E}{2}(N_C - N_{min}) + \Delta\right] = -\frac{1}{eT}\left[-\frac{\Delta E}{2}\mathrm{Int}\left(\frac{\Delta}{\Delta E}\right) + \Delta\right], \tag{4.26}$$

$$\kappa^{QL} = k_B \gamma \left(\frac{\Delta E}{k_B T}\right)^2 \frac{e^{-\Delta E/k_B T}}{1 + 4\cosh^2(\Delta/2k_B T) e^{-\Delta E/k_B T}}, \tag{4.27}$$

where $\gamma \equiv \frac{\Gamma^l \Gamma^r}{\Gamma^l + \Gamma^r}$. Here, $N_C = max(|N - N_{min}|)$, $Int(x)$ is the integer part of x for $x > 0$ and minus the integer part of $|x|$ for $x < 0$ [14].

4. Coulomb oscillations and thermopower

X. Zianni [14] has plotted the above calculated conductance and thermal conductance, as shown in Fig. 4.2. Fig. 4.2a shows the curves, computed from Eqs. (4.18) and (4.25), which are indistinguishable and hence

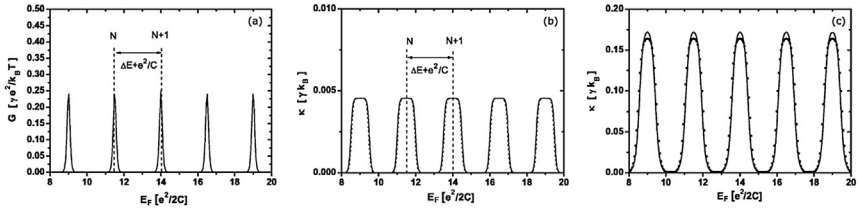


Figure 4.2 (a) Calculated conductance, G, for a series of equidistant, nondegenerate levels with separation $\Delta E = 0.5e^2/2C$ and for $k_BT = 0.05e^2/2C$. Level-independent tunnel rates have been assumed. (b) Calculated electron thermal conductance, κ (*solid line*), for a series of equidistant, nondegenerate levels with separation $\Delta E = 0.5e^2/2C$ and for $k_BT = 0.05e^2/2C$, i.e., ratio $\Delta E/k_BT = 10$. Level-independent tunnel rates have been assumed. For comparison, κ^{QL} (dots) is also shown. (c) As in panel (b) for $\Delta E/k_BT = 5$, with $\Delta E = 0.5e^2/2C$ and $k_BT = 0.1e^2/2C$. *This figure is adopted from Ref. [14].*

only one curve is shown. The calculated thermal conductance is plotted in Fig. 4.2b and c, for two values of the ratio $\Delta E/k_BT$ in the quantum limit. The thermal conductance calculated from Eq. (4.27), κ^{QL}, is plotted together with the thermal conductance, κ, extracted from Eq. (4.6) using Eqs. (4.18)–(4.20) as a more general case for comparison. In Fig. 4.2b, where $\Delta E/k_BT = 10$, κ^{QL}, and κ are in perfect agreement. In Fig. 4.2c, where $\Delta E/k_BT = 5$, κ^{QL} starts deviating from κ. This is because when the energy-level spacing approaches the thermal energy, the quantum limit assumptions become less accurate. Hence, the quantum limit analytical formula holds satisfactorily when the energy-level spacing is at least one order of magnitude higher than the thermal energy [14].

In Fig. 4.2, it is shown that in the quantum limit the conductance and the thermal conductance exhibit periodic Coulomb blockade oscillations. The peaks occur each time an extra electron enters in the dot and they are separated by intervals: $\Delta E_F = \Delta E + e^2/C$. The same periodicity originates from the fact that $|\Delta|$ is a periodic function of E_F varying from zero to a maximum at intervals of ΔE_F. G and κ depend on Δ through the same function of the hyperbolic cosine as can be seen in Eqs. (4.25) and (4.27). The peaks in κ and G occur at the values of E_F for which $\Delta = 0$ holds. The dependence of the maxima of G and κ on the characteristic parameters can be deduced from Eqs. (4.25) and (4.27), and it can be seen that the two transport properties exhibit different behaviour: G_{max} decreases linearly with increasing thermal energy and it is nearly independent of the

energy-level spacing; k_{max} depends on both the thermal energy and the energy-level spacing through their ratio, $\Delta E/k_B T$, and it decreases rapidly with decreasing temperature and increasing energy-level spacing [14].

The thermal conductance behaviour is dominated by the effect of quantum confinement due to the dependence on the ratio of the energy-level spacing over the thermal energy, $\Delta E/k_B T$. This is analytically described by Eq. (4.27) and it is graphically shown by comparing the data plotted in Fig. 4.2. The electron thermal conductance of QDs decreases nearly exponentially with decreasing temperature. The dependence on the energy level separation, ΔE, in Eq. (4.27) shows an equally fast decrease as the size of the dot decreases. Hence, it is explicitly shown that quantum confinement is responsible for the fast decrease of the electron thermal conductance [14].

Oscillations can be found in thermopower [15]. Since Δ depends linearly on E_F, Eq. (4.26) tells us that the thermopower in the quantum limit is a piece-wise linear function of the Fermi energy, with a derivative $dS/dE_F = 1/eT$ which is twice as large as the classical slope $1/2eT$ [15]. Two kinds of oscillations are observed in the thermopower: the *long-period oscillations* are due to changes in the number of electrons on the dot in the ground state. Their periodicity $\Delta E_F = \Delta E + e^2/C$ is determined by differences in ground-state energies; and the *short-period oscillations* have periodicity $\delta E_F = \Delta E$ determined by energy differences between the ground state and excited states of the QD for a constant number of electrons on the dot. Since $|\Delta \leq \frac{1}{2}(\Delta E + e^2/C)|$, Eq. (4.26) reduces to $S = -\Delta/eT$ if $e^2/C \leq \Delta E$. The thermopower thus has only the long-period oscillations (consisting of a sawtooth with periodicity $\Delta E + e^2/C$ and amplitude $(\Delta E + e^2/C)/eT$), if the charging energy is less than the level spacing. The short-period oscillations appear as soon as $e^2/C > \Delta E$. For $e^2/C \gg \Delta E$, the short-period oscillations are a fine structure on the envelope $S_{envelope} = -\Delta_{min}/2eT$, obtained from Eq. (4.26) by replacing $Int(x)$ by x.

C. Beenakker and A. Staring [15] studied the thermopower of a QD and illustrated the thermopower calculated from Eqs. (4.19) and (4.26) in Fig. 4.3. Fig. 4.3a presents the results computed from Eq. (4.19) for parameters between the classical and the quantum regime. The quantum limit is shown in Fig. 4.3b. The conductance, calculated via Eq. (4.25), is plotted in the same figure, for comparison (dotted curve). The long-period oscillations in thermopower correspond well to that of the conductance, but a corresponding fine structure in the conductance, to the short-period oscillation, is invisible. This is because, even though theoretically it can be extracted from Eq. (4.25) with periodicity δE [15], its amplitude is exponentially small on the scale of the conductance peak itself [15].

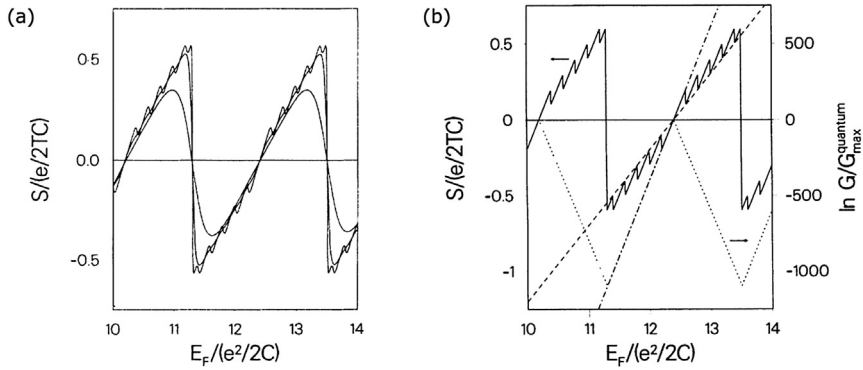


Figure 4.3 (a) Development of fine structure on the thermopower oscillations on lowering the temperature from $k_BT = (0.2–0.05–0.01) \times e^2/2C$. The curves are computed from Eq. (4.19), for a series of equidistant nondegenerate levels with $\Delta E = 0.2e^2/2C$, taking level-independent tunnel rates. (b) Low-temperature limit of the thermopower oscillations of panel (a) (solid curve). The conductance oscillations are shown for comparison (*dotted*). The curves are computed from Eqs. (4.18) and (4.19), for a series of equidistant nondegenerate levels with $\Delta E = 0.2e^2/2C, k_BT = 0.0001e^2/2C$, and taking level-independent tunnel rates. The *dash-dotted line* has slope $dS/dE_F = 1/eT$, and the *dashed line* has a slope which is twice as small [15].

A. Staring et al. [17] presented a clear experimental demonstration of the key characteristics of the thermopower oscillations of a QD, defined in the two-dimensional electron gas in a $GaAs_x/AlGa_{1-x}As$ heterostructure, using a current heating technique. As shown in Fig. 4.4a, they compared measurements of the Coulomb blockade oscillations as a function of the gate voltage

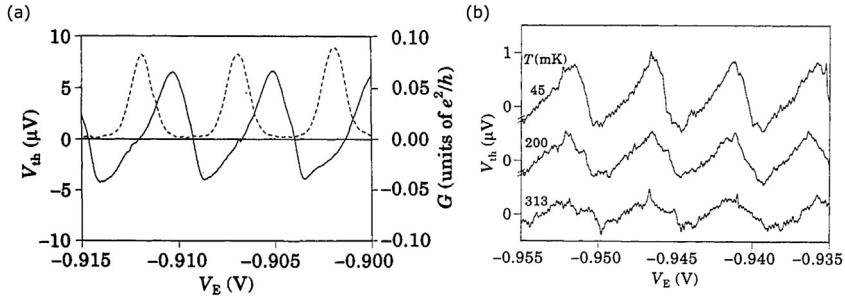


Figure 4.4 (a) Thermovoltage V_{th} at a heating current of 58 nA (*solid line*) and conductance (*dashed line*) as a function of V_E at a lattice temperature of $T = 45$ mK. (b) Thermovoltage V_{th} as a function of V_E at lattice temperature of $T = 45$, 200 and 313 mK, obtained using a heating current of 18 nA. *This figure is adopted from [17].*

V_E in the thermovoltage (solid) and conductance (dashed) of the dot, at a lattice temperature of $T = 45$ mK. Clearly, the thermovoltage V_{th} (and therefore S) oscillates periodically. The period is equal to that of the conductance oscillations and corresponds to depopulation of the dot by a single electron. In contrast to the conductance oscillations, which consist of a series of symmetric peaks separated by gate voltage regions where the conductance is suppressed, the thermovoltage oscillations have a distinct sawtooth-like lineshape. In addition, the oscillations are approximately centred on the positive slop of the thermovoltage oscillations, with the steeper negative slope occurring between two conductance peaks. No shift in gate voltage is found when the thermovoltage and thermopower are measured alternatingly [17]. When both charge and heat are exclusively carried by electrons, it has been shown both for diffusive and ballistic transport that the thermopower S is related to the energy derivative of the conductance G [18]:

$$S = \left.\frac{\Delta V}{T_C - T_L}\right|_{I=0} = -\frac{\pi^2 k_B^2}{3e}(T_C + T_L)\frac{\partial \ln G}{\partial E_F}. \tag{4.28}$$

where μ is the chemical potential of the contacts relative to 2-dimensional electron gas (2DEG) regions of interest. It is assumed that features in the thermopower are not subject to thermal broadening, and that the electrons are noninteracting [18]. This is also called the *Mott relation*.

A. Staring et al. [17] also measured the lattice temperature dependence of the thermovoltage oscillations, as shown in Fig. 4.4b, where traces of V_{th} vs. V_E are given for three different temperatures (obtained using heating current of $I = 18$ nA). The sawtooth lineshape of the thermopower oscillations is most pronounced at the lowest lattice temperature of $T = 45$ mK and is gradually replaced by a more symmetric lineshape as T is increased to 313 mK. If the temperature is further increased to 750 mK, the oscillations are no longer observable using $I = 18$ nA, due to the large noise level. However, they still observed the oscillation up to $T \leq 1.5K$, when larger heating currents are used (not shown). The thermopower measurements, therefore, suggest an effective approach to deducing the electron temperature in a 2DEG [17].

The thermopower has been measured in small dots with discrete level spectrum [19], chaotic dot [20], carbon nanotubes and molecules [21] and dots close to metallic (quasi-continuous) limit [17], and calculated for various mesoscopic systems [15].

5. The effect of degeneracy

A twofold degeneracy of the energy levels (e.g., spin degeneracy) has been investigated by X. Zianni [14] and Beenakker et al. [15]. Coulomb interaction lifts this degeneracy. The distribution functions in this case depend on whether there are an even or an odd number of electrons on the dot.

X. Zianni calculated the conductance and the thermal conductance in Ref. [14], and the results are shown in Fig. 4.5a and b, respectively. The Coulomb blockade peaks of the electrical conductance [15] and of the thermal conductance are nonsymmetric and they appear in doublets with mirror symmetry [14]. The peaks of the conductance occur each time an extra electron enters the dot, when $\Delta = 0$. The first peak of the doublet appears when an odd electron is added and the second peak appears when an even electron is added. The peaks are separated by $\Delta E_F = \Delta E + e^2/C$ when an odd electron enters the dot and by $\Delta E_F = e^2/C$ when an even electron enters the dot. In Fig. 4.5b, the thermal conductance through doubly degenerate levels is plotted together with that from Eq. (4.27) so that the effect of spin degeneracy becomes more evident. (It should be noted that ΔE and the amount of confinement are the same in both cases.) It is due to the double degeneracy of the lower (upper) neighbouring energy level that offers an additional channel of conduction; when an odd (even) electron is added, the left (right)

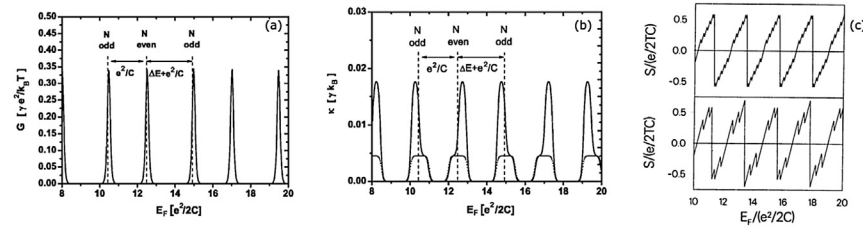


Figure 4.5 (a) Calculated conductance, G, for a series of equidistant, doubly degenerate levels with separation $\Delta E = 0.5e^2/2C$ and for $k_BT = 0.05e^2/2C$ [14]. (b) Calculated electron thermal conductance, κ (solid line), for a series of equidistant, doubly degenerate levels with separation $\Delta E = 0.5e^2/2C$ and for $k_BT = 0.05e^2/2C$. The dotted curve is for nondegenerate levels and is shown for comparison [14]. (c) Effect of spin degeneracy on the fine structure. The curves are computed from Eq. (4.19), at a temperature $k_BT = 0.001e^2/2C$, for a series of equidistant levels with level-independent tunnel rates. In the top panel, the levels are nondegenerate, with spacing $\Delta E = 0.2e^2/2C$ (as in Fig. 4.3). In the bottom panel, the levels are twofold degenerate, with spacing $\Delta E = 0.4e^2/2C$ (so that the average density of states is the same) [15].

half of the peak is enhanced. Due to spin degeneracy, more channels of conduction contribute to the thermal conductance and give an enhancement of the thermal conductance [14].

Fig. 4.5c shows the effect of spin degeneracy of the energy levels on the thermopower [15]. The average density of states is the same in the top and bottom panel, but in the top panel, the levels are nondegenerate, whilst in the bottom panel, each level is twofold degenerate. This change in the energy-level spectrum of the QD has essentially no effect on the frequency of the long-period oscillations, whilst the frequency of the short-period oscillations is changed by a factor of 2. Notice also that the fine structure differs from one oscillation to the other in the case of spin-degenerate levels, reflecting the fact that the excitation spectrum is different if there are an even or odd number of electrons on the dot [15].

6. Power factor and figure of merit

The knowledge of the conductance and the thermopower allows us to compute another important thermoelectric characteristic – the thermoelectric figure-of-merit

$$ZT = \frac{S^2 G T}{\kappa}. \tag{4.29}$$

for a QD and the corresponding power factor Q, which is defined in terms of the electric power P generated by ΔT as

$$P = V^2/R = Q(\Delta T)^2, Q = S^2 G, \tag{4.30}$$

where $R = 1/G$ is the electric resistance of the QD. Within the linear response regime discussed above, Q and ZT are the keys to characterising the performance of a thermoelectric device. From the expressions of the transport coefficients in Eqs. (4.25)–(4.27), specified for the two-terminal case, Erdman et al. [22] have computed Q and ZT analytically within the quantum limit.

Within the quantum limit, Q is found to be [22]

$$Q = \frac{\gamma}{4k_B T^3 \cosh^2(\Delta/2k_B T)} \left(\Delta + \frac{N_C - N_{min}}{2} \Delta E \right)^2, \tag{4.31}$$

where $\gamma = \left(\frac{\Gamma^l \Gamma^r}{\Gamma^l + \Gamma^r} \right)$, N_C is N at the largest value of $|N - N_{min}|$. As for G, the power factor Q is dominated by a fast decrease, given by the term

$\cosh^{-2}(\Delta/2k_BT)$, thus becoming vanishingly small within a few k_BT around $\mu = \mu_N$ (here, μ is the electrochemical potential of the dot) [see Fig. 4.6a]. In fact, the fine structure, given by the terms with N_J, is not visible in Fig. 4.6a due to the rapid suppression given by the $\cosh^{-2}(\Delta/2k_BT)$ terms in Eq. (4.31). Differently from G, the power factor vanishes at $\mu = \mu_N$, due to the fact that in this point the thermopower $S = 0$. So as μ moves away from μ_N, Q increases quadratically due to the linear growth of the thermopower with μ, and then it rapidly decreases within a few k_BT due to the $\cosh^{-2}$ $(\Delta/2k_BT)$ term. Hence, there are two symmetric peaks around $\mu = \mu_N$, within a few k_BT. These double peaks are the dominant feature of Fig. 4.6a and identify the optimal values of Δ (and consequently of μ) to obtain the absolute maximum power P_{peak}, namely when the power factor Q is maximum, $Q = Q^*$. From Eq. (4.31), we obtain that Q is maximum for values Δ^* of Δ such that [22]

$$\frac{\Delta^*}{2k_BT} = \coth\left(\frac{\Delta^*}{2k_BT}\right) \tag{4.32}$$

The numerical solution is $\Delta^* \simeq \pm 2.40 k_BT$, which corresponds to $\mu = \mu_N \pm 2.40\ k_BT$. This result does not depend on any energy scale of the system except for k_BT. The value of Q^* of Q in these points is [22]

$$Q^* \simeq 0.44\frac{\gamma k_B}{T}, \tag{4.33}$$

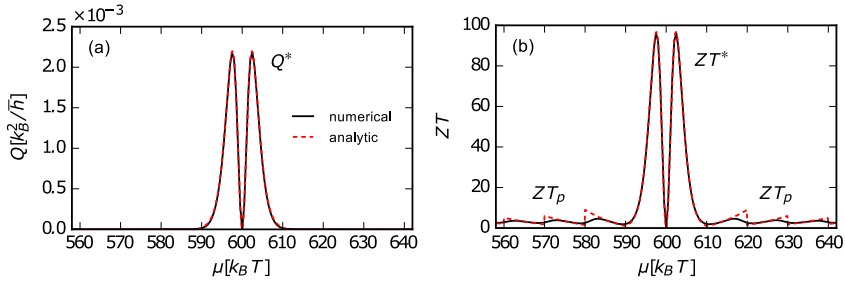


Figure 4.6 Power factor Q (a) and figure of merit ZT (b) are plotted as a function of the electrochemical potential μ. For both quantities the analytical quantum limit [given by Eqs. 4.31 and 4.35] is plotted as a red (grey in print version) dashed curve, whilst the numerically calculated result is plotted as a black solid curve. All curves are calculated at $E_C = 50k_BT$, $\Delta E = 10k_BT$, and $\hbar\Gamma_1(p) = \hbar\Gamma_2(p) = (1/100)k_BT$. *This figure is adopted from Ref. [22].*

so that the peaks of the power factor only depend on γ and on the reference temperature. In conclusion, if one wants to extract maximum power from this system, one must choose $\mu = \mu_N \pm 2.40 k_B T$ [22]. This can be confirmed by the calculation of ZT, which reaches a maximum at these same values of the electrochemical potential.

Erdman et al. [22] have also calculated the figure of merit ZT in the quantum limit. They obtained a more manageable analytical expression, by expanding K to the first order in $4\cosh^{-2}(\Delta/2k_BT)/e^{\Delta E/k_BT}$, as shown below [22]

$$ZT = \frac{1}{4}\left(\frac{\Delta}{\Delta E}\right)^2 \frac{e^{\Delta E/k_BT}}{\cosh^2\left(\frac{\Delta}{2k_Bt}\right)}, \tag{4.34}$$

which, together with Eq. (4.31), implies that ZT is directly proportional to Q and therefore it has the same double peak structure at $\mu = \mu_N \pm 2.40 k_B T$. This is clearly shown in Fig. 4.6b where ZT is plotted as a function of μ. The value of ZT in these points is [22,23]

$$(ZT)^* \approx 0.44 \frac{e^{\Delta E/k_BT}}{(\Delta E/k_BT)^2}. \tag{4.35}$$

This shows that in the limit $\Delta E/k_BT \to \infty$, $ZT \to \infty$. This is consistent with Mahan and Sofo's observation [24] that a narrow transmission function yields $ZT \to \infty$. Furthermore, these peaks in ZT correspond to peaks in Q, so in these points, P_{max} and η (P_{max}) are maximised simultaneously. Instead, when $N_C \neq N_{min}$, ZT has a discontinuity every time $\mu = e(N,p)$ with $p \neq N$, which means with a ΔE spacing.

The fine structure is the origin of the sawtooth oscillations of Fig. 4.6b. The value of ZT in each $\mu = e(N,p)$ is given by

$$ZT = \begin{cases} 3\dfrac{|N-P|+1}{|N-P|-1} & \text{if} \quad |N-P| \geq 2, \\ 1, & \text{if} \quad |N-P| = 1. \end{cases} \tag{4.36}$$

The height of these peaks, as opposed to ZT^*, has no dependence on the parameters of the system. The highest peak is obtained for $|N-p| = 2$, where $ZT_{p=N\pm2} = 9$. For values of p distant from N, the height of the peak decreases to an asymptotic value of $ZT_\infty = 3$ [22].

The dimensionless figure of merit ZT is a measure of the efficiency of a device to convert heat into electricity and vice versa. Good thermoelectric materials are considered to be those with $ZT > 3$ at room temperature [23]. Early theoretical work pointed out the possibility of enhancing ZT in structures of reduced dimensionality due to increased density of states near the Fermi level [25]. Mahan and Sofo [24] predicted a maximisation of the thermoelectric efficiency in materials with a delta-like density of states. This suggests that QDs and molecular junctions are promising candidates for good thermoelectric materials.

7. Violation of the Wiedemann–Franz law

For macroscopic samples of ordinary metals, the Wiedemann–Franz law provides a universal relation between the two conductances by stating that the Lorenz ratio

$$L \equiv \frac{\kappa}{GT'} \tag{4.37}$$

is a constant given by the Lorenz number $L_0 = (\pi^2/3) \times (k_B/e^2)$. L is also a thermoelectric coefficient in Eq. (4.1). It is a consequence of Fermi liquid theory, which is applicable when screening renders Coulomb interactions sufficiently weak. The Wiedemann–Franz law indicates that both charge and heat currents are supported by the same underlying scattering mechanisms with only weak energy dependence [26]. The situation is fundamentally different in mesoscopic systems in which level quantisation and Coulomb interaction drastically affect transport. Deviations from the Wiedemann–Franz law have been predicted for tunnelling transport through QDs for weak coupling [14,23], in the Kondo regime [14,27] and for open dots [28].

Erdman et al. [22] computed the Lorenz ratio for an interacting quantum-dot-based heat engine, using Eqs. (4.18)–(4.20), and they found

$$L = \frac{L_0}{\pi^2}\left(\frac{\Delta E}{k_B T}\right)^2 12\exp(-\Delta/k_B T)\cosh^2\left(\frac{\Delta}{k_B T}\right), \tag{4.38}$$

Obviously, the Wiedemann–Franz Law is strongly violated. The Lorenz ratio increases exponentially with Δ. At $\Delta = 0$, it is exponentially smaller than L_0 thanks to the factor $(\Delta E/k_B T)^2/\exp(\Delta E/k_B T)$. This violation has also been investigated in various reports [23,26,29].

Kubala et al. discussed the violation of the Wiedemann–Franz law in a single-electron transistor [26], and showed that the Coulomb interaction with weak tunnel couplings affects the Wiedemann–Franz law in two ways. First, the finite charging energy to add or remove an electron to or from the island suppresses some transport processes. This dramatically affects the charge and thermal conductance individually, but leaves the Wiedemann–Franz law untouched since the same transport processes are suppressed for both electric and thermal conductance. Second, however, Coulomb interaction leads to a strong energy dependence of the scattering processes. This yields, in general, a violation of the Wiedemann–Franz law [26].

Fig. 4.7 shows the analytical results for the dimensionless thermoelectric coefficients, $g_V, g_T = M,K$ (here, g_V, g_T are G and L in Eq. 4.1), and the Lorenz ratio L normalised by L_0, as a function of gate voltage for various temperatures, as well as the temperature dependence of L/L_0 for various gate voltages. Temperature and gate-voltage dependence of the Lorenz ratio can be elucidated by deriving analytical expressions for various limits. (i) In the high-temperature regime, $E_C/k_BT \ll 1$, Coulomb oscillations are washed out, i.e., there is no gate-voltage dependence. To calculate

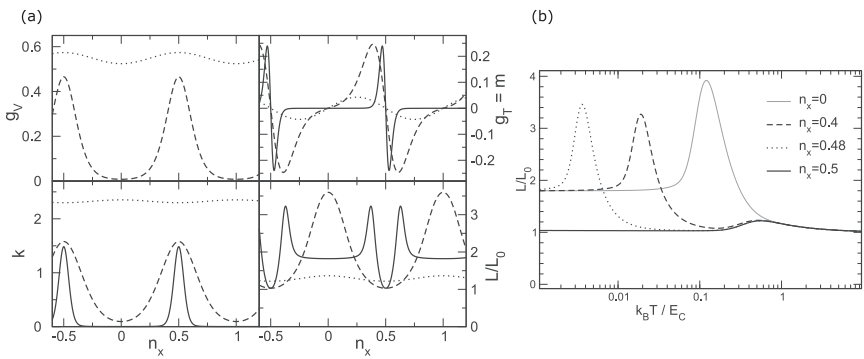


Figure 4.7 (a) Coulomb blockade oscillations of thermoelectric coefficients and the Lorenz ratio. For high temperatures ($k_BT = E_C/2$ – *dotted line*), oscillations are washed out. In the sequential-tunnelling regime ($k_BT = E_C/10$ – *dashed line*), the Lorenz ratio is given by Eq. 4.40 around each resonance. For low temperatures ($k_BT = E_C/40$ – *solid line*), the new universal Lorenz ratio $9/5L_0$ is reached in the co-tunnelling regime. (b) Temperature dependence of Lorenz ratio for different gate voltages. Two maxima separate different tunnelling regimes. The rise to the maxima starts at $k_BT \approx E_C$ and $k_BT \approx \Delta_0$, respectively. For $n_x = 0 \Leftrightarrow \Delta_0 = E_C$, the two maxima coincide, whereas at resonance ($n_x = 5 \Leftrightarrow \Delta_0 = 0$) the lower maximum is moved to the left. *The figure is adapted from Ref. [26].*

corrections to the Wiedemann–Franz law in this regime, they expanded the gate-voltage average of all thermo-electric coefficients in powers up to $(E_C/k_BT)^2$ to find

$$\frac{L}{L_0} = 1 + \frac{2}{\pi^2} E_C/k_BT - \frac{1+24\alpha}{3\pi^2}(E_C/k_BT)^2. \tag{4.39}$$

Deviations from the Wiedemann–Franz law are visible before Coulomb oscillations set in, see Fig. 4.7b, where for $k_BT \geq E_C$, the curves coincide for all gate voltages whilst $L > L_0$. (ii) In the on-resonance low-temperature regime, $E_C/k_BT \gg 1$ but $\Delta N \ll 1$ for one N (say $N = 0$), transport is dominated by sequential tunnelling and only the charge states 0 and 1 occur. The sequential-tunnelling contribution then yields

$$\frac{L}{L_0} = 1 + (\Delta/k_BT)^2/(2\pi)^2. \tag{4.40}$$

The Wiedemann–Franz law is only fulfilled for a vanishing charging energy gap, $\Delta = 0$, with corrections quadratic in Δ/k_BT away from resonance. These corrections indicate that the contribution of each transported particle to the heat current scales with the charging energy gap Δ instead of temperature k_BT as in bulk transport. In the off-resonance low-temperature regime, transport is dominated by co-tunnelling. Because of the weak energy dependence of the co-tunnelling scattering rate, proportionality between charge and heat conductance is recovered, however, with a different pre-factor, as $L/L_0 = 9/5$.

8. Nonlinear regime

So far, we have discussed the thermoelectric properties of a QD in the linear regime. However, as discussed for instance in Ref. [30], nonlinear effects are important in nanoscopic setups, since the temperature difference is applied across very small elements of the order of tens or hundreds of nanometers. As far as heat-to-work conversion is concerned, there is a practical reason to consider the nonlinear response, namely the efficiency and power output, of a QD as a heat engine, may increase with increasing temperature difference. Furthermore, for systems with time-reversal symmetry, the efficiency at maximum power can overcome the limit of $\eta_C/2$ only beyond the linear response [22].

Following the discussions by Erdman et al. [22], this section presents the discussions of the thermoelectric properties in the nonlinear regime of a

two-terminal QD system. First, because of the charge current conservation, the charge current from the left to the right is equal to that from the right to the left, $I \equiv J_l = -J_r$. Then, the average reservoir temperature is defined as $\overline{T} = (T_1 + T_2)/2$, which determines the typical thermal energy scale of the system beyond linear response (all energies will be given units of $k_B\overline{T}$). Furthermore, $\Delta\mu \equiv \Delta\mu_2 = eV$, with V applied voltage, $\Delta T \equiv \Delta T_2$, and assume equidistant energy levels with spacing given by ΔE. In order to describe the potential drop between the QD and the two reservoirs, the set of energy levels is assumed to be shifted as $E_p\ (V) = E_p + (1 - \theta_0)eV$, where $0 \leq \theta_0 \leq 1$ is the fraction of potential V that drops over the tunnel barrier which couples reservoir 2 to the QD.

In the nonlinear regime, the thermopower (Seebeck coefficient) can be defined as follows

$$S = -\frac{V}{\Delta T}|I = 0, \tag{4.41}$$

i.e., as the ratio between the thermovoltage V that develops as a result of a finite ΔT applied, at open circuit ($I = 0$). In Fig. 4.8a, S is plotted as a function of μ for various values of $\delta T/\overline{T}$. The black solid curve (calculated for $\delta T/\overline{T} = 10^{-4}$) is the linear response reference that is well approximated by the expression given by Eqs. (4.25)–(4.27) [22]. As discussed in Section 4, the black solid curve presents main oscillations of period $\Delta E + 2E_C$, and a fine structure with a ΔE spacing [15]. Since an equidistant energy spectrum

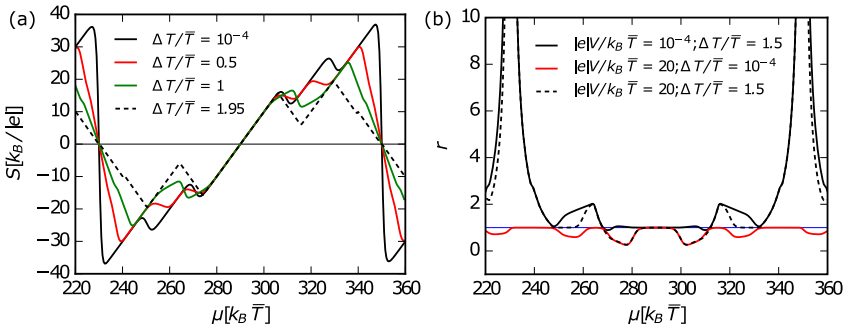


Figure 4.8 (a) Nonlinear thermopower S plotted as a function of μ for various values of $\Delta T/\overline{T}$, and $\Delta E = 20k_B\overline{T}$, $E_C = 50k_B\overline{T}$, $\hbar\Gamma^l(p) = \hbar\Gamma^r(p) = 0.01k_B\overline{T}$, $\theta_0 = 1/2$. (b) Ratio $r = (\Pi + V/2)/(\overline{T}S)$ (for the Peltier coefficient) and of $\Delta T/\overline{T}$ (for the thermopower), for the same parameter values as in (a). The blue (dark grey in print version) thin line is the reference $r = 1$. *This figure is adopted from Ref. [22].*

has been chosen, all curves share a number of features with the linear response reference. Namely, (i) S crosses zero with positive slope at the main transition energies μ_N and is periodic with periodicity $\Delta E + 2E_C$; (ii) in the range of μ considered, S is anti-symmetric with respect to $\mu = 290k_B\overline{T}$; (iii) S vanishes for values of μ in the middle points between two dominant transitions μ_N and μ_{N+1} (in Fig. 4.8a, such points are located at $\mu = 230k_B\overline{T}$ and $\mu = 350k_B\overline{T}$.) Moreover, since we set $\theta_0 = 1/2$, the linear increase of S for $\mu \simeq \mu_N$ does not depend on the ratio $\Delta T/\overline{T}$, i.e., it is well described by the linear response proportionality coefficient $1/(|e|\overline{T})$. Interestingly, such features (except for the fine structure oscillations) can be understood in terms of a noninteracting model, which also explains the reduction of the negative slope of S at the middle points as the ratio $\Delta T/\overline{T}$ increases [22].

Let us now discuss the behaviour of S when departing from the linear response regime. Fig. 4.8a shows that for all values of ΔT the thermopower deviates from the linear response curve only for μ above $310k_B\overline{T}$ (or below $270k_B\overline{T}$). A sharp departure already occurs at $\Delta T/\overline{T} = 0.5$ (red [grey in print version] curve). This can be understood from the fact that, for $\mu > 310k_B\overline{T}$, S is of the order of 15 $k_B/|e|$ which corresponds to a value of the thermovoltage ($V = -7.5k_B\overline{T}/|e|$) such that $|eV| \gg k_BT$ [22]. Note that $\mu = 310\ k_B\overline{T}$ roughly corresponds to the first step of the fine structure in linear response. In particular, whilst the first step hardly moves by increasing ΔT from its position in linear response, the second step, occurring at $\mu = 330k_B\overline{T}$ in the linear response, shifts to a smaller value for increasing $\Delta T/\overline{T}$, eventually disappearing or merging with the first step. This behaviour may be attributed to the combination of the following two effects. On one hand, the thermovoltage V, which determines the transport energy window, depends on μ and increases with ΔT according to the definition Eq. (4.41). On the other hand, an increase of $\Delta T/\overline{T}$ moves the lowest temperature (T_1) towards absolute zero, thus sharpening the Fermi distribution function $f_1(E)$. This last effect is also responsible for the sharpening of $S(\mu)$ as $\Delta T/\overline{T}$ increases. Note furthermore that the extremal values of S decrease as $\Delta T/\overline{T}$ increases [22].

Let us now consider the nonlinear Peltier coefficient defined, for a given voltage V, as

$$\Pi = \frac{J}{I}\Big|\Delta T = 0. \tag{4.42}$$

The aim is to assess the failure of the Onsager reciprocity relation $\Pi = TS$, which holds in the linear response regime. Beyond linear response, for a single-level noninteracting QD, one finds a 'corrected' reciprocity relation, namely $\Pi + V/2 = TS$ in the case where $\theta_0 = 1/2$. To single out the effect of interactions in a multi-level QD, in Fig. 4.8b, the ratio $r = (\Pi + V/2)/(TS)$ is plotted as a function of μ for various values of $\Delta T/T$ and $|e|\ V/(k_BT)$ ($\Delta T/T$ is used to compute S, whilst $|e|V/(k_B\mathrm{T})$ to compute Π). Fig. 4.8b shows that the ratio r departs significantly from 1, the linear response result (blue [dark grey in print version] thin line), only far enough from the dominant transition energy $\mu_{N=3} = 290\ k_B\overline{T}$. In particular, when Π is in the linear response regime and S is not (black solid curve) the strong deviations occurring for μ around the middle points between dominant transition energies, namely $\mu = 230k_B\overline{T}$ and $\mu = 350k_B\overline{T}$, can be explained by a two-level noninteracting model. However, the deviations occurring in the range of values of μ between $250k_B\overline{T}$ and $330\ k_B\overline{T}$ can be attributed to interaction effects. In the opposite case, where S is in the linear response and Π is not (red [grey in print version] curve), the deviations of r from 1 are entirely due to interaction effects and r takes values between 0 and 1 in the entire range of values of μ. When S and Π are beyond linear response the two behaviours discussed above co-exist giving rise to the black dashed curve.

In order to properly describe nonlinear thermoelectric effects in mesoscopic structures, a nonlinear scattering matrix theory of thermoelectric transport has been developed recently [31–33]. It has been applied to study heat transport through QDs where, e.g., a deviation from the Wiedemann–Franz law was found [34]. It was also used to investigate the magnetic field asymmetry of nonlinear transport coefficients [35]. Experimentally, the nonlocal thermoelectric response of a ballistic four-terminal structure has been investigated [34]. Further studies theoretically analysed nonlinear thermoelectric transport through molecular junctions [36,37].

9. Output power and efficiency

An important feature of nonlinear thermoelectrics is that the figure of merit ZT is no longer sufficient to characterise the thermoelectric performance [30,38]. Instead, one has to rely on quantities such as the maximal power, the efficiency at maximum power [38–40] or the maximal efficiency at a given output power [41].

Under steady-state conditions, the output power P of a multi-terminal QD system is given by the sum of all the heat currents

$$P = \sum_{\alpha=1}^{N} J_\alpha. \tag{4.43}$$

If $P > 0$, the system behaves as a heat engine, i.e., converting heat into work. In this situation, the efficiency η is defined as the ratio between the output power and the total heat current absorbed by the system

$$\eta = \frac{P}{\Sigma_\alpha + J_\alpha +}, \tag{4.44}$$

where the sum over α^+ runs over all positive heat currents. For a two-terminal system the efficiency cannot exceed the Carnot efficiency, defined as $\eta_C = 1 - T_1/T_2$, with $T_2 > T_1$. In addition, for a multi-terminal system, η cannot go beyond the two-terminal Carnot efficiency calculated using the hottest and coldest temperatures amongst $T_1, T_2, \ldots, T_N$ [42].

For a two-terminal system within the linear response regime, i.e., when the charge and heat currents depends linearly on the temperature and electrochemical potential differences, both the output power and efficiency can be written in terms of the transport coefficients in Eqs. (4.18)–(4.20). Defining $\Delta T \equiv \Delta T_2 > 0$, we have the following relations [22]:

$$P_{max} = \frac{1}{4} Q \Delta T^2, \tag{4.45}$$

$$\eta(P_{max}) = \frac{\eta_C}{2} \frac{ZT}{ZT + 2'} \tag{4.46}$$

$$\eta_{max} = \eta_C \frac{\sqrt{1 + ZT} - 1}{\sqrt{1 + ZT} + 1}. \tag{4.47}$$

As we can see in above equations, both $\eta(P_{max})$ and η_{max} are monotonous growing functions of ZT; the only restriction imposed by thermodynamics is $ZT \geq 0$. When $ZT = 0$, both $\eta(P_{max})$and η_{max} vanish, whilst for $ZT \to \infty$, $\eta_{max} \to \eta_C$, and $\eta(P_{max}) \to \eta_{CA}$, where $\eta_{CA} = \eta_C/2$ is the so-called Curzon–Ahlborn efficiency in the linear response [22]. Erdman et al. [22] have studied the above relations of power and efficiency in Eqs. (4.45)–(4.47) in both linear and nonlinear regimes, and plotted their results as in Fig. 4.9.

Fig. 4.9a presents the maximum efficiency η_{max}, obtained by maximising the efficiency η with respect to the applied voltage V, at given ΔT [22]. η_{max}

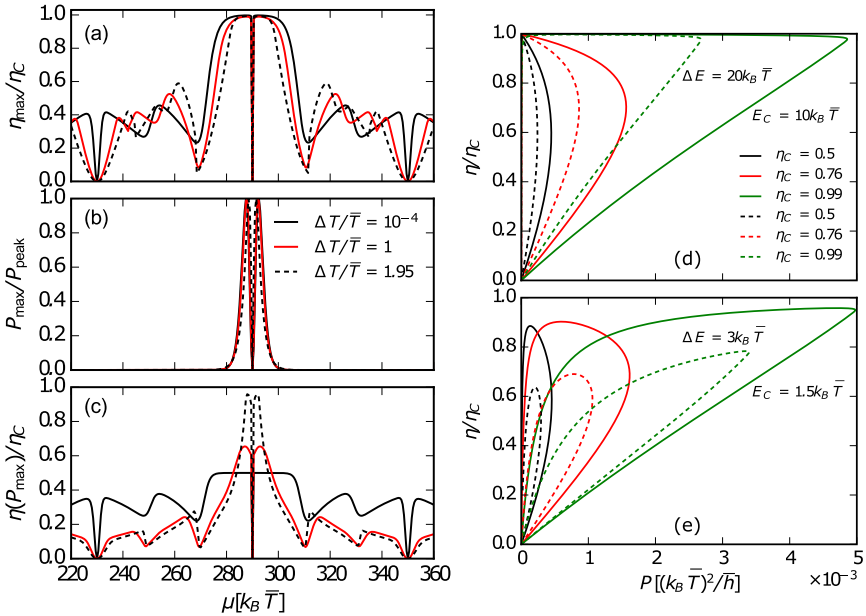


Figure 4.9 (a) Maximum efficiency, normalised to Carnot's efficiency, (b) maximum output power P_{max}, normalised to its peak value P_{peak} and (c) efficiency at maximum power η (P_{max}), normalised to η_C, plotted as a function of μ for various values of $\Delta T/\overline{T}$, for the same parameter values as in Fig. 4.8d and e show the correlation between efficiency and output power calculated by Erdman et al. [22], both for the doubly degenerate interacting case (*solid curves*) and the nondegenerate noninteracting case (*dashed curves*). Curves are obtained by increasing the value of V from zero to the thermovoltage value, corresponding to the open-circuit situation, for various values of η_C. Panel (d) refers to $\Delta E = 20k_B\overline{T}$ (with $E_C = 0$, *dashed curves*) and to $E_C = 10k_B\overline{T}$ (with $E_C = 0$, solid curves), whilst panel (e) refers to $\Delta E = 3k_B\overline{T}$ (with $E_C = 0$, *dashed curves*) and to $E_C = 1.5k_BT$ (with $E_C = 0$, solid curves). *This figure is adopted from Ref. [22].*

is plotted, normalised to η_C, as a function of μ for different values of $\Delta T/\overline{T}$. All plots show pairs of peaks close to μ_N, whose maximum is very close to η_C, and secondary peaks of smaller height. The solid black curve, relative to the linear response regime ($\Delta T/\overline{T} = 10^{-4}$), is related through Eq. (4.47) to the plot of ZT [Fig. 4.6b] discussed in Section 6. For the black curve, a pair of maxima approaching η_C occur at $\mu = \mu_N 2.40 k_B\overline{T}$, whilst η_{max} vanishes at the dominant transition energies (and at the middle points between two dominant transition energies). Moreover, a fine structure of secondary peaks, with spacing ΔE, appears for intermediate values of μ. Moving away from the linear response, the main observation is that an

increase of $\Delta T/\overline{T} > 0.1$ produces only quantitative changes to the curves. As shown in Fig. 4.9a, the main peaks of η_{max} are still approximately located at $\mu = \mu_N \pm 2.40 k_B \overline{T}$ and approaching the Carnot efficiency, whilst the peaks' width reduces slightly with increasing ΔT, and the fine structure of the secondary peaks gets simply distorted [22].

Fig. 4.9b shows another important quantity in heat-to-work conversion, the maximum output power generated P_{max}, which is obtained by maximising the output power with respect to the applied voltage V. It turns out that P_{max} exhibits pairs of peaks approximately located at $\mu = \mu_N \pm 2.40 k_B \overline{T}$ whose height increases approximately quadratically with ΔT, as long as $\Delta T/\overline{T}$ is not too close to 2. Interestingly, Fig. 4.9b shows that the maximum output power P_{max}, when normalised to its peak value P_{peak}, only very weakly depends on the ratio $\Delta T/\overline{T}$. In particular, P_{max}/P_{peak} is well approximated by the linear response result, whose analytical expression is obtained by substituting Eq. (4.31) into Eq. (4.45).

Fig. 4.9c shows the efficiency at maximum power $\eta(P_{max})$, which is calculated by taking, for each value of μ, the value of V that maximises the power. $\eta(P_{max})$ is plotted, as a function of μ, for various values of the ratio $\Delta T/\overline{T}$. By increasing such a ratio starting from the linear response [solid black curve, related to the plot of ZT in Fig. 4.6b through Eq. (4.47)], one finds that the peak values, again occurring approximately at $\mu = \mu_N \pm 2.40 k_B \overline{T}$, increase well above $\eta_C/2$ (the upper limit for the linear response). On the contrary, the efficiency at maximum power for values of μ away from μ_N, relative to the fine structure, decreases with increasing $\Delta T/\overline{T}$ beyond linear response but is only slightly different moving from $\Delta T/\overline{T} = 1$ (red [grey in print version] curve) to $\Delta T/\overline{T} = 1.95$ (black dashed curve).

Fig. 4.9d and e show how the value of the power P and the efficiency η evolve by increasing the applied voltage V from zero (where both P and η vanish) to the thermovoltage (where P vanishes as a consequence of the fact that the charge current vanishes) [22]. In particular, panel (d) refers to the case $\Delta E = 2 k_B \overline{T}$ (dashed curves) and $E_C = 10\ k_B \overline{T}$ (solid curves), whilst panel (e) refers to the case $\Delta E = 3\ k_B \overline{T}$ (dashed curves) and $E_C = 1.5\ k_B \overline{T}$ (solid curves). In the linear response (when $\eta_C \ll 1$), the power reaches its maximum when the efficiency is nearly equal to $\eta_C/2$. By increasing η_C, for all the curves in the figure, both maximum power and efficiency at maximum power increase. For large values of η_C (green [dark grey in print version] curves), the efficiency remains close to the Carnot efficiency when V increased beyond the point of maximum power. The general feature is that

in the interacting case (solid curves) the power is much larger than in the associated noninteracting situation (dashed curves). When ΔE and E_C are of the same order as $k_B\overline{T}$ [Fig. 4.9e], both the maximum power and the efficiency at maximum power are increased in the interacting case as compared with the noninteracting case. A remarkable property of both regimes discussed in panels (d) and (e) is that in the strongly nonlinear regime, the maximum power is obtained for values where the efficiency is high and close to the maximum efficiency [22].

10. Applications

Mesoscopic solid-state physics can help to overcome the limitations of current thermoelectric materials by providing powerful and highly efficient heat engines operating at the nanoscale. Soon after initial studies on the thermopower of basic mesoscopic structures like quantum point contacts [43] and QDs [15,17,44], there were first proposals pointing out that structures of reduced dimension can give rise to an increased thermoelectric figure of merit ZT as compared to bulk structures made from the same materials [25,45]. Similar ideas where brought forward by Mahan and Sofo [24] who showed that sharp spectral features give rise to high thermoelectric performance as characterised by a high value of ZT. Nanoscale conductors such as QDs naturally provide these sharp spectral features. Hence, they are promising candidates for thermoelectric heat engines [39,46–52]. Recently, Josefsson et al. experimentally demonstrated a particle exchange heat engine based on a QD embedded into a semiconductor nanowire, with a device as shown in Fig. 4.10a [53]. They directly measure the engine's steady-state electric power output and combine it with the calculated electronic heat flow to determine the electronic efficiency η. At the maximum power conditions, the $\eta$of their device is in agreement with the Curzon–Ahlborn efficiency and that the overall maximum η is in excess of 70% of the Carnot efficiency whilst maintaining a finite power output. Their results demonstrate that thermoelectric power conversion can, in principle, be achieved close to the thermodynamic limits [53]. More experimental and analytical studies were carried out on this heat engine device (in Fig. 4.10a), to optimise its power and efficiency and to investigate nonlinear and second-order effects [54].

Among the different nanoscale devices, QDs with discrete energy levels have stood out because of their use as tunable energy filters [55]. Based on that effect, two-terminal QD-based refrigerators [56,57], heat pumps [58–61], heat diodes [62], heat valves [63,64] and heat converters [44,65] have been proposed which are predicted to attain a high efficiency. Prance

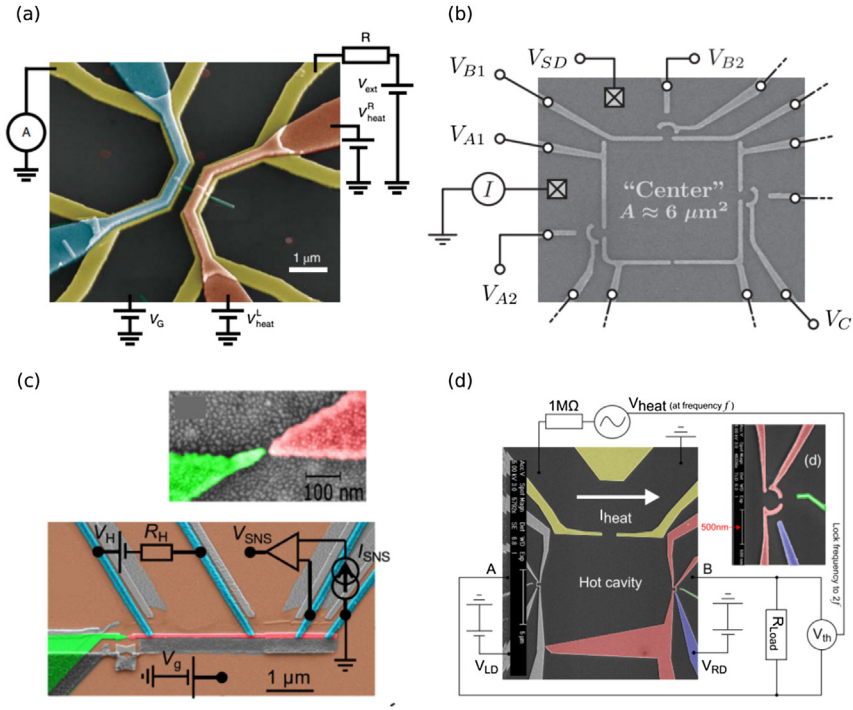


Figure 4.10 (a) False-coloured SEM image of a nominally identical device to the one used in the quantum dot heat engine experiment by Josefsson et al. [53]. It consists of an InAs/InP nanowire quantum dot (green [light grey in print version], the narrow wire at the centre) in contact with metallic leads (yellow [grey in print version], the gates with Y-shaped ends at the bottom layer). Heaters (blue, the left U-shaped gate at the top layer, and red , the right U-shaped gate at the top layer) run over the contact leads and are insulated from them by a layer of high-k oxide. One of the heaters (red) is used in the experiment for thermal biasing, and the other (blue) is unused. (b) SEM image of a typical quantum dot refrigerator device, experimentally demonstrated by Prance et al. [57]. NiCr/Au gates were patterned on the surface of a GaAs/AlGaAs heterostructure using electron beam lithography. A 2DEG 90 nm below the surface, with a carrier density of $n = 1.37 \times 10^{11}$ cm^2, is contacted by annealed AuGeNi contacts. The surface gates can define an enclosed, central 2DEG region with quantum dots to the left (dot A) and above (dot B). The right dot is unused in this experiment. (c) False-coloured SEM image of a typical device reported by Dutta et al. [63]. The source is coloured in red (the top horizontal line at the centre of the bottom figure, and the right triangle at the top figure), the drain in green (the top triangle-shaped area at the left bottom corner of the bottom figure, and the left triangle at the top figure), and the superconducting leads in blue (dark grey in print version). The circuit diagram shows the heat transport setup. The longer (2.5 μm) Superconductor-Normal Metal-Superconductor (SNS) junction is used as a heater driven by a constant DC battery and the shorter (700 nm) SNS junction is used as a thermometer. (d) False colour SEM image of the quantum dot energy harvester device, demonstrated by Jaliel et al. [66]. Ti/Au gates were patterned on the surface of GaAs/AlGaAs heterostructure using

et al. experimentally demonstrated the feasibility of electronic cooling of a 2DEG using energy selective tunnelling through QDs [57]. At an ambient electron temperature of 280 mK, their QD refrigerator device as shown in Fig. 4.10b, consisted of two QDs besides a central cavity, can cool the isolated 2DEG of 6 μm below 190 mK. It has the potential to be used as a general platform for extra-low-temperature measurements, with devices operating within it probed by noninvasive charge sensors [57]. Dutta et al. demonstrated gate control of electronic heat flow in a thermally biased single-QD junction, as shown in Fig. 4.10c. The electron temperature maps taken in the immediate vicinity of the junction, as a function of the gate and bias voltages applied to their device, reveal clearly defined Coulomb diamond patterns that indicate a maximum heat transfer at the charge degeneracy point. The nontrivial bias and gate dependence of this heat valve results from the quantum nature of the dot at the heart of device and its strong coupling to leads [63].

Another popular use of QDs' discrete electronic states as energy filters is energy harvesting. For such a purpose, quantum devices require that the energy source is separated from the electrical circuit, so no charge is extracted from it [50]. This can be accomplished in three-terminal devices where a hot terminal injects heat but no charge into the setup, thus driving a charge current between two cold reservoirs. There have been a number of proposals for these kinds of energy harvesters [42,67–85]. Three-terminal heat engines based on Coulomb-coupled QDs [68,69] have been realised experimentally recently [86–88]. Due to their design, they are however limited to low power. A three-terminal energy harvester based on two resonant-tunnelling QDs with different energy levels overcomes this issue. It can in principle reach Carnot efficiency and can be optimised to achieve

electron-beam lithography. The 2DEG was 110 nm below the surface, with a carrier concentration of $n = 1.35 \times 10^{11}$ cm^2, and was contacted by annealed AuGeNi contacts. The surface gates define a cavity of 90 μm^2 area at the central 2DEG region with two quantum dots, respectively, on the left and the right sides, and a heating channel on the top. As shown in the top right inset (d), the quantum dots, of 310 nm diameters, are constructed of three barrier gates (red, the three gates forming a rough circle at the left side of the inset (d)), one detector gate (green, the gate at the centre right with an angle on the inset (d)), and one plunger gate (blue, the gate at the bottom centre with a slight slope on the inset (d)). *In this figure, (a) is adapted from Ref. [53], (b) [57], (c) [63] and (d) [66]. Please refer to these papers for false-coloured images of these device.*

a large power in combination with a high efficiency at maximum power [75]. This has recently been experimentally demonstrated by Jaliel et al. [66]. As shown in Fig. 4.10d, their device is consisted of two high-electron mobility transistors placed on either side of a hot-electron reservoir, and a heating channel on the top. The discrete energy levels of the QDs are tuned to be aligned with low energy electron on one side and high energy electron on the other side of the hot reservoir. The QDs thus act as energy filters and allow for the conversion of heat from the cavity into electrical power. The cavity temperature can be varied by applying different heating current on the top channel. At an estimated base temperature of 75 mK, this energy harvester can generate a thermal power of 0.13 fW for a temperature difference across each dot of about 67 mK [66].

One important application of QDs is low temperature thermometers. Either for the studies of QD thermoelectricity or its applications, the temperature of electron systems plays a major role. Several low temperature thermometers have been realised using modern microlithography and nanolithography [89], such as a 'shot noise thermometer' using a tunnelling junction [90], and a 'Coulomb blockade thermometer' utilising arrays of tunnelling junctions [91]. In case of double barriers – a QD – the electron temperature can be extracted by measuring the current or conductance through the dot when transport through it is Coulomb blockaded, and fitting to the resulting temperature-dependent lineshapes [15,92]. It is a useful tool to determine the electron temperature for operating devices [93–95]. Mavalankar et al. demonstrated a noninvasive scheme for determining the temperature of the 2DEG in a GaAs/AlGaAs heterostructure, which can be run simultaneously with and independently of another operating device [96]. It consists of monitoring the charge occupation of a QD weakly tunnel-coupled to the 2DEG, using a quantum point contact (a device similar to the one in Fig. 4.10b but with a top gate above the central cavity). Electron temperatures between 120 and 951 mK were measured by this QD thermometer, and agree with what the resistance thermometer of their experimental environment have detected. It can be used to measure lower temperatures than those measurable using Coulomb blockaded transport, and is robust in the presence of a magnetic field [96].

11. Summary

In this review, we discussed in detail the thermoelectric properties of a QD. It started with the Landauer–Büttiker formalism that relates the transport properties of a conductor to the transmission probabilities between

reservoirs in local equilibrium. Then we looked into the QD model and investigated the thermoelectric coefficients in the Landauer–Büttiker formalism in the quantum limit. Both the electrical conductance and the thermal conductance exhibit periodic Coulomb blockade oscillations separated by intervals: $\Delta E_F = \Delta E + e^2/C$. The thermopower shows long-period oscillations, due to changes in the number of electrons on the dot in the ground state, and short-period oscillations when the charging energy is less than the level spacing. The spin degeneracy of the energy levels of QDs introduces doublets with mirror symmetry in oscillations of the electrical and the thermal conductance, and doubles the frequency of short-period oscillations in thermopower, without affecting that of its long-period oscillations. With the knowledge of the thermoelectric coefficients, we then discussed thermoelectric characteristics – power factor and figure of merit in the quantum limit. It demonstrates QDs are a promising candidate for thermoelectric materials and suggests how to maximise the power and efficiency of the system. By looking into the temperature and gate-voltage dependence of the Lorenz ratio, we investigated the Wiedemann–Franz law in a QD system. It is only fulfilled for a vanishing charging energy gap, $\Delta = 0$, with corrections quadratic in Δ/k_BT away from resonance in such mesoscopic systems. Then we switched to the nonlinear regime of QDs and discussed how it affects the thermoelectric coefficients, along with the quantities such as the maximal power, the efficiency at maximum power and the maximal efficiency at a given output power of a QD as a heat engine in order to sufficiently characterise its thermoelectric performance. In the strongly nonlinear regime, the maximum power of a QD heat engine is obtained for values where its efficiency is high and close to the maximum efficiency. In the end, we discussed the various theoretical proposals on how to utilise these thermoelectric characteristics of QDs, for instance, as a refrigerator, an energy harvester and a thermometer, and their recent experimental demonstrations.

The thermoelectric properties of QDs remain a hit research area. Many theoretical and experimental efforts have also been performed to investigate thermoelectric properties of other nanostructures such as superlattice [97,98], graphene nanostructures [99] and silicon nanowires with various sizes and doping levels [100,101]. With the depleting fossil fuel resources and the global warming, the potential of these nanostructures to be an alternative energy resource, specifically to be used to convert waste heat into electricity at a mesoscopic scale, continues to attract enormous research interests. However, the field still faces a number of open questions. For example, the influence of phonons has to be better understood as they not only degrade the efficiency but also make it hard to maintain a given

temperature bias across the device. In particular, it would be desirable to investigate systems where an analytic treatment of their influence can be made and to invent setups where they can be controlled in a systematic way [50]. Another interesting question addresses the thermoelectric performance of multi-terminal heat engines with broken time-reversal symmetry. Theoretically, these devices allow to get arbitrarily close to Carnot efficiency at finite output power. But how close can any physical realisations actually get? [50] Finally, the field will greatly benefit from more experimental demonstrations of theoretical ideas that have been put forward, about the applications of thermoelectric properties of QDs, and from possibly industrially applying them.

References

[1] Landauer R. Spatial variation of currents and fields due to localized scatterers in metallic conduction. IBM J Res Dev 1957;1(3):223–31.
[2] Büttiker M. Four-terminal phase-coherent conductance. Phys Rev Lett 1986;57(14):1761.
[3] Van Houten H, Molenkamp LW, Beenakker CWJ, Foxon CT. Thermoelectric properties of quantum point contacts. Semicond Sci Technol 1992;7(3B):B215.
[4] Sivan U, Imry Y. Multichannel landauer formula for thermoelectric transport with application to thermopower near the mobility edge. Phys Rev B 1986;33(1):551.
[5] Butcher PN. Thermal and electrical transport formalism for electronic microstructures with many terminals. J Phys Condens Matter 1990;2(22):4869.
[6] Stone AD. Magnetoresistance fluctuations in mesoscopic wires and rings. Phys Rev Lett 1985;54(25):2692.
[7] Washburn S. Fluctuations in the extrinsic conductivity of disordered metal. IBM J Res Dev 1988;32(3):335–46.
[8] Washburn S, Webb RA. Aharonov-bohm effect in normal metal quantum coherence and transport. Adv Phys 1986;35(4):375–422.
[9] Van Wees BJ, Van Houten H, Beenakker CWJ, Williamson JG, Kouwenhoven LP, Van der Marel D, Foxon CT. Quantized conductance of point contacts in a two-dimensional electron gas. Phys Rev Lett 1988;60(9):848.
[10] Szafer A, Stone AD. Theory of quantum conduction through a constriction. Phys Rev Lett 1989;62(3):300.
[11] Büttiker M. Absence of backscattering in the quantum hall effect in multiprobe conductors. Phys Rev B 1988;38(14):9375.
[12] Beenakker CWJ, Van Houten H. Quenching of the hall effect. Phys Rev Lett 1988;60(23):2406.
[13] Büttiker M. Transmission probabilities and the quantum hall effect. Phys Rev Lett 1989;62(2):229.
[14] Zianni X. Coulomb oscillations in the electron thermal conductance of a dot in the linear regime. Phys Rev B 2007;75(4):045344.
[15] Beenakker CWJ, Staring AAM. Theory of the thermopower of a quantum dot. Phys Rev B 1992;46(15):9667–76. https://doi.org/10.1103/PhysRevB.46.9667.
[16] Beenakker CWJ. Theory of coulomb-blockade oscillations in the conductance of a quantum dot. Phys Rev B 1991;44(4):1646.
[17] M Staring AA, W Molenkamp L, Alphenaar BW, van Houten H, A Buyk OJ, Mabesoone MAA, J Beenakker CW, T Foxon C. Coulomb-blockade oscillations in the thermopower of a quantum dot. Europhys Lett 1993;22(1):57–62. https://

doi.org/10.1209/0295-5075/22/1/011. ISSN 0295-5075, 1286-4854, http://iopscience.iop.org/0295-5075/22/1/011.
[18] Appleyard NJ, Nicholls JT, Simmons MY, Tribe WR, Pepper M. Thermometer for the 2d electron gas using 1d thermopower. Phys Rev Lett 1998;81(16):3491.
[19] Dzurak AS, Smith CG, Barnes CHW, Pepper M, Martin-Moreno L, Liang CT, Ritchie DA, Jones GAC. Thermoelectric signature of the excitation spectrum of a quantum dot. Phys Rev B 1997;55(16):R10197.
[20] Godijn SF, Möller S, Buhmann H, Molenkamp LW, Van Langen SA. Thermopower of a chaotic quantum dot. Phys Rev Lett 1999;82(14):2927.
[21] Small JP, Perez KM, Kim P. Modulation of thermoelectric power of individual carbon nanotubes. Phys Rev Lett 2003;91(25):256801. https://doi.org/10.1103/PhysRevLett.91.256801. http://link.aps.org/doi/10.1103/PhysRevLett.91.256801.
[22] Andrea Erdman P, Mazza F, Bosisio R, Benenti G, Fazio R, Taddei F. Thermoelectric properties of an interacting quantum dot based heat engine. Phys Rev B 2017;95(24):245432.
[23] Tsaousidou M, Triberis GP. Thermoelectric properties of a weakly coupled quantum dot: enhanced thermoelectric efficiency. J Phys Condens Matter 2010;22(35):355304.
[24] Mahan GD, Sofo JO. The best thermoelectric. Proc Natl Acad Sci USA 1996;93(15):7436–9. ISSN 0027-8424, 1091–6490, http://www.pnas.org/content/93/15/7436.
[25] Hicks LD, Dresselhaus MS. Effect of quantum-well structures on the thermoelectric figure of merit. Phys Rev B 1993;47(19):12727–31. https://doi.org/10.1103/PhysRevB.47.12727. http://link.aps.org/doi/10.1103/PhysRevB.47.12727.
[26] Kubala B, König J, Pekola J. Violation of the wiedemann-franz law in a single-electron transistor. Phys Rev Lett 2008;100(6):066801.
[27] Krawiec M, Wysokiński KI. Thermoelectric effects in strongly interacting quantum dot coupled to ferromagnetic leads. Phys Rev B 2006;73(7):075307.
[28] Vavilov MG, Stone AD. Failure of the wiedemann-franz law in mesoscopic conductors. Phys Rev B 2005;72(20):205107.
[29] Dutta B, Joonas TP, Antonenko DS, Meschke M, Skvortsov MA, Kubala B, König J, Winkelmann CB, Courtois H, P Pekola J. Thermal conductance of a single-electron transistor. Phys Rev Lett 2017;119(7):077701.
[30] Whitney RS. Nonlinear thermoelectricity in point contacts at pinch off: a catastrophe aids cooling. Phys Rev B 2013;88(6):064302.
[31] Sánchez D, López R. Scattering theory of nonlinear thermoelectric transport. Phys Rev Lett 2013;110(2):026804.
[32] Meair J, Jacquod P. Scattering theory of nonlinear thermoelectricity in quantum coherent conductors. J Phys Condens Matter 2013;25(8):082201.
[33] Whitney RS. Thermodynamic and quantum bounds on nonlinear dc thermoelectric transport. Phys Rev B 2013;87(11):115404.
[34] Matthews J, Sánchez D, Larsson M, Linke H. Thermally driven ballistic rectifier. Phys Rev B 2012;85(20):205309.
[35] Hwang S-Y, Sánchez D, Lee M, López R. Magnetic-field asymmetry of nonlinear thermoelectric and heat transport. New J Phys 2013;15(10):105012.
[36] Leijnse M, Wegewijs MR, Flensberg K. Nonlinear thermoelectric properties of molecular junctions with vibrational coupling. Phys Rev B 2010;82(4):045412.
[37] Hershfield S, Muttalib KA, Nartowt BJ. Nonlinear thermoelectric transport: a class of nanodevices for high efficiency and large power output. Phys Rev B 2013;88(8):085426.
[38] Van den Broeck C. Thermodynamic efficiency at maximum power. Phys Rev Lett 2005;95(19):190602.

[39] Esposito M, Lindenberg K, Van den Broeck C. Thermoelectric efficiency at maximum power in a quantum dot. Europhys Lett 2009;85(6):60010. https://doi.org/10.1209/0295-5075/85/60010. ISSN 0295-5075, 1286-4854, http://iopscience.iop.org/0295-5075/85/6/60010.
[40] Kay B, Saito K, Seifert U. Strong bounds on onsager coefficients and efficiency for three-terminal thermoelectric transport in a magnetic field. Phys Rev Lett 2013; 110(7):070603.
[41] Whitney RS. Most efficient quantum thermoelectric at finite power output. Phys Rev Lett 2014;112(13):130601.
[42] Mazza F, Bosisio R, Benenti G, Giovannetti V, Fazio R, Taddei F. Thermoelectric efficiency of three-terminal quantum thermal machines. New J Phys 2014;16(8): 085001. https://doi.org/10.1088/1367-2630/16/8/085001. ISSN 1367-2630, http://iopscience.iop.org/1367-2630/16/8/085001.
[43] Molenkamp LW, Van Houten H, Beenakker CWJ, Eppenga R, Foxon CT. Quantum oscillations in the transverse voltage of a channel in the nonlinear transport regime. Phys Rev Lett 1990;65(8):1052.
[44] Humphrey TE, Newbury R, Taylor RP, Linke H. Reversible quantum brownian heat engines for electrons. Phys Rev Lett 2002;89(11):116801. https://doi.org/10.1103/PhysRevLett.89.116801. http://link.aps.org/doi/10.1103/PhysRevLett.89.116801.
[45] Hicks LD, Dresselhaus MS. Thermoelectric figure of merit of a one-dimensional conductor. Phys Rev B 1993;47(24):16631–4. https://doi.org/10.1103/PhysRevB.47.16631. http://link.aps.org/doi/10.1103/PhysRevB.47.16631.
[46] Humphrey TE, Linke H. Reversible thermoelectric nanomaterials. Phys Rev Lett 2005;94:096601. https://doi.org/10.1103/PhysRevLett.94.096601. https://link.aps.org/doi/10.1103/PhysRevLett.94.096601.
[47] Murphy P, Mukerjee S, Moore J. Optimal thermoelectric figure of merit of a molecular junction. Phys Rev B 2008;78:161406. https://doi.org/10.1103/PhysRevB.78.161406. https://link.aps.org/doi/10.1103/PhysRevB.78.161406.
[48] Nakpathomkun N, Xu HQ, Linke H. Thermoelectric efficiency at maximum power in low-dimensional systems. Phys Rev B 2010;82(23):235428. https://doi.org/10.1103/PhysRevB.82.235428. http://link.aps.org/doi/10.1103/PhysRevB.82.235428.
[49] Kennes DM, Schuricht D, Meden V. Efficiency and power of a thermoelectric quantum dot device. EPL 2013;102(5):57003. https://doi.org/10.1209/0295-5075/102/57003.
[50] Sothmann B, Sánchez R, Jordan AN. Thermoelectric energy harvesting with quantum dots. Nanotechnology 2015;26(3):032001. https://doi.org/10.1088/0957-4484/26/3/032001. ISSN 0957-4484, http://iopscience.iop.org/0957-4484/26/3/032001.
[51] Marchegiani G, Braggio A, Giazotto F. Superconducting nonlinear thermoelectric heat engine. Phys Rev B 2020;101:214509. https://doi.org/10.1103/PhysRevB.101.214509. https://link.aps.org/doi/10.1103/PhysRevB.101.214509.
[52] David Mayrhofer R, Elouard C, Splettstoesser J, Jordan AN. Stochastic thermodynamic cycles of a mesoscopic thermoelectric engine. Phys Rev B 2021;103: 075404. https://doi.org/10.1103/PhysRevB.103.075404. https://link.aps.org/doi/10.1103/PhysRevB.103.075404.
[53] Josefsson M, Svilans A, Burke AM, Hoffmann EA. Sofia Fahlvik, Claes Thelander, Martin Leijnse, and Heiner Linke. A quantum-dot heat engine operating close to the thermodynamic efficiency limits. Nat Nanotechnol 2018;13(10):920–4. https://doi.org/10.1038/s41565-018-0200-5. ISSN 1748-3395.
[54] Josefsson M, Svilans A, Linke H, Martin L. Optimal power and efficiency of single quantum dot heat engines: theory and experiment. Phys Rev B 2019;99:235432.

https://doi.org/10.1103/PhysRevB.99.235432. https://link.aps.org/doi/10.1103/PhysRevB.99.235432.
[55] Haupt F, Martin L, Calvo HL, Classen L, Splettstoesser J, Wegewijs MR. Heat, molecular vibrations, and adiabatic driving in non-equilibrium transport through interacting quantum dots. Physica Status Solidi (B) 2013;250(11):2315–29.
[56] Edwards HL, Niu Q, Georgakis GA, de Lozanne AL. Cryogenic cooling using tunneling structures with sharp energy features. Phys Rev B 1995;52(8):5714–36. https://doi.org/10.1103/PhysRevB.52.5714. http://link.aps.org/doi/10.1103/PhysRevB.52.5714.
[57] Prance JR, Smith CG, Griffiths JP, Chorley SJ, Anderson D, Jones GAC, Farrer I, Ritchie DA. Electronic refrigeration of a two-dimensional electron gas. Phys Rev Lett 2009;102(14):146602. https://doi.org/10.1103/PhysRevLett.102.146602. http://link.aps.org/doi/10.1103/PhysRevLett.102.146602.
[58] Moskalets M, Büttiker M. Dissipation and noise in adiabatic quantum pumps. Phys Rev B 2002;66(3):035306.
[59] Arrachea L, Moskalets M, Martin-Moreno L. Heat production and energy balance in nanoscale engines driven by time-dependent fields. Phys Rev B 2007;75(24):245420.
[60] Rey M, Michael Strass, Kohler S, Peter H, Sols F. Nonadiabatic electron heat pump. Phys Rev B 2007;76(8):085337.
[61] Juergens S, Haupt F, Moskalets M, Splettstoesser J. Thermoelectric performance of a driven double quantum dot. Phys Rev B 2013;87(24):245423.
[62] Ruokola T, Ojanen T. Single-electron heat diode: asymmetric heat transport between electronic reservoirs through coulomb islands. Phys Rev B 2011;83(24):241404.
[63] Dutta B, Majidi D, Talarico NW, Lo Gullo N, Courtois H, Winkelmann CB. Single-quantum-dot heat valve. Phys Rev Lett 2020;125:237701. https://doi.org/10.1103/PhysRevLett.125.237701. https://link.aps.org/doi/10.1103/PhysRevLett.125.237701.
[64] Jussiau É, Manikandan SK, Bhandari B, Jordan AN. Thermal control across a chain of electronic nanocavities. 2021.
[65] Muralidharan B, Grifoni M. Performance analysis of an interacting quantum dot thermoelectric setup. Phys Rev B 2012;85(15):155423.
[66] Jaliel G, RK Puddy, Sánchez R, Jordan AN, Sothmann B, Farrer I, Griffiths JP, Ritchie DA, Smith CG. Experimental realization of a quantum dot energy harvester. Phys Rev Lett 2019;123(11):117701.
[67] Entin-Wohlman O, Imry Y, Aharony A. Three-terminal thermoelectric transport through a molecular junction. Phys Rev B 2010;82(11):115314. https://doi.org/10.1103/PhysRevB.82.115314. http://link.aps.org/doi/10.1103/PhysRevB.82.115314.
[68] Sánchez R, Büttiker M. Optimal energy quanta to current conversion. Phys Rev B 2011;83(8):085428. https://doi.org/10.1103/PhysRevB.83.085428. http://link.aps.org/doi/10.1103/PhysRevB.83.085428.
[69] Sothmann B, Sánchez R, Jordan AN, Büttiker M. Rectification of thermal fluctuations in a chaotic cavity heat engine. Phys Rev B 2012;85(20):205301. https://doi.org/10.1103/PhysRevB.85.205301. http://link.aps.org/doi/10.1103/PhysRevB.85.205301.
[70] Sothmann B, Büttiker M. Magnon-driven quantum-dot heat engine. Europhys Lett 2012;99(2):27001. https://doi.org/10.1209/0295-5075/99/27001. ISSN 0295-5075,1286-4854, http://iopscience.iop.org/0295-5075/99/2/27001.
[71] Ruokola T, Ojanen T. Theory of single-electron heat engines coupled to electromagnetic environments. Phys Rev B 2012;86(3):035454. https://doi.org/10.1103/PhysRevB.86.035454. http://link.aps.org/doi/10.1103/PhysRevB.86.035454.
[72] Jiang J-H, Entin-Wohlman O, Imry Y. Thermoelectric three-terminal hopping transport through one-dimensional nanosystems. Phys Rev B 2012;85(7):075412.

https://doi.org/10.1103/PhysRevB.85.075412. http://link.aps.org/doi/10.1103/PhysRevB.85.075412.

[73] Jiang J-H, Entin-Wohlman O, Imry Y. Three-terminal semiconductor junction thermoelectric devices: improving performance. New J Phys 2013;15(7):075021. https://doi.org/10.1088/1367-2630/15/7/075021. ISSN 1367-2630, http://iopscience.iop.org/1367-2630/15/7/075021.

[74] Machon P, Eschrig M, Belzig W. Nonlocal thermoelectric effects and nonlocal onsager relations in a three-terminal proximity-coupled superconductor-ferromagnet device. Phys Rev Lett 2013;110(4):047002. https://doi.org/10.1103/PhysRevLett.110.047002. http://link.aps.org/doi/10.1103/PhysRevLett.110.047002.

[75] Jordan AN, Sothmann B, Sánchez R, Büttiker M. Powerful and efficient energy harvester with resonant-tunneling quantum dots. Phys Rev B 2013;87(7):075312. https://doi.org/10.1103/PhysRevB.87.075312. http://link.aps.org/doi/10.1103/PhysRevB.87.075312.

[76] Sothmann B, Sánchez R, Jordan AN, Büttiker M. Powerful energy harvester based on resonant-tunneling quantum wells. New J Phys 2013;15(9):095021. https://doi.org/10.1088/1367-2630/15/9/095021. ISSN 1367-2630, http://iopscience.iop.org/1367-2630/15/9/095021.

[77] Bergenfeldt C, Peter Samuelsson, Sothmann B, Flindt C, Büttiker M. Hybrid microwave-cavity heat engine. Phys Rev Lett 2014;112(7):076803. https://doi.org/10.1103/PhysRevLett.112.076803. http://link.aps.org/doi/10.1103/PhysRevLett.112.076803.

[78] Donsa S, Andergassen S, Held K. Double quantum dot as a minimal thermoelectric generator. Phys Rev B 2014;89(12):125103. https://doi.org/10.1103/PhysRevB.89.125103. http://link.aps.org/doi/10.1103/PhysRevB.89.125103.

[79] Mazza F, Valentini S, Bosisio R, Benenti G, Giovannetti V, Fazio R, Taddei F. Separation of heat and charge currents for boosted thermoelectric conversion. Phys Rev B 2015;91(24):245435. https://doi.org/10.1103/PhysRevB.91.245435. http://link.aps.org/doi/10.1103/PhysRevB.91.245435.

[80] Sánchez R, Sothmann B, Jordan AN. Chiral thermo- electrics with quantum Hall edge states. Phys Rev Lett 2015;114(14):146801. https://doi.org/10.1103/PhysRevLett.114.146801. http://link.aps.org/doi/10.1103/PhysRevLett.114.146801.

[81] Sánchez R, Sothmann B, Jordan AN. Heat diode and engine based on quantum Hall edge states. New J Phys 2015;17(7):075006. https://doi.org/10.1088/1367-2630/17/7/075006. ISSN 1367-2630, http://iopscience.iop.org/1367-2630/17/7/075006.

[82] Hofer PP, Sothmann B. Quantum heat engines based on electronic Mach-Zehnder interferometers. Phys Rev B 2015;91(19):195406. https://doi.org/10.1103/PhysRevB.91.195406. http://link.aps.org/doi/10.1103/PhysRevB.91.195406.

[83] Bosisio R, Fleury G, Jean-Louis Pichard, Gorini C. Nanowire-based thermoelectric ratchet in the hopping regime. Phys Rev B 2016;93(16):165404. https://doi.org/10.1103/PhysRevB.93.165404. http://link.aps.org/doi/10.1103/PhysRevB.93.165404.

[84] Szukiewicz B, Ulrich E, Karol I, Wysokiński. Optimisation of a three-terminal nonlinear heat nano-engine. New J Phys 2016;18(2):023050. https://doi.org/10.1088/1367-2630/18/2/023050. ISSN 1367-2630, http://stacks.iop.org/1367-2630/18/i=2/a=023050.

[85] Jiang J-H, Imry Y. Near-field three-terminal thermoelectric heat engine. Phys Rev B 2018;97(12):125422. https://doi.org/10.1103/PhysRevB.97.125422. https://link.aps.org/doi/10.1103/PhysRevB.97.125422.

[86] Thierschmann H, Sánchez R, Sothmann B, Arnold F, Heyn C, Hansen W, Buhmann H, Laurens W. Molenkamp. Three-terminal energy harvester with coupled quantum dots. Nat Nanotechnol 2015;10(10):854–8. https://doi.org/

10.1038/nnano.2015.176. ISSN 1748-3387, http://www.nature.com/nnano/journal/v10/n10/full/nnano.2015.176.html.

[87] Roche B, Roulleau P, Jullien T, Jompol Y, Farrer I, Ritchie DA, Glattli DC. Harvesting dissipated energy with a mesoscopic ratchet. Nat Commun 2015;6. https://doi.org/10.1038/ncomms7738. http://www.nature.com/ncomms/2015/150401/ncomms7738/full/ncomms7738.html.

[88] Hartmann F, Pfeffer P, Höfling S, Kamp M, Worschech L. Voltage fluctuation to current converter with coulomb-coupled quantum dots. Phys Rev Lett 2015;114(14):146805. https://doi.org/10.1103/PhysRevLett.114.146805. http://link.aps.org/doi/10.1103/PhysRevLett.114.146805.

[89] Giazotto F, Heikkilä TT, Luukanen A, Savin AM, Pekola JP. Opportunities for mesoscopics in thermometry and refrigeration: physics and applications. Rev Mod Phys 2006;78:217–74. https://doi.org/10.1103/RevModPhys.78.217. https://link.aps.org/doi/10.1103/RevModPhys.78.217.

[90] Spietz L, Lehnert KW, Siddiqi I, Schoelkopf RJ. Primary electronic thermometry using the shot noise of a tunnel junction. Science 2003;300(5627):1929–32. https://doi.org/10.1126/science.1084647. ISSN 0036-8075, https://science.sciencemag.org/content/300/5627/1929.

[91] Pekola JP, Hirvi KP, Kauppinen JP, Paalanen MA. Thermometry by arrays of tunnel junctions. Phys Rev Lett 1994;73:2903–6. https://doi.org/10.1103/PhysRevLett.73.2903. https://link.aps.org/doi/10.1103/PhysRevLett.73.2903.

[92] Foxman EB, Meirav U, McEuen PL, Kastner MA, Klein O, Belk PA, Abusch DM, Wind SJ. Crossover from single-level to multilevel transport in artificial atoms. Phys Rev B 1994;50(19):14193.

[93] Rossi A, Ferrus T, Williams DA. Electron temperature in electrically isolated si double quantum dots. Appl Phys Lett 2012;100(13):133503. https://doi.org/10.1063/1.3697832.

[94] Gasparinetti S, Deon F, Biasiol G, Sorba L, Beltram F, Giazotto F. Probing the local temperature of a two-dimensional electron gas microdomain with a quantum dot: measurement of electron-phonon interaction. Phys Rev B 2011;83:201306. https://doi.org/10.1103/PhysRevB.83.201306. https://link.aps.org/doi/10.1103/PhysRevB.83.201306.

[95] Venkatachalam V, Hart S, Pfeiffer L, West K, Yacoby A. Local thermometry of neutral modes on the quantum hall edge. Nat Phys 2012;8(9):676–81. https://doi.org/10.1038/nphys2384. ISSN 1745-2481.

[96] Mavalankar A, Chorley SJ, Griffiths J, Jones GAC, Farrer I, Ritchie DA, Smith CG. A non-invasive electron thermometer based on charge sensing of a quantum dot. Appl Phys Lett 2013;103(13):133116.

[97] Colpitts T, Venkatasubramanian R, Siivola E, Brooks O'Q. Thin-film thermoelectric devices with high room-temperature figures of merit. Nature 2001;413.

[98] Zide JMO, Vashaee D, Bian ZX, Zeng G, Bowers JE, Shakouri A, Gossard AC. Demonstration of electron filtering to increase the seebeck coefficient in $in_{0.53}ga_{0.47}Asin_{0.53}ga_{0.28}al_{0.19}As$ superlattices. Phys Rev B 2006;74:205335. https://doi.org/10.1103/PhysRevB.74.205335. https://link.aps.org/doi/10.1103/PhysRevB.74.205335.

[99] Dollfus P, Nguyen VH, Jérôme Saint-Martin. Thermoelectric effects in graphene nanostructures. J Phys Condens Matter 2015;27(13):133204. https://doi.org/10.1088/0953-8984/27/13/133204.

[100] Tahir-Kheli J, Yu J-K, Goddard III WA, Boukai AI, Bunimovich Y, Heath JR. Silicon nanowires as efficient thermoelectric materials. Nature 2008;451.

[101] Narducci D. A special issue on silicon and silicon-related materials for thermoelectricity. Eur Phys J B 2015;88.

CHAPTER FIVE

Single-electron sources

Masaya Kataoka*
National Physical Laboratory, Teddington, United Kingdom
*Corresponding author: masaya.kataoka@npl.co.uk

In the last few decades, significant effort has been made to develop solid-state devices at the nanoscale aiming to control a small number of electrons. This has led to a class of device technologies that transport electrons one by one, so-called 'single-electron sources'. These types of 'dynamic' device have slightly different goals to conventional 'static' single-electron devices (their main function is to keep a certain number of electrons in a confined space). One potential application of single-electron sources is the primary metrological standard for the ampere, the unit for electrical current in the international system of units (SI). In 2019, the SI ampere (symbol 'A') was re-defined based on the fixed value of elementary charge, $e = 1.602\ 176\ 634 \times 10^{-19}$C, where C = A·s [1]. Accurately transferring electrons one by one would produce a quantised current following the equation.

$$I = ef, \tag{5.1}$$

where f is the frequency of electron transfer. Following the re-definition, this method can be used for a realisation of the ampere, although the required electron-transfer accuracy for metrological applications ($\sim$1 part in 10^8) has not yet been achieved. Another application of single-electron sources lies in 'electron quantum optics' experiments. Electron quantum optics uses electrons instead of photons or atoms for experiments analogous to those in the field of photonic or atomic quantum optics. Initial electron quantum optics experiments were conducted using continuous sources such as quantum point contacts or quantum Hall edge states, but these have been in recent years replaced by single-electron sources, which enables clock-controlled generation of electron wave packets. In this chapter, we will first look at various types of single-electron source. We will then discuss the development and current status of metrological applications. We conclude this chapter by introducing a few examples of electron-quantum optics experiments.

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00012-5
© 2021 Elsevier Ltd.
All rights reserved.

1. Types of single-electron source

In this section, we introduce seven different types of single-electron source. There are two main types of operating mode that these single-electron sources can perform. The first is to transfer one electron (or some other fixed number of electrons) at a time between source and drain contacts (or Fermi seas). The second is to emit one electron at a time (or in some cases, an electron followed by a hole) into an electron waveguide (quasi-one-dimensional channel, such as a quantum Hall edge state). In this second type of operating mode, it is important that single-electron excitation is kept coherent for a certain amount of time whilst it travels through the waveguide so that operations can be performed on the electron. Some single-electron sources introduced in the following can operate in both modes, whilst the others are specialised for one type of mode. Although the operating temperature may not be specified individually in the following parts, all single-electron sources introduced in this chapter require cryogenic temperatures ranging from $\sim$20 K down to $\sim$10 mK.

1.1 Turnstile quantum dot single-electron transfer

One way to transfer electrons one by one between two electron reservoirs is to use a turnstile motion of tunnel barriers formed between the reservoirs and a quantum dot. This type of single-electron transfer through a semiconductor quantum dot was demonstrated by Kouwenhoven et al. in 1991 [2]. Fig. 5.1 shows schematics of turnstile operation. In (a), a quantum dot is set up with N electrons. The electrochemical potential of the $(N+1)$th state is set between the electrochemical potentials (Fermi energies) $\mu_{\mathrm{L}}, \mu_{\mathrm{R}}$ of the left (source) and right (drain) reservoirs under an applied source–drain bias. Initially, both tunnel barriers are set to be opaque so that electrons cannot tunnel into or out of the dot. In order to start turnstile operation, firstly the left (entrance) barrier is lowered so that an electron tunnels into the dot (b). At the same time, the right (exit) barrier is raised. Once the $(N+1)$th state is occupied, no other electrons can tunnel into the dot as the electrochemical potential of the $(N+2)$th state is above the Fermi energies of both reservoirs. The electrons trapped in the dot cannot tunnel out as tunnelling to the left reservoir is blocked by occupied states in that reservoir, and tunnelling to the right reservoir is blocked by the high right barrier. Then, the left barrier is raised to isolate the dot from the reservoirs (c). Finally, as shown in (d), as the left barrier is raised further, the right barrier is lowered.

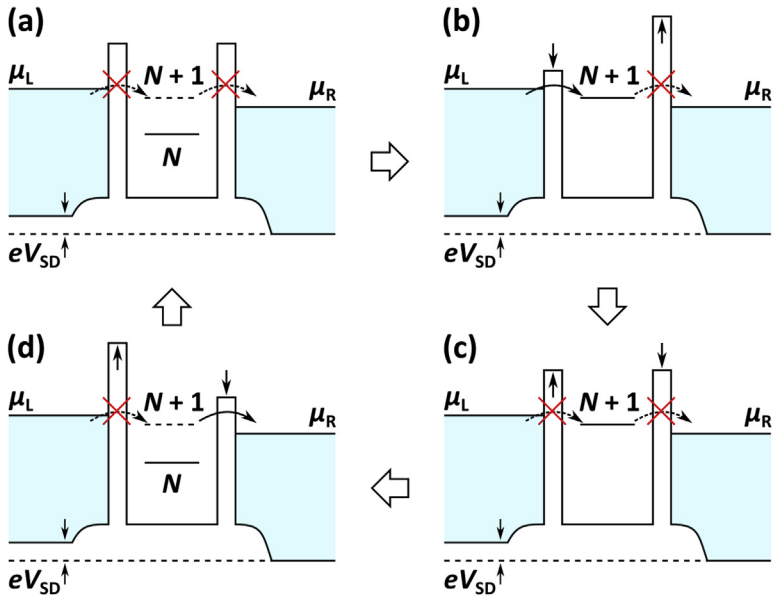


Figure 5.1 Schematics showing turnstile single-electron transfer through a quantum dot. (a) A quantum dot containing N electrons is set up with the electrochemical potential of the $(N + 1)$th state between those of the reservoirs. Tunnelling is blocked by high barriers. (b) The left barrier is lowered to allow for an electron to enter the dot. (c) The left barrier is raised to block tunnelling. (d) The right barrier is lowered to let the electron tunnel into the right reservoir. This completes a transfer of one electron from the left to right. *Adapted from Fig. 1 in Ref. [2].*

At this stage, an electron in the dot tunnels out into the right reservoir, completing one-electron transfer from left to right. Note that at least one of the barriers is high at all stages of the cycle to prevent the source–drain bias V_{SD} from driving electrons from left to right in an uncontrolled manner. For repeated operations, the right barrier is raised to go back to (a).

In Fig. 5.2, as an illustration, current I_{SD} through the quantum dot driven by turnstile motion is plotted against the bias voltage V_{SD} applied between the source and drain contacts (note that the curve is not from actual experimental data). The bias can change the number of electrons transferred per cycle, because a larger bias allows multiple electrons to tunnel into and out of the dot. The I–V curve in Fig. 5.2 shows current plateaux (marked by red arrows) corresponding to $I_{SD} = nef$, where n is the average number of electrons transferred per cycle, and f is the frequency of the operation. Here, $f = 10$ MHz, producing $ef \approx 1.6$ pA. Some current plateaux appear to be missing because the energy levels of the dot are not aligned correctly to yield

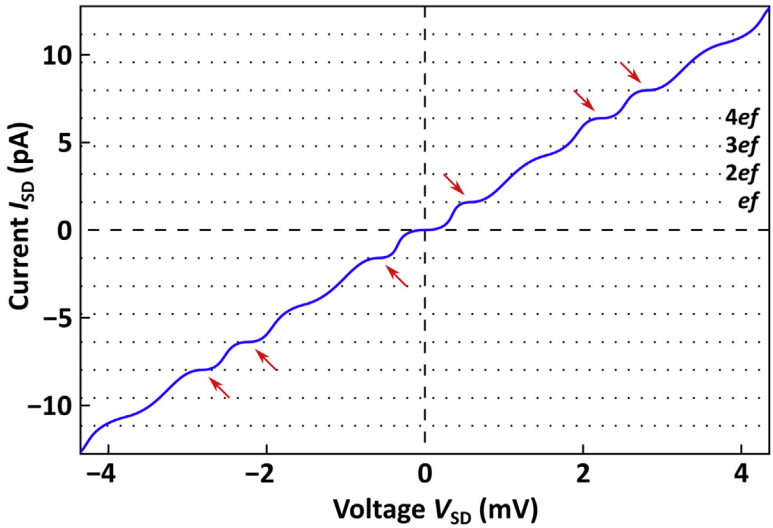


Figure 5.2 An illustration of *I–V* characteristics of turnstile electron transfer (not actual experimental data). Red arrows indicate quantised current at *nef*, with $f = 10$ MHz. *Adapted from Fig. 2 in Ref. [2].*

the turnstile mechanism shown in Fig. 5.1. The maximum frequency that can be used for turnstile electron transfer is limited by the RC time constant of the tunnel barriers. For typical turnstile operations, the minimum tunnel resistance R of the barriers is of the order of 1 MΩ. The capacitance C of the barrier is of the order of 10^{-16}F. These give an RC time constant of the order of 10^{-10}s(or 100 ps). Ref. [2] discusses how this RC constant limits the accuracy of electron transfer. The probability of unwanted tunnel events can be written as $P = \exp[-(1/f)/RC]$. For $f = 10$ MHz, $P = \exp(-1000) \approx 0$ and is negligible. But for $f = 1$ GHz, $P = \exp(-10) \approx 5 \times 10^{-5}$, and becomes nonnegligible for metrological applications, which requires an accuracy at a level of 1 part in 10^8.

1.2 Multi-junction single-electron pump

In the turnstile pumps mentioned above, the direction of electron flow is determined by the bias applied between the source and drain contacts. In 1992, Pothier et al. demonstrated a different type of single-electron transfer, in which the relative phase of ac gate signals determines the direction of electron flow [3]. Their device consists of two metallic charge islands placed between three tunnel junctions in series as shown in Fig. 5.3a. The charge

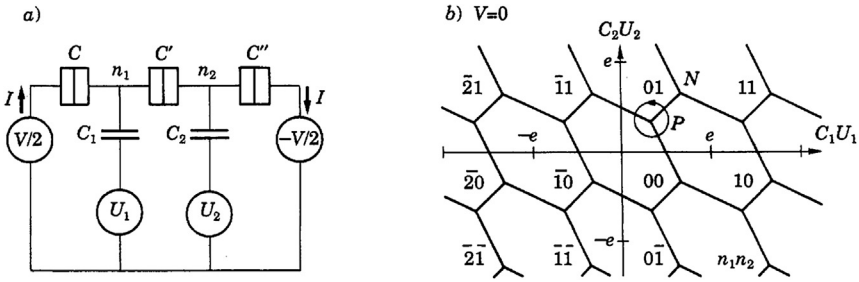


Figure 5.3 (a) Experimental circuit of a three junction single-electron pump. (b) Charge stability diagram of the two islands. The numbers within each hexagon show the number of electrons (n_1, n_2) in each island relative to the hexagon (00). The bar above the number represents a negative number. In order to perform single-electron transfer, the plunger gate voltages U_1 and U_2 are driven so that the operation point circles around a triple point, such as the one marked as P. *Reproduced with permission from Ref. [3].*

stability diagram of the double islands within the phase space of the voltages U_1 and U_2 controlling the number of electrons (n_1 and n_2) on each island is shown in Fig. 5.3b. Each hexagon represents a region with a certain combination of n_1 and n_2 as indicated by their relative numbers compared to the hexagon (00) [4]. In order to transfer electrons, the plunger gate voltages U_1 and U_2 are driven so that the operating point circles around triple points P in the direction as shown by a circle with an arrow in Fig. 5.3b. Starting from $(n_1, n_2) = (0, 0)$, one electron tunnels from the left reservoir into the left island [(1,0)], then one electron transfers from the left island to the right island [(0,1)]. Finally, one electron in the right island tunnels out to the right reservoir [(0,0)] to complete a transfer of one electron.

The *I*–*V* characteristics with two different values of the phase shift Φ between the driving signals and with no RF driving signals are shown in Fig. 5.4. It shows that there is a range of source–drain bias *V* within which the number of electrons transferred per cycle is quantised at one. It also shows that the direction of current flow can be switched by simply changing Φ by π. Similar single-electron transfer has been realised in other material systems, such as graphene [5] and two dopant atoms embedded in a Si nanowire [6]. Whilst the speed of an accurate pump operation is limited by the *RC* time constant of tunnel junctions, just as in the turnstile single-electron transfer, one advantage is that the number of junctions can be increased to suppress electron transfer error due to co-tunnelling. We will discuss the implementation of seven-junction single-electron transfer in the Section 'Quantum current standard' later.

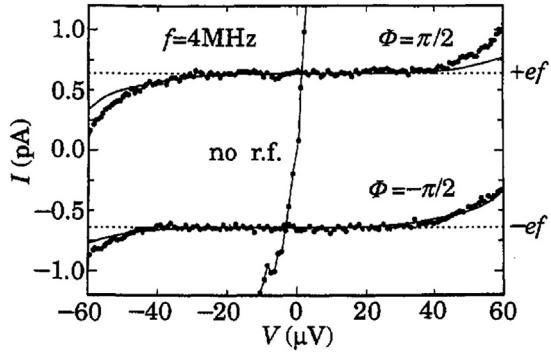


Figure 5.4 *I–V* characteristics of a three-junction pump. Curves for two values of phase shift Φ and for no RF driving signal are shown. *Reproduced with permission from Ref. [3].*

1.3 Superconductor-normal metal hybrid turnstile

In 2008, Pekola et al. introduced a new type of single-electron turnstile based on a metal–superconductor hybrid structure [7]. As shown in Fig. 5.5, their device consists of a normal metal charge island sandwiched by superconducting electrodes, with insulating oxide layers forming tunnel junctions between the superconductor and metal (SiNiS structure). Note that the inverted structure with a superconducting charge island with normal metal electrodes (NiSiN) is also possible, but this structure suffers from Joule heating.

In order to understand the operation of this hybrid turnstile device, we look at the charge stability diagram in the phase space of normalised gate voltage (n_g) and source–drain bias (V). In Fig. 5.6a, the stability diagram of a normal metal charge island with normal metal electrodes is shown. In this case, the stable $n = 0$ and 1 regions are connected by a single point at $V = 0$. The lines labelled as '0,LF' and '1,LB' ('1,RF' and '0,RB')[1] mark where the electrochemical potential of the $n = 1$ state lines up with the electrochemical potential of the left (right) reservoir. Driving n_g between $n = 0$ and 1 along the horizontal bold line goes through a region (un-shaded) where an uncontrolled amount of forward current flows, and hence the current is not quantised. For a SiNiS device with a superconducting gap Δ in the leads, the stability diagram opens up by 4Δ, and an overlapped region is created between the $n = 0$ and $n = 1$ regions (Fig. 5.6b). In this overlapped region, the charge status is locked. Starting from $n = 0$ along the

[1] The 'F' denotes 'forward' tunnelling, whilst 'B' denotes 'backward' tunnelling.

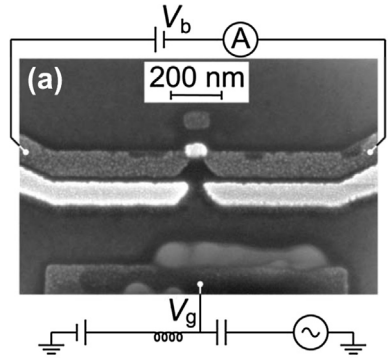


Figure 5.5 A scanning electron micrograph of a SiNiS turnstile device fabricated by shadow evaporation technique. Normal metal (AuPd, brighter part) and superconductor (Al, darker part) are evaporated at different angles through the same resist mask to form the accurately aligned structure. *Reproduced with permission from Ref. [8].*

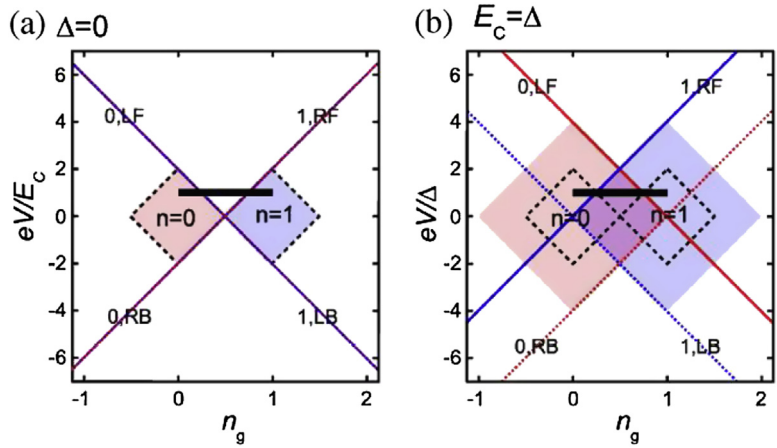


Figure 5.6 Charge stability diagram of (a) normal-metal charge island with normal-metal electrodes and (b) normal-metal charge island with superconducting electrodes. V is the source–drain bias, E_C is the unit of charging energy of the island and Δ is the superconducting gap. n_g is the normalised gate voltage that changes the number of electrons on the island. *Reproduced with permission from Ref. [9].*

horizontal bold line, when n_g crosses the '1,RF' line into the overlapped region, the charge state of the island stays at $n = 0$. When n_g crosses the '0,LF' line into the $n = 1$ un-overlapped area, an electron enters from the left reservoir into the charge island. When the direction of n_g changes and it crosses the '0,LF' line into the overlapped area, the charge state stays at

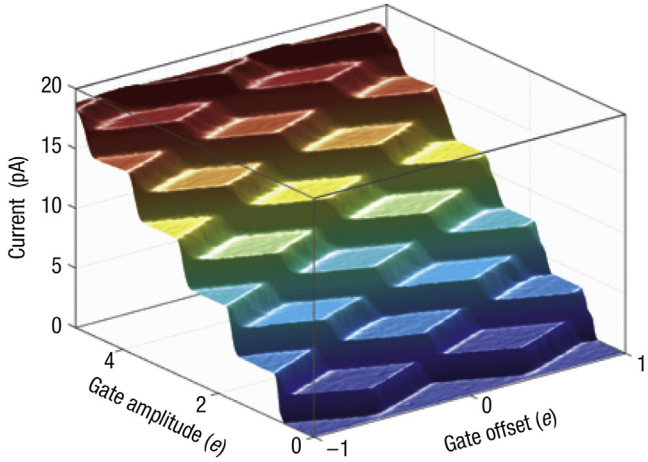


Figure 5.7 Current plateaux produced by a hybrid turnstile device as a function of gate amplitude and gate offset. Operation frequency was 12.5 MHz. *Reproduced with permission from Ref. [7].*

$n = 1$. When n_g crosses the '1,RF' line into the $n = 0$ un-overlapped area, an electron in the charge island tunnels out onto the right reservoir, completing a single-electron transfer. The direction of electron transfer can be reversed by switching the sign of the source–drain bias. If the amplitude of the ac gate signal is increased to extend the region in the stability diagram, a larger number of electrons can be transferred per cycle. Current plateaux generated by a hybrid turnstile device are shown in Fig. 5.7.

Since this device architecture is also based on tunnel junctions as turnstile quantum dots and multi-junction pumps, one might expect that the RC time constant is the limiting factor. However, consideration of other error mechanisms needs to be taken into account. An optimum choice of the source–-drain bias is found to be roughly $V = \Delta/e$ in order to minimise thermal errors [7]. Pump errors due to Andreev tunnelling (Cooper-pair tunnelling through the junction) can be avoided by choosing the sample with a larger charging energy E_C than the superconducting gap (i.e. $E_C > \Delta$)[10]. In this regime, the limiting process was found to be co-tunnelling of a Cooper pair and a single electron [10]. In order to suppress this, the tunnel junctions need to be opaque and the pump frequency has to be limited to $f \sim 6$ MHz($ef \sim 1$ pA).

1.4 Surface acoustic wave single-electron transfer

The single-electron transfer mechanisms described so far rely on electrostatic confinement by a lithographically defined charge island or quantum dot.

The island is connected to leads through tunnel barriers. The RC time constant of the barriers restricts the maximum operation frequency to around 10 MHz, limiting the scope for metrological application as quantum current standards. In 1996, Shilton et al. demonstrated a new method of single-electron transfer using surface acoustic waves [11]. In piezoelectric materials, such as lithium niobate and gallium arsenide, surface acoustic waves can be generated by applying microwaves to an inter-digitated transducer (see Fig. 5.8) at the resonant frequency $f = v_{SAW}/\lambda$, where v_{SAW} is the velocity of surface acoustic waves ($\sim$2800 m/s in GaAs) and λ is the periodicity of the fingers of the inter-digitated transducer. For $\lambda = 1$ μm, $f \approx 2.8$ GHz. Due to the piezoelectricity, surface acoustic waves are accompanied by electrostatic potential waves. This fast oscillating potential can be used to transport electrons.

Shilton et al. sent surface acoustic waves through a depleted one-dimensional channel fabricated in a GaAs/AlGaAs heterostructure. As surface acoustic waves travel through the one-dimensional channel, the accompanying potential waves pick up electrons from the source two-dimensional electron gas as shown in Fig. 5.9. Initially, multiple electrons may be trapped in the moving quantum dot formed by the combined potential of surface

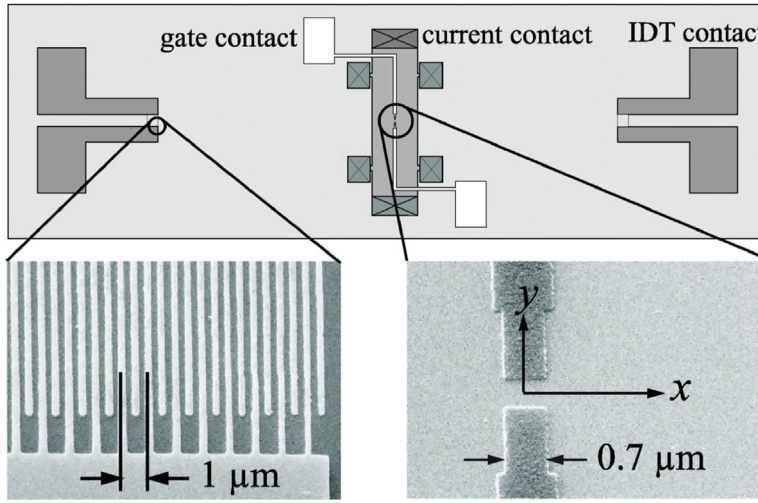


Figure 5.8 Schematics of a surface acoustic wave single-electron pump device made on a GaAs/AlGaAs heterostructure. Inter-digitated transducers (labelled 'IDT') are placed on the sides (only one of them is used for pumping electrons). At the centre of the chip, a quasi-one-dimensional channel is formed by surface split gates in a two-dimensional electron gas. *Reproduced with permission from Ref. [12].*

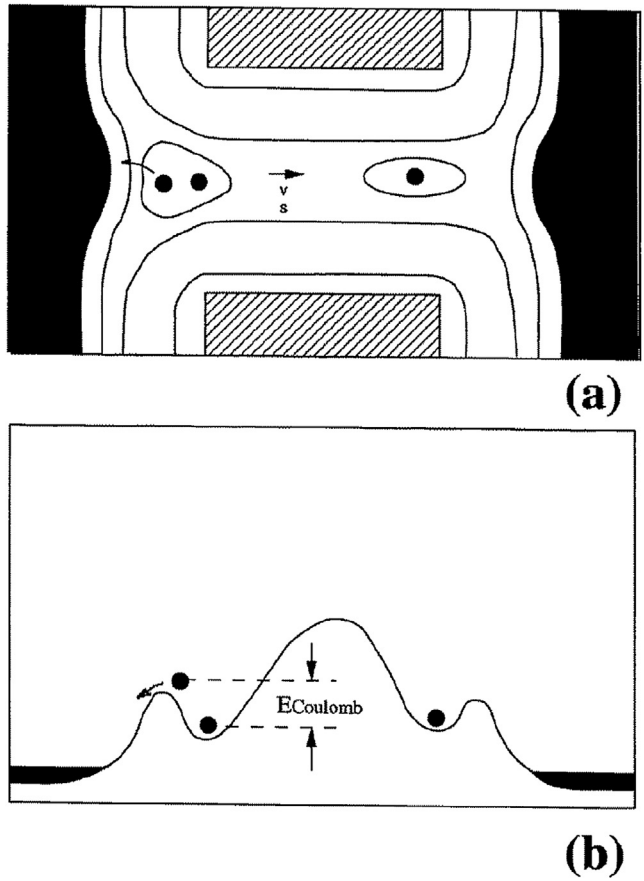


Figure 5.9 (a) Contour profile of electrostatic potential formed by surface acoustic waves and a quasi-one-dimensional confinement potential. Black circles represent electrons trapped in moving quantum dots. (b) Profile of electrostatic potential along the centre of the channel, showing an electron squeezed out of the moving quantum dot. *Reproduced with permission from Ref. [11].*

acoustic waves and one-dimensional confinement. As the quantum dot potential rises with the channel potential, electrons are squeezed out of the dot and tunnel back to the source reservoir. Under optimised conditions, only one electron is left in the moving quantum dot, which is carried to the drain reservoir. Fig. 5.10 shows acoustoelectric current quantisation as the number of electrons transported is varied by the split-gate voltage [13]. Here, $f = 2.71625$ GHz, and hence $ef \approx 435$ pA. Although the operation frequency is limited to a narrow bandwidth of the transducer resonance,

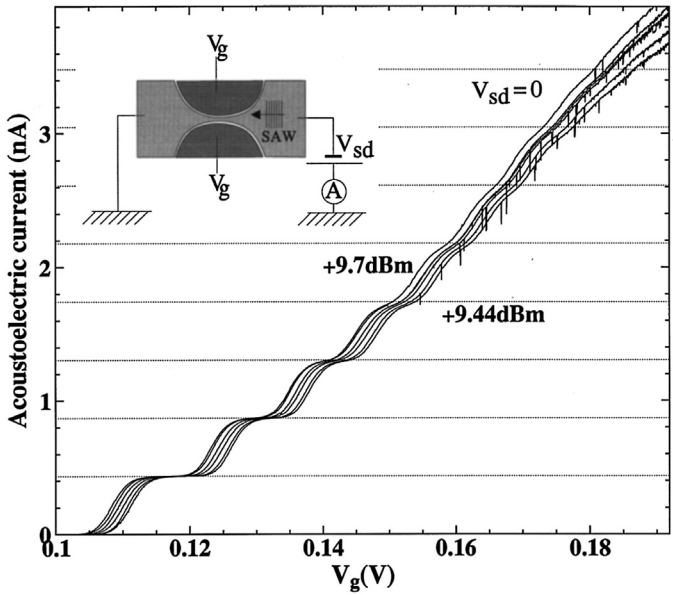


Figure 5.10 Acoustoelectric current quantisation produced by surface acoustic waves through a quasi-one-dimensional channel (shown in the inset). Operation frequency was 2.71625 GHz, producing $ef \approx 435$ pA. Multiple curves with varied microwave power applied to the transducer are shown. *Reproduced with permission from Ref. [13].*

quantised charge pumping at gigahertz frequencies can be achieved due to the absence of lithographically formed tunnel barriers.

The method described above periodically transfers electrons trapped in each potential well created by surface acoustic waves. There is a related method that uses pulses of surface acoustic waves to transfer single electrons controllably between two locations. In 2011, Hermelin et al. [14] and McNeil et al. [15] demonstrated transfer of one electron between two quantum dots connected by a one-dimensional channel. Fig. 5.11 shows the method used by McNeil et al. As shown in (a), a GaAs/AlGaAs two-dimensional electron gas system is split into two parts by the long horizontal Separation Gate. In the upper part, two quantum dots (LQD and RQD) are connected by a one-dimensional channel, which is depleted by applied gate voltages. In the lower part, a quantum point contact detector monitors the charge state of each quantum dot. Inter-digitated transducers are placed on both sides of the device. Quantum dots are initialised by a sequence of barrier and plunger gate operations (Fig. 5.11b). In this example, one electron is trapped in LQD and is brought close to the channel potential, whilst RQD

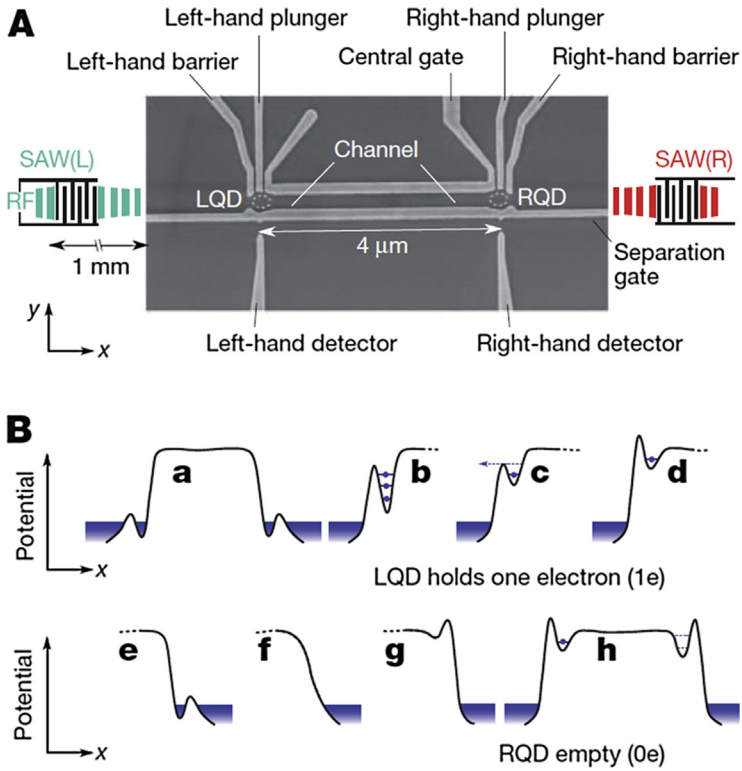


Figure 5.11 (a) A scanning electron micrograph of a GaAs/AlGaAs device for single-electron transfer between two quantum dots by a surface acoustic wave pulse. Left and right quantum dots are labelled as LQD and RQD, respectively. A transducer is placed on each side of the device. (b) **a**–**d**: An initialisation gate sequence to trap one electron in LQD. **a**. Start with the dot populated by the Fermi sea. **b**. The barrier is raised to trap multiple electrons above the Fermi energy. **c**. the plunger pushes up the dot energy to depopulate the quantum dot leaving only one electron. **d**. The barrier is raised further to bring the electrochemical potential of the populated dot state close to the channel potential. **e**–**h**: An initialisation gate sequence to prepare empty RQD. **e**. Start with the dot populated by the Fermi sea. **f**. The plunger is raised to deplete the quantum dot completely (at this point confinement potential is lost). **g**. Raise the barrier to form a quantum dot confinement potential close to the channel potential. **h**. Drop the plunger to form a deep quantum dot potential ready to receive an electron. *Reproduced with permission from Ref. [15].*

is prepared empty. Then, a short microwave pulse (its duration ranging from 10 ns to a few hundred ns) is applied to the left transducer. This generates a pulse of surface acoustic waves that takes out the electron from LQD, carrying it through the channel into RQD. When a successful transfer of the electron is made, the charge detector on LQD registers the loss of an electron (a

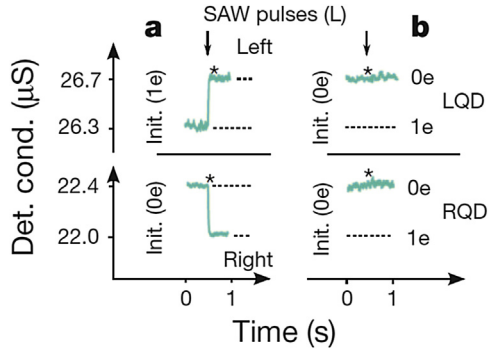


Figure 5.12 Time traces of quantum point contact charge detector conductance before and after a pulse of surface acoustic waves from the left transducer (sent at the time marked by *). (a) Successful electron transfer from LQD to RQD. (b) A control experiment starting with no electrons in both quantum dots. There is no change in their charge status after a pulse of surface acoustic waves. *Reproduced with permission from Ref. [15].*

step increase in conductance) whilst the RQD charge detector registers a gain of an electron (a step decrease in conductance) as in Fig. 5.12a. In order to confirm that the population and depopulation of the quantum dots do not happen randomly, a control experiment was performed starting with no electrons in both quantum dots (Fig. 5.12b). In this case, even after a surface acoustic wave pulse was sent, the charge detectors showed no change. A study of the correlation between the initial states and outcome confirmed successful electron transfer. It was also demonstrated that the same electron can be moved repeatedly back and forth between the two quantum dots by sending surface acoustic wave pulses alternately from left and right transducers.

1.5 Tunable barrier quantum dot pump

Another type of single-electron pump that has achieved gigahertz operation is the tunable barrier quantum dot pump. This pump overcomes the RC time constant limitation by modulating its entrance barrier between completely open $(R \ll h/e^2)$ and completely closed $(R \gg h/e^2)$. Tunable barrier single-electron pumping was first demonstrated by Blumenthal et al. by modulating both entrance and exit barriers in 2007 [16]. The method was later simplified to single-parameter operation (modulating just the entrance barrier) in 2008 [17,18].

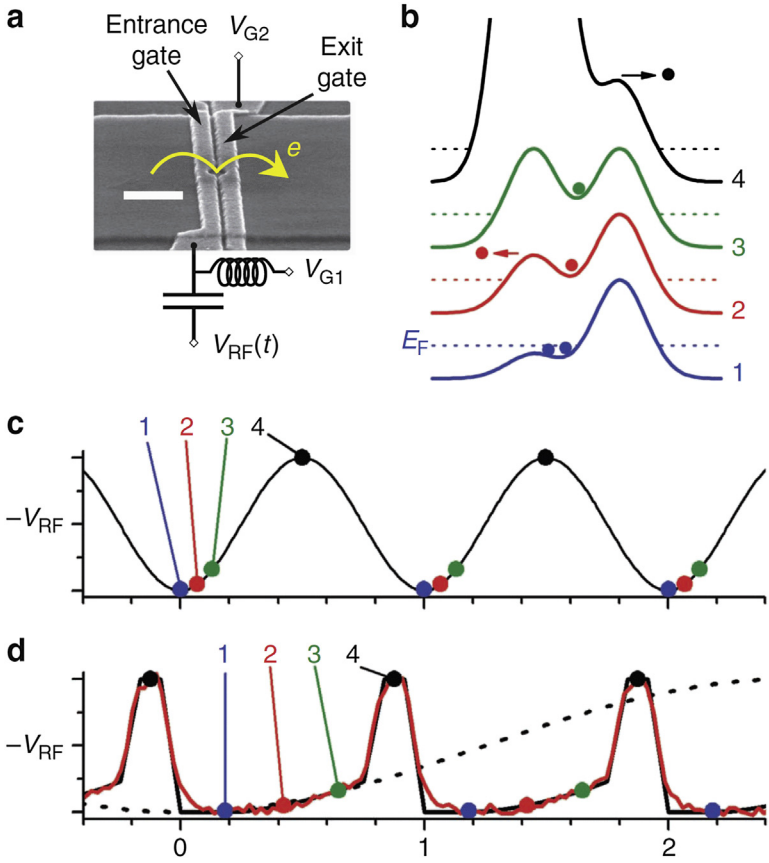


Figure 5.13 (a) Scanning electron micrograph of a GaAs/AlGaAs tunable-barrier quantum dot pump. The white scale bar is 500 nm. (b) Pump sequence 1–4. Electrostatic potential profile along the centre of channel is plotted. Filled circles represent electrons trapped in the quantum dot. (c) Sinusoidal pump excitation waveform V_{RF} plotted as a function of normalised time t/T on the x axis, where T is the period of the waveform. The numbered points correspond to the pump sequence plotted in (b). (d) Tailored pump waveform to slow down the decay cascade part (1–3). The solid black curve corresponds to a programed waveform, whilst the red (grey in print version) curve is an oscilloscope trace of the output of arbitrary waveform generator. The dotted line is a sinusoidal wave with a larger period for comparison. *Reproduced with permission from Ref. [19].*

The operation mechanism of a tunable barrier pump is shown in Fig. 5.13. The device consists of two gates (entrance and exit barriers) that cut across a narrow channel of two-dimensional electron gas. Between the gates, a quantum dot is formed that traps electrons. While there are various designs, the one shown in Fig. 5.13a shows a circular cutout in the gates

for a quantum dot, whilst other designs may form a dot in the gap between two straight barrier gates [17]. The voltage on the exit barrier gate is set to create a high potential barrier, forbidding tunnelling between the source and drain reservoirs. To start pumping, the RF signal on the entrance barrier gate lowers the entrance barrier below the Fermi energy. At this point, the potential confinement is lost, and electrons flood into the area where a quantum dot is to be formed (Fig. 5.13b-1). As the entrance barrier rises (Fig. 5.13b-2), the newly formed quantum dot traps multiple electrons. The rising entrance barrier lifts up the energy of trapped electrons, squeezing them out of the dot back to the source reservoir. This process can be approximated by the decay cascade model [20]. As electrons leave the dot, the electrochemical potential of the highest occupied state drops in a step-like manner, resulting in corresponding reduction in the tunnel coupling. When the tunnel rate becomes sufficiently small compared to the duration of the decay cascade process, the electron number in the dot is fixed (Fig. 5.13b-3). The entrance barrier is raised further, and eventually the electrons trapped in the dot are emptied over the exit barrier onto the drain reservoir (Fig. 5.13b-4). Then the cycle is repeated.

As will be discussed later in the Section 'Quantum current standard', the record accuracy of single-electron transfer has been achieved by tunable-barrier quantum dot pumps. Towards the improvement of current quantisation accuracy, two important experimental observations have been made. Firstly, a magnetic field applied perpendicular to the plane of the quantum dot improves the quantisation accuracy of current plateaux as shown in

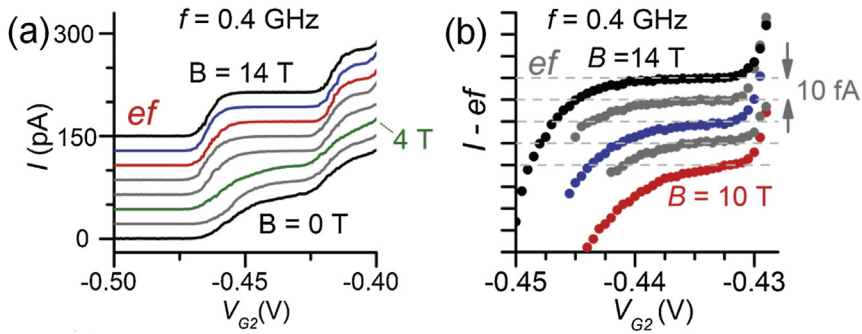


Figure 5.14 (a) Effect of perpendicular magnetic field on current plateaux. Pump current is plotted as a function of the exit-barrier gate voltage V_{G2} measured with applied magnetic fields B ranging from 0 to 14 T. Curves are offset vertically. (b) Zoomed in region around *ef* between $B = 10$ and 14 T, showing how a larger magnetic field 'flattens' the current quantisation plateau. *Reproduced with permission from Ref. [21].*

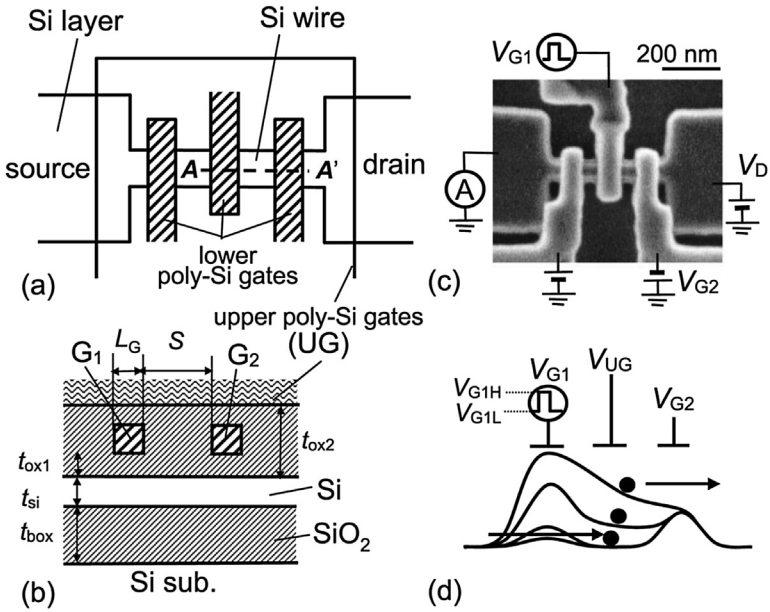


Figure 5.15 Silicon tunable-barrier single-electron pump. (a) Top-view schematics of the device. (b) Cross-sectional schematics of the device. (c) Scanning electron micrograph of the device. (d) Potential profiles along the centre of the channel, showing pumping behaviour (filled circles representing electrons). *Reproduced with permission from Ref. [18].*

Fig. 5.14 [21–23]. For metrological applications, the device operation point must be chosen where the magnitude of the current produced is robust against external fluctuations. This requires the current to show a flat plateau (at the *ef* value) when a device parameter is swept. Fig. 5.14 shows improvement of the 'flatness' of the current plateau as the magnetic field is increased. Secondly, by using a tailored waveform to slow down the movement of the entrance barrier deliberately during the decay cascade process (Fig. 5.13d), flat current quantisation plateaux are maintained at high pumping frequencies, where the accuracy degrades if pure sine waves are used. The experiments shown in Figs. 5.13 and 5.14 were made on a GaAs/AlGaAs heterostructure.

Tunable-barrier quantum dot pumps have been also realised in silicon nanowire systems [18,24] as shown in Fig. 5.15. With a silicon device in which electrons are pumped through a trap level rather than in a lithographically defined quantum dot, operation frequencies up to ~7 GHz have been reported [25].

1.6 Mesoscopic capacitor

A mesoscopic capacitor is a tiny charge island (quantum dot) capacitively coupled to a gate and tunnel coupled to an electron reservoir. By oscillating the voltage on the gate, it can emit electrons into the reservoir or absorb electrons from the reservoir (i.e., emit conduction band holes into the reservoir) alternately. In 2007, Fève et al. demonstrated coherent emission of single electrons (and single holes) from a mesoscopic capacitor [26]. Their experiment was performed in a magnetic field so that a quantum dot was coupled to a quantum Hall edge state as shown in Fig. 5.16a. A square waveform applied to the plunger gate with an amplitude matched to the energy-level spacing Δ of the dot oscillates the energy level of the dot up and down around the Fermi energy of the reservoir (Fig. 5.16a①–③). When the filled state rises above the Fermi energy, an electron is emitted into the edge state in the reservoir (②). When the empty state is brought back below the Fermi energy, an electron in the edge state tunnels into the dot, emitting a conduction band hole into the edge state (③).

Fig. 5.16b shows the measured average current in the time domain for different values of transmission D through the barrier. Larger D leads to smaller decay time τ. Here, the repetition rate is $f = 31.25$ MHz. Note that Fig. 5.16b shows a fast time-resolved measurement of current, in contrast to the relatively accurate but slow dc-current measurements reviewed so far. For this system, it is important to find a balance between D and f. For certain types of measurements (such as shot-noise measurements), it is advantageous to make f as large as possible, as it enhances the size of the signal. In order to complete electron (and hole) emission within the half cycle, the transmission D needs to be increased accordingly. On the other hand, making D too large weakens charge quantisation in the dot (energy levels become broadened and eventually cannot be resolved). In Ref. [26], $D \approx 0.2$ was found to provide optimal working conditions for $f = 180$ MHz. Mesoscopic capacitors are used in various 'electron quantum optics' experiments as described in the Section 'Electron quantum optics'.

1.7 Leviton

Up until now, we have discussed sources of single electrons in which the electrons are initially confined in well-defined energy levels in a quantum dot. We now discuss a more exotic type of charge excitation which does not involve confinement of electrons. When a time-dependent bias voltage

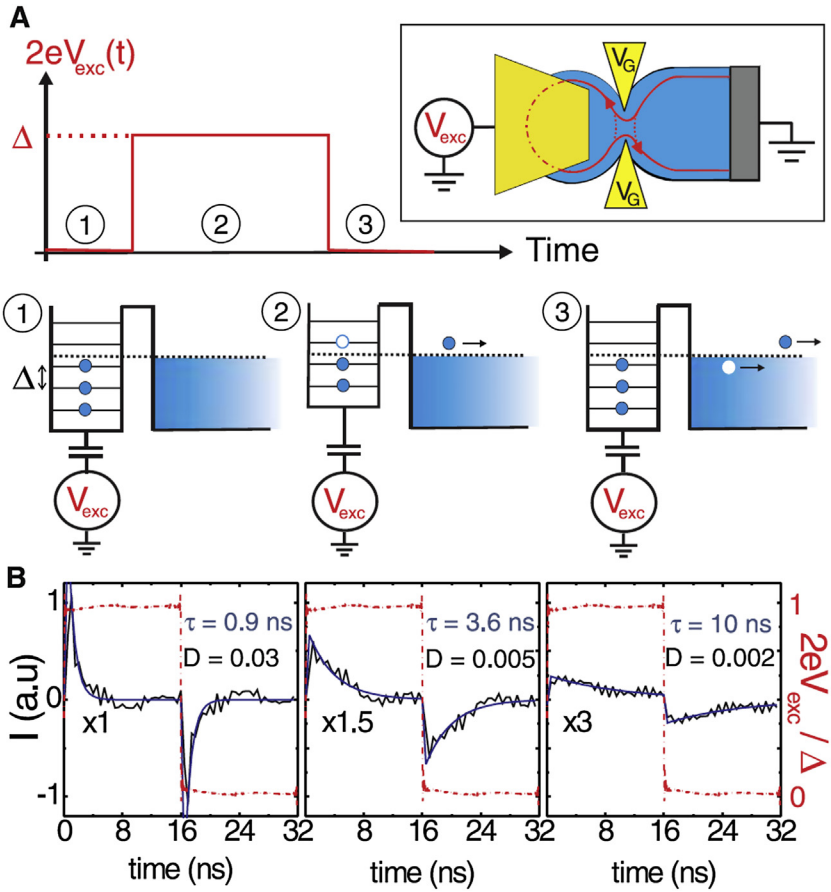


Figure 5.16 (a) Mechanism of electron hole emission by a mesoscopic capacitor. Top-right: schematics of the device. Arrowed red (grey in print version) line represents a quantum Hall edge channel. Top-left: Square waveform applied to the gate capacitively coupled to the quantum dot. Bottom: Phases of device operation. ① Energy levels in the quantum dot are filled up to the Fermi energy of the reservoir. ② When the highest occupied energy level is pushed above the Fermi energy, an electron is emitted into the edge state in the reservoir. ③ When the empty state is brought back below the Fermi energy, an electron tunnels into the dot from the reservoir, emitting a conduction band hole into the edge state. (b) Time-domain measurements of average current for different values of transmission *D*. The red (grey in print version) dotted trace is the applied excitation signal corresponding to the red square (dark grey in print version) waveform shown in (a). *Reproduced with permission from Ref. [26].*

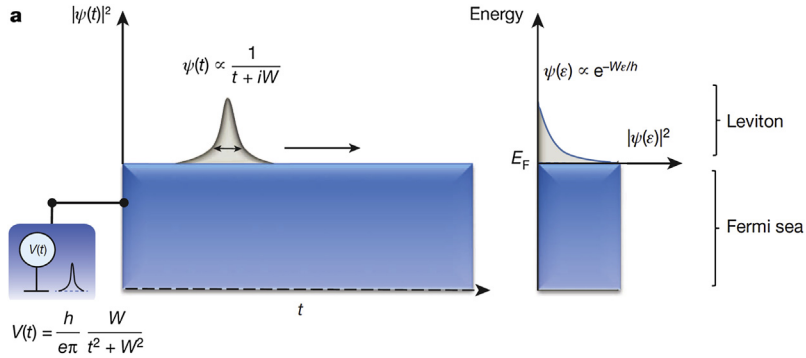


Figure 5.17 Leviton excitation. Left panel (time domain): Lorentzian pulse $V(t)$ is applied to a Fermi sea (represented by a blue [grey in print version] region), exciting a leviton $\psi(t)$. Note that the voltage pulse is applied to the source of a conductor, not to gates as in the previous examples. Right panel (energy domain): Leviton excitation $\psi(\varepsilon)$ only exists in the domain above the Fermi energy E_F. *Reproduced with permission from Ref. [27].*

$V(t)$ is applied to a conductor such that $n = \frac{e}{h}\int_{-\infty}^{\infty} V(t)dt$ is an integer, an integer number (n) of charges $Q = ne$ are injected. Generally, this also excites a neutral mixture of electrons above the Fermi energy and conduction band holes below the Fermi energy. However, Levitov et al. showed that the use of Lorentzian-shaped pulses suppresses these neutral excitations, and only electron excitations are generated [28–30]. Experimental demonstration of the predicted 'minimal' excitations (named 'Leviton') was performed by Dubois et al. in 2013 [27]. Fig. 5.17 illustrates excitation of levitons. A voltage pulse $V(t) = \frac{h}{e\pi}\frac{W}{t^2+w^2}$, where $2W$ is the width of the Lorentzian pulse, is applied to a Fermi sea. This generates an excitation $\psi(t) \propto \frac{1}{t+iW}$ in the time domain, and $\psi(\varepsilon) \propto e^{-W\varepsilon/h}$ in the energy (ε) domain.

The absence of hole excitations during leviton excitations was confirmed by characterising the energy distribution $\widetilde{f}(\varepsilon)$ of excitations. Such information can be deduced from shot-noise measurements across a partition barrier with a dc bias (Fig. 5.18). As shown in Fig. 5.18a, when electron and hole excitations are created in the left reservoir, they contribute to the partition noise by the barrier with a transmission $0 < D < 1$. If the dc bias $V_{\mathrm{d.c.}}$ is applied such that the Fermi energy of the left reservoir is raised by $h\nu$ (ν is the frequency of applied ac excitation), then the energy of hole excitation

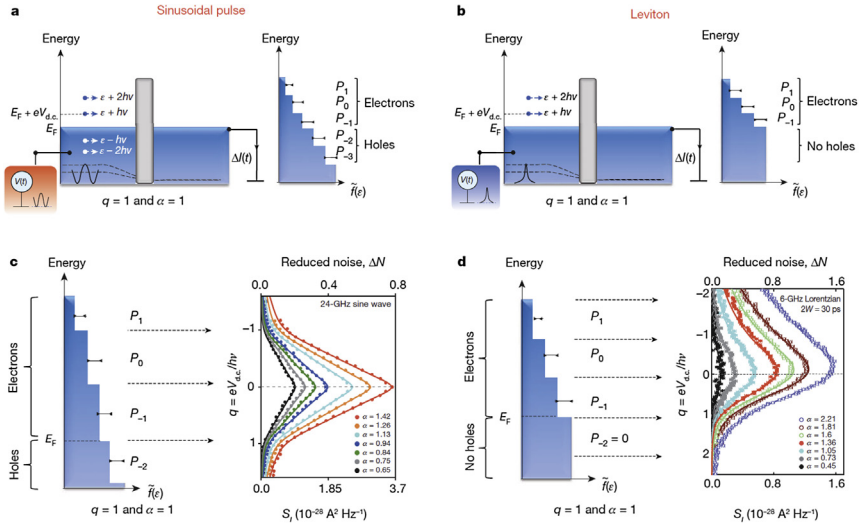


Figure 5.18 (a) Electron and hole excitations by sinusoidal drive (b) Electron excitations (no holes) by Lorentzian pulses. (c) and (d) Spectroscopic study of shot noise under sinusoidal drive [(c)] and Lorentzian pulses [(d)]. *Reproduced with permission from Ref. [27].*

within $h\nu$ of the Fermi energy in the left reservoir becomes higher than the Fermi energy of the right reservoir. At zero temperature, there is no electron excitation in this energy range in the right reservoir; hence, the transmission of the holes in the left reservoir becomes forbidden, and hence they do not contribute to the shot noise. Similarly, if $V_{\mathrm{d.c.}}$ lowers the Fermi energy of the left reservoir by $h\nu$, the transmission of electron excitations within $h\nu$ of the Fermi energy becomes forbidden, not contributing to the partition noise. Fig. 5.18c shows the shot-noise measured as a function of $V_{\mathrm{d.c.}}$ for sinusoidal excitation. As expected, the suppression of noise at a larger $|V_{\mathrm{d.c.}}|$ is symmetric around the Fermi energy, suggesting the excitations of both electrons and holes. Fig. 5.18b and d are for the case of leviton excitation. Here, the suppression of noise is asymmetric around the Fermi energy, suggesting the absence of hole excitations.

2. Quantum current standard

As mentioned in the introduction to this chapter, one of the envisaged applications of single-electron sources is the primary metrological standard

for electrical current. In this section, we discuss the effort to improve and verify the accuracy of current quantisation by single-electron sources.

2.1 Realisation of SI ampere

Before 2019, the definition of the SI ampere was based on Ampere's force law that requires mechanical force measurements. Such experiments were difficult to perform accurately in the laboratory. The 26th General Conference on Weights and Measures (CGPM) held in November 2018 adopted a resolution to revise the definition of SI units [31]. Following this decision, four (kilogram, mole, kelvin and ampere) of the seven SI base units were re-defined effective from 20 May 2019. In fact, the new SI does not directly define these units. Rather, it fixes the values of seven defining constants (a set of fundamental constants and technical constants): $\Delta\nu_{Cs}$ (the unperturbed ground state hyperfine transition frequency of ^{133}Cs), c (the speed of light in vacuum), h (the Planck constant), e (the elementary charge), k (the Boltzmann constant), N_A (the Avogadro constant) and K_{cd} (the luminous efficacy of monochromatic radiation of frequency 540×10^{12} Hz) (see Table 5.1. Note, the values of $\Delta\nu_{Cs}$, c and K_{cd} were already fixed before 2019). The definition of all seven SI base units are derived from these defining constants [1].

The ampere is now defined by the following description [1]:

The ampere, symbol A, is the SI unit of electric current. It is defined by taking the fixed numerical value of the elementary charge e to be $1.602\ 176\ 634 \times 10^{-19}$ when expressed in the unit C, which is equal to A·s, where the second is defined in terms of $\Delta\nu_{Cs}$.

Appendix 2 to the ninth SI Brochure [32] lists three methods of practical realisation of the ampere:

1. by using Ohm's law, the unit relation $A = V/\Omega$, and using practical realisations of the SI derived units the volt V and the ohm Ω, based on the Josephson and quantum Hall effects, respectively; or

Table 5.1 Defining constants of the SI base units.

Symbol	Value	Unit
$\Delta\nu_{Cs}$	9 192 631 770	Hz
c	299 792 458	m/s
h	$6.626\ 070\ 15 \times 10^{-34}$	Js
e	$1.602\ 176\ 634 \times 10^{-19}$	C
k	$1.380\ 649 \times 10^{-23}$	J/K
N_A	$6.022\ 140\ 76 \times 10^{23}$	mol^{-1}
K_{cd}	683	lm/W

2. by using a single electron transport (SET) or similar device, the unit relation A = C/s, the value of e given in the definition of the ampere and a practical realisation of the SI base unit the second s; or
3. by using the relation $I = CdU/dt$,[2] the unit relation A = FV/ s, and practical realisations of the SI derived units the volt V and the farad F and of the SI base unit seconds.

At the time of writing, as described later, single-electron sources are not yet accurate enough to be used as primary standards. The methods (a) and (c) are most commonly used in metrology laboratories. Especially, in laboratories where cryogenic measurement facilities are available, a combination of Josephson voltage standard and quantum Hall resistance standard [method (a)] can be used to realise the ampere at a level of 1 part in 10^8 [33].

2.2 Quantum metrological triangle

Since the values of Planck constant and elementary charge are fixed under the new SI system, the SI volt and ohm can be also realised independently by the Josephson effect [34] and quantum Hall effect [35]. In these effects, the voltage U and resistance (Hall resistance) R are quantised by the following formulas $U = \frac{1}{K_J} n_J f_J$ and $R = \frac{1}{n_K} R_K$, where $K_J = \frac{2e}{h}$ is the Josephson constant, $R_K = \frac{h}{e^2}$ is the von-Klitzing constant, n_J and n_K are integers and f_J is the frequency of the microwave driving the Josephson array. If a quantum current standard by SET is realised following the formula $I = n_S e f_S$, where n_S is an integer, f_S is the frequency of single-electron transfer, then these three quantum electrical standards form a so-called 'quantum metrological triangle' (Fig. 5.19) [36,37]. The quantum metrological triangle experiment is a consistency test to detect hidden errors that may exist in the realisation of the units. Here, we introduce correction factors ε_J, ε_K and ε_s in each unit realisation, i.e.,

$$U = \frac{1}{(1+\varepsilon_J)K_J} n_J f_J, \tag{5.2}$$

$$R = \frac{1}{n_K}(1+\varepsilon_K) R_K, \tag{5.3}$$

$$I = n_S(1+\varepsilon_S) e f_S. \tag{5.4}$$

Using Ohm's law $U = RI$ and considering that $\varepsilon_J, \varepsilon_K, \varepsilon_S \ll 1$,

[2] C is the capacitance of a capacitor, and U is the voltage applied across the capacitor.

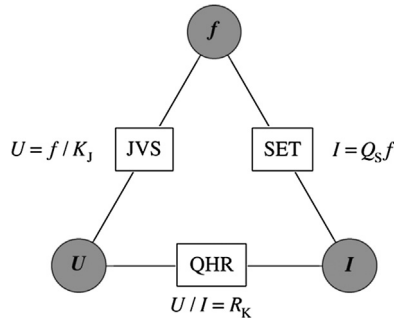


Figure 5.19 Metrological triangle. The quantities *f* (frequency), *U* (voltage) and *I* (current) are connected by the three quantum electrical standards through Ohm's law $U = RI$. *Reproduced with permission from Ref. [36].*

$$\frac{n_J n_K f_J}{2 n_S f_S} = 1 + \varepsilon_J + \varepsilon_K + \varepsilon_S. \tag{5.5}$$

The left-hand side of this equation consists of known integers and frequencies that can be measured with negligible uncertainties. These parameters can be set so that the left term becomes 1. Then, the outcome of the quantum metrological triangle experiment can be expressed as [36]

$$1 = 1 + \Delta_{QMT} \pm \sigma_{QMT}, \tag{5.6}$$

where $\Delta_{QMT} = \varepsilon_J + \varepsilon_K + \varepsilon_S$ and σ_{QMT} is the uncertainty of the experiment. If Δ_{QMT} is found to be less than σ_{QMT}, then the quantum metrological triangle closes within the uncertainty σ_{QMT}, and no correction terms are needed for the realisation of electrical units. In reality, the realisations of the volt and ohm by the Josephson and quantum Hall effects have been well established by metrology institutes for many decades. In the 2014 CODATA adjustments published in 2016 [38], the possible values of corrections ε_J and ε_K were reduced to ~ 2 parts in 10^8. Therefore, the accuracy of single-electron sources needs to be improved to a level of 1 part in 10^8 in order to conduct a meaningful quantum metrological triangle experiment.

2.3 Capacitor charging experiment with a multi-junction pump

In 1999, Keller et al. at the National Institute of Standards and Technology (NIST) demonstrated a capacitance standard based on single-electron transfer [39]. Their device consists of seven tunnel junctions with charge islands in between (see Fig. 5.20a). In their earlier work, they had shown that such a

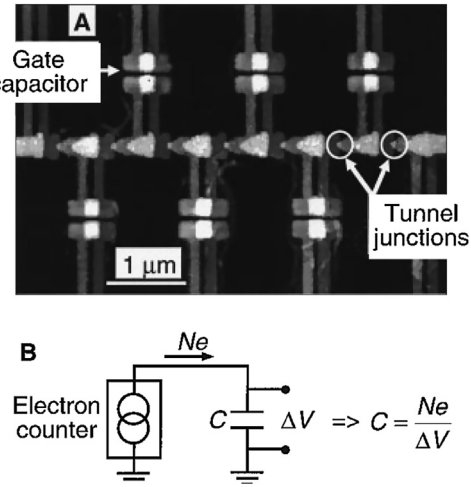


Figure 5.20 (a) Scanning force micrograph of NIST seven-junction pump. (b) Schematic representation of a capacitance standard by single-electron transfer. *Reproduced with permission from Ref. [39].*

device can transfer electrons accurately, although at a relatively slow rate of a few MHz, with an error at a level of 1 part in 10^8 [40]. Their idea is to charge a capacitor with N electrons and measure the voltage change ΔV, which defines the capacitance as $C = \frac{N_e}{\Delta V}$ (Fig. 5.20b). Their experimental realisation using a cryogenic capacitor is shown in Fig. 5.21a. In order to avoid the charging of the island as the pump transfers electrons, the voltage of the island is monitored by a single-electron transistor (marked by 'E' in the figure). The feedback circuit adjusts the voltage V across the cryogenic capacitor to keep the charge island at a virtual ground. This means that all the electrons transferred appear at the cryogenic capacitor, and not at the stray capacitance ($\sim$10 pF). After pumping N electrons, the voltage change ΔV is measured to determine C. Fig. 5.21b shows an ac bridge circuit to compare C with a room temperature capacitor C_{ref}, following $\frac{C}{C_{\mathrm{ref}}} = \frac{V_1}{V_2}$.

In 2007, the NIST team produced a full analysis of the uncertainty budget of this experiment [41]. The cryogenic capacitor was also measured using a calculable capacitor [42], giving a capacitance value in terms of the SI farad, which we assign C_0 here. Comparing C and C_0 is another way to form the quantum metrological triangle (see Ref. [36] for more details). Their final result is[3]

[3] Note that ε_{K} is not involved in this equation.

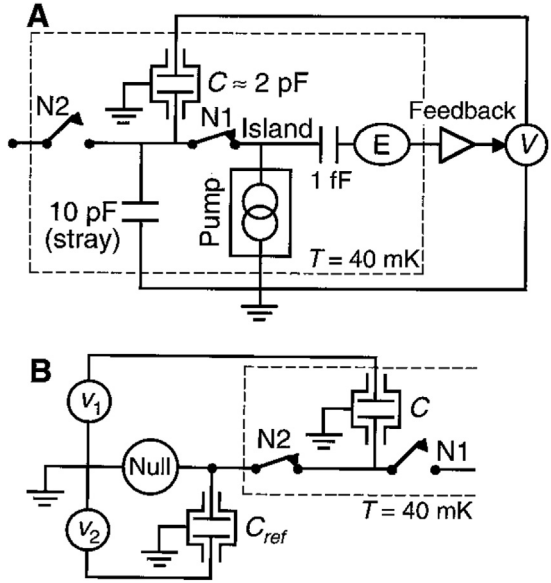


Figure 5.21 (a) Experimental circuit for capacitor charging using single-electron pump. (b) Experimental circuit for comparing the cryogenic capacitor with a room-temperature reference capacitor. *Reproduced with permission from Ref. [39].*

$$\frac{C_0}{C} = 1 + \varepsilon_J + \varepsilon_S = 1 + (-0.10 \pm 0.92) \times 10^{-6}. \tag{5.7}$$

They argued that ε_J should be sufficiently small compared to the uncertainty in their results, and therefore concluded that this was a test of the single-electron pump with $\varepsilon_S < 0.92 \times 10^{-6}$.[4]

2.4 Current quantisation accuracy of tunable-barrier pumps

As mentioned earlier, multi-junction single-electron pumps cannot be operated accurately at frequencies much larger than a few MHz, and hence the current generated is too small to be used directly for a quantum metrological triangle experiment as it was originally envisaged. Surface acoustic wave devices were promising for generating a large enough current, but their current quantisation has not been good enough to date (the best reported value

[4] Later, a team at Physikalisch-Technische Bundesanstalt (PTB) found that, in attempting to perform the quantum metrological triangle using a similar method, the extrapolation of the capacitance bridge data taken in the kHz frequency range to the dc regime could not be performed with the accuracy of 1 part in 10^6 or better [43].

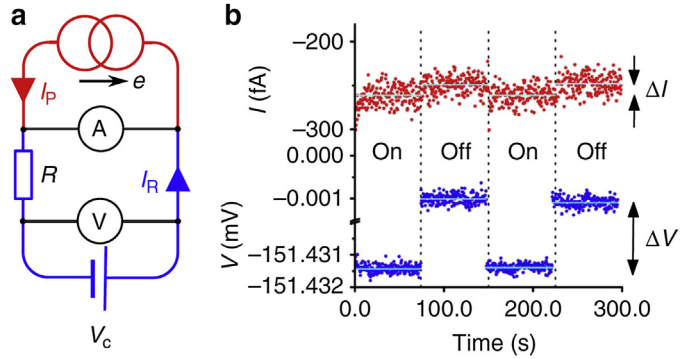


Figure 5.22 (a) Experimental circuit for traceable current measurements using a standard resistor (R), a voltmeter (V) and a transimpedance amplifier (A). (b) Current readings (*I*) of the transimpedance amplifier and voltage readings (*V*) of the voltmeter during on–off cycles. *Reproduced with permission from Ref. [19].*

is at a level of 1 part in 10^4 [44]). A tunable-barrier quantum dot pump also showed a similar accuracy (at a level of 1 part in 10^4) at ~500 MHz when its operation was first demonstrated [16]. Later, as mentioned in the Section 'Tunable barrier quantum dot pump', it was found that the use of a large magnetic field and tailored pump-driving waveform can considerably improve the accuracy of current quantisation in GaAs devices. These pumps can be operated at frequencies around 1 GHz, producing a current level of order of 100 pA. At the same time, the techniques of measuring such a small current have been improved over time, enabling high-precision measurements of electron pump current.

The most common method for testing the accuracy of tunable-barrier electron pumps is to perform high-precision measurements of their current, traceable to the Josephson voltage standard and quantum Hall resistance standard. An underlying assumption is that the uncertainty of the ampere realised by the voltage and resistance primary standards is much smaller than the uncertainty in the accuracy of single-electron pumps tested.[5]

Fig. 5.22a shows an experimental circuit adopted by Giblin et al. in 2012 [19]. An electron pump (shown as a current source at the top) generates a

[5] It should be noted that this experiment does not test the closure of the quantum metrological triangle, as Δ_{QMT} in Eq. (5.6) is dominated by ε_s under this assumption, and it becomes larger than σ_{QMT} as the precision of the experiments is made higher. In order to test the closure of the quantum metrological triangle, the current quantisation accuracy of the electron pump must be independently verified.

current I_P. This current is compared with a reference current I_R generated by a voltage source V_c and a standard resistor R with a nominal value of 1 GΩ. The difference ΔI of these currents is measured by a transimpedance amplifier. The voltage drop ΔV across the standard resistor and transimpedance amplifier in series is measured by a voltmeter. The actual value of the resistor was determined with an uncertainty of 0.8 parts in 10^6 traceable to a quantum Hall primary resistance standard. The voltmeter was calibrated traceably to a Josephson primary voltage standard with an uncertainty of 0.3 parts in 10^6. In order to remove instrumental drift from the measurements (e.g., offsets in the voltmeter and transimpedance amplifier), both the electron pump current (I_P) and reference current (I_R) were switched on and off repeatedly as shown in Fig. 5.22b. Both ΔI and ΔV were measured as the difference between the averaged values over on and off periods. The on–off cycles were repeated 96 times in order to reduce the total uncertainty of the measurements to 1.2 parts in 10^6. The pump current was calculated as follows:

$$I_P = \frac{\Delta V}{R} + \Delta I + \frac{r}{R}\Delta I, \tag{5.8}$$

where $r \approx 10$ kΩ was the input impedance of the transimpedance amplifier. Since $\frac{r}{R} \leq 10^{-4}$ and I_R was chosen so that $\frac{|\Delta I|}{I_P} < 10^{-4}$, the contribution by the third term on the right-hand side of the equation became negligible. The results when the pump was operated at 945 MHz are shown as red (dark grey in print version) data points in Fig. 5.23. These data points form a flat plateau at a level of ~1 part in 10^6, and their averaged value (shown as the

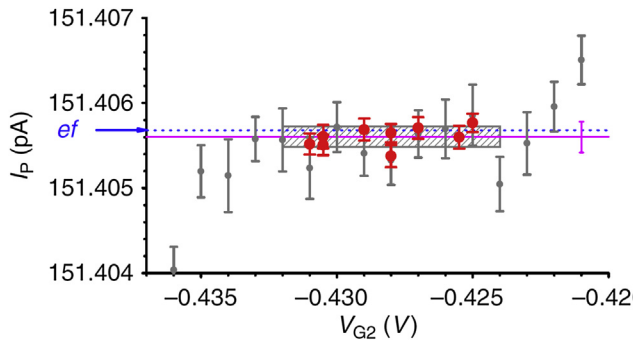


Figure 5.23 High-precision measurements of a current quantisation plateau by a tunable-barrier quantum dot pump. Red (dark grey in print version) data points were taken with 96 on–off cycles. Grey data points were taken with lower precision with 17 cycles. *Reproduced with permission from Ref. [19].*

magenta horizontal line) agrees with *ef* within the total uncertainty of 1.2 parts in 10^6.

The work mentioned above demonstrated that tunable-barrier quantum dot pumps can match the accuracy level of multi-junction pumps, but with a higher operation frequency so that the generated current can be directly measured. Since then, further progress has been made, and with the help of improvements in precision current measurements [45], a tunable-barrier pump has demonstrated a current quantisation accuracy at 0.16 parts in 10^6 [46]. The two examples of accurate tunable-barrier pumps mentioned here were made on GaAs/AlGaAs heterostructure systems, but comparable results have been achieved with Si devices [24,47]. This gives a scope for a universality test to check whether the current quantisation depends on material systems.

Another important development that is anticipated is a capability to detect pump errors. In order to test the quantum metrological triangle, the accuracy of single-electron transfer must be confirmed independently from the Josephson voltage and quantum Hall resistance standards. Error detection methods such as the one shown in Fig. 5.24 are being developed. In this scheme, multiple pumps (three in this example, denoted as P1, P2 and P3) are connected in series with islands (node 1 and node 2) in between. The charge on the nodes is monitored by sensitive charge detectors (D1 and D2). When all three pumps are operated simultaneously and correctly transfer one electron per cycle, the charge on the nodes does not change. But if one of the pumps, say P1, fails to pump an electron, then the number of electrons on the node 1 (n_1) decreases by 1, whilst that on the node 2 (n_2)

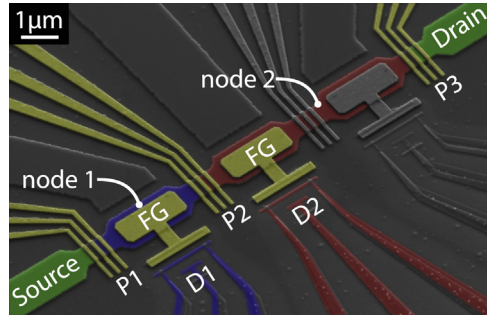


Figure 5.24 Scanning electron micrograph of pump-error detection device with multiple tunable-barrier pumps (P1, P2 and P3) in series. Single-electron transistor charge detectors (D1 and D2) monitor a variation of the charges in the nodes caused by pump errors. *Reproduced with permission from Ref. [48].*

does not change. If two pumps, P1 and P2 fail to pump an electron, then n_1 does not change, but n_2 decreases by 1. If only P2 fails to pump an electron, then n_1 increases by 1, whilst n_2 decreases by 1, and so on. This way, one can track the error of individual pumps. We cannot distinguish between the case with no pump error and the case with all three pumps having the same error (either failed to pump an electron, or pump the same excess number of electrons). However, if the error rate of each individual pump is sufficiently small, then the probability of having the same error for all three pumps becomes small enough to be negligible. By tracking the pump errors, corrections can be applied to the generated current. Fricke et al. argue that with 50 kHz charge detector bandwidth, 1 GHz pump frequency and the single-pump error rate of 1×10^{-6} or smaller, a relative uncertainty less than 1 part in 10^8 should be achievable [48].

3. Electron quantum optics

Quantum optics allows for exploitation of the quantum information carried by photons by coherently manipulating them or creating entanglement, etc. Experimental setups may consist of mirrors, beam splitters, detectors, polarisers and single-photon sources. Analogous experiments can be performed for other types of quantum particles, such as atoms, ions, as well as electrons. The initial concept of electron quantum optics in solid-state devices was developed using quasi-one-dimensional conductors, such as quantum point contacts and quantum Hall edge states, as the sources of electrons [49–52]. Applying a bias V_B to these conductors injects electrons at a rate of eV_B/h. These are continuous sources, and the timing of electron injection cannot be controlled. As single-electron sources have become available, electrons can now be emitted on demand as wave packets, with very little overlap of wavefunctions between the electrons in consecutive cycles. This gives greater control in manipulating electrons, for example, by controlling the difference in arrival times between the electrons emitted from different sources. Also, the efficiency of single-electron emission is generally good. In some systems, as described in the previous section, the probability of emission error can be smaller than 1 part in 10^6. These advantages provide scope for constructing compact electronic quantum optics system on chip. In the following, we introduce several important developments in electron quantum optics.

3.1 Partition noise measurements in Hanbury Brown and Twiss geometry

The current produced by an accurately quantised single-electron source is noiseless. There is no random fluctuation in the number of electrons emitted that can lead to noise. Splitting this current into two paths by a beam splitter introduces noise, as the partitioning is a stochastic process caused by the random nature of electron tunnelling through a barrier with a finite transmission T. Furthermore, the noise in the two split currents should be perfectly anti-correlated, because if an electron transmits through the barrier, there is no electron going into the reflected path, and vice versa. This test can be made by measuring the partition noise in a Hanbury Brown and Twiss geometry.

In optics, the Hanbury Brown and Twiss effect is used to test light sources by measuring the correlation signal between two detectors placed at the end of the paths after the beam from the source is split. An example of applications is a test for a single-photon source. If the source is truly emitting one photon at a time, and if the lengths of beam paths between the beam splitter and detectors are the same, then the probability of detecting photons simultaneously in both detectors should be zero. In the electronic version of Hanbury Brown and Twiss experiments, the detectors are replaced by noise measurements. Fig. 5.25 shows an experimental setup performed using a mesoscopic capacitor as a single-electron/hole source [53]. The source emits an electron (filled red [grey in print version] circle) and hole (open red [light grey in print version] circle) alternately into an edge state (marked by ①). This flow of particles meets a partition barrier formed by the quantum point contact and is split into the reflected (③) and transmitted (④) paths. If the transmission of the barrier is T, and if the effect of thermal excitation can be ignored (i.e., in the zero temperature limit), then the low-frequency correlations between the current noise in the outputs ③ and ④ can be written as[6]

$$S_{3,4} = -4e^2 fT(1-T), \tag{5.9}$$

where f is the frequency of the source. Fig. 5.26 shows the measurement of $S_{3,4}$ as a function of T, confirming the $T(1-T)$ dependence with the maximum noise at $T = 0.5$. Here, the full visibility $S_{3,4} = -e^2 f$ was not achieved as the thermal excitations of the Fermi sea (i.e., randomly excited

[6] In a system where only one type of particle (electron or hole) is emitted per cycle, this equation becomes $S_{3,4} = -2e^2 fT(1-T)$.

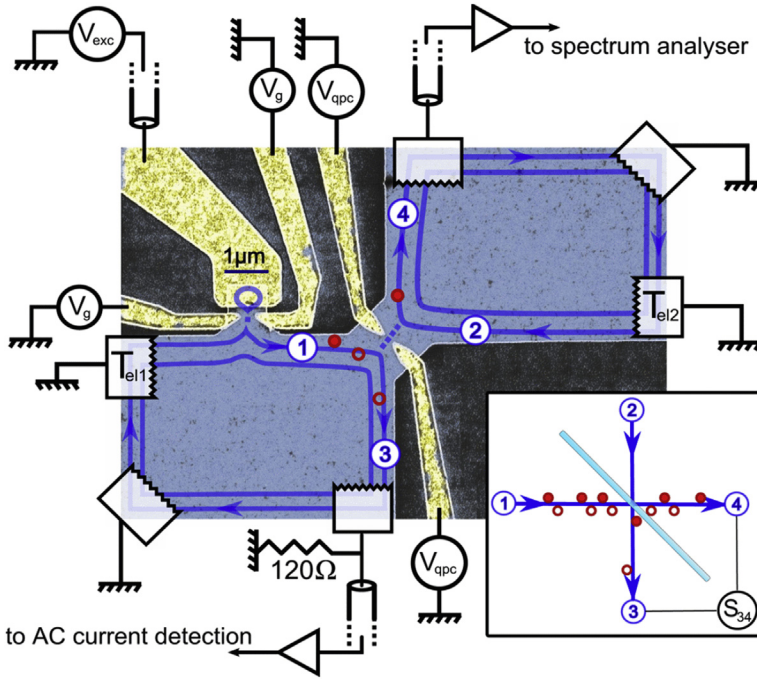


Figure 5.25 Experimental configuration of partition noise measurements using a mesoscopic capacitor. Inset shows a simplified schematic. *Reproduced with permission from Ref. [53].*

electrons) reduce the degree of noise correlation.[7] Similar partition noise experiments have been performed for levitons generated by Lorentzian pulses [27] and hot electrons emitted by a tunable-barrier pump [54]. These partition noise measurements play a role analogous to photon detectors in optics experiments, and are used as a fundamental tool in electron quantum optics experiments.

3.2 Indistinguishability test by Hong–Ou–Mandel effect

When two identical particles from different sources arrive at a beam splitter at the same time, two cases of outcome are expected depending on the type of the particles, as illustrated in Fig. 5.27. If they are bosons, they take the same path (bunching) and the coincidence counts $\langle N_3 N_4 \rangle$ (simultaneous detection by the detectors 3 and 4) become zero. The fluctuations of the

[7] The electron temperature of the Fermi sea was estimated to be 150 mK in their experiment.

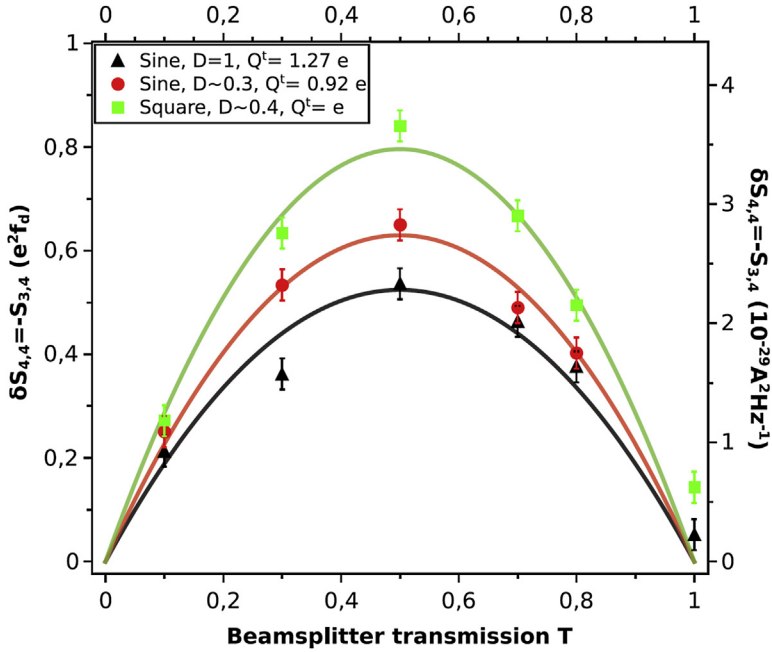


Figure 5.26 Cross-correlation $S_{3,4}$ (which is equivalent to the inverse of the excess auto-correlations $\delta S_{4,4}$) measured as a function of the transmission T of the beam splitter. Plots show three different source conditions for particle emission (sine-or square-wave drive and transmission D through the barrier at the source). *Reproduced with permission from Ref. [53].*

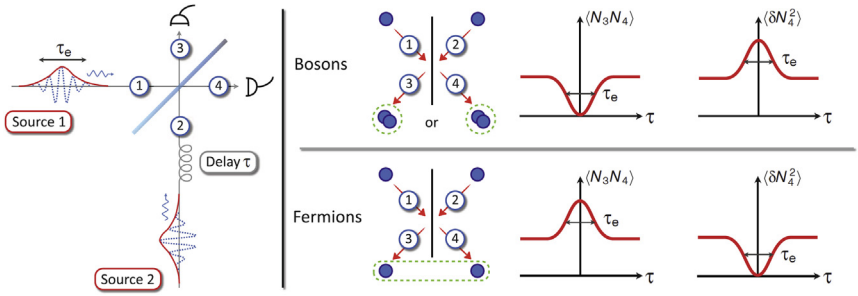


Figure 5.27 Hong–Ou–Mandel interferometry. Left: the geometry of the interferometer. Right: Cases for bosons showing bunching and fermions showing anti-bunching. *Reproduced with permission from Ref. [55].*

number of particles detected in each detector ($\langle \delta N_3^2 \rangle$ or $\langle \delta N_4^2 \rangle$) increase. On the other hand, in the case of fermions, the particles take different paths (anti-bunching) and the coincidence counts increase, whilst the fluctuations in each detector become zero. In a general case for fermions with the beam splitter transmission $T = \frac{1}{2}$, if each source emits a wave packet $|\phi_i\rangle$ $(i = 1, 2)$, the probability for the two particles to take different paths can be written as $P(1,1) = \frac{1}{2}[1 + |\langle \phi_1 | \phi_2 \rangle|^2]$. The probability for the two particles to take a same path is $P(0,2) + P(2,0) = \frac{1}{2}[1 - |\langle \phi_1 | \phi_2 \rangle^2|]$.[8] This two-particle interference, so-called Hong–Ou–Mandel effect, can be used to test the indistinguishability of the particles, to demonstrate coherence, and to extract the information about wave packet size.

An electronic Hong–Ou–Mandel interferometry [56–58] has been realised using mesoscopic capacitors as shown in Fig. 5.28 [55]. Two sources (1 and 2) inject an electron and a hole, in turn, into a beam splitter. The

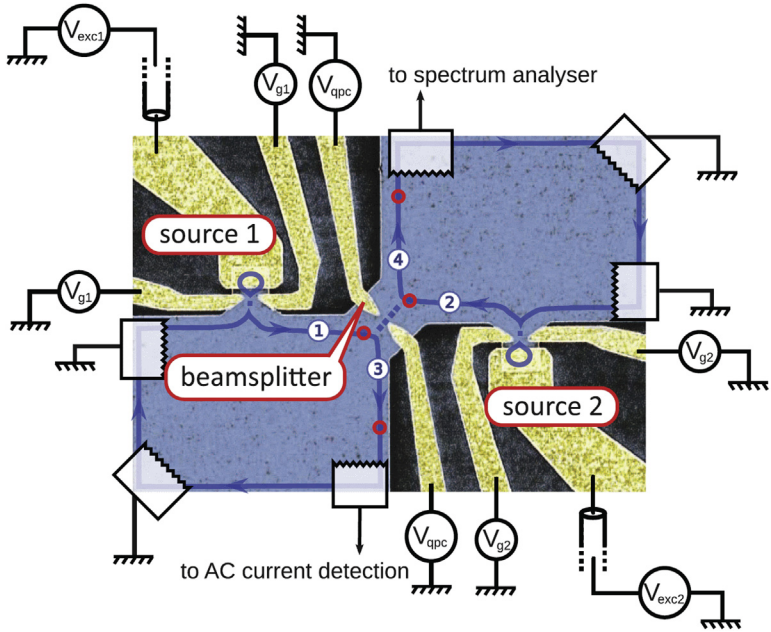


Figure 5.28 Experimental setup for a Hong–Ou–Mandel interferometry using two mesoscopic capacitors as sources. *Reproduced with permission from Ref. [55].*

[8] In the case of bosons, the sign before $|\langle \phi_1 | \phi_2 \rangle|^2$ in these equations is opposite.

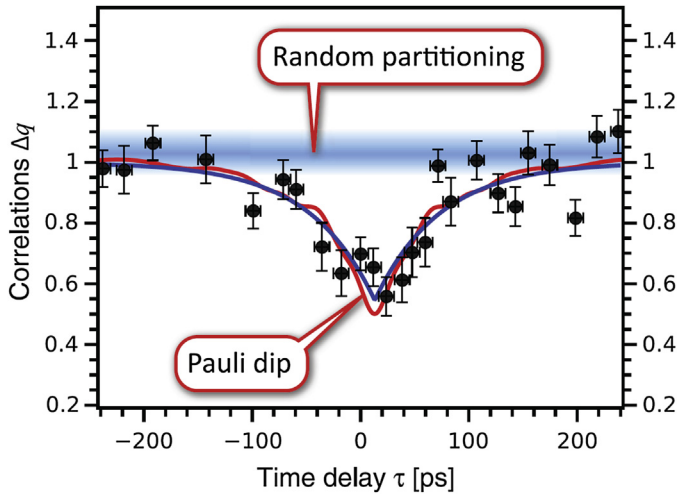


Figure 5.29 Normalised excess noise Δq plotted as a function of time delay τ. *Reproduced with permission from Ref. [55].*

relative arrival time of particles at the detector can be controlled by varying the phase difference between the ac driving signals (V_{exc1} and V_{exc2}) on the sources. The normalised excess noise $\Delta q(=\frac{S_{44}}{e^2 f}-1$, where S_{44} is the noise measured at contact 4) is plotted as a function of time delay τ in Fig. 5.29. This shows a dip (Pauli dip) near $\tau = 0$ due to the anti-bunching effect. Since $\langle \delta N_3^2 \rangle = \langle \delta N_4^2 \rangle = \frac{1}{2}[1 - |\langle \phi_1 | \phi_2 \rangle^2|]$, the width of the dip reflects the width of the electron/hole wave packets. A Hong–Ou–Mandel interferometry was also performed on levitons and the expected wave packet shape in the time domain was observed [27].

3.3 Quantum tomography

We have seen in the previous two sections how partially transmitting gates combined with correlated noise measurements can yield information about the quantum states of electrons emitted from single-electron sources. However, these do not provide full information about the electron states. For example, it does not provide information about how the wavefunction evolves over time as the electron travels through the conductor. Here, we discuss an experimental technique so-called 'quantum tomography', which allows us to obtain the full information on electron states.

In order to visualise the full quantum states, the so-called 'Wigner function' (or 'Wigner quasi-probability distribution') is frequently used. The Wigner function is defined (in the position x and momentum p phase space) as

$$W(x,p)=\frac{1}{h}\int_{-\infty}^{\infty}\psi^{*}\left(x+\frac{x'}{2}\right)\psi\left(x-\frac{x'}{2}\right)e^{\frac{ipx'}{h}}dx'. \tag{5.10}$$

It can be also plotted in the energy E and time t phase space as

$$W(E,t)=\frac{1}{h}\int_{-\infty}^{\infty}\psi^{*}\left(E-\frac{\varepsilon}{2}\right)\psi\left(E+\frac{\varepsilon}{2}\right)e^{\frac{it\varepsilon}{h}}d\varepsilon. \tag{5.11}$$

The main goal of the electron quantum tomography presented here is to map out the Wigner function on the energy–time phase space.

The method of quantum tomography is somewhat similar to the tomographic imaging used in medical applications. Projected distributions (marginal distributions) on rotating axis are taken at many angles, and the original distribution is reconstructed from the data (equivalent to looking at an object from different angles and working out its shape). In a photonic system, a popular method is homodyne tomography [59], which mixes the incoming test photon beam with a strong coherent photon field, or a laser beam (a local oscillator) at a beam splitter [illustrated in Fig. 5.30b]. The laser beam is used to trigger the generation of test photons, so there is a fixed phase relation between them. Because of the high intensity of the laser, it provides a rigid phase reference. At the beam splitter, the two beams interfere and are redirected to two photon detectors depending on the relative phase difference of the local oscillator and test photon fields. The difference in the number of photons detected by the detectors is proportional to the test photon field, and hence the quadrature amplitude (a position on the rotating axis) can be determined with an appropriate scaling factor [59]. Measurements are repeated a number of times to accumulate the statistics of quadrature amplitude, from which the probability distribution (marginal distribution) is produced. Changing the relative phase of the two beams changes the angle of projected axis on the momentum position phase space. Marginal distributions are collected at many different angles, and the Wigner function is reconstructed by a process called the inverse Radon transformation.

For an electronic system, the fermionic statistics forbids creation of a strong coherent field, and therefore the method for the photonic homodyne tomography cannot be directly applied. Instead, Grenier et al. proposed a

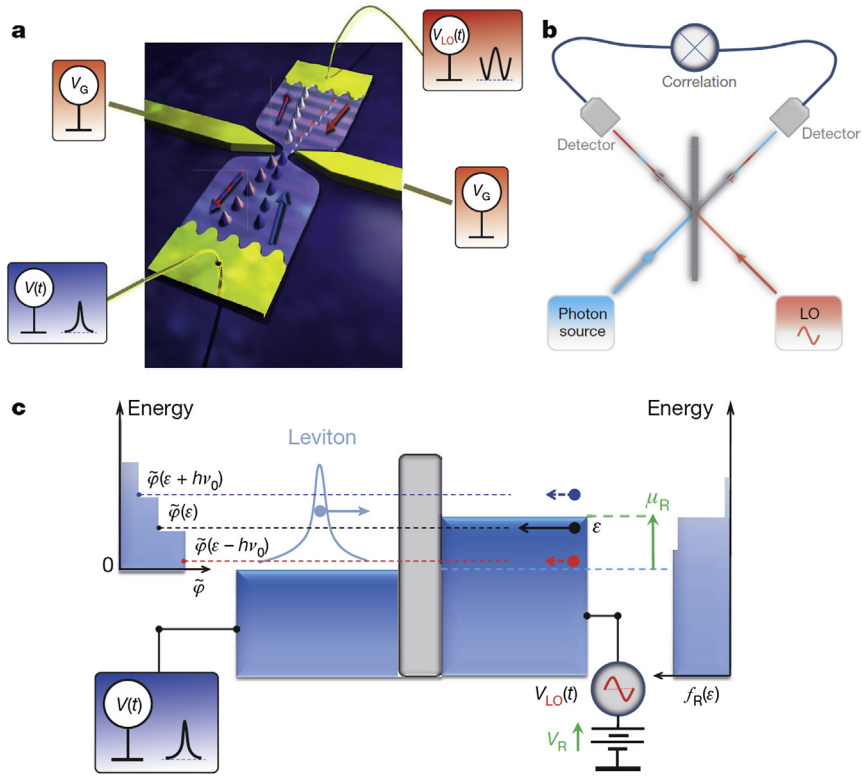


Figure 5.30 (a) Illustration for experimental setup of quantum tomography on levitons. (b) Setup for photonic quantum tomography using a local oscillator (LO). (c) Experimental configuration in energy space. The leviton is injected into the beam splitter from the left contact. The Fermi sea on the right is used as a local oscillator. *Reproduced with permission from Ref. [60].*

method to interrogate the components of the density matrix using dc and ac biases of a Fermi sea [61]. The Wigner function is then constructed from the density matrix. Jullien et al. adopted this technique and performed quantum tomography on levitons as shown in Fig. 5.30 [60]. A leviton is injected into a tunnel barrier from the left contact in Fig. 5.30c. In order to extract the diagonal components, the Fermi sea on the right of the barrier is biased by a dc voltage V_R. As we have seen earlier in the Section 'Indistinguishability test by Hong–Ou–Mandel effect', the presence of the Fermi sea up to the energy eV_R suppresses the partition noise of the incoming leviton in the energy range $[0, eV_R]$. The total partition noise is reduced by the factor $1 - \int_0^{eV_R} \tilde{\varphi}^*(\varepsilon)\tilde{\varphi}(\varepsilon)d\varepsilon$, where $\tilde{\varphi}(\varepsilon)$ is the Fourier transform of the leviton

wavefunction $\varphi(t)$ in the time domain. The diagonal part of the energy density matrix $\left|\widetilde{\varphi}(\varepsilon)\right|^2$ is obtained by differentiating the noise with respect to V_R.

In order to obtain the off-diagonal components, an ac voltage $V_{LO}(t) = \eta_{LO}(h\nu/e)\cos(2\pi\nu(t-\tau))$ is applied with $\eta_{LO} \ll 1$, frequency ν and time delay τ. This ac voltage plays an analogous, but slightly different, role to the local oscillator in photonic experiments. Due to the weak ac excitation, electrons incoming to the barrier from the right reservoir are put into a superposition of the states at their initial energy ε and energies shifted by $\pm h\nu$, i.e., $\widetilde{\varphi}_R(\varepsilon) = \widetilde{\varphi}(\varepsilon) + \frac{1}{2}\eta_{LO}e^{i2\pi\nu\tau}\widetilde{\varphi}(\varepsilon + h\nu) - \frac{1}{2}\eta_{LO}e^{-i2\pi\nu\tau}\widetilde{\varphi}(\varepsilon - h\nu)$.[9] This superposition is used to probe the off-diagonal components of the density matrix. The derivative of the partition noise with respect to V_R can be written as follows:

$$\frac{d\langle\Delta Q^2\rangle}{dV_R} \propto \eta_{LO}\cos(2\pi\nu\tau)\left[\widetilde{\varphi}^*(eV_R)\widetilde{\varphi}(eV_R + h\nu) - \widetilde{\varphi}^*(eV_R)\widetilde{\varphi}(eV_R - h\nu)\right]. \quad (5.12)$$

Since the leviton excitation is only in the energy domain above the Fermi energy, the equation above provides $\widetilde{\varphi}^*(eV_R)\widetilde{\varphi}(eV_R + h\nu)$ for the range $0 \le eV_R < h\nu$. When eV_R is increased to the next energy range $h\nu \le eV_R < 2h\nu$, Eq. (5.12) gives $\widetilde{\varphi}^*(eV_R)\widetilde{\varphi}(eV_R + h\nu) - \widetilde{\varphi}^*(eV_R)\widetilde{\varphi}(eV_R - h\nu)$. The first term can be determined as the second term is the same as the term $\widetilde{\varphi}^*(eV_R)\widetilde{\varphi}(eV_R + h\nu)$ in the previous energy range. Higher energy ranges are probed with similar steps. The components at higher harmonics are deduced by increasing the frequency of the local oscillator. In the experiment by Jullien et al., levitons were periodically generated at a frequency of $\nu_0 = 6$ GHz. The frequency of the local oscillator was chosen accordingly as $\nu = k\nu_0$, where $k = 0,1,2$. Higher harmonics were not accessible due to the small signal size. An available part of the Wigner function $W(E,t)$ in the energy-time domain was reconstructed from the energy density matrix, which is plotted along with a theoretical calculation in Fig. 5.31. Quantum tomography was also performed on electron/hole excitations generated by a mesoscopic capacitor [62], and on hot electrons from a tunable-barrier pump, using a time-dependent barrier instead of an oscillating Fermi sea [63,64].

[9] Here, we use an approximation for the probability of keeping the initial energy to be $1 - \eta_{LO}^2 \approx 1$.

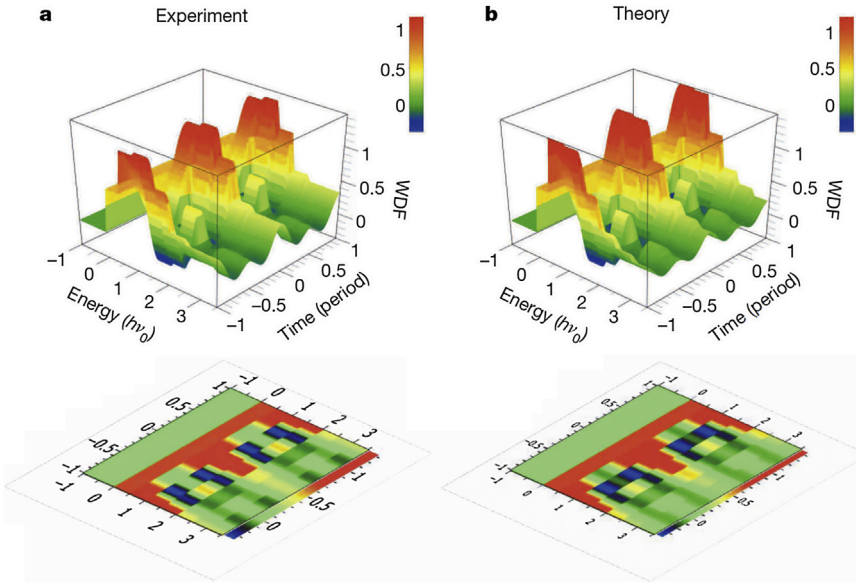


Figure 5.31 (a) Wigner function of the levitons reconstructed from the experiment. (b) Theoretical calculations of the leviton Wigner function. *Reproduced with permission from Ref. [60].*

3.4 De-coherence and relaxation in wave packet transport

In order to exploit the coherent transport of single-electron wave packets, their de-coherence and energy relaxation mechanisms need to be understood. As the electrons travel through a waveguide, they interact with their environment such as a surrounding Fermi sea and lattice vibrations (phonons), leading to a loss of coherence. Here, we introduce a few examples of work that studied such mechanisms.

Freulon et al. investigated charge fractionalisation using the Hong–Ou–Mandel interferometry when electrons (and holes) were injected into a quantum Hall edge state [65]. Capacitive interactions with the inner edge state, into which carriers were not injected, break up the wave packet into charge mode and neutral mode, travelling at different velocities v_+ and v_-, respectively (see Fig. 5.32). They observed broadening of the Pauli dip after wave packets travelled for 3 μm, estimating the time separation between two modes as 70 ps. A similar charge fractionalisation was reported in the time-of-flight measurements of plasmonic excitations through a gate-defined quasi-one-dimensional channel, but after travelling a distance of over 20 μm [66].

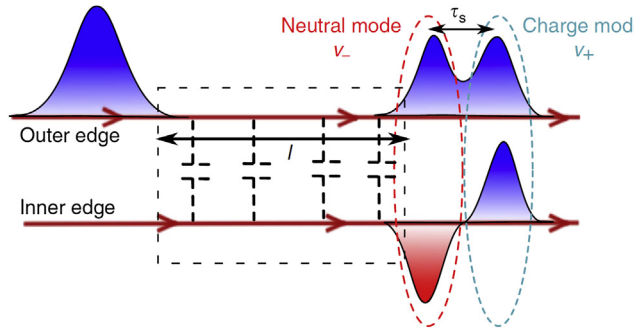


Figure 5.32 Charge fractionalisation into the charge and neutral modes due to capacitive interactions between edge states. *Reproduced with permission from Ref. [65].*

Whilst electron transport near the Fermi surface is protected from energy relaxation due to the lack of available states to which the electrons can relax, hot electrons emitted from a tunable-barrier pump can suffer from energy relaxation due to a large phase space available. Fletcher et al. observed electron relaxation by the emission of longitudinal-optical (LO) phonons [67]. In GaAs, LO phonon dispersion is flat at $\hbar\omega_{\mathrm{LO}} = 36$ meV around the Γ point. The energy spectrum of hot electrons was measured by a detector potential barrier placed 3 μm away from the source (Fig. 5.33). At a magnetic field of $B = 12$ T, most of the electrons maintain their original emission energy when arriving at the detector (Fig. 5.33d), but at $B = 6$ T, they arrive after losing $\hbar\omega_{\mathrm{LO}}$ or $2\hbar\omega_{\mathrm{LO}}$ by emitting one or two LO phonons (Fig. 5.33f). The phonon emission probability of an electron wave packet is strongly dependent on the overlap of the initial (unperturbed) and final (relaxed) states [68]. Suppression of phonon emission can be achieved by magnetic confinement, which shrinks the extent of the electron wavefunction or by softening the edge potential, which increases the distance between the initial and final states. An LO phonon scattering length over 100 μm has been reported [69].

One key component that is widely available in the field of photonic quantum optics, but has not been yet realised with single-electron sources is the Mach–Zehnder interferometer. Haack et al. proposed to use such an interferometer to measure the coherence length, a distance over which single-particle interference can be observed (Fig. 5.34) [70]. This requires not only coherence whilst electrons travel the interference path, but also an ability to control the overlap of the split wave packets at the second beam splitter by controlling either the path lengths or the electron velocity.

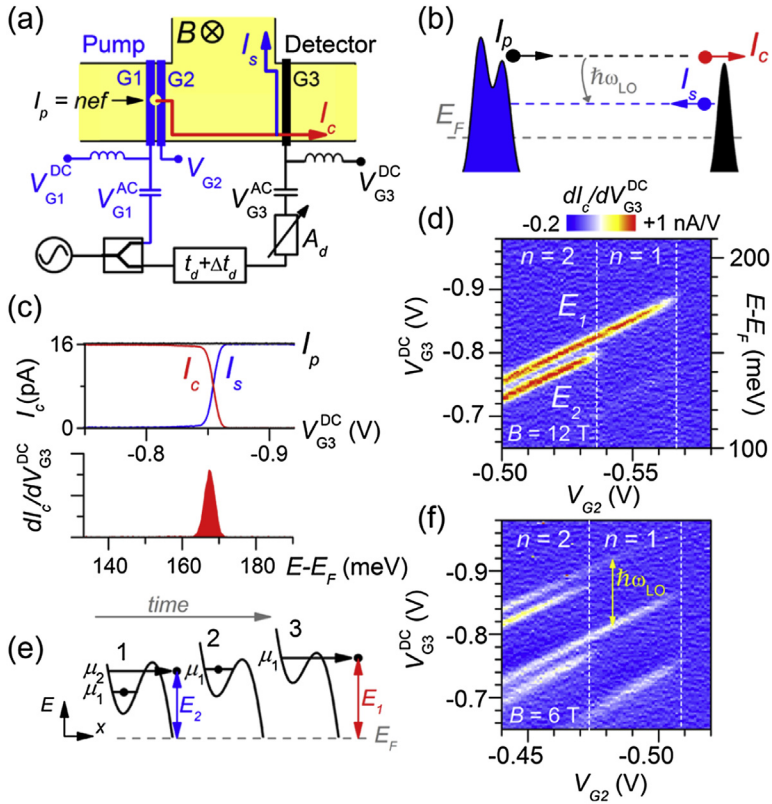


Figure 5.33 Energy spectroscopy of hot-electron wave packets. (a) Experimental setup. (b) Energy diagram. Electrons are emitted from a tunable-barrier pump on the left. The detector potential barrier on the right selects the transmission of electrons depending on their energy. (c) The threshold of transmitted (I_c) or reflected (I_s) currents indicates the electron energy. (d) Measured energy spectrum at $B = 12$ T. (e) Schematics of the dot potential describing the energy E_1 and E_2 of the last and second-last electrons emitted. (f) Measured energy spectrum at $B = 6$ T. *Reproduced with permission from Ref. [67].*

Such a single-electron interferometry would enable applications that exploit the coherent nature of electron wavepacket transport.

4. Summary

In this chapter, we examined the development of solid-state single-electron sources spanning the last 3 decades. We discussed seven different methods to transfer electrons one by one in a controlled manner. We then looked at the development of two types of application in which these

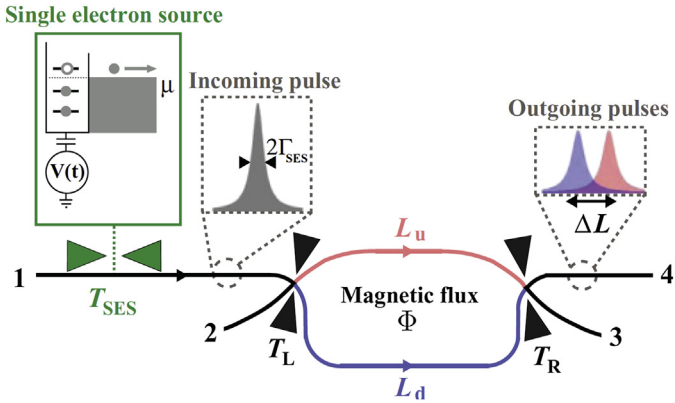


Figure 5.34 Electron wavepacket Mach–Zehnder interferometry proposed by Haack et al. The visibility is affected by the overlap of the two wavepackets at the second beam splitter. *Reproduced with permission from Ref. [70].*

devices are used: quantum current standard and electron quantum optics. With regard to the quantum current standard application, at the time of writing, accurate tunable-barrier quantum dot pumps that satisfy metrological needs (i.e., with an error rate less than $\sim$1 part in 10^8) have not yet been reported. However, considering that this field has been rapidly developed in the last decade, the realisation of single-electron-based primary ampere standards may not be far in the future. As for the electron quantum optics, this field is still at the stage of developing methods to characterise excitations and interactions using schemes analogous to photonic experiments. As the technology matures, we may see cryogenic nanoelectronic devices in which these methods are applied to control and process signals at ultrafast speed (on picosecond timescales) [71], or even to exploit the quantum information encoded onto electrons [72].

It should be noted that there are many other interesting and important developments that could not be included here. Additionally, readers may find that the descriptions of work presented here are not detailed enough. For those who would like to know more about the development of single-electron sources, the following review articles are recommended [9,36,37,72–77].

Acknowledgements

I would like to thank my colleagues, Stephen Giblin, Patrick See, Jonathan Fletcher, Alessandro Rossi, Shak Fernandes and Christian Glattli for providing valuable comments on the manuscript.

References

[1] BIPM, the 9th edition of the SI Brochure. 2019. https://www.bipm.org/utils/common/pdf/si-brochure/SI-Brochure-9.pdf.

[2] Kouwenhoven LP, Johnson AT, van der Vaart NC, Harmans CJPM, Foxon CT. Quantized current in a quantum-dot turnstile using oscillating tunnel barriers. Phys Rev Lett 1991;67:1626.

[3] Pothier H, Lafarge P, Urbina C, Esteve D, Devoret MH. Single-electron pump based on charging effects. Europhys Lett 1992;17:249.

[4] van der Wiel WG, De Franceschi S, Elzerman JM, Fujisawa T, Tarucha S, Kouwenhoven LP. Electron transport through double quantum dots. Rev Mod Phys 2002;75:1.

[5] Connolly MR, Chiu KL, Giblin SP, Kataoka M, Fletcher JD, Chua C, Griffiths JP, Jones GAC, Fal'ko VI, Smith CG, Janssen TJBM. Gigahertz quantized charge pumping in graphene quantum dots. Nat Nanotechnol 2013;8:417.

[6] Roche B, Riwar R-P, Voisin B, Dupont-Ferrier E, Wacquez R, Vinet M, Sanquer M, Splettstoesser J, Jehl X. A two-atom electron pump. Nat Commun 2013;4:1581.

[7] Pekola JP, Vartiainen JJ, Möttönen M, Saira O-P, Meschke M, Averin DV. Hybrid single-electron transistor as a source of quantized electric current. Nat Phys 2008;4:120.

[8] Kemppinen A, Kafanov S, Pashkin YA, Tsai JS, Averin DV, Pekola JP. Experimental investigation of hybrid single-electron turnstiles with high charging energy. Appl Phys Lett 2009;94:172108.

[9] Pekola JP, Saira O-P, Maisi VF, Kemppinen A, Möttönen M, Pashkin YA, Averin DV. Single-electron current sources: toward a redefined definition of the ampere. Rev Mod Phys 2013;85:1421.

[10] Averin DV, Pekola J. Nonadiabatic charge pumping in a hybrid single-electron transistor. Phys Rev Lett 2008;101:066801.

[11] Shilton JM, Talyanskii VI, Pepper M, Ritchie DA, Frost JEF, Ford CJB, Smith CG, Jones GAC. High-frequency single-electron transport in a quasi-one-dimensional GaAs channel induced by surface acoustic waves. J Phys: Condens Matter 1996;8:L531.

[12] Ahlers FJ, Kieler OFO, Sağol BE, Pierz K, Siegner U. Quantized acoustoelectric single electron transport close to equilibrium. J Appl Phys 2006;100:093702.

[13] Cunningham J, Talyanskii VI, Shilton JM, Pepper M, Kristensen A, Lindelof PE. Single-electron acoustic charge transport on shallow-etched channels in a perpendicular magnetic field. Phys Rev B 2000;62:1564.

[14] Hermelin S, Takada S, Yamamoto M, Tarucha S, Wieck AD, Saminadayar L, Bäuerle C, Meunier T. Electrons surfing on a sound wave as a platform for quantum optics with flying electrons. Nature 2011;477:435.

[15] McNeil RPG, Kataoka M, Ford CJB, Barnes CHW, Anderson D, Jones GAC, Farrer I, Ritchie DA. On-demand single-electron transfer between distant quantum dots. Nature 2011;477:439.

[16] Blumenthal MD, Kaestner B, Li L, Giblin S, Janssen TJBM, Pepper M, Anderson D, Jones G, Ritchie DA. Gigahertz quantized charge pumping. Nat Phys 2007;3:343.

[17] Kaestner B, Kashcheyevs V, Amakawa S, Blumenthal MD, Li L, Janssen TJBM, Hein G, Pierz K, Weimann T, Siegner U, Schumacher HW. Single-parameter nonadiabatic quantized charge pumping. Phys Rev B 2008;77:153301.

[18] Fujiwara A, Nishiguchi K, Ono Y. Nanoampere charge pump by single-electron ratchet using silicon nanowire metal-oxide- semiconductor field-effect transistor. Appl Phys Lett 2008;92:042102.

[19] Giblin SP, Kataoka M, Fletcher JD, See P, Janssen TJBM, Griffiths JP, Jones GAC, Farrer I, Ritchie DA. Towards a quantum representation of the ampere using single electron pumps. Nat Commun 2012;3:930.
[20] Kashcheyevs V, Kaestner B. Universal decay cascade model for dynamic quantum dot initialization. Phys Rev Lett 2010;104:186805.
[21] Fletcher JD, Kataoka M, Giblin SP, Park S, Sim H-S, See P, Ritchie DA, Griffiths JP, Jones GAC, Beere HE, Janssen TJBM. Stabilization of single-electron pumps by high magnetic fields. Phys Rev B 2012;86:155311.
[22] Wright SJ, Blumenthal MD, Gumbs G, Thorn AL, Pepper M, Janssen TJBM, Holmes SN, Anderson D, Jones GAC, Nicoll CA, Ritchie DA. Enhanced current quantization in high-frequency electron pumps in a perpendicular magnetic field. Phys Rev B 2008;78:233311.
[23] Kaestner B, Leicht C, Kashcheyevs V, Pierz K, Siegner U, Schumacher HW. Single-parameter quantized charge pumping in high magnetic fields. Appl Phys Lett 2009;94:012106.
[24] Zhao R, Rossi A, Giblin SP, Fletcher JD, Hudson FE, Möttönen M, Kataoka M, Dzurak AS. Thermal-error regime in high-accuracy gigahertz single-electron pumping. Phys Rev Appl. 2017;8:044021.
[25] Yamahata G, Giblin SP, Kataoka M, Karasawa T, Fujiwara A. High-accuracy current generation in the nanoampere regime from a silicon single-trap electron pump. Sci Rep 2017;7:45137.
[26] Fève G, Mahé A, Berroir J-M, Kontos T, Plaçais B, Glattli DC, Cavanna A, Etienne B, Jin Y. An on-demand coherent single-electron source. Science 2007;316:1169.
[27] Dubois J, Jullien T, Portier F, Roche P, Cavanna A, Jin Y, Wegscheider W, Roulleau P, Glattli DC. Minimal-excitation states for electron quantum optics using levitons. Nature 2013;502:659.
[28] Levitov LS, Lee H, Lesovik GB. Electron counting statistics and coherent states of electric current. J Math Phys 1996;37:4845.
[29] Ivanov DA, Lee HW, Levitov LS. Coherent states of alternating current. Phys Rev B 1997;56:6839.
[30] Keeling J, Klich I, Levitov LS. Minimal excitation states of electrons in one-dimensional wires. Phys Rev Lett 2006;97:116403.
[31] "Resolutions of the CGPM", the 26th meeting of the CGPM: 13-16 November 2018. https://www.bipm.org/utils/common/pdf/CGPM-2018/26th-CGPM-Resolutions.pdf.
[32] Mise en pratique for the definition of the ampere and other electric units in the SI. BIPM; 2019. Appendix 2 of the 9th edition of the SI Brochure, https://www.bipm.org/utils/en/pdf/si-mep/SI-App2-ampere.pdf.
[33] Brun-Picard J, Djordjevic S, Leprat D, Schopfer F, Poirier W. Practical quantum realization of the ampere from the elementary charge. Phys Rev X 2016;6:041051.
[34] Josephson BD. Possible new effects in superconductive tunnelling. Phys Lett 1962;1:251.
[35] Klitzing Kv, Dorda G, Pepper M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. Phys Rev Lett 1980;45:494.
[36] Keller MW. Current status of the quantum metrology triangle. Metrologia 2008;45:102.
[37] Scherer H, Camarota B. Quantum metrology triangle experiments: a status review. Meas Sci Technol 2012;23:124010.
[38] Mohr PJ, Newell DB, Taylor BN. CODATA recommended values of the fundamental physical constants: 2014. Rev Mod Phys 2016;88:035009.

[39] Keller MW, Eichenberger AL, Martinis JM, Zimmerman NM. A capacitance standard based on counting electrons. Science 1999;285:1706.
[40] Keller MW, Martinis JM, Zimmerman NM, Steinbach AH. Accuracy of electron counting using a 7-junction electron pump. Appl Phys Lett 1996;69:1804.
[41] Keller MW, Zimmerman NM, Eichenberger AL. Uncertainty budget for the NIST electron counting capacitance standard, ECCS-1. Metrologia 2007;44:505.
[42] Jackson JD. A curious and useful theorem in two-dimensional electrostatics. Am J Phys 1999;67:107.
[43] Scherer H, Schurr J, Ahlers FJ. Electron counting capacitance standard and quantum metrology triangle experiment at PTB. Metrologia 2017;54:322.
[44] Janssen J-T, Hartland A. Recent measurements of single electron transport of surface acoustic wave devices at the NPL. IEEE Trans Instrum Meas 2001;50:227.
[45] Drung D, Krause C, Becker U, Scherer H, Ahlers FJ. Ultrastable low-noise current amplifier: a novel device for measuring small electric currents with high accuracy. Rev Sci Instrum 2015;86:024703.
[46] Stein F, Scherer H, Gerster T, Behr R, Götz M, Pesel E, Leicht C, Ubbelohde N, Weimann T, Pierz K, Schumacher HW, Hohls F. Robustness of single-electron pumps at sub-ppm current accuracy level. Metrologia 2017;54:S1.
[47] Yamahata G, Giblin SP, Kataoka M, Karasawa T, Fujiwara A. Gigahertz single-electron pumping in silicon with an accuracy better than 9.2 parts in 10^7. Appl Phys Lett 2016;109:013101.
[48] Fricke L, Wulf M, Kaestner B, Hohls F, Mirovsky P, Mackrodt B, Dolata R, Weimann T, Pierz K, Siegner U, Schumacher HW. Self-referenced single-electron quantized current source. Phys Rev Lett 2014;112:226803.
[49] Henny M, Oberholzer S, Strunk C, Heinzel T, Ensslin K, Holland M, Schonenberger C. The fermionic Hanbury Brown and Twiss experiment. Science 1999;284:296.
[50] Oliver WD, Kim J, Liu RC, Yamamoto Y. Hanbury Brown and Twiss-type experiment with electrons. Science 1999;284:299.
[51] Ji Y, Chung Y, Sprinzak D, Heiblum M, Mahalu D, Shtrikman H. An electronic Mach-Zehnder interferometer. Nature 2003;422:415.
[52] Neder I, Ofek N, Chung Y, Heiblum M, Mahalu D, Umansky V. Interference between two indistinguishable electrons from independent sources. Nature 2007; 448:333.
[53] Bocquillon E, Parmentier FD, Grenier C, Berroir J-M, Degiovanni P, Glattli DC, Plaçais B, Cavanna A, Jin Y, Féve G. Electron quantum optics: partitioning electrons one by one. Phys Rev Lett 2012;108:196803.
[54] Ubbelohde N, Hohls F, Kashcheyevs V, Wagner T, Fricke L, Kästner B, Pierz K, Schumacher HW, Haug RJ. Partitioning of on-demand electron pairs. Nat Nanotechnol 2015;10:46.
[55] Bocquillon E, Freulon V, Berroir J-M, Degiovanni P, Plaçais B, Cavanna A, Jin Y, Féve G. Coherence and indistinguishability of single electrons emitted by independent sources. Science 2013;339:1054.
[56] Jonckheere T, Rech J, Wahl C, Martin T. Electron and hole Hong-Ou-Mandel interferometry. Phys Rev B 2012;86:125425.
[57] Ol'khovskaya S, Splettstoesser J, Moskalets M, Büttiker M. Shot noise of a mesoscopic two-particle collider. Phys Rev Lett 2008;101:166802.
[58] Féve G, Degiovanni P, Jolicoeur T. Quantum detection of electronic flying qubits in the integer quantum Hall regime. Phys Rev B 2008;77:035308.
[59] Smithey DT, Beck M, Raymer MG, Faridani A. Measurement of the wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum. Phys Rev Lett 1993;70:1244.

[60] Jullien T, Roulleau P, Roche B, Cavanna A, Jin Y, Glattli DC. Quantum tomography of an electron. Nature 2014;514:603.
[61] Grenier C, Hervé R, Bocquillon E, Parmentier FD, Plaçais B, Berroir JM, Fève G, Degiovanni P. Single-electron quantum tomography in quantum Hall edge channels. New J Phys 2011;13:093007.
[62] Bisognin R, Marguerite A, Roussel B, Kumar M, Cabart C, Chapdelaine C, Mohammad-Djafari A, Berroir J-M, Bocquillon E, Plaçais B, Cavanna A, Gennser U, Jin Y, Degiovanni P, Fève G. Quantum tomography of electrical currents. Nat Commun 2019;10:3379.
[63] Fletcher JD, Johnson N, Locane E, See P, Griffiths JP, Farrer I, Ritchie DA, Brouwer PW, Kashcheyevs V, Kataoka M. Continuous-variable tomography of solitary electrons. Nat Commun 2019;10:5298.
[64] Locane E, Brouwer PW, Kashcheyevs V. Time-energy filtering of single electrons in ballistic waveguides. New J Phys 2019;21:093042.
[65] Freulon V, Marguerite A, Berroir J-M, Plaçais B, Cavanna A, Jin Y, Fève G. Hong-Ou-Mandel experiment for temporal investigation of single-electron fractionalization. Nat Commun 2015;6:6854.
[66] Roussely G, Arrighi E, Georgiou G, Takada S, Urdampilleta M, Ludwig A, Wieck AD, Armagnat P, Kloss T, Waintal X, Meunier T, Bäuerle C. Unveiling the bosonic nature of an ultrashort few-electron pulse. Nat Commun 2018;9:2811.
[67] Fletcher JD, See P, Howe H, Pepper M, Giblin SP, Griffiths JP, Jones GAC, Farrer I, Ritchie DA, Janssen TJBM, Kataoka M. Clock-controlled emission of single-electron wave packets in a solid-state circuit. Phys Rev Lett 2013;111:216807.
[68] Emary C, Dyson A, Ryu S, Sim H-S, Kataoka M. Phonon emission and arrival times of electrons from a single-electron source. Phys Rev B 2016;93:035436.
[69] Johnson N, Emary C, Ryu S, Sim H-S, See P, Fletcher JD, Griffiths JP, Jones GAC, Farrer I, Ritchie DA, Pepper M, Janssen TJBM, Kataoka M. LO-phonon emission rate of hot electrons from an on-demand single-electron source in a GaAs/AlGaAs heterostructure. Phys Rev Lett 2018;121:137703.
[70] Haack G, Moskalets M, Splettstoesser J, Büttiker M. Coherence of single-electron sources from Mach-Zehnder interferometry. Phys Rev B 2011;84:081303(R).
[71] Johnson N, Fletcher JD, Humphreys DA, See P, Griffiths JP, Jones GAC, Farrer I, Ritchie DA, Pepper M, Janssen TJBM, Kataoka M. Ultrafast voltage sampling using single-electron wavepackets. Appl Phys Lett 2017;110:102105.
[72] Bäuerle C, Glattli DC, Meunier T, Portier F, Roche P, Roulleau P, Takada S, Waintal X. Coherent control of single electrons: a review of current progress. Rep Prog Phys 2018;81:056503.
[73] Grenier C, Hervè R, Fève G, Degiovanni P. Electron quantum optics in quantum Hall edge channels. Mod Phys Lett B 2011;25:1053.
[74] Kaestner B, Kashcheyevs V. Non-adiabatic quantized charge pumping with tunable-barrier quantum dots: a review of current progress. Rep Prog Phys 2015;78:103901.
[75] Kaneko N-H, Nakamura S, Okazaki Y. A review of the quantum current standard. Meas Sci Technol 2016;27:032001.
[76] Splettstoesser J, Haug RJ. Special issue: single-electron control in solid-state devices. Phys Status Solidi B 2017;254(3).
[77] Giblin SP, Fujiwara A, Yamahata G, Bae M-H, Kim N, Rossi A, Möttönen M, Kataoka M. Evidence for universality of tunable-barrier electron pumps. Metrologia 2019;56:044004.

CHAPTER SIX

Noise measurements in semiconductor nanodevices

Glattli D. Christian*
SPEC, CEA Saclay Nanoelectronics Group, Service de Physique de l'Etat Condense, Universite Paris-Saclay, CEA Saclay, Gif-sur-Yvette, France
*Corresponding author: christian.glattli@cea.fr

1. Introduction

This chapter is about noise measurements in quantum coherent semiconductor nanodevices. Quantum coherent conductors are characterised by their small size such that a carrier crossing the conductor preserves its phase coherence. This requires no inelastic energy exchange with photons, phonons, plasmons or other dynamical excitations and no phase averaging by random noise. Such a highly quantum regime is best obtained at a very low temperature which lowers the number of thermal excitations. The quantum coherent regime can be realised in metals or in doped semiconductors of sub-micron size. In this book, we concentrate on semiconductor nanodevices. Although metallic systems have played an important role in our understanding of electronic quantum noise, doped semiconductors and in particular 2D electron systems realised at the interface of modulation doped heterojunctions have been key tools to test quantum noise theories, namely by the realisation of Quantum Point Contacts (QPCs). Thus, we will restrict this review on shot noise in semiconductor nanodevices.

By quantum shot noise, we mean the noise resulting from the stochastic arrival of carriers on the contact of a conductor under the application of voltage to another contact. Each carrier generates a current pulse; the sum of these random pulses leads to macroscopic measurable current fluctuations. The shot noise is proportional to the mean current and proportional to the carrier charge. When no voltage is applied, we are left with equilibrium noise, called thermal noise or Johnson Nyquist noise. It results from the stochastic thermal population of electronic states within the energy range k_BT around the Fermi energy for a temperature T. We will show a continuous

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00003-4
© 2021 Elsevier Ltd.
All rights reserved.

cross-over between thermal and shot noise when the voltage V is so qV is approximately K_BT, q being the carrier charge. We will not consider the noise extrinsic to electronic quantum transport resulting from moving impurities in the conductor. This noise often shows a $1/f$-like spectral noise density and is due to conductance fluctuations and varies quadratically with current.

The study of the quantum current fluctuations in a conductor, or quantum shot noise, is a quite recent field of quantum condensed matter. It emerged in the mid-90s just after the theoretical works of Lesovik [4], Martin and Landauer [5] and Büttiker [27] which advocated that new quantum information is contained in current fluctuations and not obtained by measuring conductance only. The pioneering experiments demonstrating the possibility of realising accurate current noise measurements by Reznikov et al. [6] and by Kumar et al. [7] have been followed by a considerable number of noise experiments and have triggered an even larger number of theoretical publications. In pioneering theoretical works, the analogy between the quantum noise of electrons transferred through a conductor and the noise of photons propagating in an optical medium has pushed towards the development of Electron Quantum Optics where a QPC was used to mimic optical beam splitters. This has pushed experimentalists to develop single electron sources, analog to single photon sources. Here again, shot noise has been a key tool. In the absence of single electron detectors to detect where the injected electron arrived in the conductor, statistical measurements are done by measuring the resulting current fluctuations. We will see in this chapter that quantum shot noise can be used to measure the (fractional) charge of electronic carriers, to perform the energy spectroscopy of electronic excitations, to determine unambiguously the transmission probability of electronic quantum modes and to count the number of electron and hole generated by AC excitations and even to detect THz radiation.

There are already complete reviews on shot noise [1–3] particularly on the theoretical aspects. Concentrating on the quite large number of experimental studies of shot noise using semiconductor devices, this chapter has no room to make a comprehensive review of all theoretical publications and the author apologises for only quoting few works addressing fundamental aspects of noise, leaving aside some inspiring and promising new ideas.

2. The physics of quantum shot noise

2.1 The binomial statistics of quantum shot noise

As explained in the introduction, we will consider conductors shorter than the quantum coherence length. Electron thus propagates elastically (energy conserved) from one reservoir to another reservoir where they suffer inelastic scattering and de-coherence. The reservoirs are usually wide regions of the conductor which form the contacts. To give a physical understanding of shot noise, we will start with an ideal conductor made of a single quantum mode, i.e., a 1-D conductor. This can be realised using a single wall carbon nanotube (SWCN) or using a QPC transmitting a single electronic quantum mode. Disregarding spin for simplicity, we will discuss the effect of 1-D scattering and finally the scattering between modes for more general 2-D or 3-D conductors. Interactions will be first disregarded and the important role played by the Fermi statistics will be emphasised. The Fermi statistics leads to two spectacular consequences: conductance quantisation and noiseless conduction, both effects arising simultaneously.

To show this, let us start with the ideal one-dimensional (or single mode) conductor connected between two contacts shown in Fig. 6.1. At equilibrium, each reservoir emits electrons in the wire from the bottom of the conduction band to the Fermi energy. Electrons, not being scattered in the lead, are free to propagate and reach the opposite contact where they

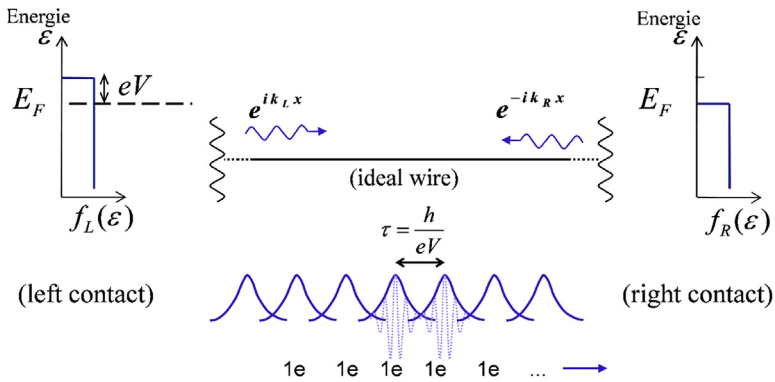


Figure 6.1 Ideal leads are connected to left (L) and right (R) contacts. The waves emitted by the left reservoir in the range E_F, $E_F + eV$ can be transformed into a compact set of orthogonal wave packets separated in time by h/eV (or spatially separated by $v_F h/eV$ with v_F the Fermi velocity). The Fermi statistics ensures the conductance is quantified in units of e^2/h and the electron flow is noiseless.

are absorbed. The left and right currents exactly compensate to give zero net current. Applying a DC voltage V to the left reservoir increases its electrochemical potential by eV and gives a net current whilst no electron is emitted by the right reservoir in the same energy range eV. The current is given by:

$$I = \frac{e^2}{h} V, \tag{6.1}$$

and the conductance is given by the conductance quantum e^2/h. Here, spin is disregarded. Throughout this section, we assume the temperature is zero. A doubling of the conductance may occur for spin degeneracy, in absence of an external magnetic field. The remarkable conductance quantum e^2/h results from two robust ingredients: *Heisenberg and Pauli principles*. Indeed, let us consider electrons with energy ε in the range $[E_F, E_F + eV]$. Following the approach of [5], by integration over the energy range eV one can perform a unitary transformation of the right moving plane waves defined at each energy ε into a series of right moving orthogonal wave packets whose separation in time is exactly h/eV according to Heisenberg uncertainty. As the Pauli principle requires single occupancy of each wave packet, the current is $(1 \times e)\ eV/h$ and the conductance e^2/h. If electrons were Bosons, no particle current quantisation would occur as the wave packet occupation number is not forced to be one.

A second spectacular manifestation of Fermi statistics is the absence of current noise. Indeed, the regular single electron injection by the left reservoir at a pace eV/h leads to a fundamentally noiseless current at all frequencies. Defining the current as a sum of the mean value I and a fluctuation, $I(t) = I + \Delta I(t)$:

$$\left\langle (\Delta I(t))^2 \right\rangle = 0, \tag{6.2}$$

One must stress that *conductance quantisation and noiseless current are linked* and reflect the powerful Pauli principle. This requires the absence of electron backscattering for each electronic mode, i.e., prefect transmission. This can be realised in a QPC where the number of fully transmitted electronic modes can be controlled by a gate, see Refs. [8,9] for conductance quantisation and [6,7] for noise suppression in QPCs.

To understand and identify the physical relevant ingredients responsible for shot noise in conductors with minimal complexity, we consider the previous ideal 1D wire in which a single point scatterer is placed in the middle. The scattering now gives rise to a finite transmission probability D of the

electronic wave through the wire. As the current injected by the left reservoir is $e.(eV/h)$, the current flowing through the wire becomes the following:

$$I = D\frac{e^2}{h}V \tag{6.3}$$

This gives the Landauer formula for the conductance of a single mode conductor: $G = De^2/h$. If spin degeneracy is included, the conductance is doubled (Fig. 6.2).

It is easy to derive the low frequency shot noise of the current due to scattering by using an electron counting approach. To take into account finite temperature, a rigorous derivation using second quantisation is given in the next section. We consider a large number N_0 of electrons emitted during a time τ by the left contact. It is given by $N_0 = eV\tau/h$. As each incoming electron can be either transmitted or reflected by the scatter, the number N of transmitted electron obeys the binomial probability distribution $P(N)$:

$$P(N) = \binom{N_0}{N} D^N (1-D)^{N_0-N}, \tag{6.4}$$

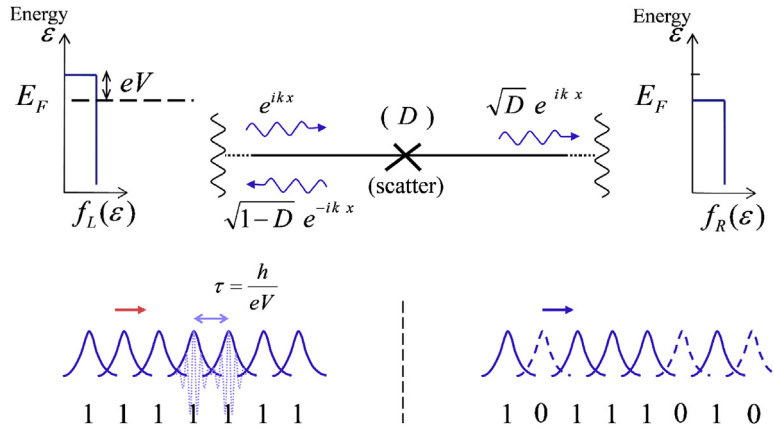


Figure 6.2 Ideal lead plus one point scatter. Electrons are now transmitted with probability D and reflected with probability $R = (1 - D)$. On the top of the figure, plane waves $\propto \exp ikx$ are indicated, whilst in the bottom, the orthogonal Martin-Landauer wavepackets are displayed. As the sequence of incoming electrons is noiseless, the statistics of electron is given by a binomial law characterising the partitioning of electrons by the scatter.

where $\begin{pmatrix} N_0 \\ N \end{pmatrix} = \frac{N_0!}{N!(N_0-N)!}$ is the binomial factor. Measuring $P(N)$ gives the Full Counting Statistics (FCS) of electrons. This is a very difficult experimental task and this has been done only in a classical regime using quantum dots; see section 4.1.5. The mean electron number and the variance of the fluctuation $\Delta N = N - \langle N \rangle$ with respect to the mean $\langle N \rangle$ are the following:

$$\langle N \rangle = DN_0, \tag{6.5}$$

$$\langle \Delta N^2 \rangle = D(1-D)N_0, \tag{6.6}$$

Restoring electrical units for the mean current $I = eN/\tau$ and introducing the low frequency Nyquist measurement bandwidth $\Delta f = 1/2\tau$, we obtain the zero frequency current fluctuations [4,5,27]:

$$\langle (\Delta I)^2 \rangle = 2e\langle I \rangle D(1-D)\Delta f, \tag{6.7}$$

Taking the limit of small transmission $D << 1$, Eq. (6.7) gives the Schottky shot noise formula [21] valid for the Poissonian statistics regime and applicable to tunnel junctions or for p-n junctions in macroscopic semiconductors. It is convenient to define the Fano factor F as the ratio of the actual shot noise normalised to the Schottky noise:

$$F = 1 - D \tag{6.8}$$

A Fano factor smaller than one is characteristic of a binomial law, an example of sub-Poissonian statistics. This is the hallmark of all quantum conductors and this is a direct consequence of the Fermi statistics which makes the incoming electron flow noiseless (at zero temperature) so that the only possible noise is due to scattering. For unit transmission $F = 0$, we recover (at zero temperature) the noiseless conduction associated with conductance quantisation.

Now we are ready to generalise these ideas to the case of an arbitrary phase coherent conductor transmitting several electronic modes. In an electron scattering wave approach, the conductor is viewed as an electronic waveguide even though it contains many scattering impurities given rise to complex interference effects. Nevertheless, one can always define some generalised transmission and reflection coefficients D_λ and $R_\lambda = 1 - D_\lambda$, respectively. Typically, the number of orbital modes propagating at the Fermi energy is respectively $N = 2W/\lambda_F$ and $\pi(W/\lambda_F)^2$ for a 2D or 3D

conductor of width W. The transmission coefficients D_λ are the eigenvalues of the product of two matrices $S_{RL}^{\dagger} S_{RL}$ where the elements $S_{RL} = \{s_{R,L;n,m}\}$ relate the amplitudes of transmission of an electronic wave emitted by the left reservoir from a mode m to that transmitted to and absorbed in the right reservoir in mode n. S_{RL} is an $N \times N$ sub-matrix of the general $2N \times 2N$ S matrix of the conductor. Finally, when spin degeneracy is considered, a factor 2 can be added to take into account double occupancy of each mode. With this decomposition, each effective mode λ contributes independently to the mean current and to the shot noise. We get, including spin degeneracy:

$$G = 2\frac{e^2}{h}\left(\sum_{\lambda} D_\lambda\right), \tag{6.9}$$

$$S_I = 2e\frac{e^2}{h} V\left[\sum_{2\lambda} D_\lambda(1 - D_\lambda)\right] = 2eI \cdot F, \tag{6.10}$$

where we have introduced the spectral density of current fluctuations: $S_I \Delta f = \langle (\Delta I)^2 \rangle$ and $I = GV$. The Fano factor now becomes the following:

$$F = \frac{\sum_{\lambda} D_\lambda(1 - D_\lambda)}{\sum_{\lambda} D_\lambda}, \tag{6.11}$$

The shot noise formulae (6.10) and (6.11) can be understood in terms of N independent single mode conductors $\{\lambda\}$ connected in parallel, as shown in Fig. 6.3, each generating shot noise given by Eq. (6.7) with $D = D_\lambda$. The temporally incoherent current fluctuations of each conductor add quadratically leading to Eq. (6.10) (Fig. 6.4).

The main message that we can extract from Eq. (6.11) is that the shot noise of coherent conductors is always sub-Poissonian because the Fano factor is always lower than one. Let us emphasise that this is due to the two following ingredients: (1) noiseless emission of electrons by reservoirs due to Fermi statistics and (2) the binomial (or multinomial) sub-Poissonian statistics of electrons partitioned between transmitted and reflected modes.

A widely studied general conductor is a diffusive metal (size larger than the elastic mean free path of electrons, but still showing coherent scattering). There are so many scattering centres that it is out of question to ab initio

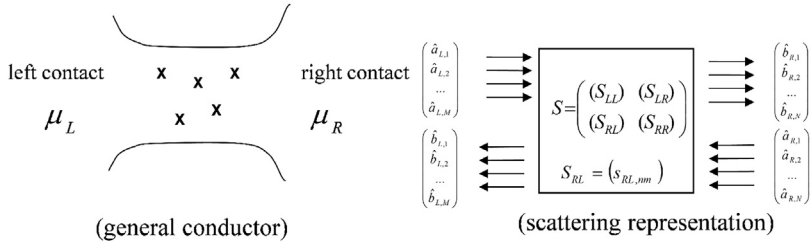


Figure 6.3 Left: general conductor (2-D or 3-D). The small crosses indicate scatterers whose scattering properties can be known ab initio or can be even totally random. Right: a general conductor is described by a scattering matrix S relating the transmission and reflection amplitudes of incoming electron waves and outgoing waves. It is made of four block matrices describing $L \rightarrow L$ or $R \rightarrow R$ amplitudes of reflection probability and $L \rightarrow R$ or $R \rightarrow L$ amplitudes of transmission probability. Despite the possibility of very complex scattering, the conductor can always be described by a finite number of transmission coefficients D_λ whose number is equal to that expected for a clean waveguide of the same transverse dimensions.

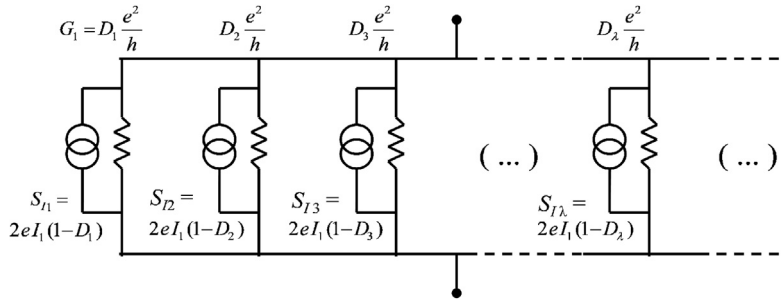


Figure 6.4 The shot noise of a general conductor can be viewed as several independent current shot noise sources generated by single mode conductors with transmission D_λ connected in parallel. This leads to the total conductance $G = \frac{e^2}{h}\left(\sum_\lambda D_\lambda\right)$, spin not included, and to the shot noise Fano factor $F = \frac{\sum_\lambda D_\lambda(1-D_\lambda)}{\sum_\lambda D_\lambda}$.

compute the S matrix for a specific sample. However, when little is known about a system containing a large number of parameters distributed at random (here the transmission), it is often possible to say something about the most probable statistics of the parameters. This is indeed the case for a diffusive system of conductance G, where the probability distribution $P(D_\lambda)$ of the transmission is given by the so-called Dorokhov distribution [12]:

$$P(D_\lambda) = \frac{G}{2e^2/h}\frac{1}{2D_\lambda\sqrt{1-D_\lambda}} \tag{6.12}$$

The Fano factor is as follows [13,14,16]:

$$F = \frac{\int_0^1 dDP(D)D(1-D)}{\int_0^1 dDP(D)D} = \frac{1}{3}, \tag{6.13}$$

This leads to a significant reduction of noise, whilst the mean transmission, given by the ratio of the mean free path to the sample length $\sim l_{el}./L \ll 1$, would suggest naively $F \simeq 1$. Contrary to our naive intuition, the distribution of transmission eigenvalues is not a Gaussian distribution centred around the mean value but is a bimodal distribution. In a diffusive system, about 63% of the modes are well transmitted and contribute almost no noise, whilst the 33% remaining are weakly transmitted and contribute to Poissonian noise.

Similar ensemble statistics of transmission probability distributions have been obtained for electrons transmitted through a chaotic ballistic cavity (or electron billiard) giving $F = 1/4$ [15–17] and for electron transmission through a random interface: $F = 1/2$. Examples will be given in Sections 4.1.3 and 4.1.4.

2.2 Quantum scattering approach to shot noise

Here, we describe the theoretical approach which allows shot noise predictions to be made for finite temperature and finite frequency using a second quantisation approach. To simplify the presentation, we will derive an expression for the 1-D conductor considered in 2.1. In subsection 2.2.1, I will consider the so-called DC Shot Noise at finite temperature which corresponds to measuring the low frequency current fluctuation of conductors biased by a DC voltage, in subsection 2.2.2, the high frequency shot noise when the frequency is higher than the temperature and finally and in subsection 2.2.3, the low frequency Photo-Assisted Shot Noise (PASN) when the conductors are experiencing AC voltages at frequency ν higher than the temperature. Examples for each regime will be given in Part.2.4.

2.2.1 DC shot noise

To calculate the conductance and noise, we first use a simple theoretical framework, following the approach of Refs. [10] and [11]. This allows the calculation of the current and its fluctuation in a quantum framework. To avoid heavy notation and make the physics more transparent, let us consider

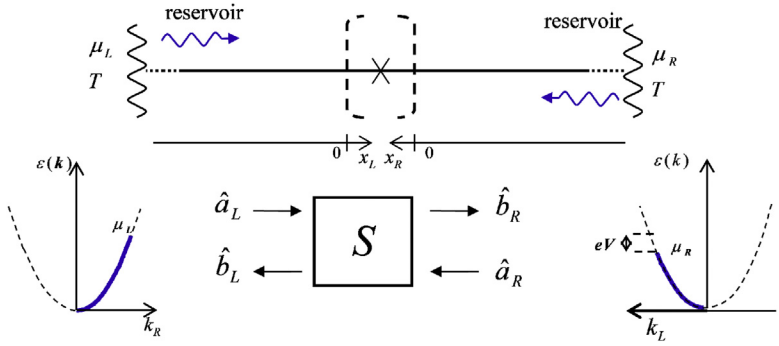


Figure 6.5 The shot noise of a general conductor can be viewed as independent current shot noise sources generated by single mode conductor of transmission D_λ connected in parallel. This leads to conductance $G = \frac{e^2}{h}\left(\sum_\lambda D_\lambda\right)$, spin not included, and Fano factor $F = \frac{\sum_\lambda D_\lambda(1-D_\lambda)}{\sum_\lambda D_\lambda}$.

the following short 1D conductor shown in Fig. 6.5. The conductor lies within the dashed brackets. It represents a scattering region which scatters the electronic waves emitted by left (L) and right (R) reservoirs. The reservoirs are described by semi-infinite ideal leads connected, at infinity, to contacts injecting electrons with chemical potential $\mu_{R,L}$ and temperature T. The left and right ideal leads behave as perfect reservoirs. They emit electrons towards the scattering region, the conductor, according to their chemical potential and temperature and they are able to absorb all electrons scattered or transmitted by the scattering region. Indeed, the leads being semi-infinite and ideal, an electron moving away from the scattering region never comes back. So defined, the reservoirs can be viewed as black-body sources and sinks of electrons.

Let us start by defining various useful physical quantities. The electron kinetic energy in the left reservoir (semi-infinite lead at $x_L < 0$) is $\varepsilon(k_L) = \frac{\hbar^2 k_L^2}{2m}$ in the left lead and similarly $\varepsilon(k_R) = \frac{\hbar^2 k_R^2}{2m}$ in the right lead. As coherent scattering is elastic, an electron scattered from left to right, or vice versa, keeps its energy and thus its momentum $k_R = k_L = k(\varepsilon)$. The associated velocity is as follows: $v_L = v_R = v(\varepsilon) = \hbar k(\varepsilon)/m$. Finally, a voltage drop V imposed by the external circuit gives the electro-chemical potential difference $\mu_L - \mu_R = eV$.

An electronic wave incoming from the left at energy ε and normalised with respect to energy gives the scattering state:

$$\begin{aligned}\varphi_L(\varepsilon;x) &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\hbar v(\varepsilon)}}\left(e^{ik(\varepsilon)x_L} + s_{LL}(\varepsilon)e^{-ik(\varepsilon)x_L}\right) \text{ for } x_L \le 0 \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\hbar v(\varepsilon)}}\, s_{RL}(\varepsilon)e^{-i(k(\varepsilon)x_R} \text{ for } x_R \le 0 \end{aligned} \tag{6.14}$$

where we have introduced the scattering amplitudes. s_{LL} is the reflection amplitude at the left side and s_{RL} is transmission amplitude from left to right. Similarly, the wavefunction of an electron emitted from the right reservoir at energy ε is described by the scattering state:

$$\begin{aligned}\varphi_R(\varepsilon;x) &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\hbar v(\varepsilon)}}\, s_{LR}(\varepsilon)e^{-i(k(\varepsilon)x_L} \text{ for } x_L \le 0 \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\hbar v(\varepsilon)}}\left(e^{ik(\varepsilon)x_R} + s_{RR}(\varepsilon)e^{-ik(\varepsilon)x_R}\right) \text{ for } x_R \le 0 \end{aligned} \tag{6.15}$$

s_{RR} is the reflection amplitude at the right side and s_{LR} is transmission amplitude from right to left. The four scattering amplitudes define the scattering matrix shown in Fig. 6.5

$$S = \begin{pmatrix} s_{LL} & s_{LR} \\ s_{RL} & s_{RR} \end{pmatrix} \tag{6.16}$$

which relates the left/right incoming wave amplitudes (a_L, a_R) to the left/right outgoing wave amplitudes (b_L, b_R). In order to calculate the current noise which results from the fluctuation of the population of electronic waves, one has to replace amplitudes by the second quantised fermionic operators $\widehat{a}_L$ and $\widehat{a}_R$. The operators act separately on the occupation numbers of the left and right reservoirs. For example, $\widehat{a}_L(\varepsilon)$ is the annihilation operator which acts on the left reservoir state $|n_L(\varepsilon)>$ and 'checks' if the state is occupied or not ($n_L(\varepsilon) = 1$ or 0). The annihilation operators $\widehat{a}_{L,R}$ and their hermitic conjugate, the creation operators $\widehat{a}^{\dagger}_{L,R}$, obey the useful relations:

$$\left\{\widehat{a}^{\dagger}_{\alpha}, \widehat{a}_{\beta}\right\} = \delta_{\alpha,\beta}\delta(\varepsilon' - \varepsilon) \tag{6.17}$$

$$\left\langle \widehat{a}^{\dagger}_{\alpha}\widehat{a}_{\beta}\right\rangle = f_{\beta}(\varepsilon)\delta_{\alpha,\beta}\delta(\varepsilon' - \varepsilon) \tag{6.18}$$

In the above equations, $\{.\}$ denotes the anti-commutator and $\langle . \rangle$ the thermodynamic average of the quantum expectation value of $\widehat{n}_\beta(\varepsilon) = \widehat{a}^\dagger_\alpha \widehat{a}_\beta$ which identifies to $f_\beta(\varepsilon) = 1/(1+\exp(\varepsilon - \mu_\beta)/k_B T)$ the Fermi distribution of reservoir $\beta = R$ or L. The last step is to write the fermion operator $\widehat{\psi}(x,t)$ from which all quantities can be calculated:

$$\widehat{\psi}(x,t) = \int d\varepsilon (\widehat{a}_L(\varepsilon)\varphi_L(\varepsilon;x) + \widehat{a}_R(\varepsilon)\varphi_R(\varepsilon;x)) e^{-i\varepsilon t/\hbar} \tag{6.19}$$

Here $x = x_{L(R)}$ when considering the scattering states on the left (right) side and $\varphi_{L(R)}$ are given by Eqs. (6.14) and (6.15). Defining

$$\begin{pmatrix} \widehat{b}_L \\ \widehat{b}_R \end{pmatrix} = \begin{pmatrix} s_{LL} & S_{LR} \\ S_{RL} & S_{RR} \end{pmatrix} \cdot \begin{pmatrix} \widehat{a}_L \\ \widehat{a}_R \end{pmatrix} \tag{6.20}$$

The fermion operator on the left side of the scattering region is written as follows:

$$\widehat{\psi}(x_L;t) = \int d\varepsilon \frac{1}{2\pi\hbar v(\varepsilon)} (\widehat{a}_L(\varepsilon) e^{ik(\varepsilon)x_L} + \widehat{b}_L(\varepsilon) e^{-ik(\varepsilon)x_L}) e^{-it/\hbar} \text{ for } x_L \leq 0 \tag{6.21}$$

A similar expression holds on the right side $x_R \leq 0$ by exchanging $L \to R$. Now, we can define the current operator $\widehat{I}(x;t) = \widehat{\psi}^\dagger(x;t) \frac{-i\hbar}{m} \frac{\partial}{\partial x} \widehat{\psi}(x;t)$ which is explicitly written:

$$\widehat{I}(x,t) \simeq -\frac{e}{h} \int d\varepsilon d\varepsilon' \left\{ a^\dagger_R(\varepsilon') a_R(\varepsilon) - b^\dagger_R(\varepsilon') b_R(\varepsilon) \right\} e^{i(\varepsilon'-\varepsilon)(t+x/v_F)/\hbar} \tag{6.22}$$

where the $\simeq$ sign means that we have anticipated that only energies close to the Fermi energy will contribute and we have replaced all velocities by the Fermi velocity v_F. From Eqs. (6.18) and (6.22), we can calculate the mean current:

$$I = \langle \widehat{I}(x;t) \rangle = \frac{e}{h} \int d\varepsilon (f_L(\varepsilon) - f_R(\varepsilon)) D(\varepsilon) \tag{6.23}$$

where we have used the unitarity relation $S^\dagger S = 1$ of the scattering matrix S and defined $D(\varepsilon) = |s_{RL}|^2 = |s_{LR}|^2$, the transmission probability. Eq. (6.23)

is the general form of the Landauer formula which explicitly gives for constant $D(\varepsilon) = D$ and electro-chemical potential difference $eV = \mu_L - \mu_R$.

$$I = D\frac{e^2}{h}V \tag{6.24}$$

Now, let us look at the current fluctuations. They are best observed when looking at the spectral density of the current noise $S_I(\omega)$. In practice (see Part. 2.3), for low frequencies, an analog to digital converter is used and the computer calculates the fast Fourier transform of the two-time current correlator $2I(t)I(t+\tau)$ over time difference τ whilst making the average over t (for DC voltage bias, the statistics is stationary and independent of the time reference). For the present quantum system, the spectral density $S_I(\omega)$ of current fluctuation as a function of frequency ω is as follows:

$$S_I(\omega) = 2\int d\tau \left\langle \widehat{I}(t)\widehat{I}(t+\tau)\right\rangle e^{i\omega\tau} \tag{6.25}$$

Here, the current operators at two different times do not usually commute, and the order in the product may matter. This issue will be discussed in the high-frequency shot noise part and is not relevant for low frequencies discussed here. Expression (6.25) involves the quantum statistical average of four fermion operators, which, using relation (6.18), leads to products of the form $f_\alpha(1 - f_\beta)$. These products mean that, in order to have a current fluctuations, electronic transitions from occupied to unoccupied states, or vice versa, must occur. As a consequence, the computation of current fluctuations to the left side of the scattering region from Eq. (6.25) gives for $\omega = 0$:

$$\begin{aligned} S_I(0) = 2\frac{e^2}{h}\int d\varepsilon \big\{ &(1 - R(\varepsilon))^2 f_R(\varepsilon)(1 - f_R(\varepsilon) + D(\varepsilon)^2 f_L(\varepsilon)(1 - f_L(\varepsilon) \\ &+ D(\varepsilon)R(\varepsilon))f_L(\varepsilon)(1 - f_R(\varepsilon)) + R(\varepsilon)D(\varepsilon)f_R(\varepsilon)(1 - f_L(\varepsilon))\big\} \end{aligned} \tag{6.26}$$

Here, $R = |s_{LL}|^2 = |s_{RR}|^2 = 1 - D$ is the reflection probability. The first two terms, of the form $f_\alpha(1 - f_\alpha)$, represent the current noise due to the fluctuating electron population in the left ($\alpha = L$) and right ($\alpha = R$) reservoirs due to thermal excitations. The last two terms represent the noise of electrons partitioned between left and right reservoirs by the scattering (for example, the third term implies that an electron emitted from the

reservoir (L) and transmitted to the right (probability Df_L) 'has to check' that no electron emitted from the right reservoir has been reflected (probability $R(1-f_R)$). The cancellation of the noise which occurs if the two states are occupied implies electron anti-bunching. For transmission not dependent on energy, integration over energy gives the following:

$$S_I(0) = \frac{e^2}{h}\left\{4k_BTD^2 + 2eVD(1-D)\coth(eV/2k_BT)\right\} \tag{6.27}$$

At zero temperature, this gives partition noise $\propto D(1-D)$ only:

$$S_I(0) = 2eD(1-D)\frac{e^2}{h}V = 2eI(1-D) = 2eIF\Bigg\} \tag{6.28}$$

One recovers the Fano factor (Eq. 6.8) introduced in subsection 2.2.1. At finite temperature but zero voltage V

$$S_I(0) = 4k_BT\frac{De^2}{h} = 4Gk_BT \tag{6.29}$$

one recovers the Johnson–Nyquist formula of pure thermal noise, with the conductance $G = D\frac{e^2}{h}$. Interestingly, comparing Eq. (6.27) and Eq. (6.29), one sees that the conductance $\propto D$ which occurs in the thermal noise expression comes from the identity $D = D^2 + D(1-D)$. This reveals a mixed contribution of thermal noise from the reservoir states $\propto D^2$ and of partition noise (or shot noise) from scattering by the conductor $\propto D(1-D)$. For small D, the case of a tunnel junction, for instance, the thermal noise is thus shot noise like.

For both finite voltage and temperature, a cross-over occurs between the Johnson–Nyquist regime and the pure shot noise regime when the voltage crosses the value $V = 2k_BT/e$. This is shown in Fig. 6.6:

To end this part devoted on DC shot noise (i.e., the low frequency non-equilibrium noise due to static DC voltage bias), we give the expression for a general conductor characterised by the set of transmission probabilities $\{D_\lambda\}$:

$$S_I(0) = \frac{e^2}{h}\left\{4k_BT\sum_\lambda D_\lambda^2 + \sum_\lambda D_\lambda(1-D_\lambda)2eV\coth(eV/2k_BT)\right\} \tag{6.30}$$

whilst the conductance is $\frac{e^2}{h}\sum_\lambda D_\lambda$.

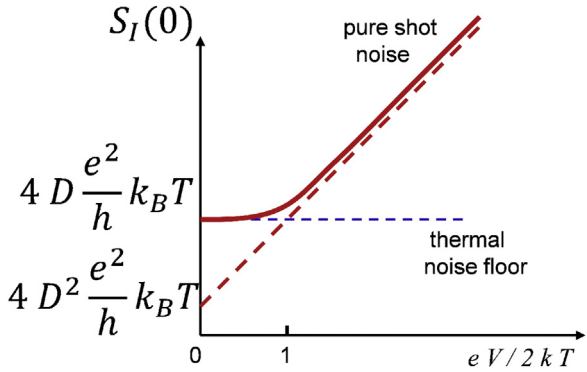


Figure 6.6 Evolution of the finite temperature low frequency shot noise with voltage for a single mode conductor of transmission *D*, spin disregarded.

Measuring the DC shot noise is a means to get information about a conductor which is not given by only making conductance measurements. The conductance is a measure of the transmission of electronic wave; the better they are transmitted, the larger the conductance. Shot noise measures the dual, particle like, property of the carriers by measuring not the number of particles transferred in a given measurement time but the fluctuations of this number. Despite this strong qualitative difference with conductance, combining shot noise and conductance gives two independent relation of the transmissions: a sum $S = \sum_\lambda D_\lambda$ and a product $P = \sum_\lambda D_\lambda(1 - D_\lambda)$. In the case of two modes, this allows us to unambiguously determine each transmission D_1 and D_2.

2.2.2 Finite frequency shot noise

Here, we address the regime where the current noise is measured at frequencies ν much higher than characteristic thermal frequency k_BT/h. In the previous section, we have seen that shot noise sets in when the voltage V is greater than the temperature ($eV > k_BT$). As a consequence, the high-frequency regime implies that ν becomes comparable to the characteristic voltage frequency eV/h and to k_BT/h. eV/h is the frequency of the Martin–Landauer wavepackets [5] mentioned above. We thus expect that, in addition to providing information on scattering not given by conductance, high-frequency shot noise will also provide information about the fundamental nature of quantum conduction.

To derive the high frequency shot noise, we start with expression (6.25) where $\omega = 2\pi\nu$. By changing ω into $-\omega$ in this expression, we arrive at a

similar expression in which the order of the two current operators has been reversed, i.e., $\langle \widehat{I}(t)\widehat{I}(t+\tau)\rangle \rightarrow \langle \widehat{I}(t+\tau)\widehat{I}(t)\rangle$. The physical meaning of the two correlators will become clear after deriving their expressions. As one can go from one choice of correlator to another one by changing ω into $-\omega$, we will start with Eq. (6.25). This gives, assuming energy independent transmission

$$S_I(\omega) = 2\frac{e^2}{h}\int d\varepsilon \big\{(1-R)^2 f_L(\varepsilon)(1-f_L(\varepsilon-\hbar\omega)) + D^2 f_R(\varepsilon)(1-f_R(\varepsilon-\hbar\omega)) + RDf_L(\varepsilon)(1-f_R(\varepsilon-\hbar\omega)) + DRf_R(\varepsilon)(1-f_L(\varepsilon-\hbar\omega))\big\} \tag{6.31}$$

One observes from the terms $f_\alpha(\varepsilon)(1-f_\beta(\varepsilon-\hbar\omega))$ that current fluctuations occur when electrons occupying a state at energy ε can find an empty state at energy $\varepsilon-\hbar\omega$. This suggests that measuring the high frequency noise given by the above relation for $\omega > 0$ is connected to the emission of photons in the external circuit [18,19].

Indeed, the integration of energy for the first two terms corresponding to thermal fluctuations in the reservoirs gives rise to a photon Bose–Einstein distribution:

$$N(\omega) = \frac{1}{\hbar\omega}\int d\varepsilon\, f_\alpha(\varepsilon)(1-f_\alpha(\varepsilon-\hbar\omega)) = \frac{1}{\exp \hbar\omega/k_BT - 1} \tag{6.32}$$

Changing the sign of ω gives the following identity: $(-\omega)N(-\omega) = (\omega)(1+N(\omega))$. This suggests an interpretation of noise measurements in terms of spontaneous photon emission to and stimulated photon absorption from the external circuit (this will be discussed in Part. 2.3). For positive ω, the current noise is called the emission noise and is given by the following:

$$S_I^{em.}(\omega) = 2\frac{e^2}{h}\big[2D^2\hbar\omega N(\omega) + D(1-D)\{(eV+\hbar\omega)N(\omega+eV/\hbar) + (\hbar\omega-eV)N(\omega-eV/\hbar)\}\big] \tag{6.33}$$

The so-called absorption noise is obtained by replacing $\omega \rightarrow -\omega$, i.e., $S_I^{abs.}(\omega) = S_I^{em.}(-\omega)$. In properly designed experiments, see Part.2.3, one measures S_I^{em}. For zero voltage, one recovers the finite frequency Johnson–Nyquist noise:

$$S_I^{em.}(\omega) = 4\frac{e^2}{h}D\hbar\omega N(\hbar\omega) = 4G\hbar\omega N(\omega) \tag{6.34}$$

This relation was derived in 1928 by Nyquist [22], considering the exchange of photons between two resistors connected via a microwave transmission line with matched characteristic impedance. It tells us that, when the frequency is much larger than temperature, the thermal radiation of photons is improbable and the noise vanishes. For zero temperature and voltage $V > 0$, only the last term of Eq. (6.32) remains and one finds the following [4]:

$$\begin{aligned} S_I^{em.}(\omega) &= 2(eV - \hbar\omega)\frac{e^2}{h}D \quad \text{for } eV > h\overline{\omega} \\ &= 0 \quad \text{for } eV < h\overline{\omega} \end{aligned} \tag{6.35}$$

The absorption noise $S_I^{abs.}$ is obtained by replacing $N(.) \rightarrow N(.) + 1$ in Eq. (6.33). As a consequence, we obtain the Kubo formula valid for finite temperature and applied voltage [20]:

$$S_I^{abs.} - S_I^{em.} = 4G\hbar\omega \tag{6.36}$$

with $G = e^2/h$ for the 1-D spinless conductor considered here. For a general conductor with electronic transmissions $\{D_\lambda\}$, the finite frequency shot noise is obtained by replacing in Eq. (6.33) the terms D^2 and $D(1 - D)$ by $\sum_\lambda D_\lambda$ and $\sum_\lambda D_\lambda(1 - D_\lambda)$.

2.2.3 Photon-assisted shot noise

Instead of looking at the emission current noise at high frequencies, which requires using difficult high-frequency techniques, see Section 3.2, one can get information on the high-frequency noise physics by simply irradiating a contact at high frequency. Here, photons will induce electronics electronic transitions in the Fermi sea. The photo-induced electron-hole pairs will generate a current noise observable at low frequency. We call this (zero-) frequency shot noise: PASN.

To understand the origin of PASN shot noise, let us consider again the simple situation of a spin-less 1-D conductor characterised by a transmission D. A voltage $V_{ac}(t) = v_{ac} \cos \omega t$ is applied on, say, the left contact. Using the approach developed of Ref. [24], the AC potential modulates the phase [23] of the electrons emitted by the left reservoir before they arrive at the scatterer. For an electron emitted by the left contact at energy ε below the Fermi energy, the wavefunction $\propto e^{ikx}e^{-i\omega t}$ becomes $\propto e^{ikx}e^{-i\omega t}e^{-i\phi(t)}$ where

$$\phi(t) = \int_{-\infty}^{t} dt' eV_{ac}(t')/\hbar \tag{6.37}$$

By developing the periodic phase term into its Fourier components

$$e^{-i\phi(t)} = \sum_{l=-\infty}^{l=+\infty} p_l e^{-il\omega t} \tag{6.38}$$

one observes that p_l represents the amplitude of probability of promoting an electron initially at energy ε to energy $\varepsilon + l\hbar\omega$. From this coherent energy scattering, one can define the new annihilation operators $\widehat{\widehat{a}}_L(\varepsilon)$ describing electrons arriving at the conductor at energy ε from the initial operator $\widehat{a}_L(\varepsilon)$ acting on the left reservoir:

$$\widehat{\widehat{a}}_L(\varepsilon) = \sum_l p_l \widehat{a}_L(\varepsilon - l\hbar\omega) \tag{6.39}$$

This tells us that the amplitude of probability of finding an electron incoming on the conductor at energy ε is given by the amplitude of the probability to have this electron initially occupying a state at energy $\varepsilon - l\hbar\omega$ multiplied by the amplitude of the probability of being scattered with an energy increment $l\hbar\omega$.

With these tools, the computation of the current and its fluctuations is simply obtained by replacing $\widehat{a}_L(\varepsilon)$ by $\widehat{\widehat{a}}_L(\varepsilon)$ in expression (6.22). A straightforward calculation of the zero frequency PASN where both a DC voltage and an AC voltage are applied to the left contact, i.e., $V_L(t) = V + v_{ac}\cos\omega t$, gives [23,25]:

$$S_I^{PASN}(V) = \frac{e^2}{h}\left\{4k_BTD^2 + \sum_l P_l 2D(1-D)(eV - l\hbar\omega)\coth(eV - l\hbar\omega)/2k_BT\right\} \tag{6.40}$$

where $P_l = |p_l|^2 = J_l(ev_{ac}/\hbar\omega)^2$ is the probability to absorb ($l > 0$) or emit ($l < 0$) l photons and $J_l(x)$ denotes an integer Bessel function of order l. Comparing this to the DC shot noise S_I^{DC} given by expression (6.27) when only a DC voltage V is applied on the left contact, we obtain the following useful expression:

$$S_I^{PASN}(V) = \sum_l P_l\, S_I^{DC}(V - l\hbar\omega/e) \tag{6.41}$$

This relation also holds for a general conductor where S_I^{DC} is given by Eq. (6.30). The physical interpretation is that performing noise measurements with the AC potential superimposed on the DC potential is like performing in parallel many DC shot noise measurements with the DC voltage

V shifted $l\hbar\omega/e$ and weighted by the probability P_l. In particular, as the zero temperature DC shot noise is $\propto |V|$, the zero bias singularity is replicated each time the DC voltage V is equal to $\hbar\omega/e$. We will see in Section 4.3.3 that the Josephson-like relation $V = \hbar\omega/e^*$ can be exploited to measure the charge e^* of fractional excitations in interacting systems.

3. Noise measurement techniques

3.1 Low frequency shot noise measurement techniques

The shot noise generated by coherent semiconductor nanodevices is in general very small. In order to maintain a coherent regime, the bias voltage applied to the conductor exceeds at most a 100 μV. As the conductance scales with $e^2/h = 1/(25.8\ \mathrm{k\Omega})$, the conductance quantum, this gives a dc current $I \simeq 4$ nA. The current noise spectral density to be detected is then a fraction of the largest noise a conductor can give: the Poissonian noise $S_I^{Pois.} = 2eI \simeq 13.10^{-28}\mathrm{A}^2/\mathrm{Hz}$. Convincing measurements requires a few percent accuracy and therefore a current noise sensitivity of $10^{-29}\mathrm{A}^2/\mathrm{Hz}$ or even better.

Technically, using a current amplifier is not a convenient approach because of the extra current noise injected into the sample by the feedback resistor. In almost all known set-ups, it is preferred to simply convert the current into a voltage across a resistor R and measure the voltage noise spectral density $S_V = R^2 S_I$. The resistor is however limited by bandwidth considerations. For low frequency noise measurements, R is at most 10 kΩ leading to a required voltage noise sensitivity of $\delta S_V = 10^{-23}\mathrm{V}^2/\mathrm{Hz}$. This is orders of magnitude smaller than the typical voltage noise $S_{V_n} \simeq 10^{-19}\mathrm{V}^2/\mathrm{Hz}$ of the best commercial low noise amplifiers. For a finite detection bandwidth Δf, increasing the noise sensitivity requires a long integration time τ of the Gaussian voltage fluctuations such that $S_{V_n}\sqrt{2/(\Delta f \tau)} \leq \delta S_V$. For a 100 kHz bandwidth, this requires an integration time $\tau = 1000$s. The bandwidth $\Delta f = 1/2\pi R C_L$ being limited by the lead capacitance C_L (100 kHz requires $C_L < 60$ pF for $R = 10$ kΩ) means that long acquisition times requiring good stability are typical in shot noise measurements.

Designing an efficient noise measurement is potentially a complex issue; we give in the following the most direct approach in order to find the best low noise detection set-up matched to the sample to be measured. This is based on the concept of noise temperature. Indeed, if we are able to measure the thermal noise of the conductor whose shot noise is to be measured, then

we have the sensitivity to measure its shot noise. This is clear from Fig. 6.6 and from Eq. (6.30). Let R be the resistance of the sample to be measured by an amplifier characterised by a voltage noise S_{V_n} and a current noise S_{In}. The total voltage noise referred to the amplifier input is as follows:

$$S_V^{in} = 4Rk_BT_s + S_{Vn} + R^2 S_{In} \tag{6.42}$$

Here, T_s is the sample noise temperature, i.e., the temperature at which the sample would generate an equilibrium thermal noise equal to the actual noise. For zero voltage bias V, $T_s = T$ the sample temperature, and, for $V >> 2k_BT/e$, $T_s \simeq F\frac{eV}{2k_B}$ according to Eq. (6.10). Let us now define the amplifier noise temperature T_{amp} referred to the sample resistance R_s or the resistance $R \simeq R_s$ used to convert current fluctuations into voltage fluctuations:

$$T_{amp} = \frac{S_{Vn}}{4k_BR} + \frac{RS_{In}}{4k_B} \tag{6.43}$$

The best noise detection is obtained for smaller values of T_{amp}, possibly close to T_s. The resistance $R = R_{opt}$ which minimises T_{amp} in 6.43 gives the best noise temperature T_{opt}. for noise detection:

$$R_{opt.} = \sqrt{S_{Vn}/S_{In}} \tag{6.44}$$

$$T_{opt.} = \frac{1}{2k_B}\sqrt{S_{Vn}S_{In}} \tag{6.45}$$

The noise characteristics of a selection of a few room temperature amplifiers are shown in Table 6.1. The selection of the best amplifier will be made so that the resistance used to convert the current noise into the voltage noise, usually close to the sample resistance R, matches the optimal resistance R_{opt}. The lead capacitance and input capacitance is another constraint affecting

Table 6.1 Characteristics of some commercial low noise amplifiers. The first two are ready to use amplifiers, the last four are monolithic op amp chips.

Reference	Voltage noise	Current noise	R_{opt}	T_{opt}	1/f	BW
SA-220 (NF)	$(0.45\ \text{nV})^2/\text{Hz}$	$(100\ \text{fA})^2/\text{Hz}$	5 kΩ	1.8K	—	80 MHz
LI75A (NF)	$(1.4\ \text{nV})^2/\text{Hz}$	$(14\ \text{fA})^2/\text{Hz}$	100 kΩ	710 mK	300 Hz	1 MHz
AD743	$(2.9\ \text{nV})^2/\text{Hz}$	$(6.9\ \text{fA})^2/\text{Hz}$	420 kΩ	720 mK	20 Hz	4.5 MHz
OPA627	$(4.5\ \text{nV})^2/\text{Hz}$	$(2.5\ \text{fA})^2/\text{Hz}$	1.8 MΩ	410 mK	100 Hz	16 MHz
OPA657	$(4.8\ \text{nV})^2/\text{Hz}$	$(1.3\ \text{fA})^2/\text{Hz}$	3.7 MΩ	225 mK	600 Hz	1.6 GHz
MAX4475/8	$(4.5\ \text{nV})^2/\text{Hz}$	$(0.5\ \text{fA})^2/\text{Hz}$	9 MΩ	80 mK	600 Hz	10 MHz

the detection bandwidth Δf. Reaching a mK noise temperature accuracy δT_n for 1 K amplifier noise temperature T_{amp} requires a product $\Delta f \tau$ of $\simeq 10^6$, according to the following relation resulting from Gaussian fluctuations:

$$\delta T_n = \left(T_s + T_{amp}\right) \frac{\sqrt{2}}{\sqrt{\Delta f \tau}} \tag{6.46}$$

As usually, $T_{amp} >> T_s$, having the lowest T_{opt}. and matching the sample resistance is highly desirable. If the sample resistance is much lower than R_{opt}, it is possible to put N amplifiers in parallel. As $S_{In} \rightarrow S_{In}/\sqrt{N}$ and $S_{Vn} \rightarrow S_{Vn}\sqrt{N}$, this divides R_{opt} by N whilst leaving T_{opt}. unchanged. For a DC connection, this also reduces the amplifier input offset voltage by $\sqrt{N}$. The maximum N is limited by frequency requirements as the input capacitance grows accordingly.

A considerable improvement in the reliability of shot noise measurements is achieved by performing cross-correlation measurements. The sample currents and the resulting voltage fluctuations are amplified by two distinct amplifiers. Then the cross-spectrum is performed, eliminating the uncorrelated amplifier voltage noise. Fig. 6.7 shows a possible setup. To understand the advantages of using cross-correlation, let us consider how uncorrelated current fluctuations $\delta I, \delta I_{n1}$ and δI_{n2} arising from the sample and the current noise of amplifiers 1 and 2, respectively, contribute to the input voltages $\delta V_{1,2}$ of each amplifier.

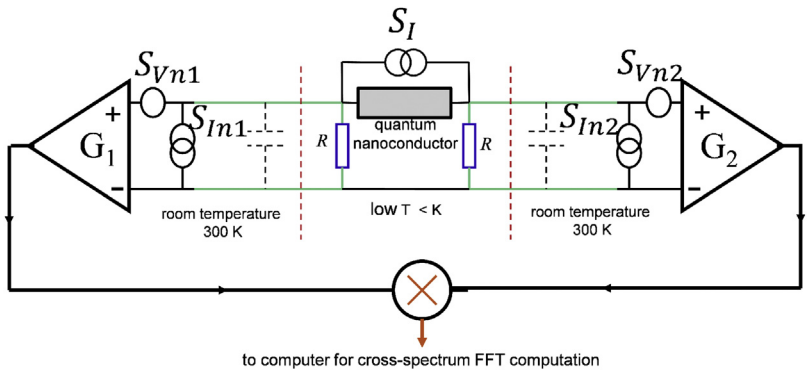


Figure 6.7 Cross-correlation noise measurement circuit used for low frequency noise measurements.

One finds

$$\delta V_1 = \frac{RR_s}{2R+R_s}\delta I + \frac{R(R+R_s)}{2R+R_s}\delta I_{n1} + \frac{R^2}{2R+R_s}\delta I_{n2} \tag{6.47}$$

$$\delta V_2 = -\frac{RR_s}{2R+R_s}\delta I + \frac{R(R+R_s)}{2R+R_s}\delta I_{n2} + \frac{R^2}{2R+R_s}\delta I_{n1} \tag{6.48}$$

One observes that the sample noise gives fluctuations of opposite sign, whilst the amplifier current noise contributes positively. Adding the voltage noise of each amplifier gives the following auto-correlation S_{V1} and cross-correlation $S_{V1,V2}$ spectral density of the detected current fluctuations.

$$S_{V1} = S_{Vn1} + \left(\frac{RR_s}{2R+R_s}\right)^2 S_I + \left(\frac{R(R+R_s)}{2R+R_s}\right)^2 S_{In1} + \left(\frac{R^2}{2R+R_s}\right)^2 S_{In2} \tag{6.49}$$

$$S_{V_1,V_2} = -\left(\frac{RR_s}{2R+R_s}\right)^2 S_I + \frac{R^3(R+R_s)}{(2R+R_s)^2}(S_{In2}+S_{In1}) \tag{6.50}$$

An expression similar to (6.49) holds for S_{V2}. One sees that the amplifier voltage noise, the dominant part of the noise, disappears from the cross-correlation. This subtraction makes the noise measurement much more reliable. Regarding noise accuracy, after averaging over a time τ, we get $S_{V1(2)}\sqrt{2}/\sqrt{\Delta f \tau}$ for autocorrelation and $S_{V1(2)}/\sqrt{\Delta f \tau}$ for the cross-correlation. The factor 2 reduction is due to the doubled detection. An important difference between cross and auto detection is that the sample noise gives rise to negative correlation, whilst all contribution to auto correlation is positive. This qualitative difference allows us to easily distinguish the sample noise from the thermal noise contribution of the measuring circuit which all contribute positively.

A further improvement is the use cryogenic amplification by using a cold High Electron Mobility Transistor (HEMT). When cooled down to few Kelvin, these transistors based on a 2D electron gas in a GaAs show larger gain and lower noise. Moreover, this allows us to place the first amplification stage, the HEMT, closer to the sample, thus reducing the lead capacitance and increasing the measurement bandwidth. However, because they have been designed for microwave (GHz) communication, they show large $1/f$ noise below 1 MHz. To recover the white noise regime, the common strategy is to place a high Q inductor L in parallel with the lead capacitance C_L,

such that the resonant frequency $1/2\pi\sqrt{LC_L}$ is a few MHz [26]. In practice, a good HEMT is the ATF-34143 transistor from Avago. With a voltage gain of 6–8, an input voltage noise of $\simeq 0.22$ nV/Hz$^{1/2}$ is achieved at 4 K with negligible current noise (due to the small <1 pF gate capacitance). Complete technical details have been given by several teams in Refs. [30–34,36]. Converting current fluctuations over a 10 kΩ resistor provides 22 fA/Hz$^{1/2}$ and after averaging, a time $\tau = 10^6/\Delta f$ (i.e., 10 s for 100 kHz bandwidth) provides a 4.5 10^{-31} A^2/Hz noise sensitivity using cross-correlation. Expressed in terms of noise temperature, this represents a resolution of 0.09 mK on the thermal noise of a 10 kΩ resistor in few seconds. Finally, noise calibration is an important task in order to perform an accurate comparison between the quantum noise generated by semiconductor nanodevices and the quantum noise model. The most reliable way is to calibrate the auto-correlation noise of the two detection channels using equilibrium Johnson–Nyquist noise. This is done by heating the sample and its resonant circuit and recording the difference between the power spectral density measure at temperature T minus the power spectral density at base temperature. This procedure allows us to know in detail the resonant circuit characteristics and to check the linear variation of the noise with temperature, the slope providing the calibration. An example is shown in Fig. 6.8.

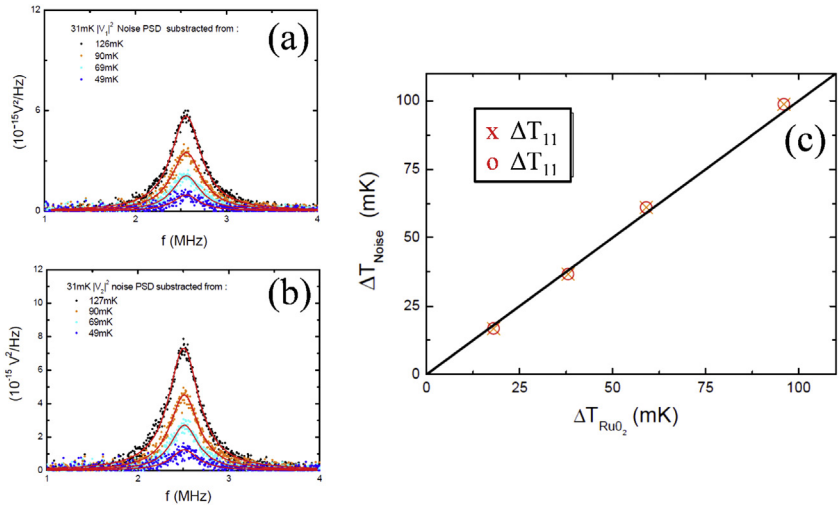


Figure 6.8 Johnson–Nyquist noise calibration: in (a) and (b), the difference of auto-correlation noise measured at temperature T minus that measured at base temperature $T_0 = 30$ mK is plotted for each detection channel (1) and (2). Comparison with $G^2_{1(2)} 4k_B(T - t_0)\Re Z(f)$ where Z(f) is the impedance of the resonant circuit provides a calibration of the gain $G_{1(2)}$ of each amplification chain. In (c), the noise recorded shows the expected linear variation versus temperature difference. *From Ref. [33].*

3.2 High frequency shot noise measurement techniques

Here, we address the high frequency regime. This is the regime where the frequency ν is such that the photon energy $h\nu$ is comparable to or higher than the thermal energy $k_B T_s$ where T_s is the device temperature. For $T = 100$ mK, this corresponds to 2.08 GHz, the microwave regime (incidently, the frequency is close to the carrier frequency used in telecommunications for cell phones, bluetooth, wifi, etc). It is best to use a quantum optics language where optical photons are replaced by microwave photons, as already suggested in Section 2.2.2. Indeed, the information between two ends of a circuit is mediated by photons which propagate as T.E.Modes in coaxial lines with a small characteristic impedance $Z_c = 50\Omega$. Let us consider a nanodevice of resistance R close to the quantum of resistance connected at $x = 0$ to a low noise microwave amplifier by a coaxial transmission line of length L. A current fluctuation $\delta I(t)$ generated by the sample excites a voltage fluctuation $\delta V(t, x = 0) = Z_c \delta I(t)$ which starts to propagates towards the detector at a speed $c/\sqrt{(\varepsilon)}$ where ε is the dielectric constant in the coaxial line. The finite impedance makes the notion of a current or voltage source as well as current noise and voltage noise inconvenient. It is best to reason in terms of electromagnetic power and noise power $\propto \delta I(t)\delta V(t)$. The microwave power emitted by the noisy sample and the power detected by the amplifier are related by a scattering matrix characterising the propagation through the connecting circuit in terms of reflection, transmission and a possible scattering to other electromagnetic modes. We note that, as early as 1928, Nyquist considered the thermal noise emitted by a resistor in terms of microwave photons and established the famous Nyquist thermal noise formula valid at all frequencies [22]. The noise power that can radiate a resistor of resistance R and noise temperature T_s in the frequency range $[\nu, \nu + \delta f]$ is as follows:

$$\delta P = N(\nu) h\nu \delta f \tag{6.51}$$

where $\nu = \Omega/2\pi$ is the Bose–Einstein distribution given by relation (6.32). To establish this relation, Nyquist assumed a characteristic impedance matching the sample resistance $Z_c = R$ such that the thermal noise emitted by the resistor is perfectly carried away by the transmission line. In practice, the situation is very different for the noise radiated by a quantum conductor as R is close to impedance quantum h/e^2, whilst Z_c is close to the vacuum impedance $\sqrt{\mu_0/\varepsilon_0}$ (their ratio being the fine structure constant). Here, most emission noise power is reflected back into the sample and few can be

detected in the external circuit. In a scattering language, using the reflection amplitude $S_{ss} = (Z_c - R)/(Z_c + R)$, the transmission coefficient of the noise power, from sample to detector, is as follows:

$$|S_{ds}|^2 = \frac{4RZ_c}{(R + Z_c)^2} = 1 - |S_{ss}|^2 \tag{6.52}$$

The small ratio Z_c/R makes the emitted noise power only a few % of the available power.

A second aspect making high frequency noise measurements difficult is the noise radiated by the amplifier used to detect the emitted noise. As considered in section 2.3.1, it is best to write the noise power in terms of equivalent noise temperature. To be more specific, let us define T_d^i the noise power radiated to the sample by the amplifier (this corresponds to the current noise of the amplifier) and T_d^v the noise which is added due to the amplification process (the voltage noise), whilst T_s is the sample noise temperature. The noise temperature measured is as follows:

$$T_m = |S_{ds}|^2 T_s + \left(1 - |S_{ds}|^2\right) T_d^i + T_d^v \tag{6.53}$$

The first term is the noise emitted by the sample and the second term the noise emitted by the detector and reflected back to it.

From a fundamental point of view, relation (6.53) makes a direct connection to the quantum relations between sample noise and detector noise as discussed in Refs. [18,20]. Using the Kubo formula relating the emission $S(\nu)$ and absorption $S(-\nu)$ noise power of the sample of conductance $1/R$ in the limit where $|S_{ds}|^2 \simeq 4R_s/R << 1$ and writing the amplifier current noise power as $k_B T_d^i = N_d(\nu)h\nu$, one gets the following:

$$T_m = Z_0[S_I(\nu)(N_d(\nu) + 1) - S_I(-\nu)N_d(\nu)] + T_d^{(\nu)} \tag{6.54}$$

Cryogenic microwave amplifiers were first developed by engineers working for radio-astronomers. They are now widely used for superconducting qubit applications. Cooled to 4 K, they can show total noise temperature $T_d^i + T_d^v$ less than 1.8 K in the 2–4 GHz range whilst dissipating only a few mW for good integration into a cryogenic environment. At 4 GHz, this corresponds to less than 10 photons added by the amplification and to $N_d \simeq 5$ in relation (6.54) as for properly designed amplifiers $T_d^i \simeq T_d^v$ and only half of the added photons are radiated towards the sample.

A good practice is to protect the sample from the amplifier microwave noise. This can be done by using a nonreciprocal device called an RF

circulator as shown in Fig. 6.9. With this trick, only the emission noise $S_I(\nu)$ is measured:

$$T_m = Z_c S_I(\nu) + T_d^{\nu} = T_s |S_{ds}|^2 + T_d^{\nu} \tag{6.55}$$

For an amplifier noise temperature $T_d = 2$ K, integrating the noise on a 100 MHz bandwidth gives a noise sensitivity of $\simeq 28$ μK in 100 s, according to relation (6.46). Despite the weak coupling $|S_{ds}|^2 \simeq 1\%$, this allows the measurement of sample noise temperature variations of $\simeq 3$ mK, i.e., a fraction of the sample's Johnson–Nyquist noise. This is thus good enough to measure the shot noise which, by nature, is always larger than the thermal noise.

Finally, we would like to mention a strategy employed in order to increase the coupling between the high impedance quantum conductor and the low impedance circuit. The technique uses an impedance transformer. For measurements with narrow bandwidth, one can use high-Q resonators. Although the frequency bandwidth shrinks which leads to slower acquisition of information, the increase in coupling is more important, giving rise to a net gain in sensitivity. However, a narrow bandwidth can be a drawback when obtaining information on the frequency dependence of the

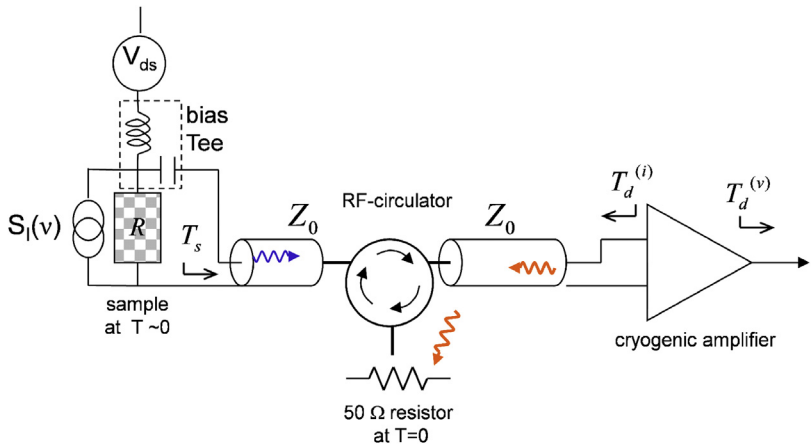


Figure 6.9 High frequency noise measurements: The shot noise power generated by the sample, resistance R, biased by a DC voltage V_{ds} is emitted in the coaxial line of characteristic impedance $Z_c = 50\Omega$. To prevent the amplifier noise from reaching the sample and interfering with measurements, an RF circulator is used. This nonreciprocal device directs the sample noise to the amplifier, whilst the amplifier noise is directed to 50Ω load at $T = 0$ and absorbed. The load, being at the lowest temperature, emits no thermal photons to the sample.

noise. Another solution is to use impedance transformation based on a series of $n = 4$ coaxial lines with increasing impedance. This was used in the experiment of Ref. [90] which will be discussed later. The principle is similar to optical anti-reflective coating used for glasses. The simplest impedance transformer is a geometrical series of characteristic impedances. The continuous limit gives the exponential horn. The so-called Chebyshev series and the constant reflection (also called binomial) series work better. Here, a large bandwidth can be obtained (from one octave to a decade in frequencies). Practically, using coplanar waveguides in the GHz range, the effective impedance viewed by the sample can reach 300Ω–1 kΩ.

4. Shot noise in semiconductor nanodevices

4.1 Quantum suppression of shot noise

The hallmark of the quantum noise of nanoconductors is a sub-Poissonian noise, i.e., a shot noise which is always lower than the Schottky Poissonian noise observed in tunnel and p-n junctions, as well as vacuum tubes. This lower noise is characterised by the Fano factor. We will address here the case of QPCs for which complete noise suppression can be observed, then we will describe examples of quantum shot noise in clean SWCNs and in graphene p-n junction. We will address the noise reduction which arises from materials exhibiting statistical properties of their transmission eigenvalues such as diffusive conductors or chaotic quantum billiards, all realised in 2DEGs in semiconductors. We will also address the finite frequency shot noise suppression due to Fermi statistics when the bias voltage is smaller than the energy needed for an electron to emit a photon.

4.1.1 Low frequency shot noise in Quantum Point Contacts and in other one-dimensional like nanoconductors

The archetype of a ballistic conductor is a QPC [8,9]. This is a short constriction realised in a 2D electron gas whose width can be tuned by the voltage of a split gate. This allows the selection of the number of propagating ballistic electronic modes and the control of their transmission. When only a single (spin degenerate) mode is transmitted, the QPC behaves as the ideal wire with a scatterer considered in Figs. 6.2 and 6.5 of Section 2.2. Early attempts to measure the low frequency shot noise in a QPC were affected by a strong conductance noise. This conductance noise arises from the random motion of charge impurities in the host semiconductor which makes the potential energy of the QPC fluctuate. According to the energy dependence of the

QPC transmission probabilities [27], the derivative of the transmission D_n of the *nth* transmitted mode varies as $dD_n/d\varepsilon \propto D_n(1-D_n)$. The spectral density of conductance fluctuations $S_G(f) \propto (D_n(1-D_n))^2 S_\varepsilon(f)$ where $S_\varepsilon(f) \propto 1/f$ is the energy fluctuation spectral density. For a DC voltage bias V_{ds} leading to a current $I = GV_{ds}$, the sample generates a current noise $S_I^{1/f} = S_G(f)V_{ds}^2$. This extrinsic current noise, not related to shot noise, is characterised by a quadratic variation with DC bias and has a $1/f$ frequency dependence. In 1990, Li et al. [28] reported 100 Hz to 100 kHz noise measurements at $T = 4$ Kelvin. They tried to estimate the white noise on top of a large $1/f$ conductance noise and they observed that this white noise was below the Schottky noise indicating a possible quantum reduction of shot noise. A similar piece of work, over the same frequency range but at lower temperatures ($T = 1.4$ K), was done in 1991 by Liefrink et al. [29]. Here, again the current noise was not due to shot noise but to conductance fluctuations. No $1/f$ noise was found but a Lorentzian frequency dependence was found which is characteristic of a two-level charged impurity, switching close to the QPC and modulating its conductance. To observe the white noise, and to reveal the fermionic nature of the ballistic electron shot noise, required either working at higher frequency or at much lower temperatures (to reduce impurity noise), or both. The first measurement of electronic quantum shot noise came with the work of Reznikov et al. [6]. Noise measurements on a QPC were done at a very high frequency, 8–18 GHz, at $T = 1.5$ K. Although the frequency was high, it was still lower than the temperature, as $k_BT/h = 31$ GHz. A voltage bias V_{ds} as high as 1 mV ($eV_{ds}/h = 241$ GHz) was used, so that low frequency shot noise formulae were still valid for comparison to the data. The linear variation of the noise with voltage bias (and thus with current) shown in the inset of Fig. 6.10a indicates that current shot noise and not extrinsic $1/f$ noise is measured. In the work of Reznikov et al., the current noise measurements definitely show a strong quantum reduction of shot noise on conductance plateaus supporting the predictions. This work was soon followed by another, more quantitative test of the scattering shot noise formula (6.30) by Kumar et al. [7]. Here, shot noise was measured at low frequency and a very low temperature $T \simeq 30$ mK. Indeed, a simple scaling argument shows that, for a given frequency, the $1/f$ noise decreases more rapidly with temperature than the shot noise does. For a given temperature T, to observe shot noise, it is enough to vary the voltage from $V_{ds} = k_BT/h$ to, say, $V_{ds} = 10\ k_BT/h$, so the voltage or the current, as well as the shot noise, scales linearly with

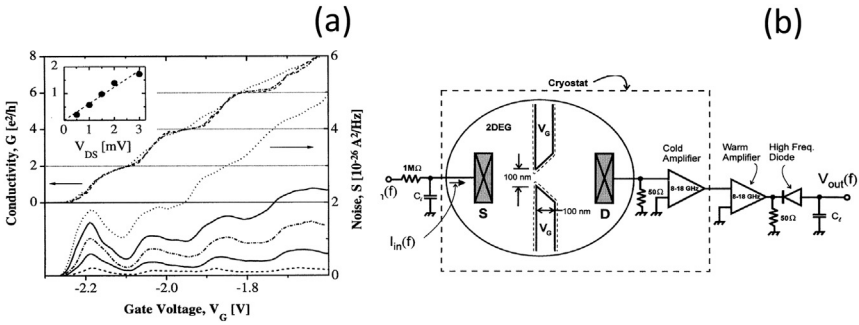


Figure 6.10 Quantum Point Contact shot noise: (a) Noise spectral density $S(v)$ and normalised linear conductance G versus gate voltage V_G. The noise is measured for $V_{ds} = 0.5$, 1, 1.5, 2 and 3 mV. Inset: Dependence of the first peak height (same scale as in main figure) on injection voltage V_{ds}. The dashed straight line is the predicted behaviour. The conductance is shown for $V_{ds} = 0.5$, 1.5 and 3 mV. (b) the high frequency set-up used for the measurements. Although high frequencies are used, the low frequency shot noise regime still holds as $f < k_B T/h$ and $< eV_{ds}/f$. *Adapted from Ref. [6].*

temperature. On the other hand, the $1/f$ current noise shows a different scaling. The $1/f$ conductance noise is known to roughly decrease linearly with temperature, whilst the current noise varies as I^2. Combining these two dependencies, one sees that the $1/f$ current noise scales as T^3. So reducing the temperature by a factor of 100 reduces the ratio of $1/f$ noise to shot noise by a factor of 10^4. The cross-over between white shot noise and $1/f$ noise, expected to be at a few 100 kHz in Li et al. experiment at $T = 4$ Kelvin, is expected to occur at a few 100 Hz at $T = 30$ mK.

Expressed in term of noise temperature $T^* = S_I/4Gk_B$, G being the sample conductance, the shot noise relation (6.30) is written as follows:

$$T^* = T\left(1 + F\left[\frac{eV_{ds}}{2k_BT}\coth\left(\frac{eV_{ds}}{2k_BT}\right) - 1\right]\right) \tag{6.56}$$

where the Fano factor is $F = 1 - D_1$ when only a single mode of transmission D_1 contributes to the QPC conductance. In Fig. 6.11a, Eq. (6.56) is reproduced by the shot noise data with no adjustable parameters for a transmission $D_1 = 0.50$ for various temperatures, and in (b) for the base temperature and various values of transmission D_1. In the temperature units chosen for shot noise and temperature, the slope of the linear variation provides a direct measurement of the Fano factor. Finally in Fig. 6.11c, the extracted values of the Fano factor are plotted for the first two QPC modes. An improved agreement is found if calculated heating effects are included or

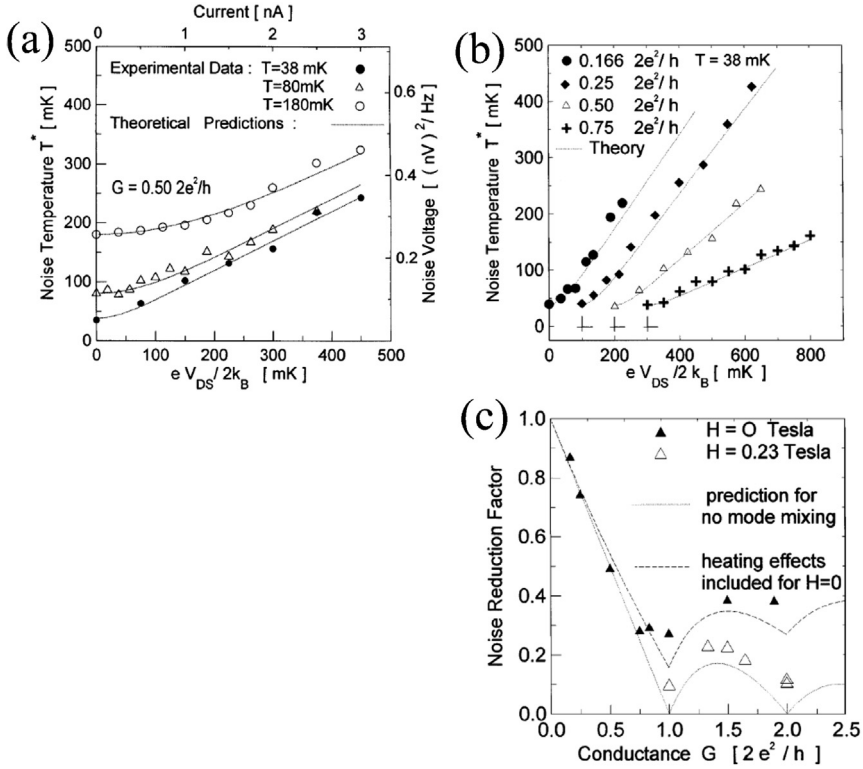


Figure 6.11 Quantum Point Contact shot noise: (a) Thermal to shot noise cross-over for transmission $D_1 = 0.5$, $D_{n>1} = 0$. The shot noise is expressed in terms of noise temperature $T^* = S_I/4Gk_B$. A smooth transition between thermal noise and shot noise occurs for $eV_{ds} = 2k_BT$. Solid curves are a comparison with predictions with no adjustable parameters. (b) Evolution of the shot noise for various transmission values D_1. Solid curves are a comparison with predictions with no adjustable parameters. For the present case, where only one channel is transmitted, the slope of the linear variation measures the Fano factor $F = 1 - D_1$. (c) Fano factor variations for the first two transmitted modes. Small heating effects can be reduced using a small magnetic field. *Adapted from Ref. [7].*

if a small magnetic field is used to suppress heating effects. The 90% suppression of shot noise on a conductance plateau gives a strong qualitative message: a perfectly conducting channel is noiseless as the Pauli exclusion principle regulates electrons.

The Fano factor of QPCs revealing the Fermi suppression of shot noise on conductance plateaus, as shown in Fig. 6.11c, has been further measured by several groups, improving the accuracy and the measurement techniques

[30–32]. It has also been observed in QPCs realised in high-mobility 2DEGs in InGaAs/InGaAsP heterojunctions by Ref. [37]. We will come back to QPCs later when considering the interaction leading to the 0.7 conductance anomaly and the high frequency and PASN.

Beyond the scope of semiconductor nanodevices, it is worth mentioning the case of Atomic Point Contacts (APCs). An atomic contact forms a 3-D constriction of atomic width between two metallic conductors. The electron elastic mean free path in good metals, being several tens of nm and the electron coherence length being even larger at low temperature, means the hierarchy of lengths is similar to that found in QPCs. The electron transport properties are expected to be similar, despite the lack of fine tuning of transmission. In APCs, the electronic transmission modes are replaced by atomic orbitals. Not surprisingly, quantum noise suppression similar to that observed in QPCs has been observed in APCs [38–42]. The knowledge of the sum $\sum_n D_n$ from conductance and of $\sum_n D_n(1-D_n)$ from noise gives important information to determine the orbital transmission D_n particularly for normal (nonsuperconducing) metals.

Another system showing a quantum suppression of noise due to Fermi statistics is the case of SWNTs. An SWNT can be seen as a graphene sheet rolled up to form a narrow tube of a few nm diameters. Because of the periodic boundary conditions imposed by the 2D graphene sheet, 1D sub-bands well separated in energy form so that at the Fermi energy a single electronic mode carries the current. This mode is fourfold degenerate with twofold valley degeneracy coming from the underlying graphene honeycomb carbon lattice structure and, in zero magnetic field, a twofold degeneracy due to spin. The conductance of a clean nanotube, with no impurity scattering, is therefore $4e^2/h$. Whilst clean SWNT can be grown with good yield using CVD methods, the difficulty is to realise good contacts with metallic electrodes. A good contact requires a perfect overlap between the evaporated metal and the nanotube. However, even if this is realised, the electronic wavefunction in the SWNT does not match the electronic wavefunction in the metallic contact as the Fermi wavelength and Fermi velocity are different. The mismatch induces finite reflection and a weaker transmission. However if, by chance, the amount of reflection is the same at both ends of the nanotube, the latter forms an electronic Fabry–Perot resonator which, at resonance, can show units of conductance $4e^2/h$. This unit of transmission should be confirmed by a quantum suppression of shot noise, this was demonstrated in Ref. [43] (Figs. 6.12 and 6.13).

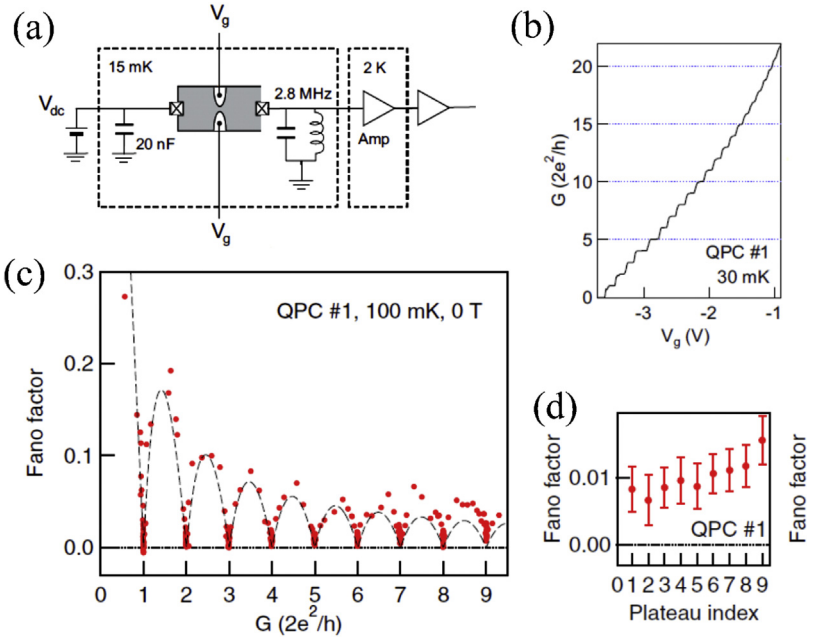


Figure 6.12 Quantum Point Contact shot noise: (a) measurement set-up using resonant circuit and cryogenic HEMT. (b) QPC conductance showing more than 15 accurately quantised plateaus. (c) Fano factor and (d) data showing near 99% suppression of shot noise on plateaus. Deviations from perfect noise suppression can be accounted for by thermal effects. *Adapted from Ref. [35].*

4.1.2 Shot noise and the 0.7 ($2e^2$/h) QPC anomaly

It may be surprising to observe that the noninteracting Landauer–Buttiker picture of quantum transport agrees so well with experimental data, whilst the correlation energy $V_c = e^2/4\pi\varepsilon a$ is comparable to the electron kinetic energy E_F, a being the mean distance between electrons. The reason is an effect of the Fermi statistics which, when the measurement energy ($\varepsilon = eV$ or k_BT) is small, suppresses a large number of interaction channels for electron–electron collisions. This has given rise to the concept of the Landau quasi-particle, which is a screened electron, i.e., a noninteracting Fermionic quasi-particle, which propagates freely during the time $\tau_{e-e} \propto (E_F/\varepsilon)^2$. One can say that the Landauer-Buttiker picture of quantum transport does not apply to electrons but to Landau quasi-particles. Indeed, by definition, the quantum coherence time is ultimately limited by the Landau quasi-particle lifetime for which inelastic electron–electron collisions set in. This picture works well except for one-dimensional conductors, for which

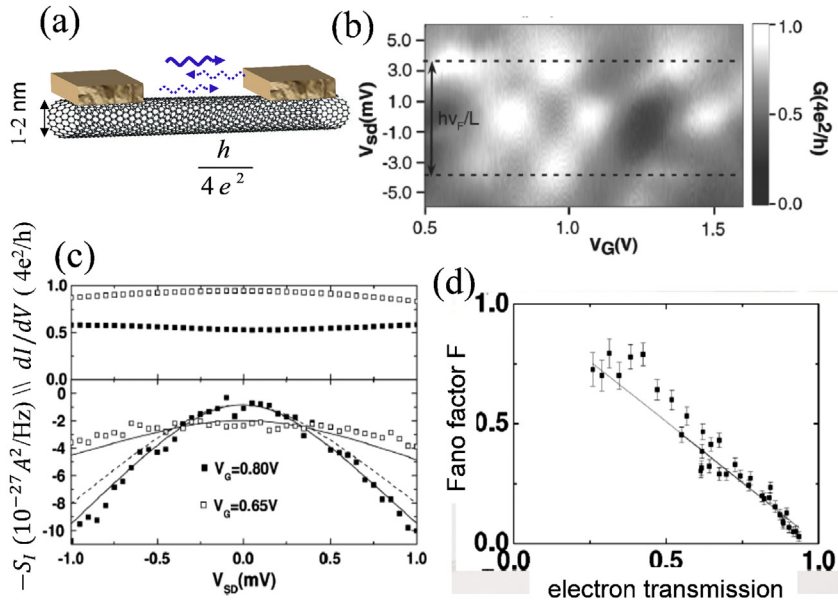


Figure 6.13 Single Wall Nanotube shot noise: (a) Schematic view of an SWNT with two metallic contacts evaporated on the top. The nanotube is deposited on a highly doped Si/SiO^2 substrate which is used as a gate to change de electron density; (b) Grey colour plot showing the Fabry–Perot-like conductance resonances versus gate voltage and source–drain voltage; (c) bottom: cross-correlation shot noise and (top) the associated measured conductance versus bias voltage. High conductance is associated with lower noise, manifesting the quantum suppression of shot noise; (d) Fano factor extracted from shot noise measurements. The 1-D variation versus transmission is in full agreement with quantum shot noise predictions. *Adapted from Ref. [43].*

the concept of Landau quasi-particle fails. For example, a long 1D wire, length L, like a carbon nanotube, is shown to be described by a Luttinger liquid for energies $\varepsilon > \hbar v_F/L$. It is not then surprising to find that for a QPC transmitting a single orbital mode, the absence of electronic screening restores the buried interaction effects. Indeed, interaction effects give rise to a weak plateau-like structure at a conductance $\simeq 0.7(2e^2/h)$. This is the so-called 0.7 QPC anomaly. Already observed, but not discussed, in one of the first observations of QPC conductance quantisation [8], the anomaly has been shown to be universal and was studied by Thomas et al. [44]. Applying a parallel magnetic field, Thomas et al. showed that the 0.7 anomaly may result from a finite spontaneous spin polarisation at zero magnetic field, probably due to exchange interaction effects. The 0.7 anomaly has also been observed in other short 1D structures like cleaved-edge-overgrowth

one-dimensional wires [45], ballistic hole quantum wires [70], Silicon quantum wires [46], QPCs made in GaN/AlGaN heterostructures [47] and in $In_{0.75}Ga_{0.25}As$ quantum wires [48]. Regarding theoretical understanding, there is currently no consensus. Possible scenarios are spontaneous spin polarisation [49], spin waves [50], acoustic phonons [51] and magnetic impurity formation [52] leading to Kondo-like effect [53–55]. Detailed experimental studies which aim to elucidate the nature of the anomaly are spin polarisation measurements [44,56], temperature studies [53,57], thermopower [58] and thermal conductance measurements [59].

Shot noise measurements have been done to probe the underlying spin degeneracy breaking by interactions. Indeed, for a QPC transmitting a single spin degenerate orbital mode, the spin up and spin down transmissions $D_\uparrow = D$ and $D_\downarrow = D_\uparrow = D$ contribute equally to the shot noise and conductance to give a Fano factor $F = 1 - D$ and a conductance $G = D(2e^2/h)$. If the spin degeneracy is broken, one gets

$$F = \frac{D_\uparrow\left(1 - D_\uparrow\right) + D_\downarrow\left(1 - D_\downarrow\right)}{D_\uparrow + D_\downarrow} \tag{6.57}$$

whilst the conductance is $G = (D_\uparrow + D_\downarrow)e^2/h$.

Measuring both conductance and shot noise provide a determination of both transmissions and a confirmation of broken spin degeneracy. The first experiment was done by Roche et al. [60]. As the 0.7 anomaly is a (relatively) high temperature physics effect, measurements were done in the temperature range of 300–900 mK requiring some thermal corrections to extract the Fano factor. A parallel magnetic field was added to increase the spin degeneracy and let the 0.7 plateau evolve to 0.5. More accurate measurements were later done by DiCarlo et al. [30]. Fig. 6.14 shows conductance and Fano factor measurements done in Refs. [60] and [61]. One observes a clear reduction of the Fano factor associated with the QPC conductance anomaly. Later measurements have been done by Nakamura et al. [62] and more recently by Jeong [63] (Fig. 6.14).

4.1.3 Low frequency shot noise in diffusive conductors

When considering a conductor of width W, which is large compared with the Fermi wavelength λ_F, the number of independent modes participating in conduction is large $\simeq (2W/\lambda_F)^{D-1}$ where D here is the spatial dimension of the conductor. When there is a large number of impurity scatterers yielding an elastic mean free path shorter than the conductor length, ab initio

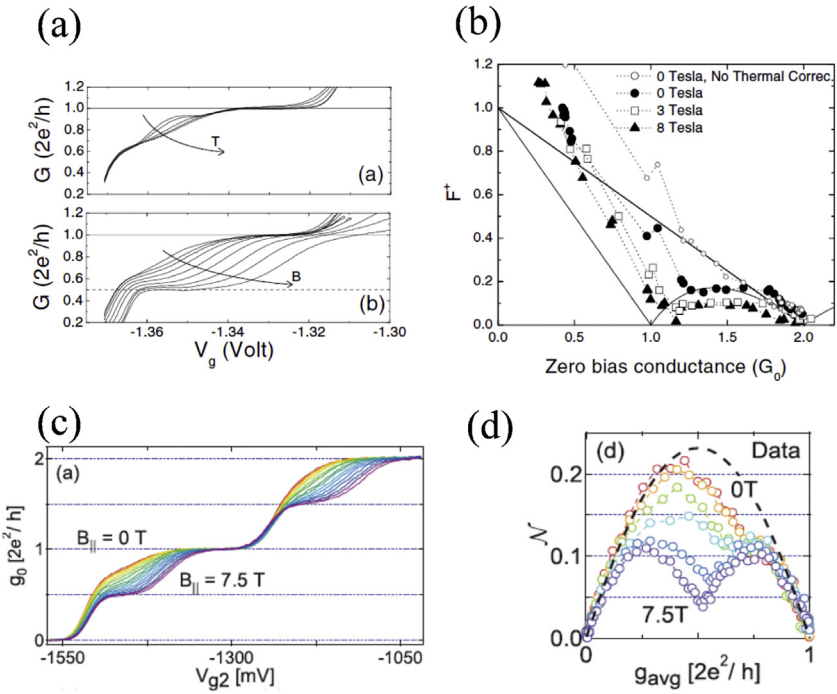


Figure 6.14 0.7 QPC anomaly probed by shot noise: (a) top graph: evolution of the QPC conductance with gate voltage for different temperatures ranging from 273 to 779 mK, bottom graph: QPC conductance for different parallel magnetic fields from 0 to 8 T; (b) Fano factor measured versus QPC conductance for various parallel magnetic fields. The location of the minimum Fano factor evolves from 0.7 to 0.5 when increasing the field. The solid lines indicate the expected Fano factor for full degeneracy limit and the full spin splitting limit; (c) conductance versus gate voltage for different magnetic fields; (d) The partition factor F $(D_\uparrow + D\downarrow)$ is plotted versus conductance for various magnetic fields. *Figures (a) and (b) are adapted from Ref. [60] and (c) and (d) from [61].*

prediction of the conductance is almost impossible. It is best to think that the distribution of transmission D_λ of electrons converges statistically towards the most probable distribution whose probability distribution is given by Eq. (6.12). This is expected to give a shot noise with mean Fano factor $F = 1/3$. This value was derived in a coherent regime using random matrix theory [13,16] and the same value was found beyond the coherent regime providing electronic transport is elastic [12,14]. Counter intuitively, the value $F = 1/3$ is expected to hold even if the mean transmission $\langle D_\lambda \rangle << 1$, a regime where we would have naively predicted pure Poissonian noise ($F = 1$), as for a tunnel junction. The reason is that, according to

Eq. (6.12), the distribution is bimodal. One can interpret the 1/3 value of the Fano factor as if two third of the modes transmit almost perfectively, whilst the remaining one third are transmitting very weakly, generating Poissonian noise.

A first experimental demonstration of a finite Fano factor with a value compatible with this picture was done by Liefrink et al. [64] using a narrow wire realised in a 2DEG in a GaAs/Ga(Al)As heterojunction. The wire of width $W = 0.5$ μm and length $L = 16$ μm was defined using lateral depletion with a negatively biased gate voltage. The bulk 2DEG 2 μm elastic mean free path, l_{el}., was reduced to $l_{eff.} \simeq 1.2$ μm because of wire boundary roughness. The estimated coherence length at the 4.2 K base temperature was shorter than the conductor length, but the expected inelastic length was much longer; data are shown in Fig. 6.15. More accurate measurements have been done using metallic wires by Steinbach et al. [65]. As was already noticed in the work by Kumar et al. [7], heating effects can give rise to an apparent shot noise which varies linearly with current (or bias voltage) at large bias, and the reason is due to electron heating. By increasing the bias voltage, the electron–electron interaction starts to rise, and over a characteristic electron–electron interaction length l_{e-e}, an electron temperature T_e can be defined. This electron temperature can be higher than the phonon temperature T imposed by the cooling system. Its local value is given by a balance between local Joule heating and electronic thermal conduction

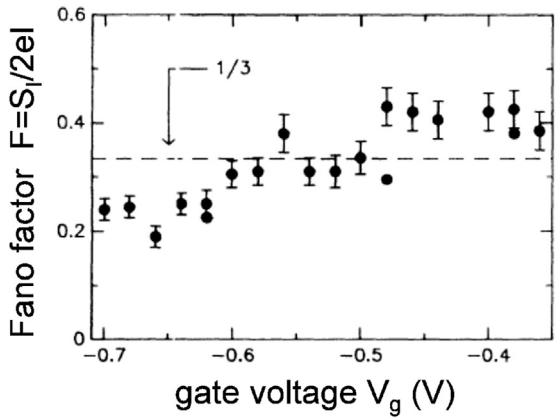


Figure 6.15 Fano factor observed for a narrow wire defined using gate depletion in a 2DEG in a GaAs/Ga(Al)As heterojunction. Varying the gate voltage makes the Fano factor fluctuate around the 1/3 expected value for the regime of incoherent diffusive electronic transport. *Adapted from Ref. [64].*

obeying the Wiedemann–Franz law, considering the large contacts as heat sinks thermalised to phonons. In a diffusive wire of length $L > l_{e-e}$, heating effects give rise to a high-voltage noise variation $S_I = 2eI\sqrt{3}/4$ and a pseudo Fano factor $\widehat{F} = \sqrt{3}/4$, whilst for $L < l_{e-e}$, the quantum Fano factor $F = 1/3$ is expected.

Measurements in the so-called interaction regime where electronic heating sets in were reported in Ref. [65]. They are shown in Fig. 6.16a. Later, similar measurements have been done by Henny et al. [66] in a regime where electron heating effects can be smaller and the Fano factor $F = 1/3$ can be recovered. The results are shown in Fig. 6.16b.

A Fano factor of $F = 1/3$ is also expected for the case of graphene which forms a gapless semiconductor whose conduction and valence bands are described by ultra-relativistic Dirac dispersion relations. Interestingly, for a graphene mono-layer with no disorder, the conduction, near the neutrality point, occurs via a large number of evanescent modes. The conductivity $\sigma = (L/W)G$, where L, W and G are, respectively, the sample length, width and conductance, is given by Ref. [67]:

$$\sigma = \frac{4e^2}{h}\frac{L}{W}\left(g_0 + \sum_{n=1}^{n=N>>1}\frac{1}{\cosh \pi n L/W^2}\right) \tag{6.58}$$

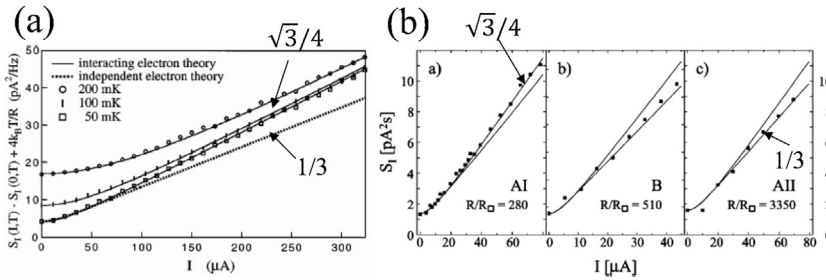


Figure 6.16 Fano factors observed for narrow metallic wires. (a) For a wire length of 30 μm the shorter electron–electron interaction length leads to electron heating. The balance between Joule heating and electronic conduction leads to a noise rising linearly with current, leading to the (apparent) Fano factor $\sqrt{3}/4$. The data accurately fit the prediction. (b) The noise measurements done on a metallic wire of 0.9 μm length show the importance of reservoirs whose resistance is as small as possible. By reducing the contact pad resistance versus the wire resistance per square (left to right graph), one goes from a regime where noise is dominated by electron–electron interaction (heating) to a regime where noise reveals the independent electron regime with Fano factor $F = 1/3$. *(a) Adapted from Ref. [65]. (b) Adapted from Ref. [66].*

where the first term g_0, which is 0 or 1/2 depending on boundary conditions, is negligible for large W/L. For a large sample, the second term is dominant and represent a sum of evanescent modes which all contribute significantly to the conductance, leading to the conductivity:

$$\sigma = \frac{4e^2}{\pi h} \tag{6.59}$$

Interestingly, although the set of evanescent mode transmissions $D_n = \frac{1}{\cosh\ \pi nL/W^2}$ is deterministic, its distribution $P(D_n)$ is identical to the probabilistic distribution of a disordered metallic conductor $P(D_\lambda)$ given by Eq. (6.12). Therefore, for a nondisordered graphene monolayer, one expects the same Fano factor $F = 1/3$ as for a disordered conductor. Indeed, the limit $F = 1/3$ is reached for $W/L > 3$, whilst F increases for a smaller ratio W/L. On the other hand, for large W/L, F is expected to decrease to $F \simeq 0.15$ when the Fermi energy moves away from the neutrality point by several tens of $v_{Fermi}h/W$ [67]. When adding disorder, theoretical predictions become difficult and one has to rely on numerical simulations. For example, in Ref. [68], it is found that, for moderate to strong disorder, the Fano factor is nearly constant. Its value is not universal and depends on the strength and the correlation length of disorder. Typically, F varies between 0.25 and 0.4 away from the neutrality point. In general, a smooth potential disorder tends to decrease F. However, when disorder is strong, it increases noise particularly close to the neutrality point as the modes are more evanescent leading to more Poissonian-like noise (Fig. 6.17).

Experiments has been done by several groups. DiCarlo et al. [69] performed shot noise measurements on a graphene monolayer with $W/L \simeq 5.7$ and found $0.35 < F < 0.38$. The Fano factor seems not to depend on the gate voltage (i.e., on the Fermi energy) as predicted for clean monolayer but more consistent with numerical simulations [68] for disordered conductors. Danneau et al. [70] on the contrary observe a Fano factor slightly above 1/3 at the neutrality point and strongly dependent on gate voltage, decreasing away from the neutrality point in qualitative agreement with clean graphene predictions of Ref. [67]. Later, Tan et al. [71] measured the shot noise of graphene nanoribbons (small W) and found a constant $F \simeq 0.4$ except very near the neutrality point where it increases to 0.7. Fig. 6.17 shows Fano factor measurements by DiCarlo et al. (a) and in (b) measurements by Danneau et al. We have included in (c) measurements done by Mostovov et al. [72], in the laboratory of the present Author, and which were taken on

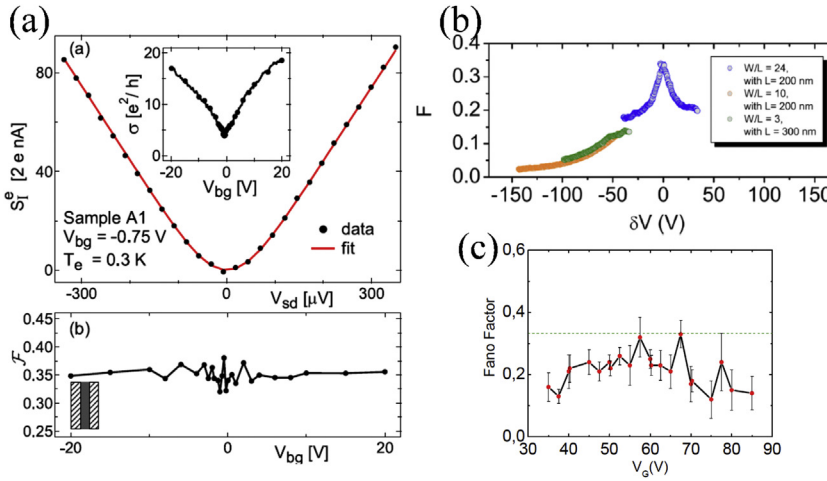


Figure 6.17 Fano factors observed in graphene monolayers. (a) Top graph: Shot noise versus bias voltage measured very close to the neutrality point. The inset shows the conductance variation versus the back-gate voltage used to change the density. Bottom graph: variation of the Fano factor versus gate, $W/L = 5.7$. (b) Fano factor versus back gate measured for several samples all designed with $W/L > 3$. The variations of F seem in qualitative agreement with clean graphene predictions [67]. (c) Fano factor observed for a narrow graphene constriction with $W/L = 4$. Here, side gates instead of back gates are used to vary the graphene density in the constriction where the Fano factor is measured. *(a) Adapted from Ref. [69]. (b) Adapted from Ref. [70]. (c) Data from [72].*

a 800 nm × 200 nm ($W/L = 4$) constriction patterned on a wide graphene monolayer. Here, a pair of side gates is used to modulate the charge density in the constriction. Fano factors between 0.2 and 0.3 are observed with a maximum value at the neutrality point.

To sum up, both theory and experiments confirm sub-Poissonian noise, the hallmark of electronic quantum shot noise but unlike for ordinary diffusive metal no universal Fano factor value is expected for disordered graphene monolayers.

4.1.4 Low frequency shot noise in large quantum dots viewed as quantum billiards

In this part, we address the case of another interesting quantum conductor whose electronic mode transmissions can be described by a statistical ensemble of transmissions whose most probable distribution $P(D_\lambda)$ is known and gives an average Fano factor $F = 1/4$.

Here, the quantum conductor forms a so-called quantum billiard. Usually realised in a 2D electron system formed in GaAs/GaAlAs heterojunctions, this semiconductor nanodevice is designed as a wide quantum dot connected to wide 2D electron leads by QPCs. It differs from a quantum dot (which have a much smaller area) by having a quasi-continuum of states, whilst a quantum dot shows discrete states; also the charging energy is negligible. The name *quantum billiard* means that the system forms the quantum analog of a classical billiard studied in the theory of classical chaotic systems. Let us start with classical billiards. When considering the classical motion of a particle at constant velocity in a closed area (like a ball on a billiard table), its trajectory is very rarely periodic. This is only for some limited billiard shapes, the circle, the square and some triangles, that trajectories are periodic. Introducing a small perturbation to the regular shape immediately leads to the appearance of a large set of chaotic trajectories. By chaotic, one means nonperiodic trajectories and more importantly that there is a significant difference between the two trajectories resulting from a tiny difference in the initial conditions. The difference between the two trajectories increases exponentially with the so-called Lyapunov exponent. When considering the system as quantum, the chaotic billiard differs from the nonchaotic billiard by its energy level distribution. Here, the quantum levels avoid each other, whilst in a nonchaotic system, the number of level crossings is large due to the existence of a mechanical invariant resulting from the separation of degrees of freedom. Here, we are not interested in the distribution of energy levels but in the distribution of the transmission eigenvalues. With performing two tiny holes in the billiard periphery to inject and collect particles which have travelled through the billiard, the predicted billiard transmission eigenvalues follow a bimodal statistical distribution. For two symmetric openings, transmitting the same number of modes $N_L = N_R = N$ [15,16] (Fig. 6.18).

$$P(D_\lambda) = \frac{N}{\pi} \frac{1}{\sqrt{D_\lambda(1 - D_\lambda)}} \tag{6.60}$$

This represents a bimodal transmission, very similar to that of diffusive systems, see Eq. (6.12). This gives an average conductance $G = (N/2)(2e^2/h)$ and a Fano factor $F = 1/4$, whilst 1/3 is found for diffusive systems. If the number of modes transmitted by the two apertures are different, one expects a Fano factor: $F = N_L N_R/(N_L + N_R)^2$. This has been experimentally measured by Oberholzer et al. [73,74] as shown in Fig. 6.18. They

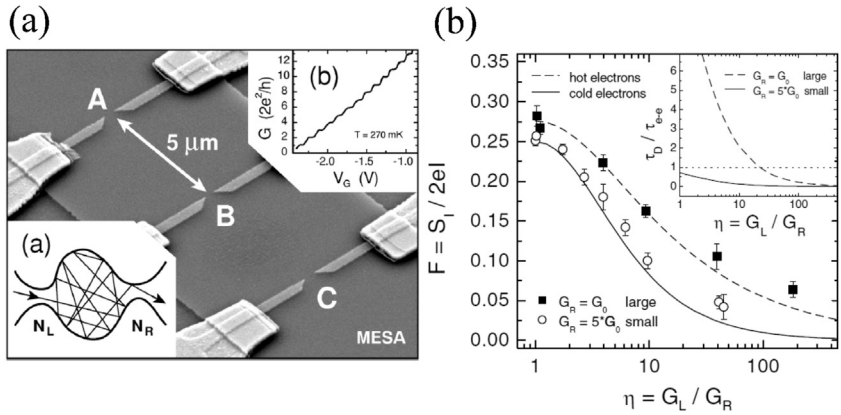


Figure 6.18 Shot noise in a chaotic quantum billiard. (a) sample: a large quantum dot formed in a 2DEG in GaAs/GaAlAs heterojunction is defined by etching and by negatively voltage biased split gates. Two Quantum point Contacts are used as input and output ports for studying the transmission through the quantum billiard (the third QPC is not used) (b) Variation of the measured Fano factor as a function of the asymmetry parameter characterising the ratio of the number of modes transmitted by the left and right QPCs. The solid curve is the prediction, open circles are data. For symmetrical transmission, the Fano factor is found to be 1/4. The dashed upper curve and filled square data points correspond to a regime where electron heating effects become important. *Adapted from Ref. [73].*

considered two regimes. The cold electron regime for which the electron dwell time in the chaotic cavity is smaller than the electron–electron interaction time and the hot electron regime for which electron self-heating inside the billiard must be included (Fig. 6.19).

4.1.5 Low frequency shot noise and Full Counting Statistics

Counting in real time the number of electrons that have passed through a quantum conductor is an extremely difficult task because of the lack of sensitive fast electron detectors. Indeed, for a spin degenerate single mode quantum conductor with transmission 0.5 and biased at a voltage 1 μV, the mean current is 0.4 nA corresponding to the transfer of a single electron every 0.4 ns, whilst the fastest detection is presently limited to a fraction of a μs. If this high-speed detection were possible, it would allow us to measure the FCS, of electrons, as was suggested in 1992 by Levitov and Lesovik for the binomial statistics of a QPC [75,77], see $P(N)$ given by Eq. (6.4). This also raises the theoretical issue of the quantum detection of current [76,77] at the single electron level [78].

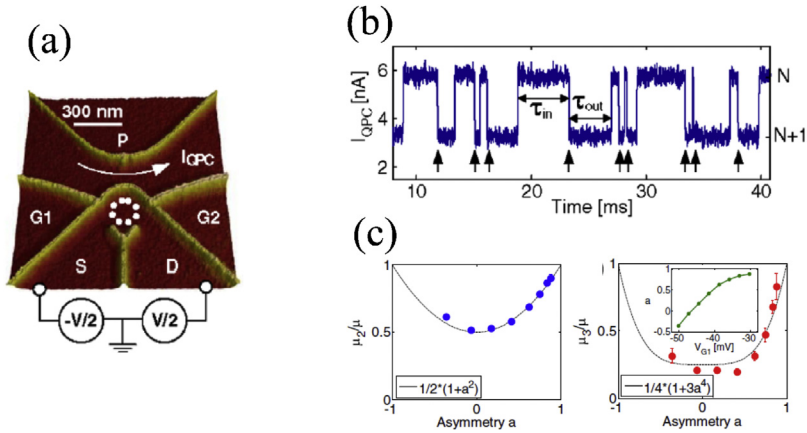


Figure 6.19 Full counting statistics: (a) a single quantum dot is used, whilst a QPC measures the charge in the dot. (b) real-time change of the electron number in the dot for a finite drain source current (c) left: inferred shot noise (the second moment) versus the asymmetry between left and right tunnel barriers, right: third moment of the current fluctuations. Solid line curves are predictions. *The figures are adapted from Ref. [83].*

Counting slow electrons is however possible and has been done by several groups using conductors made of a series of weakly coupled quantum dots. In this regime, electrons are transferred from a dot to the next dot every millisecond. Physically, this corresponds to a sequential tunnelling regime, described by classically combining the tunnelling quantum probabilities using a master equation as, at each step, the electrons lose their coherence. The measurement of the charge of a quantum dot is usually made using a close-by QPC. Indeed, the charge on the quantum dot modulates the transmission of the QPC as first shown by Field et al. [79]. Real-time fluctuations of the quantum dot charge were later measured [80–82] allowing the study of the charge fluctuations arising from a DC current passing through the dot. Later, Gustavsson [83] and Fujisawa [84] used quantum dots in series to measure the moments of the FCS distribution. In Ref. [83], see Fig. 6.19, the second and third moments of $P(N)$ were inferred from the distribution probability of the charge in the Quantum dot, whilst in Ref. [84], two quantum dots in series allowed bidirectional measurement of the current. The second moment of the distribution gives the shot noise $S_I = e^2(\delta N)^2/T_{meas}$. and the third order noise is $C_I = e^3(\delta N)^3/T_{meas}$. characterising the skewness of $P(N)$.

Other attempts to measure the FCS have been based on direct measurements of the current fluctuations. However, detection of the third moment is very difficult as it requires a very long time of acquisition to reach the appropriate noise sensitivity. Measurements have been made using tunnel junctions by Reulet et al. [85] and later by Bomze and Reznikov [86].

4.2 High frequency effect in shot noise

4.2.1 High frequency quantum suppression of shot noise in semiconductor nanodevices

Fermi statistics possesses several twists to silence electrons. As shown above in Part 2.2, it regulates the flow of electron at the frequency eV/h for a given drain–source bias voltage V. This leads to a noiseless electron flow in the absence of scattering (as is the case on a QPC conductance plateau). Here, we will show that, even in presence of scattering where we expect a stochastic transmission of electrons following binomial statistics, the noise is suppressed when the current fluctuations are observed at frequency $\nu > eV/h$. This occurs because an electron emitted in the energy range eV above the Fermi energy cannot find an empty state at energy $eV - h\nu$ to generate a current fluctuation when $h\nu > eV$. This finite frequency Fermi (or Pauli) blocking is what yields Eq. (6.35). Measuring high frequency current noise is like measuring the emission power of incoherent radiation in the measuring circuit. Also, the noise suppression can be alternatively interpreted as due to the fact that electrons do not have enough energy to emit a single photon when $eV < h\nu$.

Historically, the absence of shot noise at high frequency for voltage $V < h\nu/e$ was first observed in a metallic diffusive conductor by Schoelkopf et al. [87]. Then, Deblock et al. used two metallic superconducting–normal–superconducting junctions as quantum noise emitter and detector to measure the emission noise $S_I(\Omega)$ and absorption noise $S_I(-\Omega)$ [88] of electron quasi-particles, respectively, above and below the superconducting gap. Using semiconductor nanodevices, the high frequency noise of a QPC was later detected using a quantum dot by Onac et al. [89], and the energy level separation Δ of the quantum dot allowed the probing of the shot noise at frequency $\nu = \Delta/h$ via photon absorption photon emission processes. A full quantitative test of the finite frequency shot noise theory for a QPC using low noise cryogenic microwave amplifiers in the 4–8 GHz frequency range was done by Zakka-Bajjani et al. [90] (Fig. 6.20).

A similar clear quantum suppression of shot noise for voltages below the detection noise frequency has been also obtained by Gustavsson et al. [91]

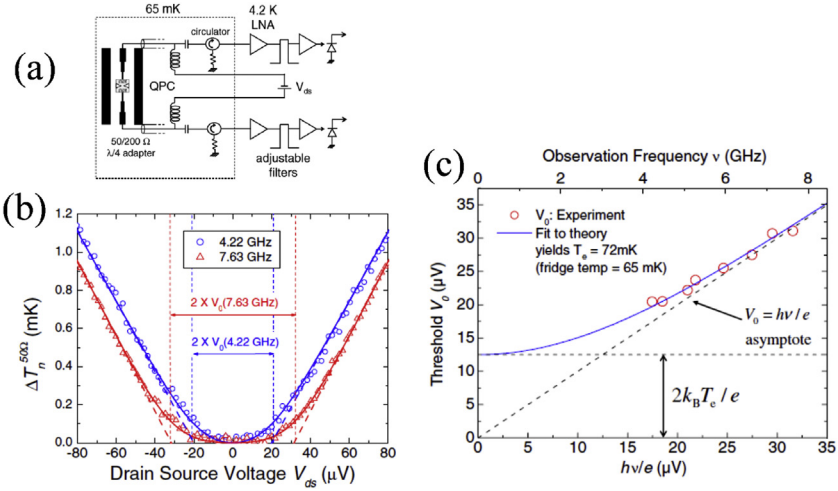


Figure 6.20 High frequency shot noise of a QPC. (a) high frequency measuring circuit. Two identical cryogenic Low Noise Amplifiers (LNAs) separately amplify the noise fluctuations. Microwave circulators are used to prevent the amplifier noise from reaching the sample and to ensure measurement of the QPC emission noise only. Because of the large difference between the sample impedance and the microwave impedance standard, a $\lambda/4$ 200Ω/50Ω coplanar strip line impedance transformer is used to improve the signal-to-noise ratio. (b) Shot noise, expressed in noise temperature units, measured versus bias voltage for QPC transmission 0.5 and two frequencies $\nu = 4.22$ and $\nu = 7.63$ GHz. Noise vanishes for a threshold voltage $\simeq h\nu/e$. The electronic temperature is 70 mK. The solid lines are a comparison with theory with no adjustable parameters. (c) evolution of the threshold voltage versus the noise detection frequency which scales as $h\nu/e$ including weak thermal corrections. *Adapted from Ref. [90].*

using a double quantum dot coupled to a QPC as a frequency selective detector for microwave radiation see Fig. 6.21. Thanks to the high quantum dot tunnel barriers, the slow electron tunnel rate allowed for real-time detection of the passage of electrons through the double dot. The photo-assisted electron flow being proportional to the QPC emission shot noise enabled its measurement. Tuning the double dot energy level allowed the selection of the frequency at which the shot noise power spectrum is detected.

4.2.2 Photo-assisted quantum shot noise

Here, we consider low frequency shot noise measurements performed in a nonequilibrium situation where both a DC and an AC bias voltage are applied to the contact of a semiconductor nanodevice: $V(t) = V_{dc} + V_{ac} \cos \Omega t$. In this case, it was shown in part 2.2.2 that the low frequency

shot noise is given by Eq. (6.41) (Fig. 6.21).which we recall here for convenience:

$$S_I^{PASN}(V_{dc}) = \sum_{l=-\infty}^{+\infty} P_l S_I^{DC}(V_{dc} - l\hbar\Omega/e) \tag{6.61}$$

Here, S_I^{DC} means the low frequency shot noise that one would measure when only the DC voltage bias is applied ($V_{ac} = 0$ and $P_l = \delta_{l,0}$). For a sinewave drive of finite amplitude V_{ac}, the photo-probabilities are given by $P_l = J_l(eV_{ac}/\hbar\Omega)^2$ where $J_l(x)$ is the integer Bessel function of order l. It is worth noting that Eq. (6.41) (or Eq. 6.61) takes this simple form only when assuming a negligible dependence on energy of the transmission. In the case of a quantum dot, the energy level separation gives a strong modulation of the transmission with energy and the PASN expression (6.41)

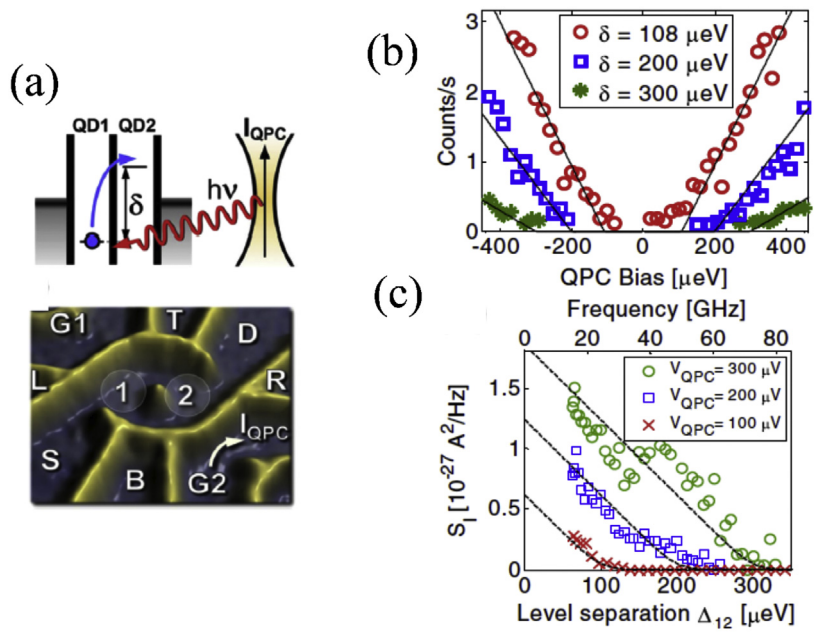


Figure 6.21 Quantum detection of high frequency QPC shot noise; (a) Top: Schematic energy diagram for operating a double quantum dot (DQD) as a high-frequency noise detector; Bottom: sample used in the measurement, with two QDs (marked by 1 and 2) and the nearby QP; (b) Electron count rate versus QPC bias for different values of detuning; (c) Noise spectrum of the QPC, extracted from the data in (b). The data show a clear suppression of quantum shot noise for bias below the detection energy set by the energy level separation. *Adapted from Ref. [91].*

needs to be reconsidered. Here, we concentrate on examples of quantum shot noise measurements in semiconductor nanodevices for which the energy dependence is considered to be weak.

Historically, PASN was first observed in diffusive metallic wires by Schoelkopf et al. [92] and in a metallic superconducting normal junction by Kozhevnikov et al. [93]. The first observation in a semiconductor nanodevice, a QPC, was done by Reydellet et al. [94]. The QPC allowed the tuning of the transmission to allow a more accurate test of PASN theory including heating effects. The results of Reydellet et al. [94] are shown in Fig. 6.22. Here, the noise is expressed in units of noise temperature defined as $S_I^{PASN}/4k_BG$ where $G = 2e^2/h\sum_n D_n$ is the QPC conductance and D_n the QPC mode transmissions. From Eq. (6.61), the noise temperature is given by the following (Fig. 6.22):

$$T_N = T\left(J_0(\alpha)^2 + \frac{\sum_n D_n^2}{\sum_n D_n}\left[1 - J_0^2(\alpha)\right]\right) + \sum_{l=1}^{+\infty} \frac{l h\nu}{k_B} J_l^2(\alpha) \frac{\sum_n D_n(1-D_n)}{\sum_n D_n} \tag{6.62}$$

with $\nu = \Omega/2\pi$ and $a = eV_{ac}/h\nu$.

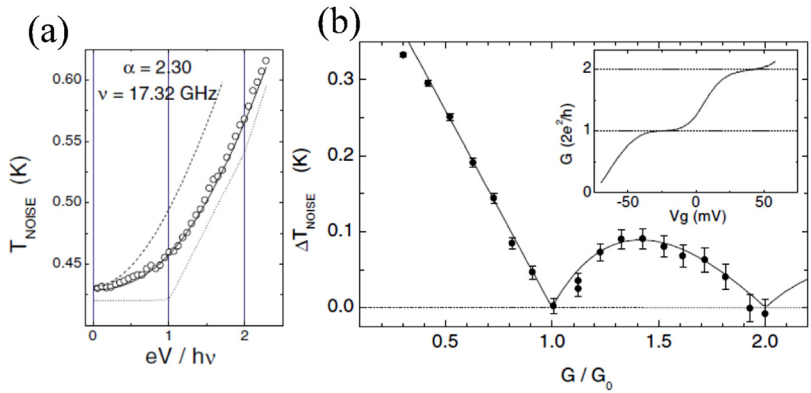


Figure 6.22 Photo-assisted shot noise of a QPC. (a) Noise temperature as a function of DC bias voltage in units $eV/h\nu$ with 17.32 GHz ac excitation. The measured electron temperature was 229 mK and the AC excitation $\alpha = 2.3$. The dotted line is the expected photon-assisted noise at zero temperature with $\alpha = eV_{ac}/h\nu = 2.3$. For comparison with pure thermal effect, the dashed line is the nonphoton-assisted shot noise with an electron temperature $T = 430$ mK which equals the zero bias noise temperature. The continuous line is the photon-assisted noise calculated using Eq. (6.61) with no adjustable parameters. (b) Zero voltage bias photo-assisted shot noise, expressed in noise temperature units, versus the QPC conductance in unit of $G_0 = 2e^2/h$ when applying a 17.32 GHz ac excitation with $\alpha = 2.3$. The solid line is Eq. 6.62 with no adjustable parameters. *Adapted from Ref. [94].*

In the experiment of Ref. [94], the ratio $h\nu/k_BT$ was not high and the zero bias DC shot noise singularity, expected to be replicated when $V = h\nu/e$, gave thermally broad singularities in the PASN measured. Sharper features have been observed later using a tunnel junction by Gabelli et al. [95] and at lower temperature using a QPC by Dubois et al. [96] to compare sinewave excitations with periodic Lorentzian excitation in the context of levitons.

The PASN can also be used as a detector of microwave radiation. Jompol et al. [97] used the incoherent radiation emitted by a QPC generating shot noise to generate a PASN on a nearby QPC. Indeed, relation (6.61) is not restricted to monochromatic excitation. Starting with the monochromatic excitation case, one can re-write the set of discrete photo-absorption probabilities as a microwave energy distribution $P(E) = \sum_{l=-\infty}^{l=+\infty} \delta(E - lh\nu)$, and Eq. (6.61) can be equivalently re-written by replacing the discrete sums over l by an integral over E, i.e., $\sum_l P_l \rightarrow \int P(E)dE$. In the case where the excitation is not monochromatic but incoherent radiation, the phase given by Eq. (6.37) now acquires a random time variation and

$$P(E) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} d\tau e^{iE\tau/\hbar} \left\langle e^{i(\phi(\tau)-\phi(0))} \right\rangle \tag{6.63}$$

where the bracket means a statistical average over phase correlation fluctuations. Using cross-correlation noise measurement, Ref. [97] showed that the amplitude of photo-assisted noise generated by the QPC detector is proportional to the noise temperature of the QPC emitter and is thus able to measure the incoherent radiation power emitted by the QPC. The measurements and semiconductor nanostructure used are shown in Fig. 6.23.

PASN has been recently used for THz radiation detection. In Ref. [98], Parmentier et al. illuminated a graphene nanoribbon placed at the centre of a bow-tie antenna by a frequency swept THz source and recorded the resulting PASN. To analyse data, a Fano factor $F = 1/3$ was used to account for the diffusive quantum transport of the graphene nanoribbon and heating effects were included. Fig. 6.24 shows an example of the noise recorded whilst sweeping the THz frequency from $\simeq$ 50 GHz to 1.5 THz. The oscillations of power with frequency are related to bow-tie antenna resonances (Fig. 6.24).

Finally, the physics of PASN has been used to provide a new type of measurement of fractional charges in the Fractional Quantum Hall Effect (FQHE) regime. A shot noise singularity is observed [99] when the DC

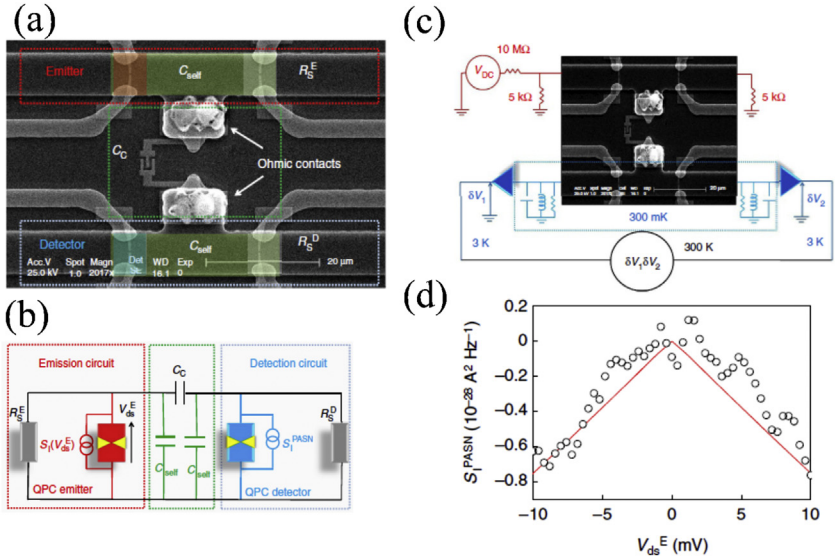


Figure 6.23 (a) SEM image of the sample. Four QPCs are used on two disconnected 2DEG strips. The left QPC on the upper (lower) strip is the QPC Emitter (Detector). The QPCs on the right are operated on conductance plateaus to serve as noiseless resistors. The two floating Ohmic contacts in the centre are used to transfer, by capacitive coupling, the potential fluctuations originating from the QPC Emitter shot noise to potential fluctuations generating photo-assisted shot noise in the QPC Detector. (b) measurement principle and (c) measurement circuit. (d) cross-correlated photo-assisted shot noise detected as a function of the voltage bias of the QPC emitter. *Adapted from Ref. [97].*

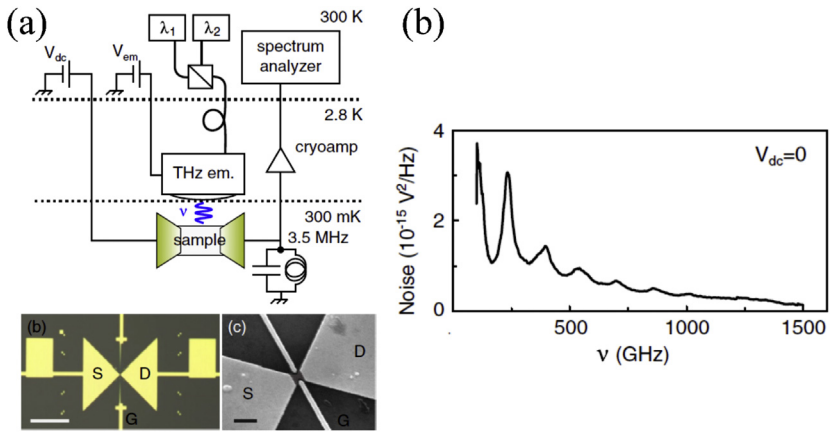


Figure 6.24 THz radiation detection using photo-assisted shot noise. (a) principle of the experiment. The frequency swept THz source is based on the beating of two solid state Laser diode frequencies (wavelength λ_1 and λ_1). An optical fiber brings the radiation to a cryogenic low-temperature-grown GaAs mixer which emits a THz radiation of frequency $\nu = C(\lambda_1^{-1} - \lambda_2^{-1})$ towards the bow-tie antenna of the sample. The images show the bow-tie antenna (left) and the graphene nanoribbon at the centre of the antenna (right). (b) Photo-assisted shot noise detected versus the radiation frequency. *Adapted from [98].*

voltage obeys the Josephson relation $e^*V = hf$ for $e^* = e/3$ and $e/5$, see Section 4.3. For single electron sources based on voltage pulses, which have been developed in electron quantum optics, the physics of PASN provides a good framework to understand the periodic generation of levitons which are time resolved minimal-excitation quasi-particles [96]. This will be discussed in Section 4.4.

4.3 Fractional Quantum Hall Effect: shot noise measurements of fractional charges

Quantum shot noise measures the variance of the fluctuations of carriers arriving in a given contact. From a quantum point of view, this gives information on the particle-like nature of the carriers as opposed to information on the wave nature of carriers provided by conductance. For given statistics, the intensity of noise is proportional to carrier charge (the common picture is to think of the noise produced by rain on a roof. The larger the water droplets, the stronger the noise). To understand this, the simplest situation to consider is when the passage of the charge carriers through the conductor follows Poissonian statistics. Following the approach of Section 2.1, let us consider the average number $\overline{N}$ of charge q transferred during a time τ. The variance is $(\Delta\overline{N})^2 = \overline{N}$ where $\Delta N = N - \overline{N}$ is a fluctuation. Restoring electrical units, the mean current and the variance of its fluctuations are, respectively, $\overline{I} = q\overline{N}/\tau$ and $(\Delta\overline{I})^2 = q^2(\Delta\overline{N^2})/\tau^2$. This gives the following:

$$\left(\Delta\overline{I}\right)^2 = 2q\overline{I}\Delta f = S_I\Delta f \tag{6.64}$$

where the Nyquist bandwidth $\Delta f = 1/2\tau$ has been introduced. For binomial statistics corresponding to finite transmission D, the noise is reduced according to Eq. (6.7) ($q = e$). One sees from the above relation 6.64 that measuring Poissonian shot noise and mean current provides the most accurate way to determine the carrier charge. We will show a direct application in the Quantum Hall regime to determine the fractional charge of anyonic excitations.

4.3.1 Edge channels and shot noise in the fractional quantum Hall regime

Before addressing shot noise, let us briefly explain what the Quantum Hall Effect regime is and how shot noise can be used to probe this regime.

When an electron is placed in a magnetic field B, it is well known that it performs a circular motion, called a cyclotron orbit, in the plane perpendicular to the magnetic field. In two dimensions, the motion is described by a harmonic oscillator-like Hamiltonian which gives rise to regularly spaced energy levels $E_n = (n - 1/2)\hbar\Omega_c$, the so-called Landau levels. $\Omega_c = eB/m$ is the cyclotron frequency for charge e and mass m. Let B be parallel to the Z-axis and (x, y) the coordinates in the plane. In a uniform plane, the cyclotron orbit energy is the same for all possible locations of the centre (X_c, Y_c) of the cyclotron orbits in the plane. This gives rise to a large degeneracy of the Landau levels. Indeed, in two dimensions, one expects that the energy depends on two quantum numbers, whilst for Landau levels, it depends on the single quantum number n. The missing quantum number is encoded by the choice to define the centre of the cyclotron orbits. This choice is not infinite as, from the equations of motion, one can show that the coordinates of the centre of cyclotron orbit obey the uncertainty relation

$$\Delta X_c \Delta Y_c = \hbar/eB \tag{6.65}$$

This resembles the uncertainty relation $\Delta p \Delta q = \hbar$ for one-dimensional motion with coordinate q and momentum p. Indeed, as one degree of freedom is frozen in the cyclotron energy, the dynamic left is of one-dimensional type. One can say that, within the Landau level n, the real space becomes a phase space with a magnetic field-tunable Planck constant $\hbar/eB$. In the same way that in one dimension the mean area occupied by a quantum state in the phase space is h; the area of a quantum state within a Landau level is h/eB. Thus, for an area S, the degeneracy is equal to the number $N = BS/\phi_0$ of flux quantum $\phi_0 = h/e$ in the plane. Introducing Fermi statistics, one sees that full occupation of *nth* Landau levels is obtained when the electron density n_s corresponds to an integer filling factor $\nu = n_s/n_\phi = n$ where $n_\phi = eB/h$ is the density of flux quantum in a plane (typically, $n_\phi = 4.13\ 10^{15}\mathrm{m}^{-2}$ for B = 1 T). The combination of Landau level quantisation and Fermi statistics gives rise to a spectacular quantisation of the Hall resistance, whose classical value $R_H = B/en_s$ becomes in quantum units $R_H = (h/e^2)(n_\phi/n_s)$. For n filled Landau levels [100] (Fig. 6.25).

$$R_H = \frac{1}{n}\frac{h}{e^2} \tag{6.66}$$

According to the above analogy between the dynamics within a Landau level and that of a one-dimensional conductor, the contribution e^2/h of each Landau level to the Hall conductance is not surprising, but the striking

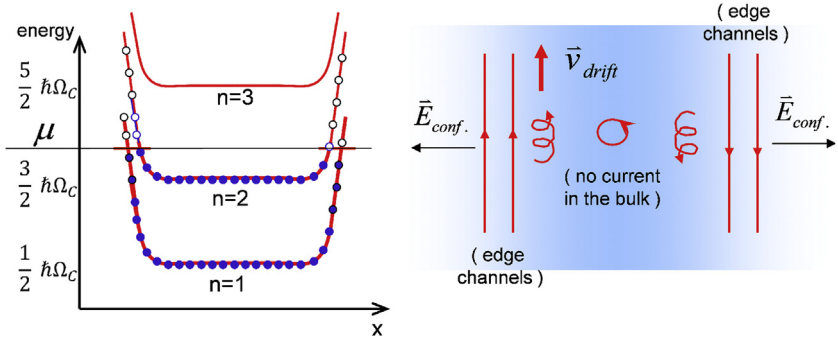


Figure 6.25 Sketch of a strip of a 2D electron gas in a strong magnetic field. Left: the Landau level energy is bend by the confining potential on the strip edges. When the level crosses the chemical potential, this defines gapless conduction modes called edge channels. For n filled Landau levels in the bulk, there are n edge channels, $n = 2$ here. Right: sketch of the edge channels. As the drift velocity of the cyclotron orbits changes sign with the confining electric field, electrons on edge channels are forced into chiral motion. For a wide channel, no electron backscattering is possible ensuring perfect conductance quantisation.

difference is the extreme robustness of the quantisation. The reason is the absence of electron backscattering. This raises the question of finding where the current flows in the Quantum Hall regime. Paradoxically, electrons in a Landau level are not moving as the centre of the cyclotron orbits is frozen. The whole system is an insulator. However, if an electric field E_y is applied, say in the Y direction, this induces a cyclotron orbit drift velocity $v_D = E_y/B$ in the X direction. Now, if one considers a finite strip, parallel to the X axis as shown in Fig. 6.25, the necessary sharply rising confining electric field perpendicular to the strip edge makes electron drift. This static field generates a permanent current and, as electric fields are opposite, the top and bottom edge currents have opposite directions. The confining potential lifts the Landau levels energy. Their crossing with the Fermi energy defines edge channels from which current can flow through the system. The large distance between opposite edges prevents backscattering, ensuring perfect conductance. To make a connection with recent developments, the Quantum Hall effect is the first example of a topological insulator giving rise to edge states with quantised conductance. An important feature to note is the incompressibility of the Quantum Hall effect for integer filling factor (in the bulk). Indeed, because of Fermi statistics, one cannot change the density without promoting an electron from the last occupied Landau level to the next unoccupied Landau Level, leaving a hole free to move elsewhere.

This costs a finite energy ($\overline{\Omega}_c$ for even ν or $g^*\mu_B B$ for odd ν) making the system incompressible, g^* is a Landé factor (Fig. 6.26).

Now that we know how and where the current flows in the Quantum Hall Effect regime, we are ready to consider shot noise measurements. Fig. 6.26a shows schematically two limiting cases which happen when using the electric field created by the split gate of a QPC to force electron backscattering between opposite edge channels of a quantum Hall conductor strip. When the QPC creates an insulating barrier between the left and right parts of the strip, the conductance, measured when collecting the transmitted current I_t at the upper right contact whilst applying a bias voltage V_{ds} between the upper left contact and the upper right contact, is low, $= De^2/h$, as $D << 1$. In this Strong Backscattering regime (SB),

$$S_I = 2eI_t \tag{6.67}$$

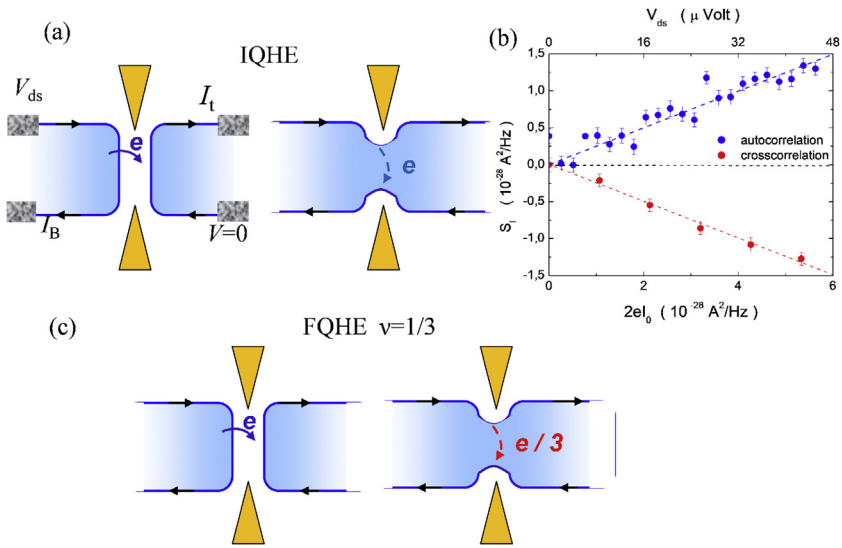


Figure 6.26 (a) strong backscattering and weak backscattering in the Integer Quantum Hall regime. A left to right transfer of electrons with Poisson statistics occurs for transmission $D << 1$. For $1 - D << 1$, a top to bottom transfer of electrons with Poisson statistic occurs; (b) The chirality of edge states allows the performance both auto-correlation (blue [dark grey in print version] data) and cross correlation (red [grey in print version] data). One observes that $(\Delta I_t)^2 = \overline{-\Delta I_t \cdot \Delta I_B}$. One can note that cross-correlation data are less noisy and more accurate than auto-correlation measurements. (c) Weak and strong backscattering in the Fractional Quantum Hall regime. For weak backscattering, the top to bottom transfer of carriers occurs with $e/3$ carriers which are the only low energy excitations expected to be transferred through the $\nu = 1/3$ correlated quantum Hall system.

One can find however another regime of shot noise with Poissonian statistics when creating the condition for very Weak Backscattering (WB) ($D \simeq 1$). Let $I_0 = (e^2/h)V_{ds}$ be the incoming current injected by the upper left contact. For WB, the transmitted current is large, $I_t = DI_0 \simeq I_0$, whilst the backscattered current $I_B = (1-D)I_0 << I_0$ is weak. It corresponds to rarely reflected electrons with Poissonian statistics. Equivalently, I is made of rarely transmitted holes (missing electrons) with Poissonian statistics.

$$S_I = 2eI_B \tag{6.68}$$

The separation between transmitted and reflected electrons is possible, thanks to the chiral nature of transport. This allows the performance of auto and cross-correlation current noise measurements as shown in Fig. 6.26b.

Now let us consider the case where electrons do not entirely fill a Landau level. For partial filling, one expects energy gapless excitations. However, it was found that for certain rational fractions of the filling factor a gap occurs making the system incompressible and the Hall conductance fractionally quantised. This is the FQHE [101]. The most robust and best understood case is for a one-third filling of the first Landau level. The FQHE is due to electron–electron interaction. For filling factor $\nu = 1/3$, a topologically ordered quantum phase is favoured in which each electron jealously occupies three quantum states (i.e., the area of three flux quanta). In the same way, the Pauli exclusion at $\nu = 1$ reduces the repulsive Coulomb interaction by preventing electron to approach each other, here electrons collectively cooperate, respecting a three flux quanta social distancing to reduce interaction and stabilise their ground state. An elementary hole-like excitation consists in emptying a quantum state, leaving a hole with fractional charge $-e/3$ [102]. This has an energy cost $\Delta \simeq e^2/4\pi\varepsilon\varepsilon_0 l_c$ where $l_c = (\hbar/eB)^{1/2}$ which roughly corresponds to the Coulomb energy to create a disk of charge $-e/3$. The gap makes the Hall system incompressible and the Hall resistance is quantised as $3h/e^2$.

How to detect and measure the fractional charges $e^* = e/3$? The shot noise given by Eq. 6.64 with $q = e^*$ provides an accurate way, free of artefacts and of unknown parameters. As schematically shown in Fig. 6.26c, one has to distinguish between two regimes. In the SB regime, the QPC creates an insulating barrier. The insulating region destroys the collective correlations and it is unlikely that fractional excitations will cross the barrier (in vacuum, general principles forbid the possibility of nonelementary charge).

Then in this regime, one expects $S_I = 2eI_t$. In the other limit, however, the very weak perturbation produced by the coupling does not destroy correlation whilst the only excitations allowed to cross the incompressible fluid have fractional charge $q = e^* = e/3$. One expects (Fig. 6.27).

$$S_I = 2e^* I_B \left(\coth \frac{e^* V_{ds}}{2k_B T} - \frac{2k_B T}{e^* V_{ds}} \right) \tag{6.69}$$

This relation has been first derived by Kane et al. in Ref. [103] in the frame of chiral Luttinger liquid theory. It is equivalent to relation 6.68 but with $e \rightarrow e/3$ and thermal effects included. Indeed, as shown by Wen in Ref. [130], there is a deep connection between fractional edge channels and the concept of Tomonaga–Luttinger liquids [104,105] which occurs in one-dimensional systems when the interaction is not negligible. In a classical hydrodynamic approach and for a $\nu = 1/(2m + 1)$ Laughlin state, Wen has considered the periphery deformations of the 2D quantum Hall conductor which preserves the total area (like a 2D droplet of an ordinary incompressible liquid). He showed that the edge channel dynamics is that of chiral Luttinger Liquids. When backscattering occurs, the conductance shows a power law variation with temperature and voltage V_{ds}. The weakest backscattering

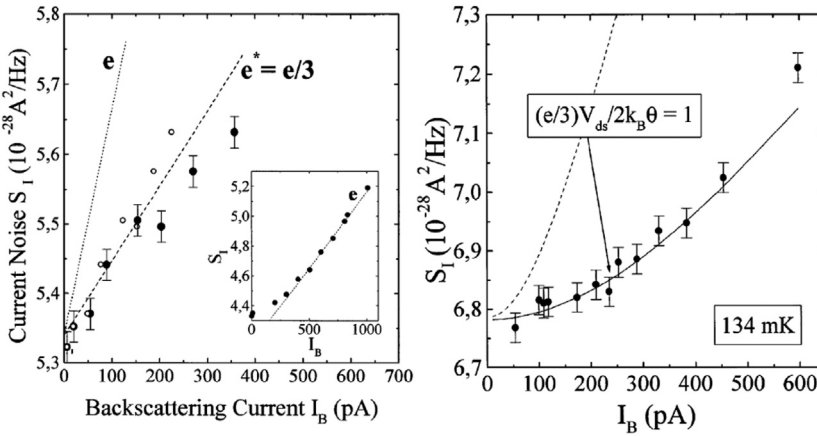


Figure 6.27 Shot noise measurements made in the weak backscattering regime for $\nu = 1/3$ by Saminadayar et al. [108]. Left graph: linear shot noise variation versus backscattered current. The electronic temperature is 42 mK and the reflection probability is kept low $1 - D \simeq 4\%$. Right graph: shot noise for higher temperature: 134 mK, in order to observe the thermal to shot noise cross-over occurring when $e * V_{ds} \simeq 2k_B T$. Here, $1 - D \simeq 0.18$. The solid curve is a comparison to Eq. 6.69 with no adjustable parameters.

makes the system evolve into an effective strong backscattering regime at the lowest temperature and voltage. For a large bias, Kane et al. showed that WB Poissonian noise leads to Eq. 6.69, whilst for low bias, the SB regime leads to $q = e$ and noise given by Eq. 6.67, including thermal effects. These two limiting cases were treated perturbatively [103,106]; however, an exact solution for intermediate coupling interpolating between WB and SB regimes has been derived by Fendley et al. and can be found in Ref. [107].

4.3.2 Low frequency DC shot noise measurements in the fractional quantum Hall regime

Two groups have measured shot noise in the WB regime. Saminadayar et al. made cross-correlation noise measurements [108]. A wide QPC was used to realise a well-defined reconstructed filling factor 1/3 at the centre of the QPC, whilst away from de QPC, the bulk filling factor was 2/3. De Picciotto et al. performed auto-correlation noise measurements on a sample with bulk filling factor 1/3 and used a narrow QPC to induce WB [26]. Fig. 6.27 shows the experimental data of ref. [108]. At the lowest temperature (42 mK), a linear variation of the noise with current I_B gives the expected charge $e/3$. The figure on the right shows the thermal to shot noise cross-over for a larger temperature. The cross over is found to occur when $(e/3)V_{ds} = 2k_BT$ as expected from Eq. 6.69. The solid curve is a comparison to data with no adjustable parameters.

Later, the shot noise has been used to determine the charge $e^* = 1/5$ at filling factor $\nu = 2/5$ by Reznikov et al. [109]. For this filling factor, two co-propagating edge channels are expected. As shown in Eq. 6.69, when sweeping the gate voltage to more negative values, the conductance, measured through the QPC, starts with the quantised value $0.4e^2/h$ and then decreases signaling the reflection of a first 2/5 edge channel. This ends by the appearance of a second plateau at $0.333\ e^2/h$ and finally the conductance decreases all the way to zero as the last channel is reflected. To analyse the data in the WB regime of the first fractional edge channel, the transmission is defined as $D = (g - 1/3)/(2/5 - 1/3)$ with $g = Gh/e^2$, the reduced conductance. The backscattered current entering in Eq. 6.69 is then $I_B = (1 - \mathrm{D})\ (2/5 - 1/3)\ (e^2/V_{ds})$. With this choice of transmission, the shot noise data measures a charge $e^* = e/5$.

The measurements of charge $e/3$ and $e/5$, starting from the $\nu = 2/5$ quantum Hall state, have been recently made using PASN [99]. Not only the low frequency shot noise similar to those performed in Refs. [26,108,109] was reproduced, but the shot noise measured under

microwave irradiation showed singularity at Josephson frequency $f = e^*V/h$ for both $e^* = e/3$ and $e/5$. This will be discussed later in section 2.4.3.3.

An intriguing question is the evolution of the quasi-particle charge with transmission. Indeed, the chiral Luttinger theory predicts the two limiting cases at $\nu = 1/3$: charge $e^* = e/3$ for WB and $e^* = e$ for strong backscattering SB. The zero temperature exact theory of Ref. [107] which describes the transition from WB to SB has been extended to finite temperature [110], allowing a better comparison of Luttinger approaches to experiments [111]. For example, in the SB regime, the chiral Luttinger perturbative theory predicts a strong nonlinear variation of the QPC differential conductance with temperature $dI/dV(V_{ds} = 0, T) \propto (T/T_B)^4$ and with voltage $dI/dV(V_{ds} = V_{ds}, T = 0) \propto (V_{ds}/V_B)^4$. This is hardly observed in experiments and most experiments observe a smoother nonlinearity (almost quadratic). Here, $T_B \simeq eV_B/k_B$ is the strength of the coupling which depends on the QPC gate voltage. A careful look at the finite temperature exact solution of Ref. [110] is very instructive. In Ref. [112], it was shown that in the typical temperature range used in experiments, an effective power law with exponent ranging from $\simeq 1.8$ to $\simeq 3$ is expected, a result in much better agreement with experiments [111,112]. This is shown in Fig. 6.29. One observes that this is only in the extremely SB limit where the zero voltage conductance is < 0.001 $(e^2/3h)$ that the conductance exponent with voltage approaches the value 4 corresponding to the perturbative limit so often mentioned in the literature. If experiments hardly reach the extreme SB regime for which charge $e^* = e$ is expected, this raises the question of how the charge given by shot noise evolves from $e/3$ to e. Indeed, as the Luttinger approach is too complex, experimentalists use a simplified, and thus approximate, description of shot noise based on the noninteracting shot noise theory. In Ref. [113], the following noninteracting shot noise formula is used (Fig. 6.28).

$$S_I = 2e^* I \frac{1}{N} \sum_{1}^{N)} \left(1 - \frac{g/g_0}{e^*/e)}\right) \left[\coth\left(\frac{e^* V}{2k_B T}\right) - \frac{k_B T}{e^* V}\right] \tag{6.70}$$

In this expression, g/g_0 is the differential conductance in units of the conductance quantum, and the effective charge e^* is a free parameter to determine by the best fit with shot noise data. The sum over N represents a mean Fano factor resulting from the integration over N points of the dI/dV characteristics up to the current I (Fig. 6.29).

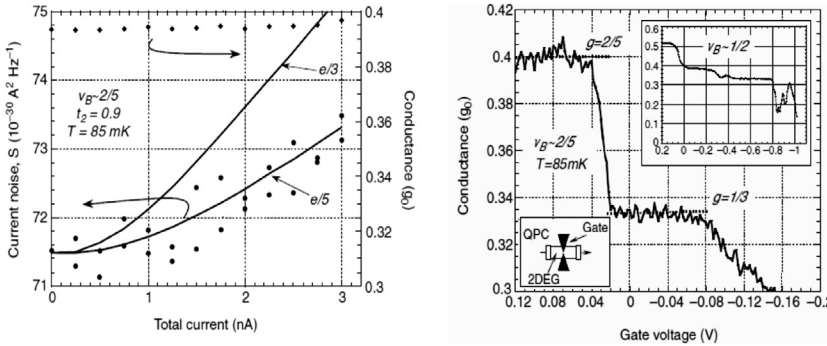


Figure 6.28 Left graph: Shot noise measurements made in the weak backscattering regime for $\nu = 2/5$ by Reznikov et al. [109]. The temperature is 85 mK and the reflection probability is kept low, $1 - D \simeq 10\%$. The solid curve is a comparison to Eq. 6.69 for charge $e^* = e/5$. Right graph: QPC conductance versus gate voltage starting at $\nu = 2/5$. A second plateau signals the formation of local $\nu = 1/3$ state in the middle of the QPC.

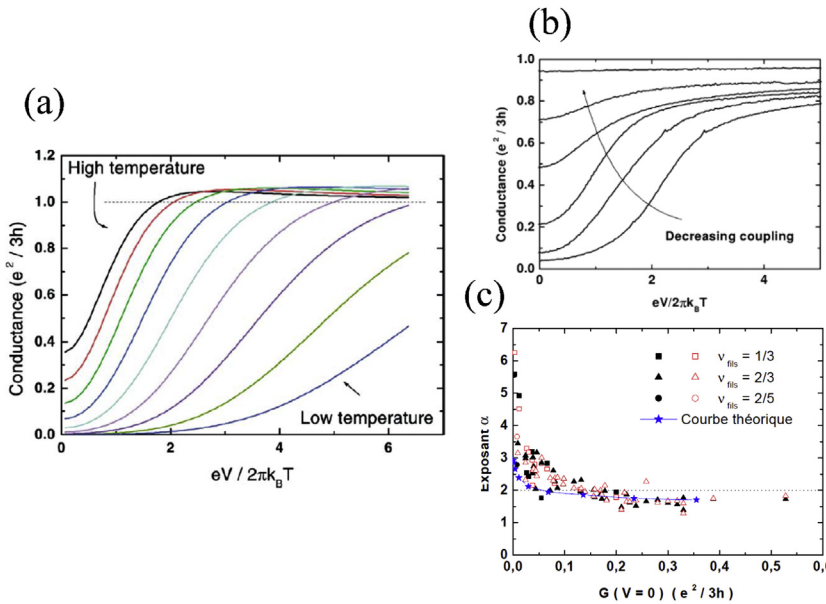


Figure 6.29 (a): conductance versus bias voltage of a QPC at $\nu = 1/3$ calculated using the exact finite temperature expressions of ref. [110]. The coupling parameter T_B is chosen as $\ln(T/T_B)$ varying from 3 to 0.6 in steps of 0.3 for temperature T. (b): QPC conductance measured for bulk filling factor $\nu = 1/3$. (c) effective exponent α describing the power law of the differential conductance with voltage $dI/dV \propto (V/V_B^{\alpha})$. Open and black symbols correspond to measurements. The solid blue (light grey in print version) curve, with blue (dark grey in print version) stars, corresponds to effective exponents extracted from the exact solution displayed in figure (a). *All figures are adapted from [112].*

Fig. 6.30a shows the charge extracted from data following this empirical approach for filling factors 1/3 and 2/5. It is interesting to compare these findings with the evolution of the charge obtained from the analysis of the exact finite temperature noise calculation of Ref. [110]. This has been done by Trauzettel et al. [114]. Here, the following noninteracting shot noise relation is used in which the charge e^* is a free parameter to be determined. Expressing the energy-dependent $\mathscr{T}(E)$ transmission in units of $G_{1/3} = e^{-2}/3h$ from the differential conductance $dI/dV = T\,(eV)G_{1/3}$, the noninteracting WB shot noise expression used is (Fig. 6.30)

$$S_{WB} = 2e^* I_{BS}(V)\mathscr{T}(eV)\coth\left(\frac{e^* V}{2k_B T}\right) + 4G_{1/3}k_B T\mathscr{T}(eV)^2 \qquad (6.71)$$

and the SB noninteracting shot noise expression used is:

$$S_{SB} = 2e^* I(V)(1-\mathscr{T}(eV))\coth\left(\frac{e^* V}{2k_B T}\right) + 4G_{1/3}k_B T\mathscr{T}(eV)^2 \qquad (6.72)$$

Here, $I_{BS} = Ve^2/3h - I$. Fig. 6.30b shows the charge extracted from the exact solution in Ref. [114] versus the zero voltage bias conductance. The filled circles are obtained using Eq. 6.71 and the open circles using Eq. 6.72.

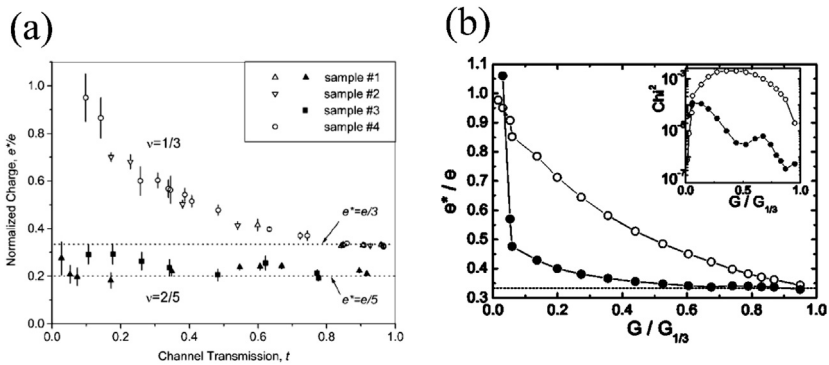


Figure 6.30 (a) Evolution of the charge from shot noise data where the charge is extracted from the noninteracting shot noise expression 6.71. (b) The charge extracted from the exact solution in Ref. [114] is displayed versus zero voltage bias conductance. The filled circles are obtained using Eq. 6.71 and the open circles using Eq. 6.72. The inset shows the estimation of the χ^2 which measures the quantitative difference between the simplified noninteraction expression and the exact solution. We see that, for a zero bias conductance between $G_{1/3}$ and $\simeq 05G_{1/3}$, the charge remains very close to $e/3$ and Eq. 6.71 is accurate whilst for conductance lower than $\simeq 0.05G_{1/3}$, one should use Eq. 6.72 with a charge very close to e. *(a) Adapted from Ref. [113]. (b) Adapted from Ref. [114].*

The inset shows the estimate of the χ^2 which measures the quantitative difference between the simplified noninteraction expression and the exact solution. We see that, for a zero bias conductance down to $\simeq 0.5 G_{1/3}$, the charge remains very close to $e/3$ and the WB expression given by Eq. 6.71 remains accurate. From this analysis, the authors of Ref. [114] conjectured that the measured noise in Ref. [113] could have been fitted quite well by the exact field theoretic solution for all values of zero bias transmission without any adjustment of e^*. For conductance lower than $\simeq 0.05 G_{1/3}$, there is sharp transition from the WB to SB expression and one should use Eq. 6.72 with a charge very close to e. A recent discussion on the use of the noninteracting formula for the fractional regime can be found in Ref. [115].

The shot noise of fractional quasi-particles $e/3$ has been further studied [116,117] in more complex geometries including several QPCs in series to look at a possible bunching of quasi-particles. Indeed, the quasi-particles are expected to be anyons, intermediate between fermions or bosons, and unlike fermions they may bunch as do photons (Fig. 6.31).

At filling factor $\nu = 5/2$, a correlated Fractional quantum Hall liquid occurs corresponding to the condensation of composite fermions. The quasi-particles are described by a two-component wavefunction and are believed to follow non-Abelian statistics. The quasi-particle charge was measured by Dolev et al. [118] and found to be $e/4$ in agreement with theory.

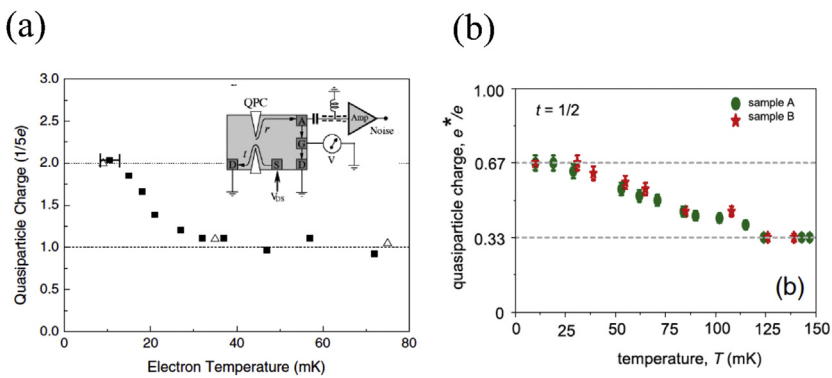


Figure 6.31 (a) Evolution of the charge for weak backscattering at $\nu = 2/5$ from auto-correlation noise measurements versus temperature. The extracted charge is shown to increase at low temperature. The inset shows the measurement circuit. (b) Same as (a) but for filling factor $\nu = 2/3$ and noise measured at transmission 1/2. *(a) Adapted from Ref. [119]. (b) Adapted from Ref. [120].*

At very low temperatures, shot noise measurements made in Refs. [119–121] seem to indicate that, for some fractional states involving several co- or counter-propagating modes, the charge e^* increases and almost doubles when lowering the temperature. This was observed at filling factors 2/5 and 3/7 [119], 2/3 [120] and even 5/2 [121]. This was interpreted as quasi-particle bunching, a condensation of quasi-particles not expected from standard theories. Fig. 6.31 shows the charge extracted from auto-correlation noise measurements for the filling factor $\nu = 2/5$ and for filling factor $\nu = 2/3$. When lowering the temperature, an increase of shot noise is observed which is attributed to an increase of the quasi-particle charge. Recently, measurements done at $\nu = 2/5$ may change the interpretation of a low temperature increase of the quasi-particle charge [122]. Cross-correlation shot noise measurements were made and, surprisingly, the auto-correlation noise and the cross-correlation noise behaved differently. This was observed in two high mobility samples. The effect occurs for $nu = 2/5$ and 2/3 but not for $\nu = 1/3$. As shown in Fig. 6.32b for $\nu = 2/5$, the cross-correlation remains constant down to the lowest temperature, whilst the auto-correlation noise increases and doubles. It is known that the cross-correlation measures the noise exclusively due to the partitioning of the carriers by the QPC between left and right contacts, whilst the auto correlation is the sum of the partition noise plus the extra noise coming from the injecting contacts (like thermal noise). The cross-correlation measures the charge, whilst the auto-correlation measures charge plus something else. One may interpret this increase of auto-correlation noise as an increase of the carrier temperature which is most likely to occur when lowering the base temperature due to lower electron–phonon thermal coupling. However, a detailed mechanism is not yet understood. This calls for more measurements of the carrier charge in this low temperature regime in order to confirm an actual doubling of the carrier charge (Fig. 6.32).

The use of shot noise in the Fractional Quantum Hall effect has been a very fruitful mean of investigation. It has been used to provide evidence for the appearance of neutral modes corresponding to filling factors for which the coexistence of co-propagating and counter-propagating edge modes are expected [123–126]. In particular, the first evidence has been observed of the long sought counter-propagating neutral modes at filling factor $\nu = 2/3$ [123]. Also, shot noise has been recently used to understand the fractional excitations occurring at the interface of $\nu = 1$ an $\nu = 1/3$ quantum Hall states [127,128].

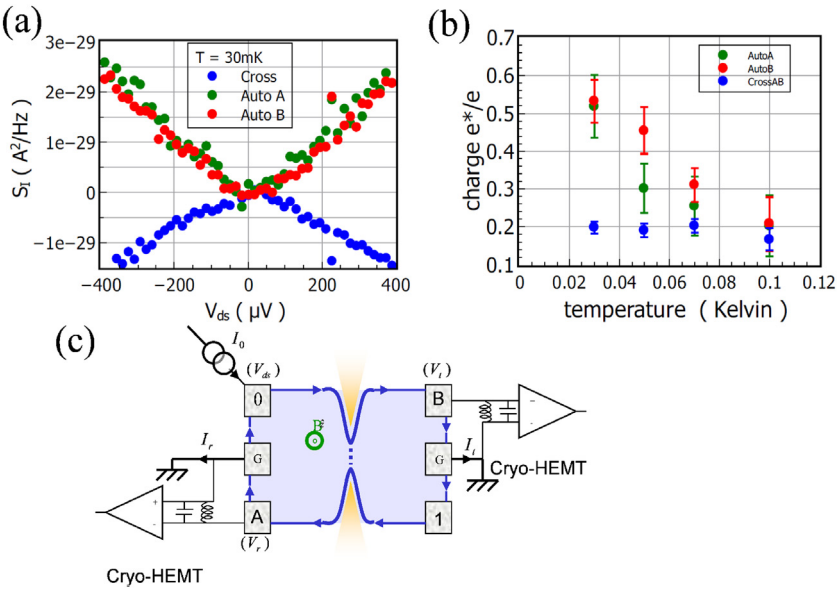


Figure 6.32 (a) The red (light grey in print version) and green (grey in print version) data corresponding to the auto-correlated noise fluctuations of backscattered and transmitted current, respectively, both measured at a low temperature of $T = 30$ mK. The cross-correlation spectrum (blue [dark grey in print version] data) is by contrast negative as expected for quasi-particle partitioning by a QPC. One observes that the amplitude of variation with bias voltage of the autocorrelation is much higher than that of the cross-correlation. The quasi-particle charge obtained from autocorrelation alone is thus overestimated. In (b) is shown the variation of the quasi-particle charge at $\nu = 2/5$ extracted from both type of measurement. One observes that the cross-correlation provides a constant charge $e/5$ with temperature. (c) measurement principle allowing simultaneous measurement of auto- and cross-correlated shot noise.

4.3.3 *New measurements of fractional charge with photo-assisted shot noise and high frequency shot noise*

In part 2.4.2.2, which is devoted to PASN, it was shown that, when a DC voltage and an AC voltage at frequency f are superimposed on a contact of a quantum conductor, the resulting low frequency shot noise is given by an expression involving a sum of DC shot noise $S_I^{DC}(V_{dc})$ with shifted voltage $V_{dc} - lhf/e$ and weighted by the probability P_l to absorb (emit) l photons. In particular, the (zero temperature) zero bias DC shot noise singularity, $S_I^{DC} \propto |V_{dc}|$, is replicated for voltages lhf/e. This phenomenon reminds us of the AC Josephson effect of superconducting tunnel junctions [129], where the singular zero bias supercurrent is replicated at Josephson voltages $V_J = lhf/2e$ and gives the celebrated Shapiro steps with a carrier charge $2e$. Both

PASN and Shapiro steps result from the same photo-assisted phenomena: the energy gained by a carrier of charge e^* transferred between two conductors biased at voltage V_{dc} can be compensated by absorbing a photon of energy $h\nu = e^* V_{dc}$ such that the system is brought back to a state similar to the zero voltage state. Hence, if a singularity occurs at zero bias, it is replicated for voltages obeying the Josephson relation:

$$e^* V_J = lhf \tag{6.73}$$

with $e^* = 2e$ for the AC Josephson effect or the PASN of a superconducting–normal junction [93] and with $e^* = e$ for the PASN measurements discussed in Section 4.2.2.

What about the fractional quantum Hall regime? In 1991, X. G. Wen already proposed a Josephson relation for fractional charges for the high frequency noise in the Fractional Quantum Hall effect [130], see also [106]. PASN formulas analogous to Eq. (6.61) but using a fractional e^* have been derived in Refs. [131–133] and a Josephson relation has also been found for the photo-assisted current [134].

The first experimental evidence of a fractional Josephson relation was observed by Kapfer et al. [99] for charges $e/5$ and $e/3$ using a high-mobility sample at bulk filling factor $\nu = 2/5$. Similar to the experiment of Reznikov et al. [109], the QPC can induce a WB regime of $\nu = 2/5$ inner fractional edge channels, or realise, at a more negative QPC gate voltage, a local $\nu = 1/3$ state and induce a WB regime. In these two regimes, the DC shot noise was measured in Ref. [99] at fractional charges $e^* = e/5$ and $e/3$, respectively, consistent with the measurements of Ref. [109]. By superimposing an AC microwave voltage, Kapfer et al. measured the PASN. To better illustrate the shot noise singularity and hence the Josephson relation, they constructed the following excess noise ΔS_I which is defined by the following (Fig. 6.33):

$$\begin{aligned}\Delta S_I(V_{dc}) &= S_I^{PASN}(V_{dc}) - P_0 S_I^{DC}(V_{dc}) \\ &= P_1 S_I^{DC}\left(V_{dc} - h_f/e^*\right) + P_{-1} S_I^{DC}(V_{dc} + hf/e^*)\end{aligned} \tag{6.74}$$

With this trick, one is left with only pure photo-assisted contributions (here, the ± 2 photon contributions are neglected). Experimentally, this consists of measuring the noise, with both a DC and an AC voltage applied, and subtracting the noise with only the DC voltage applied, weighted by the appropriate P_0, the probability to not emit nor absorb a photon. Fig. 6.33 shows the results obtained for the WB of charge $e/5$ (similar measurements

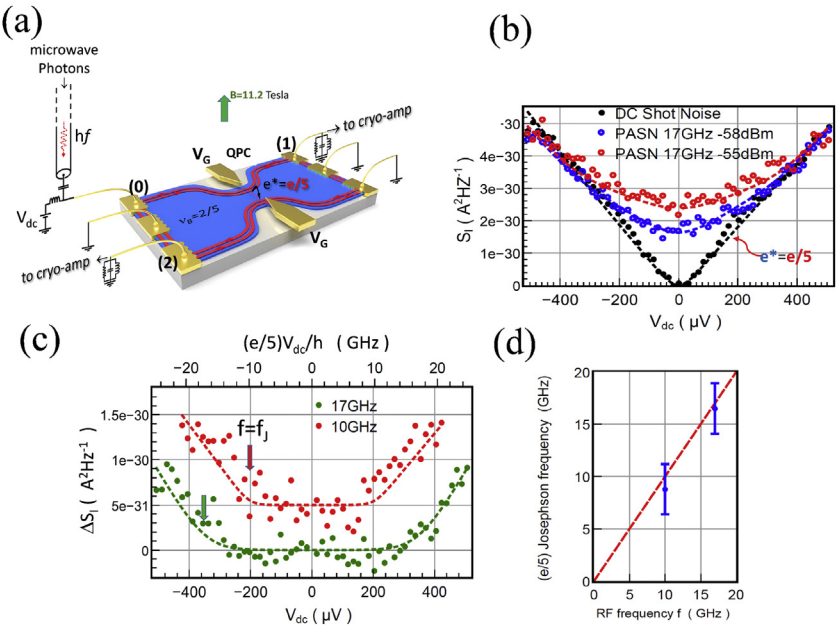


Figure 6.33 Fractional charge measured through a Josephson relation, (a): schematic of the sample including the microwave irradiation, the cross-correlation noise measurements and the QPC inducing a weak backscattering at $\nu = 2/5$. (b): cross-correlated shot noise measurements versus DC voltage for no RF excitation (black symbols) and for two different RF excitations (blue [dark grey in print version] and red [grey in print version] symbols). The dotted lines are a comparison with theory using $e^* = e/5$. The electron temperature is 30 mK. (c) Excess noise, defined by Eq. 6.74 for 17 and 10 GHz. The onset voltage above which the noise rises, the Josephson voltage, is shown to increase with frequency. (d) the measured onset voltage in units of Josephson frequency is plotted versus the excitation frequency. *Adapted from [99].*

are done for charge $e/3$). In a single experiment, one has two different determinations of the fractional charge: the 'old' one, based on charge granularity, and the 'new' one, based on a comparison of frequency to voltage resulting from the single photon absorption/emission of a fractional charge. Note that the new measurement confirms the absence of the puzzling doubling of the charge $e/5$ at the lowest temperature. Finally, recent high frequency shot noise measurements have been made by Bisognin et al. [180]. Measurements of charge $e/3$ at filling factor 4/3 and 2/3 were made by observing the suppression of high frequency shot noise at low DC voltage when the quasiparticle energy $(e/3)V_{dc}$ is smaller than the energy hf required to emit a single microwave photon in the external circuit. The physics behind the noise

suppression is similar to that observed for charge e by Zakka-Bajjani et al. in Ref. [90] and discussed in Section 4.2.1. It also provides a determination of the fractional charge based on a comparison of voltage to frequency.

4.4 Investigating two-particle correlation and interference with shot noise measurements

This part will discuss experiments done with electrons which may be considered an analog of quantum optics experiments where, in the analogy, electrons replace photons and the optical medium (interferometers, beam splitters and optical fibers) is replaced by ballistic quantum conductors. This analogy was first put forward in seminal papers on shot noise by Buttiker [10] and by Martin and Landauer [5] both of which established the basis of the, now rising field, of Electron Quantum Optics. Whilst the measurement of a current essentially gives information on the flow of individual carriers (how many electrons pass through the circuit), the shot noise involves the correlation $I(t)I(t+\tau)$ of two currents at different time. It thus gives information about the presence of two carriers for the relative time τ or distance $v_F\tau$. We immediately see a strong difference between photons, which are bosons, and electrons, which are fermions. One can have several photons in a beam at a given time which enhances the beam fluctuations as they are bosons, whilst one cannot have two electrons close because of Pauli exclusion. So, shot noise is sensitive to quantum statistics. But it is even better than that: it is a direct measure of the particle exchange. This is best shown by considering the following noise measurements made using a 4-terminal quantum conductor shown in Fig. 6.34. Each lead $\alpha = 1$ to 4 contains a single mode and an assumption of zero temperature is made for simplicity. One is interested in measuring the current correlation between leads 2 and 4, whilst a DC voltage V can be applied to leads 1 or 3 or to both 1 and 3. From the scattering theory of quantum noise, the low frequency cross-correlated current fluctuation, when applying voltage V to contact 1 only, is given [10] by

$$S^1_{2,4} = -2\frac{e^2}{h}eV\left(s_{21}s^*_{21}s_{41}s^*_{41}\right) \tag{6.75}$$

and when applying voltage V to contact 3 only is given by

$$S^3_{2,4} = -2\frac{e^2}{h}eV\left(s_{23}s^*_{23}s_{43}s^*_{43}\right) \tag{6.76}$$

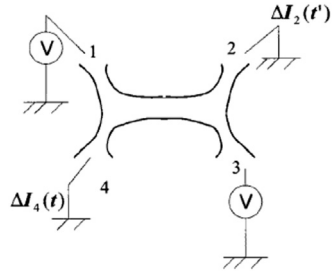


Figure 6.34 Four contact phase coherent conductor used to give evidence about the role of particle exchange in cross-correlated shot noise measurements following [10].

where $s_{j,i}$ is the amplitude of transmission from input lead i to output lead j. The interesting point emphasised in Ref. [10] is that applying voltage V to both contacts simultaneously gives a noise $S_{2,4}^{1+3} \neq S_{2,4}^{1} + S_{2,4}^{3}$, and the difference reveals a nonseparable exchange term due to particles coming from 1 to 3 and de-localised in leads 2 and 4 [10]:

$$S_{2,4}^{1+3} - S_{2,4}^{1} - S_{2,4}^{3} = -2\frac{e^2}{h} eV\left(s_{21}s_{23}^{*}s_{43}s_{41}^{*} + s_{23}s_{21}^{*}s_{41}s_{43}^{*}\right) \tag{6.77}$$

We see that quantum noise reveals particle exchange and can be used to probe quantum statistics. Also an important difference between $S_{2,4}^{1+3}$ and $S_{2,4}^{1}$ and $S_{2,4}^{3}$ is that the last two can be put in the form of a product of probabilities $|s_{21}|^2|s_{41}|^2$ and $|s_{23}|^2|s_{43}|^2$, respectively, emphasising the statistical partitioning of single particles between output leads 2 and 4, whilst the former cannot be factorised as a product of probabilities and hence implies two-particle quantum entanglement. It is thus highly valuable to study current cross-correlations in quantum conductors. This is not only a safer way to measure noise, as advocated in the part devoted to measurement techniques, or for fractional charge measurements, but this is also a way to probe deep quantum correlations. Samuelsson et al. [135] have proposed an experimental scheme in order to measure two-particle Aharonov–Bohm interference using cross-correlated shot noise. Its realisation will be discussed in the next subsection (Figs. 6.35 and 6.36).

4.4.1 Hanbury Brown and Twiss shot noise experiments using DC voltage sources

The electronic analog of the celebrated Hanbury Brown and Twiss (HBT) quantum optics experiment, where a photon beam is split by a beam splitter

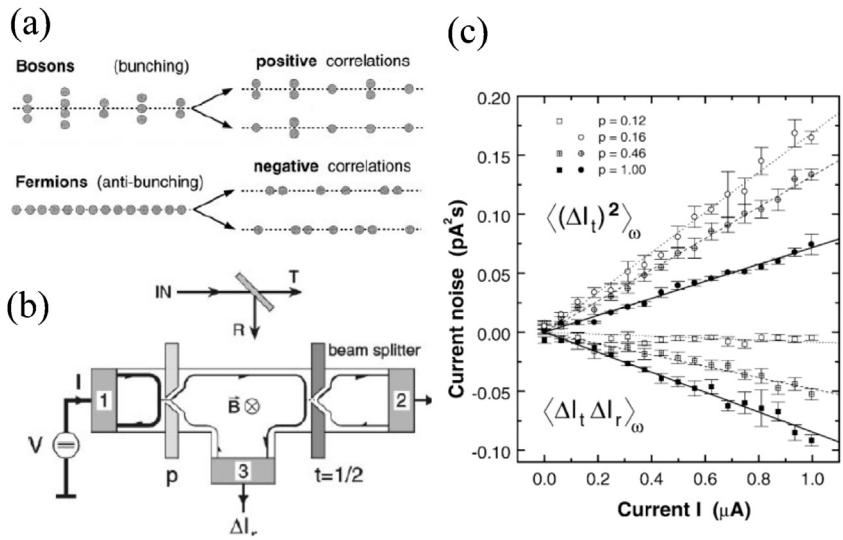


Figure 6.35 Electronic analog of Hanbury Brown and Twiss (HBT) experiment (a) sketch of photon bunching and electron anti-bunching (b) schematic of the sample used. A first QPC on the left transmits electrons injected from contact (1) with probability p. The second QPC on the right is the HBT beam splitter used to partitioned the electron beam. The cross-correlated current noise is measured at contacts (2) and (3); (c) The positive autocorrelations and the corresponding negative cross-correlations of the current noise plotted are versus the incoming current. For $P = 1$ their amplitudes are equal, whilst for $p << 1$, the cross-correlation vanishes and the Fano factor of autocorrelation approaches one, as expected for a Poissonian statistics. *Adapted from Refs. [137] and [136].*

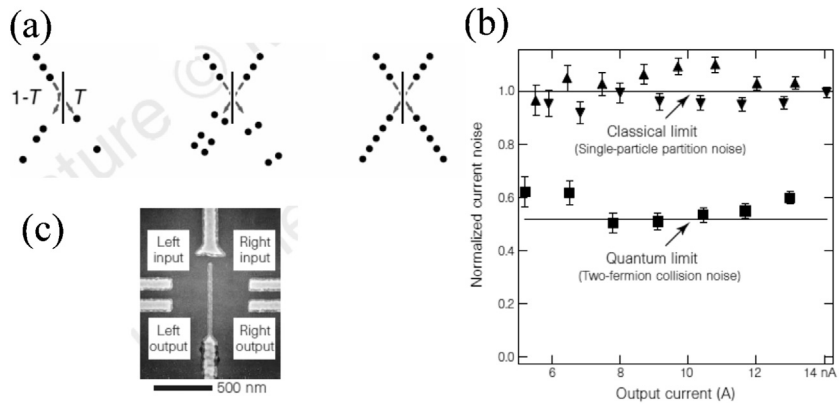


Figure 6.36 Quantum interference in electron collisions (a) sketch of photon bunching and electron anti-bunching (c) Sample used in the experiment. Left and right contacts inject electrons towards the electronic beam splitter defined by the middle vertical gate. Two separate contacts collect the current fluctuations. (b) black triangles are noise data when the voltage is applied to one contact only (single particle partition noise), squares are noise data for voltage applied to both contacts, showing the quantum reduction due to electron interference. The two-particle noise is not zero because of beam splitter imperfections. *Adapted from Ref. [142].*

and the intensity correlations of output beams are measured, has been realised by Henny et al. [136] and Oberholzer et al. [137] using electrons propagating along chiral edge channels in the Quantum Hall Regime and a QPC is used as the electron beam splitter. The noiseless incoming electron beam was made fluctuating by introducing another QPC before the QPC beam splitter. The results are shown in Fig. 6.35. For a noiseless incoming beam, the cross-correlated noise is negative but of same amplitude as the auto-correlated noise, whilst, when diluting the incoming beam and so making it noisy, the auto-correlated noise rises, whilst the amplitude and negative correlation decrease. This provides a firm test of our understanding of electron shot noise. A similar experiment was simultaneously done by Oliver et al. [138]. The negative cross-correlation of the current of these experiments confirms the anti-bunching of electrons as opposed to photon bunching. A two-electron collision experiment for testing electron anti-bunching inspired by the measurement protocol leading to Eq. 6.77 has been done by Liu et al. [142] and is shown in Fig. 6.36.

To understand these experiments does not require any more physics once we know from auto-correlation noise measurements [6,7] that an incoming beam is noiseless. For example in Refs. [136,137], for $P = 1$, the negative correlations in the previous experiments can also be viewed as resulting from conservation of the noiseless incoming current, a positive current fluctuation of the transmitted current being compensated by a negative fluctuation of the reflected current. Similar reasoning can explain the results for $P < 1$ (Fig. 6.37).

A much more informative, but extremely difficult, experimental approach is to observe two-particle interference effects in cross-correlated shot noise. Indeed, from Eq. 6.77, one sees that the sum of products of scattering amplitudes in the nonseparable exchange term could allow the observation of two-particle interference. A smart scheme to do this was proposed by Samuelsson et al. [135]. It consists of designing an experimental set up in which two particles each coming from independent voltage sources can give rise to two-particle interference, whilst no single-particle interference can occur. An experimental implementation has been realized by Neder et al. [139].

Let us concentrate on Fig. 6.37a which shows the set-up originally considered by Samuelsson et al. [135]. The dashed lines with arrows at their end stand for chiral edge channels along which electrons propagate. The oblique thick dashed lines represent electron beam splitters realised using QPCs. Two voltage sources S1 and S2 inject electrons. The amplitude of

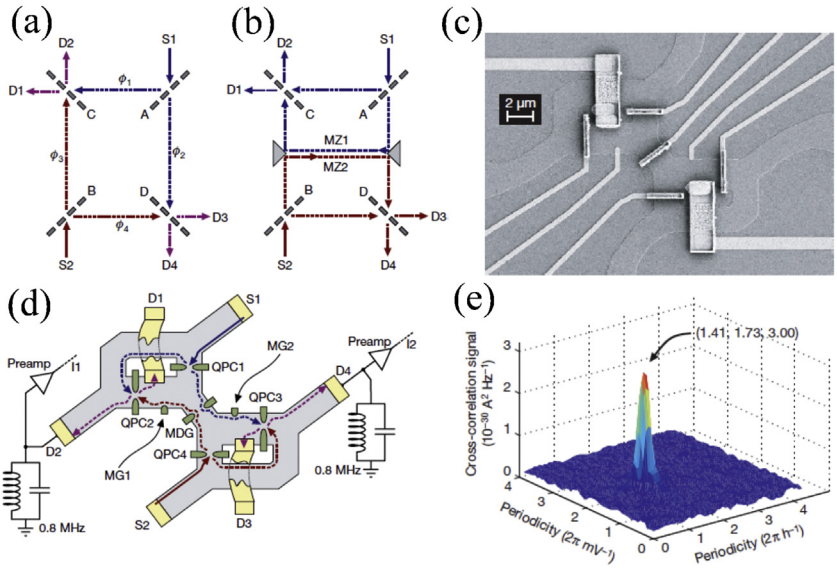


Figure 6.37 Two-particle interference. (a) original two-particle interference scheme borrowed from Ref. [135]. (b) A centre QPC splits the main loop into an upper loop and lower loop. For closed QPC, the set-up defines two Mach–Zehnder interferometers which allow the probing of single-particle interference. When open, one comes back to the topology of figure (a), no single-particle interference is possible but two-particle interference can be probed by correlating the outputs D2 and D4. (c) SEM image of the sample; (d) sample schematics, and (e) Fourier transform of the periodic cross-correlated noise oscillations versus flux and gate voltage used to vary the phase demonstrating two-particle interference. *Adapted from Ref. [139].*

transmission and reflection through a beam splitter are $t = 1/\sqrt{2}$ and $r = i/\sqrt{2}$. We are interested in the cross-correlated current fluctuations between output D2 and D4. They arise from the coincidence probability to find an electron at D2 and D4. Including the phases ϕ_i, accumulated along the path joining the input and output beam splitters, the two-particle probability amplitude is as follows: $\left(-\frac{1}{2}\right)e^{i\phi_1}\left(-\frac{1}{2}\right)e^{i\phi_4} + \left(\frac{1}{2}\right)e^{i\phi_2}\left(\frac{1}{2}\right)e^{i\phi_3}$ and the two-particle coincidence probability is

$$P_{\mathrm{D2,D4}}^{\mathrm{S1,S2}} = \frac{1}{8}\left(1 + \cos\Phi_{Total}\right) \tag{6.78}$$

where $\phi_{Total} = \phi_1 + \phi_4 - \phi_2 - \phi_3$ is the Aharonov–Bohm phase $\phi_{Total} = BS$ accumulated in the presence of a perpendicular magnetic field B in the

loop of area S formed by the four paths. The detailed analysis of Refs [135] and [139] shows that the cross-correlated noise is proportional to the above coincidence probability. On the contrary, when looking at the average current exiting at any output D_i, no Aharonov–Bohm phase, and hence no interference, occurs as there is not a closed single particle path available. For better control, in the experiment of Neder et al., the large Aharonov–Bohm loop is split into two loops by inserting a QPC in the middle; see Fig. 6.37b. When the QPC is closed, one has two separate upper and lower Mach–Zehnder interferometers in which conductance measurements can check and ensure that single-particle interference occurs. When the middle QPC is open, no single particle interference is observed, as expected, but tiny oscillations in the cross-correlated shot noise are observed when varying Φ_{Total} with magnetic flux or by changing the main loop area with a nearby gate.

4.5 Shot noise measurements for Electron Quantum Optics

4.5.1 Current noise of single electron pump and sources

The field of electron quantum optics has recently risen with the advent of on demand single electron sources. Single electron sources are described in Chapter 11 of this book by Masaya Kataoka. Here, we will concentrate on coherent single electron sources. By coherent we mean that the single electron is injected in single, well defined, quantum state. This differs, for instance, from metallic electron turnstiles which inject electrons into a wide mixture of noncoherent quantum states. A coherent source requires the injection of electrons from a discrete energy level of a quantum dot (or mesoscopic capacitor) or to be able to inject an electron in the form of a well-defined time resolved quantum state of the Fermi sea: a Leviton. The coherent on-demand single electron sources play a role analogous to single photon sources in quantum optics and therefore are very relevant to electron quantum optics. We refer the reader to the latest reviews of this field made in Refs. [143,144]. To analyse the quality of these sources, HBT experiments as currently performed in Quantum Optics for single photon sources, see for example Refs [140,141] (Fig. 6.38).

Beyond electron quantum optics, we will start by illustrating how current noise measurements can be done to quantify the accuracy of single electron pumps aimed at providing current standards. In this case a QPC electron, beam splitter is not used for HBT correlations, but the direct emission noise of the pump is measured. In the work of Maire et al. [145], the pump is realised in a narrow wire, and a quantum dot is formed by confining

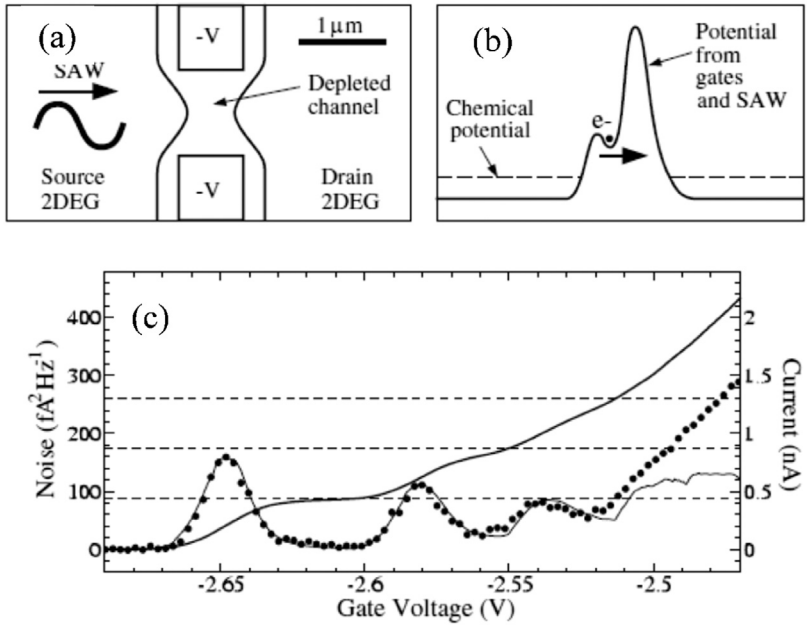


Figure 6.38 Shot noise of a SAW-based electron pump. (a) A short 1D wire is formed in a 2D electron gas using negatively biased split gates on top of a GaAs/GaAlAs heterojunction. A surface acoustic wave creates a moving potential well which captures an electron and acts as a pump. (b) Energy diagram of the moving quantum dot formed by the combination of the SAW potential well and split gate transverse confinement. (c) Average current and associated shot noise observed. On a current plateau the noise vanishes. *Adapted from [146].*

electrons between two gates, one of which is driven by a voltage at frequency $f_p \simeq 400$ MHz. If p is the probability of transferring an electron at operating frequency f_p, the mean current is $I_p = pef_p$ and the current noise is expected to follow binomial statistics with $S_I = 2ef_p p(1-p)$. The experiment showed good agreement with the binomial model [145]. Similar shot noise measurements were done for Surface Acoustic Wave (SAW)-based single electron pumps by Robinson et al. [146]. The pump was operated at $f_p = 3$ GHz and the noise measured at MHz frequency to avoid $1/f$-like impurity noise [147]. Fig. 6.38 shows the quantisation plateaus of the current typical of SAW-based electron pumps and the shot noise of the pump versus the split gate voltage used to confine electrons in a narrow wire.

We will now consider on-demand coherent single electron sources that are highly suitable for Electron Quantum Optics. There are two different types. The first type is the Mesoscopic Capacitor Single Electron Source

(MC-SES) [148]. This is a quantum dot which periodically injects into a quantum conductor an energy resolved electron followed by an energy resolved hole situated above and below the Fermi energy of the conductor, respectively. The second type is the Leviton Single Electron Source (L-SES) [96]. Here, a short voltage pulse V(t) with a Lorentzian time shape applied to a contact injects from the external circuit a single electron into a quantum wire. This requires a quantised Faraday flux, i.e., $\int_{-\infty}^{\infty} V(t)dt = h/e$ such that the resulting current pulse $I(t) = e^2/hV(t)$ corresponds to a single electron, i.e., $\int_{-\infty}^{\infty} I(t)dt = e$. In the L-SES, the electron is injected in remarkable minimal excitation state, a time resolved quasi-particle of the Fermi sea which has been predicted by Levitov et al. [77] and Keeling et al. [152] and was dubbed a leviton. The two sources show an interesting duality regarding their energy and temporal properties. The L-SES source injects electrons in the form of a time-resolved Lorentzian wavepacket just above the Fermi energy with an exponentially decreasing energy distribution. The MC-SES injects energy resolved electrons, with a Lorentzian energy distribution, well above the Fermi energy, whilst the electron wavepacket shows an exponential decay in time. This is illustrated in Fig. 6.39. There are important differences regarding noise. The MC-SES has a finite probability of emitting an electron (or a hole) which generates charge noise that can be directly measured at finite frequency at the output of the source. The L-SES source injects one electron with certainty for each Lorentzian pulse, but imperfections in the pulse shape may give rise to a small amount (few %) of neutral excitations which require the use of a QPC beam-splitter and HBT measurements to be observed. A full analytical calculation of the noise of the MC-SES has been done by Albert et al. [150] and agrees perfectly with the experimental results (Fig. 6.39).

Fig. 6.40d shows the noise measurements of the MC-SES made by Mahé et al. [149]. The source is made from a Quantum Dot coupled to a quantum Hall edge channel via a QPC with tunable transmission D. A key parameter is the electron emission time given by $\tau = h/\Delta(1/D - 1/2)$ where Δ is the dot energy level spacing. A gate above the dot, driven by a square wave at frequency f, pushes alternatively, by capacitive coupling, the last occupied energy level above and below the Fermi energy of the edge channel to emit an electron and then a hole, see Fig. 6.40a. When closing the QPC, τ increases and becomes comparable to the injection period $T = 1/f$ introducing errors. The emission probability is given by $P = \tanh(T/4\tau)$ and the expected charge noise measured in a time $T/2$ is $\propto 2e^2fP(1 - P)$. In the experiment, because of the capacitive nature of the device, the noise

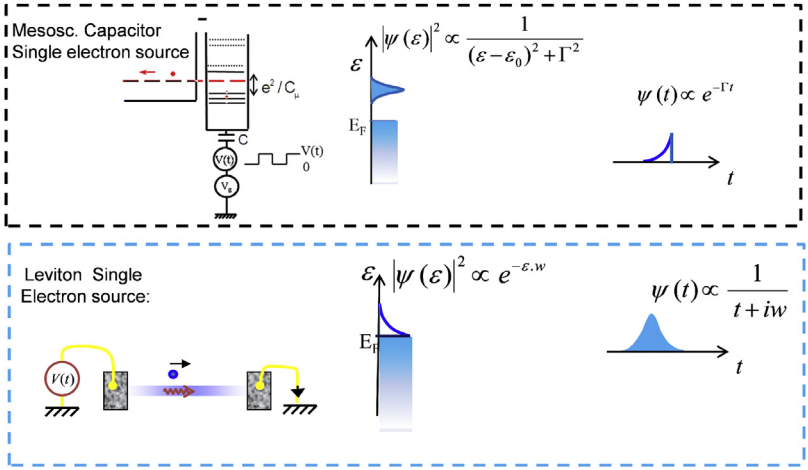


Figure 6.39 Duality between the MC-SES and the L-SES. Top: Mesoscopic Capacitor electron source [148]: electrons are emitted from the last occupied energy level of a quantum dot above the Fermi energy. The level broadening gives a Lorentzian energy distribution. The wavepacket is semi-exponential in the time domain; Bottom: Leviton single electron source [96]. Lorentzian voltage pulses inject electrons in a minimal excitation quantum state: a leviton. This is a time resolved electron source with a Lorentzian wavepacket and semi-exponential energy distribution just above the Fermi energy.

measurements have been done in a frequency range 1.2–1.8 GHz around the drive frequency $f = 1.5$ GHz. These finite frequency measurements reveal an interesting fundamental extra noise, a phase noise called quantum jitter noise, related to the random emission time coming from the tunnelling process. The data of Fig. 6.40d are compared with calculation which was validated in Refs. [150] (Fig. 6.40).

4.5.2 *Hanbury Brown and Twiss experiments and on-demand single electron sources*

An HBT experiment consists of splitting an incident beam of electrons with a beam splitter and analysing the transmitted particle number fluctuations. This is a very useful tool to characterise single electron sources. Indeed, for electronic systems, which are described by Fermi statistics, the quantum partition shot noise gives, for a periodically driven source, a direct measure of the total number of the incoming electron-like and hole-like excitations emitted by the source:

$$S_I = 2ef_d D(1-D)(N_e + N_h) \tag{6.79}$$

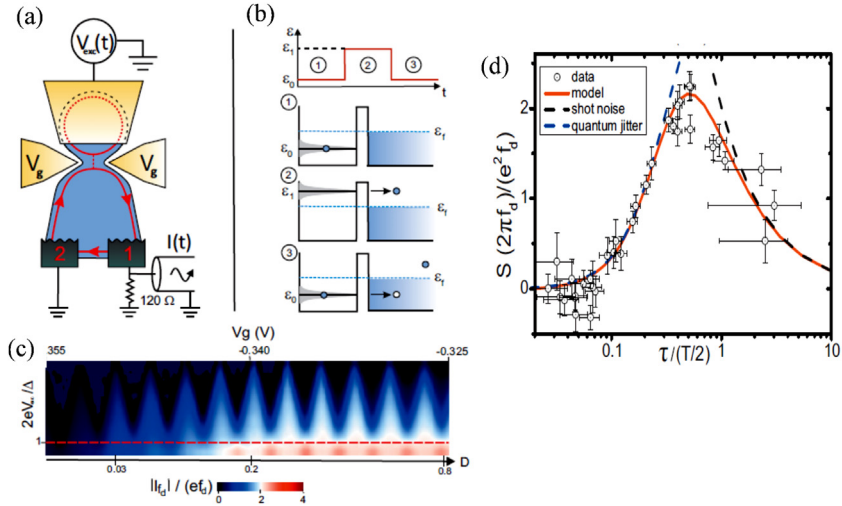


Figure 6.40 High frequency noise of a Mesoscopic Capacitor single Electron Source. (a) Sketch of the sample. The 2DEG is in blue (grey in print version), the chiral edge channel is in red (dark grey in print version). The top gate is driven by a square wave voltage $V(t)$. The resulting emission of electrons followed by holes gives rise to an AC current which is measured at the same frequency. A QPC is used to tune the transmission D of electrons from the dot to the chiral lead. (b) energy diagram showing the principle of energy resolved single electron/hole emission. (c) colour plot of the AC current versus the QPC gate voltage and the amplitude of the squarewave in units of Δ, the Dot energy level separation. The units of current are ef, $f = 1.5$ GHz. (d) Current noise data measured in the frequency range 1.2–1.8 GHz. They are compared with the shot noise (black dashed line), quantum jitter noise (blue [dark grey in print version] dashed line) and with the model (orange [grey in print version] line) validated later by analytical [150] and numerical works. *Adapted from [149].*

where D is the transmission of the beam splitter and N_e, N_h the number of incoming electron and hole excitation per period $1/f_d$. The expression holds for low temperatures: $k_B T << hf_d$. This relation was discussed in the work of Reydellet et al. [94], where the authors interpreted the measured PASN as if the electron-hole pairs generated by a sine-wave voltage behave as independent beams of electron and of holes, each contributing to partition noise, leading to Eq. 6.79. Keeping only single photon processes for small drive amplitude $V_{ac} < hf_d/e$, the number of electrons and holes was explicitly $N_e = N_h = J_1\,(eV_{ac}/hf_d)^2$. Whilst it was shown that electrons and holes are not fully independent but linked by quantum coherence [151], relation 6.79 still holds in the general situation when both electron and holes are generated. This result was later derived by Keeling et al. [152] in their work on minimal excitation states and by others [153,155,156].

The first HBT experiment showing the one by one partitioning of electron/hole emitted by an on-demand single electron source, an MC-S, was done by Bocquillon et al. [153]. Here, the capacitive nature of the source implied $N_e = N_h$. The authors observed a clear $D\,(1-D)$ variation of the low frequency measured with the transmission D of the QPC used for partitioning. The set-up and the results are shown in Figs. 11.25 and 11.26 of Chapter 11 of this book devoted to single electron sources. The shot noise was however a bit lower than expected, underestimating the amount of electrons and holes emitted by the source. This reduction of noise was properly interpreted as due to the anti-bunching effect of electrons or holes colliding with thermal electron and hole excitations coming from the unused input port of the beam splitter. In other words, one has to make the substitution $N_e + N_h \rightarrow N_e + N_h - \int_{-\infty}^{\infty}(n_e(\varepsilon)f(\varepsilon) + n_h(\varepsilon)\overline{f}(\varepsilon))d\varepsilon$ in Eq. 6.79. Here, $n_{e/h}(\varepsilon)d\varepsilon$ represents the number of electrons or holes in the energy range $[\varepsilon, \varepsilon + d\varepsilon]$ and $f(\varepsilon)$ and $\overline{f}(\varepsilon)$, respectively, the thermal distribution of electrons and holes of the unused beam splitter input.

An HBT-like set-up was also used as a key tool to provide the first experimental evidence of Levitons in Ref. [96]. Levitons are time-resolved minimal excitations states, originally predicted and discussed by Levitov and collaborators in Refs. [77,152,154]. They carry an integer electronic charge and are characterised by a complete absence of hole excitations. According to Eq. 6.79, their noise is minimal as expected for the partitioning of electrons not accompanied by electron-hole excitations. In the work of Dubois et al. [96], periodic Lorentzian voltage pulses $V(t)$ are applied to a contact at a frequency f_d. The amplitude and width of the pulse determines the charge q per pulse per electronic quantum mode. Indeed, the instantaneous current injected in a given electronic mode is $I(t) = (e^2/h)V(t)$ and the injected charge is $q = (e/h)\int eV(t)dt$. A single electron is emitted, $q = e$, when the Faraday flux $\int V(t)dt$ is tuned to the quantum flux value h/e. In the experiment, the injected electrons were sent to a QPC transmitting a single mode such that the partitioning generates a measurable low frequency shot noise. Expressing the excess noise ΔS_I as the difference between the noise measured when voltage pulses are applied and the noise measured when only the DC voltage $V_{dc} = \langle V(t)\rangle$ is applied gives a measure of the electron-hole excitations. Indeed, according to Eq. 6.79 (Fig. 6.41).

$$\begin{aligned}\Delta S_I &= 2e^2 f_d D(1-D)(N_e + N_h - n)\\ &= 2e^2 f_d D(1-D)\Delta N_{eh}\end{aligned} \tag{6.80}$$

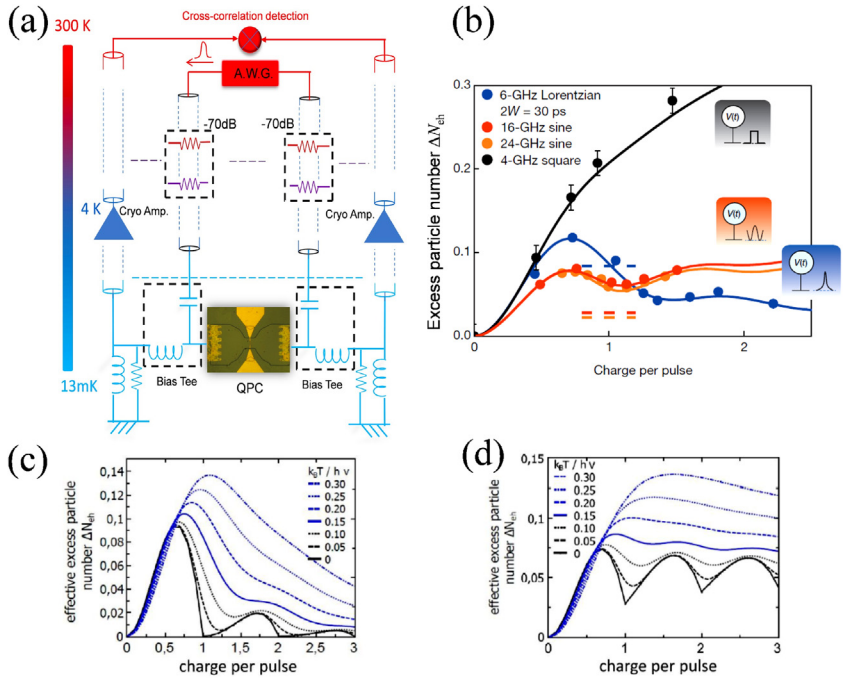


Figure 6.41 Levitons: minimal excitation single electron sources probed using HBT noise correlations. (a) Experimental set-up: an Arbitrary Wave Generator (AWG) sends, to the left contact of a 2DEG, periodic Lorentzian voltages pulses of 30 ps width at 6 GHz repetition frequency. In the middle of the sample, a QPC partitions the injected electrons. This generates current fluctuations which are collected at the left and right contacts. Two matched resonators, with 2.5 MHz centre frequency and 450 kHz bandwidth, filter the current noise which is amplified whilst a computer performs the cross-correlation; (b) Shot noise measured in units of effective excess particle number ΔN_{eh} for different voltage pulse shapes versus the charge per pulse. For $q = e$, Lorentzian voltage pulses give no more noise (blue [dark grey in print version] data) than a calculable thermal noise level indicated by the horizontal blue [dark grey in print version] dashed line. In contrast, sine-wave pulses generate an extra noise (orange [grey in print version] data and dashed line) revealing electron-hole excitations. The solid line are predictions with no adjustable parameters; (c) noise predicted for Lorentzian pulses in units of excess particle number for finite temperature (ν denotes here the repetition frequency); (d) same as (c) but for sine-wave pulses. *Adapted from [96].*

where $n = q/e$ is the charge injected per period $T = 1/f_d$. Minimal excitation states, $\Delta N_{eh} = 0$, occur only for Lorentzian pulses with integer charge, i.e., $eV(t) = \frac{h}{\pi W}\sum_k \frac{1}{1+(t-kT)^2/W^2}$. This is not the case for integer charge sine-wave pulses $eV(t) = hf_d(1-\cos 2\pi f_d t)$ or for square wave pulses. Also

Lorentzian pulses with $q \neq e$ contain hole excitations meaning that injecting a noninteger charge can only be done at the expense of large number of superpositions of electron and hole excitations, a manifestation of the so-called dynamical orthogonality catastrophe [157,158]. The experimental results of Dubois et al. [96] are shown in Fig. 6.41b. The excess shot noise measured ΔS_I is expressed in units of effective excess particle number $\Delta N_{eh} \equiv \Delta S_I / 2e^2 f_d D(1-D)$ from Eq. (6.80). Here, the sign $\equiv$ stresses that the equality is correct only in the limit $hf_d/k_BT << 1$ and finite temperature corrections have to be included. When thermal noise contributions are included, one observes that, for $q = 1$, Lorentzian pulses do not add extra noise, demonstrating the existence of levitons, whilst sine or square wave pulses give rise to extra noise: the injected electron is accompanied by a neutral cloud of electron-hole pairs. Fig. 6.41(c) and (d) show the computed excess noise for Lorentzian and sine-wave pulses at different temperatures. At zero temperature, one expects a local minimum in the noise for integer q. For finite temperature, the minimum is shifted to higher values in agreement with observations.

HBT correlation has been also used to characterise the emission of few electron sources based on a nonadiabatic quantised charge pump by Ubbelohde et al. [159]. The sample is shown in Fig. 6.42a. The energy diagram describing the pumping and the resulting pumped current is shown in (b). The electron source is made from a nanowire etched in a 2DEG with two gates creating barriers to define a quantum dot. The voltage of the left gate, the entrance gate, is periodically modulated. On the first half period, the input barrier is lowered to capture electrons in the lowest energy level of the dot. In the second half period, the barrier increases, whilst the dot energy level is pushed above the right barrier to eject the electrons, so realising efficient pumping. Note that, contrary to the leviton source or the M-C source, electrons are injected at a very high energy here. Quantised current plateaus at $i_p = ef_d$ and $2ef_d$ indicate stable and regular pumping of single and two electrons, respectively. The sample is placed under strong perpendicular magnetic such that the electrons are emitted in chiral edge channels. They propagate towards a third gate which plays the role of a beam splitter to partition the electrons. Here, the beam splitter transmission strongly depends on the energy of the emitted electrons. HBT correlations are obtained from cross-correlation measurements of the fluctuations of transmitted and reflected currents. For single electron emission, the authors observed the expected single electron binomial partition, where

the cross-correlated noise, takes the value $-0.5e^2 f_d$ for half transmission (transmitted current $0.5ef_d$). For two-electron emission, an important question is to determine if the electrons are independently transmitted or if the electron pair shows correlation. Let p_2 (p_0) be the probability of two electrons transmitted (reflected) and p_1 for a split electron pair. The transmitted current is $I_T = ef_d(p_1 + 2p_2)$, whilst the cross-correlated noise is $S_X = -2e^2 f_d(p_0(1-p_0) + p_2(1-p_2) + 2p_0p_2)$. Combined with probability conservation, measuring both current and noise allows the determination of all probabilities. The authors of [159] found two distinct regimes. For a slow sine-wave gate drive, the electrons showed statistically independent scattering resulting in a Poisson binomial distribution with $p_0 = T_a T_b$ and $p_2 = (1-T_a)(1-T_b)$ where $T_{a,b}$ are the transmission coefficients for each electron. This case corresponds to the green region in Fig. 6.42e defined by $\sqrt{p_0} + \sqrt{p_1} \leq 1$, see Ref. [160]. In figure (e), the data of figure (c) corresponds to the blue points which follow the limit $T_a = T_b$. However, for a nonadiabatic voltage, whose time shape is indicated in the inset on top of figure (d), correlated pair emission (red squares in (e)) was found, probably due to interaction effects. Also, the peaks in the noise for transmitted currents $\simeq 0.5$ and $\simeq 1.5ef_d$ indicate electron bunching (Fig. 6.42).

To sum up this part, all the above examples show that HBT noise correlation is an important tool to access the physics of single electron sources.

4.5.3 Electronic Hong Ou Mandel experiments using on-demand single electron sources

An electronic Hong Ou Mandel (HOM) experiment is the electronic analog of the celebrated HOM photonic quantum optics experiment [161]. It differs from HBT experiments in that particles which are emitted from two independent sources are mixed in a beam splitter and a controlled time delay is introduced between the sources. In the experiments with electrons, as the technology of electron coincidence detection is not available, unlike in quantum optics, the HOM correlations are obtained from statistical measurements of the cross-correlated fluctuations of the transmitted and reflected electron beams. This provides similar information. When the delay τ between the arrival of two electrons on the beam splitter is zero, the two particles interfere giving information about their quantum statistics (fermion or boson). When τ is much larger than the time width W of the wavepacket, independent partitioning of the particles occurs regardless of their quantum statistics. For $\tau \simeq W$, the HOM correlation gives information about the size of particle wavepackets.

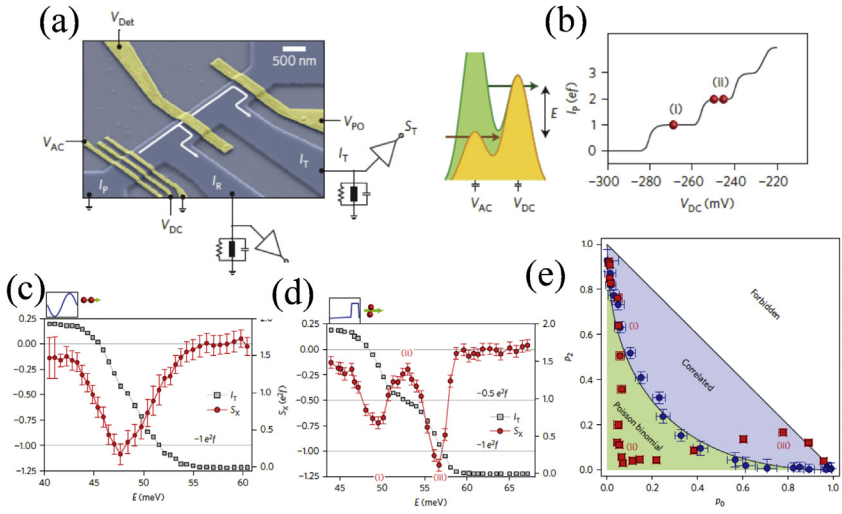


Figure 6.42 HBT experiment with single electron pump. (a) sample used. The gates controlled by voltages V_{ac} and V_{dc} are used to operate the electron pump. The white line stand for chiral edge channels; The centre gate forms the beam splitter which partitions electrons between transmitted and reflected leads; (b) Energy diagram of the loading (orange [light grey in print version]) and unloading (green [grey in print version]) phase of the dynamic quantum dot. The pump current is shown on the right versus V_{dc}. Current plateaus at ef_f and $2ef_d$ are observed; (c) noise and transmitted current versus the energy barrier of the beam splitter for adiabatic two-electron pumping; (d) noise and transmitted current versus the energy barrier of the beam splitter for adiabatic two-electron pumping; (e) the filled blue (dark grey in print version) circles correspond to probabilities $p_{0,2}$ deduced from the measurements presented in (c) and the filled red (grey in print version) squares to measurement done in (d). The latter indicate correlated scattering due to interactions. *Adapted from Ref. [159].*

An HOM experiment requires the physical separation of the two input and the two output channels of a beam splitter. In quantum conductors, this is realised by using the chiral edge channels which form in a strong perpendicular magnetic field. The similarity between the photonic and electronic set-up allows the treatment of both photonic (bosonic) correlations and electronic (fermionic) correlations on the same footing, as shown in Fig. 6.43. The beam splitter is characterised by transmission amplitude t and reflection amplitude ir, with r, t real and $t^2 = T$ and $r^2 = R = 1 - T$. To understand the nature of HOM interference, let us consider the coincidence probability of finding two particles and $\tau = 0$. The amplitude of probability b (1,2) to find, say, particle (1) and (2) on the top and bottom outputs, respectively, is given by the sum of two contributions. Either (1)

arrived at the top input and (2) at the bottom, input amplitude a (1, 2) or particle (2) arrived at the top input and (1) at the bottom, amplitude a (2, 1). These two-particle paths give the following:

$$\begin{aligned} b(1,2) &= t^2 a(1,2) + (ir)^2 a(2,1) \\ &= \left(T - Re^{i\Theta_S}\right) a(1,2) \end{aligned} \tag{6.81}$$

We have used the indistinguishably of the particles: $a(1,2) = a(2,1)e^{i\Theta_S}$ and the statistical angle $\Theta_S = 0$ or $= \pi$ which characterises bosonic and fermionic statistics, respectively. The coincidence probability $P(1,2) = |b(1,2)|^2$ is, for a semi-transparent beam splitter as follows:

$$P(1,2) = \frac{1}{2}(1 - \cos\Theta_S) \tag{6.82}$$

Photon bunching leads to no coincidence, whilst Pauli exclusion leads to perfect coincidence.

The first HOM experiment was done by Bocquillon et al. [162]. Two MC-SES sources, realised in a 2DEG, were used with energy level separation of 1.8 Kelvin at 100 mk temperature. The sources, synchronously, driven at $f_d = 2.1$ GHz, alternately inject a single electron and a single hole into the outer edge channel of the three co-propagating edge channels which form at filling factor $\nu = 3$. A centre QPC, used as a beam splitter and tuned at transmission 1/2, performs the two-particle quantum mixing. No cross-correlation was used, but only the auto-correlation of low frequency fluctuations of the current measured at one beam-splitter output. Suppression of noise for $\tau = 0$, as expected for fermionic anti-bunching was observed but was not complete due to the Coulomb interaction of the injected carriers with the unused co-propagating edge channels. Quantitative analysis and discussion can be found in Refs. [163–165] (Fig. 6.43).

Levitons provide single electron wavepackets with good immunity regarding de-coherence as these excitations are injected at the Fermi energy. Dubois et al. [96] performed HOM correlation of Levitons. Here, the HOM shot noise perfectly vanished for zero time delay proving the perfect indistinguishability of the levitons sent from two independent contacts. Fig. 6.44 shows the results obtained for single charged levitons in Ref. [96] and for doubly charged levitons in Ref. [166]. For levitons of charge e, the noise is $\propto (1 - |\langle\psi_1(0)|\psi_1(\tau)\rangle|^2)$ whilst for levitons with charge $2e$, the noise is $\propto (2 - |\langle\psi_1(0)|\psi_1(\tau)\rangle|^2 - |\langle\psi_2(0)|\psi_2(\tau)\rangle|^2)$ where $\psi_1(t)$ and $\psi_2(t)$ are the first and second orthogonal time-resolved Lorentzian wavepackets as

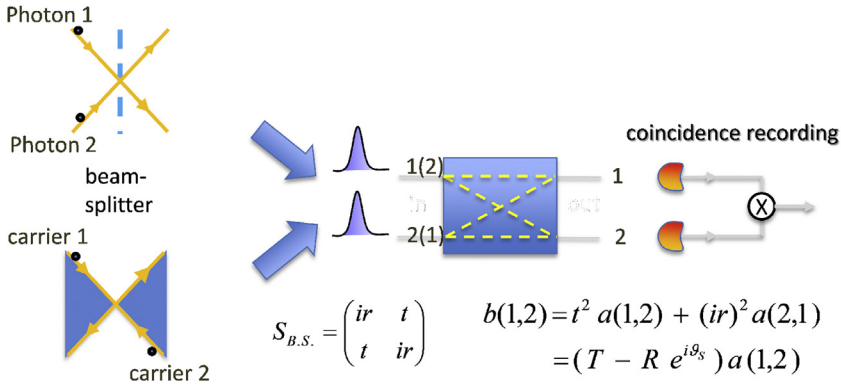


Figure 6.43 Principle of HOM correlation. Using chiral edge channels, quantum conductors can realise experiments similar to photonic experiments, where a QPC is used as a beam splitter for incoming carriers and chiral edge channels ensure separation of the input and output carrier beams. Using a beam splitter scattering matrix, one can treat photons and electrons in a similar way. However, they have different statistical angles Θ_S when exchanging two particles: $\Theta_S = 0$ for photons, whilst $\Theta_S = \pi$ for fermions. This leads to no coincidence events for photons (for $R = T = 1/2$) and full coincidence events for fermions.

described in Refs. [166,167]. The quantities $|\langle\psi_{1(2)}(0)|\psi_{1(2)}(\tau)\rangle|^2$ are the second-order correlation function $g_2(\tau)$ familiar to quantum optics and introduced for fermions in Refs. [164,168]. The leviton source provides a convenient approach to explore HOM correlations with multiple electrons.

HOM correlations have been also made using voltage pulse shapes injecting nonminimal excitation such as single electron sine wave sources in Ref. [96], supplementary information, see also [166,167]. Interestingly, the cloud of neutral excitation surrounding the injected electrons was shown to give perfect anti-bunching and zero HOM noise at $\tau = 0$. HOM noise measurements with sinewave pulses have also been made using a tunnel junction instead of a QPC [169,170]. More exotic injection schemes have also been considered. In Ref. [171], Glattli et al. theoretically considered the injection of a pseudorandom binary sequence of Levitons carrying integer charge. They calculate the HOM noise for two identical sequences of levitons but with a time delay τ between the sequences. At $\tau = 0$, a vanishing noise is observed. This shows that coherent random binary sequences of levitons could be used for quantum information applications in electron quantum optics. Interestingly, when $\tau = kT$, a multiple $k \neq 0$ of the injection period, the HOM shows a 50% dip that can be understood as due to

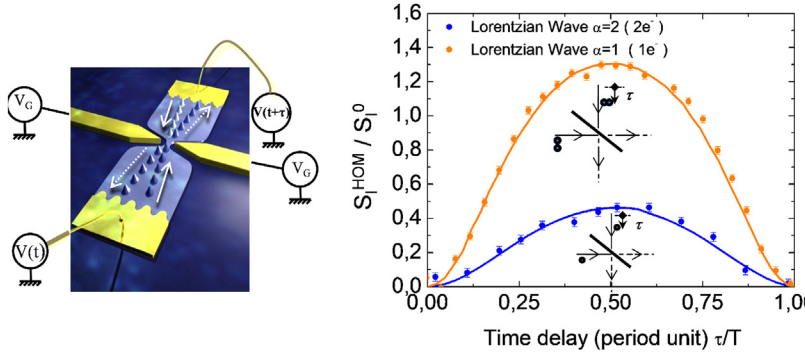


Figure 6.44 Electronic Hong Ou Mandel correlation for charge *e* and 2*e* levitons, adapted from Refs. [96] and [166]. Left: sample and measurement principle. The sample is a 2DEG cooled to 30 mK. 6 GHz periodic Lorentzian voltage pulses of 30ps width are applied to the contacts. A QPC tuned to a transmission of 0.5 partitions the incoming levitons. Cross-correlation of the resulting HOM current fluctuations is measured; Right: HOM shot noise measurements. The noise is normalised to independent (classical) partition noise. Perfect indistinguishability is shown by the vanishing shot noise for time delay $\tau = 0$ and $\tau = T$, where T is the 166 ps injection period. The graph shows HOM measurements for doubly and single charge levitons. The noise is a maximum for $\tau = T/2$ for which levitons are at most distinguishable. The solid lines are a comparison with predictions with no adjustable parameters.

the 50% probability to find both zero leviton or one leviton at each beam splitter inputs (Fig. 6.44).

HOM experiments have also been used to study the fractional quantum Hall effect where it was predicted that injecting charge e levitons into a $\nu = 1/3$ fractional edge channel leads to fermionic correlation whilst signs of anyonic statistics may have been naively expected [172]. HOM correlations could also be studied in a 2D topological insulator [173] or with bogoliubov quasi-particles in hybrid superconductor/normal conductors [174].

4.5.4 Quantum State Tomography using shot noise

A final application of electronic HOM correlation is electron Quantum State Tomography (QST) [177]. The aim of quantum tomography is to have the most complete description of a wavefunction ϕ. Because of Heisenberg uncertainties, a quantum measurement can give information on a wavefunction with respect to only one of the two conjugate variables (energy or time or momentum or position) giving, for example, $\psi(t)$ or $\psi(E)$ depending on the chosen variable. The aim of QST is to obtain the pseudo-probability distribution called the Wigner function which describes

the wavefunction as a function of the two conjugate variables $W(t, E)$. It is given by the following:

$$W(t,E) = \int_{-\infty}^{\infty} d\delta \langle \phi^*(E+\delta/2)\phi(E-\delta/2)\rangle e^{-i\delta t/\hbar} \tag{6.83}$$

A similar expression can be written by interchanging energy and time variables in the integral. Performing QST requires a lot of measurements done in different basis to 'slice' the wavefunction and to provide a complete representation. To do that, it is more convenient to use the energy representation of the wavefunction ϕ and determine the energy density matrix $\langle \phi^*(E')\phi(E)\rangle$.

Remarkably, HOM-like shot noise measurements can be used to determine the energy density matrix. The idea developed by Grenier et al. [175], see also [176], has been experimentally implemented for the QST of a Leviton by Jullien et al. [33,177]. An important simplification is brought by the periodic injection of levitons at frequency ν_0. This implies that all energy dependencies vary by step multiple of $h\nu_0$, such that only the diagonal elements of the energy density matrix $\langle \phi^*(E+kh\nu_0)\phi(E)\rangle$ have to be measured, k integer. Let us see how shot noise can help to do this. The diagonal element $\langle \phi^*(E)\phi(E)\rangle$ is the energy distribution of an electron. To measure it, one can send the leviton to be analysed at, say, the left input of a QPC beam splitter and use the anti-bunching with electrons emitted by the right reservoir and look at the resulting partition noise suppression. If $S_I^0 = 2e^2\nu_0 D(1-D)$ is the partition noise of levitons alone and V_R the voltage applied on the right reservoir, the anti-bunching gives the noise

$$S_I = S_I^0\left(1 - \int_0^{eV_R} d\varepsilon |\phi(\varepsilon)|^2\right) \tag{6.84}$$

The term in the integral denotes the components of leviton partition noise which are suppressed by anti-bunching with the electrons emitted by the right reservoir in the energy range eV_R. Differentiating the noise with respect to voltage provides the energy spectroscopy of the leviton and a measure of the diagonal part of the energy density matrix. The results of the shot noise spectroscopy are shown in Fig. 6.45a which gives the diagonal part of the energy density matrix. How to access the nondiagonal elements? The idea is to use the same recipe but putting the electrons from the right reservoir in a coherent superposition of energies $\varepsilon, \varepsilon + kh\nu_0$ and $\varepsilon - kh\nu_0$ such that the right reservoir electrons anti-bunch with the periodic

levitons at those energies. This is done by superimposing onto the DC voltage V_R a weak AC voltage $eV_{Lo}(t) = \eta_{Lo} h\nu_0 \cos 2k\pi\nu_0(t+\tau)$ with $\eta_{Lo} << 1$. The difference between the noises measured when V_{Lo} is on and V_{Lo} is off gives access to the off-diagonal elements of the density matrix and is given by the following expression (Fig. 6.45) :

$$\Delta S_I^k = S_I^0 2k\eta_{Lo} \cos 2\pi\nu_0\tau \times \int_0^{eV_R} (\phi(\varepsilon)^*\phi(\varepsilon + kh\nu_0) - \phi(\varepsilon)^*\phi(\varepsilon - kh\nu_0)) \tag{6.85}$$

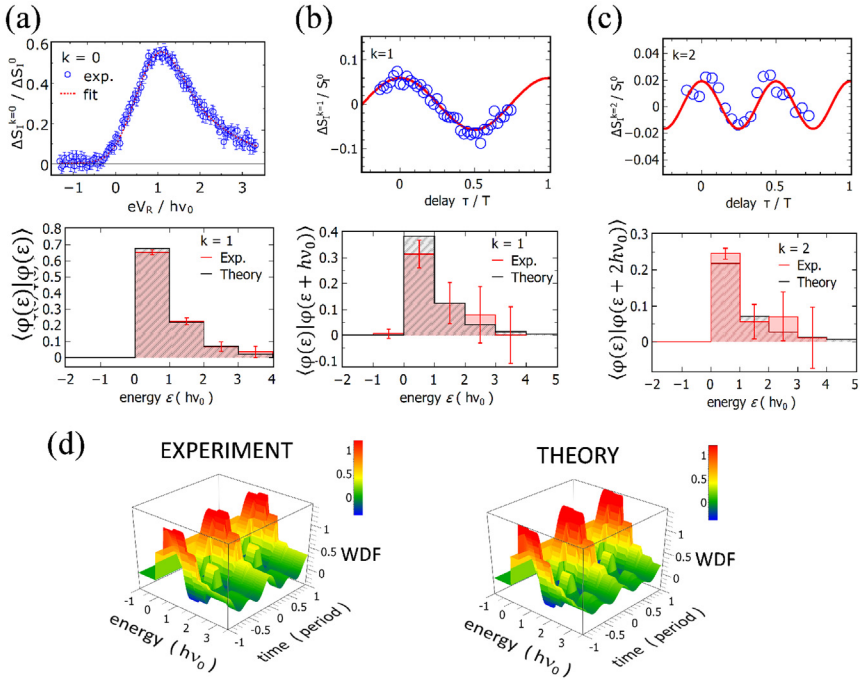


Figure 6.45 Quantum State Tomography of a Leviton: (a) shot noise measurements (top) from which the Leviton energy distribution (bottom) can be extracted. Note the absence of components for negative energy, a hallmark of levitons: (b) Excess shot noise measurements (top) versus time delay τ for DC voltage $eV_R = h\nu_0$ for $k = 1$. The oscillation demonstrates coherence and validates Eq. 6.85. Repeating the measurements for various V_R gives the off-diagonal energy matrix element (bottom) for $k = 1$; (c) same for $k = 2$. (d) 3D representation of the experimental and theoretical Wigner Distribution Function versus the conjugate time and energy variables. *Adapted from [177].*

In the above equations, 6.84 and 6.85, thermal effects are neglected. This is reasonable for the experimental conditions of Jullien et al. where the 6 GHz injection frequency and the 30 mK electronic temperature gave a large ratio $h\nu_0/k_BT = 9.7$.

Figs. 6.45(b) and (c) show the experimental results for $k = 1$ and $k = 2$. The clear oscillation of the excess noise versus τ measured when $eV_R = h\nu_0$ provides evidence of strong quantum coherence. Varying V_R and differentiating gives the off-diagonal elements. In (d) the (partially) re-constructed Wigner function W*(t, E)* is shown. It agrees well with the theoretical partial Wigner function when contributions $k \geq 3$ are disregarded. The results of Jullien et al. [177] provide the first example of experimental QST of an electron in condensed matter.

QST experiments have been reproduced for sinewave pulses and Lorentzian pulses carrying charge e and $2e$ by Bisognin et al. [178]. Here, the smaller 5 GHz injection frequency and larger 80 mK electron temperature giving a smaller ratio $h\nu_0/k_BT \simeq 3$ implied to include the strong effects of thermal excitations. QST has also been done not using noise but using an energy and time dependent barrier to sample the wavefunction in time and energy by Fletcher et al. [179].

5. Conclusion

When a current flows through a conductor, fundamental fluctuations of the current appear which are due to the discrete nature of the charge carrier. The measurement of this shot noise brings important information about the nature (diffusive, ballistic, chaotic, etc.) of the electron transport through the conductor which is not obtained by measuring just conductance. Shot noise has been used as a remarkable diagnostic of quantum transport in semiconductor nanodevices. Similarly, shot noise provides an accurate way to measure the charge of the carrier, and has been found an efficient tool to understand the nature of the quantum phases of the fractional quantum Hall effect. All this has led to the development of ultra-sensitive and very reliable low frequency and high frequency noise measurements.

The electronic current noise being analogous to the photon noise, the study of shot noise has led to the development of electron quantum optics whose aim is to exploit the analogy between electrons and photons. This has motivated physicists to invent the electronic analog of single photon sources. These single electron sources have enabled the study of quantum transport at the level of a single electron. By periodically injecting single electrons into

the conductor, the measurement of the resulting current noise gave remarkable information on the sources. This has also enabled new experiments with electrons analogous to the HBT and to the HOM experiments for photons as well as the possibility to perform the quantum tomography of the wave function of a single electron. The statistics of the electron flow in a quantum conductor is the generalisation of current noise measurements. Whilst some experiments have started to address this topic, measurement tools adapted to investigate a fully quantum regime are missing. This would probably require the detection of single electrons arriving in each contacts of a conductor.

During the last 20 years, the measurement of shot noise became a standard tool for the understanding of the physics of quantum conductor. We believe that shot noise measurements will participate to new exciting developments of quantum transport and electron quantum optics in the future.

References

[1] Blanter Y, Büttiker M. Shot noise in mesoscopic conductors. Phys Rep 2000;336:1.
[2] Kobayashi K. What can we learn from noise?–Mesoscopic nonequilibrium statistical physics. Proc Jpn Acad Ser B 2016;92:204.
[3] Beenakker C, Schönenberger C. Quantum shot noise. Phys Today 2003;56:37–42.
[4] Lesovik GB. Excess quantum noise in 2D ballistic point contacts. Pis'ma Zh Eksp Teor Fiz 1989;49:513. JETP Lett. 49, 592 (1989).
[5] Martin T, Landauer R. Wave packet approach to noise in multichannel mesoscopic systems. Phys Rev B 1992;45:1742.
[6] Reznikov M, Heiblum M, Shtrikman H, Mahalu D. Temporal correlation of electrons: suppression of shot noise in a ballistic quantum point contact. Phys Rev Lett 1995;75:3340.
[7] Kumar A, Saminadayar L, Glattli DC, Jin Y, Etienne B. Experimental test of the quantum shot noise reduction theory. Phys Rev Lett 1996;76:2778.
[8] van Wees BJ, van Houten H, Beenakker CWJ, et al. Quantized conductance of point contacts in a two dimensional electron gas. Phys Rev Lett 1988;60:848.
[9] Wharam DA, Thornton TJ, Newbury R, Pepper M, Ahmed H, Frost JEF, Hasko DG, Peacock DC, Ritchie DA, Jones GAC. One-dimensional transport and the quantisation of the ballistic resistance. J Phys C: Solid State Phys 1988;21:L209.
[10] Buttiker M. Scattering theory of current and intensity noise correlations in conductors and wave guides. Phys Rev B 1992;46:12485.
[11] Lesovik GB, Sadovskyy IA. Scattering matrix approach to the description of quantum electron transport. Phys-Usp 2011;54:1007.
[12] Dorokhov ON. On the coexistence of localized and extended electronic states in the metallic phase. Solid State Commun 1984;51:381.
[13] Beenakker CWJ, Buttiker M. Suppression of shot noise in metallic diffusive conductors. Phys Rev B 1992;46:1889.
[14] Nagaev KE. On the shot noise in dirty metal contacts. Phys Lett 1992;169:103–7.
[15] Jalabert RA, Pichard J-L, Beenakker CWJ. Universal quantum signatures of Chaos in ballistic transport. Europhys Lett 1994;27:255.
[16] Beenakker CWJ. Random-matrix theory of quantum transport. Rev Mod Phys July 1997;69(3).

[17] Nazarov YV. In: Cerdeira HA, Kramer B, Schon G, editors. Quantum dynamics of submicron structures. NATO ASI series E291. Dordrecht: Kluwer; 1995.
[18] Lesovik GB, Loosen R. On the detection of finite-frequency current fluctuations. JETP Lett 1997;65:295.
[19] Creux M, Crépieux A, Martin T. Finite-frequency noise cross correlations of a mesoscopic circuit: a measurement method using a resonant circuit. Phys Rev B 2006;74:115323.
[20] Gavish U, Levinson Y, Imry Y. Detection of quantum noise. Phys Rev B 2000;62(R):R10637.
[21] Schottky W. Uber spontane Stromschwankungen in verschiedenen Elektrizitatsleitern. Ann Phys (Berlin) 1918;362:541.
[22] Nyquist H. Thermal agitation of electric charge in conductors. Phys Rev 1928;32:110.
[23] Lesovik GB, Levitov LS. Noise in an ac biased junction: nonstationary Aharonov-Bohm effect. Phys Rev Lett 1994;72:538.
[24] Moskalets M, Büttiker M. Floquet scattering theory of quantum pumps. Phys Rev B 2002;66:205320.
[25] Pedersen MH, Buttiker M. Scattering theory of photon-assisted electron transport. Phys Rev B 1998;58:12993.
[26] De-Picciotto R, Reznikov M, Heibrum M, Umansky V, Bunin G, Mahalu D. Direct observation of a fractional charge. Nature 1997;389:162.
[27] Buttiker M. Quantized transmission of a saddle-point constriction. Phys Rev 1990;B41:7906.
[28] Li Y éP, Tsui DC, Heremans JJ, Simmons JA. Low-frequency noise in transport through quantum point contacts. Appl Phys Lett 1990;57:774.
[29] Liefrink F, Scholten AJ, Dekker C, Eppenga R, van Houten H, Foxon CT. Low-frequency noise of quantum point contacts in the ballistic and quantum Hall regime. Physica B 1991;175:213.
[30] DiCarlo L, Zhang Y, McClure DT, Marcus CM, Pfeiffer LN, West KW. Rev Sci Instrum 2006;77:073906.
[31] Hashisaka M, Yamauchi Y, Nakamura S, Kasai S, Kobayashi K, Ono T. Measurement for quantum shot noise in a quantum point contact at low temperature. J Phys Conf Ser 2008;109:012013.
[32] Arakawa T, Nishihara Y, Maeda M, Norimoto S, Kobayashi K. Cryogenic amplifier for shot noise measurement at 20 mK. Appl Phys Lett 2013;103:172104.
[33] Jullien T. Mesoscopic few-electron voltage pulse source. Université Paris Sud - Paris XI; 2014. https://tel.archives-ouvertes.fr/tel-01158785.
[34] Hashisaka M, Ota T, Yamagishi M, Fujisawa T, Musraki K. Cross-correlation measurement of quantum shot noise using homemade transimpedance amplifier. Rev Sci Instrum 2014;85:054704.
[35] Muro T, Nishihara Y, Norimoto S, Ferrier M, Arakawa T, Kobayashi K, Ihn T, Rössler C, Ensslin K, Reichl C, Wegscheider W. Finite shot noise and electron heating at quantized conductance in high-mobility quantum point contacts. Phys Rev B 2016;93:195411.
[36] Yang W, Wei J. Cryogenic amplifier with low input-referred voltage noise calibrated by shot noise measurement. Chin Phys B 2018;27(6):060702.
[37] Nishihara Y, Nakamura S, Kobayashi K, Ono T, Kohda M, Nitta J. Shot noise suppression in InGaAs/InGaAsP quantum channels. Appl Phys Lett 2012;100(20):203111.
[38] Kumar M, Avriller R, Levy Yeyati A, van Ruitenbeek JM. Detection of vibration-mode scattering in electronic shot noise. Phys Rev Lett 2012;108:146602.

[39] Vardimon R, Klionsky M, Tal O. Experimental determination of conduction channels in atomic-scale conductors based on shot noise measurements. Phys Rev B 2013; 88(R):161404.
[40] Chen R, Wheeler PJ, Di Ventra M, Natelson D. Enhanced noise at high bias in atomic-scale Au break junctions. Sci Rep 2015;4:4221.
[41] Tewari S, Sabater C, Kumar M, Stahl S, Crama B, van Ruiten-beek JM. Fast and accurate shot noise measurements on atomic-size junctions in the MHz regime. Rev Sci Instrum 2017;88(9):093903.
[42] Mohr M, Jasper-Toennies T, Weismann A, Frederiksen T, Garcia-Lekue A, Ulrich S, Herges R, Berndt R. Conductance channels of a platform molecule on Au(111) probed with shot noise. Phys Rev B 2019;99:245417.
[43] Herrmann LG, Delattre T, Morfin P, Berroir J-M, Plaçais B, Glattli DC, Kontos T. Shot noise in Fabry-Perot interferometers based on carbon nanotubes. Phys Rev Lett 2017;99:156804.
[44] Thomas KJ, Nicholls JT, Simmons MY, Pepper M, Mace DR, Ritchie DA. Possible spin polarization in a one-dimensional electron gas. Phys Rev Lett 1996;77:135.
[45] de Picciotto R, Pfeiffer LN, Baldwin KW, West KW. Temperature-dependent 0.7 structure in the conductance of cleaved-edge-overgrowth one-dimensional wires. Phys Rev B 2005;72:033319.
[46] Bagraev NT, Buravlev AD, Klyachkin LE, et al. Quantized conductance in silicon quantum wires. Semiconductors 2002;36:439.
[47] Chou HY, Luscher S, Goldhaber-Gordon D, et al. High-quality quantum point contacts in GaN/AlGaN heterostructures. Appl Phys Lett 2005;86:073108.
[48] Simmonds PJ, Sfigakis F, Beere HE, et al. Quantum transport in $In_{0.75}Ga_{0.25}As$ quantum wires. Appl Phys Lett 2008;92:152108.
[49] Spivak B, Zhou F. Ferromagnetic correlations in quasi-one-dimensional conducting channels. Phys Rev B 2000;61:16730.
[50] Reimann SM, Koskinen M, Manninen M. End states due to spin-Peierls transition in quantum wires. Phys Rev B 1999;59:1613.
[51] Seelig G, Matveev KA. Electron-phonon scattering in quantum point contacts. Phys Rev Lett 2003;90:176804.
[52] Rejec T, Meir Y. Magnetic impurity formation in quantum point contacts. Nature 2006;442:900.
[53] Cronenwett SM, Lynch HJ, Goldhaber-Gordon D, et al. Low-temperature fate of the 0.7 structure in a point contact: a Kondo-like correlated state in an open system. Phys Rev Lett 2002;88:226805.
[54] Meir Y, Hirose K, Wingreen NS. Kondo model for the "0.7 anomaly" in transport through a quantum point contact. Phys Rev Lett 2002;89:196802.
[55] Meir Y. The theory of the '0.7 anomaly' in quantum point contacts. J Phys: Condens Matter 2008;20:164208.
[56] Graham AC, Thomas KJ, Pepper M, Cooper NR, Simmons MY, Ritchie DA. Interaction effects at crossings of spin-polarized one-dimensional subbands. Phys Rev Lett 2003;91:136404.
[57] Kristensen A, Bruus H, Hensen AE, et al. Bias and temperature dependence of the 0.7 conductance anomaly in quantum point contacts. Phys Rev B 2000;62:10950.
[58] Appleyard NJ, Nicholls JT, Simmons MY, Tribe WR, Pepper M. Thermometer for the 2D electron gas using 1D thermopower. Phys Rev Lett 1998;81:3491.
[59] Chiatta O, Nicholls JT, Proskuryakov Y, Lumpkin N, Farrer I, Ritchie DA. Quantum thermal conductance of electrons in a one-dimensional wire. Phys Rev Lett 2006;97:056601.
[60] Roche P, Segala J, Glattli DC, et al. Fano factor reduction on the 0.7 conductance structure of a ballistic one-dimensional wire. Phys Rev Lett 2004;93:116602.

[61] DiCarlo L, Zhang Y, McClure DT, Reilly DJ, Marcus CM, Pfeiffer LN, West KW. Shot-noise signatures of 0.7 structure and spin in a quantum point contact. Phys Rev Lett 2006;97:036810.
[62] Nakamura S, Hashisaka M, Yamauchi Y, Kasai S, Ono T, Kobayashi K. Conductance anomaly and Fano factor reduction in quantum point contacts. Phys Rev B 2009;79: 201308(R.
[63] Jeong H. Shot noise suppression in a quantum point contact with short channel length. Chin Phys Lett 2015;32:077301.
[64] Liefrink F, Dijkhuis JI, de Jong MJM, Molenkamp LW, van Houten H. Experimental study of reduced shot noise in a diffusive mesoscopic conductor. Phys Rev B 1994; 49(R):14066.
[65] Steinbach AH, Martinis JM, Devoret MH. Observation of hot-electron shot noise in a metallic resistor. Phys Rev Lett 1996;76:3806.
[66] Henny M, Oberholzer S, Strunk C, Schönenberger C. 1/3-shot-noise suppression in diffusive nanowires. Phys Rev B 1999;59:2871.
[67] Tworzydlo J, Trauzettel B, Titov M, Rycerz A, Beenakker CWJ. Sub-poissonian shot noise in graphene. Phys Rev Lett 2006;96:246802.
[68] Lewenkopf CH, Mucciolo ER, Castro Neto AH. Numerical studies of conductivity and Fano factor in disordered graphene. Phys Rev B 2008;77:081410(R).
[69] DiCarlo L, Williams JR, Zhang Y, McClure DT, Marcus CM. Shot noise in graphene. Phys Rev Lett 2008;100:156801.
[70] Danneau R, Wu F, Craciun MF, Russo S, Tomi MY, Salmilehto J, Morpurgo AF, Hakonen PJ. Shot noise in ballistic graphene. Phys Rev Lett 2008;100:196802.
[71] Tan ZB, Puska A, Nieminen T, Duerr F, Gould C, Molenkamp LW, Trauzettel B, Hakonen PJ. Shot noise in lithographically patterned graphene nanoribbons. Phys Rev B 2013;88:245415.
[72] Mostovov A. Quantum shot noise in graphene. Ph-D Thesis. Université Pierre et Marie Curie - Paris VI; 2014.
[73] Oberholzer S, Sukhorukov EV, Strunk C, Schönenberger C, Heinzel T, Holland M. Shot noise by quantum scattering in chaotic cavities. Phys Rev Lett 2001;86:2114.
[74] Oberholzer S, Sukhorukov EV, Schönenberger C. Crossover between classical and quantum shot noise in chaotic cavities. Nature 2002;415:765.
[75] Levitov LS, Lesovik GB. Charge-transport statistics in quantum conductors. JETP Lett 1992;55:555.
[76] Levitov LS, Lesovik GB. Quantum measurement in electric circuit, cond-mat/ 9401004. 1994.
[77] Levitov LS, Lee H, Lesovik GB. Electron counting statistics and coherent states of electric current. J Math Phys(NY) 1996;37:4845.
[78] Averin DV, Sukhorukov EV. Counting statistics and detector properties of quantum point contacts. Phys Rev Lett 2005;95:126803.
[79] Field M, Smith CG, Pepper M, Ritchie DA, Frost JEF, Jones GAC, Hasko DG. Measurements of Coulomb blockade with a noninvasive voltage probe. Phys Rev Lett 1993;70:1311.
[80] Lu W, Ji Z, Pfeiffer L, West KW, Rimberg AJ. Real-time detection of electron tunnelling in a quantum dot. Nature (London) 2003;423:422.
[81] Fujisawa T, Hayashi T, Hirayama Y, Cheong HD, Jeong YH. Electron counting of single-electron tunneling current. Appl Phys Lett 2004;84:2343.
[82] Bylander J, Duty T, Delsing P. Current measurement by real-time counting of single electrons. Nature (London) 2005;434:361.
[83] Gustavsson S, Leturcq R, Simovič B, Schleser R, Ihn T, Studerus P, Ensslin K, Driscoll DC, Gossard AC. Counting statistics of single electron transport in a quantum dot. Phys Rev Lett 2006;96:076605.

[84] Fujisawa T, Hayashi T, Tomita R, Hirayama Y. Bidirectional counting of single electrons. Science 2006;312:1634–6.
[85] Reulet B, Senzier J, Prober DE. Environmental effects in the third moment of voltage fluctuations in a tunnel junction. Phys Rev Lett 2003;91:196601.
[86] Bomze Y, Gershon G, Shovkun D, Levitov LS, Reznikov M. Measurement of counting statistics of electron transport in a tunnel junction. Phys Rev Lett 2005; 95:176601.
[87] Schoelkopf RJ, Burke PJ, Kozhevnikov AA, Prober DE, Rooks MJ. Frequency dependence of shot noise in a diffusive mesoscopic conductor. Phys Rev Lett 1997;78:3370.
[88] Deblock R, Onac E, Gurevich L, Kouwenhoven LP. Detection of quantum noise from an electrically driven two-level system. Science 2003;301:203.
[89] Onac E, Balestro F, Willems van Beveren LH, Hartmann U, Nazarov YV, Kouwenhoven LP. Using a quantum dot as a high-frequency shot noise detector. Phys Rev Lett 2006;96:176601. Published 1 May 2006.
[90] Zakka-Bajjani E, Ségala J, Portier F, Roche P, Glattli DC, Cavanna A, Jin Y. Experimental test of the high-frequency quantum shot noise theory in a quantum point contact. Phys Rev Lett 2007;99:236803.
[91] Gustavsson S, Studer M, Leturcq R, Ihn T, Ensslin K, Driscoll DC, Gossard AC. Frequency-selective single-photon detection using a double quantum dot. Phys Rev Lett 2007;99:206804.
[92] Schoelkopf RJ, Kozhevnikov AA, Prober DE, Rooks MJ. Observation of "photon-assisted" shot noise in a phase-coherent conductor. Phys Rev Lett 1998;80:2437.
[93] Kozhevnikov AA, Schoelkopf RJ, Prober DE. Observation of photon-assisted noise in a diffusive normal metal-superconductor junction. Phys Rev Lett 2000;84:3398.
[94] Reydellet L-H, Roche P, Glattli DC, Etienne B, Jin Y. Quantum partition noise of photon-created electron-hole pairs. Phys Rev Lett 2003;90:176803.
[95] Gabelli J, Reulet B. Dynamics of quantum noise in a tunnel junction under ac excitation. Phys Rev Lett 2008;100:026601.
[96] Dubois J, Jullien T, Portier F, Roche P, Cavanna A, Jin Y, Wegscheider W, Roulleau P, Glattli DC. Minimal-excitation states for electron quantum optics using levitons. Nature 2013;502:659.
[97] Jompol Y, Roulleau P, Jullien T, Roche B, Farrer I, Ritchie DA, Glattli DC. Detecting noise with shot noise using on-chip photon detector. Nat Commun 2015;6:6130.
[98] Parmentier FD, Serkovic-Loli LN, Roulleau P, Glattli DC. Photon-assisted shot noise in graphene in the terahertz range. Phys Rev Lett 2016;116:227401.
[99] Kapfer M, Roulleau P, Santin M, Farrer I, Ritchie DA, Glattli1 DC. A Josephson relation for fractionally charged anyons. Science 2019;363:846–9.
[100] Klitzing Kv, Dorda G, Pepper M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. Phys Rev Lett 1980; 45:494.
[101] Tsui DC, Stormer HL, Gossard AC. Two-dimensional magnetotransport in the extreme quantum limit. Phys Rev Lett 1982;48:1559.
[102] Laughlin RB. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. Phys Rev Lett 1983;50:1395.
[103] Kane CL, Fisher MPA. Nonequilibrium noise and fractional charge in the quantum Hall effect. Phys Rev Lett 1994;72:724–7.
[104] Tomonaga S. Remarks on Bloch's method of sound waves applied to many-fermion problems. Prog Theor Phys (Kyoto) 1950;5:544.
[105] Luttinger JM. An exactly soluble model of a many-fermion system. J Math Phys 1963; 4:1154.

[106] Chamon C de C, Freed DE, Wen XG. Tunneling and quantum noise in one-dimensional Luttinger liquids. Phys Rev B 1995;51:2363–79.
[107] Fendley P, Ludwig AWW, Saleur H. Exact nonequilibrium dc shot noise in Luttinger liquids and fractional quantum Hall devices. Phys Rev Lett 1995;75:2196.
[108] Saminadayar L, Glattli DC, Jin Y, Etienne B. Observation of the e/3 fractionally charged Laughlin quasiparticle. Phys Rev Lett 1997;79:2526.
[109] Reznikov M, de Picciotto R, Griffiths TG, Heiblum M, Umansky V. Observation of quasiparticles with one-fifth of an electron's charge. Nature 1999;399:238.
[110] Fendley P, Saleur H. Nonequilibrium DC noise in a Luttinger liquid with an impurity. Phys Rev B 1996;54:10845.
[111] C Glattli D, Rodriguez V, Perrin H, Roche P, Jin Y, Etienne B. Shot noise and the Luttinger liquid-like properties of the FQHE. Phys E Low-dimens Syst Nanostruct 2000;6:22.
[112] Roche P, Rodriguez V, Christian Glattli D. Quantum Hall effect, chiral Luttinger liquids and fractional charges. Compt Rendus Phys 2002;3:717.
[113] Griffiths TG, Comforti E, Heiblum M, Stern A, Umansky V. Evolution of quasiparticle charge in the fractional quantum Hall regime. Phys Rev Lett 2000;85:3918.
[114] Trauzettel B, Roche P, Glattli DC, Saleur H. Effect of interactions on the noise of chiral Luttinger liquid systems. Phys Rev B 2004;70:233301.
[115] Feldman DE, Heiblum M. Why a noninteracting model works for shot noise in fractional charge experiments. Phys Rev B 2017;95:115308.
[116] Comforti E, Chung YC, Heiblum M, Umansky V. Multiple scattering of fractionally charged quasiparticles. Phys Rev Lett 2002;89:066803.
[117] Comforti E, Chung YC, Heiblum M, Umansky V, Mahalu D. Bunching of fractionally charged quasiparticles tunnelling through high-potential barriers. Nature 2002; 416:515.
[118] Dolev M, Heiblum M, Umansky V, Stern A, Mahalu D. Observation of a quarter of an electron charge at the : = 5/2 quantum Hall state. Nature 2008;452:829.
[119] Chung YC, Heiblum M, Umansky V. Scattering of bunched fractionally charged quasiparticles. Phys Rev Lett 2003;91:216804.
[120] Bid A, Ofek N, Heiblum M, Umansky V, Mahalu D. Shot noise and charge at the 2/3 composite fractional quantum Hall state. Phys Rev Lett 2009;103:236802.
[121] Dolev M, Gross Y, Chung YC, Heiblum M, Umansky V, Mahalu D. Dependence of the tunneling quasiparticle charge determined via shot noise measurements on the tunneling barrier and energetics. Phys Rev B 2010;81:161303.
[122] M. Kapfer et al. in preparation Kapfer M. Dynamic of excitations of the Fractional quantum Hall effect : fractional charge and fractional Josephson frequency. Ph-D thesis. Physics: Universite Paris-Saclay; 2018.
[123] Bid A, Nissim O, Inoue H, Heiblum M, Kane CL, Umansky V, Mahalu D. Observation of neutral modes in the fractional quantum Hall regime. Nature 2010;466:585.
[124] Dolev M, Gross Y, Sabo R, Gurman I, Heiblum M, Umansky V, Mahalu D. Characterizing neutral modes of fractional states in the second Landau level. Phys Rev Lett 2011;107:036805.
[125] Gross Y, Dolev M, Heiblum M, Umansky V, Mahalu D. Upstream neutral modes in the fractional quantum Hall effect regime: heat waves or coherent dipoles. Phys Rev Lett 2012;108:226801.
[126] Inoue H, Grivnin A, Ronen Y, Heiblum M, Umansky V, Mahalu D. Proliferation of neutral modes in fractional quantum Hall states. Nat Commun 2014;5:4067.
[127] Hashisaka M, Ota T, Muraki K, Fujisawa T. Shot-noise evidence of fractional quasiparticle creation in a local fractional quantum Hall state. Phys Rev Lett 2015;114: 056802.

[128] Cohen Y, Ronen Y, Yang W, Banitt D, Park J, Heiblum M, Mirlin A, Gefen Y, Umansky V. Synthesizing a $\nu = 2/3$ fractional quantum Hall effect edge state from counter-propagating $\nu = 1$ and $\nu = 1/3$ states. Nat Commun 2019;10:1920.
[129] Shapiro S. Josephson currents in superconducting tunneling: the effect of microwaves and other observations. Phys Rev Lett 1963;11:80–2.
[130] Wen XG. Edge transport properties of the fractional quantum Hall states and weak-impurity scattering of a one-dimensional charge-density wave. Phys Rev B 1991;44: 5708–19.
[131] Crépieux A, Devillard P, Martin T. Photoassisted current and shot noise in the fractional quantum Hall effect. Phys Rev B 2004;69:205302.
[132] Bena C, Safi I. Emission and absorption noise in the fractional quantum Hall effect. Phys Rev B 2007;76:125317.
[133] Chevallier D, Jonckheere T, Paladino E, Falci G, Martin T. Detection of finite-frequency photoassisted shot noise with a resonant circuit. Phys Rev B 2010;81: 205411.
[134] Safi I, Sukhorukov EV. Determination of tunneling charge via current measurements. Eur Phys Lett 2010;91:67008.
[135] Samuelsson P, Sukhorukov EV, Buttiker M. Two-particle Aharonov-Bohm effect and entanglement in the electronic Hanbury Brown-Twiss setup. Phys Rev Lett 2004;92:026805.
[136] Oberholzer MHS, Strunk C, Heinzel T, Ensslin K, Holland M, Schonenberger C. The fermionic Hanbury Brown and Twiss experiment. Science 1999;284:296.
[137] Oberholzer S, Henny M, Strunk C, Schönenberger C, Heinzel T, Ensslin K, Holland M. The Hanbury Brown and Twiss experiment with fermions. Physica E (Amsterdam) 2000;6:314.
[138] Oliver W, Kim J, Liu R, Yamamoto Y. Hanbury Brown and Twiss-type experiment with electrons. Science 1999;284:299.
[139] Neder I, Ofek N, Chung Y, Heiblum M, Mahalu D, Umansky V. Interference between two indistinguishable electrons from independent sources. Nature 2007; 448:pages333–337.
[140] Michler P, Kiraz A, Becher C, Schoenfeld WV, Petroff PM, Zhang L, Hu E, Imamoglu A. A quantum dot single-photon turnstile device. Science 2000;290: 2282–5.
[141] Lounis B, Moerner WE. Single photons on demand from a single molecule at room temperature. Nature 2000;407:491–3.
[142] Liu RC, Odom B, Yamamoto Y, Tarucha S. Quantum interference in electron collision. Nature 1998;391:263–5.
[143] Bauerle C, Glattli DC, Meunier T, Portier F, Roche P, Roulleau P, Takada S, Waintal X. Coherent control of single electrons: a review of current progress. Rep Prog Phys 2018;81:056503. p.33.
[144] Special issue: single-electron control in solid-state devices. Edited by J. Splettstoesser and R. J. Haug Phys Status Solidi B 2017;254(3).
[145] Maire N, Hohls F, Kaestner B, Pierz K, Schumacher HW, Haug RJ. Noise measurement of a quantized charge pump. Appl Phys Lett 2008;92:082112.
[146] Robinson AM, Talyanskii VI. Shot noise in the current of a surface acoustic-wave-driven single-electron pump. Phys Rev Lett 2005;95:247202.
[147] Robinson AM, Talyanskii VI, Pepper M, Cunningham JE, Linfield EH, Ritchie DA. Measurements of noise caused by switching of impurity states and of suppression of shot noise in surface-acoustic-wave-based single-electron pumps. Phys Rev B 2002;65:045313.

[148] Fève G, Mahé A, Berroir J-M, Kontos T, Plaçais B, Glattli DC, Cavanna A, Etienne B, Jin Y. An on-demand coherent single-electron source. Science 2007; 316:1169–72.
[149] Mahé A, Parmentier FD, Bocquillon E, Berroir J-M, Glattli DC, Kontos T, Plaçais B, Fève G, Cavanna A, Jin Y. Current correlations of an on-demand single-electron emitter. Phys Rev B 2010;82:201309(R).
[150] Albert M, Flindt C, Büttiker M. Accuracy of the quantum capacitor as a single-electron source. Phys Rev B 2010;82:041407(R).
[151] Rychkov VS, Polianski ML, Büttiker M. Photon-assisted electron- hole shot noise in multiterminal conductors. Phys Rev B 2005;72:155326.
[152] Keeling J, Klich I, Levitov LS. Minimal excitation states of electrons in one- dimensional wires. Phys Rev Lett 2006;97:116403.
[153] Bocquillon E, Parmentier FD, Grenier C, Berroir J-M, Degiovanni P, Glattli DC, Placais B, Cavanna A, Jin Y, Feve G. Electron quantum optics: partitioning electrons one by one. Phys Rev Lett 2012;108:196803.
[154] Ivanov DA, Lee HW, Levitov LS. Coherent states of alternating current. Phys Rev B 1997;56:6839.
[155] Dubois J, Jullien T, Grenier C, Degiovanni P, Roulleau P, Glattli DC. Integer and fractional charge Lorentzian voltage pulses analyzed in the framework of photon-assisted shot noise. Phys Rev B 2013;88:085301.
[156] Grenier C, Dubois J, Jullien T, Roulleau P, Glattli DC, Degiovanni P. Fractionalization of minimal excitations in integer quantum Hall edge channels. Phys Rev B 2013;88:085302.
[157] Anderson PW. Infrared catastrophe in Fermi gases with local scattering potential. Phys Rev Lett 1967;18:1049–51.
[158] Lee HW, Levitov L. Orthogonality catastrophe in a mesoscopic conductor due to a time-dependent flux. 1993. http://arxiv.org/abs/cond-mat/9312013.
[159] Ubbelohde N, Hohls F, Kashcheyevs V, Wagner T, Fricke L, Kästner B, Pierz K, Schumacher HW, Haug RJ. Partitioning of on-demand electron pairs. Nat Nanotechnol 2015;10:46–9.
[160] Hassler F, Suslov MV, Graf GM, Lebedev MV, Lesovik GB, Blatter G. Wave-packet formalism of full counting statistics. Phys Rev B 2008;78:165330.
[161] Hong CK, Ou ZY, Mandel L. Measurement of subpicosecond time intervals between two photons by interference. Phys Rev Lett 1987;59:2044–6.
[162] Bocquillon E, Freulon V, M Berroir J-, Degiovanni P, Plaçais B, Cavanna A, Jin Y, Feve G. Coherence and indistinguishability of single electrons emitted by independent sources. Science 2013;339:1054–7.
[163] Wahl C, Rech J, Jonckheere T, Martin T. Interactions and charge fractionalization in an electronic Hong-Ou-Mandel interferometer. Phys Rev Lett 2014;112:046802.
[164] Bocquillon E, et al. Electron quantum optics in ballistic chiral conductors. Ann Phys (Berlin) 2014;526(No. 1–2):1–30.
[165] Freulon V, Marguerite A, Berroir J-M, Pla ais B, Cavanna A, Jin Y, Fève G. Hong-Ou-Mandel experiment for temporal investigation of single-electron fractionalization. Nat Commun 2015;6:6854.
[166] Glattli DC, Roulleau P. Hanbury-Brown Twiss noise correlation with time controlled quasi-particles in ballistic quantum conductors. Phys E (Amsterdam, Neth.) 2016;82:99.
[167] Glattli DC, Roulleau P. Levitons for electron quantum optics. Phys Status Solidi B 2017;254:1600650.
[168] Moskalets M. Single-electron second-order correlation function G(2) at nonzero temperatures. Phys Rev B 2018;98:115421.

[169] Vanević M, Gabelli J, Belzig W, Reulet B. Electron and electron-hole quasiparticle states in a driven quantum contact. Phys Rev B 2016;93:041416(R).
[170] Yue XK, Yin Y. Normal and anomalous electron-hole pairs in a quantum conductor driven by a voltage pulse. Phys Rev B 2019;99:235431.
[171] Glattli DC, Roulleau P. Pseudorandom binary injection of levitons for electron quantum optics. Phys Rev B 2018;97:125407.
[172] Ronetti F, Vannucci L, Ferraro D, Jonckheere T, Rech J, Martin T, Sassetti M. Crystallization of levitons in the fractional quantum Hall regime. Phys Rev B 2018;98: 075401.
[173] Ferraro D, Wahl C, Rech J, Jonckheere T, Martin T. Electronic Hong-Ou-Mandel interferometry in two-dimensional topological insulators. Phys Rev B 2014;89: 075407.
[174] Ferraro D, Rech J, Jonckheere T, Martin T. Nonlocal interference and Hong-Ou-Mandel collisions of single Bogoliubov quasiparticles. Phys Rev B 2015;91:075406.
[175] Grenier C, Hervé R, Bocquillon E, Parmentier FD, Plaçais B, Berroir J-M, Fève G, Degiovanni P. Single-electron quantum tomography in quantum Hall edge channels. New J Phys 2011;13:093007.
[176] Ferraro D, Feller A, Ghibaudo A, Thibierge E, Bocquillon E, Feve G, Grenier C, Degiovanni P. Wigner function approach to single electron coherence in quantum Hall edge channels. Phys Rev B 2013;88:205303.
[177] Jullien T, Roulleau P, Roche B, Cavanna A, Jin Y, Glattli DC. Quantum tomography of an electron. Nature 2014;514:pp603–607.
[178] Bisognin R, Marguerite A, Roussel B, Kumar M, Cabart C, Chapdelaine C, Mohammad-Djafari A, Berroir J-M, Bocquillon E, Plaçais B, Cavanna A, Gennser U, Jin Y, Degiovanni P, Fève G. Quantum tomography of electrical currents. Nat Commun 2019;10:3379.
[179] Fletcher JD, Johnson N, Locane E, See P, Griffiths JP, Farrer I, Ritchie DA, Brouwer PW, Kashcheyevs V, Kataoka M. Continuous-variable tomography of solitary electrons. Nat Commun 2019;10:5298.
[180] Bisognin R, Bartolomei H, Kumar M, Safi I, Berroir J-M, Bocquillon E, Plaçais B, Cavanna A, Gennser U, Jin Y, Feve G. Microwave photons emitted by fractionally charged quasi-particles. Nat Commun 2019;10:1708.

CHAPTER SEVEN

Electrical and superconducting transport in topological insulator nanoribbons

Morteza Kayyalha[1], Leonid P. Rokhinson[2,3,*] and Yong P. Chen[2,3]
[1]Department of Electrical Engineering, Pennsylvania State University, University Park, PA, United States
[2]Department of Physics and Astronomy, Purdue University, West Lafayette, IN, United States
[3]Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, United States
*Corresponding author: leonid@purdue.edu

1. Introduction

In solid state physics, materials are divided into two distinct categories: metals, such as gold, and insulators, such as SiO_2. This distinction is based on whether the Fermi energy resides inside the conduction/valence bands (metals) or inside the bandgap (insulators). However, discovery of the integer quantum Hall (QH) effect, which is insulating in bulk whilst conducting along the edges, rattled this distinction. The study of the QH effect has led to a new classification of materials: topological materials [1,2]. In the QH insulator, an external magnetic field breaks the time reversal symmetry and creates Landau levels. The QH states belong to different topological classes compared to the trivial insulators. When insulators from two different topological classes are placed next to each other, the insulating gap must close at their boundary giving rise to edge modes (in two dimensions) and surface states (in three dimensions). In a sense, comparing the topology of a nontrivial band structure and a trivial one is analogous to comparing a rubber band and a Möbius strip; one cannot be smoothly deformed into the other without breaking [2]. The closing of the bandgap at the boundary of topological and trivial insulators is hence analogous to the 'breaking' of the rubber band and converting it into the Möbius strip. This 'topological transition' makes the edge/surface states independent of the material details and insensitive to any smooth deformation of the system. Over the past few years, several topological states of matter, including two-dimensional (2D) [3] and three-dimensional (3D) topological insulators (TIs) [1,2], topological crystalline insulators [4], topological Dirac semimetals and Weyl semimetals [5,6], have been theoretically predicted and experimentally

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00004-6
© 2021 Elsevier Ltd.
All rights reserved.

verified. In this chapter, we focus on 3D TIs, including TI nanoribbons (TINRs), and discuss their electrical and superconducting transport properties.

TIs belong to a class of topological materials where time reversal symmetry is preserved. A strong spin–orbit coupling (SOC) and band inversion in the TI results in the appearance of a nontrivial insulating bulk bandgap and topologically protected surface states. The topological protection refers to the immunity of surface states against backscattering by nonmagnetic impurities such as crystalline defects, and surface roughness [1,2,7]. These topological surface states (TSSs) are spin-helical states with linear Dirac fermion-like energy–momentum dispersion. Early examples of 3D TIs are Bi_2Se_3, Bi_2Te_3, Sb_2Te_3 and their compounds [1,2]. These materials are amongst the most commonly used thermoelectric materials and have been studied extensively over the years [8]. However, only recently, topological origin of thermoelectric and electronic properties became understood. Other examples of 3D TIs include strained HgTe [9] and SmB_6 [10].

Fig. 7.1a and b depict schematics of the surface band dispersion in momentum space and spin texture of the TSS in a real space, respectively. Due to the strong SOC, spin of charge carriers in the TSS is locked perpendicular to the momentum and tangential to the surface. Therefore, the spin-helical TSS may give rise to the current/momentum-induced spin polarisation and, hence, have many potential applications in spintronics [11,12]. Moreover, ferromagnetic doping of the TIs can open an exchange gap in TSS, leading to the quantum anomalous Hall effect [13,14] or topological magnetoelectric effect [15]. The TSS coupled to a conventional superconductor may give rise to the topological superconductivity and Majorana modes [1,16], top candidates to create a fault-tolerant qubit.

This chapter is organised as follows: we will first overview some of the electrical transport signatures of TSS in 3D TIs. We will then introduce

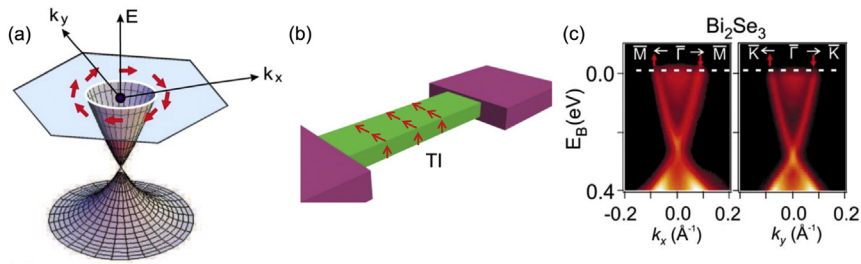


Figure 7.1 (a) Schematic of spin-helical surface energy dispersion of a typical 3D TI. (b) Schematic of the spin texture in a 3D TI surface state. Spin is always tangential to the surface and perpendicular to the momentum. (c) Band structure of Bi_2Se_3 measured by ARPES. *Credit: Adapted from Refs. [1,17,42].*

the TINR as an interesting candidate for the observation of topological transport. Finally, we will present recent progress in combining 3D TINR and superconductors in pursuit of topological superconductivity and Majorana modes.

2. Overview of electrical transport in TI

The band dispersion of a 3D TI has been measured by an angle resolved photoemission spectroscopy (ARPES) on Bi_2Se_3 samples [1,17]. Fig. 7.1c shows the ARPES data, where linear spin-polarised Dirac dispersion of the TSS is demonstrated. In a perfect TI bulk is insulating and electrical current is carried by surface states. In reality, bulk carriers contribute to the conduction due to unintentional impurity doping and thermal excitations. There are several methods to separate the contributions of TSS and bulk carriers. Doping a conventional Bi_2Te_3, which is highly electron doped, with Sb shifts the chemical potential into the bandgap and compensates for the unintended doping. It was found that ternary and quaternary tetradymite TI materials, such as $(Bi,Sb)_2Te_3$ and $BiSbTeSe_2$, exhibit a very low bulk contribution [18–21]. For instance, the sheet resistance R_{sh} in $BiSbTeSe_2$ does not change with the thickness (see Fig. 7.2a) [22], suggesting that R_{sh} is completely dominated by the TSS contribution at $T = 2$ K and the bulk is insulating. We note that in conventional metals sheet resistance is inversely proportional to the thickness.

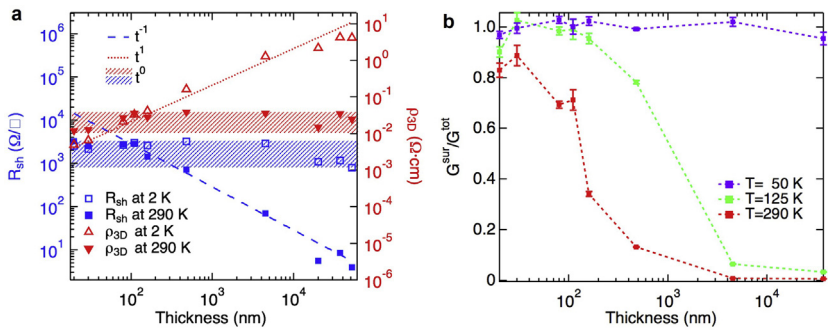


Figure 7.2 (a) Sheet resistance R_{sh} and 3D resistivity ρ_{3D} in $BiSbTeSe_2$ versus thickness at $T = 2$ K and $T = 295$ K. At $T = 2$ K, R_{sh} exhibits a 2D behaviour and does not change with the thickness. The dashed blue (dark grey in print version) line and the horizontal red (grey in print version) band represent the expected behaviours for R_{sh} and ρ_{3D} of a 3D bulk conductor, respectively. (b) TSS contribution to the total conductance (G^{sur}/G^{tot}) versus thickness at three temperatures. *Credit: Ref. [22].*

2.1 Temperature dependence of the electrical conductivity

For TIs with the Fermi energy residing within the gap, we expect that contribution of bulk conduction will be suppressed at low temperatures. Therefore, the temperature dependence of the electrical conductivity can provide insight into the nature of electrical transport and, consequently, reveal the contribution of the TSS. The surface contribution to the total conductance can be extracted from the temperature-dependent electrical conductivity using a simple model developed in Ref. [23]. In this model, the total sheet conductance is a sum of a thermally activated bulk conductance and a metallic surface conductance $G^{tot}(T) = G^{bulk}(T) + G^{sur}$, where $G^{bulk} = t\sigma_{0b}\exp\left(-\frac{E_g}{k_B T}\right)$, t is the thickness of the TI, σ_{0b} is the high temperature conductivity of the bulk, E_g is the bulk gap, k_B is the Boltzmann constant and T is the temperature. The metallic surface contribution can be modelled as $G^{sur} = 1/(R_{s0} + AT)$, where R_{s0} accounts for the residual resistance at $T = 0$ K and A introduces electron–phonon coupling [23]. The model predicts that for temperatures $k_B T < E_g$, G^{tot} is dominated by G^{sur} and increases with decreasing temperature. This model was employed by Y. Xu et al. [22] to estimate the surface to bulk contribution in BiSbTeSe$_2$ films. Fig. 7.2b plots the extracted surface contribution to the total conductance versus thickness. The authors observed that at low temperatures or in thin flakes conductance is mostly dominated by the surface contribution.

2.2 3D TI in a perpendicular magnetic field

2.2.1 Low magnetic field range: weak anti-localisation

In a strong spin orbit-coupled system, such as a TI, magnetoconductance is expected to exhibit weak anti-localisation (WAL). WAL originates from destructive quantum interference of self-intersecting scattering paths due to spin–momentum coupling. When a small magnetic field is applied, time reversal symmetry is broken and WAL is suppressed. Therefore, at low magnetic fields (B) conductance decreases with an increase of magnetic field, and negative magnetoconductance $\Delta\sigma(B) = \sigma(B) - \sigma(0)$ with a characteristic cusp at $B = 0$ T is observed. The WAL can be modelled using the Hikami–Larkin–Nagaoka formula [24]:

$$\Delta\sigma(B) = -\alpha\frac{e^2}{2\pi^2\hbar}\left[\ln\left(\frac{\hbar}{4el_\varphi^2 B}\right) - \Psi\left(\frac{\hbar}{4el_\varphi^2 B} + \frac{1}{2}\right)\right], \tag{7.1}$$

where Ψ is the digamma function, l_φ is the phase coherence length and α is a numeric prefactor. For a linear Dirac dispersion of the TSS, which has a Berry phase of π, $\alpha = -0.5$. In a TI film, however, there are two parallel surfaces (top and bottom), each contributing -0.5 to α. Therefore, for transport which involves both TSS $\alpha = -1$. The value of α was recently extracted from magnetoconductance measurements in gate-tunable BiSbTeSe$_2$ flakes by Y. Xu et al. [22,25]. Fig. 7.3 depicts the gate dependence of the longitudinal resistance R_{xx} in BiSbTeSe$_2$ flakes at zero magnetic field, where a peak in the resistance is observed close to the back-gate voltage $V_{bg} = -60$ V. This gate voltage highlights the position of the chemical potential where the total charge in the system is neutral and, hence, is called a charge neutrality point (CNP). Fig. 7.3b and c depict magnetoconductance and the extracted α and l_φ in the TI flakes. When both the top and the bottom surfaces are electron doped, which is the case for $V_{bg} > -30$ V, $\alpha \approx -1.2$. Near the charge CNP or for the case when the top surface is electron doped and the bottom surface is hole doped ($V_{bg} < -30$ V), α is further reduced. The reason behind this deviation is not clear to us.

2.2.2 High magnetic field range: half-integer QH effect

When a TI is exposed to a strong perpendicular magnetic field, Landau levels are formed. In conventional semiconductors, because of the parabolic nature of the band dispersion, the Landau quantisation produces equidistant energy levels and gives rise to the integer QH effect, manifested by quantised Hall conductivity and zero longitudinal conductivity [26–30]. For relativistic carriers with linear Dirac dispersion, Landau levels are not equidistant and, more importantly, the Hall conductivity is quantised at half integers

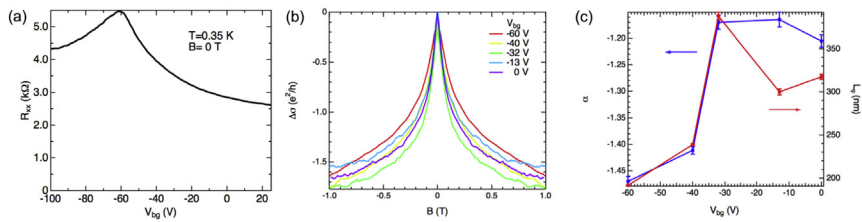


Figure 7.3 Longitudinal resistance R_{xx} as a function of the back-gate voltage V_{bg} at $T = 0.35$ K and $B = 0$ T. (b) Magnetoconductance $\Delta\sigma$ versus magnetic field for various back-gate voltages, marking the weak anti-localisation effect. (c) α and l_φ versus V_{bg}. All measurements are done in BiSbTeSe$_2$ flakes obtained by the scotch tape exfoliation from a bulk crystal. *Credit: Refs. [22,25].*

[15,22,31]. A half-integer QH effect was first observed in gate-tunable $BiSbTeSe_2$ flakes by Y. Xu et al. [22] and was later reproduced by others [32,33]. Fig. 7.4a presents a longitudinal resistance $R_{xx} = V_{2,3}/I_{1,4}$, where subscripts correspond to different terminals as shown in the inset, and a Hall resistance $R_{xy} = V_{5,3}/I_{1,4}$ as functions of the back-gate voltage V_{bg} at $T = 0.35$ K in the same $BiSbTeSe_2$ flake as the one shown in Fig. 7.3. Fig. 7.4b depicts the calculated longitudinal (σ_{xx}) and Hall conductivity (σ_{xy}) as functions of V_{bg}. At $V_{bg} = 0$ V, the sample is n-doped (both top and bottom surfaces) and due to the large thickness of the flake, the back-gate control over the top surface is negligible. Therefore, the top surface most likely remains n-doped throughout the V_{bg} range used in this study. When both the top and the bottom surfaces are electron doped ($V_{bg} > -60$ V), well-developed quantised plateaus in σ_{xy} accompanied by vanishing σ_{xx} are observed. The quantised values of σ_{xy} are integer values of e^2/h, where e is the electron charge and h is the Planck constant. This indicates that each surface of the TI flake contributes half-integer quantised conductance to the measured total conductivity. Moreover, the lack of gate control on the top surface ensures that the contribution from the top surface is always $\frac{1}{2}e^2/h$. The observed plateaus in this experiment can be understood as parallel contributions from the top surface ($\nu_t e^2/h$) and the bottom surface ($\nu_b e^2/h$). In other words, $\sigma_{xy}^{tot} = \sigma_{xy}^{top} + \sigma_{xy}^{bottom} = (\nu_t + \nu_b)e^2/h = (N_t + N_b + 1)e^2/h$, where $\nu_{t(b)} = N_{t(b)} + \frac{1}{2}$ is the Landau level filling factor and $N_{t(b)}$ is an integer number. In electron-doped bottom-gated TIs, ν_t is always ½

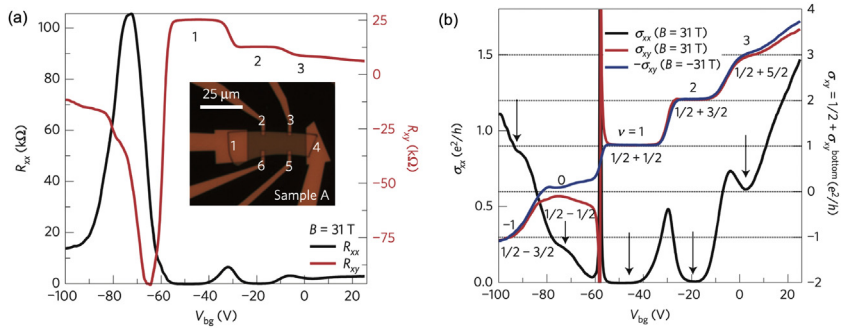


Figure 7.4 (a) Longitudinal (R_{xx}) and Hall (R_{xy}) resistance as functions of the back-gate voltage (V_{bg}) at $B = 31$ T and $T = 0.35$ K. Inset shows the optical image of the device where different electrodes are marked. (b) Calculated 2D longitudinal σ_{xx} and Hall σ_{xy} conductivities from (a). *Credit: Ref. [22].*

and, hence, $\sigma_{xy}^{tot} = (N_b + 1)e^2/h$ is always quantised. The top and bottom Landau level filling factors are marked for each plateau in Fig. 7.4b. In order to control v_t, one needs to fabricate dual gated structures where both the top and bottom surfaces are controlled independently. This was accomplished by Yang Xu et al. in Ref. [34].

3. Electrical transport in TI nanoribbons

In this section, we introduce quantum interference effects in TSSs. Unlike bulk samples, in mesoscopic systems, e.g., TINR, the low temperature phase coherence length is comparable to the length scale of the sample and phase coherence transport can be observed. TINRs resemble a metallic hollow cylinder and electron trajectories can self-interfere when electrons move around the cross-sectional circumference. When a magnetic field along the length of a TINR is applied, quantum interference is affected by the magnetic flux [35]. As a result, the interference patterns will give rise to quantum oscillations of magnetoresistance as a function of the magnetic flux, with a characteristic period of the flux quantum, $\Phi_0 = h/e$ [36–39]. Such oscillations are known as Aharonov–Bohm (AB) effects and have been observed in prototypical TINRs including Bi_2Te_3 [40] and Bi_2Se_3 [41]. Note that a nanoribbon in this notation refers to a nanowire with a rectangular cross-section.

In the TINRs, the confinement of the surface states along the circumference will result in discretisation of modes around the circumference (we define this direction to be along the y axis as shown in Fig. 7.5a). Unlike conventional metallic hollow cylinders, the spin-helical nature of TSS in the TINRs ensures that the spin of carriers is always tangential to the surface and perpendicular to the momentum. This spin-momentum locking enforces a particle moving around the perimeter of the TINR to pick up an additional AB phase (Berry phase of π). In other words, we must use anti-periodic boundary conditions for TINRs when we consider the quantisation of the transversal modes. Therefore, the transversal modes in TINRs are described as follows [36,37]:

$$k_{y,n} = \frac{2\pi}{C}\left(n + \frac{1}{2}\right), \tag{7.2}$$

where $k_{y,n}$ is the transverse momentum along the circumference corresponding to n, C is the circumference of the TINR and n is an integer. For

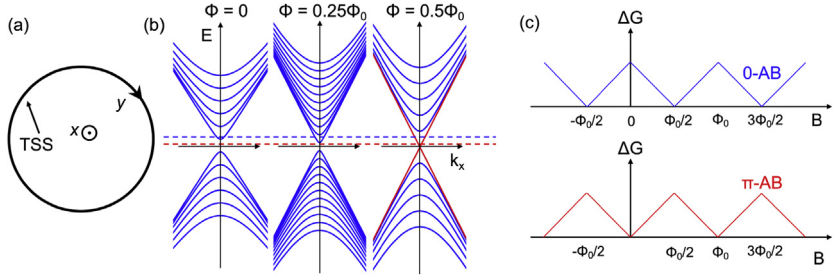


Figure 7.5 (a) Cross-section of a TI nanoribbon. We define y direction along the circumference and longitudinal x direction along the nanoribbon length. (b) Energy dispersion of the TI nanoribbon surface modes (E versus longitudinal momentum k_x) for three magnetic fluxes. (c) The predicted magnetoconductance ΔG versus B exhibiting Aharonov–Bohm (AB) oscillations with phase of 0 (blue [dark grey in print version] curve) and π (red [grey in print version] curve). 0-AB (π-AB) oscillations correspond to the Fermi energy E_F, as shown by the dashed blue (dark grey in print version) (red [grey in print version]) horizontal line in (B). *Credit: Reproduced from Ref. [42].*

any given k_y mode, the energy dispersion is also described using the linear Dirac dispersion as follows:

$$E_n(k_x) = \pm \hbar v_F \sqrt{k_x^2 + k_{y,n}^2}, \tag{7.3}$$

where k_x is the longitudinal momentum along the length of the TINR and v_F is the Fermi velocity of surface modes. When an axial magnetic field is applied, the electron wavefunction will pick up an additional AB phase of $2\pi\Phi/\Phi_0$ and quantised transverse modes will take the form of [36,37,42]

$$k_{y,n} = \frac{2\pi}{C}\left(n + \frac{1}{2} - \frac{\Phi}{\Phi_0}\right), \tag{7.4}$$

where $\Phi = B/S$ is the magnetic flux and S is the cross-sectional area of the TINR. Fig. 7.5b plots the energy dispersion of various k_y modes as a function of k_x for three different magnetic fluxes. Unique to the TINR, the energy dispersion is gapped at zero magnetic flux as well as at even multiples of $\Phi_0/2$. This is a result of the quantum confinement effect along the circumference. As the flux deviates from even multiples of $\Phi_0/2$, the energy gap decreases and closes at odd multiples of $\Phi_0/2$, restoring the linear Dirac dispersion. Therefore, an axial magnetic flux drives a periodic topological transition of the surface modes, from a system with all topologically trivial doubly degenerate transversal modes at even multiples of $\Phi_0/2$ to a system

with a single nondegenerate mode for $k_y = 0$ and doubly degenerate modes for $k_y > 0$ at odd multiples of $\Phi_0/2$. The one-dimensional mode at $k_y = 0$ is spin helical with linear energy–momentum dispersion and is topologically protected against backscattering by nonmagnetic impurities. The energy gap between the two adjacent modes at $k_x = 0$ is also given by $\Delta E = \hbar v_F \Delta k_y$, where $\Delta k_y = k_{y,n+1} - k_{y,n} = 2\pi/C$ [42].

The unique magnetic flux dependence of the energy dispersion has an interesting consequence in the gate tunable magnetoconductance of the TINR. Let us take a look at the case where the chemical potential (controlled by the gate voltage) is tuned to the red (grey in print version) dashed line in Fig. 7.5b. At $T = 0$ K, the magnetoconductance of the TINR will be larger at $\Phi = (n + 1/2)\Phi_0$ compared to $\Phi = (2n)\Phi_0/2$, because at $\Phi = (n + 1/2)\Phi_0$, the topological mode at $k_y = 0$ will contribute to the conductance, whereas at $\Phi = (2n)\Phi_0/2$, the chemical potential is inside the gap and there is no mode to contribute to the conductance. Therefore, the magnetoconductance will have maxima at $(n+1/2)\Phi_0$ and minima at $(n)\Phi_0$. A schematic representation of this type of AB oscillation (π-AB) is shown in Fig. 7.5c. Unlike the conventional AB effects, where magnetoconductance maxima are at zero and even multiples of $\Phi_0/2$ (Fig. 7.5c blue [grey in print version] curve), for the TINR at this specific gate voltage, the 1D topological mode gives rise to a π phase shift in the AB oscillations. This phenomenon is called the π-AB effect and is a result of the additional π phase shift introduced by the spin rotation of the 1D topological mode around the perimeter of the TINR. If we change the position of the chemical potential, for instance, to the dashed blue (dark grey in print version) line in Fig. 7.5b, at $\Phi = (2n)\Phi_0/2$, there are two modes contributing to the transport, whereas at $\Phi = (2n + 1)\Phi_0/2$, there is only one mode. We note here that all the modes in Fig. 7.5b are doubly degenerate except for the topological mode with the linear energy dispersion. Therefore, the magnetoconductance for this chemical potential will have maxima at even multiples of $\Phi_0/2$ and minima at odd multiples of $\Phi_0/2$. This type of magnetoconductance oscillations resembles the conventional AB effect (0-AB) without any phase shift, as shown with the blue (dark grey in print version) curve in Fig. 7.5c. Therefore, in the TINR, we can alternate the phase of the AB oscillations by changing the chemical potential of the TINRs. This 0 to π transition is periodic with respect to k_F with a period of $\Delta k_y = 2\pi/C$, which determines the energy separation (ΔE) of the modes at $k_x = 0$. We note that, experimentally, the temperature broadening and

Fermi energy fluctuations along the length of the nanoribbon should be smaller than the energy gap ΔE between the adjacent modes in Fig. 7.5b. The π-AB oscillations of the magnetoconductance were experimentally measured in Bi_2Se_3 nanoribbons in Ref. [43]. Jauregui et al. [42] reproduced the π-AB oscillations in Bi_2Te_3 nanoribbons and further demonstrated a clear periodic 0 to π alternation with respect to k_F, driven by the gate voltage. Fig. 7.6a and b summarises the experimental results corresponding to Ref. [42]. The gate-tunable AB oscillations are analysed by labelling each peak (dip) in the magnetoconductance at $\Phi = \Phi_0/2$ ($\Phi = 0$) with an integer (N). This procedure is done for back-gate voltages above V_0, where

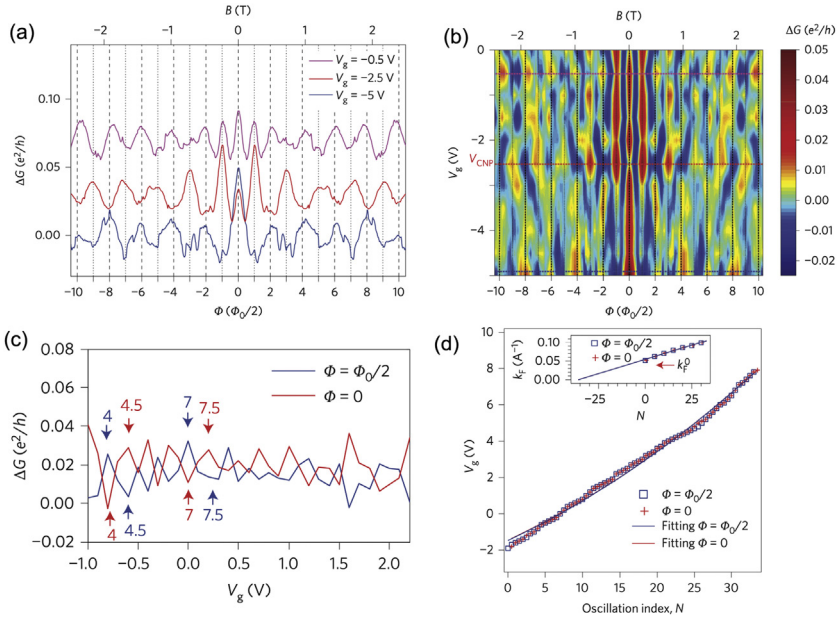


Figure 7.6 (a) Magnetoconductance (ΔG, with a smooth background subtracted) as a function of the magnetic flux (Φ) at three back-gate voltage V_g's for Bi_2Te_3 nanoribbons. The nanoribbon exhibits a 0-AB effect at $V_g = -5$ V and a π-AB effect for $V_g = -2.5$ V. (b) Colour plot of ΔG as functions of Φ and V_g, exhibiting 0 to π transition of the AB oscillations with V_g, which controls the chemical potential. The transition from 0 to π is periodic with respect to Δk. (c) ΔG versus V_g for $\Phi = \Phi_0/2$ and $\Phi = 0$. The corresponding N and $N + 1/2$ indexes are marked in the figure. (d) Each back-gate voltage corresponding to each N or $N + 1/2$ from (c) is plotted as a function of the N. Inset: k_F versus N exhibiting a linear behavior. This linear behaviour reveals that the oscillations (N) are periodic with respect to k_F with a periodicity given by Δk_y. *Credit: Ref. [42].*

V_0 is the smallest gate voltage where a transition from 0- to π-AB oscillations is observed. Similarly, each dip (peak) in the magnetoconductance at $\Phi = \Phi_0/2$ ($\Phi = 0$) is labelled with $N + 1/2$. Fig. 7.6c plots the back-gate voltage (V_g) as a function of the extracted oscillation index (N). For a back-gated TINR, the sheet carrier density n_s is given by

$$n_s = C_{ox}\left(V_g - V_{CNP}\right)/e, \tag{7.5}$$

where C_{ox} is the parallel plate capacitance of the oxide layer per unit area. Moreover, for a 2D band, the carrier density is also given by the following:

$$n_s = \left(k_F^2 - k_{F,0}^2\right)/4\pi, \tag{7.6}$$

where k_F is the Fermi wavevector of the surface modes and $k_{F,0}$ is the Fermi wavevector in the bulk valence band. We note that in our description, the TSS of the TINR is treated as a 2D plane with anti-periodic boundary conditions, and hence a 2D carrier density can properly describe the system. Assuming $k_F = k_{F,0} + N\Delta k_y$ with $\Delta k_y = 2\pi/C$, $k_{F,0}$ and C_{ox} can be extracted by fitting Eqs. (7.5) and (7.6) to V_g versus N data plotted in Fig. 7.6d. From this fit, $k_{F,0} \sim 0.05\ \text{Å}^{-1}$ and $C_{ox} \sim 100\ \text{nF/cm}^2$ (for $SrTiO_3$ substrate) are obtained. To see the periodic nature of the oscillation index N with respect to k_F, the calculated $k_F = \sqrt{4\pi C_{ox}/e(V_g - V_0) + k_{F,0}^2}$ is plotted as a function of N in Fig. 7.6d. For $k_F < k_{F,0}$ ($V_g < V_0$), the magnetoconductance is dominated by the bulk valence band conduction, carriers are not localised to the TSS, and no AB oscillations are observed.

As we discussed earlier, at $T = 0$ K and for the dashed red (grey in print version) line in Fig. 7.5b, the conductance oscillations are due to the ballistic spin-helical TSS. Considering the ballistic nature of the $k_y = 0$ mode, one would expect the amplitude of the AB oscillations to be e^2/h, i.e., ideally the amplitude of ΔG in Fig. 7.6a should be quantised (e^2/h). However, all the experimentally measured AB oscillations in the TINRs [40–43] reported $\Delta G \sim 0.1e^2/h$, which is significantly smaller than the theoretically expected value. The exact mechanism giving rise to the reduction of the AB amplitude is not clear, even though disorder and temperature broadening as well as invasive contacts, whose reflection and transmission coefficients can influence the four-terminal magnetoconductance measurements, have been considered as the underlying reasons.

3.1 Ballistic transport from the temperature dependence of AB oscillations

The other interesting property of AB oscillations is their temperature dependence. The AB effect will have a different characteristic temperature dependence depending on whether transport is ballistic or diffusive [44]. Previous studies have shown that the amplitude of the AB oscillations in the diffusive regime exhibits a $T^{-1/2}$ dependence [40,41], whereas temperature dependence is exponential for ballistic nanoribbons [45]. In the TINRs, the amplitude of AB oscillations of the magnetoconductance is determined by the 1D topological mode with $k_y = 0$. Since this mode is topologically protected against backscattering, we expect to observe an exponential T-dependence of the AB amplitude. Fig. 7.7a depicts a colour

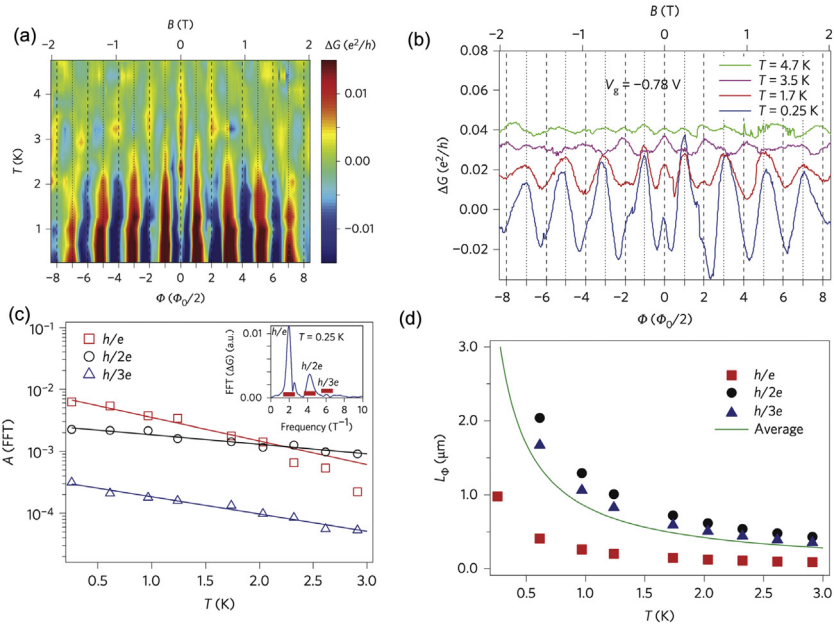


Figure 7.7 (a) Magnetoconductance (ΔG) versus temperature and magnetic flux in Bi_2Te_3 nanoribbons. (b) A few cuts of ΔG versus Φ of (a) for different temperatures. (c) temperature dependence of FFT amplitude A for the first (h/e), second ($h/2e$) and third ($h/3e$) harmonics. Inset is the FFT of the data of (b) at T = 0.25 K. The peak amplitude (A) is defined as the integrated area over the corresponding interval represented by the red (light grey in print version) horizontal line. (d) Temperature dependence of the phase coherence length (L_φ) obtained from $L_\varphi = jC/2Tb_j$, where j is the j^{th} harmonic in (c) and b_j is obtained from fitting A to e^{-b_jT}. l_φ exhibits T^{-1} dependence. An average between the three data sets is shown as a solid green (grey in print version) line. *Credit: Ref. [42].*

map plot of magnetoconductance in Bi_2Te_3 nanoribbon as functions of an axial magnetic field and temperature [42]. The amplitude of AB oscillations, corresponding to the first, second and third harmonics obtained from a fast Fourier transfer (FFT) of the data in (a), are also plotted in Fig. 7.7b. The amplitude of all three harmonics shows an exponential decay with the increasing temperature, revealing the ballistic nature of transport along the TINR length. Theoretically, the amplitude of FFT (A_j), where j is the j^{th} harmonic, is proportional to $e^{-b_j T}$, where the decay factor (b_j) is found to be $b_j = jC/(2Tl_\varphi)$ [45,46] and l_φ is the phase coherence length. Experimentally, it is found that b_j is a nonlinear function of j. Although the underlying origin of this deviation from a linear behaviour is not well established, similar deviations were attributed to thermal averaging. An extracted $l_\varphi = jC/(2Tb_j)$ displays a T^{-1} dependence [42]. This -1 exponent differs from 0.4 to 0.5 observed in Refs. [41,47–49] and, most likely, indicates weaker coupling of the TINRs to the environment [44].

4. Superconducting transport in TI nanoribbons

The TSSs are topologically nontrivial at zero magnetic field and, thus, TI/superconductor (TI/S) heterostructures are ideal to form a topological superconducting phase. Many theoretical works have studied the existence of Majorana modes in 3D TIs [16,38,39,50–55]. Additionally, TI/S interfaces, including S/TI/S Josephson junctions, have been experimentally demonstrated by many groups [56–64]. Superconducting transport in TIs is usually studied in a Josephson geometry, where two closely spaced superconductors are placed on top of a TI flake to form a Josephson junction. Fu and Kane [16] first pointed out that the induced superconductivity in the TSSs can give rise to zero-energy Majorana excitations. These excitations are named Majorana modes because their mathematical formulation is identical to Majorana fermions. The energy dispersion of the surface states in TIs is

$$E(k_F) = \pm \hbar v_F |k_F|. \tag{7.7}$$

This band structure is ideal to form a topological superconducting phase (which gives rise to the Majorana modes), because for any chemical potential inside the bulk gap, the spin-helical 'spinless' regime is automatically accessed, and the TI/S interface is topologically nontrivial at zero magnetic field [16,39]. In contrast, in semiconductor nanowires the spin-helical states emerge at large magnetic fields [65–70].

Fu and Kane [16] derived the following energy-phase relation for the Andreev bound states in a zero length Josephson junction (JJ):

$$E(\phi) = \pm \Delta_0 \cos\frac{\phi}{2}, \tag{7.8}$$

where Δ_0 is the superconducting gap at zero temperature, and ϕ is the phase difference across the JJ. Olund et al. [54] studied the effect of the channel length and of the out-of-gap states on the supercurrent; they derived

$$E(\phi) = \pm \Delta_0 \cos\left(\frac{E(\phi)L}{\hbar v_F} \pm \frac{\phi}{2}\right), \tag{7.9}$$

where L is the channel length, for a JJ with an arbitrary channel length. Expression (7.9) becomes (7.8) in a short-channel limit $L \to 0$. Ghaemi et al. [71] studied the effect of impurities; they found that the presence of impurities will act as an effective increase of the channel length, replacing L with L_{eff} in (7.9). Current phase relation (CPR) can be calculated by taking derivative of the energy with respect to the phase:

$$I(k_y, \phi) = -\frac{e}{h}\sum_{E_n \geq 0}\frac{\partial E_n}{\partial \phi}\tanh\frac{E_n}{2k_B T}. \tag{7.10}$$

Eq. (7.10) is derived assuming that the temporal changes of ϕ are slow. Hence, the Fermi distribution function (the tanh term) is included to account for this slow varying phase. Fig. 7.8a and b plot the energy spectrum of Andreev bound states and the corresponding CPR for topological and trivial junctions (transmission coefficient $D = 0.2$ is assumed for the trivial junction). Backscattering from nonmagnetic impurities is prohibited for TSS, which leads to the energy levels crossing at $\phi = \pi$, and, consequently, a 4π-periodic CPR [16]. In conventional JJs, the CPR is 2π-periodic [72]. The two energy branches have different fermion parity, and in order to measure 4π periodicity, the measurements have to be performed faster than the quasi-particle poisoning time [73]. In the presence of quasi-particle poisoning, the CPR of a topological JJ is expected to be skewed (due to the ballistic transport) but 2π periodic. In a short trivial junction, the energy-phase relation is given by [72]:

$$E(\phi) = \pm \Delta_0 \sqrt{1 - D\sin^2\frac{\phi}{2}}, \tag{7.11}$$

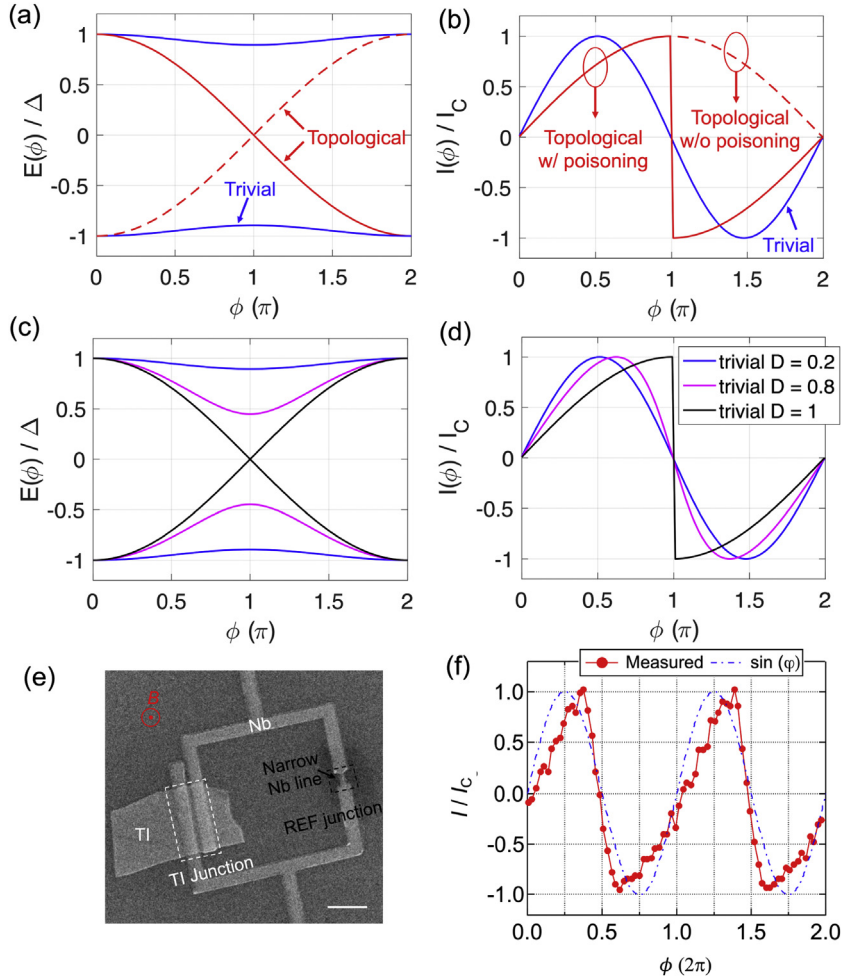


Figure 7.8 (a) Energy spectrum E of the Andreev bound states (ABS) versus the phase difference ϕ between the two superconductors in topological (red curve) and trivial (blue curve) Josephson junctions. (b) The current I, normalised by I_c (the critical current) versus ϕ for topological (red [grey in print version] solid and dashed curves) and trivial (blue [dark grey in print version] curve) junctions. Because of the quasi-particle poisoning, the topological junction may exhibit a highly skewed two π periodic $I - \phi$ [73]. (c,d) The energy spectrum E (c) and the normalised current I/I_c (d) in a trivial junction for various junction transparencies D. (e) A scanning electron microscope (SEM) image of an asymmetric superconducting quantum interference device (SQUID) used to measure the current phase relation (CPR). Scale bar is 1 μm. (f) The current phase relation (symbols): the normalised current (I/I_C) of the TI-based JJ versus the phase ϕ measured at temperature $T = 20$ mK. Dashed blue (grey in print version) curve is a reference sinusoidal function. *Credit: (e) and (f) are taken from Ref. [64].*

where D is the junction transparency. In Fig. 7.8c and d $E-\phi$ and $I-\phi$ are plotted for three trivial junctions with different transparencies D. For a perfectly ballistic junction, $E-\phi$ and $I-\phi$ curves resemble those of the topological junction with quasi-particle poisoning (slow varying phase). Therefore, observation of highly skewed CPR, even though a necessary step, is not sufficient evidence for the existence of topological superconductivity. In fact, to reveal the 4π-periodic nature of the CPR high frequency measurements, with frequencies above the quasi-particle poisoning rate, are required.

The CPR of the TI JJ can be measured in an asymmetric quantum interference device (SQUID), shown in Fig. 7.8e. The asymmetric SQUID consists of a reference (REF) junction with a known sinusoidal CPR in parallel with a TI-based junction. If the critical current of the REF junction is much larger than that of the TI junction, the magnetic flux modulation of the SQUID critical current will directly probe the CPR of the TI JJ. We have measured the CPR of $BiSbTeSe_2$ flakes using this technique; the results are plotted in Fig. 7.8f. We observe that the CPR is highly skewed, which is consistent with our understanding of ballistic TSS with a slow varying ϕ (controlled by the magnetic flux threading the SQUID) [64]. In contrast, in a TINR the $k_y = 0$ 1D surface mode only appears at $\Phi = \left(n + \frac{1}{2}\right)\Phi_0$ and, hence, CPR in zero magnetic field is not expected to be 4π periodic.

4.1 Temperature dependence of the critical current in TI nanoribbons

Temperature dependence of the critical current I_c, which is the maximum value of $I(\phi)$ in Eq. (7.10) at any given temperature, is a versatile probe to study the induced superconductivity. $I_c(T)$ dependence has been investigated in S/TI/S JJs based on $BiSbTeSe_2$ nanoribbons with Nb as superconducting leads [61,63]. The $I_c(T)$ reveals an anomalous enhancement of the critical current at temperatures below $0.2T_c$, where T_c is the critical temperature of the JJ. We will refer to this transition temperature as T^*. Fig. 7.9a displays I_c versus T in a TINR at three different back-gate voltages. Starting from T_c, I_c increases smoothly with reducing temperature down to T^*, below which I_c exponentially increases as T approaches zero. In general, any model which considers two characteristic energy gaps in the system

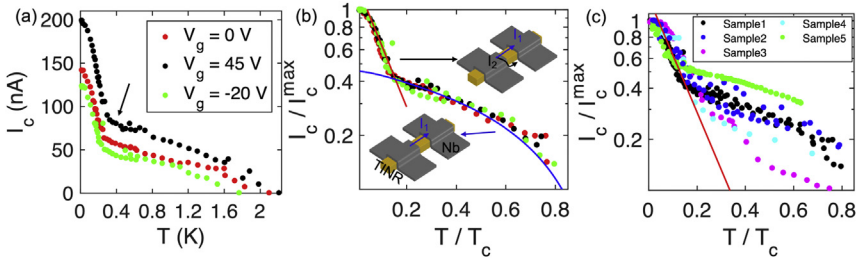


Figure 7.9 (a) Temperature dependence of I_c for different back-gate voltages (V_g's) in $BiSbTeSe_2$ nanoribbons. (b) Normalised I_c/I_c^{max} versus normalised T/T_c in a log-linear scale. The solid blue (dark grey in print version) curve is a fit to the short junction model and the solid red (grey in print version) curve is an exponential fit. Inset: cartoon of the TI nanoribbon junction depicting I_1 and I_2. (c) I_c/I_c^{max} versus T/T_c in a log-linear scale for five different $BiSbTeSe_2$ nanoribbons. Solid line is an exponential fit. *Credit: Ref. [63].*

can reproduce our observed temperature dependence. In the original paper [68], T^* was associated with superconductivity being induced into the bottom surface of the nanoribbon (the side which is not in contact with Nb), consistent with the observation that for $T < T^*$ Little–Park oscillations (discussed below) are observed [66]. The observed I_c versus T is radically different from that of the *short* JJs, where I_c saturates at low temperatures without exhibiting any exponential behaviour [72,74]. In contrast, for *long* JJs, I_c increases exponentially with decreasing temperature [72,75–78]. We note that short refers to a junction with $L \ll \xi$, where ξ is the superconducting phase coherence length, whilst $L \geq \xi$ is considered a long junction. The model assumes that the increase in I_c versus decreasing T for $T^* < T \leq T_c$ followed by an exponential enhancement of I_c for $T < T^*$ as observed in Fig. 7.9 accounts for a transition from a short to long junction limit at T^*, where superconductivity extends to the bottom TSS. The total current $I = I_1 + I_2$ is carried by the top TSS (I_1 in the inset of Fig. 7.9b) and by the bottom TSS (I_2). The supercurrent corresponding to I_2 involves modes with large k_y that extend around the circumference and go from the top left superconductor to the top right one. The circumference of the nanoribbon $C \geq \xi$ and these modes are in the long junction regime. I_1 is governed by modes with small k_y, which are in a ballistic regime, and is well described by Eq. (7.10). I_2 has an exponential T-dependence [72,78]:

$$I_2 \propto \exp\left(-\frac{k_B T}{\delta}\right), \tag{7.12}$$

where δ is the characteristic energy gap of the modes in the long junction and is given by

$$\delta = \frac{\hbar v_F}{2\pi d}, \tag{7.13}$$

where d is the effective length. I_2 becomes appreciable for small $T < \delta / k_B$. From the fit to experimental data, we extract $\delta \sim 0.08\Delta_0$, corresponding to the effective length of $d = 1.2$ μm, twice larger than $C \sim 0.7$ μm. This anomalous temperature enhancement has been observed in multiple $BiSbTeSe_2$ nanoribbon JJs (see Fig. 7.9c). Yet another theoretical explanation to the observed temperature dependence was put forward by Talantsev [79], who argues that the sharp transition in $I_c - T$ curves might be related to the linear Dirac energy dispersion of the TINR rather than the winding of the TSS. Future theoretical and experimental investigations might be needed to clarify the origin of anomalous critical current enhancement at low temperatures.

4.2 Aharonov–Bohm effect in TI nanoribbon Josephson junctions

In the presence of an axial magnetic field, the critical current in the junction is expected to exhibit AB oscillations, similar to the conductance oscillations in the normal state discussed earlier. Because of the unique quantisation of the transverse momentum k_y (Eq. 7.4) and the band dispersion in the TINR, a transition between 0- and π-AB effect is also predicted to occur in the superconducting state [38,39,51,80,81]. In the superconducting state, oscillations are due to the interference of Cooper pairs with charge $2e$ and flux is quantised in units of $\Phi_{s0} = h/2e$, half the flux in the normal state (this phenomenon is also known as the Little–Park effect). In the presence of ballistic Andreev bound states, including Majorana bound states, oscillations with h/e period are predicted [51,82].

Fig. 7.10 shows a nanoribbon device with four normal (N, Ti/Au) and four superconducting (S, Nb) contacts [83]. For N contacts, the AB oscillations with a period $\Delta B_n = 0.48$ T are observed. The phase of the AB oscillations changes from π for $V_g = 14.5$ V to 0 for $V_g = 16$ V and back to π for $V_g = 17.5$ V, similar to the AB oscillations discussed earlier (Fig. 7.7). This gate-induced periodic $\pi - 0 - \pi$ transition unambiguously confirms TSS transport in this TI nanoribbon. Now we compare AB oscillations in the normal regime with oscillations of the critical current in a JJ formed between two S contacts in the same nanoribbon. In Fig. 7.10c, differential

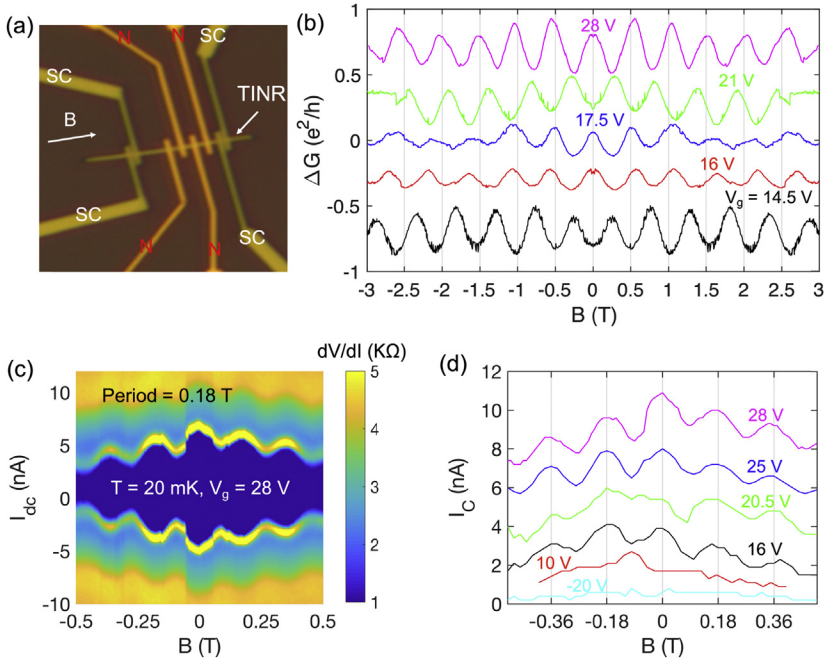


Figure 7.10 (a) An optical microscope image of the TI nanoribbon with Ti/Au normal (N) contacts and Nb superconducting (SC) contacts. (b) Magnetoconductance ΔG versus the axial magnetic field measured using the normal contacts for various back-gate voltages (V_g's) at $T = 20$ mK. Curves are shifted vertically for clarity. (c) Colour map of the two-terminal differential resistance dV/dI as functions of I_{dc} and B measured at $V_g = 28$ V and $T = 20$ mK. (d) The critical current I_c versus B measured for different V_g's at $T = 20$ mK. Quantum interference patterns in I_c with period of $\Delta B_s \sim 0.18$ T are observed. *Credit: Ref [83].*

resistance is plotted as a function of dc current I_{dc} and the magnetic field B at $T = 20$ mK and $V_g = 28$ V for a JJ. The blue (light grey in print version) region corresponds to the $R = 0$ state, and its boundary (I_c) oscillates with magnetic field. In Fig. 7.10d, oscillations of I_c versus the axial magnetic field are plotted for several gate voltages. Three features stand out in the superconducting transport: first, the interference patterns are AB-like (or SQUID-like) and, hence, are different from the Fraunhofer patterns measured in the perpendicular magnetic fields. Second, the period of the oscillations is $\Delta B_s \sim 0.18$ T, which is > 2 times smaller than the period of 0.48 T in the normal state. The fact that $\Delta B_s \lesssim \Delta B_n/2$ may suggest that oscillations are due to Cooper pair interference. There are however several other possible factors that can contribute to the period reduction in the S state,

including magnetic flux focusing and supercurrent extending to the outer surface of Nb contacts, thus effectively increasing the enclosed area. Third, the phase of the supercurrent oscillations remains zero regardless of the applied gate voltage. It is not clear why the π-AB oscillations do not show up in the superconducting state. However, we speculate that superconducting condensate is very effective in screening the gate field and, thus, the gate is not as efficient in tuning the Fermi level in superconducting state. Overall, the AB-like quantum interference patterns observed in the TINR provide yet another signature of the surface superconductivity. We note here that further experimental investigation, e.g., direct measurement of the 4π-periodic CPR, is necessary to conclude whether topological superconductivity and Majorana modes can indeed exist in the TINR-based Josephson junctions.

5. Summary and outlook

In this chapter we summarised recent progress in electrical and superconducting transport of TINRs. In particular, we analysed signatures of topological transport in temperature and magnetic field dependence of the electrical conductivity and the supercurrent. Furthermore, AB interferences of TSSs were discussed in both normal and superconducting states. These signatures provide strong evidence for the existence of TSS and importance of TSS transport in TINRs. The TSSs have already found applications in spintronics, and TI/S hybrid structures are one of the prime candidates for the observation of the Majorana modes, which are predicted to exhibit non-Abelian statistics. Despite significant efforts over the past decade, Majorana modes have not been conclusively observed. Further progress may require higher quality TIs (fewer bulk states, lower disorder and reduced amount of impurities) and cleaner, more transparent TI/S interfaces. YPC and LPR acknowledge partial support by Quantum Science Center (QSC), a US Department of Energy (DOE) National Quantum Information Science Research Center, during the writing of this chapter.

References

[1] Hasan MZ, Kane CL. *Colloquium*: topological insulators. Rev Mod Phys 2010;82: 3045–67.
[2] Qi X-L, Zhang S-C. Topological insulators and superconductors. Rev Mod Phys 2011;83:1057–110.
[3] Maciejko J, Hughes TL, Zhang S-C. The quantum spin Hall effect. Annu Rev Condens Matter Phys 2011;2:31–53.

[4] Ando Y, Fu L. Topological crystalline insulators and topological superconductors: from concepts to materials. Annu Rev Condens Matter Phys 2015;6:361–81.
[5] Yan B, Felser C. Topological materials: Weyl semimetals. Annu Rev Condens Matter Phys 2017;8:337–54.
[6] Armitage NP, Mele EJ, Vishwanath A. Weyl and Dirac semimetals in three-dimensional solids. Rev Mod Phys 2018;90:015001.
[7] Moore JE. The birth of topological insulators. Nature 2010;464:194–8.
[8] Goldsmid HJ. Introduction to thermoelectricity. 2nd ed. Springer; 2010.
[9] König M, Molenkamp LW, Qi X, Zhang S. Quantum spin Hall insulator state in HgTe quantum wells. Science 2007;318:766–71.
[10] Wolgast S, Kurdak Ç, Sun K, Allen JW, Kim DJ, Fisk Z. Low-temperature surface conduction in the Kondo insulator SmB_6. Phys Rev B 2013;88:180405.
[11] Fan Y, Wang KL. Spintronics based on topological insulators. Spin 2016;6:1640001.
[12] Tian J, Hong S, Miotkowski I, Datta S, Chen YP. Observation of current-induced, long-lived persistent spin polarization in a topological insulator: a rechargeable spin battery. Sci Adv 2017;3:e1602531.
[13] Yu R, Zhang W, Zhang H-J, Zhang S-C, Dai X, Fang Z. Quantized anomalous Hall effect in magnetic topological insulators. Science 2010;329:61–4.
[14] Chang C-Z, Zhang J, Feng X, Shen J, Zhang Z, Guo M, Li K, Ou Y, Wei P, Wang L-L, Ji Z-Q, Feng Y, Ji S, Chen X, Jia J, Dai X, Fang Z, Zhang S-C, He K, Wang Y, Lu L, Ma X-C, Xue Q-K. Experimental observation of the quantum anomalous Hall effect in a magnetic topological insulator. Science 2013;340:167–70.
[15] Qi X-L, Hughes TL, Zhang S-C. Topological field theory of time-reversal invariant insulators. Phys Rev B 2008;78:195424.
[16] Fu L, Kane CL. Superconducting proximity effect and Majorana fermions at the surface of a topological insulator. Phys Rev Lett 2008;100:096407.
[17] Xia Y, Qian D, Hsieh D, Wray L, Pal A, Lin H, Bansil A, Grauer D, Hor YS, Cava RJ, Hasan MZ. Observation of a large-gap topological-insulator class with a single Dirac cone on the surface. Nat Phys 2009;5:398–402.
[18] Ren Z, Taskin AA, Sasaki S, Segawa K, Ando Y. Large bulk resistivity and surface quantum oscillations in the topological insulator Bi2 Te2 Se. Phys Rev B 2010;82: 241306.
[19] Taskin AA, Ren Z, Sasaki S, Segawa K, Ando Y. Observation of Dirac holes and electrons in a topological insulator. Phys Rev Lett 2011;107:016801.
[20] Xiong J, Luo Y, Khoo Y, Jia S, Cava RJ, Ong NP. High-field Shubnikov-de Haas oscillations in the topological insulator Bi_2 Te_2 Se. Phys Rev B 2012;86:045314.
[21] Arakane T, Sato T, Souma S, Kosaka K, Nakayama K, Komatsu M, Takahashi T, Ren Z, Segawa K, Ando Y. Tunable Dirac cone in the topological insulator Bi_2-xSb xTe_3-ySey. Nat Commun 2012;3:636.
[22] Xu Y, Miotkowski I, Liu C, Tian J, Nam H, Alidoust N, Hu J, Shih C-K, Hasan MZ, Chen YP. Observation of topological surface state quantum Hall effect in an intrinsic three-dimensional topological insulator. Nat Phys 2014;10:956–63.
[23] Gao BF, Gehring P, Burghard M, Kern K. Gate-controlled linear magnetoresistance in thin Bi_2Se_3 sheets. Appl Phys Lett 2012;100:212402.
[24] Hikami S, Larkin AI, Nagaoka Y. Spin-orbit interaction and magnetoresistance in the two dimensional random system. Prog Theor Phys 1980;63:707–10.
[25] Xu Y. Quantum transport in three-dimensional topological insulators. Purdue University; 2019.
[26] Klitzing KV, Dorda G, Pepper M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistaqnce. Phys Rev Lett 1980; 45:494–7.
[27] Girivin SM. Topological aspects of low dimensional systems. Springer; 2000.
[28] Prange RE, Girvin SM, Klitzing KV, editors. The quantum Hall effect. 2nd ed. Springer; 1989.

[29] den Nijs M. Quantized Hall conductance in a two dimensional periodic potential. Phys A Stat Mech Appl 1984;124:199–210.
[30] Laughlin RB. Quantized Hall conductivity in two dimensions. Phys Rev B 1981;23: 5632–3.
[31] Fu L, Kane CL. Topological insulators with inversion symmetry. Phys Rev B 2007;76: 045302.
[32] Yoshimi R, Tsukazaki A, Kozuka Y, Falson J, Takahashi KS, Checkelsky JG, Nagaosa N, Kawasaki M, Tokura Y. Quantum Hall effect on top and bottom surface states of topological insulator $(Bi_{11}$-x Sbx$)_2$ Te_3 films. Nat Commun 2015;6:6627.
[33] Zou W, Wang W, Kou X, Lang M, Fan Y, Choi ES, Fedorov AV, Wang K, He L, Xu Y, Wang KL. Observation of Quantum Hall effect in an ultra-thin $(Bi_{0.53}Sb_{0.47})_2Te_3$ film. Appl Phys Lett 2017;110:212401.
[34] Xu Y, Miotkowski I, Chen YP. Quantum transport of two-species Dirac fermions in dual-gated three-dimensional topological insulators. Nat Commun 2016;7:11434.
[35] Aronov AG, Sharvin YV. Magnetic flux effects in disordered conductors. Rev Mod Phys 1987;59:755–79.
[36] Zhang Y, Vishwanath A. Anomalous Aharonov-Bohm conductance oscillations from topological insulator surface states. Phys Rev Lett 2010;105:206601.
[37] Bardarson JH, Brouwer PW, Moore JE. Aharonov-Bohm oscillations in disordered topological insulator nanowires. Phys Rev Lett 2010;105:156803.
[38] Cook A, Franz M. Majorana fermions in a topological-insulator nanowire proximity-coupled to an s-wave superconductor. Phys Rev B 2011;84:201105.
[39] Cook AM, Vazifeh MM, Franz M. Stability of Majorana fermions in proximity-coupled topological insulator nanowires. Phys Rev B 2012;86:155431.
[40] Xiu F, He L, Wang Y, Cheng L, Te Chang L, Lang M, Huang G, Kou X, Zhou Y, Jiang X, Chen Z, Zou J, Shailos A, Wang KL. Manipulating surface states in topological insulator nanoribbons. Nat Nanotechnol 2011;6:216–21.
[41] Peng H, Lai K, Kong D, Meister S, Chen Y, Qi XL, Zhang SC, Shen ZX, Cui Y. Aharonov-Bohm interference in topological insulator nanoribbons. Nat Mater 2010; 9:225–9.
[42] Jauregui LA, Pettes MT, Rokhinson LP, Shi L, Chen YP. Magnetic field-induced helical mode and topological transitions in a topological insulator nanoribbon. Nat Nanotechnol 2016;11:345–51.
[43] Cho S, Dellabetta B, Zhong R, Schneeloch J, Liu T, Gu G, Gilbert MJ, Mason N. Aharonov–Bohm oscillations in a quasi-ballistic three-dimensional topological insulator nanowire. Nat Commun 2015;6:7634.
[44] Hansen AE, Kristensen A, Pedersen S, Sørensen CB, Lindelof PE. Mesoscopic decoherence in Aharonov-Bohm rings. Phys Rev B 2001;64:045327.
[45] Dufouleur J, Veyrat L, Teichgräber A, Neuhaus S, Nowka C, Hampel S, Cayssol J, Schumann J, Eichler B, Schmidt OG, Büchner B, Giraud R. Quasiballistic transport of Dirac fermions in a Bi_2Se_3 nanowire. Phys Rev Lett 2013;110:186806.
[46] Dragoman M, Konstantinidis G, Tsagaraki K, Kostopoulos T, Dragoman D, Neculoiu D. Graphene-like metal-on-silicon field-effect transistor. Nanotechnology 2012;23:305201.
[47] Ning W, Du H, Kong F, Yang J, Han Y, Tian M, Zhang Y. One-dimensional weak antilocalization in single-crystal Bi_2 Te_3 nanowires. Sci Rep 2013;3:1564.
[48] Matsuo S, Koyama T, Shimamura K, Arakawa T, Nishihara Y, Chiba D, Kobayashi K, Ono T, Chang CZ, He K, Ma XC, Xue QK. Weak antilocalization and conductance fluctuation in a submicrometer-sized wire of epitaxial Bi_2Se_3. Phys Rev B Condens Matter 2012;85:075440.

[49] Hamdou B, Gooth J, Dorn A, Pippel E, Nielsch K. Aharonov-Bohm oscillations and weak antilocalization in topological insulator Sb_2Te_3 nanowires. Appl Phys Lett 2013; 102:223110.
[50] Ioselevich PA, Feigel'Man MV. Anomalous Josephson current via Majorana bound states in topological insulators. Phys Rev Lett 2011;106:077003.
[51] Ilan R, Bardarson JH, Sim HS, Moore JE. Detecting perfect transmission in Josephson junctions on the surface of three dimensional topological insulators. New J Phys 2014; 16:053007.
[52] Potter AC, Fu L. Anomalous supercurrent from Majorana states in topological insulator Josephson junctions. Phys Rev B 2013;88:121109.
[53] Tkachov G, Hankiewicz EM. Spin-helical transport in normal and superconducting topological insulators. Phys Status Solidi B 2013;250:215–32.
[54] Olund CT, Zhao E. Current-phase relation for Josephson effect through helical metal. Phys Rev B 2012;86:214515.
[55] Tkachov G. Magnetoelectric Andreev effect due to proximity-induced nonunitary triplet superconductivity in helical metals. Phys Rev Lett 2017;118:016802.
[56] Lee JH, Lee G-H, Park J, Lee J, Nam S-G, Shin Y-S, Kim JS, Lee H-J. Local and nonlocal Fraunhofer-like pattern from an edge-stepped topological surface Josephson current distribution. Nano Lett 2014;14:5029–34.
[57] Wiedenmann J, Bocquillon E, Deacon RS, Hartinger S, Herrmann O, Klapwijk TM, Maier L, Ames C, Brüne C, Gould C, Oiwa A, Ishibashi K, Tarucha S, Buhmann H, Molenkamp LW. 4π-periodic Josephson supercurrent in HgTe-based topological Josephson junctions. Nat Commun 2015;7:10303.
[58] Kurter C, Finck ADK, Hor YS, Van Harlingen DJ. Evidence for an anomalous current-phase relation in topological insulator Josephson junctions. Nat Commun 2015;6:7130.
[59] Xu J-P, Wang M-X, Liu ZL, Ge J-F, Yang X, Liu C, Xu ZA, Guan D, Gao CL, Qian D, Liu Y, Wang Q-H, Zhang F-C, Xue Q-K, Jia J-F. Experimental detection of a Majorana mode in the core of a magnetic vortex inside a topological insulator-superconductor $Bi_2Te_3/NbSe_2$ heterostructure. Phys Rev Lett 2015;114:017001.
[60] Sun H-H, Zhang K-W, Hu L-H, Li C, Wang G-Y, Ma H-Y, Xu Z-A, Gao C-L, Guan D-D, Li Y-Y, Liu C, Qian D, Zhou Y, Fu L, Li S-C, Zhang F-C, Jia J-F. Majorana zero mode detected with spin selective Andreev reflection in the vortex of a topological superconductor. Phys Rev Lett 2016;116:257003.
[61] Jauregui LALA, Kayyalha M, Kazakov A, Miotkowski I, Rokhinson LPLP, Chen YPYP. Gate-tunable supercurrent and multiple Andreev reflections in a superconductor-topological insulator nanoribbon-superconductor hybrid device. Appl Phys Lett 2018;112:093105.
[62] Ghatak S, Breunig O, Yang F, Wang Z, Taskin AA, Ando Y. Anomalous fraunhofer patterns in gated Josephson junctions based on the bulk-insulating topological insulator $BiSbTeSe_2$. Nano Lett 2018;18:5124–31.
[63] Kayyalha M, Kargarian M, Kazakov A, Miotkowski I, Galitski VM, Yakovenko VM, Rokhinson LP, Chen YP. Anomalous low-temperature enhancement of supercurrent in topological-insulator nanoribbon Josephson junctions: evidence for low-energy Andreev bound states. Phys Rev Lett 2019;122:047003.
[64] Kayyalha M, Kazakov A, Miotkowski I, Khlebnikov S, Rokhinson LP, Chen YP. Highly skewed current-phase relation in superconductor-topological insulator-superconductor Josephson junctions. Npj Quantum Mater 2020;5:7.
[65] Lutchyn RM, Sau JD, Das Sarma S. Majorana fermions and a topological phase transition in semiconductor-superconductor heterostructures. Phys Rev Lett 2010;105: 077001.

[66] Oreg Y, Refael G, von Oppen F. Helical liquids and Majorana bound states in quantum wires. Phys Rev Lett 2010;105:177002.
[67] Mourik V, Zuo K, Frolov SM, Plissard SR, Bakkers EPAM, Kouwenhoven LP. Signatures of Majorana fermions in hybrid superconductor-semiconductor nanowire devices. Science 2012;336:1003–7.
[68] Das A, Ronen Y, Most Y, Oreg Y, Heiblum M, Shtrikman H. Zero-bias peaks and splitting in an Al–InAs nanowire topological superconductor as a signature of Majorana fermions. Nat Phys 2012;8:887.
[69] Albrecht SM, Higginbotham AP, Madsen M, Kuemmeth F, Jespersen TS, Nygård J, Krogstrup P, Marcus CM. Exponential protection of zero modes in Majorana islands. Nature 2016;531:206–9.
[70] Zhang H, Liu C-X, Gazibegovic S, Xu D, Logan JA, Wang G, Van Loo N, Bommer JDS, De Moor MWA, Car D, Op Het Veld RLM, Van Veldhoven J, Koelling S, Verheijen MA, Pendharkar M, Pennachio DJ, Shojaei B, Lee JS, Palmstrøm CJ, Bakkers EPAM, Das Sarma S, Kouwenhoven LP. Quantized Majorana conductance. Nature 2018;556:74–9.
[71] Ghaemi P, Nair VP. Effect of impurities on the Josephson current through helical metals: exploiting a neutrino paradigm. Phys Rev Lett 2016;116:037001.
[72] Golubov AA, Kupriyanov MY, Il'ichev E. The current-phase relation in Josephson junctions. Rev Mod Phys 2004;76:411–69.
[73] Peng Y, Pientka F, Berg E, Oreg Y, von Oppen F. Signatures of topological Josephson junctions. Phys Rev B 2016;94:085409.
[74] Likharev KK. Superconducting weak links. Rev Mod Phys 1979;51:101–59.
[75] Dubos P, Courtois H, Pannetier B, Wilhelm FK, Zaikin AD, Schon G. Josephson critical current in a long mesoscopic S-N-S junction. Phys Rev B 2001;63:064502.
[76] Angers L, Chiodi F, Montambaux G, Ferrier M, Guéron S, Bouchiat H, Cuevas JC. Proximity dc squids in the long-junction limit. Phys Rev B 2008;77:165408.
[77] Ting Ke C, V Borzenets I, Draelos AW, Amet F, Bomze Y, Jones G, Craciun M, Russo S, Yamamoto M, Tarucha S, Finkelstein G. Critical current scaling in long diffusive graphene-based Josephson junctions. Nano Lett 2016;16:4788–91.
[78] V Borzenets I, Amet F, Ke CT, Draelos AW, Wei MT, Seredinski A, Watanabe K, Taniguchi T, Bomze Y, Yamamoto M, Tarucha S, Finkelstein G. Ballistic graphene Josephson junctions from the short to the long junction regimes. Phys Rev Lett 2016;117:237002.
[79] Talantsev. Classifying induced superconductivity in atomically thin Dirac-cone materials. Condens Matter 2019;4:83.
[80] de Juan F, Bardarson JH, Ilan R. Conditions for fully gapped topological superconductivity in topological insulator nanowires. SciPost Phys 2019;6:060.
[81] Bercioux D, Cayssol J, Vergniory MG, Calvo MR. Topological matter: lectures from the topological matter school 2017. Springer; 2018.
[82] Fornieri A, Amado M, Carillo F, Dolcini F, Biasiol G, Sorba L, Pellegrini V, Giazotto F. A ballistic quantum ring Josephson interferometer. Nanotechnology 2013;24:245201.
[83] Kayyalha M. Electrical, thermoelectric, and phase coherence transport in two-dimensional materials. PhD Dissertation. Purdue University; 2018.

CHAPTER EIGHT

Silicon qubit devices

Simon Schaal[1,2] and M. Fernando Gonzalez-Zalba[3,*]
[1]London Centre for Nanotechnology, University College London, London, United Kingdom
[2]Intel Components Research, Intel Corporation, Hillsboro, OR, United States
[3]Quantum Motion Technologies, London, United Kingdom
*Corresponding author: mg507@cam.ac.uk

1. Introduction

1.1 Moore's law

Over the last 5 decades, we have undergone a revolution in information technology and electronics driven by the innovation of a single circuit element – the transistor. As a result of these innovations, transistor density in integrated circuits has doubled every 2 years over the last 50 years, following a trend predicted by Gordon Moore, the co-founder of Intel, in Moore's law. State-of-the-art complementary metal–oxide–semiconductor (CMOS) circuits have reached dimensions of only a few nanometres. In 2019, some foundries started production of the 5 nm node, where a single atom being in the wrong place could strongly affect circuit performance. A fundamental physical limit has therefore been reached and the semiconductor industry has formally acknowledged that Moore's law is coming to an end [1]. Consequently, nowadays, processors are predominately becoming more powerful by using multi-core architectures. Supercomputers consist of a network of such processors and can perform more than 10^{15} floating-point operations per second. Still, state-of-the-art supercomputers are not powerful enough to tackle some relevant computationally hard problems where no algorithms exist that provide a solution within polynomial time. Examples include simulating quantum chemistry, many-body physics and optimisation tasks. A significant increase in computing power or new ideas is required to address these problems. Interestingly, defects and quantum effects limiting stability and down-scaling of conventional silicon-based processors are promising candidates for realising a quantum computer capable of solving computationally hard problems that seem intractable using conventional computers.

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00005-8
© 2021 Elsevier Ltd.
All rights reserved.

1.2 Quantum computation

With the idea in mind, that there will be a physical limit in the density of transistors, researchers had already started to investigate alternative methods to conventional computation in the 1980s. Among the first to imagine that a quantum system could be used to perform a computation was Richard Feynman. He introduced the idea of performing quantum simulations by mapping a quantum system of interest onto a controllable quantum device. Only a few years later, the foundation of a universal (general purpose) quantum computer with capabilities beyond traditional computation was laid out by David Deutsch showing that the Church–Turing principle of universal computation and quantum theory are compatible.

The building block of a quantum computer is a so-called quantum bit (qubit), realised using a quantum mechanical two-level system. As a result of the quantum principle of superposition, any mixture of the two states, typically referred to as 0 and 1, is allowed, making quantum computing appear similar to an analogue computer but in two dimensions given the additional degree of freedom of the quantum mechanical phase. However, when a qubit is measured, only either a 0 or 1 is obtained (Fig. 8.1). Repeated measurements enable the exploration of the probabilistic nature of these projective quantum measurements with $|\alpha|^2$ and $|\beta|^2$ being the probability to measure 0 or 1, respectively. Any superposition state $|\psi\rangle$ of a qubit can be visualised on a so-called Bloch sphere (sphere of radius 1) where the north and south pole correspond to the logical states 0 and 1.

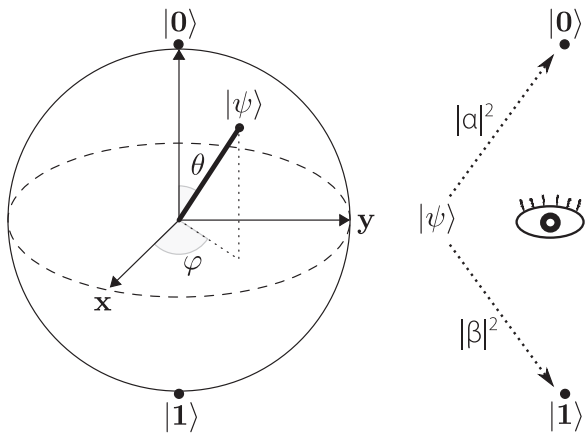


Figure 8.1 ***Qubit.*** Illustration of a qubit state $|\psi\rangle$ in a Bloch sphere. In a measurement of a superposition state, either 0 or 1 is obtained.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle \tag{8.1}$$

Superposition gives rise to a quantum parallelism, which can be illustrated by considering an n-qubit superposition state which generates a vast 2^n dimensional Hilbert space

$$|\psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i|i\rangle. \tag{8.2}$$

Moreover, qubits can be entangled where they act as a group, rather than being fully isolated. These effects illustrate the power of a quantum computer and the difficulty of classically simulating such quantum systems.

To make use of superposition and entanglement within a computation, well-tailored algorithms are required to unleash the potential computational speed up [2]. In gate-based quantum computation, an algorithm is performed by manipulating qubits in a series of gates, similar to conventional computing, realised using unitary transformations acting on single or pairs of qubits. Multiple universal sets of gates exist which any possible operation on a quantum computer can be reduced to, e.g., the controlled-NOT gate between two qubits combined with single qubit rotations forms a universal set [3]. In 1994, Shor's quantum algorithm for factorisation brought the vision of quantum computation to the world's attention, by making quantum computers a threat to common RSA public-key encryption due to an expected exponential speed-up. Later on, Grover presented an algorithm that offers a quadratic speed-up when searching an unstructured database. Moreover, as envisioned by Feynman, a quantum computer allows solutions to be found to many problems in chemistry and physics, such as understanding reactions and materials which could lead to efficient fertiliser production, novel pharmaceuticals and more efficient solar cells. Research on algorithms has shown that algorithms formulated so far are based on a small number of different concepts, e.g., quantum Fourier transform (Shor), amplitude amplification (Grover), quantum walks and quantum simulation. There is ongoing research in designing new algorithms.

1.3 Quantum computing platforms

David DiVincenzo [4] has summarised the requirements for physically implementing a qubit as follows: It needs to be a 'scalable physical system with well characterised qubits' and with the 'ability to initialise the state of the qubits' whilst remaining coherent 'much longer than the gate

operation time'. Additionally, there must be a 'universal set of quantum gates' and a 'qubit specific measurement capability' allowing manipulation and readout of every qubit reliably with high fidelity. Finding and engineering a quantum system which fulfills these somewhat contradictory criteria, of a well isolated but still on-demand well controlled system, represents the challenge of building a quantum computer.

Quantum state coherence and gate fidelity play an important role for a quantum computer due to the fragile and nondigital nature of quantum states. Even a very small error in the quantum state phase can alter the result of the computation and demonstrates that errors do not only have bit-flip character. Consequently, error detection and correction are an essential ingredient for large-scale quantum computation where errors within a computation are more likely to occur. The error rate, given by the number of gates that can be performed before an error occurs, becomes a parameter as important as the number of qubits to assess how powerful a computer can be. Protecting qubits from errors comes at the cost of a large overhead in the number of qubits: e.g., on the order of 10^8 physical qubits are expected to be required to run some of the most complex quantum algorithms fault tolerantly [2].

Many proposals on implementing a quantum computer have emerged [5]. Implementations based on superconducting circuits [6,7] and ion-traps [8] are the most promising candidates to demonstrate an advantage over conventional computers based on devices with close to 100 qubits and high fidelity but still imperfect gates. In silicon-based devices, qubits are formed using quantum states of confined electron, hole or nuclear donor spins [9,10,72]. These spin states have long coherence times and fast gate operations can be achieved. Moreover, the structures necessary to form a qubit are smaller than 100 nm and qubits could be operated at elevated temperatures of 1 K or higher [11] and integrated with conventional electronics [12,73]. More recently, computing at the surface code thresholds has been demonstrated in a two-qubit device [74]. Currently, demonstrators only consist of a few functional qubits [13], but silicon spin qubits have excellent future prospects to reach a scale large enough to tackle the most challenging and relevant problems and perform fault-tolerant quantum computations.

1.4 This chapter

In this chapter, spin qubit devices in silicon are introduced and essential concepts for the formation, manipulation and readout of such devices are discussed. Detailed reviews can be found in Refs. [14,15,75].

2. Fabrication

Nanofabrication is necessary to form structures to confine, control and readout single spins in silicon. In the following fabrication of structures for spin qubit applications in silicon are discussed. One of the most common structures is a quantum dot (QD), where confinement of single electrons is achieved.

2.1 Silicon host material

Semiconductor-based qubits have first been implemented in GaAs-based systems. GaAs-based materials can be grown with extremely high purity and atomic layer precision in heterostructures, which has led to the formation of two-dimensional electron gases (2DEGs) of very high mobility. Additional electron confinement has been reached by device patterning via etching or using voltages applied to metal structures at the surface. Confinement in all remaining dimensions has been reached resulting in the formation of so-called QDs [16].

Silicon has similar properties but offers some advantages as a host material for spin-based qubits, specifically in terms of qubit coherence and large-scale fabrication; however, some challenges had to be overcome in order to implement spin qubits in silicon devices [14]. To begin, the increased effective mass of electrons in silicon compared to other materials (e.g., factor of $\sim$3 compared to GaAs) requires scaling down of gate structure dimensions in order to reach a similar separation of the QD energy levels. Due to advances in lithography techniques, quantum devices are now routinely fabricated in university and industrial clean rooms [17–20]. Moreover, bulk silicon is a semiconductor with indirect band gap and diamond cubic symmetry leading to a sixfold degenerate conduction band minimum. This valley degeneracy could in principle be problematic when forming a quantum bit that relies on a nondegenerate spin two-level system; however, a sufficient valley splitting between 0.3 and 1.5 meV can be typically achieved in silicon MOS devices via strong confinement, electric fields and strain. In Si/SiGe, the valley splitting, typically between 0.1 and 0.3 meV, poses more serious challenges in terms of operating temperature and scalability [21].

Having overcome these challenges, the excellent mechanical and electrical properties and highly developed large-scale fabrication industry make silicon an ideal host material for fabrication of large-scale quantum processors. Additionally, silicon-based spin qubits offer very long coherence and relaxation times

compared to other semiconductor qubits. Spin relaxation due to energy exchange is characterised by the timescale T_1 which is typically mediated by phonons. Phonons cannot directly induce spin relaxation of pure spin states; however, spin–orbit coupling, that arises for electron spins in a lattice potential and electron spins confined in QDs [14], couples spin levels to orbital levels lifting phonon-related relaxation rules. Silicon has no piezoelectric phonons and an overall weak spin–orbit coupling, resulting in a long spin relaxation time on the order of seconds in electron spin qubit devices at cryogenic temperature [11,22,23,76]. Finally, silicon has a low natural abundance of nuclear spin carrying isotopes ^{29}Si (5%) and can be purified to mostly (e.g., 800 ppm) nuclear spin free isotopes ^{28}Si and ^{30}Si. The hyper-fine interaction between a spin qubit and constantly fluctuating spins in the host lattice provides a large contribution to de-coherence time described by T_2 (irreversible loss of coherence), which is strongly suppressed in the 'semiconductor vacuum' of isotopically purified silicon, where T_2 can be charge noise limited [24] and reach values up to 28 ms [28]. Consequently, even at cryogenic temperature up to 1.45 K, operation of a spin qubit formed using a single electron spin is possible [11].

2.2 Si-MOS

In planar Si-MOS devices, QDs are formed at the silicon/oxide interface using metal gate structures at the surface (without etching of the silicon layer) similar to planar transistors. Fig. 8.2 shows an illustration of a Si-MOS QD device [25] including a summary of processing steps developed for university clean room fabrication [20].

To begin, a field oxide (SiO_2) is grown on a bare (high resistivity) silicon wafer which is typically done by controlled oxidation in a furnace. Using micro-fabrication patterning steps, involving spinning of a photo-resist, writing of a pattern using optical lithography followed by oxide wet etching, regions for ion implantation are defined. Ion implantation is the basis to form electron reservoirs and ohmic contacts. Via such reservoirs, transport measurements can be performed and single electrons can be loaded onto a QD device. After ion implantation, the contaminated oxide is removed followed by micro-patterning of the active device area, where the field oxide is removed and a thin gate oxide (typically 5–10 nm SiO_2) of low trap density is deposited. A final micro-fabrication patterning, metal deposition and lift-off step provide metal contacts for bonding after which the chip is ready for nanofabrication of the QD gate structures.

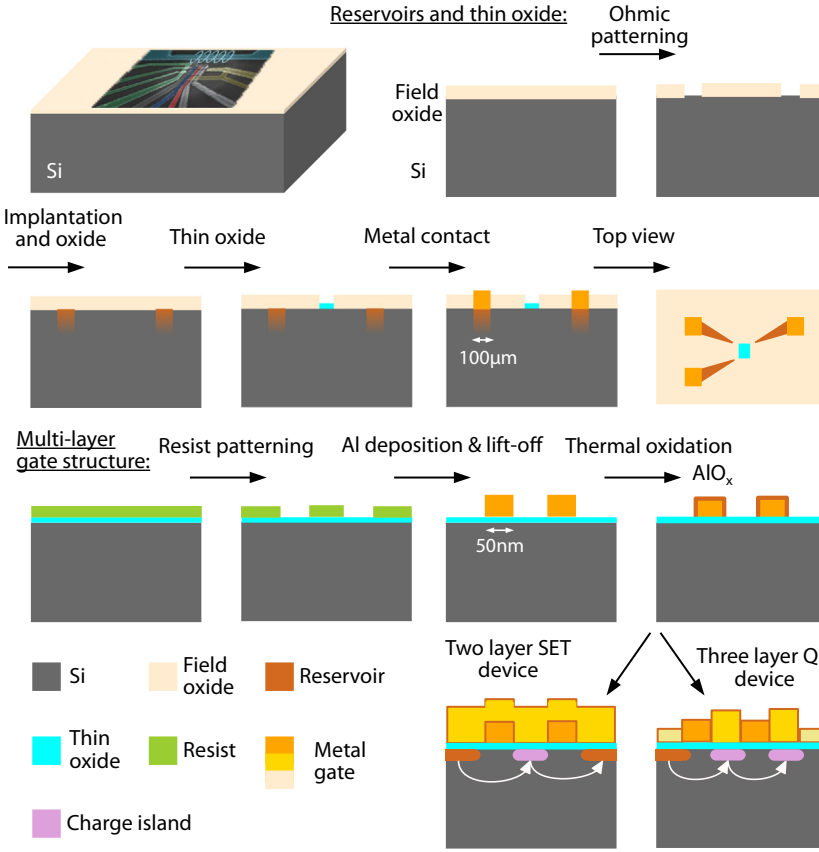


Figure 8.2 ***Planar Si-MOS device fabrication.*** Simplified micro-fabrication steps to form electron reservoirs and a high-quality thin oxide. On the thin-oxide multi-layer gate, metal structures are fabricated using nanofabrication to form SETs and QDs. *Based on Ref. [20].*

Development of a technique for fabrication of thin and dense gate structures has been a key milestone for reliable QD formation in planar Si-MOS devices [26]. To begin, PMMA resist is spun and a first set of gate structures is patterned using electron-beam patterning (typical separation and width < 50 nm) followed by deposition of a 30-nm-thick aluminium (Al) film. In a lift-off process, Al metal is removed in any region that has been left covered with resist, whilst the Al gate pattern remains. The set of Al gates is oxidised on a hotplate, where a thin AlO_x oxide layer is formed, providing electrical isolation and allowing for fabrication of subsequent and overlapping gates. Additional structures are fabricated in the same process, but typically the Al layer thickness is increased by ≈ 20 nm for each additional layer to

avoid breaking of gate structures at points of overlap. Such multi-layer gate fabrication allows formation of a single-electron transistor (SET) and a dense dimensional chain of QDs. A SET consists of an overlapping top gate for accumulation and barriers for depletion, whilst a QD structure consists of multiple slightly overlapping barrier and accumulation gates. Exemplary devices fabricated using this technique can be found in Refs. [20,26–28].

2.3 Si/SiGe

Rather than using a bulk Si wafer, QDs can be formed in a similar way in a Si/SiGe heterostructure. In such a heterostructure, QDs are commonly formed in a Si quantum well buried 50 nm beneath the surface between two layers of $Si_{0.7}Ge_{0.3}$. Typically, a global top gate is used for the accumulation of a 2DEG in the quantum well and fine gate structures provide further confinement through depletion [29].

Fig. 8.3 illustrates the fabrication process of such depletion-mode devices. To begin, a thin dielectric (typically Al_2O_3) is deposited using atomic layer deposition. A micro-mesa structure is etched to prevent accumulation of charges in regions far from the QD device. To make ohmic contacts to the Si quantum well, the oxide is selectively removed in a patterning and etching step, followed by ion implantation. In a lift-off step as illustrated in Fig. 8.2, fine metal gates are deposited (Al, Al/Ti or Au) for ohmic contacts and depletion gates that run up to the edge of the micro-mesa. Subsequently, thick metal gates are deposited thick enough to fully climb the micro-mesa and connect bonding pads at one end with the fine gates at the other end. To form the top gate for accumulation, first another layer of dielectric is deposited followed by gate deposition in a lift-off step. Exemplary devices fabricated using this technique can be found in Refs. [13,31]. Recently, Si/SiGe QD devices using a dense multi-layer gate structure of confinement and accumulation gates similar to Si-MOS devices have been demosntrated. Such accumulation-mode devices have shown excellent control over electron confinement and electron tunnelling rates [19,77].

2.4 SOI

Many of the processes introduced in the previous subsections are incompatible with industrial scale CMOS fabrication technology, such as electron-beam patterning, deposition of Al films and metal lift-off. Using fully depleted silicon-on-insulator (SOI) nanowire technology, QDs have been formed using purely optical and dry-etching techniques in a transistor (wrap-around gate) geometry. Fabrication processes for such single-electron devices have been developed at CEA-Leti [17,18] on a 300 mm

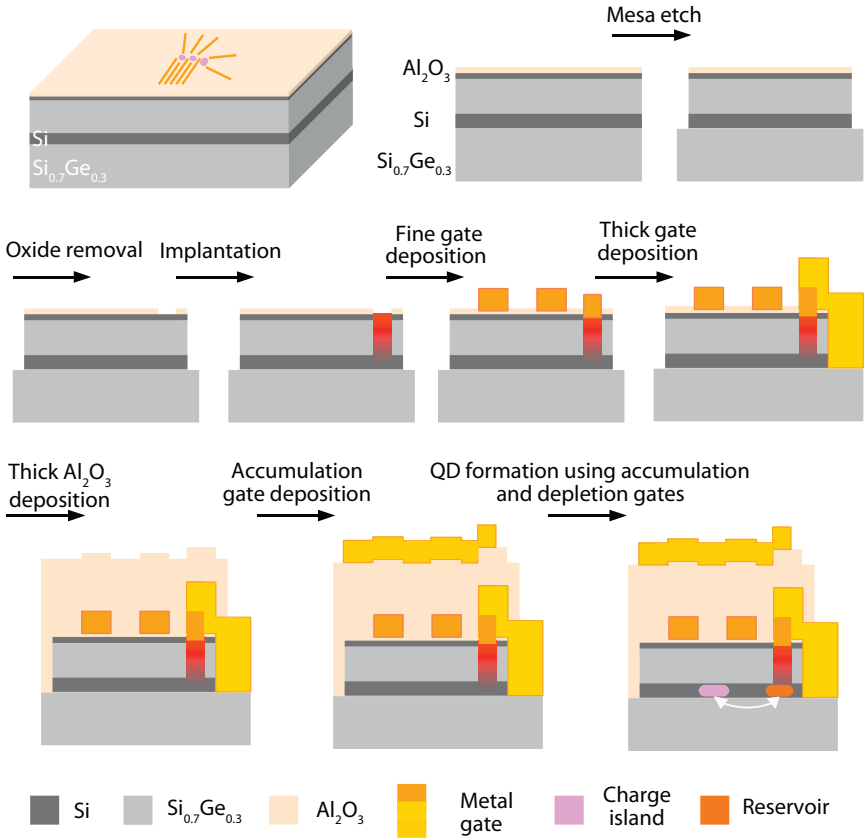


Figure 8.3 ***SiGe QD device fabrication.*** Simplified fabrication steps to form a micro-mesa around the device, followed by ohmic contact implantation, depletion and accumulation gate deposition separated by aluminum oxide. *Based on Ref. [30].*

CMOS platform using 193 nm DUV lithography. One of the key achievements to define structures small enough to form QDs using these optical techniques has been resist trimming.

Fabrication is based on an SOI wafer consisting of an 850 μm thick Si substrate, 145 μm thick buried oxide and ~10 nm thick Si layer at the surface. To form a nanowire transistor device, a nanowire is etched followed by gate deposition and source/drain formation as shown in Fig. 8.4. For each patterning step using DUV lithography, a bottom anti-reflective coating (BARC) and resist is spun onto the SOI wafer. The smallest feature size that can be obtained using DUV lithography is 80 nm; however, combined with resist trimming, features down to 5 nm can be obtained as illustrated in Fig. 8.4. After the resist (and BARC) is trimmed to the desired width, a

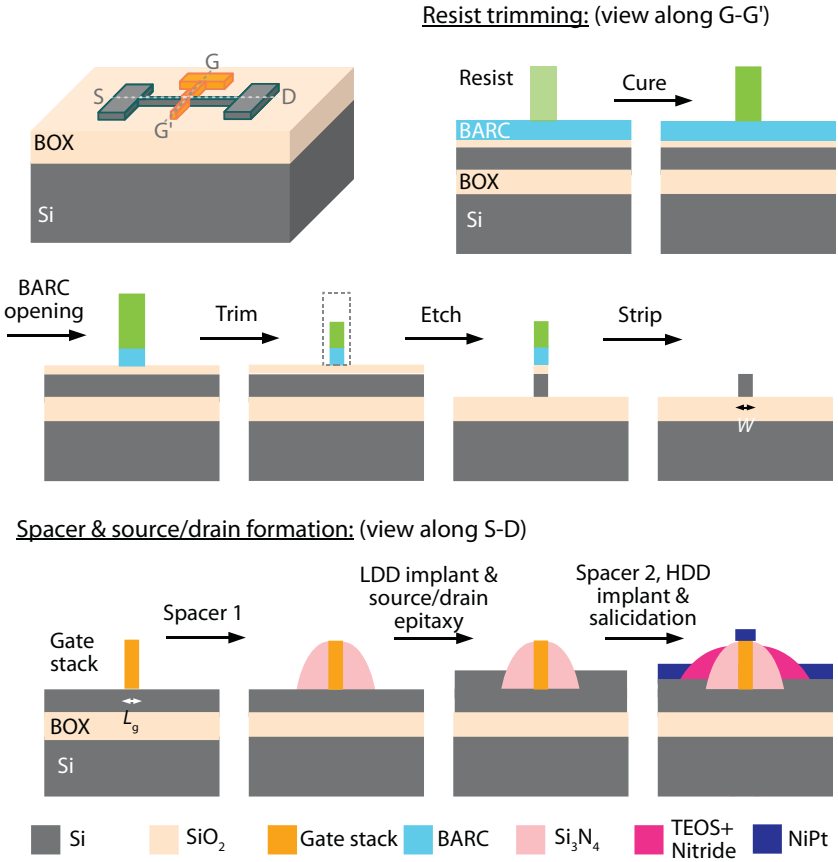


Figure 8.4 ***Si nanowire transistor fabrication.*** Nanowire and gate patterning is performed using a DUV lithography and resist trimming process. Formation of tunnel barriers is achieved using spacers. Spacer and source/drain formation is illustrated. *Based on Refs. [17,18].*

nanowire is etched and any remaining resist is removed. A gate is formed using a similar DUV lithography and resist trimming process. The gate stack consists of a thin 0.8 nm layer SiO_2, 1.9 nm layer of high-k dielectric HfSiON followed by a 5 nm thick layer of TiN and 50 nm of poly-Si with an equivalent oxide thickness of 1.3 nm. To fabricate source and drain reservoirs, at first silicon–nitride spacers are formed at both sides of the gate followed by silicon epitaxy and low-dose doping. Finally a second set of spacer and salicidation enables contacts to be made to the device using standard tungsten interconnects followed by Cu contact vias. Devices with a nanowire width W and gate length L_g ranging from 10 µm down to 10 nm have been fabricated using these techniques.

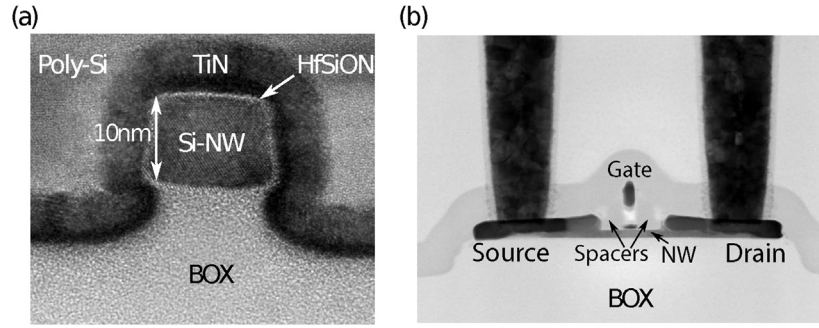


Figure 8.5 ***Nanowire transistor device.*** (a) TEM images of a nanowire transistor device along the gate (G-G′ in Fig. 8.4) and (b) source to drain (S-D in Fig. 8.4). *Images courtesy Dr. S. Barraud, CEA Leti, Grenoble.*

Transmission electron microscope images of a silicon nanowire and gate stack (cross-sectional view along G-G′) and an image of a contacted device (cross-sectional view along S-D) are shown in Fig. 8.5a and b, respectively. When a positive voltage is applied to the gate of such nanowire transistor, a current is measured between the source and drain contacts. Close to threshold, transport channels form only at the top corners of the nanowire due to the enhanced electric field. This corner effect is the basis for the formation of the so-called corner QDs. Combined with disorder due to surface roughness and charge traps, such channels turn into localised QDs at cryogenic temperature [32]. Depending on the nanowire width and voltage applied to the Si substrate that acts as a back gate, a single or a double QD can be formed in a single nanowire. Fabrication using industrial processes can provide routes to reproducible large-scale fabrication and integration of quantum devices with conventional electronics and circuits.

3. Silicon spin qubits

3.1 Single-spin qubit

A qubit based on a single spin 1/2 is one of the most natural two-level quantum systems. It is formed by the Zeeman energy split spin-up and down states $|0\rangle = |\downarrow\rangle$ and $|1\rangle = |\uparrow\rangle$ subject to an external magnetic field B_z (usually > 1 T) with Hamiltonian

$$H_0 = g\mu_B B_z \widehat{S}_z = \frac{\hbar\omega_0}{2}\,\sigma_z, \tag{8.3}$$

where g is the electron g-factor, μ_B the Bohr magneton, $\widehat{S}$ the spin operator and σ_z the Pauli z matrix. Quantum information processing using such

spin 1/2 particles confined in semiconductor devices has been proposed by Loss-DiVincenzo [33] for electron spins in QDs and Kane [34] for donor nuclear spins. In the following, operation of an electron spin qubit is discussed in detail through the example of a QD-based qubit. Many of these concepts also apply to donor-based qubits. In addition to a spin 1/2 electron, donors in silicon devices also have a nuclear spin that can be used as an additional qubit or memory for quantum information processing.

3.1.1 Qubit control and operation

Coherent oscillations between the Zeeman-split spin 1/2 levels can be achieved using electron spin resonance (ESR). In ESR, an ac magnetic field B_{ac} (perpendicular to the constant external field B_z) at frequency $\omega \approx g\mu_B B_z/\hbar$, close to the spin splitting (typically at microwave frequencies), drives coherent spin oscillations. This can be shown by analysing the time-dependent Hamiltonian in an external field B_z and drive field B_{ac} given by the following:

$$H(t) = \frac{\hbar\omega_0}{2}\sigma_z + \hbar\Omega\cos(\omega t + \delta)\sigma_x \tag{8.4}$$

with $\Omega = g\mu_B B_{\mathrm{ac}}/\hbar$ and $\Omega \ll \omega_0,\ \omega$. The dynamics due to this Hamiltonian are solved by transforming into the reference frame rotating at frequency ω of the ac field B_{ac} and neglecting faster oscillating components which leads to an effective Hamiltonian

$$H_{\mathrm{rot}} = \frac{\hbar}{2}\begin{pmatrix} \Delta & \Omega e^{-i\delta} \\ \Omega e^{i\delta} & -\Delta \end{pmatrix}. \tag{8.5}$$

In this effective Hamiltonian, z-rotations are performed at angular frequency given by $\Delta = \omega_0 - \omega$, whilst rotations around an axis in the x–y plane (determined by the phase δ) are performed at angular frequency Ω (determined by B_{ac}). Fig. 8.6a shows an exemplary QD spin qubit device where electron spins are confined in the QDs at the Si/SiO_2 interface using voltages applied to the accumulation gates at the surface (see Section 1.2 for fabrication details and other ways to form QDs). Spin manipulation is achieved using a shorted on-chip wave guide (strip line) that ends very close to the QDs providing a local oscillating magnetic field. Multiple spins can be controlled with the same strip line (as it is close enough), where local g-factor anisotropies and/or hyperfine interaction [28,35] can lead to a distribution in manipulation frequencies. High fidelity control (99.6% and better) of multiple spins has been achieved [28]; however, the frequency of coherent spin

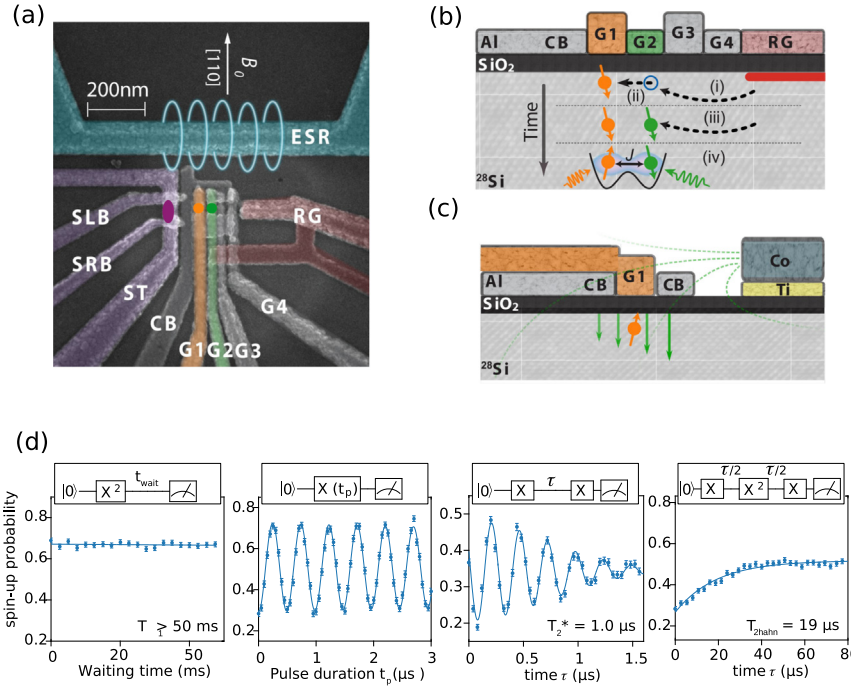


Figure 8.6 ***Spin qubit device operation and characterisation.*** (a) False-coloured SEM micrograph of a two-qubit QD device including a microwave strip line for control and single-electron transistor for readout. (b) Cross-sectional view along the qubit device. Operation is illustrated over time: electron spins are loaded and initialised (i–iii) followed by spin manipulation and interaction (iv). (c) Cross-sectional view along G1 of a device with micro-magnet instead of an ESR line. (d) Set of experiments for spin qubit characterisation. Measurement of the relaxation time, manipulation speed through Rabi oscillations, de-phasing time in a Ramsey sequence and de-coherence time in a Hahn echo experiment. *(a and b) Modified from Ref. [25] with permission from the author. (d) Modified from Ref. [13] with permission from the author.*

oscillations is so far limited to ∼ 1 MHz due to limited magnitude of B_{ac} that can be delivered to the location of the spins before a further increase in signal power results in reduced spin coherence.

Additionally, electron spins can be manipulated using electric fields via electric dipole spin resonance (EDSR) through mechanisms that indirectly couple spin to electric fields. EDSR is achieved using spin–orbit coupling [36] and asymmetries in the local environment, which can be in the form of a local change in g-factor [35], a changing local nuclear spin environment [37] or using a micro-magnet [31,38]. The latter is the most successful approach as it can provide local field changes on the order of $\Delta B_z = 30$ mT between spins resulting in coherent oscillations up to

30 MHz [24,39]. An exemplary micro-magnet is shown in Fig. 8.6c and manipulation is achieved using an oscillating voltage on G1 that translates into an effective oscillating magnetic field in the rest frame of the electron spin which is moved quickly through an anisotropic environment.

Qubit operation and characterisation is performed in pulsed experiments (see Fig. 8.6b). First spins are initialised using spin-selective tunnelling [40] or by loading a random spin from a reservoir followed by waiting for time $t_{\text{wait}} \gg T_1$ until the spin relaxes. In a multi-dot system, high-fidelity initialisation can also be achieved by pulsing to a specific relaxation hotspot [13] (typically close to a state anti-crossing). Once spins are initialised, single qubit gates can be performed using ESR or EDSR. Typical experiments for characterising important spin qubit parameters are shown in Fig. 8.6d. By initialising a spin-up state followed by a measurement after time t_{wait}, the relaxation time Ti is obtained from the exponential decay of the spin-up probability. Coherent (Rabi) oscillations between the spin-up and down state can be observed in an experiment where the duration t_p of a microwave drive is varied. The in-plane de-phasing time T_2^* is obtained from the free induction decay in a Ramsey sequence, where the spin is initialised, brought into a superposition state for a varying free evolution time τ, and then projected onto the z axis followed by a measurement. The de-coherence time T_2 is obtained in a Hahn-echo sequence, which is similar to the Ramsey experiment but includes a spin flip in the middle of the free evolution time which results in re-focussing of the spin. Multiple spin flips can be incorporated for further enhancement of the coherence time known as Carr–Purcell–Meiboom–Gill sequence. De-phasing and de-coherence times of $T_2^* = 120$ μs and $T_2^{\text{CPMG}} = 28$ ms for electron spins in QDs [28], and $T_{2e}^* = 270$ μs, $T_{2e}^{\text{CPMG}} = 560$ ms and $T_{2n}^* = 1.75$ s and $T_{2n}^{\text{CPMG}} = 35$ s for the electron and nuclear spin of donors [23] in isotopically purified ^{28}Si have been demonstrated combined with control fidelities exceeding 99.6% in all implementations and reaching up to $> 99.9\%$ using pulse optimisation [41], demonstrating that a spin 1/2 systems in silicon provides an excellent basis for spin qubits and memories. Recently, spin qubits formed using hole states in gate-defined germanium QDs have additionally shown to provide an excellent basis for qubits with de-coherence times $T_2^* > 140$ ns and $T_2^{\text{CPMG}} = 100$ μs and with control fidelities exceeding 99% exploiting a large spin-orbit coupling for fast manipulation [78].

3.1.2 Two-qubit gate

When two spins are brought close together, spin singlet (S) and triplet (T_0, T_+, T_-) states are formed. Two electron spins in a double QD device with

finite tunnel coupling (allowing a transfer from one dot to the other) can either interact when confined in the same dot (0,2) or when distributed over both neighboring dots (1,1). For an even number of electron spins, transfer from one dot to the other becomes spin dependent due to the Pauli exclusion principle leading to higher energies for (0,2) triplets. The spin Hamiltonian of a coupled double QD can be formally derived in second quantisation including contributions from the on-site energy H_ε, the Zeeman energy H_z, valley degeneracy H_v, Hubbard-style Coulomb repulsion H_U and tunnel coupling H_t. Assuming sufficient valley splitting and ignoring the higher energy (0,2) triplets, the Hamiltonian can be reduced into the basis of $\left\{(1,1):|\uparrow,\uparrow\rangle,|\uparrow,\downarrow\rangle,|\downarrow,\uparrow\rangle,|\downarrow,\downarrow\rangle;(0,2):\frac{1}{\sqrt{2}}|0,\uparrow\downarrow-\downarrow\uparrow\rangle\right\}$ and written as [42,43]

$$H=\begin{pmatrix} E_z & 0 & 0 & 0 & 0 \\ 0 & \delta E_z/2 & 0 & 0 & t_e \\ 0 & 0 & -\delta E_z/2 & 0 & -t_e \\ 0 & 0 & 0 & -E_z & 0 \\ 0 & t_e & -t_e & 0 & U-\varepsilon \end{pmatrix} \tag{8.6}$$

with E_z being the Zeeman energy, δE_z the difference in Zeeman energy between the two dots, t_c the tunnel coupling, ε the energy de-tuning between both dots and U the charging energy to move both electrons into the same dot. An illustration of the resulting energy diagram is shown in Fig. 8.7a. Using a Schrieffer–Wolff transformation and assuming $U-\varepsilon \gg t_c$, the Hamiltonian can effectively (to first order) be reduced to the well-known Heisenberg interaction in the basis of individual (1,1) spin states $\{|\uparrow,\uparrow\rangle,|\uparrow,\downarrow\rangle,|\downarrow,\uparrow\rangle,|\downarrow,\downarrow\rangle\}$ [42].

$$H_{\text{int}}=\begin{pmatrix} E_z & 0 & 0 & 0 \\ 0 & \frac{\delta E_z}{2}-\frac{J}{4} & \frac{J}{2} & 0 \\ 0 & \frac{J}{2} & -\frac{\delta E_z}{2}-\frac{J}{4} & 0 \\ 0 & 0 & 0 & -E_z \end{pmatrix} \tag{8.7}$$

with $J=\frac{2t_c^2}{U-\varepsilon-\delta E_z}+\frac{2t_c^2}{U-\varepsilon+\delta E_z}$ being the Heisenberg exchange interaction strength. Based on this interaction, there are three ways to construct a two-qubit gate.

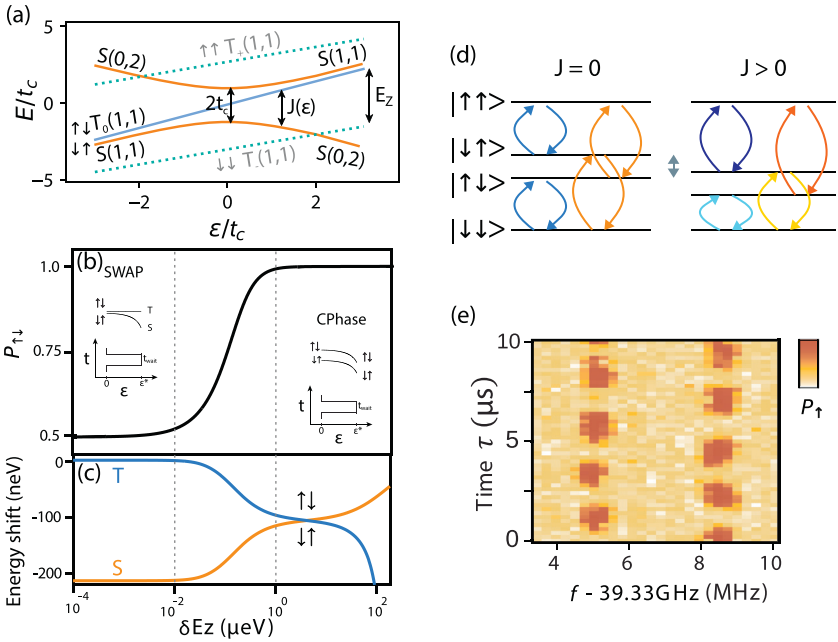


Figure 8.7 ***Two-qubit gate.*** (a) Energy level diagram of a coupled double QD as a function of energy de-tuning. (b) Probability of $|\uparrow\downarrow\rangle$ and (c) energy difference between $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ as a function of δE_z, SWAP gate is performed at small δE_z where singlet and triplet are the eigenstates, whilst a CPhase gate is obtained at large δE_z. (d) Illustration of resonant drive between two qubit states for $J = 0$ and $J > 0$. At $J > 0$ oscillations of one spin conditional on the other spin can be performed. (e) Change in drive frequency of qubit 1 as qubit 2 oscillates between two states over time, demonstrating conditional resonant driving. *(c) Modified from Ref. [43] with permission from the author. (e) Modified from Ref. [25] with permission from the author.*

Exchange gate: For Heisenberg exchange only, when $\delta E_z \ll J$, the off diagonal terms drive a flip-flop of the two spins, and singlet-triplet states are eigenstates which allows the realisation of a SWAP gate (see Fig. 8.7b). Combined with single qubit rotations along the z-axis $R_z^{(1)/(2)}(\theta)$ (of qubit 1 and 2) a CNOT gate can be constructed in five steps:

$$U_{\mathrm{CNOT}} = R_z^{(1)}(\pi/2)R_z^{(2)}(-\pi/2)U_{\mathrm{SWAP}}^{-1/2}R_z^{(1)}(\pi)U_{\mathrm{SWAP}}^{-1/2}. \tag{8.8}$$

To perform a two-qubit gate, the exchange interaction strength J can be tuned electrically via t_c or ε. The latter has so far been preferred in experiments using coupled QDs and donors [29,44] and sub-ns $\sqrt{\mathrm{SWAP}}$ operations have been achieved with a fidelity up to 90%. Gate fidelity is charge noise limited and could be improved using symmetric operation, where

exchange is tuned via t_c instead of ε, which has shown to improve noise performance [45].

Phase gate: when $\delta E_z \gg J$ dominates, $|\uparrow,\downarrow\rangle$ and $|\downarrow,\uparrow\rangle$ are eigenstates of the system and a phase of equal magnitude, but different sign is acquired over time between the states, which allows for a controlled phase gate that is achieved when the phase difference equals π (see Fig. 8.7b). The CPhase gate can be constructed in a single fast pulse by tuning J(ε) which is one of the reasons why it has been the preferred experimental implementation of a two-qubit gate in silicon so far and has been performed based on an intrinsic difference [42] in δE_z, and by using a micro-magnet [13]. With the latter technique, a rotation can be performed within 50 ns and a Bell state fidelity of $F = 89\%$ has been achieved which is a measure of the gate fidelity. More recently, a fidelity above 99.5% has been achieved [74].

Resonant gate: An alternative resonantly driven two-qubit gate for electron spins in QDs can be constructed based on the tunable exchange interaction J and spin resonance techniques (ESR/EDSR). When $J = 0$, each of the two spins can be driven unconditionally on the other spin state due to an identical energy difference as illustrated in Fig. 8.7d. For the first spin for example, transitions from $|\downarrow,\downarrow\rangle$ to $|\uparrow,\downarrow\rangle$ and $|\downarrow,\uparrow\rangle$ to $|\uparrow,\uparrow\rangle$ are driven at the same frequency. At $J > 0$, distinct drive frequencies form for each spin depending on the other spin state allowing for conditional single spin manipulation. This has been demonstrated [25,46] and is shown in Fig. 8.7e, where two different drive frequencies for qubit 1 are observed over time, whilst qubit 2 is oscillating between the up and down state. A controlled rotation with gate fidelity of up to $F = 98\%$ within 2 μs has been demonstrated. A faster gate could be performed when combined with EDSR and stronger micro-magnets [24,46].

3.2 Singlet-triplet qubit

Instead of encoding quantum information in a single spin, the singlet-triplet (S-T_0) qubit makes use of the $S_z = 0$ sub-space of the combined two electron spin states, where the symmetric triplets (T_+, T_-) are split off by an external magnetic field. This is appealing because it allows for single-qubit gates via electrically controlled exchange interaction between the two spins, which drives two qubit gates in a qubit encoded in a single spin as discussed above [29]. Moreover, the $S_z = 0$ sub-space is protected from overall magnetic field fluctuations which has led to improved coherence times especially in early GaAs-based implementations. By transforming Eq. (8.6) into

a singlet-triplet basis and making similar approximations in the regime $U-\varepsilon \gg t_c$ and neglecting higher energy states, the (1,1) singlet and triplet Hamiltonian is given by [42].

$$H_{ST_0} = \begin{pmatrix} J(\varepsilon) & \Delta E_z \\ \Delta E_z & 0 \end{pmatrix}. \tag{8.9}$$

Here $J(\varepsilon) = \frac{1}{2}(U-\varepsilon) + \sqrt{2t_c^2 + \frac{1}{4}(U-\varepsilon)^2}$ is the exchange interaction and $\Delta E_z = \delta E_z \cos\theta$ is the local Zeeman energy difference with $\theta(\varepsilon, t_c) = \arctan\left(\frac{2\sqrt{2}t_c}{U-\varepsilon-\sqrt{8t_c^2+(U-\varepsilon)^2}}\right)$ being an angle introduced by the transformation. In this system, the exchange interaction generates a phase, whilst the local Zeeman difference drives oscillations between the S and T_0 state. Another advantage of the S-T_0 qubit is the possibility of electrostatic coupling between two S-T_0 qubits due to the difference in orbital wavefunction between a singlet and a triplet. However, such a gate is sensitive to charge noise and has not shown high fidelity so far [47]. Similar to a spin 1/2 qubit, S-T_0 qubits can additionally be coupled via the exchange interaction. Fast initialisation into a singlet state is achieved through Pauli spin blockade (at $\varepsilon > 0$) and a controlled Zeeman difference δE_z for high-fidelity single qubit gates can be achieved via dynamical nuclear spin polarisation or by using a micro-magnet. The two-electron spin system of a donor coupled to a QD provides an excellent S-T_0 qubit due to the hyperfine field generated by the nuclear spin leading to a well-defined local field difference between the donor and dot electron spins. For a phosphorus donor-dot S-T_0 qubit, manipulation at 57 MHz is demonstrated [48]. Detailed methods for the operation of S-T_0 qubits can be found in Ref. [49].

3.3 Spin readout

In the previous section, different ways of forming and controlling spin qubits in silicon quantum devices were discussed. This section focusses on the mechanisms for readout of the charge and spin of a single electron in a QD or dopant device. The magnetic moment of an electron spin, which is proportional to the Bohr magneton $\mu_B = 9.274 \times 10^{-24}$ J/T, is very weak. For an electron spin, 1/2 the energy difference between spin up and down is $\sim$100 μeV in a field of 1 T. The key to readout of single spins

in nanostructures such as QDs is to convert from the spin degree of freedom to the charge degree of freedom. The energy related to a single charge in nanostructures is typically a few meV, with $e = 1.6 \times 10^{-19}$C, and movement of fractions of single charges in nanostructures can be detected through highly sensitive charge sensors.

3.3.1 Charge sensors

Very precise charge sensors capable of detecting single electron tunnelling can be formed in nanostructures that exhibit forms of quantised transport. Such sensors are either formed in one-dimensional constrictions (quantum point contact, QPC) or using charge islands, such as SETs and QDs, and are placed in close proximity to the qubit device as shown in Fig. 8.6a. When operated at the point of maximum transconductance, such devices realise very sensitive sensors to changes in the local charge distribution. Movements of single electrons on and off a nearby QD can be detected for sufficiently large capacitive coupling between the sensor and QD via changes in the current measured through the sensor [15]. Sensitivity and bandwidth can be improved by using radio frequency (rf) reflectometry techniques [50].

3.3.2 Spin-to-charge conversion

Energy or spin-selective tunnelling allows readout of the electron spin in a single QD by coupling to a reservoir as illustrated in Fig. 8.8a. The discrete spin-up and down states in the QD are split using a magnetic field B_z into $\mu_\uparrow$ and $\mu_\downarrow$ and for readout the edge of the continuous states in the reservoir, the Fermi level μ_{res}, is aligned in the centre between the up and down state. As a result of this, an electron in the spin-up state prefers to tunnel onto the reservoir, followed by a spin-down electron tunnelling back into the QD. On the contrary, an initial spin down electron remains in the QD [40,51]. Thus, for spin-selective tunnelling, the QD is pulsed from the region of stable charge configuration close to the dot-to-reservoir transition as indicated by the arrows at the (1,0) and (0,1) configuration of a double QD in Fig. 8.8c. Following the pulse, a change in the signal is observed when the spin-up electron tunnels out (after time t_{out}) of the QD, which leaves the QD empty, and remains until a spin down electron tunnels back in (t_{in}) as shown in Fig. 8.8d.

In a double QD, Pauli spin blockade allows for spin-to-charge conversion without an external reservoir, which distinguishes between the spin singlet and triplet states. This allows for readout of a single spin via a well-defined

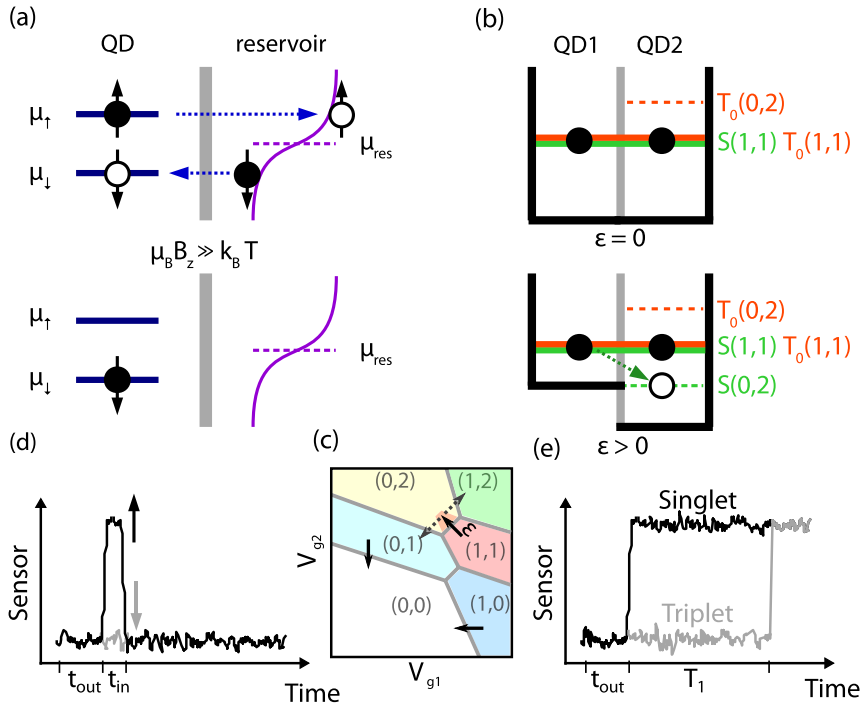


Figure 8.8 ***Spin to charge conversion.*** (a) Spin selective tunnelling of Zeeman energy split spin-up and down states in a QD to a reservoir. (b) Singlet-triplet Pauli spin blockade in a tunnel-coupled double QD. (c) Stability diagram of a double QD with charge occupation indicated by numbers and colouring. *Arrows* indicate pulses related to reservoir and spin blockade-based readout. Charge sensor signal for spin up and down in a spin-selective tunnelling (d) and singlet-triplet spin blockade measurement (e). *(a) Based on Ref. [52].*

ancillary spin [53] or direct readout of an S-T_0. This method is illustrated in Fig. 8.8b. At zero de-tuning ($\varepsilon = 0$), there is a small splitting between the $S(1,1)$ and $T_0(1,1)$ state and the $S(1,1)$ and $S(0,2)$ charge states are degenerate, whilst the $T_0(0,2)$ is split off due to Pauli exclusion and $T_{+,-}$ states are split off (not shown) due to an external magnetic field. At positive de-tuning ($\varepsilon > 0$) the (0,2), charge state becomes energetically favourable, and the $S(1,1)$ state transitions to $S(0,2)$ via elastic tunnelling (after time t_{out}),

Whilst the T_0 (1,1) state remains blocked until lifted via spin relaxation (after time T_1). The de-tuning axis ε and the region of spin blockade is indicated in Fig. 8.8c. The change in charge configuration of the singlet is detected by a nearby charge sensor as shown in Fig. 8.8e, allowing for singlet-triplet readout, when the sensor is well aligned with the double

QD charge dipole as the overall number of charges in the double QD does not change. When a double QD is additionally coupled to a reservoir, the S(0,2) state can be selectively mapped to the (1,0) charge configuration (direct enhanced latching readout) or the T_0(1,1) can be mapped to the (1,2) charge configuration (reverse enhanced latching readout) [54] as indicated by dashed arrows in Fig. 8.8c. This enhanced latching mechanism improves the signal, due to state selective change of the double QD occupation by a whole electron, and extends signal lifetime by changing from spin dominated relaxation to a meta-stable charge state.

To maximise readout fidelity in reservoir-based readout, the Zeeman splitting ($\sim$0.1 meV/T) should be much larger than the thermal energy ($\sim$0.09 meV/K). For spin-blockade-based readout, the singlet-triplet (or valley) splitting is the energy scale limiting fidelity due to relaxation. A singlet-triplet splitting on the order of 1 meV is observed [55] even at low fields. Moreover, singlet-triplet readout forms the basis for parity measurements required in error correction schemes [56].

3.3.3 Gate-based dispersive readout

Charge sensors need to be placed within a few hundreds of nanometers of qubits in order to achieve sufficient capacitive coupling. This limits the number of qubits that can be read out using a single charge sensor and leads to increased architectural complexity. Gate-based rf reflectometry is a resonant technique that turns the metal gates defining silicon qubit structures itself into a sensor by embedding into a resonant circuit. As such, additional structures to form charge sensors become unnecessary. Additionally, operation at radio or even close to microwave frequencies offer large readout bandwidth and high sensitivity [57].

In gate-based readout, a weak rf signal is applied to the QD gate which drives the resonant circuit close to resonance and allows to probe the ability of the electron with given spin to tunnel in and out of the QD based on a parametric contribution to the device admittance [58,59]. Typically, changes in the device capacitance are observed known as tunnelling and quantum capacitance [60,61]. When driven close to resonance, small changes in the device capacitance lead to a change in the phase of the rf signal reflected from the resonator. The highest sensitivity for gate-based sensing has been reported [62,63] in nanowire transistor devices as introduced in Section 1.4, which have large gate coupling due to a corner geometry. A schematic showing a cross-section of the nanowire in the region of the split gate (shown in orange) is shown in Fig. 8.9a where one of the gates is connected to a resonator and driven at an rf amplitude for readout.

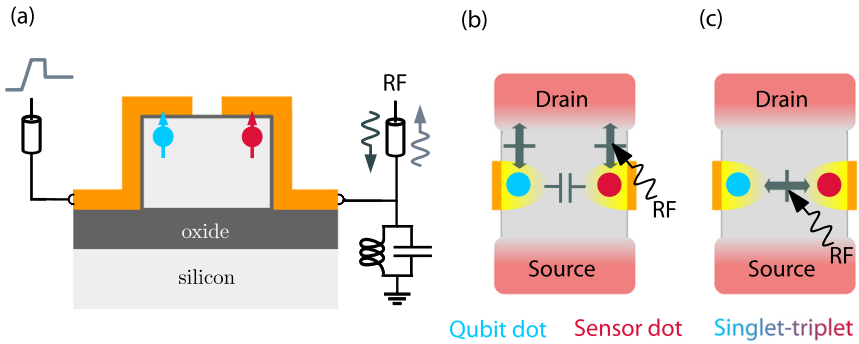


Figure 8.9 ***Gate-based readout.*** (a) Cross-section of a split-gate nanowire double QD device. One gate is connected to a pulse line and the second gate is connected to a resonator for gate-based rf readout. (b) For capacitive inter-dot coupling, the rf dot can act as a charge sensor for spin readout via spin dependent tunnelling when coupled to a reservoir. (c) When both dots are tunnel coupled, spin readout can be performed based on Pauli spin blockade without readout reservoirs.

As illustrated in Fig. 8.9b and c, any QD within a multi-dot structure can be turned into a sensor using a gate-based approach, either through capacitive coupling to a qubit and using charge sensing techniques (sensor dot), or through tunnel coupling to a qubit and using Pauli spin blockade (ancillary dot). In the former case, readout is performed via spin-dependent tunnelling to a reservoir. In the latter, inter-dot tunnelling and Pauli spin blockade removes the need for a readout reservoir. Similarly, direct readout of a singlet-triplet qubit can be achieved. Single-shot spin readout using these techniques has been demonstrated with a fidelity up to 98% within 6 μs integration time [64] in the singlet-triplet basis and up to 99% within 1 ms integration time using a sensor dot [65]. Further improvements in readout fidelity and speed can be achieved using Josephson parametric amplification and readout circuit optimisation [62,63,79]. Recently, gate-based readout of multiple QDs using a single resonant circuit in a dynamic random access scheme has been demonstrated providing the basis for a compact and scalable readout scheme of many qubit devices [12].

4. Future developments

Excellent spin qubits have been demonstrated in silicon devices with long coherence times and high-fidelity control [24,28]. Moreover, a programmable two-qubit processor has been realised [13] performing first quantum algorithms on a silicon device. Recently, demonstrations have

been extended to a three electron spin qubit device and four hole spin processor [77,78]. Current efforts focus on improving two-qubit gate fidelity and readout fidelity which are now above the threshold for fault-tolerant quantum computation [56,74,80]. Whilst complex devices with many QDs have been fabricated [19] and multiple architectures for a scalable silicon spin-based quantum computer have emerged [66–68], operation of multiple coupled spin qubits in devices that follow designs capable of scaling to many qubits are still to be demonstrated. The small size and short range interactions of spin qubits in silicon allow for a high qubit density but also set extraordinary requirements on gate density, gate fan-out and interconnects. Alternative approaches for longer range coupling which includes mediated exchange [69], electron shuttling [68] or microwave resonators [70] providing a larger inter-qubit spacing are being explored. Additionally, cryogenic electronics that operate close to qubit devices for efficient signal generation, routing and processing have become a promising area of research [71,73,81–83]. Overall, these developments highlight the excellent prospects of silicon spin qubit devices for realisation of large-scale quantum information processing.

Acknowledgements

The authors would like to thank S. Barraud, W. Huang, A. Dzurak, T. Watson and T. Meunier for providing important material for figures within this chapter.

References

[1] Waldrop MM. The chips are down for Moore's law. Nature 2016;530(7589):144–7. https://doi.org/10.1038/530144a.

[2] Montanaro A. Quantum algorithms: an overview. npj Quantum Inf 2016;2(1):15023. https://doi.org/10.1038/npjqi.2015.23. arXiv:1511.04206.

[3] Nielsen MA, Chuang IL. Quantum computation and quantum information. Cambridge: Cambridge University Press; 2010. https://doi.org/10.1017/CBO9780511976667.

[4] DiVincenzo DP. The physical implementation of quantum computation. Fortschr Phys 2000;48(9–11):771–83. https://onlinelibrary.wiley.com/doi/10.1002/1521-3978(200009)48:9/11%3C771::AID-PROP771%3E3.0.CO;2-E. arXiv:0002077.

[5] Ladd TD, Jelezko F, Laflamme R, Nakamura Y, Monroe C, O'Brien JL. Quantum computers. Nature 2010;464(7285):45–53. https://doi.org/10.1038/nature08812. arXiv:1009.2267.

[6] Devoret MH, Schoelkopf RJ. Superconducting circuits for quantum information: an outlook. Science 2013;339(6124):1169–74. https://doi.org/10.1126/science.1231930.

[7] Arute F, Arya K, Babbush R, Bacon D, Bardin JC, Barends R, Biswas R, Boixo S, Brandao FGSL, Buell DA, Burkett B, Chen Y, Chen Z, Chiaro B, Collins R, Courtney W, Dunsworth A, Farhi E, Foxen B, Fowler A, Gidney C, Giustina M, Graff R, Guerin K, Habegger S, Harrigan MP, Hartmann MJ, Ho A, Hoffmann M, Huang T, Humble TS, Isakov SV, Jeffrey E, Jiang Z, Kafri D, Kechedzhi K,

Kelly J, Klimov PV, Knysh S, Korotkov A, Kostritsa F, Landhuis D, Lindmark M, Lucero E, Lyakh D, Mandrà S, McClean JR, McEwen M, Megrant A, Mi X, Michielsen K, Mohseni M, Mutus J, Naaman O, Neeley M, Neill C, Niu MY, Ostby E, Petukhov A, Platt JC, Quintana C, Rieffel EG, Roushan P, Rubin NC, Sank D, Satzinger KJ, Smelyanskiy V, Sung KJ, Trevithick MD, Vainsencher A, Villalonga B, White T, Yao ZJ, Yeh P, Zalcman A, Neven H, Martinis JM. Quantum supremacy using a programmable superconducting processor. Nature 2019;574(7779): 505–10. https://doi.org/10.1038/s41586-019-1666-5.

[8] Wright K, Beck KM, Debnath S, Amini JM, Nam Y, Grzesiak N, Chen J-S, Pisenti NC, Chmielewski M, Collins C, Hudek KM, Mizrahi J, Wong-Campos JD, Allen S, Apisdorf J, Solomon P, Williams M, Ducore AM, Blinov A, Kreikemeier SM, Chaplin V, Keesan M, Monroe C, Kim J. Benchmarking an 11-qubit quantum computer. Nat Commun 2019;10(1):5464. https://doi.org/10.1038/s41467-019-13534-2. arXiv:1903. 08181.

[9] Schreiber LR, Bluhm H. Quantum computation: silicon comes back. Nat Nanotechnol 2014;9(12):966–8. http://www.nature.com/doifinder/10.1038/nnano.2014.249.

[10] Morton JJL, McCamey DR, Eriksson Ma, Lyon Sa. Embracing the quantum limit in silicon computing. Nature 2011;479(7373):345–53. https://doi.org/10.1038/nature10681.

[11] Yang CH, Leon RCC, Hwang JCC, Saraiva A, Tanttu T, Huang W, Camirand Lemyre J, Chan KW, Tan KY, Hudson FE, Itoh KM, Morello A, Pioro-Ladrière M, Laucht A, Dzurak AS. Operation of a silicon quantum processor unit cell above one kelvin. Nature 2020;580(7803):350–4. https://doi.org/10.1038/s41586-020-2171-6. arXiv:1902.09126.

[12] Schaal S, Rossi A, Ciriano-Tejel VN, Yang T-Y, Barraud S, Morton JJL, Gonzalez-Zalba MF. A CMOS dynamic random access architecture for radio-frequency readout of quantum devices. Nat Electron 2019;2(6):236–42. https://doi.org/10.1038/s41928-019-0259-5.

[13] Watson TF, Philips SGJ, Kawakami E, Ward DR, Scarlino P, Veldhorst M, Savage DE, Lagally MG, Friesen M, Coppersmith SN, Eriksson MA, Vandersypen LMK. A programmable two-qubit quantum processor in silicon. Nature 2018;555(7698):633–7. https://doi.org/10.1038/nature25766. arXiv:1708.04214.

[14] Zwanenburg FA, Dzurak AS, Morello A, Simmons MY, Hollenberg LCL, Klimeck G, Rogge S, Coppersmith SN, Eriksson MA. Silicon quantum electronics. Rev Mod Phys 2013;85(3):961–1019. arXiv:arXiv:1206.5202v1, http://link.aps.org/doi/10.1103/RevModPhys.85.961.

[15] Hanson R, Kouwenhoven LP, Petta JR, Tarucha S, Vandersypen LMK. Spins in few-electron quantum dots. Rev Mod Phys 2007;79(4):1217–65. https://doi.org/10.1103/RevModPhys.79.1217. arXiv:0610433.

[16] Davies JH. The physics of low-dimensional semiconductors. Cambridge: Cambridge University Press; 1997. https://doi.org/10.1017/CBO9780511819070.

[17] Jehl X, Niquet Y-M, Sanquer M. Single donor electronics and quantum functionalities with advanced CMOS technology. J Phys Condens Matter 2016;28(10):103001. https://doi.org/10.1088/0953-8984/28/10/103001.

[18] Barraud S, Lavieville R, Hutin L, Bohuslavskyi H, Vinet M, Corna A, Clapera P, Sanquer M, Jehl X. Development of a CMOS route for electron pumps to be used in quantum metrology. Technologies 2016;4(1):10. https://doi.org/10.3390/technologies4010010.

[19] Zajac DM, Hazard TM, Mi X, Nielsen E, Petta JR. Scalable gate architecture for a one-dimensional array of semiconductor spin qubits. Phys Rev Appl 2016;6(5): 054013. https://doi.org/10.1103/PhysRevApplied.6.054013. arXiv:1607.07025.

[20] Rossi A, Tanttu T, Hudson FE, Sun Y, Möttönen M, Dzurak AS. Silicon metal-oxide-semiconductor quantum dots for single-electron pumping. JoVE 2015;(100):1–11. https://doi.org/10.3791/52852.
[21] Lim WH, Yang CH, Zwanenburg Fa, Dzurak aS. Spin filling of valley-orbit states in a silicon quantum dot. Nanotechnology 2011;22(33):335704. https://doi.org/10.1088/0957-4484/22/33/335704. arXiv:1103.2895.
[22] Yang CH, Rossi A, Ruskov R, Lai NS, Mohiyaddin FA, Lee S, Tahan C, Klimeck G, Morello A, Dzurak AS. Spin-valley lifetimes in a silicon quantum dot with tunable valley splitting. Nat Commun 2013;4(May):1–8. https://doi.org/10.1038/ncomms3069. arXiv:1302.0983.
[23] Muhonen JT, Dehollain JP, Laucht A, Hudson FE, Kalra R, Sekiguchi T, Itoh KM, Jamieson DN, McCallum JC, Dzurak AS, Morello A. Storing quantum information for 30 seconds in a nanoelectronic device. Nat Nanotechnol 2014;9(12):986–91. https://doi.org/10.1038/nnano.2014.211. arXiv:1402.7140.
[24] Yoneda J, Takeda K, Otsuka T, Nakajima T, Delbecq MR, Allison G, Honda T, Kodera T, Oda S, Hoshi Y, Usami N, Itoh KM, Tarucha S. A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%. Nat Nanotechnol 2018;13(2):102–6. https://doi.org/10.1038/s41565-017-0014-x. arXiv:1708.01454.
[25] Huang W, Yang CH, Chan KW, Tanttu T, Hensen B, Leon RCC, Fogarty MA, Hwang JCC, Hudson FE, Itoh KM, Morello A, Laucht A, Dzurak AS. Fidelity benchmarks for two-qubit gates in silicon. Nature 2019;569(7757):532–6. https://doi.org/10.1038/s41586-019-1197-0. arXiv:1805.05027.
[26] Angus SJ, Ferguson AJ, Dzurak AS, Clark RG. Gate-defined quantum dots in intrinsic silicon. Nano Lett 2007;7(7):2051–5. https://doi.org/10.1021/nl070949k.
[27] Lim WH, Zwanenburg FA, Huebl H, Möttönen M, Chan KW, Morello A, Dzurak AS. Observation of the single-electron regime in a highly tunable silicon quantum dot. Appl Phys Lett 2009;95(24):242102. https://doi.org/10.1063/1.3272858. arXiv:0910.0576.
[28] Veldhorst M, Hwang JCC, Yang CH, Leenstra aW, de Ronde B, Dehollain JP, Muhonen JT, Hudson FE, Itoh KM, Morello A, Dzurak AS. An addressable quantum dot qubit with fault-tolerant control-fidelity. Nat Nanotechnol 2014;9(12):981–5. https://doi.org/10.1038/nnano.2014.216. arXiv: 1407.1950v1.
[29] Maune BM, Borselli MG, Huang B, Ladd TD, Deelman PW, Holabird KS, Kiselev AA, Alvarado-Rodriguez I, Ross RS, Schmitz AE, Sokolich M, Watson CA, Gyure MF, Hunter AT. Coherent singlet-triplet oscillations in a silicon-based double quantum dot. Nature 2012;481(7381):344–7. https://doi.org/10.1038/nature10707.
[30] Scarlino P. Spin and valley physics in a Si/SiGe quantum dot [Ph.D. thesis]. TU Delft; 2016.
[31] Kawakami E, Scarlino P, Ward DR, Braakman FR, Savage DE, Lagally MG, Friesen M, Coppersmith SN, Eriksson MA, Vandersypen LMK. Electrical control of a long-lived spin qubit in a Si/SiGe quantum dot. Nat Nanotechnol 2014;9(9):666. https://doi.org/10.1038/nnano.2014.153. arXiv:1404.5402.
[32] Voisin B, Nguyen V-H, Renard J, Jehl X, Barraud S, Triozon F, Vinet M, Duchemin I, Niquet Y-M, de Franceschi S, Sanquer M. Few-electron edge-state quantum dots in a silicon nanowire field-effect transistor. Nano Lett 2014;14(4):2094–8. https://doi.org/10.1021/nl500299h.
[33] Loss D, DiVincenzo DP. Quantum computation with quantum dots. Phys Rev 1998;57(1):120–6. https://doi.org/10.1103/PhysRevA.57.120. arXiv:9701055.
[34] Kane BE. A silicon-based nuclear spin quantum computer. Nature 1998;393(6681):133–7. https://doi.org/10.1038/30156. http://www.nature.com/articles/30156.

[35] Crippa A, Maurand R, Bourdet L, Kotekar-Patil D, Amisse A, Jehl X, Sanquer M, Laviéville R, Bohuslavskyi H, Hutin L, Barraud S, Vinet M, Niquet Y-M, De Franceschi S. Electrical spin driving by g-matrix modulation in spin-orbit qubits. Phys Rev Lett 2018;120(13):137702. https://doi.org/10.1103/PhysRevLett.120.137702. arXiv:1710.08690.

[36] Golovach VN, Borhani M, Loss D. Electric-dipole-induced spin resonance in quantum dots. Phys Rev B 2006;74(16):165319. https://doi.org/10.1103/PhysRevB.74.165319. arXiv:0601674.

[37] Nowack KC, Koppens FHL, Nazarov YV, Vandersypen LMK. Coherent control of a single electron spin with electric fields. Science 2007;318(5855):1430–3. https://doi.org/10.1126/science.1148092.

[38] Pioro-Ladrière M, Obata T, Tokura Y, Shin Y-S, Kubo T, Yoshida K, Taniyama T, Tarucha S. Electrically driven single-electron spin resonance in a slanting Zeeman field. Nat Phys 2008;4(10):776–9. https://doi.org/10.1038/nphys1053. arXiv:0805.1083.

[39] Takeda K, Kamioka J, Otsuka T, Yoneda J, Nakajima T, Delbecq MR, Amaha S, Allison G, Kodera T, Oda S, Tarucha S. A fault-tolerant addressable spin qubit in a natural silicon quantum dot. Sci Adv 2016;2(8):e1600694. https://doi.org/10.1126/sciadv.1600694. arXiv:1602.07833.

[40] Elzerman JM, Hanson R, Willems van Beveren LH, Witkamp B, Vandersypen LMK, Kouwenhoven LP. Single-shot read-out of an individual electron spin in a quantum dot. Nature 2004;430(6998):431–5. https://doi.org/10.1038/nature02693. arXiv:0411232.

[41] Yang CH, Chan KW, Harper R, Huang W, Evans T, Hwang JCC, Hensen B, Laucht A, Tanttu T, Hudson FE, Flammia ST, Itoh KM, Morello A, Bartlett SD, Dzurak AS. Silicon qubit fidelities approaching incoherent noise limits via pulse engineering. Nat Electron 2019;2(4):151–8. https://doi.org/10.1038/s41928-019-0234-1. arXiv:1807.09500.

[42] Veldhorst M, Yang CH, Hwang JCC, Huang W, Dehollain JP, Muhonen JT, Simmons S, Laucht A, Hudson FE, Itoh KM, Morello A, Dzurak AS. A two-qubit logic gate in silicon. Nature 2015;526(7573):410–4. https://doi.org/10.1038/nature15263. arXiv:1411.5760.

[43] Meunier T, Calado VE, Vandersypen LMK. Efficient controlled-phase gate for single-spin qubits in quantum dots. Phys Rev B 2011;83(12):121403. https://doi.org/10.1103/PhysRevB.83.121403.

[44] He Y, Gorman SK, Keith D, Kranz L, Keizer JG, Simmons MY. A two-qubit gate between phosphorus donor electrons in silicon. Nature 2019;571(7765):371–5. https://doi.org/10.1038/s41586-019-1381-2.

[45] Martins F, Malinowski FK, Nissen PD, Barnes E, Fallahi S, Gardner GC, Manfra MJ, Marcus CM, Kuemmeth F. Noise suppression using symmetric exchange gates in spin qubits. Phys Rev Lett 2016;116(11):116801. https://doi.org/10.1103/PhysRevLett.116.116801. arXiv:1511.07336.

[46] Zajac DM, Sigillito AJ, Russ M, Borjans F, Taylor JM, Burkard G, Petta JR. Resonantly driven CNOT gate for electron spins. Science 2018;359(6374):439–42. https://doi.org/10.1126/science.aao5965. arXiv:1708.03530.

[47] Shulman MD, Dial OE, Harvey SP, Bluhm H, Umansky V, Yacoby a. Demonstration of entanglement of electrostatically coupled singlet-triplet qubits. Science 2012;336(6078):202–5. https://doi.org/10.1126/science.1217692. arXiv:1202.1828.

[48] Harvey-Collard P, Jacobson NT, Rudolph M, Dominguez J, Ten Eyck GA, Wendt JR, Pluym T, Gamble JK, Lilly MP, Pioro-Ladrière M, Carroll MS. Coherent coupling between a quantum dot and a donor in silicon. Nat Commun 2017;8(1):1029. https://doi.org/10.1038/s41467-017-01113-2. arXiv:1512.01606.

[49] Botzem T, Shulman MD, Foletti S, Harvey SP, Dial OE, Bethke P, Cerfontaine P, McNeil RPG, Mahalu D, Umansky V, Ludwig A, Wieck A, Schuh D, Bougeard D, Yacoby A, Bluhm H. Tuning methods for semiconductor spin qubits. Phys Rev Appl 2018;10(5):054026. https://doi.org/10.1103/PhysRevApplied.10.054026. arXiv: 1801.03755.

[50] Reilly DJ, Marcus CM, Hanson MP, Gossard AC. Fast single-charge sensing with a rf quantum point contact. Appl Phys Lett 2007;91(16):162101. https://doi.org/10.1063/1.2794995. arXiv:0707.2946.

[51] Morello A, Pla JJ, Zwanenburg Fa, Chan KW, Tan KY, Huebl H, Möttönen M, Nugroho CD, Yang C, van Donkelaar Ja, Alves ADC, Jamieson DN, Escott CC, Hollenberg LCL, Clark RG, Dzurak AS. Single-shot readout of an electron spin in silicon. Nature 2010;467(7316):687–91. https://doi.org/10.1038/nature09392. arXiv:1003.2679.

[52] Pla JJ, Tan KY, Dehollain JP, Lim WH, Morton JJL, Jamieson DN, Dzurak AS, Morello A. A single-atom electron spin qubit in silicon. Nature 2012;489(7417): 541–5. https://doi.org/10.1038/nature11449. arXiv:1305.4481.

[53] Jones C, Fogarty MA, Morello A, Gyure MF, Dzurak AS, Ladd TD. Logical qubit in a linear array of semiconductor quantum dots. Phys Rev X 2018;8(2):021058. https://doi.org/10.1103/PhysRevX.8.021058. arXiv:1608.06335.

[54] Harvey-Collard P, D'Anjou B, Rudolph M, Jacobson NT, Dominguez J, Ten Eyck GA, Wendt JR, Pluym T, Lilly MP, Coish WA, Pioro-Ladrière M, Carroll MS. High-fidelity single-shot readout for a spin qubit via an enhanced latching mechanism. Phys Rev X 2018;8(2):021046. https://doi.org/10.1103/PhysRevX.8.021046. arXiv:1703.02651.

[55] Lai NS, Lim WH, Yang CH, Zwanenburg FA, Coish WA, Qassemi F, Morello A, Dzurak AS. Pauli spin blockade in a highly tunable silicon double quantum dot. Sci Rep 2011;1(1):110. https://doi.org/10.1038/srep00110. arXiv:1012.1410.

[56] Fowler AG, Mariantoni M, Martinis JM, Cleland AN. Surface codes: towards practical large-scale quantum computation. Phys Rev 2012;86(3):032324. https://doi.org/10.1103/PhysRevA.86.032324. arXiv:1208.0928.

[57] Gonzalez-Zalba MF, Barraud S, Ferguson aJ, Betz aC. Probing the limits of gate-based charge sensing. Nat Commun 2015;6(1):6084. https://doi.org/10.1038/ncomms7084.

[58] Petersson KD, Smith CG, Anderson D, Atkinson P, Jones GAC, Ritchie DA. Charge and spin state readout of a double quantum dot coupled to a resonator. Nano Lett 2010;10(8):2789–93. https://doi.org/10.1021/nl100663w. arXiv:1004.4047.

[59] House MG, Kobayashi T, Weber B, Hile SJ, Watson TF, van der Heijden J, Rogge S, Simmons MY. Radio frequency measurements of tunnel couplings and singlet-triplet spin states in Si:P quantum dots. Nat Commun 2015;6:8848. https://doi.org/10.1038/ncomms9848.

[60] Mizuta R, Otxoa RM, Betz AC, Gonzalez-Zalba MF. Quantum and tunneling capacitance in charge and spin qubits. Phys Rev B 2017;95(4):045414. https://doi.org/10.1103/PhysRevB.95.045414. arXiv:1604.02884.

[61] Esterli M, Otxoa RM, Gonzalez-Zalba MF. Small-signal equivalent circuit for double quantum dots at low-frequencies. Appl Phys Lett 2019;114(25):253505. https://doi.org/10.1063/1.5098889. arXiv:1812.06056.

[62] Ahmed I, Haigh JA, Schaal S, Barraud S, Zhu Y, Lee C-m, Amado M, Robinson JWA, Rossi A, Morton JJL, Gonzalez-Zalba MF. Radio-frequency capacitive gate-based sensing. Phys Rev Appl 2018;10(1):014018. https://doi.org/10.1103/PhysRevApplied.10.014018. arXiv:1801.09759.

[63] Schaal S, Ahmed I, Haigh JA, Hutin L, Bertrand B, Barraud S, Vinet M, Lee C-M, Stelmashenko N, Robinson JWA, Qiu JY, Hacohen-Gourgy S, Siddiqi I,

Gonzalez-Zalba MF, Morton JJL. Fast gate-based readout of silicon quantum dots using Josephson parametric amplification. Phys Rev Lett 2020;124(6):067701. https://doi.org/10.1103/PhysRevLett.124.067701. arXiv:1907.09429.

[64] Zheng G, Samkharadze N, Noordam ML, Kalhor N, Brousse D, Sammak A, Scappucci G, Vandersypen LMK. Rapid gate-based spin read-out in silicon using an on-chip resonator. Nat Nanotechnol 2019;14(8):742–6. arXiv:1901.00687, http://www.nature.com/articles/s41565-019-0488-9.

[65] Urdampilleta M, Niegemann DJ, Chanrion E, Jadot B, Spence C, Mortemousque P-A, Bäuerle C, Hutin L, Bertrand B, Barraud S, Maurand R, Sanquer M, Jehl X, De Franceschi S, Vinet M, Meunier T. Gate-based high fidelity spin readout in a CMOS device. Nat Nanotechnol 2019;14(8):737–41. https://doi.org/10.1038/s41565-019-0443-9. arXiv:1809.04584.

[66] Veldhorst M, Eenink HGJ, Yang CH, Dzurak AS. Silicon CMOS architecture for a spin-based quantum computer. Nat Commun 2017;8(1):1766. https://doi.org/10.1038/s41467-017-01905-6. arXiv:1609.09700.

[67] Li R, Petit L, Franke DP, Dehollain JP, Helsen J, Steudtner M, Thomas NK, Yoscovits ZR, Singh KJ, Wehner S, Vandersypen LMK, Clarke JS, Veldhorst M. A crossbar network for silicon quantum dot qubits. Sci Adv 2018;4(7). https://doi.org/10.1126/sciadv.aar3960. eaar3960. arXiv:1711.03807.

[68] Vandersypen LMK, Bluhm H, Clarke JS, Dzurak AS, Ishihara R, Morello A, Reilly DJ, Schreiber LR, Veldhorst M. Interfacing spin qubits in quantum dots and donors—hot, dense, and coherent. npj Quantum Inf 2017;3(1):34. https://doi.org/10.1038/s41534-017-0038-y. arXiv:1612.05936.

[69] Srinivasa V, Xu H, Taylor JM. Tunable spin-qubit coupling mediated by a multielectron quantum dot. Phys Rev Lett 2015;114(22):226803. https://doi.org/10.1103/PhysRevLett.114.226803.

[70] Mi X, Benito M, Putz S, Zajac DM, Taylor JM, Burkard G, Petta JR. A coherent spin-photon interface in silicon. Nature 2018;555(7698):599–603. https://doi.org/10.1038/nature25769. arXiv:1710.03265.

[71] van Dijk JPG, Charbon E, Sebastiano F. The electronic interface for quantum processors. Microprocess Microsyst 2018;66:90–101. https://doi.org/10.1016/j.micpro.2019.02.004. arXiv:1811.01693.

[72] Maurand R, Jehl X, Kotekar-Patil D, Corna A, Bohuslavskyi H, Laviéville R, Hutin L, Barraud S, Vinet M, Sanquer M, De Franceschi S. A CMOS silicon spin qubit. Nat Commun 2016;7:13575. https://doi.org/10.1038/ncomms13575. http://www.nature.com/articles/ncomms13575.

[73] Ruffino A, Yang T-Y, Michniewicz J, Peng Y, Charbon E, Gonzalez-Zalba MF. Integrated multiplexed microwave readout of silicon quantum dots in a cryogenic CMOS chip. 2021:1–14. arXiv 2101.08295, http://arxiv.org/abs/2101.08295.

[74] Xue X, Russ M, Samkharadze N, Undseth B, Sammak A, Scappucci G, Vandersypen LMK. Computing with spin qubits at the surface code error threshold 2021:1–19. arXiv:2107.00628, http://arxiv.org/abs/2107.00628.

[75] Gonzalez-Zalba MF, de Franceschi S, Charbon E, Meunier T, Vinet M, Dzurak AS. Scaling silicon-based quantum computing using CMOS technology: state-of-the-art, challenges and perspectives. 2020. arXiv:2011.11753, http://arxiv.org/abs/2011.11753.

[76] Ciriano-Tejel VN, Fogarty MA, Schaal S, Hutin L, Bertrand B, Ibberson L, Gonzalez-Zalba MF, Li J, Niquet Y-M, Vinet M, Morton JJL. Spin readout of a CMOS quantum dot by gate reflectometry and spin-dependent tunneling. PRX Quantum 2021;2(1):010353. https://doi.org/10.1103/PRXQuantum.2.010353.

[77] Takeda K, Noiri A, Nakajima T, Yoneda J, Kobayashi T, Tarucha S. Quantum tomography of an entangled three-qubit state in silicon. Nat Nanotechnol 2021;16: 965–9. https://doi.org/10.1038/s41565-021-00925-0.

[78] Hendrickx NW, Lawrie WIL, Russ M, van Riggelen F, de Snoo SL, Schouten RN, Sammak A, Scappucci G, Veldhorst M. A four-qubit germanium quantum processor. Nature 2021;591:580–5. https://doi.org/10.1038/s41586-021-03332-6. eprint: 2009.04268; ISSN:0028-0836, 7851.

[79] Ibberson DJ, Lundberg T, Haigh JA, Hutin L, Bertrand B, Barraud S, Lee C-M, Stelmashenko NA, Oakes GA, Cochrane L, Robinson JWA, Vinet M, Gonzalez-Zalba MF, Ibberson LA. Large dispersive interaction between a CMOS double quantum dot and microwave photons. PRX Quantum 2021;2(2):020315. https://doi.org/10.1103/PRXQuantum.2.020315.

[80] Harvey-Collard P, D'Anjou B, Rudolph M, Jacobson NT, Dominguez J, Ten Eyck GA, Wendt JR, Pluym T, Lilly MP, Coish WA, Pioro-Ladrière M, Carroll MS. High-fidelity single-shot readout for a spin qubit via an enhanced latching mechanism. Phys Rev X 2018;8(2):021046. https://doi.org/10.1103/PhysRevX.8.021046.

[81] Pauka SJ, Das K, Hornibrook JM, Gardner GC, Manfra MJ, Cassidy MC, Reilly DJ. Characterizing quantum devices at scale with custom Cryo-CMOS. Phys Rev Appl 2020;13(5):054072. https://doi.org/10.1103/PhysRevApplied.13.054072. ISSN: 2331-7019.

[82] Pauka SJ, Das K, Kalra R, Moini A, Yang Y, Trainer M, Bousquet A, Cantaloube C, Dick N, Gardner GC, Manfra MJ, Reilly DJ. A cryogenic CMOS chip for generating control signals for multiple qubits. Nat Electron 2021;4(1):64–70. https://doi.org/10.1038/s41928-020-00528-y.

[83] Xue X, Patra B, van Dijk JPG, Samkharadze N, Subramanian S, Corna A, Wuetz BP, Jeon C, Sheikh F, Juarez-Hernandez E, Perez Esparza B, Rampurawala H, Carlton B, Ravikumar S, Nieva C, Kim S, Lee H-J, Sammak A, Scappucci G, Veldhorst M, Sebastiano F, Babaie M, Pellerano S, Charbon E, Vandersypen Lieven MK. CMOS-based cryogenic control of silicon quantum circuits. Nature 2021; 593(7858):205–10. https://doi.org/10.1038/s41586-021-03469-4.

CHAPTER NINE

Electrical control of semiconductor quantum dot single photon sources

A.J. Bennett*
School of Engineering, Queen's Buildings, Cardiff University, Cardiff, Wales
*Corresponding author: BennettA19@cardiff.ac.uk

1. Introduction and motivation

Epitaxial single quantum dots are an advanced semiconductor technology that has formed a fertile test bed for new physics, in the solid state and in quantum technology, over the last 20 years. A number of excellent reviews focus on various aspects of their development [1–8]. They can be grown within high-quality crystals, displaying narrow linewidths and near unity internal quantum efficiencies. A single dot can be isolated by means of judicious growth conditions or post-processing, allowing it to emit non-classical light with sub-poissionan statistics [9], entangled pairs of photons [10], complex multi-photon entangled states [11] or quantum mechanically indistinguishable photons [12]. In addition, there are some areas where dots have a unique capability over competing technologies, such as single atoms, ions or point-like colour centers. Firstly, taking advantage of the versatility of the host semiconductor platform, their energy can be tuned by composition, growth or external fields. Secondly, trapped carriers in the dot can interact with 10^4–10^5 nuclear spins which facilitate the novel study of mesoscopic spin–nuclei interactions. Finally, semiconductor devices can be developed that allow the injection of carriers, controlled charging or tuning of emission. These electrically controlled devices are the focus of this chapter.

2. Diode designs for single quantum dot photon sources

2.1 Heterostructures for electric field control of quantum dots

Self-assembled InGaAs quantum dots are usually grown by molecular beam epitaxy or metal-oxide vapor phase epitaxy on planar substrates.

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00013-7
© 2021 Elsevier Ltd.
All rights reserved.

During epitaxy, the composition can be varied to create doped layers through the introduction of beryllium or carbon as an acceptor, or silicon as a donor impurity. Processing this material to add ohmic or Schottky contacts allows an externally applied bias to vary the electric field across the dot, parallel to the growth direction.

Conveniently, the arsenides also include an almost lattice matched alloy in the form of $Al_xGa_{1-x}As$, whose refractive index can be changed from 3.54 to 2.97 (at 900 nm) by increasing the aluminum fraction, x, from zero to unity. Alternating layers of high and low aluminum fraction alloys, each with an optical thickness of one quarter wavelength, produce defect-free Bragg mirrors that enhance light extraction in the vertical direction. The high degree of control over the alloy composition and doping in the vicinity of the dots can also be used to control the tunnelling of charge carriers into the dots, as we shall see. To maintain the stability of the growth conditions, it is not always convenient to vary Ga and Al flux between layers with different alloy compositions. In this case, a superlattice of periodic GaAs and AlAs layers a few nanometers thick may be stacked to approximate any alloy composition, by varying the relative thicknesses of the alternating layers.

Fig. 9.1a shows a schematic of the band structure of two diode heterostructures with an external bias applied between the doped layers, V_{ext}. The simplest device is a GaAs *p-i-n* doping profile with InGaAs quantum dots at

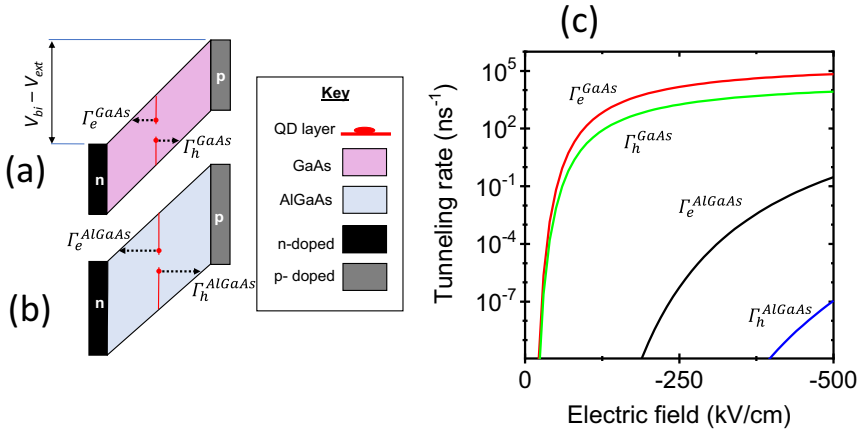


Figure 9.1 ***Tunnelling rates of electrons and holes in a* p-i-n *diode*** with a one-dimensional triangular confining potential with 3 nm dot height: (a) when the intrinsic region is GaAs. (b) when the intrinsic region is $Al_{0.75}Ga_{0.25}As$. (c) tunnelling rates as a function of applied electric field. The hole tunnelling rate is lower in each case as a result of the greater hole effective mass. *Modified from Bennett et al. [13].*

the center of the intrinsic region, as used in the first quantum LED [14]. At zero external bias, a built-in potential (V_{bi}) is dropped across the intrinsic region, which is equal to the difference in Fermi energies between the *p* and *n* regions. The energy levels within the dots can be studied in reverse bias using an external laser to photo-excite carriers within the semiconductor, some of which are captured by the dots and subsequently radiatively recombine as photoluminescence. In this case, the relative intensities of the charged and neutral states are determined by the tunnelling rates of carriers out of the dot, as illustrated by the dashed arrows in Fig. 9.1a. As a result of the lower effective mass of electrons in GaAs (0.063 m_e) compared to holes (0.52 m_e), the hole tunnelling rate is approximately an order of magnitude lower, as shown in Fig. 9.1c. To illustrate how device design can control the tunnelling rate, we show in Fig. 9.1b a *p-i-n* diode grown entirely in $Al_{0.75}Ga_{0.25}As$, in which the electron (hole) masses are increased to 0.125 m_e (0.57 m_e) and the confinement energies to 420 meV (400 meV). As a result of these changes, the tunnelling rates are reduced by several orders of magnitude at a given electric field (Fig. 9.1c). In practice, the optical quality of InGaAs quantum dots encapsulated in GaAs is superior to those in $Al_{0.75}Ga_{0.25}As$, so this second device is rarely studied; however, it does illustrate how alloy composition can be used to control tunnelling.

Fig. 9.2 presents the layer sequences and band structures of the four most used heterostructures for the control of single dots. These are the light emitting diode (a), the electron charging diode (b), the hole charging diode (c) and the energy tuning diode (d). Fig. 9.3 displays plots of photoluminescence versus electric field for typical single dots in each of these devices, each recorded at 4K with weak 850 nm laser excitation.

The first panel (a) of each figure shows the LED. Under applied reverse bias (electric field below zero), electron tunnelling from the confined states is faster, so emission from charge configurations with more holes, such as X^+, is dominant. Under an electric field more negative than a few tens of kV/cm, radiative emission is quenched, as expected from Fig. 9.1c. As the externally applied bias (V_{ext}) is increased, the band structure moves closer to the flat-band condition, where $V_{bi} - V_{ext} = 0$, current begins to flow across the device. For GaAs at 4 K, this will occur at $V_{bi} = 1.42$ V. When the current increases, carriers may be captured by the dot and radiatively recombine, which is conventionally known as electroluminescence. The linewidths of InAs dot transitions can be as low as 4.0 micro-eV at a temperature of 4 K [16]. Unequal capture rates for electrons and holes result in uncontrolled

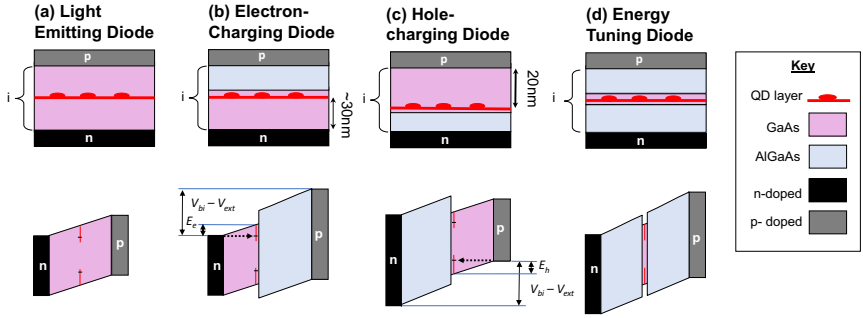


Figure 9.2 ***Device designs to control single quantum dots in diodes.*** (a) shows a *p-i-n* diode suitable for efficient injection of carriers for light emission, consisting of only GaAs and InGaAs dots, as reported by Yuan et al. [14]. (b) An electron charging device, where a high aluminium content Al barrier between the dots and the *p*-contact, inhibits hole injection into the dots, but allow bias controlled electron tunnelling, as reported by Warburton et al. [15]. The built-in (V_{bi}) and external (V_{ext}) voltages may be balanced to allow electron tunnelling into states confined by energy E_e. (c) A hole charging device, where a high aluminium content Al barrier between the dots and the *n*-contact inhibits electron injection into the dots, but allow bias controlled hole tunnelling. (d) a diode for Stark-tuning of dots, with high aluminium content AlGaAs on both sides of the active region, as reported by Bennett et al. [13].

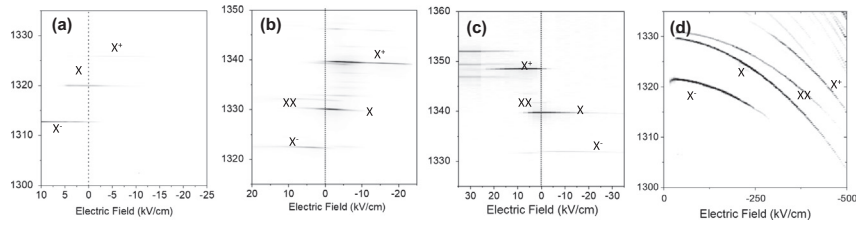


Figure 9.3 ***Photoluminescence versus electric field plots for the diodes shown in Fig. 9.2, excited with 850 nm CW diode laser.*** (a) light emitting diode, with electroluminescence dominant above 0 kV/cm. (b) electron charging diode. (c) hole charging diode. (d) Stark shift tuneable diode. *Modified from Bennett et al. [13].*

emission from several different excitonic species, for example, X^- and X, reducing the internal quantum efficiency for decay via one transition [17]. It has been shown that varying the thickness of the intrinsic region on either side of the dots balances the injection of electrons and holes. This can be favourable in certain applications, such as the emission of entangled photon pairs from the neutral exciton cascade [18]. Alternatively, a delta-doped layer of donors may be included within 10 nm of the dots to introduce an excess of one type of carrier [19].

Fig. 9.2b and c show devices that deterministically charge a dot with a finite number of electrons or holes, using compositionally unbalanced bandgaps on either side of the dots. In both cases, there is a few nanometer thick GaAs layer between the InAs dot layer and the AlGaAs barrier to maintain optical quality. For the electron (hole) charging device, there is a high aluminium content blocking layer between the dot and the p (n) contact preventing holes (electrons) tunnelling into the device. The Fermi level of the n (p) layer may be varied using V_{ext} allowing carriers to tunnel through a 30 nm thick GaAs barrier [15] into the dot. For example, increasing the magnitude of the reverse bias applied to the electron charging device from zero causes a single electron to tunnel into the dot leading to the sudden appearance of the X^- transition in the photoluminescence spectrum. The tunnelling of two electrons into the dot is energetically forbidden until the external bias is further increased, due to the Coulomb repulsion energy of the two electrons trapped in the zero-dimensional confinement potential. Similar arguments can be made for the hole tunnelling device in Fig. 9.2c. The emission spectrum as a function of external bias then reveals much information about the internal energy levels of the dot [20–22] and allows study of the coupling between a zero-dimensional confined state and the Fermi sea in the contact [23,24]. The majority of reports on these heterostructures employ Schottky contacts, for ease of fabrication [21,22], as these do not rely upon annealing or doped semiconductor contact layers, but the principle of operation is the same. The asymmetric barriers and Schottky contacts make this diode design less optimal for the generation of electroluminescence.

In Fig. 9.2a–c as the magnitude of the electric field is increased, carriers tunnel out of the dots, reducing the probability of radiative recombination and broadening the spectral width of transitions. When electric fields are applied to confined charges, they are separated in space, leading to a quadratic variation in energy, known as the Quantum Confined Stark Effect. Typically, the energy (E) of a transition can be written as $E = \varepsilon F^2 - pF + E_0$, where E_0 is the transition energy at zero field ($F = 0$). The dipole moment is p and ε is the polarisability, which describes how much field is required to separate the charges. An early surprising result was obtained by Fry et al. showing that in self-assembled InGaAs dots the electron is located towards the base of the dot, whilst the hole is nearer the apex [25]. Subsequent research on the Stark tuning of quantum dots was undertaken in the photo-current regime [26], where a laser is swept across a transition as carriers tunnel out to a contact; thus, peaks in

photocurrent spectra reveal the energy of the transitions. Photo-current studies can be useful to study the physics of quantum dots, but the absence of photon generation precludes applications in quantum photonics.

The quadratic energy level dependence of InGaAs dots coupled with the low built-in dipole moment means Stark shifts of only a few hundred micro-eV can be observed in the devices as shown in Fig. 9.2a–c, before carrier tunnelling quenches radiative emission [27]. Finding two dots with overlapping spectra, as might be required for some quantum photonics applications, requires a pre-selection of dots from a large ensemble followed by fine-tuning using the Stark effect. However, as we saw in Fig. 9.1c, the device with symmetric high-aluminium content AlGaAs barriers can apply large electric fields without significant carrier escape at a cost of reduced spectral quality from growth of dots encapsulated in the AlGaAs [13]. Fig. 9.2d introduces a useful variation on this design where the dots are located in the center of a GaAs quantum well, which is itself clad in high Aluminium content AlGaAs. This ensures the dots are grown on and encapsulated in GaAs, so retaining the preferential optical properties of this type of dot. However, the AlGaAs barriers still prevent carrier escape. Thus, dots emitting at $\sim$1.3 eV can withstand fields of hundreds of kV/cm before tunnelling dominates, enabling Stark shifts of up to 25 meV, as shown in Fig. 9.3d. Again, as the magnitude of the field is increased, the electron tunnelling becomes greater, so a preference for X^+ emission is seen. Dots which are more strongly confined, such as those emitting at 0.9 eV, can withstand even greater electric fields, displaying Stark shifts equal to 10% of the emission energy [28].

2.2 Heterostructures to increase photon collection efficiency from single quantum dots

All devices based on dots in this material system suffer from the fact the refractive index of GaAs is high (3.54 at 900 nm). Thus, photons emitting from a dipole-like transition aligned in the plane of the sample are refracted at the top surface of the semiconductor and are collected with a rather low efficiency by an objective lens above the sample. A popular solution is to embed the dots within a weak planar optical cavity formed between two Bragg mirrors, consisting of alternating quarter-wave layers of high and low refractive index AlGaAs. This redirects the emission in the vertical direction, increasing collection efficiency by a factor of up to 20 over a wavelength range defined by the quality factor of the cavity [17]. This rather simple solution to boost the photon collection can be implemented

monolithically during epitaxy, with the diode structures shown in Fig. 9.2 embedded entirely within the space between the mirrors. The thickness of this spacer must be set such that the dots are located at an antinode in the optical field of the cavity, which in the thinnest devices corresponds to the center of a one-wavelength thick GaAs spacer (such as for a GaAs LED) or the center of a half-wavelength AlGaAs spacer (in the case of an energy-tuneable diode), as shown in Fig. 9.4b. Electrical connections can be established to the diode by conduction through the mirrors, or by partially removing mirror material adjacent to the device to create intra-cavity contacts (as shown in Fig. 9.4). In designing a diode, one should take into account that GaAs/AlAs interfaces in the intrinsic region can trap photo-created charges and distort the tuning/charging behaviour of the device within certain bias ranges.

Post-growth processing of the semiconductor stack into a device can be undertaken by standard photolithography, etching and contacting schemes to create a device like that shown in Fig. 9.4. Apertures in the top metal contact may be used to isolate the emission from a single dot, or the density may be kept low by optimising growth conditions so that individual dots can be separately addressed by a diffraction limited optical spot in a confocal microscope.

Increasing the speed of operation of electrically controlled devices is an important area of research. The intrinsic speed limit of a diode is determined by the resistance–capacitance product of the diode. This can be minimised by reducing the thickness of the intrinsic region and the area of the diode (capacitance arising from the dielectric constant, diode area and intrinsic

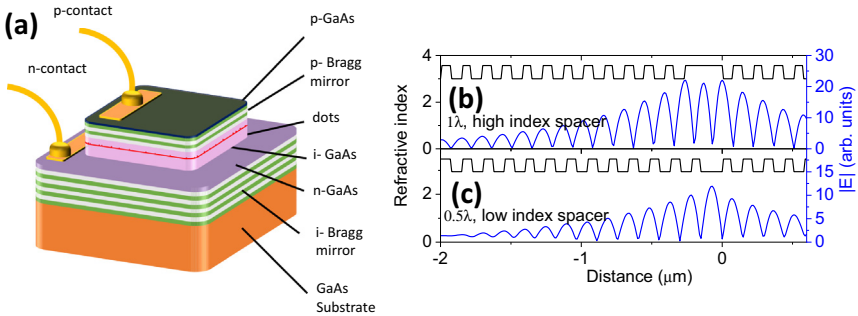


Figure 9.4 ***Design of planar-micro-cavity LEDs.*** (a) typical processed device design with intra-cavity contacts. Refractive index and electric field profile along the vertical direction in a planar micro-cavity, where the spacer may consist of (b) GaAs or (c) high aluminium composition AlGaAs.

region thickness) and reducing the resistance (arising from the material resistance and electrical contacts). Beyond design of the diode itself, consideration must be given to the electronics used to drive the system so that the impedance of connections is matched. This is non-trivial in devices held at 4 K and driven by room temperature electronics. Solutions include the development of low temperature electronics integrated with the diodes [29] and the use of an impedance matching circuit next to the diode. Operation of LEDs with a size of 40 × 40 microns at speeds of a few GHz is easily achieved [30] and recent optimisation of processing has led to devices capable of reaching tens of GHz [31].

3. Control of internal energy levels in quantum dots

3.1 Electric field control of neutral transitions

The s-shell of a single dot typically displays emission from four transitions: (i) decay from exciton to vacuum, X, (ii) decay from biexciton to exciton, XX, (ii) decay from positive trion to hole, X^+ and (iv) decay from negative trion to electron, X^-. The data shown in the previous figures show that the X has a larger polarisability than that of the XX, which is typical. Therefore, the difference between the X and XX energies, known as the binding energy, varies parabolically, but the binding energy at zero electric field differs from dot to dot.

The neutral exciton state consists of a doublet separated by a small energy called the Fine Structure Splitting (FSS) which is commonly denoted *s*. This energy difference between the eigenstates of the exciton is determined by the shape and strain of a quantum dot, crystal inversion asymmetry and the exchange interaction between electron and hole spin. In an ensemble of dots with nonzero mean FSS, the majority of the dots have their eigenstates aligned along the crystal axis. There is no fine structure associated with the X^-, X^+ or XX states as electrons, holes or both carriers in these states have net zero spin. Therefore, the fine structure manifests itself as a small energy difference in the orthogonal linearly polarised photons of the XX transition, and an equal magnitude (but opposite sign) energy splitting of the X transition (Fig. 9.5a). The FSS of self-assembled dots has received a great deal of attention due to its technological importance in the generation of entangled photon pairs from the neutral exciton cascade [32–35].

It is now well established that the application of electric field parallel to the growth direction causes a change in the FSS, and this is most easily seen

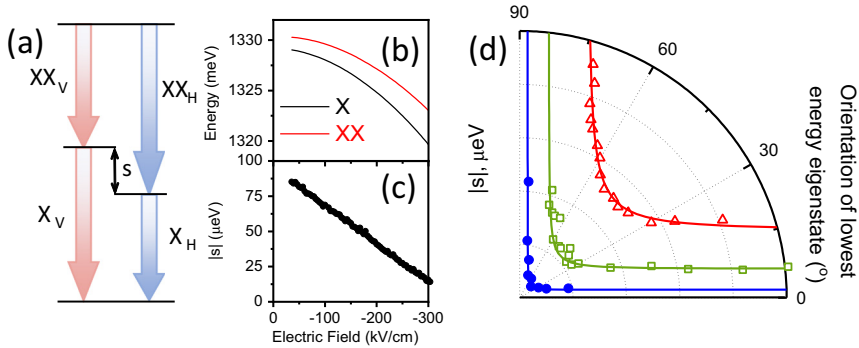


Figure 9.5 ***Vertical electric field control of the fine structure splitting (s) in the neutral exciton state.*** (a) Energy level diagram of the neutral exciton cascade in a dot. Biexciton (XX) and exciton (X) photon polarisations are correlated for large *s* and are observed to be entangled when *s* is reduced. (b) X and XX photon energies as a function of electric field in an energy tuneable device, such as Fig. 9.2d. (c) Linear change in *s* as a function of vertical electric field strength. (d) avoided crossing of the excitonic states as minimum in s is approached for three dots with different minima. As the minimum is approached, the polarisation of the photon from the lowest energy X rotates relative to the crystal orientation. *Modified from Bennett et al. [37].*

in the devices that can apply the greatest field, such as the energy tuneable diode shown in Fig. 9.2d. Intuitively, it is not surprising that the FSS varies as the electron and hole separation is changed, but it was not until it was experimentally observed that theorists considered this mechanism [36]. Several surprising observations can be made. Firstly, the FSS varies linearly with applied field in dots with $s >> s_0$, as shown in Fig. 9.5c. This is a consequence of both exciton eigenstates having identical polarisability and marginally different dipole moments. There is a remarkable homogeneity in the rate at which FSS varies within an ensemble of self-assembled dots, being $\gamma = 0.285 \pm 0.019$ μeV KV^{-1}cm [37]. Furthermore, studies of InGaAs/GaAs dots grown by different methods have also revealed identical rates of FSS change with electric field applied parallel to the growth direction [38,39]. However, dots encapsulated in other materials such as AlGaAs have a different rate of change of FSS with field, but the shift appears to always be linear [40]. Secondly, on tuning the FSS with electric field, there exists a finite minimum FSS, s_0, in each quantum dot. This s_0 value is not correlated with the excitonic emission energy, the FSS at zero field or the polarisability, or any other parameter, suggesting it is an intrinsic property of the (randomly varying) dot composition, strain and size.

Fig. 9.5d shows the behaviour of three typical dots with different s_0 as a function of applied electric field. When the FSS is large relative to the minimum value, the excitonic eigenstates are orientated along the crystal axes of the GaAs, as revealed by the polarisation of the emitted photons. As a field is applied, s falls linearly and the polarisation is fixed until it approaches s_0, at which point the linear polarisation of the transition rotates by 90 degrees in the plane. In other words, the two eigenstates (at mean energy E_0) undergo an avoided crossing with field, F, similar to the coupling of two classical oscillators with a spring, following the form of Eq. (9.1). The minimum $|s| = s_0$ is seen at a field of $F = F_0$. θ is a real angle, which gives the orientation of the linear eigenstates relative to the crystal axes of the GaAs crystal.

$$E\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = \begin{pmatrix} E_0 & -\frac{s_0}{2} \\ -\frac{s_0}{2} & E_0 - \gamma(F - F_0)\end{pmatrix}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$$

$$|s| = \sqrt{\gamma^2(F - F_0)^2 + s_0^2} \quad \theta_{\pm} = \pm\tan^{-1}\left(\frac{s_0}{\gamma(F - F_0) \pm |s|}\right) \tag{9.1}$$

Although many methods exist to modify the FSS in a quantum dot, for instance, using lateral electric field, magnetic field, strain or the dynamic Stark effect [41–46], applying an electric field along the vertical direction is arguably the most practical. It is also the easiest method to vary the FSS on a rapid timescale relative to the radiative lifetime of the exciton, which means it can be used to modify the energy levels of a qubit stored in the exciton before it radiatively decays [47].

When the FSS is small enough to prevent an observer distinguishing which path decay has occurred from in Fig. 9.5a, the photon polarisations of X and XX are entangled [37]. This means that the polarisation of each separate X and XX photon is not determined at the point of emission, but the act of measuring one photon uniquely and instantaneously determines the other's polarisation. Entanglement is a useful resource in quantum technology enabling teleportation, secret sharing and multiparty communication. The ability of a dot to create a single entangled pair of photons is highly attractive for researchers. Once a quantum dot with suitably low s_0 is identified, the electric field can be used to turn entangled photon emission on and off. Further, with fast detectors, it is possible to sort photon pairs as a function of the time difference between their detections, selecting a subset of pairs from the cascade which are entangled, even if the FSS is larger than the

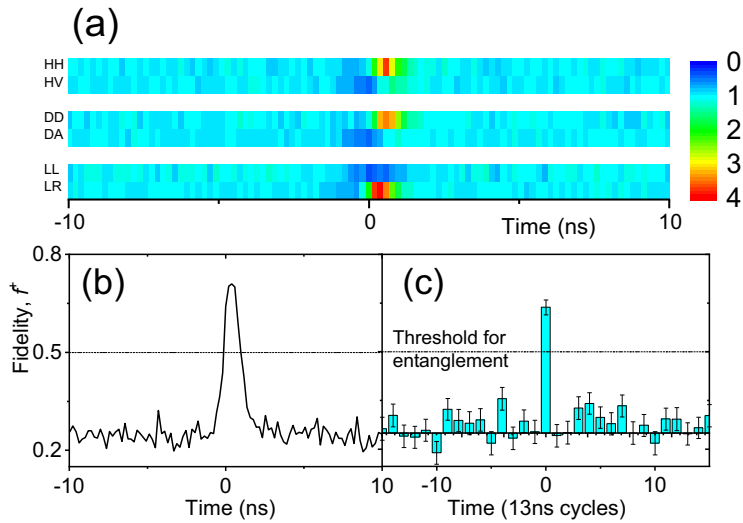


Figure 9.6 ***Entangled photon emission from an electric field tuneable device.*** If a dot can be selected with a sufficiently low minimum s, in this case 1.5 μeV, entangled photon pair emission may be observed from the XX:X cascade. (a) Time resolved photon pair correlations for same and oppositely polarised photons in three orthogonal bases. (b) Resulting fidelity to a stationary entangled photon pair for ($XX_HX_H + XX_VX_V$) under CW excitation (c) the result under pulsed excitation at 80 MHz repetition rate. In both CW and pulsed cases, the threshold of entanglement is surpassed. *Modified from Bennett et al. [37].*

radiative widths of the X transition [33]. Fig. 9.6 shows data from a dot with $s = 1.5$ micro-eV. Strong correlations between X and XX photon polarisations are seen in the two linear bases, and strong anti-correlation is seen in the circular basis (Fig. 9.6a). Under both CW (b) and pulsed excitation (c), the source fidelity to the entangled Bell state $[XX_HX_H + XX_VX_V]$ exceeds the classical limit.

3.2 Electric field control of charged transitions

An electric field applied parallel to the growth direction can also control the spin of a single charge in the dot, by modifying the energy of that spin in a magnetic field, *B*. In a magnetic field orientated along the growth direction (Faraday geometry), the eigenstates of a single electron (hole) may have spin up or down, which couple uniquely to opposite circularly polarised photons in the negative (positive) trion transitions. The energy splitting caused by the orientation of that single spin is determined by its *g*-factor such that

$\delta E = g\mu_B B$. In self-assembled InGaAs/GaAs dots, we do not observe any electric field-induced change in g-factor for this magnetic field orientation, for either the electron or the hole.

However, when the magnetic field is orientated in the plane of the sample (Voigt geometry), each trion transition splits into four parts, as illustrated in Fig. 9.7. The negatively charged trion, X^-, has its upper state split by the product of the field (B), the Bohr magneton (μ_B) and g-factor of the hole in the presence of two opposite spin electrons, which we denote as $g^{||}_{h,\,X-}$. Also, the ground state of the transition is split into two levels, which are separated by an energy proportional to the g-factor of the electron, in the absence of any other charges, denoted as $g^{||}_e$. As shown in Fig. 9.7a, this results in four allowed transitions, in which the two vertical transitions (E_1 and E_4) have one linear polarisation and the two diagonal transitions (E_2 and E_3) have the opposite linear polarisation. Similarly g-factors can be extracted from the energies of the X^+ transition in the Voigt geometry,

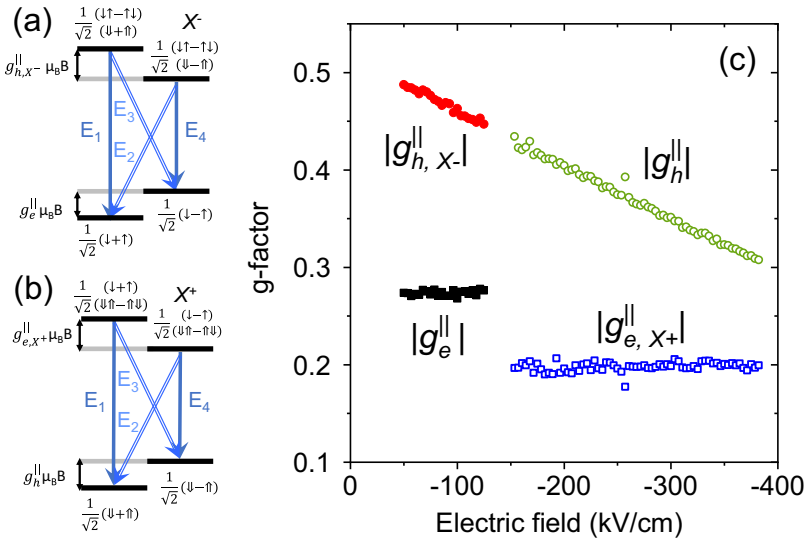


Figure 9.7 ***Electric field tuning of the Voigt geometry g-factors of single quantum dots observed through the charged exciton transitions.*** (a) Energy levels of the negatively charged exciton, X^-. The Voigt geometry g-factor is denoted $g^{||}$ for the electron (e) and hole (h) derived from the X^-/X^+ transition. (b) Energy levels of the positively charged exciton, X^+. (c) Measured changes in Voigt geometry g-factors for a typical InGaAs dot. The absolute value of the hole g-factor at zero field varies from dot to dot, but the rate of change of g-factor with field in the population of dots is remarkably homogeneous. *Modified from Bennett et al. [48].*

except that now the upper state consists of an electron spin in the presence of two opposite spin holes ($g^{||}_{e,\,X+}$) and the ground state is only a single hole ($g^{||}_{h}$). From scans such as that shown in Fig. 9.3d, using the energy tuneable devices, it is possible to separately determine the magnitude of all four of these g-factors, but not their sign. Measurements of a typical quantum dot are shown in Fig. 9.7c. We see that the electron g-factor is constant with electric field within each excitonic species, but displays a discontinuity at -120 kV/cm when the electron g-factor transitions from $g^{||}_{e}$ (determined from a single electron in the dot) to $g^{||}_{e,\,X+}$ (from an electron in the presence of two holes). In other words, the presence of the tightly confined hole wavefunctions, which have a smaller spatial distribution relative to the electrons, leads to the apparent change in the electron g-factor.

Conversely, the Voigt geometry hole g-factor varies continuously with electric field in the same manner for X^- and X^+. This is presumably a result of the larger electron wavefunction not significantly changing the hole wavefunction. The linear variation in hole g-factor may arise from the field moving the wavefunction to locations of different compositions or strain within the dot. This interpretation gives an insight into the variation in electron–hole exchange interaction seen in the neutral exciton spectrum as a function of field which we may now surmise is dominated by the behaviour of the hole. This conclusion is supported by theoretical analyses [49–51]. Measurements of a few dozen dots in this ensemble all display a tuning rate for the hole g-factor which is $(5.7 \pm 1.5) \times 10^{-4}$ cm/kV.

In most dots studied, the hole g-factor is increased as the magnitude of the reverse bias increases, but in a small fraction of the ensemble the, g-factor at zero field has an opposite sign. As the hole g-factor is then tuned to a minimum, the two eigenstates of the hole display an avoided crossing [38], which has a mathematically identical form to the variation of the FSS seen in the neutral exciton, with a rotation of the eigenstates relative to the crystal axes. This opens up the possibility that a single electrical gate could be used to manipulate a hole spin in the dot, changing not only the magnitude of the hole–spin splitting but also the orientation, allowing for complete control of the hole spin on the Bloch sphere [52].

4. Hybrid approaches to control of quantum dots

We have seen that electric field applied to quantum dots along the growth direction can be used to inject carriers, control charging, vary

absolute energies and alter internal energy levels. However, there has been limited success in achieving more than one of these outcomes in a single device. Recent work has shown combining more than one approach to controlling quantum dots can create new functionalities and realise new applications, for instance, in developing a compact electrically driven emitter with tuneable energy. Here, we review a selection of these hybrid approaches focussing on those that use electric field.

4.1 A tuneable electroluminescent quantum dot light source

An LED operates in forward bias and an energy-tuneable diode operates in reverse bias, so at first glance, these two technologies are incompatible. However, the integration of multiple devices in a single chip offers a neat solution to this problem. The concept is illustrated in Fig. 9.8a. Two diodes fabricated in close proximity are optically coupled. Diode A is driven in forward bias to emit light. The spectrum will be dominated by emission from the wetting layer at 870 nm. This emission may be coupled into a second diode B, through in-plane waveguiding between the Bragg mirrors. It is an open challenge to identify the best design for optically coupling light from diode A to diode B whilst having them controlled by independent electrical contacts. In the report of Lee et al. [53], a few micron wide trench was etched between the LEDs and some light was (inefficiently) coupled, although in later iterations the diode A surrounded diode B on all sides, giving a higher efficiency [54]. A related concept is the use of an electrically driven laser on chip [55] coupled to multiple single-dot devices.

For the report of Lee et al. [53], the heterostructure was based on quantum dots emitting at 940 nm within a 10 nm GaAs quantum well, clad with $Al_{0.75}Ga_{0.25}As$. Fig. 9.8c shows the pump diode does not pass current until $\sim$2.1 V, consistent with the bandgap of the $Al_{0.75}Ga_{0.25}As$ alloy. Above 2.1 V, wetting layer emission from diode A creates electron–hole pairs in diode B by photoluminescence. This second diode may then be tuned independently to vary the energy, FSS and g-factor of the single dots. When spectral filtering is used to remove the 870 nm wetting layer emission from A, emission from single dots in B is seen to vary as would be expected and display anti-bunched photon statistics from individual transitions. This device concept has been developed to create a pump-tune device where the second diode can minimise the FSS of a quantum dot, enabling the device to be deployed in a field test of entangled photon distribution [54].

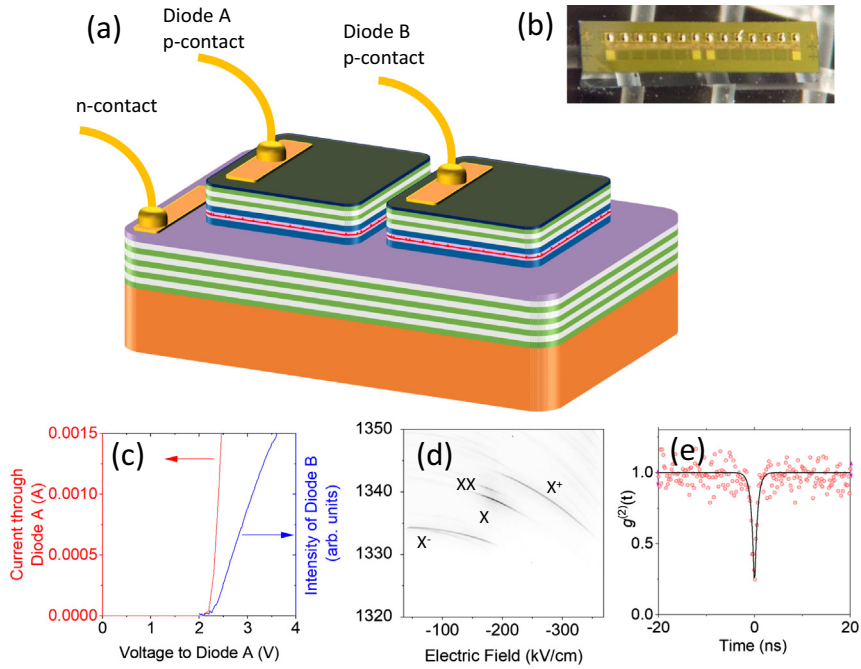


Figure 9.8 ***A device to simultaneously electrically excite and tune single dots.*** (a) Schematic of the *de minimis* device design with two independently addressable diodes in close proximity: Diode A is driven in forward bias to create electroluminescence (mainly at 870 nm) which photo-excites Diode B, which can Stark shift tuned in reverse bias. (b) Photograph of a linear array of closely packed LEDs, on a 250 micron pitch. (c) Current through A, and light emission from B, as a function of the bias on A. (d) spectrum of a dot in B as a function of the electric field on B, when A is held at fixed current. (e) autocorrelation histogram on the emission of single dot in B, confirming the suppression of classical light emission from A. *Modified from Ellis et al. [56] and Lee et al. [53].*

4.2 Combining coherent optical control with tuneable light sources

A major theme in quantum dot research in the last decade has been the use of resonant optical fields to control spins [19], to generate ultra-coherent photons [57,58] and deterministically generate indistinguishable photons [59,60]. Combining electric field control and resonant optical fields is also a convenient way to create quantum light. Many of the diodes discussed in this chapter are of a planar design with tens of microns lateral distance between etched surfaces, which reduces the polarisation scattered laser reaching the detector in a confocal microscope.

A key parameter is the distance between the emitter and nearby trap states: this is vital in minimising the linewidths of the transitions and maximising coherence [61]. In parallel with reducing the background doping level in the intrinsic region, we have observed that spectral jitter is reduced when the distance between a dot and the doped layers is increased to a few hundred nanometers. It has been reported that a fast electrical diode can be used within a feedback mechanism to lock individual transitions to an ultra-narrow laser, further reducing the transition linewidth to the radiative limit [62].

Conversely, the ability to rapidly change the transition energies of a dot at GHz speeds can be used to create a novel source of coherent single photons. When resonant with the laser, photons can either be coherently scattered (in which case their phase may be locked to the phase of laser field) or the dot may be excited, and subsequent spontaneous emission has a coherence determined by the solid state emitter. Fig. 9.9 shows the results of an experiment when the device is excited by a resonant, narrowband CW laser and the electric field across the dot is modulated at high speed to periodically shift the transition into resonance with the laser [63]. Photon autocorrelation measurements show the resulting light to be anti-bunched on a timescale determined by the radiative lifetime of the upper state. However, the spectrum is narrowed to a width of 1.3 micro-eV as the resonant excitation populates only the emissive state reducing electrostatic spectral jitter to a minimum.

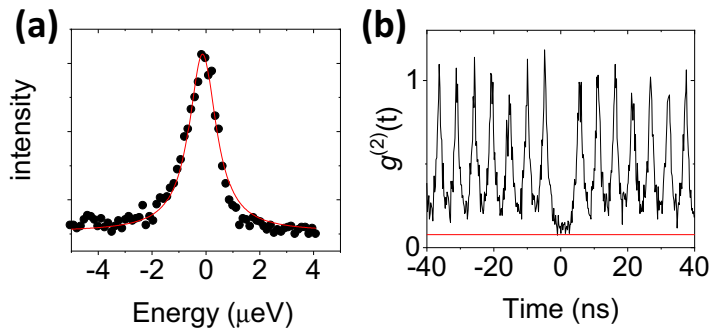


Figure 9.9 ***Resonance fluorescence in combination with fast tuning of dot energy.*** (a) shows the intensity as a function of electric field, excited with a weak narrow band CW laser. (b) An autocorrelation histogram result obtained when a diode is periodically driven with electrical pulses to bring it in and out of resonance with a CW laser, showing strong suppression of multi-photon emission. *Modified from Cao et al. [63].*

4.3 Strain and electric field tuneable quantum dots

Trotta et al. have shown that the combination of electric field and strain is an especially promising combination. In a series of works, they have demonstrated devices that combine electrical injection with strain tuning [49,64], the use of both control techniques to create tuneable entangled photon sources [65] and most impressively the universal minimisation of the FSS for any quantum dot [66]. They have shown that an electric field applied along the growth direction perturbs the symmetry of the exchange interaction in a manner orthogonal to the strain field. Thus, the two tuning knobs can be used in unison to ensure *every* dot can have a minimum FSS s_0 below the measurement resolution, and emit highly entangled photon pairs from the neutral cascade without time filtering (Fig. 9.10).

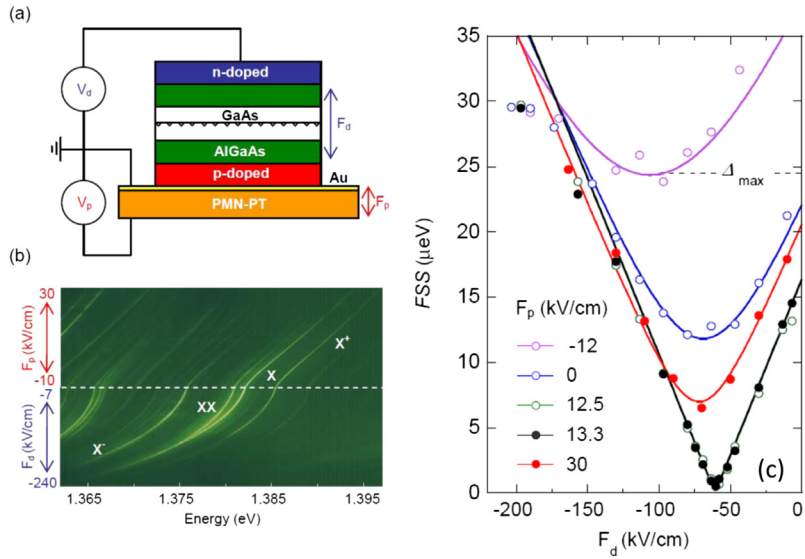


Figure 9.10 ***Strain and electric field tuneable quantum dots*** (a) Electric field tuneable device mounted upon a piezo-electronic transducer. (b) Effect of tuning the piezo-electronic field F_P and the diode electric field F_d. (c) Through simultaneous optimisation of the two fields, a minimum in fine structure splitting can always be found for each dot. *Reprinted with permission from Trotta et al. [66].*

5. Future directions

Quantum dots have been one of the most exciting research areas in solid state physics over the last 20 years. Ground-breaking work has unearthed new physics from the interaction of charges in confined systems with resonant optical lasers [22,57,58,67], ensembles of nuclear spins [7,68,69] and nano-photonic cavities [6,70,71]. Inclusion of dots within semiconductor heterostructures has played a key role in these studies, as the control of charge and energy has allowed scientists to develop new technologies and measurement techniques.

Already, devices based on self-assembled InGaAs dots are available from a number of start-up companies and have been deployed in field trials of quantum optical systems [54]. It will be interesting to see whether single photon emitters in other materials systems will be developed to the high degree of technological maturity achieved with this type of dot. Of particular interest are material systems that can operate at elevated temperatures. The toolkits developed from studying dots are enabling researchers to develop quantum light sources based on materials that can operate at elevated temperature, including colour centers in diamond [72], aluminium nitride [73], silicon carbide [74] and two-dimensional materials such as hexagonal boron nitride [74]. Already, single photon emitting diodes based on these devices have been reported [75,76], albeit with phonon-sidebands hundreds of milli-electron volts wide.

A further avenue for study is the deterministic placement of dots and cavities to create devices with higher yield. To date, the best optical quality dots are all randomly self-assembled, and it is a significant technological challenge to induce dot nucleation at precise locations without introducing defects or strain which degrade the optical properties. Recent reports have shown some progress in creating high-quality positioned dots [77] and others have developed methods to create cavities aligned to a randomly positioned dots [78–80].

Finally, future research on dots is surely going to be motivated by applications in quantum technology, whether that be in communications, sensing, imaging, computing or simulation. Some researchers are focussed on optimising dots that emit in the telecom bands at 1310 and 1550 nm [54,81] for applications in communications. Other researchers are pushing the limits of what can be achieved in terms of efficiency and indistinguishability, most recently in showing a 14-photon boson sampling experiment

[82]. Such demonstrations are tantalisingly close to demonstrating the supremacy of a photonic quantum computer over classical technology, a result which would have profound implications for our understanding of what computation means.

Credit: I would like to thank Dr J P Hadden and Dr P Androvitsaneas for critical reading of this manuscript. I would like to acknowledge Toshiba Research Europe Limited (TREL) for funding, numerous co-workers in the Quantum Information Group at TREL and the Semiconductor Physics Group at the Cavendish Laboratory (Cambridge University) whose work was essential to obtaining these results. In addition, I acknowledge EPSRC grants EP/T017813/1 and EP/T001062/1 for funding.

References

[1] Reitzenstein S, Forchel A. Quantum dot micropillars. J Phys D Appl Phys 2010;43(3).
[2] Shields AJ, Stevenson RM, Thompson RM, Yuan Z, Kardynal BE. Generation of single photons using semiconductor quantum dots. In: Nano-physics bio-electronics A new Odyssey; 2002. p. 111–46.
[3] Bennett AJ, Atkinson P, See P, Ward MB, Stevenson RM, Yuan ZL, et al. Single-photon-emitting diodes: a review. Phys Status Solidi Basic Res 2006;243(14):3730–40.
[4] Eisaman MD, Fan J, Migdall A, Polyakov SV. Invited review article: single-photon sources and detectors. Rev Sci Instrum 2011;82(7):1–25.
[5] Khitrova G, Gibbs HM, Kira M, Koch SW, Scherer A. Vacuum Rabi splitting in semiconductors. Nat Phys 2006;2(2):81–90.
[6] Buckley S, Rivoire K, Vučković J. Engineered quantum dot single-photon sources. Reports Prog Phys 2012;75(12):126503.
[7] Urbaszek B, Marie X, Amand T, Krebs O, Voisin P, Maletinsky P, et al. Nuclear spin physics in quantum dots: an optical investigation. Rev Mod Phys 2013;85(1):79–133.
[8] Lodahl P. Quantum-dot based photonic quantum networks. Quantum science and technology, Vol. 3. Institute of Physics Publishing; 2018.
[9] Michler P, Kiraz A, Becher C, Schoenfeld WV, Petroff PM, Zhang L, et al. A quantum dot single-photon turnstile device. Science 2000;290(5500):2282–5.
[10] Shields AJ, Stevenson RM, Young RJ. Entangled photon generation by quantum dots. In: Single semiconductor quantum dots; 2009. p. 227–65.
[11] Schwartz I, Cogan D, Schmidgall ER, Don Y, Gantz L, Kenneth O, et al. Deterministic generation of a cluster state of entangled photons. Science 2016;354(6311):434–7.
[12] Santori C, Fattal D, Vučković J, Solomon GS, Yamamoto Y. Indistinguishable photons from a single-photon device. Nature 2002;419:594–7.
[13] Bennett AJ, Patel RB, Skiba-Szymanska J, Nicoll CA, Farrer I, Ritchie DA, et al. Giant Stark effect in the emission of single semiconductor quantum dots. Appl Phys Lett 2010;97(3):2–4.
[14] Yuan Z, Kardynal BE, Stevenson RM, Shields AJ, Lobo CJ, Cooper K, et al. Electrically driven single-photon source. Science 2002;295(5552):102–5.
[15] Warburton RJ, Schäflein C, Haft D, Bickel F, Lorke A, Karrai K, et al. Optical emission from a charge-tunable quantum ring. Nature 2000;405(6789):926–9.
[16] Patel RB, Bennett AJ, Cooper K, Atkinson P, Nicoll CA, Ritchie DA, et al. Postselective two-photon interference from a continuous nonclassical stream of photons emitted by a quantum dot. Physical Review Letters 2008;100:207405.
[17] Bennett AJ, Unitt DC, See P, Shields AJ, Atkinson P, Cooper K, et al. Microcavity single-photon-emitting diode. Appl Phys Lett 2005;86(18):1–3.

[18] Salter CL, Stevenson RM, Farrer I, Nicoll CA, Ritchie DA, Shields AJ. An entangled-light-emitting diode. Nature 2010;465(7298):594–7.
[19] Press D, de Greve K, Mcmahon PL, Ladd TD, Friess B, Forchel A, et al. Ultrafast optical spin echo in a single quantum dot. Nat Photon 2010;4:367–70.
[20] Ediger M, Dalgarno PA, Smith JM, Gerardot BD, Warburton RJ, Karrai K, et al. Controlled generation of neutral, negatively-charged and positively-charged excitons in the same single quantum dot. Appl Phys Lett 2005;86(21):1–3.
[21] Warburton RJ, Urbaszek B, McGhee EJ, Schulhauser C, Högele A, Karrai K, et al. Charged excitons in self-assembled quantum dots. Mater Res Soc Symp Proc 2003; 737(26):95–105.
[22] Brunner D, Gerardot BD, Dalgarno PA, Wüst G, Karrai K, Stoltz NG, et al. A coherent single-hole spin in a semiconductor. Science 2009;325(5936):70–2.
[23] Kroner M, Govorov AO, Remi S, Biedermann B, Seidl S, Badolato A, et al. The nonlinear Fano effect. Nature 2008;451(7176):311–4.
[24] Kleemans NAJM, Van Bree J, Govorov AO, Keizer JG, Hamhuis GJ, Nötzel R, et al. Many-body exciton states in self-assembled quantum dots coupled to a Fermi sea. Nat Phys 2010;6(7):534–8.
[25] Fry PW, Itskevich IE, Mowbray DJ, Skolnick MS, Finley JJ, Barker JA, et al. Inverted electron-hole alignment in InAs-GaAs self-assembled quantum dots. Phys Rev Lett 2000;84(4):733–6.
[26] Ramsay AJ, Gopal AV, Gauger EM, Nazir A, Lovett BW, Fox AM, et al. Damping of exciton rabi rotations by acoustic phonons in optically excited InGaAs/GaAs quantum dots. Phys Rev Lett 2010;104(1):017402.
[27] Oulton R, Finley JJ, Ashmore AD, Gregory IS, Mowbray DJ, Skolnick MS, et al. Manipulation of the homogeneous linewidth of an individual In(Ga)As quantum dot. Physical Review B: Condensed Matter and Materials Physics 2002;66(4):1–4.
[28] Ward MB, Dean MC, Stevenson RM, Bennett AJ, Ellis DJP, Cooper K, et al. Coherent dynamics of a telecom-wavelength entangled photon source. Nat Commun 2014;5:1–6.
[29] Michaelis De Vasconcellos S, Gordon S, Bichler M, Meier T, Zrenner A. Coherent control of a single exciton qubit by optoelectronic manipulation. Nat Photon 2010; 4(8):545–8.
[30] Bennett AJ, Unitt DC, See P, Shields AJ, Atkinson P, Cooper K, et al. Electrical control of the uncertainty in the time of single photon emission events. Phys Rev B 2005; 72(3):2–5.
[31] Müller JRA, Mark Stevenson R, Skiba-Szymanska J, Shooter G, Huwer J, Farrer I, et al. Active reset of a radiative cascade for entangled-photon generation beyond the continuous-driving limit. Phys Rev Res 2020:2. 043292.
[32] Young RJ, Stevenson RM, Hudson AJ, Nicoll CA, Ritchie DA, Shields AJ. Bell-inequality violation with a triggered photon-pair source. Phys Rev Lett 2009; 102(3):1–4.
[33] Stevenson RM, Hudson AJ, Bennett AJ, Young RJ, Nicoll CA, Ritchie DA, et al. Evolution of entanglement between distinguishable light states. Phys Rev Lett 2008; 101(17):1–4.
[34] Young RJ, Stevenson RM, Atkinson P, Cooper K, Ritchie DA, Shields AJ. Improved fidelity of triggered entangled photons from single quantum dots. New J Phys 2006; 8(06):1–9.
[35] Stevenson RM, Young RJ, Atkinson P, Cooper K, Ritchie DA, Shields AJ. A semiconductor source of triggered entangled photon pairs. Nature 2006;439:178–82.
[36] Singh R, Bester G. Nanowire quantum dots as an ideal source of entangled photon pairs. Phys Rev Lett 2009;103(6):1–4.

[37] Bennett AJ, Pooley MA, Stevenson RM, Ward MB, Patel RB, De La Giroday AB, et al. Electric-field-induced coherent coupling of the exciton states in a single quantum dot. Nat Phys 2010;6(12):947–50.
[38] Bennett AJ, Pooley MA, Cao Y, Sköld N, Farrer I, Ritchie DA, et al. Voltage tunability of single spin-states in a quantum dot. Nat Commun 2013;7.
[39] Prechtel JH, Maier F, Houel J, Kuhlmann AV, Ludwig A, Wieck AD, et al. Electrically tunable hole g factor of an optically active quantum dot for fast spin rotations. Phys Rev B 2015;91(16):1–8.
[40] Pooley MA, Bennett AJ, Farrer I, Ritchie DA, Shields AJ. Engineering quantum dots for electrical control of the fine structure splitting. Appl Phys Lett 2013;103(3).
[41] Muller A, Fang W, Lawall J, Solomon GS. Creating polarization-entangled photon pairs from a semiconductor quantum dot using the optical Stark effect. Phys Rev Lett 2009;103(21):2–5.
[42] Seidl S, Kroner M, Högele A, Karrai K, Warburton RJ, Badolato A, et al. Effect of uniaxial stress on excitons in a self-assembled quantum dot. Appl Phys Lett 2006; 88(20):1–3.
[43] Seidl S, Gerardot BD, Dalgarno PA, Kowalik K, Holleitner AW, Petroff PM. Statistics of quantum dot exciton fine structure splittings and their polarization orientations. Phys E 2008;40:2153–5.
[44] Alén B, Bosch J, Granados D, Martínez-Pastor J, García JM, González L. Oscillator strength reduction induced by external electric fields in self-assembled quantum dots and rings. Phys Rev B 2007;75(4):1–7.
[45] Young RJ, Stevenson RM, Shields AJ, Atkinson P, Cooper K, Ritchie DA, et al. Inversion of exciton level splitting in quantum dots. Phys Rev B 2005;72(11):1–4.
[46] Ellis DJP, Stevenson RM, Young RJ, Shields AJ, Atkinson P, Ritchie DA. Control of fine-structure splitting of individual InAs quantum dots by rapid thermal annealing. Appl Phys Lett 2007;90(1):19–21.
[47] de la Giroday AB, Bennett AJ, Pooley MA, Stevenson RM, Skold N, Patel RB, et al. All-electrical coherent control of the exciton states in a single quantum dot. Phys Rev B 2010;82:241301.
[48] Bennett AJ, Pooley MA, Cao Y, Sköld N, Farrer I, Ritchie DA, et al. Voltage tunability of single-spin states in a quantum dot. Nat Commun 2013;4.
[49] Trotta R, Martín-Sánchez J, Wildmann JS, Piredda G, Reindl M, Schimpf C, et al. Wavelength-tunable sources of entangled photons interfaced with atomic vapours. Nat Commun 2016;7:1–7.
[50] Ding F, Singh R, Plumhof JD, Zander T, Křápek V, Chen YH, et al. Tuning the exciton binding energies in single self-assembled InGaAs/GaAs quantum dots by piezoelectric-induced biaxial stress. Phys Rev Lett 2010;104(6):067405.
[51] Plumhof JD, Křápek V, Ding F, Jöns KD, Hafenbrak R, Klenovský P, et al. Strain-induced anticrossing of bright exciton levels in single self-assembled GaAs/AlxGa1--xAs and InxGa 1-xAs/GaAs quantum dots. Phys Rev B 2011;83(12):1–4.
[52] Pingenot J, Pryor CE, Flatt ME. Method for full Bloch sphere control of a localized spin via a single electrical gate. Appl Phys Lett 2008;92(22):222502.
[53] Lee JP, Murray E, Bennett AJ, Ellis DJP, Dangel C, Farrer I, et al. Electrically driven and electrically tunable quantum light sources. Appl Phys Lett 2017;110:071102.
[54] Xiang Z-H, Huwer J, Skiba-Szymanska J, Stevenson RM, Ellis DJP, Farrer I, et al. A tuneable telecom-wavelength entangled light emitting diode. Commun Phys 2019;3:121.
[55] Stock E, Albert F, Hopfmann C, Lermer M, Schneider C, Höfl S, et al. On-chip quantum optics with quantum dot microcavities. Advanced Materials 2012:1–4.

[56] Ellis DJP, Bennett AJ, Dangel C, Lee JP, Griffiths JP, Mitchell TA, et al. Independent indistinguishable quantum light sources on a reconfigurable photonic integrated circuit. Appl Phys Lett 2018;112(21):211104.
[57] Nguyen HS, Sallen G, Voisin C, Roussignol P, Diederichs C, Cassabois G. Ultra-coherent single photon source. Appl Phys Lett 2011;99(26).
[58] Matthiesen C, Vamivakas AN, Atatüre M. Subnatural linewidth single photons from a quantum dot. Phys Rev Lett 2012;108(9):1–4.
[59] Ding X, He Y, Duan ZC, Gregersen N, Chen MC, Unsleber S, et al. On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar. Phys Rev Lett 2016;116(2).
[60] He Y, Su Z, Huang H, Ding X, Qin J, Wang C, et al. Scalable boson sampling with a single-photon device. 2016. arXiv:160304127v2.
[61] Houel J, Kuhlmann AV, Greuter L, Xue F, Poggio M, Warburton RJ, et al. Probing single-charge fluctuations at a GaAs/AlAs interface using laser spectroscopy on a nearby InGaAs quantum dot. Phys Rev Lett 2012;108(10):107401.
[62] Prechtel JH, Kuhlmann AV, Houel J, Greuter L, Ludwig A, Reuter D, et al. Frequency-stabilized source of single photons from a solid-state qubit. Phys Rev X 2013;3(4):041006.
[63] Cao Y, Bennett AJ, Ellis DJP, Farrer I, Ritchie DA, Shields AJ. Ultrafast electrical control of a resonantly driven single photon source. Appl Phys Lett 2014;105(5):051112.
[64] Trotta R, Atkinson P, Plumhof JD, Zallo E, Rezaev RO, Kumar S, et al. Nanomembrane quantum-light-emitting diodes integrated onto piezoelectric actuators. Adv Mater 2012;24(20):2668–72.
[65] Trotta R, Wildmann JS, Zallo E, Schmidt OG, Rastelli A. Highly entangled photons from hybrid piezoelectric-semiconductor quantum dot devices. Nano Lett 2014;14(6):3439–44.
[66] Trotta R, Zallo E, Ortix C, Atkinson P, Plumhof JD, Van Den Brink J, et al. Universal recovery of the energy-level degeneracy of bright excitons in InGaAs quantum dots without a structure symmetry. Phys Rev Lett 2012;109(14):147401.
[67] Vamivakas AN, Zhao Y, Lu CY, Atatüre M. Spin-resolved quantum-dot resonance fluorescence. Nat Phys 2009;5(3):198–202.
[68] Waeber AM, Hopkinson M, Farrer I, Ritchie DA, Nilsson J, Stevenson RM, et al. Few-second-long correlation times in a quantum dot nuclear spin bath probed by frequency-comb nuclear magnetic resonance spectroscopy. Nat Phys 2016;12(7):688–93.
[69] Gerardot BD, Brunner D, Dalgarno PA, Öhberg P, Seidl S, Kroner M, et al. Optical pumping of a single hole spin in a quantum dot. Nature 2008;451(7177):441–4.
[70] Faraon A, Fushman I, Englund D, Stoltz N, Petroff P, Vučković J. Coherent generation of non-classical light on a chip via photon-induced tunnelling and blockade. Nat Phys 2008;4(11):859–63.
[71] Fattal D, Santori C, Vuckovic J, Solomon GS, Yamamoto Y. Indistinguishable single photons from a quantum dot. Phys Status Solidi Basic Res 2003;238(2):305–8.
[72] Kurtsiefer C, Mayer S, Zarda P, Weinfurter H. Stable solid-state source of single photons. Phys Rev Lett 2000;85(2):290–3.
[73] Bishop SG, Hadden JP, Alzahrani FD, Hekmati R, Huffaker DL, Langbein WW, et al. Room-temperature quantum emitter in aluminum nitride. ACS Photon 2020;7(7):1636–41.
[74] Tran TT, Bray K, Ford MJ, Toth M, Aharonovich I. Quantum emission from hexagonal boron nitride monolayers. Nat Nanotechnol 2016;11(1):37–41.
[75] Mizuochi N, Makino T, Kato H, Takeuchi D, Ogura M, Okushi H, et al. Electrically driven single-photon source at room temperature in diamond. Nat Photon 2012;6(5):299–303.

[76] Fuchs F, Soltamov VA, Väth S, Baranov PG, Mokhov EN, Astakhov GV, et al. Silicon carbide light-emitting diode as a prospective room temperature source for single photons. Sci Rep 2013;3(1):1–4.
[77] Jöns KD, Atkinson P, Müller M, Heldmaier M, Ulrich SM, Schmidt OG, et al. Triggered indistinguishable single photons with narrow line widths from site-controlled quantum dots. Nano Lett 2013;13(1):126–30.
[78] Dousse A, Suffczyński J, Beveratos A, Krebs O, Lemaître A, Sagnes I, et al. Ultrabright source of entangled photon pairs. Nature 2010;466(7303):217–20.
[79] Hennessy K, Badolato A, Winger M, Gerace D, Atatüre M, Gulde S, et al. Quantum nature of a strongly coupled single quantum dot-cavity system. Nature 2007; 445(7130):896–9.
[80] Davanco M, Liu J, Sapienza L, Zhang CZ, De Miranda Cardoso JV, Verma V, et al. Heterogeneous integration for on-chip quantum photonic circuits with single quantum dot devices. Nat Commun 2017;8(1):1–23.
[81] Müller T, Skiba-Szymanska J, Krysa AB, Huwer J, Felle M, Anderson M, et al. A quantum light-emitting diode for the standard telecom window around 1,550 nm. Nat Commun 2018;9(1):1–6.
[82] Wang H, Qin J, Ding X, Chen MC, Chen S, You X, et al. Boson sampling with 20 input photons and a 60-mode interferometer in a 1014 -dimensional Hilbert space. Phys Rev Lett 2019;123(25):250503.

CHAPTER TEN

Semiconductor quantum dot solar cells

Yoshitaka Okada[1,2,*] Katsuhisa Yoshida[1,3] Yasushi Shoji[1,4] and Ryo Tamaki[1]

[1]Research Center for Advanced Science and Technology (RCAST), The University of Tokyo, Tokyo, Japan
[2]Department of Advanced Interdisciplinary Studies, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan
[3]Institute of Applied Physics, University of Tsukuba, Tsukuba, Japan
[4]Global Zero Emission Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan
*Corresponding author: okada@mbe.rcast.u-tokyo.ac.jp

1. Introduction

For an ideal solar cell (SC), in which all the nonradiative recombination processes are negligibly small, the theoretical maximum energy conversion efficiency becomes as high as ~85% by assuming fully concentrated black-body radiation at 5800K [1,2]. The maximum efficiency of a single-junction SC cell, however, is limited to the Shockley–Queisser limit of ~31% for an air mass 1.5 spectrum [3]. The main physical processes that limit the SC efficiency are the losses by thermal dissipation or thermalisation, and nonabsorption of low-energy below-bandgap photons. These losses could add up to as much as ~ 40% of incident solar energy. Thus, improving the SC efficiency means developing methods to reduce these losses.

One of the concepts that are well established today is to split the solar spectrum amongst multiple bandgap absorbers or sub-cells, e.g., tandem or multi-junction cells. The other approaches employ more advanced techniques such as hot carriers, multi-exciton generation and intermediate band (IB) absorption in low-dimensional nanostructures such as semiconductor quantum dots (QDs) [4–6]. Amongst various approaches being studied, this chapter is devoted to describing the basic principles and development of the state-of-the-art technologies for QD-based intermediate band solar cells (IBSCs).

A high-density, size-controlled array of QDs incorporated in the active region of a *p-i-n* single-junction SC has attracted significant interest as a potential IBSC, which utilises the sub-bandgap infrared (IR) photons to

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00002-2
© 2021 Elsevier Ltd.
All rights reserved.

generate additional photocurrent other than that corresponding to the valence-to-conduction band (VB-CB) optical excitation, through extra absorption via superlattice miniband states [4–6]. The QDs also offer the possibility of reducing thermal dissipation loss. If such a nanostructured IBSC SC were truly realised, its conversion efficiency can exceed the Shockley–Queisser limit of a conventional single-junction SC. The calculated maximum efficiency for an SC incorporating one IB becomes >47% under one sun, and ~63% under full solar concentration [6].

In a QD-IBSC, QDs are required to be homogeneous and small in size, and are regularly and tightly placed. This structural configuration then leads to formation of an IB or a superlattice miniband that is well separated in energy from the higher-order states [7,8]. Secondly, IB states should be partially filled, or ideally be half-filled, with electrons in order to ensure an efficient pumping of electrons by providing both the empty states to receive electrons being photo-excited from the VB, and filled states to promote electrons to the CB via absorption of additional sub-bandgap photons [9].

The proposed implementation of a QD-IBSC must accompany two-step photo-absorption (TSPA) via IB states, but it has been difficult to clearly demonstrate this operation concept *at room temperature*, which is due to, as will be discussed in more detail in the following section, relatively small optical absorption rates from the QD-IB states to the CB continuum. The rate of thermal escape of electrons from the IB to the CB as well as relaxation from the CB to the IB (the reverse process) also increases significantly at room temperature.

Research into QD-IBSCs is therefore currently undergoing two stages. The first is to develop technology to fabricate high-density and size-controlled QD arrays with low interface defect densities and sufficiently long carrier lifetimes in the optimised QD/host material bandgap system. The fabrication of QD arrays is most commonly achieved by spontaneous self-assembly of coherent three-dimensional (3D) islands in lattice-mismatched epitaxy, well known as Stranski–Krastanov (S-K) growth in molecular beam epitaxy (MBE) and metal organic vapor phase epitaxy (MOVPE). However, the number of QDs is limited by lattice strain accumulation in the single crystal material as the number of stacked QD layers is increased. In S-K growth of InAs/(Al)GaAs system, which is one of the most investigated and well-established QD material systems, misfit dislocations are generated after typically 10–15 layers of stacking [10]. Recently, strain-compensated or strain-balanced growth technique has shown

significantly improved QD quality and QD-SC characteristics even after the stacking of 50–100 QD layers by S-K growth. To date, InAs QDs in a AlGaInAs matrix on an InP substrate [11,12], InAs QDs in GaAsP [13] and in GaP [14] matrices on GaAs substrates and InAs QDs in a GaNAs matrix on GaAs(001) [15–17] and (311)B substrates [18,19] have been reported.

The second stage needed to implement a high efficiency QD-IBSC is to realise partially filled or ideally half-filled IB states in order to maximise the photocurrent generation by TSPA, which is the main operating principle of the IBSC. Strandberg and Reenaas [20] have calculated that the IB region must be partially filled by means of doping in order to achieve high efficiency, if QD-IBSC were to be operated under one sun condition. On the other, it becomes possible to sustain a reasonable population of photo-generated electrons even in a nondoped QD-IBSC, if it is operated under concentrated sunlight, typically 100–1000 suns. The optimal cell and material designs for IBSCs have also been reported [21,22].

The sub-bandgap IR photons can optically pump the electrons both from VB to IB states and from IB to CB states. However, both the rate of thermal escape, as opposed to optical excitation, of electrons from IB to CB continuum and the transition rate from CB to IB increase significantly at room temperature. If all the carriers in IB states are in electrochemical equilibrium, which is the condition defined by an independent quasi-Fermi level E_{IF}, then a clear detection of photo-excitation of electrons from IB to CB giving rise to an increase in the photocurrent at room temperature should become possible. A photocurrent production as a direct result of TSPA has been successfully studied [23–25], though the rate of this TSPA is still small at present.

In this chapter, we will first introduce a recent numerical analysis of device characteristics based on a self-consistent drift-diffusion model, for which the carrier dynamics with a particular emphasis on TSPA and recombination processes is taken into consideration. Then, the current status of the fabrication and performance of high-density In(Ga)As QD SCs by using self-organised MBE technology is discussed. As mentioned, maximising the TSPA processes via QDs is of paramount importance in a successful implementation of high-efficiency IBSCs, and for this, Fourier transform infrared (FTIR) photocurrent spectroscopy provides a powerful tool to study the IR photocurrent response via TSPA in In(Ga)As QDSCs. The details of FTIR spectroscopy characterisation of QDSC are also given.

2. Drift-diffusion analysis of quantum efficiency in QD-IBSCs

2.1 Introduction

For the characterisation of SC, the measurements of current–voltage (*J-V*) curves under sunlight illumination and of quantum efficiencies (QEs) are essential. While the basic device physics for single-junction SCs are well developed, the analysis of QD-IBSCs is not straightforward because the carrier dynamics via the IB needs to be taken into account. For example, in the light *J-V* measurements, the measured currents contain information on both the carrier dynamics via IB and the dynamics occurring within the host material, and it is difficult to distinguish between these two effects. Regarding the QE measurement, in which monochromatic light is used to resolve the wavelength dependence of carrier absorption and collection, it is not suitable for evaluating the potential of fabricated cells as IBSC because it does not provide information on the rates of TSPA.

In order to quantify the TSPA process, advanced experimental techniques have been reported [23,25–29]. Amongst them, a typical approach is the QE measurement performed under a continuous IR light [23]. The IR light allows the TSPA and also extraction of carriers captured by IB. For further quantification, the optical absorption spectrum, the density of state of QDs and also carrier transition by the nonradiative process must be taken into consideration. Therefore, the analysis based on the device simulation incorporating those effects is essential.

In this section, we introduce the drift-diffusion method for QD-IBSCs incorporating a nonradiative electron capture process from CB to IB, which then leads to a reduction of output electron current [30]. The simulation method for IBSC is first explained and then the results of simulation of light *J-V* measurement and QE are explained. We also calculate QE curves under continuous IR light irradiation.

2.2 Simulation method

We employ the 1D drift-diffusion method for IBSC [21,31] incorporating a fast electron relaxation process from the CB to the IB. Here, we assume the IB is the energy state for electrons. In this method, we should find the concentrations of electrons (n) in the CB, holes (p) in the VB and also electrons in the IB (n_I) and the electrostatic potential (ψ), self-consistently.

In the Poisson equation, the electrostatic potential ψ is given by

$$\frac{d}{dx}\left(\varepsilon\frac{d\psi}{dx}\right) = -q\left(p + N_D^+ - n - n_I - N_A^-\right), \tag{10.1}$$

where ε is the dielectric constant, q is the elementary charge and N_D^+ and N_A^- are the ionised donor and acceptor concentrations, respectively. In the steady state, the continuity equations of electrons in CB and holes in VB are given by

$$\frac{1}{q}\frac{dJ_n}{dx} = R_{CV} + R_{CI} - G_{CV} - G_{CI}, \tag{10.2}$$

$$-\frac{1}{q}\frac{dJ_p}{dx} = R_{CV} + R_{IV} - G_{CV} - G_{IV}, \tag{10.3}$$

where J_n and J_p are the electron and hole current densities, respectively, and are given by the drift-diffusion transport equations. In Eqs. (10.2) and (10.3), G_{ij} and R_{ij} denote the optical carrier generation rate and recombination/relaxation rate between i-band and j-band, where C, V and I express CB, VB and IB, respectively. The continuity equation for the IB electron is reduced to the local balance equation for each QD layer by neglecting the electron transport within IB or QD confined states. As a result, the balance equation for a QD layer can be written as

$$G_{CI} - R_{CI} = G_{IV} - R_{IV}, \tag{10.4}$$

and determines the IB electron concentration n_I for each QD layer.

In the electron transition rate between the CB and the IB (R_{CI}), we assume two different processes expressed as

$$R_{CI} = R_{CI}^{Rad} + R_{CI}^{Th}, \tag{10.5}$$

where R_{CI}^{Rad} and R_{CI}^{Th} are the radiative and the thermal relaxation rates of electrons from the CB to the IB, respectively. Whilst R_{CI}^{Rad} can be determined [21,31], we define R_{CI}^{Th} as a function of relaxation time constant τ, and thus,

$$R_{CI}^{Th} = \frac{N_I}{\tau}\left[f_C(1-f_I) - \exp\left(-\frac{E_{CI}}{k_B T}\right) f_I(1-f_C)\right], \tag{10.6}$$

where N_I is the IB effective density of state which corresponds to the spin degeneracy and sheet density of QDs divided by QD height. Additionally,

$f_C = \exp[(\mu_C - E_C)/k_B T]$ where E_C is the CB edge, μ_C is the chemical potential of CB, k_B is the Boltzmann constant and T is the temperature, respectively. The IB electron occupancy f_I is given by

$$f_I = \frac{1}{\exp\left(\frac{E_I - \mu_I}{k_B T}\right) + 1}, \tag{10.7}$$

where E_I is the energy level of the IB, and μ_I is the chemical potential of the IB. Here, we parameterise τ to mimic a fast electron relaxation process from the CB to the IB, such as phonon relaxation.

The optical generation rates for each band-to-band transition depend on the optical absorption coefficients. Here, we employ the absorption coefficient of GaAs for the CB-VB transition [32], and the coefficients of CB-IB and IB-VB transitions are simplified as

$$\alpha_{CI}(E) = \begin{cases} 10^4\ \mathrm{cm}^{-1} & \text{if } E_{CI} \le E \le E_{IV}, \\ 0\ \mathrm{cm}^{-1} & \text{otherwise}, \end{cases} \tag{10.8}$$

$$\alpha_{IV}(E) = \begin{cases} 10^4\ \mathrm{cm}^{-1} & \text{if } E_{IV} \le E \le E_{CV}, \\ 0\ \mathrm{cm}^{-1} & \text{otherwise}, \end{cases} \tag{10.9}$$

where E is the photon energy. For the window layers, we employ the GaAs absorption coefficient and the absorption energy edge is shifted to the window layer's bandgap.

Let us consider a general case that the spectrum of absorption coefficients α_{CV}, α_{CI} and α_{IV} can overlap with each other, whilst those are simplified as defined as Eqs. (10.8) and (10.9). The optical carrier generation rates are weighted by their absorption coefficients since the incident photon flux is shared by the absorption for each band-to-band transition. The weighted function for CB-VB transition is defined as

$$W_{CV}(E) = \frac{\alpha_{CV}(E)}{\alpha_{CV}(E) + (1 - f_I)\alpha_{IV}(E) + f_I \alpha_{CI}(E)}. \tag{10.10}$$

In the same way, we define $W_{IB}(E)$ and $W_{CI}(E)$ for the IB-VB and CB-IB optical absorption rates, respectively. The photon flux density $\Phi(x, E)$ at x is expressed by

$$\Phi(x, E) = \Phi_0(E) F(E, x), \tag{10.11}$$

where Φ_0 is the incident photon flux. $F(E,x)$ is a decay factor due to the photon absorption by material and given by

$$F(E,x) = \exp\left[-\int_0^x \alpha_{CV}(E)dx' + \sum_l^{N_{QD}} \int_0^x f_I \alpha_{CI}(E)\delta(x'-x_l)dx' + \sum_l^{N_{QD}} \int_0^x (1-f_I)\alpha_{IV}(E)\delta(x'-x_l)dx' \right], \tag{10.12}$$

where x_l is the position of l-th QD layer, and N_{QD} is the number of those layers. It should be noted if one employs any advanced optical confinement techniques, then F will take a more complicated form. Finally G_{CV}, G_{IV} and G_{CI} are given by

$$G_{CV}(x) = \int \alpha_{CV} W_{CV} \Phi(x) dE, \tag{10.13}$$

$$G_{IV}(x) = \int \alpha_{IV}(1-f_I) W_{IV} \Phi(x) dE, \tag{10.14}$$

$$G_{CI}(x) = \int \alpha_{CI} f_I W_{CI} \Phi(x) dE. \tag{10.15}$$

Here, we omit the argument of E for simplicity. Therefore, $W_{ij}\Phi$ is an expression for the available photon flux for the i-j band transition.

Other calculation conditions are as follows: The schematic of employed SC structure is presented in Fig. 10.1 and the band diagram at thermal

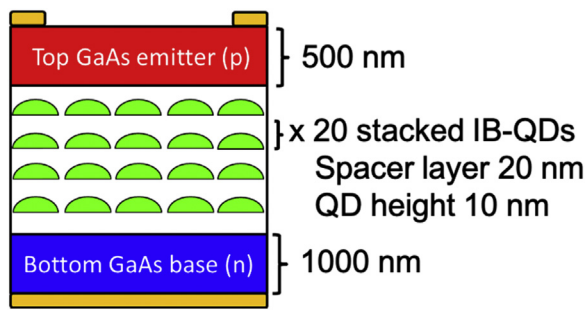


Figure 10.1 Schematic of the employed device structure. There are 20 and 120 nm thick window layers at the top and bottom emitters, respectively, though omitted in the figure for simplicity. The top and bottom window layers have 0.5 eV band offset for both CB and VB. The characteristics of SC without an IB is also plotted and denoted as Control. Acceptor density of 7.0×10^{16} cm^{-3} and donor density of 5×10^{16} cm^{-3} are assumed in the top and the bottom layers, respectively.

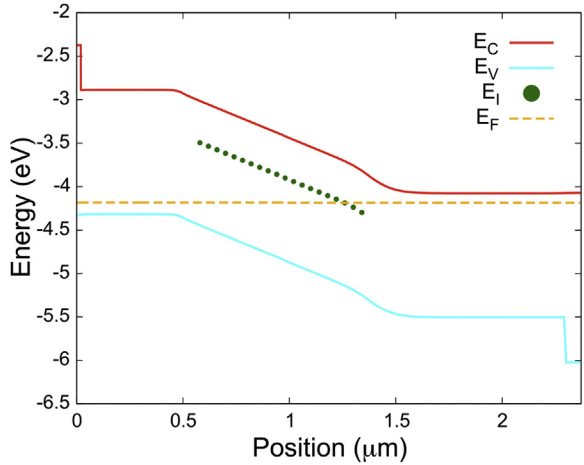


Figure 10.2 Calculated band diagram at thermal equilibrium condition at 300 K. E_C and E_V are the bottom and top band edges of CB and VB, respectively. E_I is the position of IB and E_F is the Fermi level.

equilibrium is shown in Fig. 10.2. We employ 20 and 120 nm thick window layers at both the front and the bottom sides of the cell, and CB and VB offset energies of 0.5 eV are assumed. In the calculation, GaAs material parameters at 300 K [33] are mainly employed in the host material and window layers. An acceptor density of 7.0×10^{16} cm^{-3} and a donor density of 5×10^{16} cm^{-3} are used for the top and the bottom layers, respectively. The energy gaps E_{CI} and E_{IV} are 0.43 and 0.95 eV, respectively, and the sheet density of QD layer is 10^{10} cm^{-2}. We employ radiative recombination/relaxation processes for each band-to-band transition and the coefficients of radiative recombination/relaxation rates are calculated by using absorption coefficients. The incident photon spectrum Φ_0 is assumed to be the same as that from a 5800 K blackbody for the J-V calculation and a monochromatic light for QE is assumed to have a constant photon flux density equivalent to 0.05 mA/cm^2.

2.3 Results and discussion

First, the simulated light J-V dependence on capture time τ under one sun illumination is presented in Fig. 10.3. The open-circuit voltage (V_{OC}) is significantly affected by the capture time τ. The fill factor (FF) and short-circuit current density (J_{SC}) are drastically degraded when $\tau < 10^{-11}$ s. If the

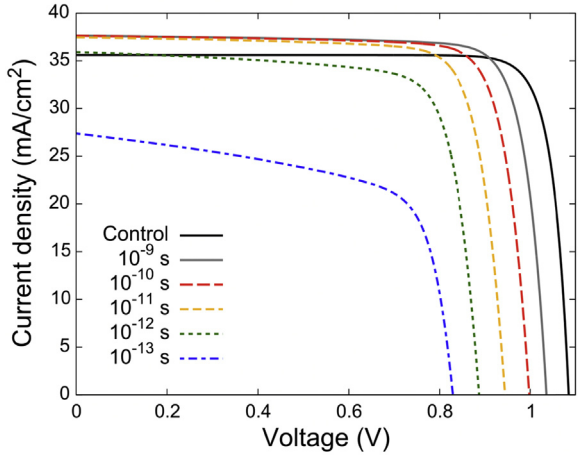


Figure 10.3 Calculated current–voltage characteristics depending on electron capture time constants by IB under one sun illumination.

lifetime τ becomes shorter, the optical excitation from CB to IB will also be smaller. In this case, the IB acts as an efficient recombination centre as is the case for the Shockley–Read–Hall (SRH) recombination.

Compared to the light *J-V* curve of the single-junction SC (Control in Fig. 10.3), the curve calculated for an ideal QD-IBSC ($\tau = 10^{-9}$ s) shows a larger J_{SC} due to the contribution from TSPA but a smaller V_{OC} than the Control cell. The reduction of V_{OC} is attributed to an increased recombination via the IB. The benefits of IBSC may seem to be small, but a further increase of J_{SC} and recovery of V_{OC} and *FF* are achievable by introducing a concentration of the sunlight. Therefore, IBSC exceeds the convergence of the Shockley–Queisser limit for the single-junction SCs. The detailed study can be found in Ref. [31].

Next, the QE dependence on τ is simulated and shown in Fig. 10.4. The QE spectrum is divided into three regions depending on the wavelength of photons as follows: The first region is in the range of 300 nm $\leq \lambda \leq$ 400 nm, where QE signals are insensitive to τ. The second region is in the range of 400 nm $\leq \lambda \leq \lambda_{CV}$, where λ_{CV} corresponds to the host material bandgap energy. The last range is $\lambda_{CV} \leq \lambda$, where only VB-IB optical excitation is available.

In the first range (300 nm $\leq \lambda \leq$ 400 nm), the incident light excites only the carriers in the top-emitter layer. Although the electrons optically generated in the top emitter layer must travel to the bottom contact, the

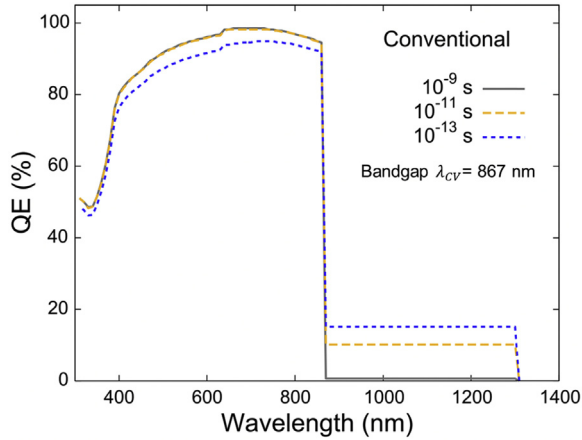


Figure 10.4 Calculated quantum efficiencies with different capture time constants τ. Note the host material bandgap (λ_{CV}) corresponds to 867 nm.

recombination in the IB and bottom emitter layers hardly occurs because few excess holes are generated in these regions. Therefore, the QE is considered to be unaffected by τ. It should also be noted that the shorter the wavelength of the incident light, the larger the absorption coefficient is in general. As a result, surface recombination at the top contact is generally responsible for the degradation of QE in this region.

In the other cases, the above scenario needs to be modified to take into account the holes generated in the IB region and the bottom emitter layer. For the second case, in the range $400\ \text{nm} \leq \lambda \leq \lambda_{CV}$, the smaller the QE becomes, the shorter the capture time constant τ is. This is understandable because when τ is short, the photo-generated electrons are efficiently captured by the IB and recombine with holes. In this case, the IB acts as an efficient recombination centre and degrades the QE.

In the last case, when the incident photon excites only the IB-VB transition ($\lambda_{CV} \leq \lambda$), the smaller the electron capture constant, the larger the QE becomes. This tendency is completely the opposite dependence to that found in the second case ($400\ \text{nm} \leq \lambda \leq \lambda_{CV}$). This behaviour can be explained as follows: Here, the incident photon only excites G_{IV}, and Eq. (10.4) is reduced to

$$-R_{CI} = G_{IV} - R_{IV}. \tag{10.16}$$

It requires $R_{CI} < 0$ to obtain a positive value of QE. This condition is satisfied as long as the quasi-Fermi level of IB (μ_I) is larger than that of

CB (μ_C). Additionally, the electron relaxation models employed here are based on the detailed balance theory where R_{CI}^{Rad} and R_{CI}^{Th} are zero in thermal equilibrium or where $\mu_C = \mu_I$, positive for $\mu_C > \mu_I$ and negative for $\mu_C < \mu_I$. Therefore, if τ is small and $R_{CI} \sim R_{CI}^{Th}$, then the net-generation rate in Eq. (10.16) is only given by the difference between G_{IV} and R_{IV}. R_{CI}^{Th} just follows the G_{IV} and R_{IV} profile and does not determine the net optical generation rate via IB. The IB electrons are then excited by thermal excitation. Otherwise, when $\mu_C < \mu_I$ and the relaxation time constant is large, R_{CI}^{Th} is negligible, then R_{CI}^{Rad} becomes the bottleneck though it is relatively small. Therefore, the net carrier transfer rate between the IB and VB, $(G_{IV} - R_{IV})$, becomes also quite small. As a result, the electrons photo-excited into the IB from the VB just recombine with holes and are hardly excited into CB.

This fact indicates that the QE values for $\lambda_{CV} < \lambda$ are negligible or almost zero if the SC exhibits an ideal IBSC operation, i.e., R_{CI}^{Rad} is dominant in R_{CI}. In this limit, the short-circuit current (I_{SC}) under the white light illumination could not be reproduced by integrating the product of QE and photon spectrum, although it is acceptable for single-junction SCs. Therefore, the QE measurement under the continuous IR light is required to analyse IBSC performance.

The calculated QE under IR illumination is presented in Fig. 10.5. The IR light is assumed to have a photon energy of 0.5 eV and power of 5.0 mW/cm^2, which can only excite the CB-IB transition. When compared to Fig. 10.4, significant changes in the QE are found in the range of

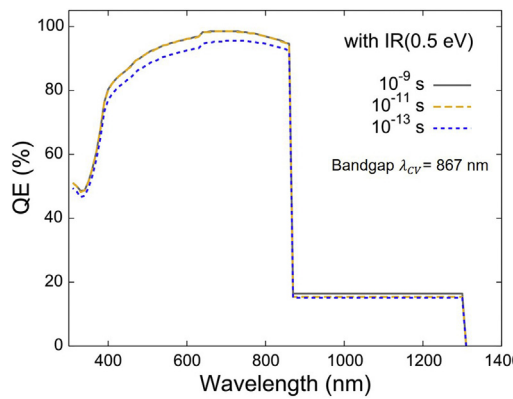


Figure 10.5 Calculated quantum efficiencies under IR light illumination (5.0 mW/cm^2) with different capture time constants τ.

$\lambda_{CV} \leq \lambda$. The QE is strongly enhanced by IR light for long capture time constants, such as $\tau = 10^{-9}$ s and 10^{-11} s, both giving values of QE exceeding those for $\tau = 10^{-13}$ s.

We take the difference, ΔQE, between QEs with and without IR light for each τ, in order to carry out a detailed analysis. The QE enhancements are also found in the range $\lambda_{CV} > \lambda$, as presented in Fig. 10.6. Note the vertical axis of Fig. 10.6 is in a logarithmic scale. It indicates that IR light suppresses the CB electron capturing as expected, and this effect becomes more significant as τ decreases. On the other hand, the QE enhancement by IR light becomes pronounced as τ increases when $\lambda_{CV} < \lambda$. It indicates that the photo-generated electrons by IB-VB excitation are efficiently extracted by IR light absorption through G_{CI} as long as τ is not too small.

We find another interesting ΔQE dependence on τ by comparing ΔQE in $\lambda_{CV} > \lambda$ with that in $\lambda_{CV} < \lambda$. When τ is long such as $\tau = 10^{-9}$ s, $\Delta QE(\lambda_{CV} < \lambda) > \Delta QE(\lambda_{CV} > \lambda)$. The opposite relation $\Delta QE(\lambda_{CV} < \lambda) < \Delta QE(\lambda_{CV} > \lambda)$, however, is found in the case with $\tau = 10^{-13}$ s. These findings show that we could quantify the cell potential for ideal IBSC operation by carrying out a comparison between $\Delta QE(\lambda > \lambda_{CV})$ and $\Delta QE(\lambda < \lambda_{CV})$. The quantification of the absorption coefficient α_{CI} by ΔQE would be difficult since ΔQE contains information about not only α_{CI} but also τ. However, it is useful to estimate the optical absorbance by the VB-IB transition.

The I_{SC} presented in Fig. 10.3 under one sun illumination cannot be reproduced by using the QE spectrum under IR light illumination. This

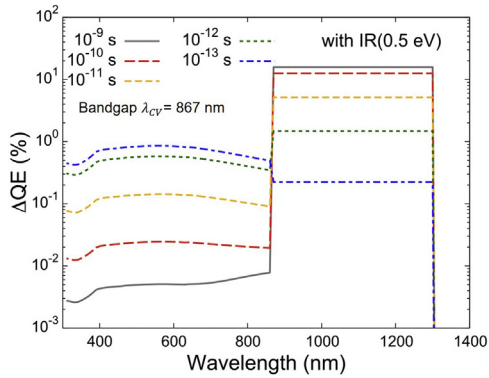


Figure 10.6 Difference of quantum efficiency ΔQE with and without IR light illumination (5.0 mW/cm^2). These results are the difference between Figs. 10.3 and 10.4.

is because the superposition principle is not acceptable for the IBSC, whereas it is valid in the single-junction SC, as pointed out above. The incident light employed in the QE measurement is weak enough to be considered as a perturbation to the device in equilibrium, whilst the sunlight modifies the IB electron concentration and the potential profile. Additionally, when τ is small, and sunlight illuminates the cell surface, electrons photo-generated by short-wavelength light ($\lambda \leq 400$ nm) are reduced by recombination through the IB since holes are generated by longer wavelength light in IB region, which does not occur under QE measurement. Hence, we should use the light J-V and QE measurements for different purposes. The former is to characterise the SC performance, and the latter is to estimate the potential of fabricated device as an IBSC. Furthermore, the ΔQE measurement is beneficial and necessary to properly understand the potential of IBSC.

3. Improvement of carrier collection efficiency in QDSCs using field-damping layers

3.1 Band structure engineering of QDSCs by using field-damping layers

As described in the previous section, the ideal operation of QD-IBSCs is that carriers generated in QDs due to photo-absorption are extracted by optical excitation (TSPA) whilst minimising thermal and electric field-assisted escape processes. It is thus desirable to place the QD layers into the region with a flat energy band. In this section, we introduce the means to control the electric field in the QD layers using a field-damping layer (FDL) and show the importance of the band structure on the operation of QDSCs [34,35]. Fig. 10.7 shows schematic diagrams of (a) p^+-i-n and (b) p^+-n^--i-n structure SCs with QDs inserted in the intrinsic layer. The order of doping concentrations for p^+, n and n^- is $\sim 10^{18}$, $\sim 10^{17}$ and $\sim 10^{16}$ cm^{-3}, respectively. In the structure shown in Fig. 10.7a, the field-assisted carrier escape process becomes dominant in the photocurrent production from QD layers because a strong electric field generated by the junction is applied to the QD layers. Since this carrier extraction is fundamentally different from the ideal QD-IBSC concept, QDs cause a decrease in V_{OC}. For the structure shown in Fig. 10.7b, an n^--FDL creates a band curvature which allows the formation of a flat-band potential in QD layers. In this case, the extraction process of the QD generated carriers would

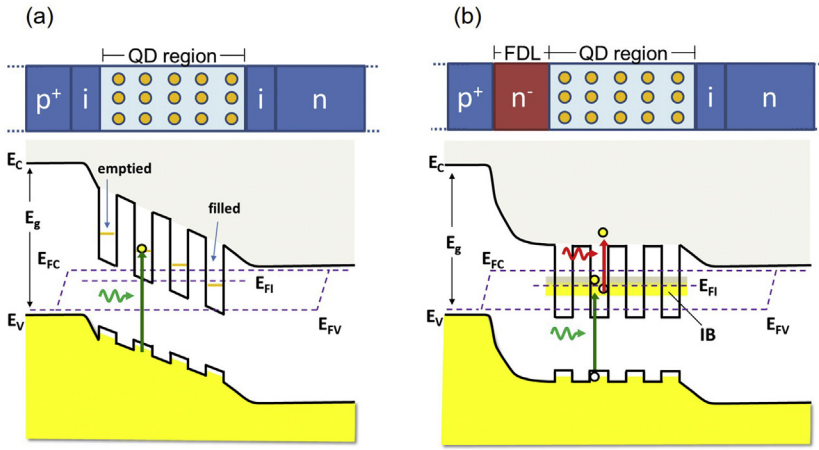


Figure 10.7 Schematic band diagram of QDSC (a) without and (b) with FDL. Light is incident from the left side of the drawing.

be limited to the thermal and optical transitions. This band structure is close to the ideal for operation of QD-IBSCs. Experimental results that demonstrate these concepts are described below.

Fig. 10.8a shows the schematic structure of QDSCs studied in our laboratory. The samples were fabricated by solid-source MBE. SCs have a

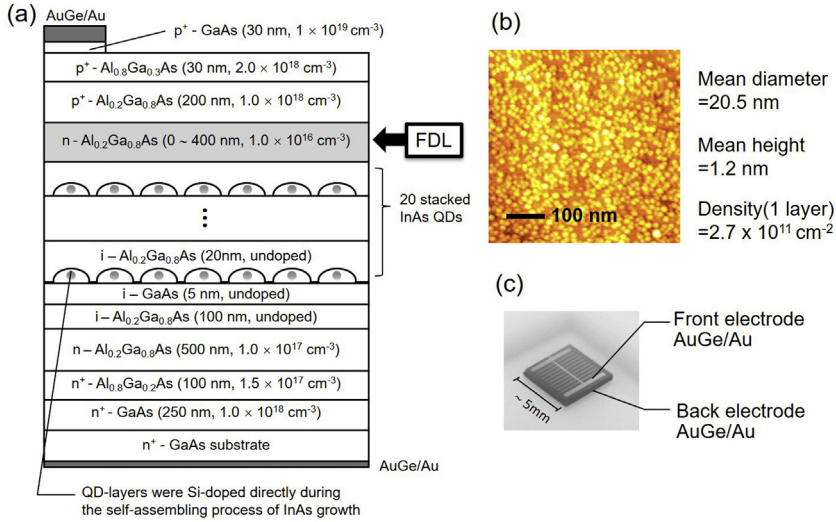


Figure 10.8 (a) Schematic structure of fabricated InAs/AlGaAs QDSCs. (b) AFM image of the InAs QDs grown on AlGaAs layer. (c) Photograph of the device after processing.

p-n^{-}*-i-n* structure using AlGaAs as a host material. The 20 stacked pairs of InAs QDs/$Al_{0.2}Ga_{0.8}As$ layers were grown in the intrinsic region at a growth temperature of 460°C. The InAs QDs were grown by S-K growth mode utilising the lattice mismatch. Fig. 10.8b shows the atomic force microscopy (AFM) image of the 2 MLs-thick InAs QDs grown on $Al_{0.2}Ga_{0.8}As$ layer. The mean diameter, height and sheet density of QDs are 20.5 nm, 1.2 nm and 2.7×10^{11} cm^{-2}, respectively. On the AlGaAs layer, the surface diffusion length of In adatoms is shorter, and as a result, relatively small and high-density QDs are obtained [36,37]. The QD layers were Si doped directly during the assembling stage of InAs growth [23,38]. The sheet density of Si doping was set to 2.7×10^{11} cm^{-2} per QD layer in order to dope approximately one Si atom per QD. This doping gives a better condition for TSPA to occur by partially occupying the IB formed by QDs with electrons. Fig. 10.8c shows a photograph of the cell. To evaluate the optical and electrical properties of the QDSCs, AuGe/Au electrodes were deposited for top and bottom Ohmic contacts. The size of the devices was approximately 5 mm × 5 mm. In the samples, the FDL inserted between p-emitter layer and multi-stacked QD layers can control the strength of the electric field applied to the QD layers by its thickness. Four types of cells with different thick FDL (0, 100, 250 and 400 nm) were prepared. Fig. 10.9 shows the calculated band diagram of QDSCs with FDLs. Although a strong electric field is applied to the QD layer for QDSC without FDL, the field is weakened as the FDL thickness increases. For the QDSC with a 400 nm-thick FDL, QD layers are placed in an almost flat-band region.

First, the effect of introducing the FDL on the electrical characteristics of QDSCs is studied. Fig. 10.10a shows the *J-V* curves under one sun (air-mass 1.5 global solar spectrum) illumination measured for 20-layer stacked InAs/AlGaAs QDSCs. The J_{SC} for QDSC with a thick FDL is higher than that of QDSC with a thinner FDL. This is due to the dependence of photoabsorption on the thickness of the FDL. Fig. 10.10b plots the V_{OC} and *FF* which peak as a function of the thickness of the FDL, the changing values being due to the control of the electric field strength by the FDL. The QDSCs with nonzero FDL show higher *FF* and V_{OC} compared to the sample without the FDL. This is because the introduction of the FDL reduces the SRH recombination by locating the QD layers away from the strong electric field region [39,40]. In this experiment, the highest *FF* and V_{OC} are obtained for QDSC with 250 nm-thick FDL. This result can be understood from the band structure around the V_{OC} operating point as shown in Fig. 10.11. For the QDSC with 400 nm-thick FDL, the electric field is

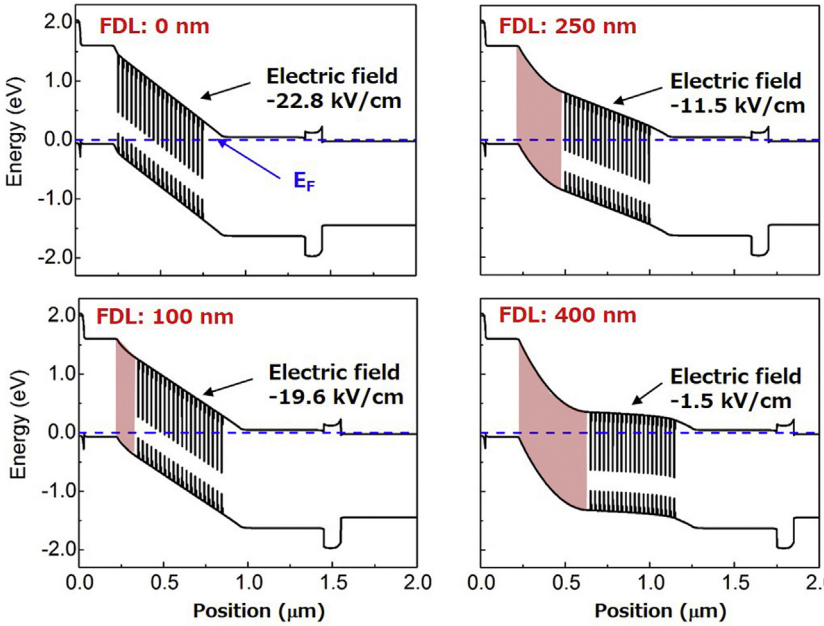


Figure 10.9 Calculated equilibrium energy band structures of 20-stacked InAs/AlGaAs QDSCs with different thickness of FDL.

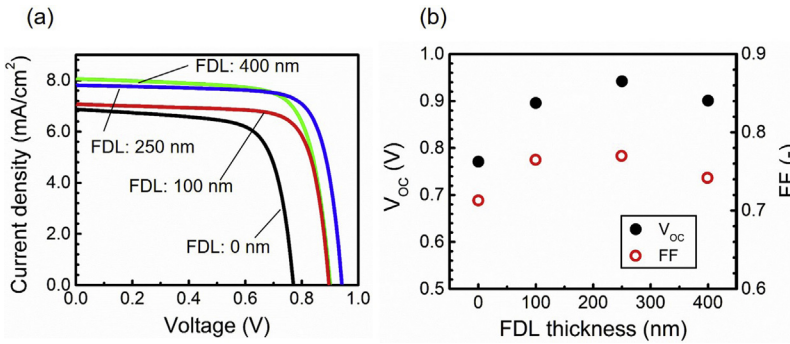


Figure 10.10 (a) Light current–voltage curves measured for 20-stacked InAs/AlGaAs QDSCs with different thickness of FDL. (b) Change of FF and V_{OC} with respect to the thickness of FDL.

excessively suppressed by the FDL, resulting in the generation of unintentional forward electric bias between the FDL and QD layer. Consequently, the collection of photo-generated carriers decreases and *FF* drops to a lower value than that of QDSC with 250 nm-thick FDL. In addition, band

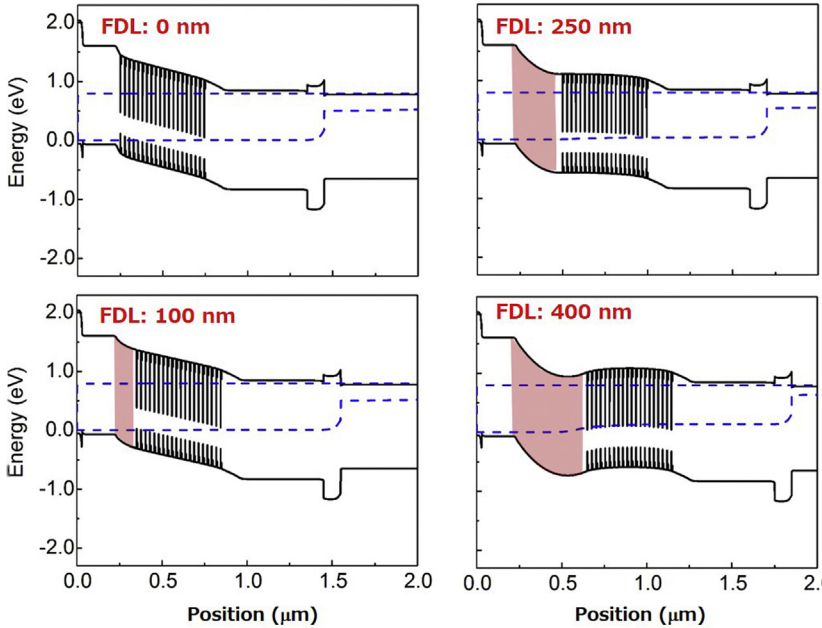


Figure 10.11 Calculated energy band structures of 20-stacked InAs/AlGaAs QDSCs with different thickness of FDL at forward bias voltage of 0.8 V.

bending due to excessive electric field suppression of FDL increases SRH recombination in the FDL, as shown in Fig. 10.12 leading to a decrease in V_{OC}.

Next, the optical properties of QDSCs with the FDL are characterised. Fig. 10.13 shows the external quantum efficiency (EQE) spectra of the QDSCs with FDLs measured under short-circuit conditions. The

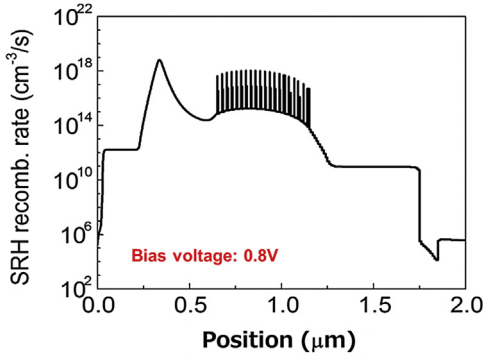


Figure 10.12 Calculated SRH recombination rate of 20-stacked InAs/AlGaAs QDSCs with 400 nm-thick FDL at forward bias voltage of 0.8 V.

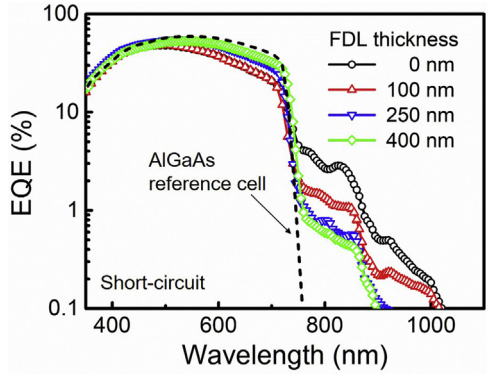


Figure 10.13 EQE spectra measured for 20-stacked InAs/AlGaAs QDSCs with different thicknesses of FDL.

measurement was performed by irradiating with monochromatic photons at $10^{16}/cm^2$. The EQE response in the wavelength region shorter than 750 nm is due to photo-absorption in the AlGaAs host material. On the other hand, the signal in the wavelength region longer than 750 nm is attributed to the contribution from 20-stacked InAs QD layers. In this measurement, since the SC is irradiated with monochromatic light, TSPA hardly occurs. Thus, the EQE signal related to the QD layers means the photocurrent is extracted by processes other than TSPA, which is undesirable in QD-IBSC operation. It can be seen that the EQE response of the QD layers decreases as the thickness of the FDL increases. This result indicates that field-assisted extraction of generated carriers can be reduced by the introduction of an FDL.

In order to detect the production of photocurrent as a direct result of the optical transition of electrons from IB to CB, a measurement system shown in Fig. 10.14 was used [23,41]. Light source (1) provides continuous-wave monochromatic photons, which generate electron-hole pairs in InAs/AlGaAs QDs. Light from a Xe lamp with a filter of air-mass 1.5 global solar spectrum placed in front of it passes through an appropriate set of filters that altogether only allow the IR region of $\lambda > 1064$ nm to pass, this is shown as light source (2) in Fig. 10.14. The illumination of low energy photons from this IR source can therefore only pump electrons from the IB to the CB, but do not have sufficient energy to pump electrons from the VB to the IB or directly from the VB to the CB. Thus, the photocurrent production due to TSPA via QDs can be estimated by measuring the difference between EQE with and without the IR illumination of the light source (2). The difference

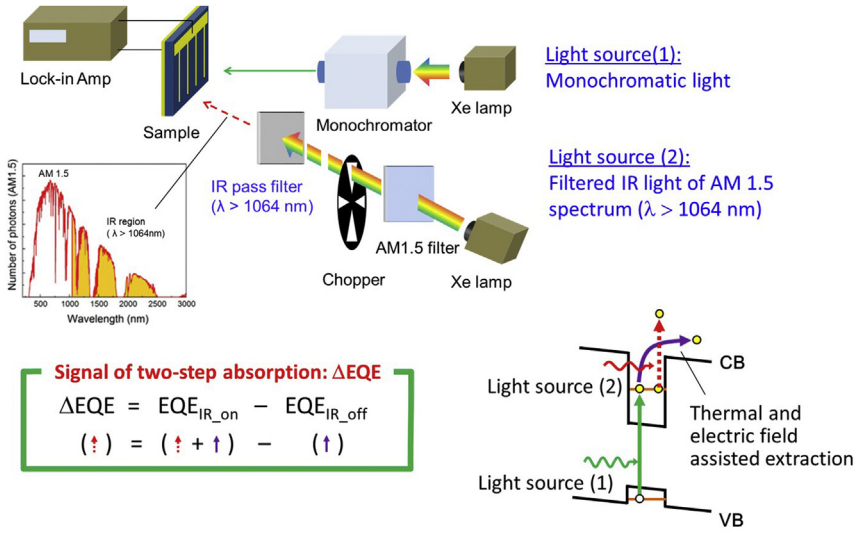


Figure 10.14 Measurement setup used to characterise photocurrent production as a direct result of optical transitions of electrons from IB to CB. One of the light sources provides continuous-wave monochromatic photons, which generate electron-hole pairs in IB (Light source (1)). Light from Xe lamp with AM 1.5 filter placed at the exit passes through a set of filters that altogether allow only IR photons of $\lambda > 1064$ nm to be transmitted (Light source (2)). The illumination of low energy photons from this IR source can then only pump the electrons from IB to CB.

in EQE (ΔEQE) by IR illumination was detected by using a lock-in amplifier synchronised with an optical chopper set at 85 Hz. Fig. 10.15 shows the difference in EQE values, ΔEQE, taken with (EQE_{IR_on}) and without (EQE_{IR_off}) IR light illumination measured for QDSCs with the FDL.

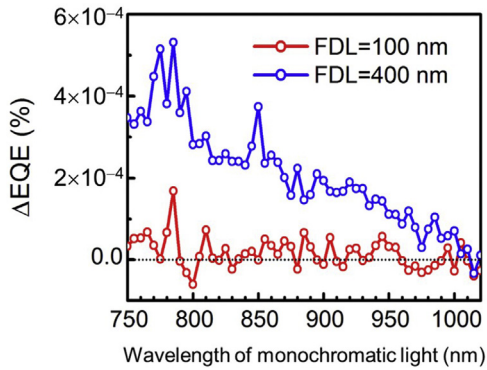


Figure 10.15 Difference in EQE values, ΔEQE, between with (EQE_{IR_on}) and without (EQE_{IR_off}) IR light illumination measured for InAs/AlGaAs QDSCs with 100 nm and 400 nm-thick FDL.

The measurement was carried out at room temperature. A clear photocurrent production due to TSPA is observed in the QDSC with 400 nm thick FDL, whereas no measurable photocurrent difference is found in the QDSC with 100 nm thick FDL. This result indicates that TSPA can be increased due to the formation of flat-band QD region.

3.2 Effect of wide-gap material capping on QDSCs using FDL

The electric field across QD layers becomes weak in QDSCs using the FDL. This affects the collection efficiency of carriers absorbed by the host material. Here, to clearly show this effect, a description is given of using an SC incorporating InGaAs QDs of a larger size. Fig. 10.16 shows the performance of an InGaAs/AlGaAs QDSC with a 250 nm-thick FDL. The basic structure of the QDSC is the same as that of Fig. 10.8a, except that the InAs material of the QD layers is replaced with $In_{0.4}Ga_{0.6}As$. $In_{0.4}Ga_{0.6}As$ QDs are larger in size than those of InAs because they have a smaller lattice mismatch with the GaAs substrate. Hence, for InGaAs/AlGaAs QDSCs, although the amount of photo-absorption by QDs increases, the probability that photo-generated carriers are trapped by the QDs also increases. From the light *J-V* curve of InGaAs/AlGaAs QDSC with 250 nm-thick FDL shown in Fig. 10.16a, the *FF* is estimated to be 0.37, indicating a low carrier collection efficiency. In Fig. 10.16b, solid and dashed line shows the EQE spectra of the FDL-inserted AlGaAs SC with and without InGaAs QDs,

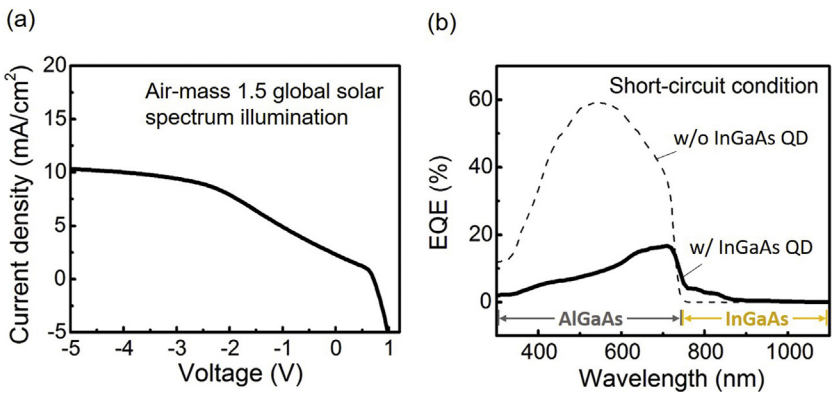


Figure 10.16 (a) Light current–voltage curve measured for 20-stacked InGaAs/AlGaAs QDSC with 250 nm-thick FDL. (b) EQE spectrum measured for 20-stacked InGaAs/AlGaAs QDSCs with 250 nm-thick FDL. Dashed line indicates the EQE spectrum of control AlGaAs cell (without InGaAs QDs).

respectively. It can be seen that the presence of InGaAs QD layers significantly reduces the EQE on FDL-inserted QDSC. This is because InGaAs QD layers trap most of the photo-carriers generated in AlGaAs host material.

In order to prevent the electrons from flowing into QDs, we studied the formation of a wide-gap fence structure in CB, as shown in Fig. 10.17a. In this structure, the path of electrons in the QD layers is limited; hence, an improvement in the carrier collection efficiency can be expected. It is necessary to selectively grow barrier layers with wide-gap material directly on the QD, and material selection is important to realise the band structure shown in Fig. 10.17a. Fig. 10.17b shows the bandgap energy and lattice constant of III-V compound semiconductors. AlAs is one of the materials with the energy gap larger than that of the host material. However, two-dimensional growth occurs like embedding the QDs when AlAs was deposited on top of the QDs since the lattice constant is almost the same as the

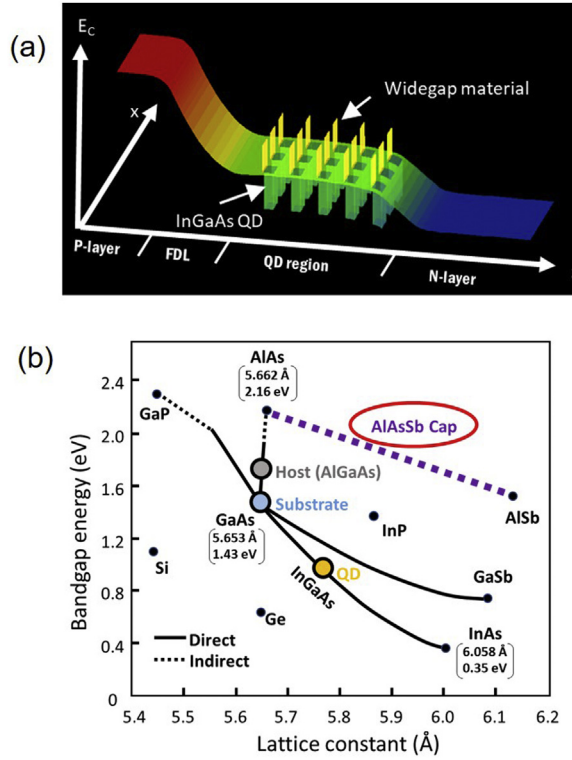


Figure 10.17 (a) Schematic CB diagram of FDL-inserted QDSC with wide-gap fence structure. (b) Bandgap energy and lattice constant of III-V compound semiconductors.

substrate material [42]. In contrast, since AlAsSb has a larger lattice constant than the substrate material (GaAs), 3D growth is possible using the SK mode. Thus, we attempted to fabricate a fence structure by capping QDs with AlAsSb using the S-K growth mode. Fig. 10.18a and b show AFM images of the $In_{0.4}Ga_{0.6}As$ QDs with and without AlAsSb cap layer grown on the AlGaAs layer at the substrate temperature of 500°C. The deposition thickness of InGaAs and AlAsSb was 6 and 3 MLs, respectively. The InGaAs QDs were formed at a growth rate of 1 ML/s under the As flux of 1.4×10^{-3} Pa, whilst the AlAsSb cap layer was fabricated at a growth rate of 0.15 ML/s under the As flux of 4.6×10^{-4} Pa and Sb flux of 2.3×10^{-5} Pa. It is observed that the shape of the 3D islands is changed by capping with AlAsSb. In particular, the structure extends in the [1–10] direction after AlAsSb capping. The cross-sectional scanning transmission electron microscope image in Fig. 10.18c shows a 3D structure of AlAsSb formed so as to cover InGaAs QDs. By stacking this structure, the band diagram close to that of Fig. 10.17 can be realised.

Finally, the effect of the introduction of the AlAsSb fence structure on the performance of QDSCs is demonstrated. Fig. 10.19 shows the light *J-V* curves measured for FDL-inserted QDSCs with and without AlAsSb fence structure. The current density of both samples indicates almost the same value at the applied reverse bias voltage of −5 V. This result means

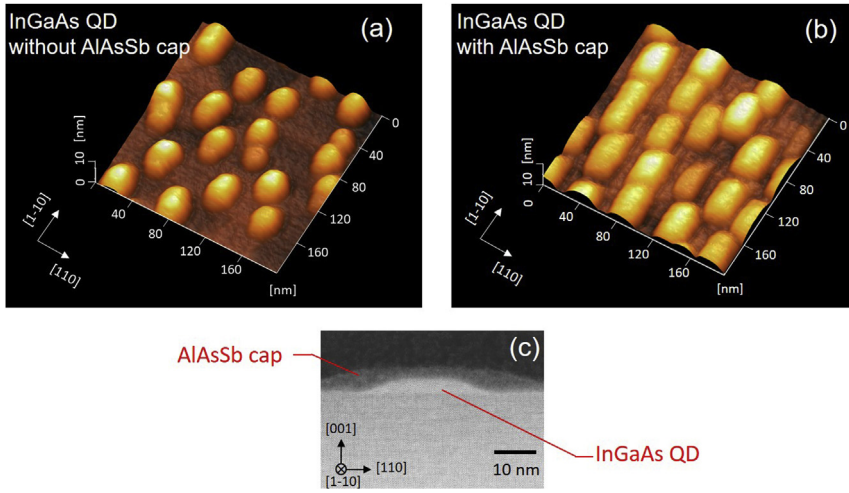


Figure 10.18 (a) AFM image of InGaAs QDs. (b) AFM image of InGaAs QDs capped by AlAsSb layer. (c) Cross-sectional STEM image of InGaAs QDs with AlAsSb cap layer.

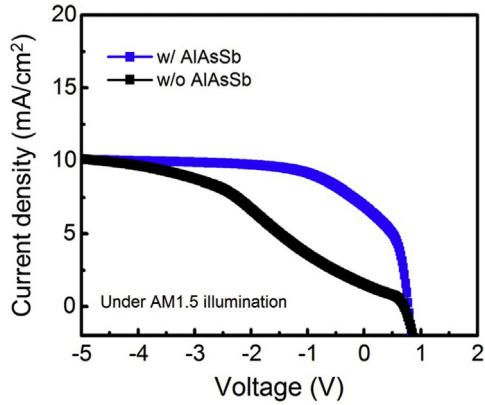


Figure 10.19 Light current–voltage curves measured for FDL-inserted InGaAs/AlGaAs QDSC with and without AlAsSb fence structure.

that when a high electric field is applied to the QD layers, both samples can avoid trapping of photo-carriers into QDs. However, it is necessary to obtain a high carrier collection efficiency without applying a high electric field to the QD layers according to the operation principle of QD-IBSCs. In this case, the carrier collection efficiency significantly affects the *FF* of QDSCs. From the light *J-V* curves, it is clear that the AlAsSb fence structure improves *FF* from 0.37 to 0.48 for 20-stacked InGaAs/AlGaAs QDSCs. Assuming that all the photo-carriers are collected at a reverse bias voltage of −5 V, the carrier collection efficiency at the operating voltage is 11.4% for the QDSC without the fence structure. In comparison, the carrier collection efficiency is improved to 46.3% for the QDSC with an AlAsSb fence structure. Furthermore, V_{OC} also improves from 0.714 to 0.787 V by the introduction of fence. These results suggest that the AlAsSb fence structure suppresses carrier capture into QDs.

4. FTIR spectroscopy of TSPA processes in QDSCs

4.1 Two-step photon absorption spectroscopy

In QDSCs, below host bandgap energy near-IR photons can generate photocurrent via inter-band VB to IB transition [15,16,43]. However, in the view of IBSC operation principle, it is not favourable because those photo-carriers escape from QD confinement states to CB continuum by thermionic emission or field-assisted tunnelling. It diminishes quasi-Fermi level splitting between the IB and CB; therefore, the V_{OC} of QDSC degrades.

On the other hand, if IR photons excite photo-carriers in QD confinement states to CB continuum via TSPA, high conversion efficiency would be possible in QDSCs by achieving gain in photocurrent and voltage preservation simultaneously. TSPA was demonstrated experimentally by using IR-biased photocurrent spectroscopy. When illuminating a QDSC simultaneously with two IR light sources with below bandgap energy photons, photo-carriers generated in QDs can be optically pumped out to the CB owing to the second step IR photon absorption. Photocurrent enhancement via TSPA in QDSCs was demonstrated at low temperature at first [26] and later at room temperature [23]. Furthermore, the IB to CB transition has been spectrally resolved as a function of second step IR photon energy [28,44], and extreme broadband photocurrent spectroscopy clarified full spectrum features of the QDSCs from visible to mid-IR spectral range [45]. In TSPA, however, thermionic emission or field-assisted tunnelling of photo-carriers from QD confinement states diminishes the gain in photocurrent at room temperature. It is indispensable to understand device optimisation strategy for high-efficiency gain via TSPA and estimate possible gain in photocurrent at actual operation conditions.

4.2 FTIR photocurrent spectroscopy of In(Ga)As QDSCs

IR photocurrent response via TSPA in In(Ga)As QDSCs was investigated by applying FTIR photocurrent spectroscopy. A schematic illustration of the optical configuration is depicted in Fig. 10.20. White-light halogen lamp was used as the bias light source to generate photo-carriers in QDs by inter-band transitions. An IR lamp equipped in an FTIR spectrometer was the IR light source. Photocurrent signal from the QDSCs was amplified by a transimpedance pre-amplifier and was detected as the interferogram synchronised with the interferometer in the FTIR spectrometer. Fourier transformation converted the interferogram to IR photocurrent intensity spectra. On the other hand, IR light power intensity spectrum was obtained by using a wavelength insensitive DLaTGS pyroelectric detector, which was calibration standard for the spectral intensity. The photocurrent signal intensity spectra were normalised by the IR light power intensity spectrum; as the result, IR photocurrent spectra of the second step IR photon absorption of TSPA could be evaluated. The benefits of FTIR technique are quick acquisition, high wavelength resolution and broadband spectral measurement range, compared with conventional monochromator scanning. FTIR

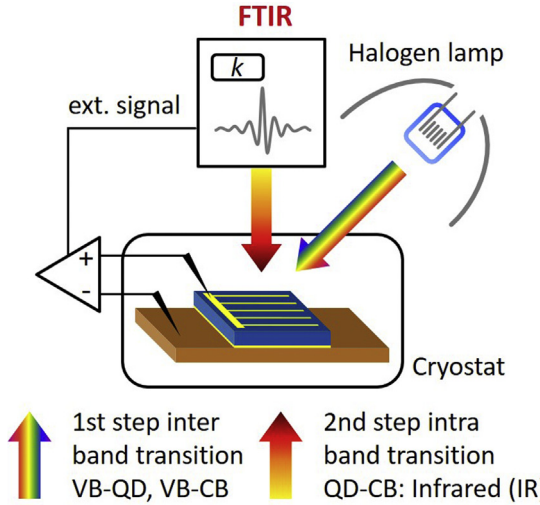


Figure 10.20 Schematic illustration of FTIR photocurrent spectroscopy setup.

photocurrent spectroscopy is suitable to perform measurements for series of samples to unveil universal relationship or optimisation strategy for QDSCs.

Low temperature IR photocurrent spectra in In(Ga)As QDSCs were investigated by FTIR photocurrent spectroscopy. IR photocurrent spectra at respective temperature on (a) InAs/GaAs, (b) InAs/$Al_{0.2}Ga_{0.8}As$ and (c) InGaAs/$Al_{0.3}Ga_{0.7}As$ QDSCs were summarised in Fig. 10.21. In each

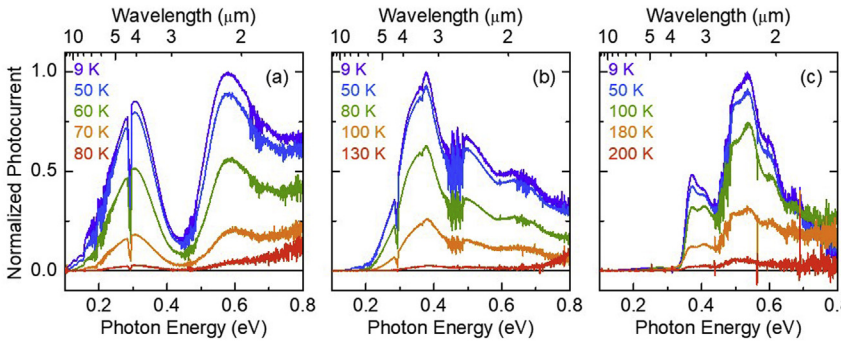


Figure 10.21 Temperature dependence of IR photocurrent spectra under white-light bias irradiation on (a) InAs/GaAs, (b) InAs/$Al_{0.2}Ga_{0.8}As$ and (c) $In_{0.4}Ga_{0.6}As/Al_{0.3}Ga_{0.7}As$ QDSCs. Photocurrent spectra were measured by FTIR photocurrent spectroscopy. The measurements were performed at short-circuit conditions. The dip at 0.3 eV is due to atmospheric CO_2 absorption band. *Reprinted with permission from Tamaki, et al. Proc. SPIE 9743; 2016, p. 974318. Copyright 2016 SPIE.*

QDSC, photocurrent intensity was normalised at the peak value obtained at 9 K. IR photocurrent intensity monotonically decreased with increasing operation temperature, indicating thermionic emission of photo-carriers diminish gain in photocurrent via TSPA. However, In(Ga)As QDSCs with wide bandgap host materials, AlGaAs, could indicate IR response at higher temperature than the narrow bandgap GaAs host solar cells. Furthermore, the IR absorption edge of the InAs/GaAs QDSC was smaller than that of wide bandgap AlGaAs host solar cells.

The IR absorption edge of the QD to CB photocurrent production was determined by fitting the IR photocurrent absorption tail at lowest temperature (9 K). In Fig. 10.22, the vertical axis was the square root of photocurrent times photon energy. The intersections between the fitting result (black solid lines) and zero base line for the vertical axis were the extracted IR absorption edge. In InAs/GaAs, InAs/$Al_{0.2}Ga_{0.8}As$ and InGaAs/$Al_{0.3}Ga_{0.7}As$ QDSCs, the absorption edges were estimated to be 0.095, 0.185 and 0.324 eV, respectively.

Temperature dependence of the peak photocurrent intensity was summarised in Fig. 10.23. Threshold temperature of the IR photocurrent response was defined as the temperature at photocurrent intensity with 10^{-2} compared with 9 K, where the QD confinement of photo-carriers competed against thermionic emission processes. In InAs/GaAs, InAs/$Al_{0.2}Ga_{0.8}As$ and InGaAs/$Al_{0.3}Ga_{0.7}As$ QDSCs, the threshold temperature were evaluated as 85, 139 and 217 K, respectively.

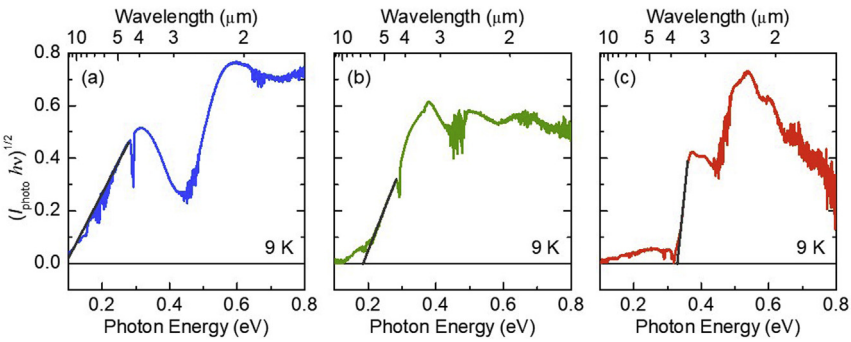


Figure 10.22 IR absorption edge fitting analysis of IR photocurrent spectra under white-light bias irradiation on (a) InAs/GaAs, (b) InAs/$Al_{0.2}Ga_{0.8}As$ and (c) $In_{0.4}Ga_{0.6}As$/$Al_{0.3}Ga_{0.7}As$ QDSCs. Intersections between the fitting results (black solid lines) and zero base line are the estimated IR absorption edges. *Reprinted with permission from Tamaki, et al. Proc. SPIE 9743; 2016, p. 974318. Copyright 2016 SPIE.*

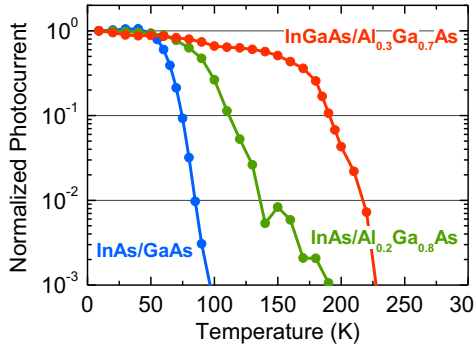


Figure 10.23 Temperature dependence of the peak intensity of IR photocurrent under white-light bias irradiation on InAs/GaAs (blue [dark grey in print version]), InAs/$Al_{0.2}Ga_{0.8}As$ (green [light grey in print version]), and $In_{0.4}Ga_{0.6}As/Al_{0.3}Ga_{0.7}As$ (red [grey in print version]) QDSCs, and the peak energies were 0.306, 0.383 and 0.370 eV, respectively. *Reprinted with permission from Tamaki, et al. Proc. SPIE 9743; 2016, p. 974318. Copyright 2016 SPIE.*

The IR absorption edge and threshold temperature were increased with increasing host material bandgap energy. Series of In(Ga)As QDSCs were investigated and the results were summarised in Fig. 10.24. As clearly shown, 'universal linear relationship' in the IR absorption edge and threshold temperature was clarified regardless of the detail device structure differences. Nevertheless, there were some differences even between the QDSCs with same host materials. It suggested that the IR absorption edge was a good parameter to represent the competition between photo-carrier

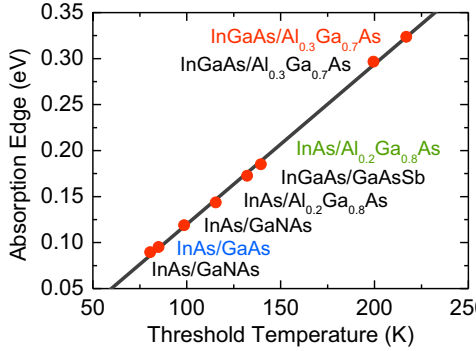


Figure 10.24 Universal linear relationship between threshold temperature and IR absorption edge in In(Ga)As QDSCs. *Reprinted with permission from Tamaki, et al. Proc. SPIE 9743; 2016, p. 974318. Copyright 2016 SPIE.*

confinements in QDs and the thermionic emission or field-assisted tunnelling processes. By extrapolating the universal linear relationship towards high temperature region, the IR absorption edge of 0.459 eV was predicted to obtain 295 K threshold temperature, which was 0.15 eV larger than the highest value in InGaAs/$Al_{0.3}Ga_{0.7}As$ QDSCs in current study. In detailed balance theory, the optimum bandgap combination in IBSCs under sun light concentration has been calculated to be 1.9 and 0.7 eV, respectively, for VB-CB and IB-CB transitions [44,46]. Therefore, the predicted IR absorption edge is within the range of optimum bandgap combinations. The universal linear relationship revealed strategy for device optimisation to achieve efficient TSPA at ambient conditions.

4.3 2D photocurrent excitation spectroscopy of In(Ga)As QDSCs

Here, TSPA process is investigated by using 2D photocurrent excitation spectroscopy. Monochromatic visible to near-IR light and a mid-IR light source equipped with an FTIR spectrometer simultaneously illuminated on the QDSCs. The monochromatic light injected photo-carriers in QDs by the first step inter-band transitions, and the mid-IR light excited photo-carriers out of the QD confinement states to CB continuum via the second step photon absorption. The photocurrent signal from QDSCs was amplified by a pre-amplifier and the interferogram synchronised with the interferometer in FTIR spectrometer was Fourier transformed to get intensity spectra at mid-IR region. Finally, photocurrent spectral maps were obtained by scanning the monochromatic light wavelengths. The photon flux spectra from the FTIR spectrometer and the monochromator were measured by wavelength insensitive pyroelectric detectors, and divided the 2D photocurrent spectral maps to obtain the relative photocurrent response for the same photon flux excitation. Moreover, absolute photocurrent intensity was calibrated by the near-IR EQE of the QDSCs under study at room temperature. As the result of above measurement and data analysis, 2D-ΔEQE maps for both the first step and the second step transitions on TSPA have been determined in detail.

2D-ΔEQE spectral maps (per IR photon, logarithmic scale) were obtained by 2D photocurrent excitation spectroscopy on InAs QDSCs at 9 K as shown in Fig. 10.25. The spectral weight was systematically blue shifted by incorporating Al due to increase in the band offset between the

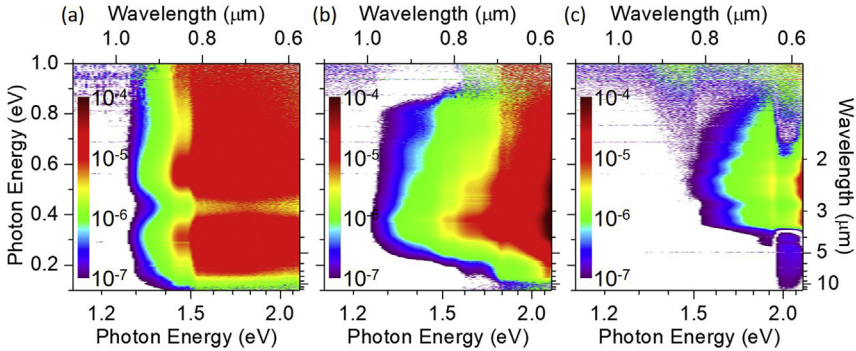


Figure 10.25 2D-ΔEQE spectral maps (per IR photon, logarithmic scale) obtained by 2D photocurrent excitation spectroscopy on (a) InAs/GaAs, (b) InAs/$Al_{0.2}Ga_{0.8}As$ and (c) $In_{0.4}Ga_{0.6}As/Al_{0.3}Ga_{0.7}As$ QDSCs at 9 K. Horizontal and vertical axes indicate first and second step photon absorption energy, respectively.

QD confinement states and CB edges. These whole spectral features indicated the potential for photocurrent production via TSPA, which can be utilised to estimate the spectral matching condition for sun light spectrum. Compared with the ΔEQE and EQE in QDSCs, the orders of magnitude difference were still remained. One reason is that the absorption coefficient for second step photon absorption was still not sufficient, which should improve by fabricating high-density multi-stacked QDs or applying light management technique [47]. The other reason is due to the lack of photo-generation in QDs. Light concentration could be a possible solution to realise photo-filling effect in QDs [48].

Nevertheless, 2D photocurrent excitation maps could be calculated by using ΔEQE maps and standard solar spectrum, as shown in Fig. 10.26. Atmospheric absorption in IR region resulted in zero photocurrent production bands in corresponding region. The estimated J_{SC} by integrating the 2D photocurrent maps were about tens of nA/cm^2, which were about six orders of magnitude smaller than that of the conventional bulk photovoltaics (tens of mA/cm^2). A feasible strategy to overcome this difference would be as follows: (i) ×10 by QD density, (ii) ×10 by light management, (iii) ×100 by QD carrier lifetime and (iv) ×100 by light concentration. This improvement is still challenging, but still in practical range for a target value to realise high-efficiency QDSCs by IBSC operation.

We have investigated IR photocurrent spectra of TSPA in In(Ga)As QDSCs by using FTIR photocurrent spectroscopy. To control the

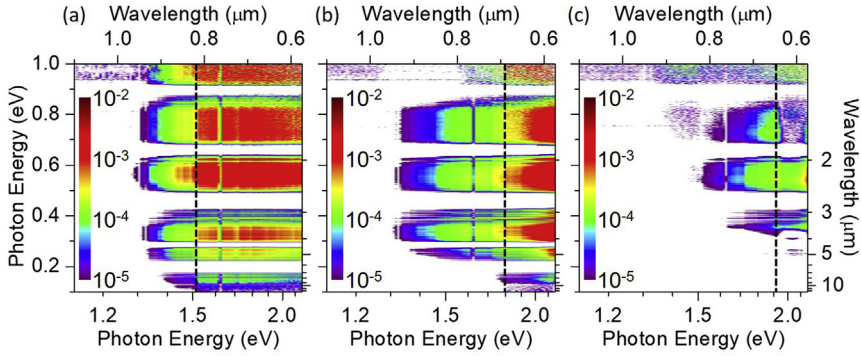


Figure 10.26 2D photocurrent excitation maps (mA/cm^2/eV, logarithmic scale) obtained by ΔEQE spectral map (a) InAs/GaAs, (b) InAs/$Al_{0.2}Ga_{0.8}As$ and (c) $In_{0.4}Ga_{0.6}As/Al_{0.3}Ga_{0.7}As$ QDSCs under AM1.5D irradiation. Horizontal and vertical axes indicate first and second step photon absorption energy, respectively. Broken lines indicate bandgap energy of the respective host materials.

photo-carrier confinement in In(Ga)As QDs, $Al_xGa_{1-x}As$ ($x = 0$, 0.2, 0.3) barrier layers were utilised as the host materials. The IR absorption edge of the QD to CB transition was evaluated at low temperature by analysing the low energy tail of the FTIR photocurrent spectra. The threshold temperature of the TSPA in In(Ga)As QDSCs was determined by observing temperature dependence of IR peak photocurrent intensity. A universal linear relationship between the threshold temperature and the absorption edge of IR photo-response clarified the quantitative target value for higher temperature operation and revealed the optimisation strategy to achieve high efficiency TSPA at ambient conditions. Furthermore, 2D photocurrent spectroscopy was conducted to investigate 2D-ΔEQE spectral maps and 2D photocurrent excitation maps under standard solar illumination. There were still orders of magnitude differences remained compared with the ideal condition, which need to be improved. Nevertheless, the predicted values were the minimum requirement to achieve high-efficiency TSPA at room temperature, and the quantitative target value is now clarified, and actual efficiency gain in QDSCs will be available in near future.

5. Conclusion

This chapter is devoted to identify the issues and factors that have been and yet need to be explored and challenged both theoretically and experimentally in order to achieve high-efficiency QD-based IBSCs. The

understanding and control of carrier dynamic processes, particularly the TSPA is paramount to realisation of high efficiency. Significant efforts and a good progress have been made in the device physics as well as practical demonstration of QD-IBSCs as presented. These achievements have been made possible largely due to advancement of high quality III-V semiconductor thin-film as well as quantum nanostructure material growth and processing and characterisation technologies.

Strong optical absorption, long carrier lifetimes, suppressed trapping of carriers back into QDs are prerequisite in IBSC materials and we have pointed out the necessity of optimising these processes in terms of material and cell structure. The difficulty in the IB approach for photovoltaics stems from the dual nature of IB: it must have optimum carrier occupancy such that it can participate in optical absorptions which promote electrons into and out-of the IB. If the occupancy in the IB is too low, the second optical transition (from the IB to the CB) is strongly bleached. Conversely, if the occupancy is too high, the first optical absorption (from the valence to the IB) is strongly bleached. Current QD materials suffer from low occupancy in the IB and need to be further optimised. For this, method to realise ideally half-filled IB states to maximise photocurrent generation by TSPA, such as by the doping of IB region, photofilling by light concentration and quantum ratchet structures [49,50], is currently heavily explored. In addition to further improvements in the material quality, the control of absorption matching in IB materials would also improve the efficiency of IBSCs in the near future.

Acknowledgements

We would like to gratefully acknowledge Professors David A. Ritchie of University of Cambridge, J-F. Guillemoles of Centre national de la recherche scientifique (CNRS), N. Ekins-Daukes of University of New South Wales, S. Tomić, T. Sogabe, N. Miyashita of RCAST, The University of Tokyo and T. Kita of Kobe University for their collaboration.
This work is supported in part by the New Energy and Industrial Technology Development Organization (NEDO) and Ministry of Economy, Trade and Industry (METI), Japan.

References

[1] Green MA. Third generation photovoltaics: advanced solar energy conversion. Springer Series in Photonics; 2003.
[2] Würfel P. Physics of solar cells: from basic principles to advanced concepts. Wiley-VCH; 2009.
[3] Shockley W, Queisser HJ. Detailed balance limit of efficiency of p-n junction solar cells. J Appl Phys 1961;32:510.
[4] Nozik AJ. Quantum dot solar cells. Physica E 2002;14:115.

[5] Nozik AJ. Multiple exciton generation in semiconductor quantum dots. Chem Phys Lett 2008;457:3.
[6] Luque A, Martí A. Increasing the efficiency of ideal solar cells by photon induced transitions at intermediate levels. Phys Rev Lett 1997;78:5014.
[7] Tomić S, Jones TS, Harrison NM. Absorption characteristics of a quantum dot array induced intermediate band: implications for solar cell design. Appl Phys Lett 2008; 93:263105.
[8] Beattie NS, See P, Zoppi G, Ushasree PM, Duchamp M, Farrer I, Ritchie DA, Tomić S. Quantum engineering of InAs/GaAs quantum dot based intermediate band solar cells. ACS Photonics 2017;4:2745.
[9] Martí A, Cuadra L, Luque A. Partial filling of a quantum dot intermediate band for solar cells. IEEE Trans Electron Dev 2001;48:2394.
[10] Solomon GS, Trezza JA, Marshall AF, Harris Jr JS. Vertically aligned and electronically coupled growth induced InAs islands in GaAs. Phys Rev Lett 1996;76:952.
[11] Okada Y, Shiotsuka N, Komiyama H, Akahane K, Ohtani N. Multi-stacking of highly uniform self-organized quantum dots for solar cell applications. In: Proceedings of the 20th European photovoltaic solar energy conference. WIP; 2005. p. 51.
[12] Akahane K, Yamamoto N, Tsuchiya M. Highly stacked quantum-dot laser fabricated using a strain compensation technique. Appl Phys Lett 2008;93:041121.
[13] Popescu V, Bester G, Hanna MC, Norman AG, Zunger A. Theoretical and experimental examination of the intermediate-band concept for strain-balanced (In,Ga)As/Ga(As,P) quantum dot solar cells. Phys Rev B 2008;78:205321.
[14] Hubbard SM, Cress CD, Bailey CG, Raffaelle RP, Bailey SG, Wilt DM. Effect of strain compensation on quantum dot enhanced GaAs solar cells. Appl Phys Lett 2008;92:123512.
[15] Oshima R, Takata A, Okada Y. Strain compensated InAs/GaNAs quantum dots for use in high-efficiency solar cells. Appl Phys Lett 2008;93:083111.
[16] Okada Y, Oshima R, Takata A. InAs/GaNAs strain-compensated quantum dot solar cell. J Appl Phys 2009;106:024306.
[17] Oshima R, Okada Y, Takata A, Yagi Y, Akahane K, Tamaki R, Miyano K. High-density quantum dot superlattice for application to high-efficiency solar cells. Physica Status Solidi (C) 2011;8:619.
[18] Shoji Y, Oshima R, Takata A, Okada Y. Structural properties of multi-stacked self-organized InGaAs quantum dots grown on GaAs (311)B substrate. J Cryst Growth 2010;312:226.
[19] Shoji Y, Narahara K, Tanaka H, Kita T, Akimoto K, Okada Y. Effect of spacer layer thickness on multi-stacked InGaAs quantum dots grown on GaAs (311)B substrate for application to intermediate band solar cell. J Appl Phys 2012;111:074305.
[20] Strandberg R, Reenaas TW. Drift-diffusion model for intermediate band solar cells including photofilling effects. Prog Photovoltaics Res Appl 2011;19:21.
[21] Yoshida K, Okada Y, Sano N. Self-consistent simulation of intermediate band solar cells: effect of occupation rates on device characteristics. Appl Phys Lett 2010;97: 133503.
[22] Sullivan JT, Simmons CB, Buonassisi T, Krich JJ. Targeted search for effective intermediate band solar cell materials. IEEE J Photovoltaics 2015;5:212.
[23] Okada Y, Morioka T, Yoshida K, Oshima R, Shoji Y, Inoue T, Kita T. Increase of photocurrent by optical transitions *via* intermediate quantum states in direct-doped InAs/GaNAs strain-compensated quantum dot solar cell. J Appl Phys 2011;109: 024301.
[24] Sablon KA, Little JW, Mitin V, Sergeev A, Vagidov N, Reinhardt K. Strong enhancement of solar cell efficiency due to quantum dots with built-in charge. Nano Lett 2011; 11:2311.

[25] Kada T, Asahi S, Kaizu T, Harada Y, Tamaki R, Okada Y, Kita T. Efficient two-step photocarrier generation in bias-controlled InAs/GaAs quantum dot superlattice intermediate-band solar cells. Sci Rep 2017;7:5865.
[26] Martí A, Antolín E, Stanley CR, Farmer CD, López N, Díaz P, Cánovas E, Linares PG, Luque A. Production of photocurrent due to intermediate-to-conduction-band transitions: a demonstration of a key operating principle of the intermediate-band solar cell. Phys Rev Lett 2006;97:247701.
[27] Elborg M, Noda T, Mano T, Jo M, Sakuma Y, Sakoda K, Han L. Voltage dependence of two-step photocurrent generation in quantum dot intermediate band solar cells. Sol Energy Mater Sol Cell 2015;134:108.
[28] Tamaki R, Shoji Y, Okada Y, Miyano K. Spectrally resolved interband and intraband transitions by two-step photon absorption in InGaAs/GaAs quantum dot solar cells. IEEE J Photovoltaics 2015;5:229.
[29] Creti A, Tasco V, Cola A, Montagna G, Tarantini I, Salhi A, Al-Muhanna A, Passaseo A, Lo- mascolo M. Role of charge separation on two-step two photon absorption in InAs/GaAs quantum dot intermediate band solar cells. Appl Phys Lett 2016;108:063901.
[30] Yoshida K, Okada Y. Numerical analysis for carrier capturing process on quantum-dot intermediate-band solar cells. [unpublished].
[31] Yoshida K, Okada Y, Sano N. Device simulation of intermediate band solar cells: effects of doping and concentration. J Appl Phys 2012;112:084510.
[32] Casey HC, Miller BI, Pinkas E. Variation of minority-carrier diffusion length with carrier concentration in GaAs liquid-phase epitaxial layers. J Appl Phys 1973;44:1281.
[33] Selberherr S. Analysis and simulation of semiconductor devices. Springer-Verlag; 1984.
[34] Martí A, Antolín E, Cánovas E, López N, Linares PG, Luque A, Stanley CR, Farmer CD. Elements of the design and analysis of quantum-dot intermediate band solar cells. Thin Solid Films 2008;516:6716.
[35] Shoji Y, Tamaki R, Datas A, Martí A, Luque A, Okada Y. Effect of field damping layer on two step absorption of quantum dots solar cells. The 6th world conference on photovoltaic energy conversion. [Kyoto, Japan].
[36] Guimard D, Nishioka M, Tsukamoto S, Arakawa Y. High density InAs/GaAs quantum dots with enhanced photoluminescence intensity using antimony surfactant-mediated metal organic chemical vapor deposition. Appl Phys Lett 2006;89:183124.
[37] Shoji Y, Okada Y. Effect of external bias on multi-stacked InAs/AlGaAs quantum dots solar cell. In: 40th IEEE photovoltaic specialists conference; 2014. Denver, Colorado.
[38] Inoue T, Kido S, Sasayama K, Kita T, Wada O. Impurity doping in self-assembled InAs/GaAs quantum dots by selection of growth steps. J Appl Phys 2010;108:063524.
[39] Morioka T, Okada Y. Dark current characteristics of InAs/GaNAs strain-compensated quantum dot solar cells. Physica E 2011;44:390.
[40] Driscoll K, Bennett MF, Polly SJ, Forbes DV, Hubbard SM. Effect of quantum dot position and background doping on the performance of quantum dot enhanced GaAs solar cells. Appl Phys Lett 2014;104:023119.
[41] Shoji Y, Akimoto K, Okada Y. Self-organized InGaAs/GaAs quantum dot arrays for use in high-efficiency intermediate-band solar cells. J Phys D: Appl Phys 2013;46: 024002.
[42] Tutu FK, Lam P, Wu J, Miyashita N, Okada Y, Lee K-H, Ekins-Daukes NJ, Wilson J, Liu H. InAs/GaAs quantum dot solar cell with an AlAs cap layer. Appl Phys Lett 2013; 102:163907.
[43] Luque A, Martí A, Stanley C, López N, Cuadra L, Zhou D, Pearson JL, McKee A. General equivalent circuit for intermediate band devices: potentials, currents and electroluminescence. J Appl Phys 2004;96:903.

[44] Tamaki R, Shoji Y, Okada Y, Miyano K. Spectrally resolved intraband transitions on two-step photon absorption in InGaAs/GaAs quantum dot solar cell. Appl Phys Lett 2014;105:073118.
[45] Datas A, López E, Ramiro I, Antolín E, Martí A, Luque A, Tamaki R, Shoji Y, Sogabe T, Okada Y. Intermediate band solar cell with extreme broadband spectrum quantum efficiency. Phys Rev Lett 2015;114:157701.
[46] Okada Y, Ekins-Daukes NJ, Kita T, Tamaki R, Yoshida M, Pusch A, Hess O, Phillips CC, Farrell DJ, Yoshida K, Ahsan N, Shoji Y, Sogabe T, Guillemoles J-F. Intermediate band solar cells: recent progress and future directions. Appl Phys Rev 2015;2:021302.
[47] Mellor A, Luque A, Tobias I, Marti A. The feasibility of high-efficiency InAs/GaAs quantum dot intermediate band solar cells. Sol Energy Mater Sol Cells 2014;130:225.
[48] Strandberg R, Reenaas TW. Photofilling of intermediate bands. J Appl Phys 2009;105: 124512.
[49] Yoshida M, Ekins-Daukes NJ, Farrell DJ, Phillips CC. Photon ratchet intermediate band solar cells. Appl Phys Lett 2012;100:263902.
[50] Vaquero-Stainer A, Yoshida M, Hylton NP, Pusch A, Curtin O, Frogley M, Wilson T, Clarke E, Kennedy K, Ekins-Daukes NJ, Hess O, Phillips CC. Semiconductor nanostructure quantum ratchet for high efficiency solar cells. Commun Phys 2018;1:7.

CHAPTER ELEVEN

Monolithic III–V quantum dot lasers on silicon

Jae-Seong Park, Mingchu Tang, Siming Chen and Huiyun Liu*
Department of Electronic and Electrical Engineering, University College London, London, United Kingdom
*Corresponding author: huiyun.liu@ucl.ac.uk

1. Introduction

With ever-growing global data traffic which is being accelerated by emergence of cloud networks and 5G communication, the development of ultra-fast and low-cost optical transmission in data centres has attracted substantial interest recently [1,2]. In this regard, Si photonics, using optical rather than electrical signals, is of great technical importance since it provides not only versatile optical functions on a single-chip but also low-cost and large-scale dense integration [3–5]. In addition, the Si-on-Insulator (SOI) platform generally adopted in Si photonics is compatible with advanced Si manufacturing and complementary metal oxide semiconductor (CMOS) technology [6]. Because of these advantages, Si photonics technology has been developed at an unprecedented rate over several decades, and 100G 4-channel coarse wavelength division multiplexing (CWDM) optical transceivers using Si photonics are commercially available now [7].

Currently available Si photonics, however, has been suffering from the absence of native light source, namely the laser, which is a critical limitation impeding dense integration and cost-effective fabrication. This is because Si has an indirect bandgap hindering efficient radiative recombination and hence lasing operation [8]. Although Si Raman lasers and Ge-based alloy lasers have been demonstrated, the performance of group IV-based lasers is still too poor to be implemented in photonic integrated circuits (PICs) and is unlikely to outperform III–V lasers in the foreseeable future [9–12]. In contrast, III–V materials with direct bandgap offer outstanding electrical and optical properties, enabling III–V lasers to be widely employed as light sources in many applications [13]. The fabrication of optical components such as modulators [14,15], waveguides [16], (de)/multiplexers [17,18] and detectors [19] is well established in the SOI platform and the integration of III–V lasers has emerged as the most promising solution to address the

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00009-5
© 2021 Elsevier Ltd.
All rights reserved.

absence of native light source in Si photonics. Current Si-based PICs generally use heterogeneous integration of high-performance III–V quantum-well (QW) lasers grown on native substrates, i.e., transferring laser devices grown on III–V semiconductor wafers onto the SOI platform [20,21]. Among these techniques, plasma-assisted bonding of a III–V die onto an SOI wafer has been widely employed, which allows evanescent coupling between III–V lasers and passive Si waveguides [22]. In this approach, III–V material grown on its native substrate is first bonded to the SOI wafer, the native substrate is then removed and the laser devices are fabricated.

Even though Si photonics with the heterogeneous integration of III–V QW lasers has achieved remarkable progress, the number of components in a single PIC is far less than that of InP-based monolithic PICs and the fabrication cost is still high [23]. This can be attributed to inherent limitations of the heterogeneous approach, such as poor scalability and the use of III–V native substrates which are much smaller and cost significantly more than Si substrates [24]. In order to truly exploit all the benefits of Si photonics, it is therefore imperative to achieve the direct integration of monolithic III–V lasers on the SOI platform, enabling cost-effective fabrication and dense integration. Notwithstanding the recent advances in the monolithic growth of III–V on Si, however, the deployment of monolithic III–V lasers on Si-based PICs is still challenging because the growth of III–V materials on Si introduces a lot of defects, such as threading dislocations (TDs) and anti-phase boundaries (APBs), due to significant material differences between III-Vs and Si [25,26]. In addition, thick buffer layers, essentially used to reduce the density of defects during the monolithic growth, make it difficult to couple the emitted light from lasers into Si waveguide on the SOI surface.

With the increasing importance of monolithic III–V growth on Si to realise integration with Si photonics, III–V lasers employing quantum dots (QDs) as active regions have recently been extensively investigated because the quantum confinement effect of the QDs offers many advantages such as temperature stability and low threshold current density [27,28]. Compared with QW structures, in particular, QDs provide robust tolerance to TDs [29], which is of great benefit for the monolithic growth of III–V on Si as a high density of TDs is inevitably generated during the growth. In other words, TDs propagating into the active region can impair only a limited number of QDs, enabling the rest of the QDs to contribute to the optical gain. Owing to these revolutionary benefits of QDs, III–V QD lasers grown

on Si are regarded as the most suitable Si-based light sources for the monolithic integration of Si photonics.

In this chapter, we focus on the progress of III–V/Si heteroepitaxy and III–V QD lasers on Si, which covers various approaches for resolving growth issues and introduces the current status of different types of III–V QD lasers on Si. Finally, we discuss the potential future directions of III–V QD lasers on Si for monolithic integration in Si photonics.

2. Advantages of quantum dot lasers on silicon

2.1 Semiconductor quantum dots

QDs are very small zero-dimensional nanostructures, whose size varies from a few to tens of nanometres. Along with the development of a new class of materials, semiconductor QDs such as InAs or InP, which can be used as an active layer of optoelectronic devices, have been extensively investigated due to their artificial atom-like properties. Arakawa [28] first theoretically proposed the use of III–V QDs as an active region in semiconductor lasers in 1982. For typical semiconductor lasers, very thin QW heterostructure effectively confines the carriers within the active region, resulting in modification of the allowed electronic states. In other words, the modified step-like density of state (DOS) causes a reduced thermal energy distribution of carriers, as opposed to the continuous distribution in bulk semiconductors, as shown in Fig. 11.1a and b. In the case of the QD active region, the carriers are strongly localised within nanometre-sized dots, causing a series of delta function-like DOS and strongly suppressed carrier diffusion in active region, as can be seen in Fig. 11.1c.

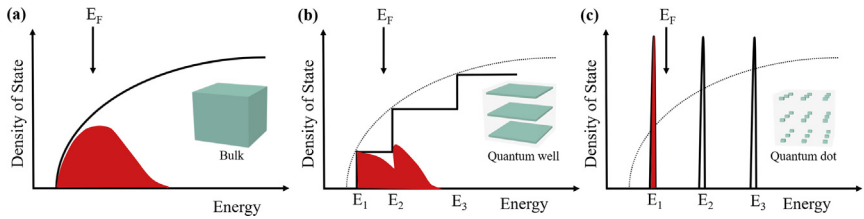


Figure 11.1 Density of states, *D(E)*, and density of occupied states per unit volume and energy, *N(e)*, corresponding to (a) bulk, (b) quantum well and (c) quantum dot semiconductors. The estimated *N(E)* is the product of *D(E)* and Fermi–Dirac distribution function, *f(E)*, at room temperature. The enclosed region of *N(E)* indicates the electron density, which shows the thermal energy distribution of carriers.

These effects contribute to the far more reduced thermal energy distribution of carriers than the QW structure, enabling temperature-insensitive characteristics.

2.2 Benefits of quantum dots over quantum wells in Si-based lasers

As previously mentioned, QD lasers could potentially provide temperature-insensitive operation which stems from the delta function-like DOS and large inter-sub-band separation between the ground state and the excited states in the conduction band of the QDs, whereas in the QW lasers, the carriers continuously occupy higher energy state as the temperature increases. Also, both the short carrier lifetime of the ground state and atom-like DOS are advantageous for population inversion, contributing to a very low threshold current (density) compared with QW lasers [30].

In addition, QD lasers are superior to QW lasers as regards dynamic properties. The linewidth enhancement factor (α_H-factor) is one of the important parameters used to explain the dynamic behaviour of lasers. The α_H-factor describes the degree of carrier-induced variation of the refractive index n of the active layer, which is defined as $\alpha_H = -4\pi/\lambda\,(\mathrm{d}n/\mathrm{d}N)(\mathrm{d}g/\mathrm{d}N)^{-1}$ where N and g are the carrier density and optical gain, respectively [31]. In general, a large value of α_H results in detrimental effects to the spectral linewidth, sensitivity to optical feedback, chirp and so on. In case of QD lasers, a perfect Gaussian energy distribution of QDs leads to a symmetric gain spectrum and differential gain. As a result, the differential refractive index change calculated via the Kramers–Kronig relation is exactly zero at the peak position [32]. In real QD lasers, however, various factors including the plasma effect, the inhomogeneous size distribution of dots and crystal defects result in a nonzero α_H value. Nonetheless, many experimental results have demonstrated a very low α_H-factor of about 0.5 and even 0.13 for p-doped QD lasers [33], which is considerably lower than that of QW lasers which typically have a value of 2 or higher [31].

QD lasers also possess a very low relative intensity noise (RIN) [34–36] and a strong tolerance to optical feedback [37–39], originating from a near-zero α_H value. Since optical feedback normally introduces undesirable effects such as increased noise, linewidth broadening and coherence collapse, optical isolators are commonly employed to reduce optical feedback [37]. In this regard, QD lasers are a very promising light source for isolator-free next-generation optical links and PICs.

Together with these fundamental advantages, the use of QD active regions provides defect-tolerable properties [27,40,41] and resultant long lifetimes [42,43] which are particularly favourable for Si-based lasers. Considering that the monolithic growth of III–V materials on Si substrates normally generates a high density of defects, the defect-tolerable III–V QD lasers are much more beneficial for monolithic Si-based lasers than QW lasers whose performance is significantly degraded in the presence of a high density of defects. For QW lasers on Si, the first AlGaAs double-heterostructure lasers monolithically grown on Si substrates with pulse operation at 77 K [44] triggered considerable attention for demonstrating III–V lasers on Si. Soon after, the first room temperature (RT) continuous-wave (cw) operation of AlGaAs/GaAs QW heterostructure lasers on Si was demonstrated by Deppe et al. [45] in 1987. Despite these pioneering developments of III–V QW lasers on Si substrates, the performance of III–V QW lasers on Si still suffers from a large threshold current density and short lifetime [46–48], mainly due to the high density of defects. On the other hand, since the growth condition and characterisation of InAs ultra-thin layers on GaAs was first reported in 1985 [49], the growth of QDs has been extensively investigated over several decades [32,50–57], leading to the first demonstration of an AlGaAs/GaAs-based GaAs QD laser on Si in 1998 [58]. With rapid advances in the growth techniques and QD laser structures, the first demonstration of InAs QD lasers on Si at the telecommunication wavelength of 1.3 μm was reported by Wang et al. [59] in 2011. The InAs/GaAs QD laser on Si at 1.3 μm produced a threshold current density (J_{th}) of 725 A/cm^2 and an output power of 26 mW at RT, attributed to the optimised growth temperature of the initial GaAs nucleation layer.

Recently, Liu et al. [60] investigated the experimental and theoretical differences between III–V QD and QW lasers grown on on-axis Si with the same density of TDs (5×10^7 cm^{-2}). In this report, the QD lasers on Si achieved a low J_{th} of 173 A/cm^2, high single facet output power of >100 mW at RT, and maximum operating temperature of 105°C, whilst the QW lasers on Si showed no lasing operation at RT, as shown in Fig. 11.2a. In Fig. 11.2b, the theoretical calculation showed that the QD laser can achieve excellent lasing behaviour with low J_{th} even in the presence of high density of TDs. In contrast, the QW lasers show very high threshold current density and require a very low density of TDs to obtain lasing operation, as shown in Fig. 11.2c. This indicates that, for a high-performance monolithic QW laser on Si, the density of TDs should be reduced to a similar level as that of the native substrate. Consequently, it was theoretically

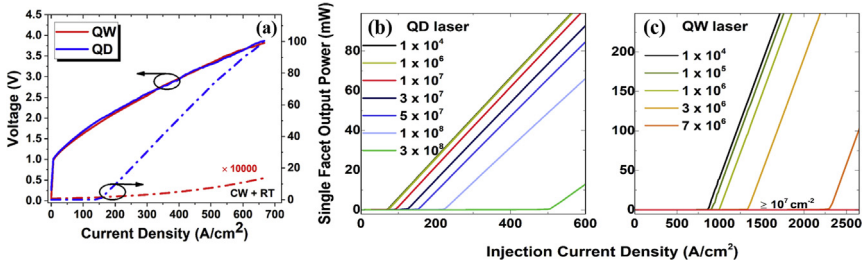


Figure 11.2 (a) Typical room temperature *L-I-V* characteristics of QD and QW lasers grown on on-axis (001) Si under the same growth conditions. Calculated *L-I* curves for an uncoated 50 μm × 3 mm (b) QD and (c) QW lasers according to the density of threading dislocation. *Reproduced with permission from [60] © 2019 IEEE.*

and experimentally confirmed that the QD lasers are ideal candidates for monolithic Si-based light sources.

3. Heteroepitaxial growth of III–V on silicon

Even though heterogeneous integration of III–V lasers on Si has been successfully demonstrated, monolithic integration is regarded as a long-term goal for low-cost and dense integration of photonic components. In order to realise high-performance Si-based monolithic III–V laser and, hence, leverage the full benefits of monolithically integrated Si photonics, the direct growth of high-quality III–V materials on Si is an important prerequisite. Unfortunately this approach poses critical problems arising from the substantial mismatch of material properties between III–V and Si, such as the coefficient of thermal expansion (CTE), lattice constant and polarity [20]. For instance, the lattice mismatch between GaAs (InP) and Si is about 4% (8%), and the CTE of GaAs ($6.6 \times 10^{-6}\,K^{-1}$) is larger than that of Si ($2.6 \times 10^{-6}\,K^{-1}$) by a factor of about 2.5.

3.1 Challenges for heteroepitaxial growth

(1) The formation of APBs originates from different crystalline polarity. During the growth of polar III-Vs on the nonpolar Si substrate, monoatomic steps on Si surface result in anti-phase domain disorder within the III–V materials [26,61]. APBs indicate that there is an incorrectly bonded plane, either III–III or V–V bonding, separating two III–V domains. In addition, the incorrect bonds are doubly charged defects, creating highly compensated semiconductor behaviour

and thus compromising the electronic properties of the epitaxial layer [62]. This planar defect also serves as a nonradiative recombination centre and leakage path, degrading the performance of laser devices and even preventing lasing operation at all.

(2) The mismatched lattice constant of 4% (8%) between GaAs (InP) and Si introduces the build-up of strain energy in the pseudomorphic layer, which is proportional to the thickness of the epilayer. If the III–V epitaxial layer grown on Si exceeds a certain thickness, the so-called critical thickness, the accumulated strain energy is released through the generation of misfit dislocations (MDs) along the heterogeneous interface. Because MDs cannot be terminated within the crystal lattice, MDs reach the edge of wafer or preferentially create TDs at the end of the MDs as they propagate towards the surface. It is well known that dislocations introduce electronic states in the semiconductor band gap, acting as nonradiative recombination centres, as well as providing a favourable path for defect diffusion [63].

(3) The large difference in CTE between III–V and Si ($6.6 \times 10^{-6}\,K^{-1}$ and $2.6 \times 10^{-6}\,K^{-1}$ for GaAs and Si, respectively) and consequent thermal stress result in the formation of thermally generated cracks during the cool-down process from the optimised GaAs growth temperature ($\sim$580°C) to RT at the end of wafer growth. The presence of thermal cracks introduces multiple deleterious effects on the quality of the III–V layers and device performance, e.g., scattering centres for light propagation, an electrical shorting path and restrictions on the possible total thickness of the epitaxial layers ($\sim$7 μm) [64].

3.2 Solutions for high-quality III–V/Si epitaxy

3.2.1 Approaches for APB-free III–V growth on on-axis Si

It is well known that growing III–V materials on offcut Si substrates with various angles from 4 degrees to 7 degrees effectively suppresses the formation of APBs [65–68]. By utilising these offcut Si substrates, high-quality III–V layers grown on Si have been widely achieved with a low threading dislocation density (TDD) and without APBs [59,69–75]. However, offcut Si substrates are incompatible with the SOI platform which uses on-axis (001) Si and a CMOS process. In order to combine the monolithic III–V light source with other photonic components, it is therefore crucial to grow high-quality III–V materials on on-axis Si substrates without APBs. Accordingly, a great deal of effort has been focussed on the epitaxial growth

of III–V layers on CMOS-compatible on-axis (001) Si substrates without APBs. To date, there are several mature strategies including patterned Si substrates [76–78] and intermediate GaAs [62,79] or GaP [80–83] buffer layers.

3.2.1.1 Patterned Si substrates

Li et al. [76] investigated GaAs growth on v-grooved (001) Si substrates. To obtain an APB-free GaAs-on-v-groove Si (GoVS) template, n-type on-axis (001) Si substrates coated in patterned SiO_2 were anisotropically etched using a KOH solution to form v-grooved structures exposing {111} crystal planes and, subsequently, GaAs nanowires were selectively grown on the v-grooved Si by using metal-organic chemical vapor deposition (MOCVD). After removing the SiO_2 pattern, a 1 μm-thick GaAs thin film was regrown at 600°C coalescing the GaAs nanowires, followed by AlGaAs/GaAs strained-layer superlattices (SLSs) and GaAs buffer layer. The coalesced GaAs film on the nanowire array achieved a flat surface with a root-mean-square (rms) deviation of 1.9 nm across a scanned area of $5 \times 5\ \mu m^2$, as shown in Fig. 11.3a and b, respectively. It was revealed that highly ordered nanowires around tiara-like structure beneath the SiO_2 sidewall confined most of stacking faults (Fig. 11.3c). The suppression of APBs was explained by the homogeneous nucleation of III–V layers on the Si {111} plane.

3.2.1.2 Buffer layer grown by MOCVD/MOVPE

Alcotte et al. [62] also reported anti-phase domain free GaAs layers grown on a nominal (001) Si substrate by MOCVD. The Si wafers, deoxidised in a SiConi chamber using a NF_3/NH_3 remote plasma, were transferred into the MOCVD chamber and subsequently annealed in H_2 ambient,

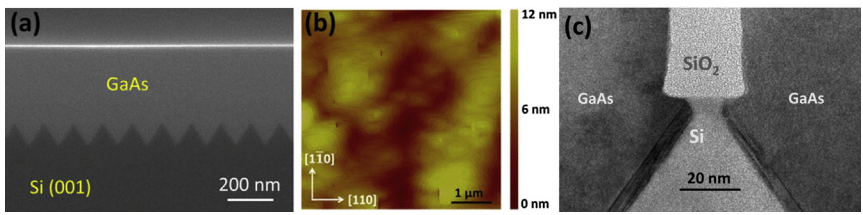


Figure 11.3 (a) Cross-sectional SEM image and (b) AFM image of a coalesced GaAs film grown on a GaAs nanowire array grown on Si. (c) Cross-sectional TEM image showing the stacking fault confined by tiara-like structure underneath SiO_2. *Reproduced with permission from Ref. [76] © 2015 AIP publishing LLC.*

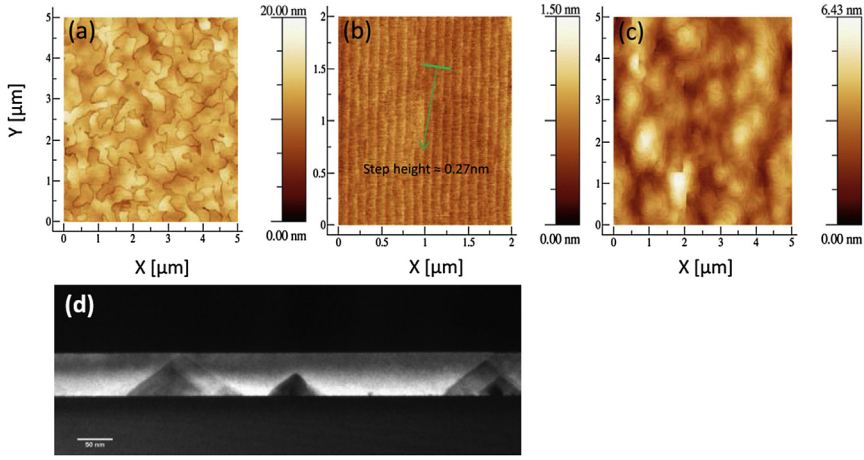


Figure 11.4 (a) 5×5 μm^2 AFM image of 400 nm-thick GaAs on unoptimised (001) Si with a high density of randomly distributed APBs. (RMS = 1.6 nm). (b) 2×2 μm^2 AFM image of optimised 0.15 degrees (001) Si with double-stepped surface. (c) 5×5 μm^2 AFM image of an APB-free 150 nm-thick GaAs grown on optimised 0.15 degrees (001) Si. (RMS = 0.8). (d) Dark field cross-sectional TEM image of GaP film grown on on-axis (001) Si. *(a), (b), and (c) Reproduced from Ref. [62], licensed under CC BY 4.0. (d) Reproduced with permission from Ref. [80] © 2010 Elsevier B.V. All right reserved.*

promoting a 2×1 surface reconstruction. The optimised Si surface was mainly comprised of double atomic steps, which are essential to inhibit the formation of APBs, and only a few monoatomic islands were found at the step edges. The 150 nm-thick GaAs layer was grown on this optimised Si surface using a conventional two-step process, resulting in low surface roughness without APBs. Unlike the GaAs film grown on un-optimised Si (001) which showed randomly distributed APBs (Fig. 11.4a), the optimised Si surface exhibited double steps (Fig. 11.4b), and GaAs films grown on optimised Si achieved a rms roughness of 0.8 nm without APBs (Fig. 11.4c), which is confirmed by AFM measurements. The APB-free GaAs on nominal (001) Si was attributed to the surface preparation, for example, plasma treatment in SiConi and H_2 ambient annealing in MOCVD, which promote the formation of double atomic steps on the Si surface.

In addition, Volz et al. [80] developed the APB-free GaP/Si template by using vapor phase epitaxy (VPE) and metal-organic vapor phase epitaxy (MOVPE). The growth of GaP on on-axis (001) Si substrates consists of a two-step process, i.e., the homoepitaxial growth of a Si buffer layer by

VPE and the heteroepitaxial growth of GaP buffer layer by MOVPE. The homoepitaxial Si buffer layer was first grown and then annealed in H_2 ambient, the second process serving an important role in the formation of double atomic steps on the on-axis Si substrate. Then, a GaP layer was grown on the homoepitaxial Si buffer layer using a two-step growth process comprised of a nucleation and an overgrowth step. As can be seen in Fig. 11.4d, a cross-sectional dark field TEM image shows the self-annihilation of APBs within the GaP buffer layers.

3.2.1.3 All-MBE grown GaAs buffer layer

Despite substantial recent development in the growth of monolithic III-Vs on CMOS-compatible on-axis (001) Si, all of the mature methods mentioned above require an MOCVD/MOVPE growth system to obtain an anti-phase domain-free virtual substrate, involving additional cost and thereby reducing the benefit of low-cost monolithic integration. Moreover, high-quality III–V QD structures are commonly grown by molecular beam epitaxy (MBE) system because it is relatively difficult for MOCVD system to grow high-quality InAs/GaAs QDs [84]. In this regard, it is desirable to achieve a one-step growth of III–V laser structures on on-axis Si (001) substrate by using an MBE system. Recently, Kwoen et al. [85] reported the elimination of APBs in all-MBE GaAs layers grown on on-axis Si (001). In this work, the effect of the composition of $Al_xGa_{1-x}As$ as a nucleation layer on the elimination of APBs was investigated. Fig. 11.5a–d show cross-sectional SEM images of GaAs layers grown on Si with $Al_xGa_{1-x}As$ and GaAs nucleation layers, with x having values of 0, 0.3, 0.5, and 0.7. Unlike the $Al_{0.7}Ga_{0.3}As$ and $Al_{0.5}Ga_{0.5}As$ nucleation layers (Fig. 11.5c

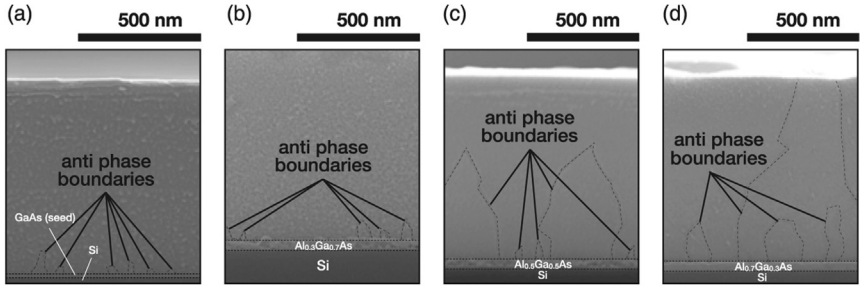


Figure 11.5 SEM images of GaAs buffer layer grown on (001) Si substrate with (a) GaAs nucleation layer, (b) $Al_{0.3}Ga_{0.7}As$ nucleation layer, (c) $Al_{0.5}Ga_{0.5}As$ nucleation layer and (d) $Al_{0.7}Ga_{0.3}As$ nucleation layer. *Reproduced with permission from Ref. [85] © 2019 The Japan Society of Applied Physics.*

and d), the $Al_{0.3}Ga_{0.7}As$ nucleation layer (Fig. 11.5b) provided a similar quality as the GaAs layer (Fig. 11.5a), without APBs. It was believed that the annihilation mechanism of APBs was mainly attributed to the self-annihilation rather than the formation of double atomic steps on Si.

3.2.2 Approaches for the reduction of threading dislocation density

High-quality III–V lasers grown on native substrates such as GaAs or InP, which are widely employed in heterogeneous integration for Si photonics, have a very low density of TDs ($<10^5$ cm^{-2}) [86]. On the other hand, III–V layers monolithically grown on Si substrates normally generate a high density of TDs (10^6–10^{10} cm^{-2}). In order to realise high-performance monolithic III–V lasers on Si, it is primarily important to reduce the density of TDs acting as electron scattering and nonradiative recombination centres [42,60]. Thus, substantial efforts have been devoted to developing growth techniques over decades, including dislocation filter layers (DFLs) [87–91], the two-step growth of a buffer layer [92], thermal cycle annealing [93–95] and so on [96–98]. For example, Tang et al. [90,91] investigated the optimisation of SLSs such as InGaAs/GaAs and InAlAs/GaAs which are commonly used now for DFLs, and successfully demonstrated the reduction in TDD down to the order of 10^6 cm^{-2}, in particular for QD lasers emitting at the telecommunication wavelength of 1.3 μm. In Fig. 11.6, it can be clearly seen that the majority of TDs are effectively confined by three sets of InGaAs/GaAs SLSs DFLs, this success being ascribed to the optimisation of the indium composition ($x = 0.18$) and GaAs thickness (10 nm) in the

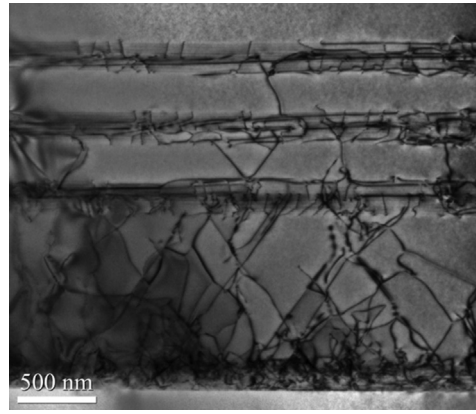


Figure 11.6 Cross-sectional TEM image of three layers of InGaAs/GaAs SLSs DFLs on GaAs/Si. *Reproduced from Ref. [91], licensed under CC BY 3.0.*

$In_xGa_{1-x}As$/GaAs SLSs as well as the growth condition of GaAs spacer layer. Also, Yang et al. [88], who investigated the effect of self-organised In(Ga, Al)As/GaAs QDs on the reduction of TDD, reported that the 10 layers of self-organised InAs QDs as DFLs were effective in reducing TDD. This can be explained by the dislocation bending below the QD islands.

3.2.3 Approaches for minimising thermal cracks

For the suppression of thermal cracks caused by a large difference of the CTE, it is important to minimise the total thickness of the buffer layer. Selective area masking can also be used to minimise crack nucleation centres [99–101]. Recently, Yang et al. [102] investigated the effect of annealing temperature, doping concentration and the thickness of the Ge buffer layer on the density of TDs. They showed that InAs/GaAs QDs grown on a Si substrate with a thin 270 nm-thick Ge buffer layer exhibited a TDD of 6.1×10^8 cm^{-2}, similar to the TDD for a 1400 nm-thick GaAs buffer layer (5.9×10^8 cm^{-2}). This result shows that use of the Ge buffer layer can reduce the total thickness of the conventional buffer layer by at least 1 μm without degrading the quality of the III–V layers. Clearly, this is a promising approach to prevent the formation of thermal cracks whilst minimising the thickness of the buffer layers.

4. Current status of III–V quantum dot lasers on silicon

4.1 Fabry–Pérot edge-emitting lasers on Si

4.1.1 High-performance QD lasers on offcut Si

Even though QD lasers have a relatively strong tolerance to TDs when compared to QW lasers, the presence of a high density of TDs still has a detrimental effect on the performance of QD laser devices, and thus minimising the density of TDs is a key issue. Since the first demonstration of telecommunication wavelength (1.3-μm) InAs/GaAs QD lasers on Si [59], rapid advances have been made in growth techniques such as using a DFL layer, a nucleation layer and so on, as mentioned in Section 3. On the basis of previous research, in 2016, Chen et al. [40] reported high-performance and high-reliability 1.3-μm InAs/GaAs QD lasers grown on Si. In this work, Si (100) substrates with a 4 degrees offcut to the [011] plane were used to prevent the formation of APBs and obtain a single domain during the growth. To reduce the density of TDs, optimised structure and growth conditions, including an AlAs nucleation layer, a multi-step GaAs

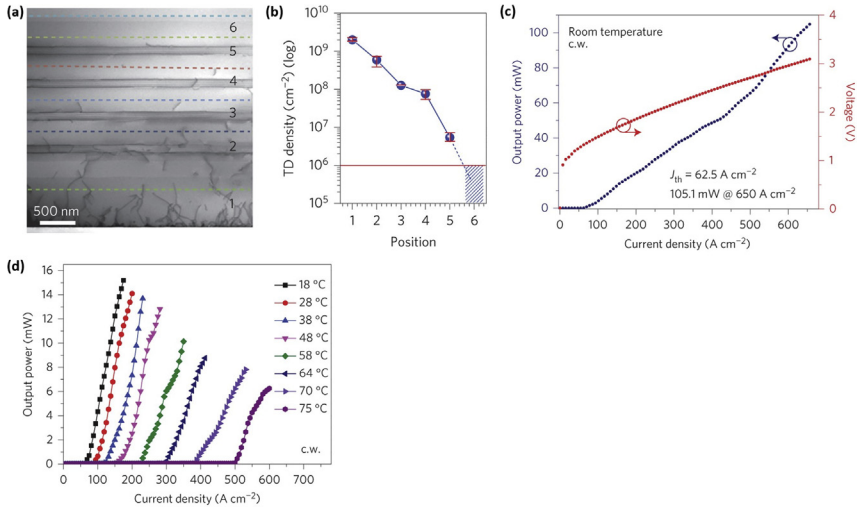


Figure 11.7 (a) Bright-field TEM image of DFLs on GaAs/Si. (b) Density of threading dislocation measured at different positions, denoted in (a). (c) Typical L-I-V characteristics for 50 × 3200 μm QD lasers on Si under cw operation at RT. (d) Temperature-dependent L-I curve for same device. *Reproduced with permission from Ref. [40] © 2016 Springer Nature.*

buffer layer and InGaAs/GaAs SLS DFLs, were employed. As shown in position 1 of Fig. 11.7a, most of the defects were confined within the 200 nm-thick GaAs buffer layer. This was achieved by growing a 6 nm-thick AlAs nucleation layer and three GaAs buffer layers (30/170/800 nm) grown at 350, 450 and 590°C, respectively. To further reduce the density of TDs, optimised $In_{0.18}Ga_{0.82}As$/GaAs SLS DFLs were introduced, resulting in a reduction of the TDD to $\sim 10^5$ cm^{-2} as shown in Fig. 11.7b. The broad area lasers fabricated using these techniques achieved a record low threshold current density of 62.5 A/cm^2 at RT, a high power of 100 mW and an operating temperature up to 75°C, as shown in Fig. 11.7c and d. Moreover, an extrapolated mean time failure was estimated to be over 100,000 h, indicating that the QD laser on Si is feasible for practical applications.

4.1.2 O-band QD lasers on on-axis Si

Over the past few years, extensive efforts have been directed towards using CMOS-compatible on-axis (001) Si substrates. Indeed, many groups have demonstrated the InAs/GaAs QD lasers on on-axis Si by adopting several solutions such as intermediate GaP [80,103–105] or GaAs [106,107] buffer

layers grown by MOCVD or MOVPE, patterned Si (GoVS) [77,108,109] and MBE-grown AlGaAs buffer layers [84,110]. For instance, Liu et al. [103] first demonstrated 1.3-μm cw QD lasers grown on on-axis Si using a commercially available GaP/Si template. The broad-area lasers ($2000 \times 20\ \mu m^2$) with as-cleaved facets, grown on a GaP/Si template and a GaAs native substrate, showed a threshold current (density) of 345 mA (862 A/cm^2) and 190 mA (475 A/cm^2), respectively. The narrow-ridge waveguide lasers on GaP/Si with high-reflection (HR) facet coating achieved thresholds as low as 30 mA and operating temperatures up to 90°C from $3 \times 750\ \mu m^2$ and $3.5 \times 1500\ \mu m^2$ devices, respectively. Following this result, Jung et al. [105] reported high-reliability 1.3-μm QD ridge waveguide lasers on GaP/Si. In this paper, the ridge-waveguide laser ($2.5 \times 1079\ \mu m^2$) with HR coating on one facet produced a threshold current as low as 6.2 mA and a single-side peak wall-plug efficiency of 21%, as shown in Fig. 11.8a. The maximum operating temperature up to 85°C with output power of ~8 mW was achieved from a QD device ($8 \times 1341\ \mu m^2$) in Fig. 11.8b. The highest injection efficiency of 87%

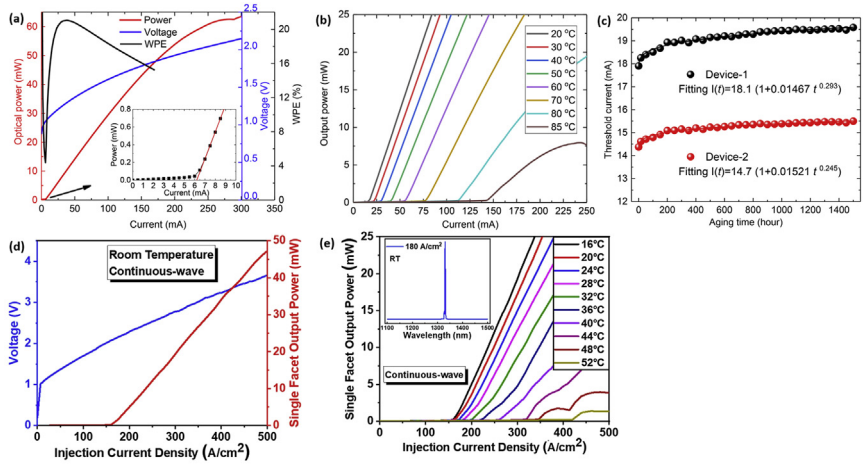


Figure 11.8 (a) L-I-V and wall-plug efficiency curves from 2.5×1079 μm QD lasers on GaP/Si with HR coating under cw operation at RT showing a threshold current of 6.2 mA and single-side peak efficiency of 21%. (b) Temperature-dependent L-I curve for 2.5×1079 μm QD lasers on GaP/Si. (c) Evolution of cw threshold current increases for two 1641 μm-long QD lasers on GaP/Si device during 1500 h lifetime test. (d) Conventional L-I-V curve and (e) temperature-dependent L-I curve from 50×3000 μm broad-area QD lasers on GaAs/Si. *(a), (b), and (c) Reproduced with permission from Ref. [105] © 2018 American Chemical Society. (d) and (e) Reproduced from Ref. [107], licensed under CC BY 4.0.*

was obtained from a QD device ($6 \times 1341\ \mu m^2$). The high-performance of QD lasers on GaP/Si was attributed to the p-type modulation doping in the QD active region and the TDD being reduced down to $8.4 \times 10^6\ cm^{-2}$. Fig. 11.8c presents a cw threshold current increase from two 1641 μm-long devices during a 1500-hour lifetime test. Only a 9.5% increase of threshold current was observed, and, moreover, the extrapolated mean time to failure was estimated to be more than a million hours (Fig. 11.8c), implying the superior reliability of QD lasers.

Also, Chen et al. [106], who employed an MOCVD-grown GaAs buffer layer on on-axis (001) Si, reported 1.3-μm InAs/GaAs QD lasers on Si. In this work, the broad-area lasers ($25 \times 3000\ \mu m^2$) without both facet coating and p-modulation doping exhibited J_{th}s of 425 A/cm^2 and 210 A/cm^2 for Si and GaAs substrate, respectively. Under cw operation, the laser devices on Si obtained a maximum single facet output power of 43 mW and operating temperatures up to 36°C. Following this result, Li et al. [107] reported an improved performance of 1.3-μm QD lasers on on-axis Si using an MOCVD-grown GaAs buffer layer. As shown in Fig. 11.8d, the broad-area laser on Si ($50 \times 3000\ \mu m^2$) showed a low J_{th} of 160 A/cm^2 at RT under cw operation and a single facet output power of 48 mW. In Fig. 11.8e, it was shown that the maximum operating temperature of 52°C was obtained, which can be further improved by the p-modulation doping technique.

For the patterned Si substrate, Norman et al. [108] demonstrated electrically pumped cw 1.3-μm QD lasers on GoVS templates. The TDD of QD lasers was measured to be $7 \times 10^7\ cm^{-2}$. The fabricated laser device ($9 \times 1200\ \mu m^2$) showed a threshold current of 81 mA and single-facet output power over 50 mW. The cw maximum operating temperature up to 80°C was achieved from the $9 \times 1200\ \mu m^2$ device with HR coating. Based on this work, Shang et al. [109] further improved the performance of QD lasers on a GoVS template. The 1450 μm-long laser devices with ridge width of 10 and 2 μm were fabricated. The $10 \times 1450\ \mu m^2$ device produced a threshold current (density) of 39 mA (270 A/cm^2) and maximum power of 75 mW, and the $2 \times 1450\ \mu m^2$ device showed threshold current (density) of 16 mA (550 A/cm^2) and maximum power of 33 mW. In addition, the dynamic characteristics of $5 \times 1450\ \mu m^2$ devices with a different probe metal design were measured. Type A and B lasers, denoting that the probe metal has no vertical overlap between the p-/n-metal contacts and a portion of the p-probe metal was placed on SiO_2/n-contact layers, respectively, exhibited 3 dB bandwidths ($f_{3\ dB}$) of 5.8 GHz (type A) and

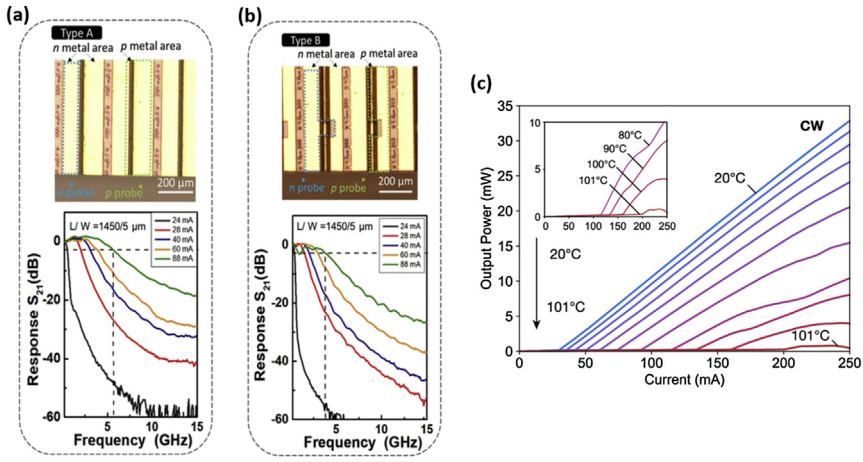


Figure 11.9 Frequency response of a (a) type A and (b) type B lasers with a cavity length of 1450 μm and width of 5 μm showing the maximum $f_{3\ dB}$ of 5.8 GHz (type A) and 3.6 GHz (type B). (c) Temperature-dependent L-I curve under cw operation from QD lasers directly grown on on-axis (001) Si. *(a) and (b) Reproduced from Ref [109], licensed under CC BY 4.0. (c) Reproduced with permission from Ref. [110] © The Optical Society.*

3.6 GHz (type B), as shown in Fig. 11.9a and b. The better performance of type A device was ascribed to the reduced parasitic capacitance.

With regard to the all MBE-grown QD lasers on on-axis Si, without using patterned or virtual templates, Kwoen et al. [110] were the first to report high-temperature cw operation of a InAs/GaAs QD laser on on-axis Si. In this work, a 40 nm-thick AlGaAs seed layer and 800 nm-thick GaAs buffer layer were directly grown on an n-type (001) Si substrate using an MBE system, and the TDD was measured to be 4.7×10^7 cm^{-2}. The fabricated laser device (7×1000 μm^2) produced a threshold current (density) of 27.6 mA (370 A/cm^2) and cw operating was obtained up to 101°C, as shown in Fig. 11.9c.

4.1.3 C-band QD lasers on on-axis Si

Whilst 1.3-μm InAs/GaAs QD lasers on Si have been intensively investigated, there are a few reports of C-band (1.55-μm) QD lasers grown on Si [111–113], which is important for mid/long-haul communications [114,115]. This is mainly because the large lattice mismatch (8%) between InP and Si generates a TDD on the order of 10^9–10^{10} cm^{-2}, which is much higher than that of GaAs on Si (10^6–10^8 cm^{-2}). Indeed, it was not

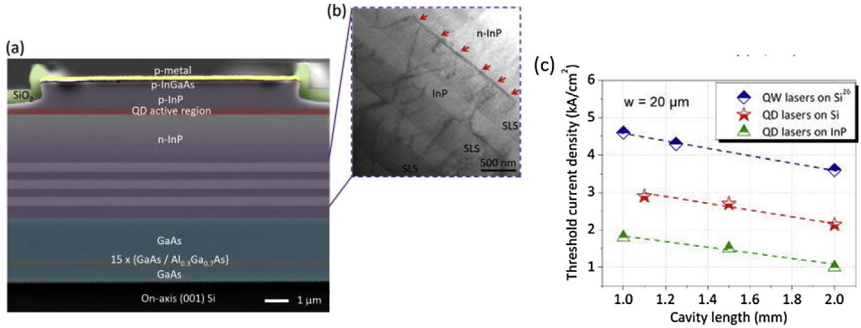


Figure 11.10 (a) Colour-enhanced cross-section SEM image of epitaxial layer. Cross-section TEM images of (b) three stacks of 10-period InGaAs/InP SLSs showing defect filtering effect denoted by *red* (dark grey in print version) arrows and (c) Threshold current density of QD lasers grown on Si and native InP substrate, compared with QW lasers grown on Si. *Reproduced with permission from Ref. [111] © AIP publishing.*

until 2018 that Zhu et al. first demonstrated an InP-based 1.5 μm InAs QD laser on Si under pulsed operation [111]. In this result, the InP-on-v-grooved-Si template was used to suppress APBs, and InP layer was grown on GaAs/Si, as illustrated in Fig. 11.10a. To reduce the density of TDs, 10 periods of $In_{0.6}Ga_{0.4}As$/InP (10/30 nm) SLSs as DFL were introduced and repeated three times with 250 nm-thick InP spacer layers, resulting in a reduction of TDD from the order of 10^9–10^{10} cm^{-2} to 1.5×10^8 cm^{-2}, as shown in Fig. 11.10b. With an active region consisting of five layers of InAs/InAlGaAs QDs, a J_{th} of 1.6 kA/cm^2 and a single facet output power of 57 mW were obtained from a 10×5000 μm^2 device. A 20×1000 μm^2 device exhibited a lasing temperature up to 80°C and a characteristic temperature of 58.7 K without p-modulation doping. In addition, InP-based QD lasers on Si and InP substrate were also compared with QW lasers on Si. The threshold current densities obtained are plotted as a function of cavity length in Fig. 11.10c. It was shown that the J_{th} of QD lasers on Si is a factor of 2 times higher than those on InP but a factor of 1.6 times lower than QW lasers on Si, indicating that QDs as an active medium are a promising approach for monolithic integration on Si.

4.2 Single mode QD edge-emitting lasers on Si

4.2.1 QD distributed-feedback lasers

In order to realise the ultimate goal of Si photonics, it is particularly important to develop an array of longitudinal single-mode lasers for advanced optical communication systems with applications such as wavelength

division multiplexing (WDM) [12,116,117]. The WDM technology transmits multiple optical signals at different wavelengths through a single optical fibre, typically with a channel (wavelength) spacing of 20 nm for a standard CWDM system. In this respect, distributed feedback (DFB) laser arrays, providing a limited number of wavelength channels, are of great importance for addressing monolithic light sources in Si photonics used for the WDM system.

In 2018, Wang et al. [118] demonstrated the first electrically pumped O-band QD DFB laser array grown on Si. In this paper, an InAs/GaAs QD laser structure was grown on an n-type (001) Si substrate with a 4 degrees off-cut towards the [011] plane. Fig. 11.11a displays a schematic illustration of the fabricated single DFB laser. Unlike conventional DFB lasers where Bragg gratings were formed within the active region, in this work, lateral gratings were employed in order to avoid complicated regrowth step and fabricate both the waveguide ridge and gratings via a single dry-etching step. To obtain single-longitudinal mode lasing, first-order sidewall gratings with a λ/4 phase shift were used; to suppress back

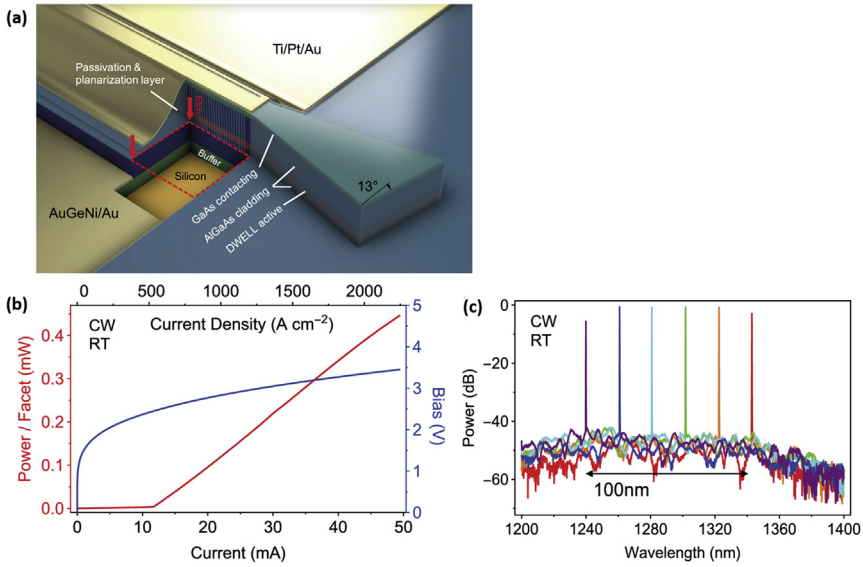


Figure 11.11 (a) Schematic illustration of a QD DFB laser on Si showing vertical layer structures, output coupler and vertical gratings. (b) Typical L-I-V characteristics of a single 1 mm-long QD DFB laser at RT. (c) Optical spectra of QD DFB laser array with different gratings showing 100 nm coverage in the O-band and a 20 nm channel spacing. *Reproduced from Ref. [118], licensed under CC BY 4.0.*

reflections, AR output couplers instead of AR coatings were adopted. The fabricated DFB laser array on Si exhibited a cw threshold current of 12 mA and side-mode suppression ratio (SMSR) of 50 dB, as can be seen in Fig. 11.11b and c, respectively. In particular, the DFB laser array achieved a broad wavelength coverage of 100 nm in the O-band and a channel (wavelength) spacing of 20 ± 0.2 nm, which are suitable for a standard CWDM system. This can be attributed to a broad gain bandwidth due to the inhomogeneity of the QDs.

In 2020, Wan et al. [119] also reported first 1.3-um QD DFB lasers grown on an APD-free GaP/Si substrate. In this work, conventional two-step MBE re-growth was employed to obtain first-order DFB gratings just above the active layer without a $\lambda/4$ phase shift. The as-cleaved DFB lasers showed lasing behaviour up to 70°C under cw operation for $2.5 \times 2000\ \mu m^2$ devices and threshold current (density) of 20 mA ($440\ A\ cm^{-2}$) for $3 \times 1500\ \mu m^2$ devices. The as-cleaved DFB laser array ($3 \times 700\ \mu m^2$) with a channel (wavelength) spacing of 13 nm and peak wavelengths from 1285 to 1338 nm exhibited SMSRs from 45 to 50 dB. In addition, a QD DFB laser array demonstrated an aggregate data rate of over 640 Gbit/s by a proof-of-concept test using external modulation.

4.2.2 QD wavelength tunable lasers

Wavelength tunable lasers are promising light sources particularly for dense wavelength division multiplexing (DWDM) systems which commonly require 0.4 or 0.8 nm channel (wavelength) spacing [120]. This is because the emission wavelength of tunable lasers can be finely altered in a controlled manner, whilst the DFB lasers need a change in device structure for wavelength tuning. Currently available tunable lasers are commonly combined with distributed Bragg reflectors [121,122], micro-ring (MR) resonators [123] and coupled cavities [124,125]. However, they are limited by a large footprint, re-growth steps in fabrication and the use of a native substrate requiring a bonding process and hence additional cost.

In 2020, Wan et al. [126] first demonstrated the InAs/GaAs QD tunable lasers grown on GaP/Si substrates using a half-wave-coupled ring resonator without epitaxial re-growth or gratings. A half-wave coupler [127] was employed in this work to achieve high SMSR. As shown in Fig. 11.12a and b, the tunable laser consisted of a conventional FP cavity coupled with two active ring resonators by two half-wave couplers. Note that the cross-sectional images of the FP cavity, MR resonator and coupler are the same as shown in Fig. 11.12c, because identical structures were

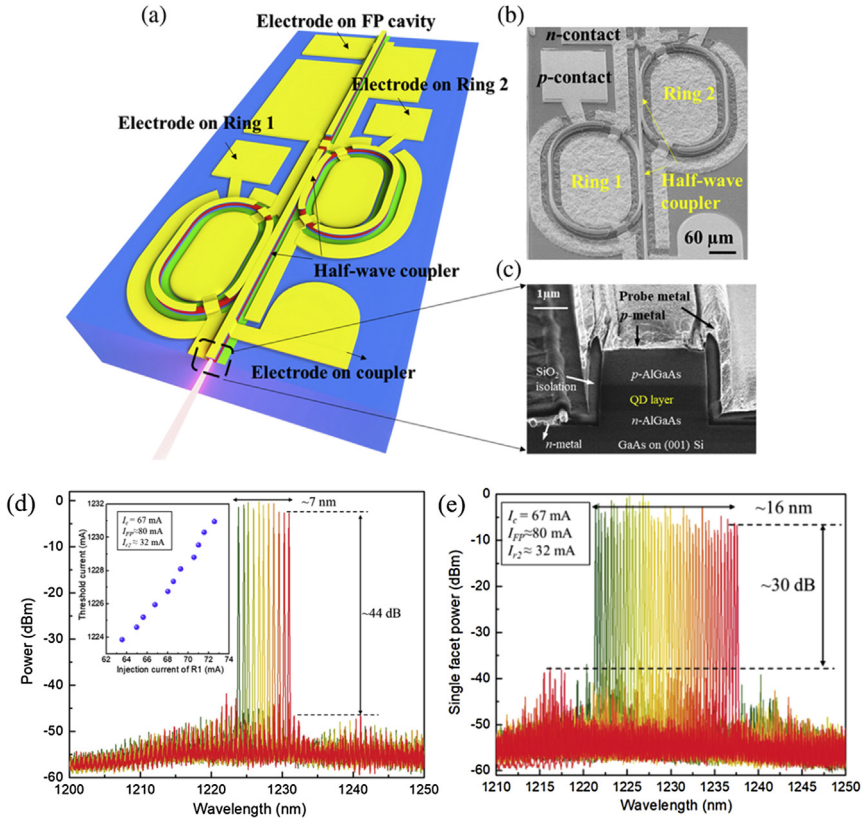


Figure 11.12 (a) schematic diagram and (b) top-view SEM image of the tunable laser system. (c) cross-sectional SEM image of the FP cavity. Superimposed tuning spectra (d) with 11-channel wavelength switching and a minimum SMSR of 45 dB and (e) with 37-channel wavelength switching and minimum SMSR of 30 dB. *Reproduced with permission from Ref. [126] © The Optical Society.*

simultaneously processed to fabricate each device. Under cw operation at RT, 11-channel (37-channel) wavelength switching was achieved with a 7 nm (16 nm) tuning range and minimum SMSR of 45 dB (30 dB), as shown in Fig. 11.12d and e, respectively. The tuning range can be potentially increased by adjusting the perimeter difference between the two ring resonators and varying the thermoelectric cooler temperature.

4.3 QD mode-locked lasers on Si

QD mode-locked lasers (MLLs) are also promising candidates for WDM light sources because the unique properties of QDs provide an inherently

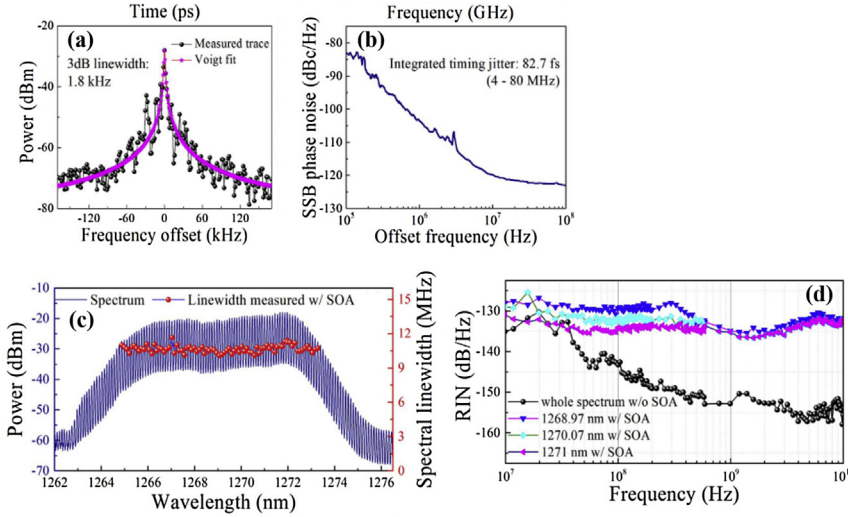


Figure 11.13 The fabricated 20 GHz QD-MLL grown on Si: (a) narrow span RF peak (RF peak occurred at 20.02 GHz, RBW: 1 kHz); (b) single-sideband noise plot showing a low timing jitter of 82.7 fs; (c) optical spectrum and corresponding optical line width of each mode within 10 dB; and (d) relative intensity noise of O-band and specific filtered wavelength channels under $I_{gain} = 180$ mA, $V_{SA} = -1.92$ V. *Reproduced with permission from Ref. [129] © The Optical Society.*

inhomogeneous broadening of the gain spectrum and ultrafast carrier dynamics [128]. Recently, Liu et al. [129] first reported 20 GHz passively mode-locked QD lasers grown directly on an on-axis Si substrate. As shown in Fig. 11.13a and b, the QD MLLs showed a narrow RF 3 dB linewidth of 1.8 kHz and a record low timing jitter of 82.7 fs from 4 to 80 MHz, respectively. In Fig. 11.13c, the square-like optical spectrum with the largest 3 dB bandwidth including 58 comb lines was obtained. The average RIN of the whole spectrum and a single wavelength channel is measured to be −152 dB/Hz and −133 dB/Hz, respectively (Fig. 11.13d). Utilizing 64 wavelength channels as optical carriers, an aggregate total transmission capacity of 4.1 Tbit/s was achieved via Nyquist four-level pulse amplitude modulation format.

4.4 QD micro-cavity lasers on Si

Micro-cavity lasers offer unique advantages such as extremely low power consumption and dense integration over FP lasers. This is because micro-cavity lasers have a small cavity volume, whilst the FP lasers require a cavity length of at least several hundred micrometres [130]. To implement these

benefits on Si photonics, various QD micro-cavity lasers grown on Si, including whispering-gallery mode micro-disk and MR, have been widely investigated in recent years [113,131–143].

4.4.1 QD micro-disk lasers

With regard to the micro-disk (MD) lasers on Si, for example, Kryzhanovskaya et al. [136] first demonstrated electrically pumped cw operation of QD MD lasers grown on Si with emission wavelength beyond 1.3 μm (1.32–1.35 μm). Fig. 11.14a presents a schematic diagram of the QD MD laser structure. The MD lasers with diameters from 14 to 30 μm, grown on (001) Si substrate with 4 degrees offcut to the [011] plane, showed cw lasing operation at RT and produced the minimum J_{th} of 600 A/cm^2. Fig. 11.14b displays representative lasing spectra of MD lasers with diameters of 14 and 30 μm. The peak wavelength of MD lasers with diameter of 14 μm is located at 1351 nm and both devices produce nearly single-mode lasing. Like the FP lasers, the research focus of MD lasers also moved towards the monolithic growth on on-axis (001) Si substrates compatible with the

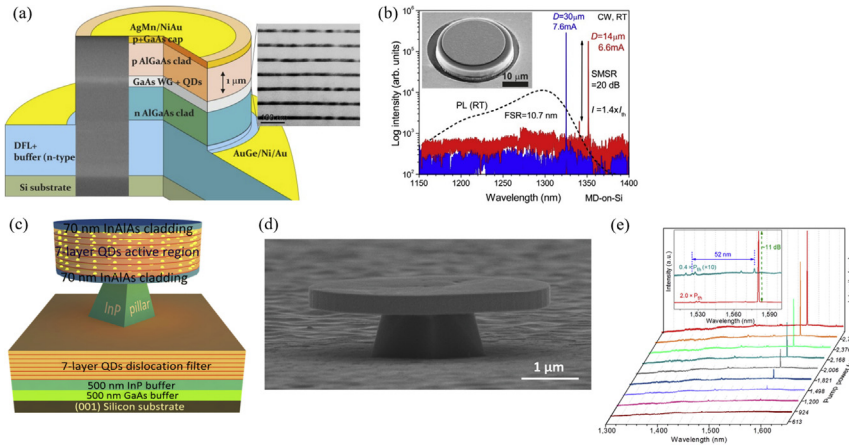


Figure 11.14 (a) schematic diagram of 1.3-μm QD micro-disk lasers on Si with an SEM image of the cross-section (left inset) and a bright-field TEM image of the QD active layer (right inset). (b) RT cw emission spectra at a current of 1.4× threshold for MD lasers on Si with diameters of 14 and 30 μm. Dashed curve presents PL spectra of QD and inset shows the SEM image of the MD laser. (c) Schematic illustration of the epitaxial structure of a 1.55-μm QD micro-disk laser on Si. (d) Side-view SEM image and (e) μ-PL spectra as a function of optical pump power for 4 μm-diameter micro-disk lasers on Si. The inset of (e) shows a comparison of emission below and above threshold. *(a) and (b) Reproduced with permission from Ref. [136] © The Optical Society. (c), (d), and (e) Reproduced with permission from Ref. [135] © 2017, American Chemical Society.*

CMOS platform. For instance, Zhou et al. [140] reported O-band InAs/GaAs QD MD lasers grown on planar on-axis Si (001) with a diameter of 1.1 μm and a very low threshold of 3 μW. In this result, the fabricated devices were used to demonstrate optically pumped cw operation at RT with emission wavelength at 1315 nm.

For the C-band MD lasers, Shi et al. [135], incorporating QD DFLs and a GaAs intermediate buffer layer on (001) planar Si for InP growth, demonstrated 1.55 μm InAs/InAlGaAs QD MD lasers on Si with a low threshold of 1.6 mW at 4.5 K under cw optical pumping. Fig. 11.14c depicts the schematic diagram of the epitaxial structures. To optically isolate the active region of the disk, an InP pillar was formed by the selective etching of the 1 μm InP buffer, as shown in Fig. 11.14d. Micro-photoluminescence (μ-PL) spectra as a function of optical pump power were measured to characterise the lasing behaviour of 4 μm QD MD lasers in Fig. 11.14e. C-band 1.578 μm peak appeared under cw optical pumping at 4.5 K, and single-mode lasing at twice the threshold was demonstrated with a high extinction ratio of ~11 dB, as shown in the inset of Fig. 11.14e. Based on this work, Shi et al. [113] improved maximum operating temperature up to 60°C. However, for both O-band and C-band QD MD lasers on on-axis Si, electrically pumped cw operation has not been demonstrated yet.

4.4.2 QD micro-ring lasers

For the MR lasers, Wan et al. [141] reported electrically pumped cw 1.3-μm QD MR lasers on on-axis (001) Si with a GoVS template. They investigated various ring structures with radii from 5 to 50 μm and ring width from 2 to 7 μm. Fig. 11.15a illustrates the fabrication process of an MR laser device. The MR lasers with a radius of 50 μm and width of 4 μm gave a threshold current of 15 mA corresponding to J_{th} of 1.2 kA/cm^2 and achieved high-temperature cw operation up to 100°C. Moreover, as shown in Fig. 11.15b, a submilliamp threshold current of 0.6 mA was obtained from a smaller device with a radius of 5 μm and width of 3 μm, corresponding to J_{th} of 0.9 kA/cm^2. Fig. 11.15c shows the average threshold current with different cavity size. The threshold current decreases as cavity size decreases, indicating that there is no significant increase of optical loss or nonradiative recombination in the small cavities. Based on this result, Wan et al. [142] investigated the statistical comparison of lasing performance in InAs QD MR lasers grown on different substrates such as on-axis (001) GaP/Si and GoVS as well as GaAs. It was shown that the average threshold current densities of MR lasers on GaP/Si and GoVS template were

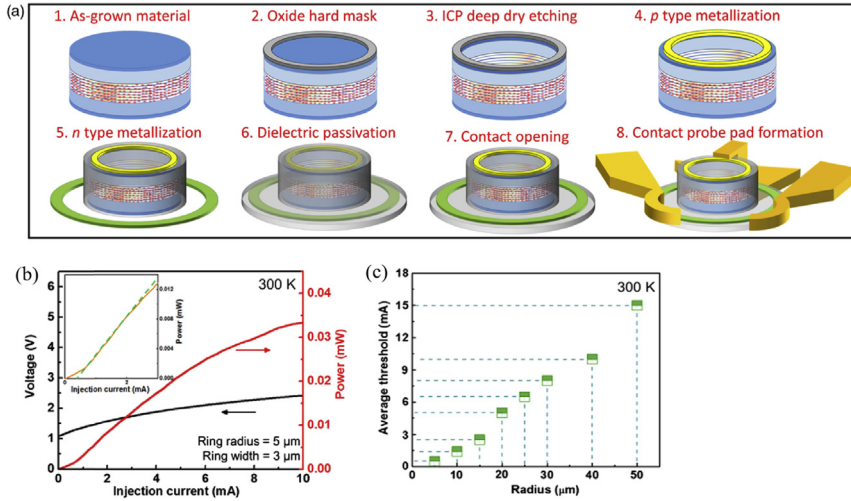


Figure 11.15 (a) The fabrication process of a micro-ring laser. (b) L-I-V characteristics of a micro-ring laser with a radius of 5 μm and a ring width of 3 μm. (c) Average threshold current with different radii at RT. *Reproduced with permission from Ref. [141] © The Optical Society.*

measured to be 0.675 kA/cm^2 and 0.875 kA/cm^2, respectively, corresponding to 2–3 times higher J_{th} than MR lasers grown on native GaAs substrate (0.325 kA/cm^2). The characteristic temperature (T_0) of various MR lasers was estimated to be in the range of 27–35 K, 25–26 K and 20–40 K for the GaAs, GoVS and GaP/Si substrates, respectively. It was revealed that, whilst the MR lasers on GaP/Si provided a better threshold current and T_0 than those on the GoVS template, the MR lasers on GoVS have advantages in terms of thinner buffer thickness and smaller RMS roughness which are beneficial for the fabrication. In addition, directly modulated QD MR lasers on GaP/Si were also demonstrated by Wan et al. [143], producing a T_0 of 103 K, threshold current of 3 mA and 3 dB bandwidth of 6.5 GHz.

4.5 QD photonic crystal lasers on Si

In addition to MD/MR lasers using a whispering-gallery mode, photonic crystal (PC) lasers grown on Si were also demonstrated recently. Zhou et al. [144] first reported the ultra-small PC membrane QD lasers grown on CMOS-compatible on-axis Si (001) substrates. Fig. 11.16a shows a schematic illustration of the fabricated InAs/GaAs QD *L*3 defects PC lasers on (001) Si, in which the lattice constant, radius of etched air hole and shift distance are denoted by a, r and 0.15a, respectively. As shown in Fig. 11.16b and c, the

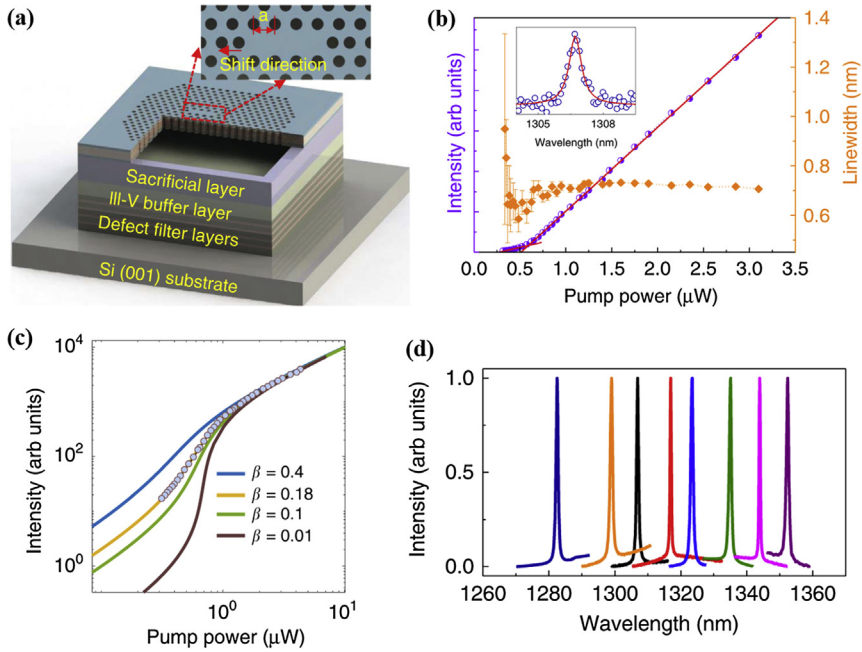


Figure 11.16 (a) Schematic diagram of the InAs/GaAs QD PC (L3 cavity) lasers grown on on-axis (001) Si substrate. (b) L-L curve and linewidth of the lasing peak at 1306 nm. (c) Logarithmic L-L plot of the PC laser. Dots indicate the experimental data and solid lines are theoretically calculated data for several β values. (d) Normalised PL spectra obtained from PC lasers with different structural parameters above lasing threshold. *Reproduced from Ref. [144], licensed under CC BY 4.0.*

PC lasers with a = 310 nm and r/a = 0.27 achieved a very low threshold of 0.6 μW and a large spontaneous emission coupling efficiency of 18% under cw optical pumping at RT, respectively. By changing the parameters of the PC lasers, moreover, a wide tuneable range of the lasing peak was demonstrated. This tuning range had a value of 70 nm and is shown in Fig. 11.16d.

5. Future directions of quantum dot lasers on Si

The role of light source in Si photonics, in general, is to provide an optical signal in a WDM system. Because a WDM system transmits multiple optical signals, each of which is carried by different wavelengths, the light sources for the WDM system require fundamental single-mode lasing at multiple wavelengths. An ideal on-chip laser for WDM should satisfy the following basic requirements [12]: (1) O-band or C-band discrete multi-

wavelength, (2) electrically pumped cw lasing, (3) fundamental longitudinal-single mode with high SMSR and (4) sufficient output power. In addition to these fundamental requirements, low power consumption, high temperature operation, at least up to 80°C, and channel (wavelength) spacing of 20 nm (0.4 or 0.8 nm) for CWDM (DWDM) applications are also necessary. Despite successful demonstrations of various QD lasers monolithically grown on Si, an on-chip laser for WDM, satisfying all the requirements for practical implementation, has not been demonstrated yet. Indeed, unlike hybrid laser integration which has already demonstrated a 400G 8-channel CWDM transmitter using silicon photonics [145], using the monolithic approach, only limited functionality of advanced QD lasers on Si has been demonstrated [118,129].

Micro-cavity lasers have also been regarded as a potential WDM light source especially due to their significant advantages such as ultra-low threshold and small foot-print [146]. Compared with a DFB laser array on Si, however, monolithic on-chip micro-cavity lasers remain to be further explored. For example, the instability of the output power hinders practical applications due to the isotropic characteristic of lasing emission and mode competition between clockwise and counter-clockwise modes [147]. The external reflector can be applied to obtain unidirectional emission, but it increases the device foot-print and complexity of the fabrication process [148]. In addition, the fundamental single-mode operation with high SMSR is required. Although parity-time symmetric [149,150] or grating integrated [151] MR lasers were demonstrated in single-mode operation, several key issues such as electrical injection, high threshold current and reliability remain to be solved.

Until now, the research for monolithic III–V QD lasers on Si has been focussed on individual device performance rather than the monolithic integration process with another optical component. With the rapidly improving performance of monolithic QD lasers, advanced strategies for fabricating QD lasers on Si compatible with the integration process are required to turn the potential of monolithically integrated Si photonics into reality.

6. Conclusion

In less than a decade, Si-based on-chip III–V lasers employing QD active region have been evolving at an unprecedented rate, and these

provide a hitherto missing component in Si photonics. Indeed, the reliability of QD lasers grown on CMOS-compatible on-axis Si substrates has already been shown to be comparable to that of QW lasers grown on a native substrate. Moreover, defect-tolerable QD lasers on Si show a superior lasing characteristic even in the presence of high densities of TDs, whilst QW lasers on Si gave no lasing operation. With the clear advantages of III–V QD lasers, a wide variety of QD lasers on Si have been demonstrated, such as Fabry–Pérot, micro-cavity, mode-locked and DFB lasers. Nonetheless, the deployment of on-chip III–V QD lasers on Si photonics has not been demonstrated yet, due to several remaining challenges. In heteroepitaxial growth, the reduction in density of TDs is the most important prerequisite for high-performance III–V lasers on Si since the device performance strongly depends on the quality of the epitaxial layers. In addition, one-step growth of APB-free III–V QD lasers on on-axis Si should be further matured, which is more cost-effective than multi-step buffer growth such as GaAs or GaP buffer grown by MOCVD/MOVPE or using a GoVS template. In terms of device fabrication, it is important to realise high-performance longitudinal-single-mode lasers or MLLs, which satisfy the requirements of a practical WDM light source. Additionally, the most recent results of III–V QD lasers on Si are limited to the conventional device fabrication method and characterisation of a single device. However, it is now needed to consider the compatibility with the overall monolithic integration process. Lastly, the light coupling issue between lasers and passive Si waveguides on SOI needs to be solved for the practical use of on-chip lasers. This is because, in general, on-chip III–V lasers emit light several micrometres above the passive Si waveguide on SOI, and this is one of the most critical obstacles that needs to be overcome to enable a monolithic approach to be advantageous over the mature heterogeneous approach.

Monolithic Si-based light sources have long been recognised to be the holy grail of Si photonics. However, owing to the considerable achievements in the development of Si-based QD lasers over recent years, the prospect of monolithic III–V QD light sources grown on Si is very promising. As we have reviewed, the next steps of on-chip III–V QD lasers will be the further development of the fabrication of integration-compatible devices and co-integration between active and passive optical components on the SOI platform. Even though Si photonics using mature heterogeneous integration is a predominant technology at this moment, it is obvious that the future of Si photonics is likely to be monolithic integration via on-chip III–V QD lasers, enabling low-cost and dense photonic integration.

Acknowledgement

This research was supported by UK EPSRC under Grants No. EP/P006973/, EP/S024441/1 and EP/T028475/1 and EPSRC National Epitaxy Facility.

References

[1] Cisco global cloud index: forecast and methodology, 2016–2021. White Paper; 2018.
[2] Cheng Q, Bahadori M, Glick M, Rumley S, Bergman K. Recent advances in optical technologies for data centers: a review. Optica 2018;5:1354–70.
[3] Soref R. The past, present, and future of silicon photonics. IEEE J Sel Top Quant Electron 2006;12:1678–87.
[4] Mayer AS, Kirkpatrick BC. Silicon photonics, frontiers in modern optics. IOS Press; 2016. p. 189–205. 190 of Proceedings of the International School of Physics "Enrico Fermi" series.
[5] Asghari M, Krishnamoorthy AV. Energy-efficient communication. Nat Photon 2011;5:268–70.
[6] Lim AEJ, Song J, Fang Q, Li C, Tu X, Duan N, Chen KK, Tern RPC, Liow TY. Review of silicon photonics foundry efforts. IEEE J Sel Top Quant Electron 2014;20:405–16.
[7] Intel. Intel® silicon photonics100G optical transceiver. 2017. https://www.intel.co.uk/content/www/uk/en/products/network-io/high-performance-fabrics/silicon-photonics.html.
[8] Liang D, Bowers JE. Recent progress in lasers on silicon. Nat Photon 2010;4:511–7.
[9] Rong H, Jones R, Liu A, Cohen O, Hak D, Fang A, Paniccia M. A continuous-wave Raman silicon laser. Nature 2005;433:725–8.
[10] Rong H, Xu S, Kuo YH, Sih V, Cohen O, Raday O, Paniccia M. Low-threshold continuous-wave Raman silicon laser. Nat Photon 2007;1:232–7.
[11] Wirths S, Geiger R, Von Den Driesch N, Mussler G, Stoica T, Mantl S, Ikonic Z, Luysberg M, Chiussi S, Hartmann JM, Sigg H, Faist J, Buca D, Grützmacher D. Lasing in direct-bandgap GeSn alloy grown on Si. Nat Photon 2015;9:88–92.
[12] Zhou Z, Yin B, Michel J. On-chip light sources for silicon photonics. Light Sci Appl 2015;4:e358.
[13] Roelkens G, Van Campenhout J, Brouckaert J, Van Thourhout D, Baets R, Romeo PR, Regreny P, Kazmierczak A, Seassal C, Letartre X, Hollinger G, Fedeli JM, Di Cioccio L, Lagahe-Blanchard C. III-V/Si photonics by die-to-wafer bonding. Mater Today 2007;10:36–43.
[14] Reed GT, Mashanovich G, Gardes FY, Thomson DJ. Silicon optical modulators. Nat Photon 2010;4:518–26.
[15] Streshinsky M, Ding R, Liu Y, Novack A, Yang Y, Ma Y, Tu X, Chee EKS, Lim AE-J, Lo PG-Q, Baehr-Jones T, Hochberg M. Low power 50 Gb/s silicon traveling wave Mach-Zehnder modulator near 1300 nm. Opt Exp 2013;21:30350–7.
[16] Heck MJR, Bauters JF, Davenport ML, Spencer DT, Bowers JE. Ultra-low loss waveguide platform and its integration with silicon photonics. Laser Photon Rev 2014;8:667–86.
[17] Liu A, Liao L, Chetrit Y, Basak J, Nguyen H, Rubin D, Paniccia M. Wavelength division multiplexing based photonic integrated circuits on silicon-on-insulator platform. IEEE J Sel Top Quant Electron 2010;16:23–32.
[18] Dong P. Silicon photonic integrated circuits for wavelength-division multiplexing applications. IEEE J Sel Top Quant Electron 2016;22:370–8.
[19] Michel J, Liu J, Kimerling LC. High-performance Ge-on-Si photodetectors. Nat Photon 2010;4:527–34.

[20] Tang M, Park JS, Wang Z, Chen S, Jurczak P, Seeds A, Liu H. Integration of III-V lasers on Si for Si photonics. Prog Quant Electron 2019;66:1–18.
[21] Liang D, Kurczveil G, Huang X, Zhang C, Srinivasan S, Huang Z, Seyedi MA, Norris K, Fiorentino M, Bowers JE, Beausoleil RG. Heterogeneous silicon light sources for datacom applications. Opt Fiber Technol 2018;44:43–52.
[22] Fang AW, Park H, Cohen O, Jones R, Paniccia MJ, Bowers JE. Electrically pumped hybrid AlGaInAs-silicon evanescent laser. Opt Exp 2006;14:9203–10.
[23] Jones R, Doussiere P, Driscoll JB, Lin W, Yu H, Akulova Y, Komljenovic T, Bowers JE. Heterogeneously integrated InP/silicon photonics: fabricating fully functional transceivers. IEEE Nanotechnol Mag 2019;13:17–26.
[24] Norman JC, Jung D, Wan Y, Bowers JE. Perspective: the future of quantum dot photonic integrated circuits. APL Photon 2018;3:030901.
[25] Kunert B, Mols Y, Baryshniskova M, Waldron N, Schulze A, Langer R. How to control defect formation in monolithic III/V hetero-epitaxy on (100) Si? A critical review on current approaches. Semicond Sci Technol 2018;33:093002.
[26] Li Q, Lau KM. Epitaxial growth of highly mismatched III-V materials on (001) silicon for electronics and optoelectronics. Prog Cryst Growth Char Mater 2017;63:105–20.
[27] Liao M, Chen S, Park JS, Seeds A, Liu H, III- V. quantum-dot lasers monolithically grown on silicon. Semicond Sci Technol 2018;33:123002.
[28] Arakawa Y, Sakaki H. Multidimensional quantum well laser and temperature dependence of its threshold current. Appl Phys Lett 1982;40:939–41.
[29] Liu H, Wang T, Jiang Q, Hogg R, Tutu F, Pozzi F, Seeds A. Long-wavelength InAs/GaAs quantum-dot laser diode monolithically grown on Ge substrate. Nat Photon 2011;5:416–9.
[30] Coleman JJ, Young JD, Garg A. Semiconductor quantum dot lasers: a tutorial. J Lightwave Technol 2011;29:499–510.
[31] Newell TC, Bossert DJ, Stintz A, Fuchs B, Malloy KJ, Lester LF. Gain and linewidth enhancement factor in InAs quantum-dot laser diodes. IEEE Photon Technol Lett 1999;11:1527–9.
[32] Bimberg D, Kirstaedter N, Ledentsov NN, Alferov ZI, Kop'ev PS, Ustinov VM. InGaAs-GaAs quantum-dot lasers. IEEE J Sel Top Quant Electron 1997;3:196–205.
[33] Duan J, Huang H, Jung D, Zhang Z, Norman J, Bowers JE, Grillot F. Semiconductor quantum dot lasers epitaxially grown on silicon with low linewidth enhancement factor. Appl Phys Lett 2018;112:251111.
[34] Capua A, Rozenfeld L, Mikhelashvili V, Eisenstein G, Kuntz M, Laemmlin M, Bimberg D. Direct correlation between a highly damped modulation response and ultra low relative intensity noise in an InAs/GaAs quantum dot laser. Opt Exp 2007;15:5388–93.
[35] Liao M, Chen S, Liu Z, Wang Y, Ponnampalam L, Zhou Z, Wu J, Tang M, Shutts S, Liu Z, Smowton PM, Yu S, Seeds A, Liu H. Low-noise 1.3 μm InAs/GaAs quantum dot laser monolithically grown on silicon. Photon Res 2018;6:1062–6.
[36] Liu AY, Komljenovic T, Davenport ML, Gossard AC, Bowers JE. Reflection sensitivity of 1.3 μm quantum dot lasers epitaxially grown on silicon. Opt Exp 2017;25:9535–43.
[37] O'Brien D, Hegarty SP, Huyet G, McInerney JG, Kettler T, Laemmlin M, Bimberg D, Ustinov VM, Zhukov AE, Mikhrin SS, Kovsh AR. Feedback sensitivity of 1.3/spl mu/m InAs/GaAs quantum dot lasers. Electron Lett 2003;39:1819–20.
[38] Duan J, Huang H, Dong B, Jung D, Norman JC, Bowers JE, Grillot F. 1.3-μm reflection insensitive InAs/GaAs quantum dot lasers directly grown on silicon. IEEE Photon Technol Lett 2019;31:345–8.
[39] Helms J, Petermann K. A simple analytic expression for the stable operation range of laser diodes with optical feedback. IEEE J Quant Electron 1990;26:833–6.

[40] Chen S, Li W, Wu J, Jiang Q, Tang M, Shutts S, Elliott SN, Sobiesierski A, Seeds AJ, Ross I, Smowton PM, Liu H. Electrically pumped continuous-wave III-V quantum dot lasers on silicon. Nat Photon 2016;10:307–11.
[41] Wu J, Tang M, Liu H, III- V. Quantum dot lasers epitaxially grown on Si substrates. In: Nanoscale semiconductor lasers. Elsevier; 2019. p. 17–39 [Chapter 2].
[42] Jung D, Herrick R, Norman J, Turnlund K, Jan C, Feng K, Gossard AC, Bowers JE. Impact of threading dislocation density on the lifetime of InAs quantum dot lasers on Si. Appl Phys Lett 2018;112:153507.
[43] Wu J, Chen S, Seeds A, Liu H. Quantum dot optoelectronic devices: lasers, photodetectors and solar cells. J Phys D Appl Phys 2015;48:363001.
[44] Windhorn TH, Metze GM, Tsaur B, Fan JCC. AlGaAs double-heterostructure diode lasers fabricated on a monolithic GaAs/Si substrate. Appl Phys Lett 1984; 45:309–11.
[45] Deppe DG, Holonyak N, Nam DW, Hsieh KC, Jackson GS, Matyi RJ, Shichijo H, Epler JE, Chung HF. Room-temperature continuous operation of p - n Al x Ga 1− x As-GaAs quantum well heterostructure lasers grown on Si. Appl Phys Lett 1987;51: 637–9.
[46] Huang X, Song Y, Masuda T, Jung D, Lee M. InGaAs/GaAs quantum well lasers grown on exact GaP/Si (001). Electron Lett 2014;50:1226–7.
[47] Wang J, Ren X, Deng C, Hu H, He Y, Cheng Z, Ma H, Wang Q, Huang Y, Duan X, Yan X. Extremely low-threshold current density InGaAs/AlGaAs quantum-well lasers on silicon. J Lightwave Technol 2015;33:3163–9.
[48] Kryzhanovskaya NV, Moiseev EI, Polubavkina YS, Maximov MV, Kulagina MM, Troshkov SI, Zadiranov YM, Lipovskii AA, Baidus NV, Dubinov AA, Krasilnik ZF, Novikov AV, Pavlov DA, Rykov AV, Sushkov AA, Yurasov DV, Zhukov AE. Electrically pumped InGaAs/GaAs quantum well microdisk lasers directly grown on Si(100) with Ge/GaAs buffer. Opt Exp 2017;25:16754–60.
[49] Goldstein L, Glas F, Marzin JY, Charasse MN, Le Roux G. Growth by molecular beam epitaxy and characterization of InAs/GaAs strained-layer superlattices. Appl Phys Lett 1985;47:1099–101.
[50] Moison JM, Houzay F, Barthe F, Leprince L, André E, Vatel O. Self-organized growth of regular nanometer-scale InAs dots on GaAs. Appl Phys Lett 1994;64: 196–8.
[51] Oshinowo J, Nishioka M, Ishida S, Arakawa Y. Highly uniform InGaAs/GaAs quantum dots (∼15 nm) by metalorganic chemical vapor deposition. Appl Phys Lett 1994; 65:1421–3.
[52] Kirstaedter N, Ledentsov NN, Grundmann M, Bimberg D, Ustinov VM, Ruvimov SS, Maximov MV, Kop'ev PS, Alferov ZI, Werner P, Gösele U, Heydenreich J, Richter U. Low threshold, large T/sub o/injection laser emission from (InGa)As quantum dots. Electron Lett 1994;30:1416–7.
[53] Kamath K, Bhattacharya P, Sosnowski T, Norris T, Phillips J. Room-temperature operation of In0.4Ga0.6As/GaAs self-organised quantum dot lasers. Electron Lett 1996;32:1374–5.
[54] Nötzel R. Self-organized growth of quantum-dot structures. Semicond Sci Technol 1996;11:1365–79.
[55] García JM, Medeiros-Ribeiro G, Schmidt K, Ngo T, Feng JL, Lorke A, Kotthaus J, Petroff PM. Intermixing and shape changes during the formation of InAs self-assembled quantum dots. Appl Phys Lett 1997;71:2014–6.
[56] Ustinov VM, Maleev NA, Zhukov AE, Kovsh AR, Egorov AY, Lunev AV, Volovik BV, Krestnikov IL, Musikhin YG, Bert NA, Kop'ev PS, Alferov ZI, Ledentsov NN, Bimberg D. InAs/InGaAs quantum dot structures on GaAs substrates emitting at 1.3 μm. Appl Phys Lett 1999;74:2815–7.

[57] Wasilewski ZR, Fafard S, McCaffrey JP. Size and shape engineering of vertically stacked self-assembled quantum dots. J Cryst Growth 1999;201:1131–5.
[58] Egawa T, Ogawa A, Jimbo T, Umeno M. AlGaAs/GaAs laser diodes with GaAs islands active regions on Si grown by droplet epitaxy. Jpn J Appl Phys 1998;37: 1552–5. Part 1.
[59] Wang T, Liu H, Lee A, Pozzi F, Seeds A. 1.3-μm InAs/GaAs quantum-dot lasers monolithically grown on Si substrates. Opt Exp 2011;19:11381–6.
[60] Liu Z, Hantschmann C, Tang M, Lu Y, Park JS, Liao M, Pan S, Sanchez A, Beanland R, Martin M, Baron T, Chen S, Seeds A, Penty R, White I, Liu H. Origin of defect tolerance in InAs/GaAs quantum dot lasers grown on silicon. J Lightwave Technol 2020;38:240–8.
[61] Hamers RJ, Tromp RM, Demuth JE. Scanning tunneling microscopy of Si(001). Phys Rev B 1986;34:5343–57.
[62] Alcotte R, Martin M, Moeyaert J, Cipro R, David S, Bassani F, Ducroquet F, Bogumilowicz Y, Sanchez E, Ye Z, Bao XY, Pin JB, Baron T. Epitaxial growth of antiphase boundary free GaAs layer on 300 mm Si(001) substrate by metalorganic chemical vapour deposition with high mobility. APL Mater 2016;4:046101.
[63] Fang SF, Adomi K, Iyer S, Morkoç H, Zabel H, Choi C, Otsuka N. Gallium arsenide and other compound semiconductors on silicon. J Appl Phys 1990;68:R31–58.
[64] Yang VK, Groenert M, Leitz CW, Pitera AJ, Currie MT, Fitzgerald EA. Crack formation in GaAs heteroepitaxial films on Si and SiGe virtual substrates. J Appl Phys 2003;93:3859–65.
[65] Fischer R, Neuman D, Zabel H, Morkoç H, Choi C, Otsuka N. Dislocation reduction in epitaxial GaAs on Si(100). Appl Phys Lett 1986;48:1223–5.
[66] Kawabe M, Ueda T. Molecular beam epitaxy of controlled single domain GaAs on si (100). Jpn J Appl Phys 1986;25:L285–7.
[67] Kawabe M, Ueda T. Self-annihilation of antiphase boundary in GaAs on Si(100) grown by molecular beam epitaxy. Jpn J Appl Phys 1987;26:L944–6.
[68] Liao M, Chen S, Huo S, Chen S, Wu J, Tang M, Kennedy K, Li W, Kumar S, Martin M, Baron T, Jin C, Ross I, Seeds A, Liu H. Monolithically integrated electrically pumped continuous-wave III-V quantum dot light sources on silicon. IEEE J Sel Top Quant Electron 2017;23:1–10.
[69] Linder KK, Phillips J, Qasaimeh O, Liu XF, Krishna S, Bhattacharya P, Jiang JC. Self-organized In0.4Ga0.6As quantum-dot lasers grown on Si substrates. Appl Phys Lett 1999;74:1355–7.
[70] Groenert ME, Leitz CW, Pitera AJ, Yang V, Lee H, Ram RJ, Fitzgerald EA. Monolithic integration of room-temperature cw GaAs/AlGaAs lasers on Si substrates via relaxed graded GeSi buffer layers. J Appl Phys 2003;93:362–7.
[71] Lee AD, Jiang Q, Tang M, Zhang Y, Seeds AJ, Liu H. InAs/GaAs quantum-dot lasers monolithically grown on Si, Ge, and Ge-on-Si substrates. IEEE J Sel Top Quant Electron 2013;19:1901107.
[72] Lee A, Jiang Q, Tang M, Seeds A, Liu H. Continuous-wave InAs/GaAs quantum-dot laser diodes monolithically grown on Si substrate with low threshold current densities. Opt Exp 2012;20:22181–7.
[73] Liu AY, Zhang C, Norman J, Snyder A, Lubyshev D, Fastenau JM, Liu AWK, Gossard AC, Bowers JE. High performance continuous wave 1.3 μ m quantum dot lasers on silicon. Appl Phys Lett 2014;104:041104.
[74] Liu AY, Srinivasan S, Norman J, Gossard AC, Bowers JE. Quantum dot lasers for silicon photonics. Photon Res 2015;3:B1–9.
[75] Liu AY, Herrick RW, Ueda O, Petroff PM, Gossard AC, Bowers JE. Reliability of InAs/GaAs quantum dot lasers epitaxially grown on silicon. IEEE J Sel Top Quant Electron 2015;21:690–7.

[76] Li Q, Ng KW, Lau KM. Growing antiphase-domain-free GaAs thin films out of highly ordered planar nanowire arrays on exact (001) silicon. Appl Phys Lett 2015; 106:072105.
[77] Wan Y, Li Q, Geng Y, Shi B, Lau KM. InAs/GaAs quantum dots on GaAs-on-V-grooved-Si substrate with high optical quality in the 1.3 μ m band. Appl Phys Lett 2015;107:081106.
[78] Li Q, Jiang H, Lau KM. Coalescence of planar GaAs nanowires into strain-free three-dimensional crystals on exact (001) silicon. J Cryst Growth 2016;454:19–24.
[79] Martin M, Caliste D, Cipro R, Alcotte R, Moeyaert J, David S, Bassani F, Cerba T, Bogumilowicz Y, Sanchez E, Ye Z, Bao XY, Pin JB, Baron T, Pochet P. Toward the III-V/Si co-integration by controlling the biatomic steps on hydrogenated Si(001). Appl Phys Lett 2016;109:253103.
[80] Volz K, Beyer A, Witte W, Ohlmann J, Nmeth I, Kunert B, Stolz W. GaP--nucleation on exact Si (0 0 1) substrates for III/V device integration. J Cryst Growth 2011;315:37–47.
[81] Grassman TJ, Carlin JA, Galiana B, Yang LM, Yang F, Mills MJ, Ringel SA. Nucleation-related defect-free GaP/Si(100) heteroepitaxy via metal-organic chemical vapor deposition. Appl Phys Lett 2013;102:142102.
[82] Beyer A, Ohlmann J, Liebich S, Heim H, Witte G, Stolz W, Volz K, Beyer A, Ohlmann J, Liebich S, Heim H, Witte G, Stolz W, Volz K. GaP heteroepitaxy on Si (001): correlation of Si-surface structure, GaP growth conditions, and Si-III/V interface structure. J Appl Phys 2012;111:083534.
[83] Warren EL, Kibbler AE, France RM, Norman AG, Stradins P, McMahon WE. Growth of antiphase-domain-free GaP on Si substrates by metalorganic chemical vapor deposition using an in situ AsH3 surface preparation. Appl Phys Lett 2015; 107:082109.
[84] Kwoen J, Jang B, Lee J, Kageyama T, Watanabe K, Arakawa Y. All MBE grown InAs/GaAs quantum dot lasers on on-axis Si (001). Opt Exp 2018;26:11568–76.
[85] Kwoen J, Lee J, Watanabe K, Arakawa Y. Elimination of anti-phase boundaries in a GaAs layer directly-grown on an on-axis Si(001) substrate by optimizing an AlGaAs nucleation layer. Jpn J Appl Phys 2019;58:SBBE07.
[86] Beanland R, Dunstan DJ, Goodhew PJ. Plastic relaxation and relaxed buffer layers for semiconductor epitaxy. Adv Phys 1996;45:87–146.
[87] Mi Z, Yang J, Bhattacharya P, Huffaker DL. Self-organised quantum dots as dislocation filters: the case of GaAs-based lasers on silicon. Electron Lett 2006;42:121–3.
[88] Yang J, Bhattacharya P, Mi Z. High-performance In0.5Ga0.5As/GaAs quantum-dot lasers on silicon with multiple-layer quantum-dot dislocation filters. IEEE Trans Electron Dev 2007;54:2849–55.
[89] Nozawa K, Horikoshi Y. Low threading dislocation density GaAs on si(100) with InGaAs/GaAs strained-layer superlattice grown by migration-enhanced epitaxy. Jpn J Appl Phys 1991;30:L668–71.
[90] Tang M, Chen S, Wu J, Jiang Q, Dorogan VG, Benamara M, Mazur YI, Salamo GJ, Seeds A, Liu H. 1.3-μm InAs/GaAs quantum-dot lasers monolithically grown on Si substrates using InAlAs/GaAs dislocation filter layers. Opt Exp 2014;22:11528–35.
[91] Tang M, Chen S, Wu J, Jiang Q, Kennedy K, Jurczak P, Liao M, Beanland R, Seeds A, Liu H. Optimizations of defect filter layers for 1.3-μm InAs/GaAs quantum-dot lasers monolithically grown on Si substrates. IEEE J Sel Top Quant Electron 2016;22:50–6.
[92] Akiyama M, Kawarada Y, Kaminishi K. Growth of single domain GaAs layer on (100)-oriented si substrate by MOCVD. Jpn J Appl Phys 1984;23:L843–5.
[93] Yamaguchi M, Yamamoto A, Tachikawa M, Itoh Y, Sugo M. Defect reduction effects in GaAs on Si substrates by thermal annealing. Appl Phys Lett 1988;53: 2293–5.

[94] Okamoto H, Watanabe Y, Kadota Y, Ohmachi Y. Dislocation reduction in GaAs on Si by thermal cycles and InGaAs/GaAs strained-layer superlattices. Jpn J Appl Phys 1987;26:L1950–2.
[95] Yamaguchi M, Tachikawa M, Itoh Y, Sugo M, Kondo S. Thermal annealing effects of defect reduction in GaAs on Si substrates. J Appl Phys 1990;68:4518–22.
[96] Yamaguchi M, Tachikawa M, Sugo M, Kondo S, Itoh Y. Analysis for dislocation density reduction in selective area grown GaAs films on Si substrates. Appl Phys Lett 1990;56:27–9.
[97] Bolkhovityanov YB, Pchelyakov OP. GaAs epitaxy on Si substrates: modern status of research and engineering. Phys Usp 2008;51:437–56.
[98] Yamaguchi M. Dislocation density reduction in heteroepitaxial III-V compound films on Si substrates for optical devices. J Mater Res 1991;6:376–84.
[99] Guo W, Date L, Pena V, Bao X, Merckling C, Waldron N, Collaert N, Caymax M, Sanchez E, Vancoille E, Barla K, Thean A, Eyben P, Vandervorst W. Selective metal-organic chemical vapor deposition growth of high quality GaAs on Si(001). Appl Phys Lett 2014;105:062101.
[100] Oh S, Jun DH, Shin KW, Choi IH, Jung SH, Choi JH, Park W, Park Y, Yoon E. Control of crack formation for the fabrication of crack-free and self-isolated high-efficiency gallium arsenide photovoltaic cells on silicon substrate. IEEE J Photovolt 2016;6:1031–5.
[101] Huang H, Ren X, Lv J, Wang Q, Song H, Cai S, Huang Y, Qu B. Crack-free GaAs epitaxy on Si by using midpatterned growth: application to Si-based wavelength-selective photodetector. J Appl Phys 2008;104:113114.
[102] Yang J, Jurczak P, Cui F, Li K, Tang M, Billiald L, Beanland R, Sanchez AM, Liu H. Thin Ge buffer layer on silicon for integration of III-V on silicon. J Cryst Growth 2019;514:109–13.
[103] Liu AY, Peters J, Huang X, Jung D, Norman J, Lee ML, Gossard AC, Bowers JE. Electrically pumped continuous-wave 1.3 μm quantum-dot lasers epitaxially grown on on-axis (001) GaP/Si. Opt Lett 2017;42:338–41.
[104] Jung D, Norman J, Kennedy MJ, Shang C, Shin B, Wan Y, Gossard AC, Bowers JE. High efficiency low threshold current 1.3 μ m InAs quantum dot lasers on on-axis (001) GaP/Si. Appl Phys Lett 2017;111:122107.
[105] Jung D, Zhang Z, Norman J, Herrick R, Kennedy MJ, Patel P, Turnlund K, Jan C, Wan Y, Gossard AC, Bowers JE. Highly reliable low-threshold InAs quantum dot lasers on on-Axis (001) Si with 87% injection efficiency. ACS Photon 2018;5:1094–100.
[106] Chen S, Liao M, Tang M, Wu J, Martin M, Baron T, Seeds A, Liu H. Electrically pumped continuous-wave 1.3 μm InAs/GaAs quantum dot lasers monolithically grown on on-axis Si (001) substrates. Opt Exp 2017;25:4632–9.
[107] Li K, Liu Z, Tang M, Liao M, Kim D, Deng H, Sanchez AM, Beanland R, Martin M, Baron T, Chen S, Wu J, Seeds A, Liu H. O-band InAs/GaAs quantum dot laser monolithically integrated on exact (0 0 1) Si substrate. J Cryst Growth 2019;511:56–60.
[108] Norman J, Kennedy MJ, Selvidge J, Li Q, Wan Y, Liu AY, Callahan PG, Echlin MP, Pollock TM, Lau KM, Gossard AC, Bowers JE. Electrically pumped continuous wave quantum dot lasers epitaxially grown on patterned, on-axis (001) Si. Opt Exp 2017;25:3927–34.
[109] Shang C, Wan Y, Norman JC, Collins N, MacFarlane I, Dumont M, Liu S, Li Q, Lau KM, Gossard AC, Bowers JE. Low-threshold epitaxially grown 1.3-μm InAs quantum dot lasers on patterned (001) Si. IEEE J Sel Top Quant Electron 2019;25:1–7.
[110] Kwoen J, Jang B, Watanabe K, Arakawa Y. High-temperature continuous-wave operation of directly grown InAs/GaAs quantum dot lasers on on-axis Si (001). Opt Express 2019;27:2681–8.

[111] Zhu S, Shi B, Li Q, Lau KM. 1.5 μm quantum-dot diode lasers directly grown on CMOS-standard (001) silicon. Appl Phys Lett 2018;113:221103.
[112] Zhu S, Shi B, Lau KM. Electrically pumped 1.5 μm InP-based quantum dot microring lasers directly grown on (001) Si. Opt Lett 2019;44:4566–9.
[113] Shi B, Zhu S, Li Q, Tang CW, Wan Y, Hu EL, Lau KM. 1.55 um Room-Temperature Lasing From Subwavelength Quantum-Dot Microdisks Directly Grown on (001) Si. Appl Phys Lett 2017;110:121109.
[114] Sugo M, Mori H, Itoh Y, Sakai Y, Tachikawa M. 1.5 μm-Long-Wavelength multiple quantum well laser on a Si substrate. Jpn J Appl Phys 1991;30:3876–8.
[115] Bhowmick S, Baten MZ, Frost T, Ooi BS, Bhattacharya P. High performance InAs/In0.53Ga0.23Al 0.24As/InP quantum dot 1.55 μm tunnel injection laser. IEEE J Quant Electron 2014;50:7–14.
[116] Heck MJR, Bauters JF, Davenport ML, Doylend JK, Jain S, Kurczveil G, Srinivasan S, Tang Y, Bowers JE. Hybrid silicon photonic integrated circuit technology. IEEE J Sel Top Quant Electron 2013;19:6100117.
[117] Pukhrambam PD. Wavelength division multiplexing laser arrays for applications in optical networking and sensing: overview and perspectives. Jpn J Appl Phys 2018;57:08PA03.
[118] Wang Y, Chen S, Yu Y, Zhou L, Liu L, Yang C, Liao M, Tang M, Liu Z, Wu J, Li W, Ross I, Seeds AJ, Liu H, Yu S. Monolithic quantum-dot distributed feedback laser array on silicon. Optica 2018;5:528–33.
[119] Wan Y, Norman JC, Tong Y, Kennedy MJ, He W, Selvidge J, Shang C, Dumont M, Malik A, Tsang HK, Gossard AC, Bowers JE. 1.3 μm quantum dot-distributed feedback lasers directly grown on (001) Si. Laser Photon Rev 2020;14:2000037.
[120] Coldren LA, Fish GA, Akulova Y, Barton JS, Johansson L, Coldren CW. Tunable semiconductor lasers: a tutorial. J Lightwave Technol 2004;22:193–202.
[121] Keyvaninia S, Roelkens G, Van Thourhout D, Jany C, Lamponi M, Le Liepvre A, Lelarge F, Make D, Duan G-H, Bordel D, Fedeli J-M. Demonstration of a heterogeneously integrated III-V/SOI single wavelength tunable laser. Opt Exp 2013;21:3784–92.
[122] Ferrotti T, Blampey B, Jany C, Duprez H, Chantre A, Boeuf F, Seassal C, Ben Bakir B. Co-integrated 1.3μm hybrid III-V/silicon tunable laser and silicon Mach-Zehnder modulator operating at 25Gb/s. Opt Exp 2016;24:30379–401.
[123] Tran MA, Huang D, Guo J, Komljenovic T, Morton PA, Bowers JE. Ring-resonator based widely-tunable narrow-linewidth Si/InP integrated lasers. IEEE J Sel Top Quant Electron 2020;26:1–14.
[124] Tsang WT. The cleaved-couple-cavity (C3) laser. Semiconduct Semimet 1985;22:257–373 [Chapter 5].
[125] Wan Y, Zhang S, Norman JC, Kennedy M, He W, Tong Y, Shang C, He J, Tsang HK, Gossard AC, Bowers JE. Directly modulated single-mode tunable quantum dot lasers at 1.3 μm. Laser Photon Rev 2020;14:1900348.
[126] Wan Y, Zhang S, Norman JC, Kennedy MJ, He W, Liu S, Xiang C, Shang C, He J-J, Gossard AC, Bowers JE. Tunable quantum dot lasers grown directly on silicon. Optica 2019;6:1394–400.
[127] He J-J, Liu D. Wavelength switchable semiconductor laser using half-wave V-coupled cavities. Opt Exp 2008;16:3896–911.
[128] Thompson MG, Rae AR, Xia M, Penty RV, White IH. InGaAs quantum-dot mode-locked laser diodes. IEEE J Sel Top Quant Electron 2009;15:661–72.
[129] Liu S, Wu X, Jung D, Norman JC, Kennedy MJ, Tsang HK, Gossard AC, Bowers JE. High-channel-count 20 GHz passively mode-locked quantum dot laser directly grown on Si with 41 Tbit/s transmission capacity. Optica 2019;6:128–34.
[130] Vahala KJ. Optical microcavities. Nature 2003;424:839–46.

[131] Wan Y, Li Q, Liu AY, Gossard AC, Bowers JE, Hu EL, Lau KM. Optically pumped 1.3 μm room-temperature InAs quantum-dot micro-disk lasers directly grown on (001) silicon. Opt Lett 2016;41:1664–7.
[132] Wan Y, Li Q, Liu AY, Chow WW, Gossard AC, Bowers JE, Hu EL, Lau KM. Sub-wavelength InAs quantum dot micro-disk lasers epitaxially grown on exact Si (001) substrates. Appl Phys Lett 2016;108:221101.
[133] Wan Y, Li Q, Liu AY, Gossard AC, Bowers JE, Hu EL, Lau KM. Temperature characteristics of epitaxially grown InAs quantum dot micro-disk lasers on silicon for on-chip light sources. Appl Phys Lett 2016;109:011104.
[134] Li Q, Wan Y, Liu AY, Gossard AC, Bowers JE, Hu EL, Lau KM. 1.3-μm InAs quantum-dot micro-disk lasers on V-groove patterned and unpatterned (001) silicon. Opt Exp 2016;24:21038–45.
[135] Shi B, Zhu S, Li Q, Wan Y, Hu EL, Lau KM. Continuous-wave optically pumped 1.55 μm InAs/InAlGaAs quantum dot microdisk lasers epitaxially grown on silicon. ACS Photon 2017;4:204–10.
[136] Kryzhanovskaya N, Moiseev E, Polubavkina Y, Maximov M, Kulagina M, Troshkov S, Zadiranov Y, Guseva Y, Lipovskii A, Tang M, Liao M, Wu J, Chen S, Liu H, Zhukov A. Heat-sink free CW operation of injection microdisk lasers grown on Si substrate with emission wavelength beyond 1.3 μm. Opt Lett 2017;42: 3319–22.
[137] Zhu S, Shi B, Li Q, Wan Y, Lau KM. Parametric study of high-performance 1.55 μm InAs quantum dot microdisk lasers on Si. Opt Exp 2017;25:31281–93.
[138] Kryzhanovskaya NV, Moiseev EI, Polubavkina YS, Maximov MV, Mokhov DV, Morozov IA, Kulagina MM, Zadiranov YM, Lipovskii AA, Tang M, Liao M, Wu J, Chen S, Liu H, Zhukov AE. Elevated temperature lasing from injection micro-disk lasers on silicon. Laser Phys Lett 2018;15:015802.
[139] Zhang B, Wei WQ, Wang JH, Wang HL, Zhao Z, Liu L, Cong H, Feng Q, Liu H, Wang T, Zhang JJ. O-band InAs/GaAs quantum-dot microcavity laser on Si (001) hollow substrate by in-situ hybrid epitaxy. AIP Adv 2019;9:015331.
[140] Zhou T, Tang M, Xiang G, Fang X, Liu X, Xiang B, Hark S, Martin M, Touraton M-L, Baron T, Lu Y, Chen S, Liu H, Zhang Z. Ultra-low threshold InAs/GaAs quantum dot microdisk lasers on planar on-axis Si (001) substrates. Optica 2019;6:430–5.
[141] Wan Y, Norman J, Li Q, Kennedy MJ, Liang D, Zhang C, Huang D, Zhang Z, Liu AY, Torres A, Jung D, Gossard AC, Hu EL, Lau KM, Bowers JE. 1.3 μm sub-milliamp threshold quantum dot micro-lasers on Si. Optica 2017;4:940–4.
[142] Wan Y, Jung D, Norman J, Shang C, MacFarlane I, Li Q, Kennedy MJ, Gossard AC, Lau KM, Bowers JE. O-band electrically injected quantum dot micro-ring lasers on on-axis (001) GaP/Si and V-groove Si. Opt Exp 2017;25:26853–60.
[143] Inoue D, Jung D, Norman J, Wan Y, Nishiyama N, Arai S, Gossard AC, Bowers JE. Directly modulated 1.3 μm quantum dot lasers epitaxially grown on silicon. Opt Exp 2018;26:7022–33.
[144] Zhou T, Tang M, Xiang G, Xiang B, Hark S, Martin M, Baron T, Pan S, Park JS, Liu Z, Chen S, Zhang Z, Liu H. Continuous-wave quantum dot photonic crystal lasers grown on on-axis Si (001). Nat Commun 2020;11:1–7.
[145] Driscoll JB, Doussiere P, Islam S, Narayan R, Lin W, Mahalingam H, Park JS, Lin Y, Nguyen K, Roelofs K, Dahal A, Venables R, Liao L, Jones R, Zhu D, Priyadarshi S, Parthasarathy B, Akulova Y. First 400G 8-channel CWDM silicon photonic integrated transmitter. IEEE Int Conf Group IV Photonics GFP 2018:1–2.
[146] Liang D, Huang X, Kurczveil G, Fiorentino M, Beausoleil RG. Integrated finely tunable microring laser on silicon. Nat Photon 2016;10:719–22.

[147] Sui S-S, Tang M-Y, Yang Y-D, Xiao J-L, Du Y, Huang Y-Z. Hybrid spiral-ring microlaser vertically coupled to silicon waveguide for stable and unidirectional output. Opt Lett 2015;40:4995–8.
[148] Mechet P, Verstuyft S, de Vries T, Spuesens T, Regreny P, Van Thourhout D, Roelkens G, Morthier G. Unidirectional III-V microdisk lasers heterogeneously integrated on SOI. Opt Exp 2013;21:19339–52.
[149] Feng L, Wong ZJ, Ma RM, Wang Y, Zhang X. Single-mode laser by parity-time symmetry breaking. Science 2014;346:972–5.
[150] Hodaei H, Miri M-A, Heinrich M, Christodoulides DN, Khajavikhan M. Parity-time-symmetric microring lasers. Science 2014;346:975–8.
[151] Arbabi A, Kamali SM, Arbabi E, Griffin BG, Goddard LL. Grating integrated single mode microring laser. Opt Exp 2015;23:5335–47.

CHAPTER TWELVE

Physics and applications of semiconductor nanowire lasers

Patrick Parkinson*
Department of Physics and Astronomy and the Photon Science Institute, University of Manchester, Manchester, United Kingdom
*Corresponding author: patrick.parkinson@manchester.ac.uk

1. Introduction

Semiconductor lasers based on bottom-up grown nanowires have been produced and characterised since the early 2000s as high-brightness, tunable and nanoscale coherent light sources which can be integrated into novel devices [1,2]. Their development forms part of the story of lasers from the gas [3] and homojunction lasers [4,5] of the 1960s, the heterostructured lasers [6,7] of the 1970s, the development of quantum-well lasers [8] in the 1980s and quantum dot lasers [9] in the 1990s. Over this period, the development of physics (e.g., plasmonics [10,11]), materials (e.g., perovskite nanowire lasers [12]) and applications (e.g., for photonic neuromorphic networks [13]) has grown hand-in-hand. For insight into developments in specific nanowire lasers systems, the reader is directed towards one of the many excellent review articles which have been published covering materials such as zinc oxide (ZnO) [14–16], III-Vs [17], perovskites [18,19], nitrides [20] or applications such as integration [21], tunability [22] or electrical injection [23].

Despite this development, the motivation for using the nanowire architecture remains unchanged. Semiconductor nanowires are produced from a single material or a heterostructure of semiconducting materials, typically with a defect-free crystal structure, with diameters ranging from 100 nm to 1 μm, and lengths ranging from 1 μm to tens of μm. Crucially, the nanowire forms both the laser cavity and the gain material, with the wire end-facets often acting as mirrors forming a Fabry–Pérot cavity [24] as shown in Fig. 12.1. As such, nanowire lasers approach the smallest possible physical volume for a photonic laser, on the order of the wavelength of light produced, creating opportunities for use as near-field and thresholdless coherent light sources.

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00010-1
© 2021 Elsevier Ltd.
All rights reserved.

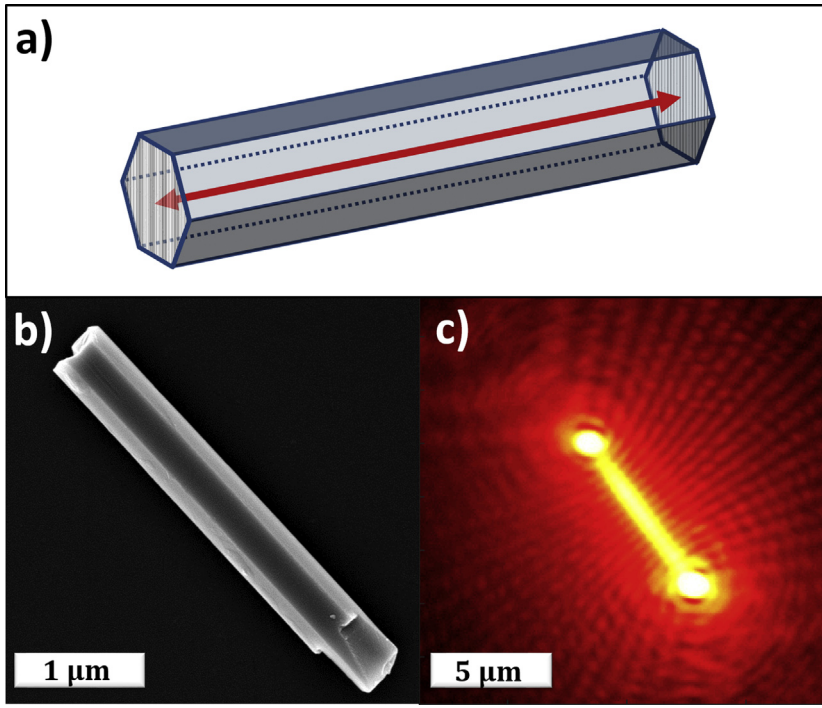


Figure 12.1 A single nanowire as a Fabry–Pérot laser cavity; (a) a schematic of the resonant cavity formed by reflection at the end facets, (b) Scanning electron microscopy image of a typical III-V nanowire transferred from the growth substrate to a test substrate and (c) a far-field image of a lasing III-V nanowire under optical excitation, showing characteristic interference fringing.

The use of nanowires also confers the processing advantages of this technology, including high scalability of growth [25], pick-and-place fabrication [26] and the ability to grow wires onto lattice-mismatched substrates due to strain relaxation [27,28]. Nanowires can be measured as-grown [29–31], transferred to a low-index substrate [32,33], embedded in a matrix [34] or transferred to a photonic circuit [26]. Whilst optical pumping is a simple route for characterisation and to achieve lasing in un-contacted nanowires, electrical pumping is the major target for development, and has been demonstrated for device applications [35–37].

Nanowire lasers have applications spanning lighting, computing, communication, sensing and in nanophotonics. In particular, three fields have emerged where nanowire lasers provide a clear advantage:

On-chip photonics. On-chip photonic or optoelectronic integrated circuitry uses photons as the fundamental information carrier, as opposed to electrons

used in conventional devices. The use of photonics potentially alleviates bandwidth and power density issues [38] and can be used to enable post-von Neumann architectures [13]; however, integrating conventional electronic circuitry (based on silicon and complementary metal-oxide-semiconductor (CMOS) processes) with optical emission elements (based on direct bandgap materials) is challenging. In particular, integrating laser light sources has proven to be the most significanat obstruction to the development of this technology; conventional large-area heteroepitaxy is limited by lattice mismatch, and flip-chip approaches [39] are slow and costly fabrication steps. Whilst off-chip laser solutions have been demonstrated [13,40], semiconductor nanowires provide a potential on-chip solution as they can be grown directly onto silicon [17,31,41,42], a transparent waveguide [43] or integrated via pick-and-place [26] or atomic force microscopy (AFM) technology [44].

Chip-to-chip communication. For medium-to-long distance communication, telecoms-band coherent sources are required to match the fibre optical bands at 1310 and 1550 nm [45]. At present, chip-to-chip communication requires high-speed connection over copper wires, giving rise to energy inefficiencies, the potential for interference and heating which scales with distance. Semiconductor nanowires were initially proposed as on-chip, coherent infrared light sources in the mid-2000s [46]. However, the development of room-temperature telecoms-band nanowire lasers has accelerated recently, with array-based [47–49], photonic crystal-based [50] and single-wire light sources [31,42,51,52] having been used to meet this technological demand. Recent developments have sought to exploit multiple axial quantum dots to approach the ultimate goal of telecoms band lasing in single nanowires [53].

Biological study. The use of semiconductor nanowires for biological applications has long been established, due to their high surface-to-volume ratio for sensing [54], highly multiplexed detection [55], as topographical guides [56] or as intracellular probes [57]. Nanowire lasers provide highly localised, nanoscale light sources which can safely pass through the cell membrane [58]. As opposed to passive nanowire lightguides [59], fully active lasers can be entirely embedded and externally optically pumped to provide a measure of intracellular refractive index change [60].

In Section 2, the basic properties and physics of nanoscale laser systems are reviewed. Research on optoelectronic nanowires is discussed in Section 3; nanowire lasers based on ZnO, III-nitrides, chalcogenide, III-V, organic materials and perovskite semiconductors are introduced along with heterostructuring and growth and fabrication. Current research topics in nanowire

lasers are introduced in Section 4, including array-based and plasmonic nanolaser structures. The current state of the art and an outlook is discussed in Section 5, including approaches for room-temperature, continuously lasing, telecom-wavelength electrically pumped wire growth.

2. Lasers

The essential components of a laser have been known since the 1950s [3,61]; an active material providing photon gain through stimulated emission of radiation, and a feedback mechanism, often a Fabry–Pérot cavity consisting of two or more mirrors. To achieve lasing, the losses incurred for one round trip of the cavity must be compensated by gain in the active material, and the cavity must be able to support a standing wave. For a complete description of laser techniques and technology, the reader is directed towards standard texts [62] and a more theoretical approach for nanowires is given in Ref. [63]; what follows is a brief outline of laser physics relevant to nanowire laser systems.

2.1 Laser fundamentals

We consider the propagation of an electromagnetic wave $E_0(z) = E_z(z)E^T$ with arbitrary transverse structure E^T, in a cavity lying in the z direction with a length of L, complex propagation constant $\widetilde{k}$ and two end mirrors with reflectivity R_1 and R_2. In a Fabry–Pérot cavity, a round-trip propagation leads to a field that can be expressed as follows:

$$E(z=2L) = E_0 R_1 R_2 \exp\left(i\widetilde{k}2L\right) \tag{12.1}$$

By separating the propagation constant into real and imaginary components – corresponding to phase constant β and loss term κ, respectively – we can write $\widetilde{k} = \beta + i\kappa$, leading to two conditions for lasing where round-trip amplitude and phase changes are compensated and $E(z = 2L) = E_0$.

$$R_1 R_2 \exp(-2L\kappa) = 1 \tag{12.2}$$

$$\Im[\exp(i2L\beta)] = 0 \tag{12.3}$$

These two conditions can be treated separately.

2.1.1 Threshold gain

The condition given in Eq. (12.2) describes the requirement on gain and loss within the gain material. The loss $\kappa = \alpha - g$, where g is gain and α is distributed loss; at threshold of operation the gain g_{th} must compensate both distributed losses α and end-facet losses, such that

$$g_{th} = \alpha - \frac{\ln R_1 R_2}{2L} \tag{12.4}$$

The distributed losses term α refers to losses associated with both intra-cavity processes (e.g., reabsorption, scattering) and extra-cavity processes associated with the evanescent field extending into the surrounding medium. It is notable that in the case where reflectivity is high $(R_{1,2} \approx 1)$ the material gain required for lasing tends towards α and does not scale with cavity length; conversely, if $R_{1,2} \ll 1$, the gain required scales inversely with cavity length. The design for gain in a nanowire is described in Section 2.1.3.

2.1.2 Laser Cavity and waveguides

The second condition – Eq. (12.3) – determines the longitudinal modes that can be supported by the cavity. This can be re-written as $2L\beta = m2\pi$ where m is a positive integer. In general, β must be numerically calculated for a given waveguide mode; however, for transverse electromagnetic modes, the approximation $\beta = 2\pi n_{\text{eff}}/\lambda_0$ can be used, where λ_0 is the free-space wavelength of the light and n_{eff} is the effective modal refractive index. Combining these, we find the condition on cavity length

$$2L = \frac{m\lambda_0}{n_{\text{eff}}} \tag{12.5}$$

It can be seen that for nanoscale lasers, as L approaches λ_0, the separation in wavelength between neighbouring modes which satisfy Eq. (12.5) increases. This has implications for producing homogeneous nanowire lasers, as it imposes tight requirements on nanowire length to align longitudinal modes with the peak of the gain spectrum. For many semiconductor materials used for nanowire lasers, the effective refractive index is relatively high: for instance, a group refractive index [17] $n_{\text{grp}} = 4.85$ and a phase index [64] $n_{\text{eff}} \approx 3.3$ have been calculated for a HE_{11} mode in a gallium arsenide (GaAs) nanowire with diameter of around 360 nm. This means that waves are effectively confined within the nanowire in vacuum

($n_{vac} = 1$) which acts as an optical scale waveguide. The modal refractive index varies with both nanowire diameter and transverse structure of the mode (and of the nanowire if radial heterostructuring is present), as shown in Fig. 12.2b. This confinement gives rise to two important features – the range of transverse modes supported by the waveguide and an enhancement in the field strength within the active gain material, which is described by Γ, the confinement factor.

Transverse modes describe the form of the distribution of electric field within a cavity in the plane perpendicular to the propagation direction. For realistic situations, such as hexagonal or faceted nanowires, these must be calculated numerically [24,64]. However, an analytical treatment for the simplified case of a cylindrical waveguide in vacuum is often enlightening and provides qualitatively similar distributions. A number of classes of transverse modes can be hosted by nanowire laser cavities: TE (electric field perpendicular to the direction of propagation), TM (magnetic field perpendicular to the direction of propagation) and HE modes (hybrid modes, with nonzero electric and magnetic fields in the direction of propagation). Cylindrically symmetric modes are enumerated with two integers – p, referring to the radial order and l referring to the angular order. Transverse modes can be written as a product of a Gaussian and a Laguerre polynomial [62,65], such that the transverse field E^T is given as

$$E^T_{l,p}(\rho, \phi) \propto \exp(il\phi)\exp\left(-\rho^2\right)\rho^l L^l_p(2\rho) \tag{12.6}$$

where $\rho = r/w$ is the normalised radial co-ordinate, r is the radial co-ordinate, w is the beam waist, ϕ is the angular co-ordinate and L^l_p is a Laguerre polynomial. In Eq. (12.6), the first term, represents the angular momentum of the state, the second term is the Gaussian-type radial extent and the latter terms provide the spatial structure. Fig. 12.2a shows the most important transverse modes for a GaAs nanowire laser; note that, in this more complete simulation, the symmetry of certain modes is broken by the presence of a SiO_2 substrate indicated by a horizontal line below the nanowire in the figure. In the case of broken symmetry, an additional term can be used to denote whether the mode has a greater field in the substrate or in vacuum.

For nanowire lasers made from a single homogeneous gain material in vacuum, the transverse mode has relatively limited impact on gain and hence the performance of the laser, other than determining the cutoff diameter below which propagation cannot be supported [16]. However in heterostructured devices or those with anisotropic surrounding media, the

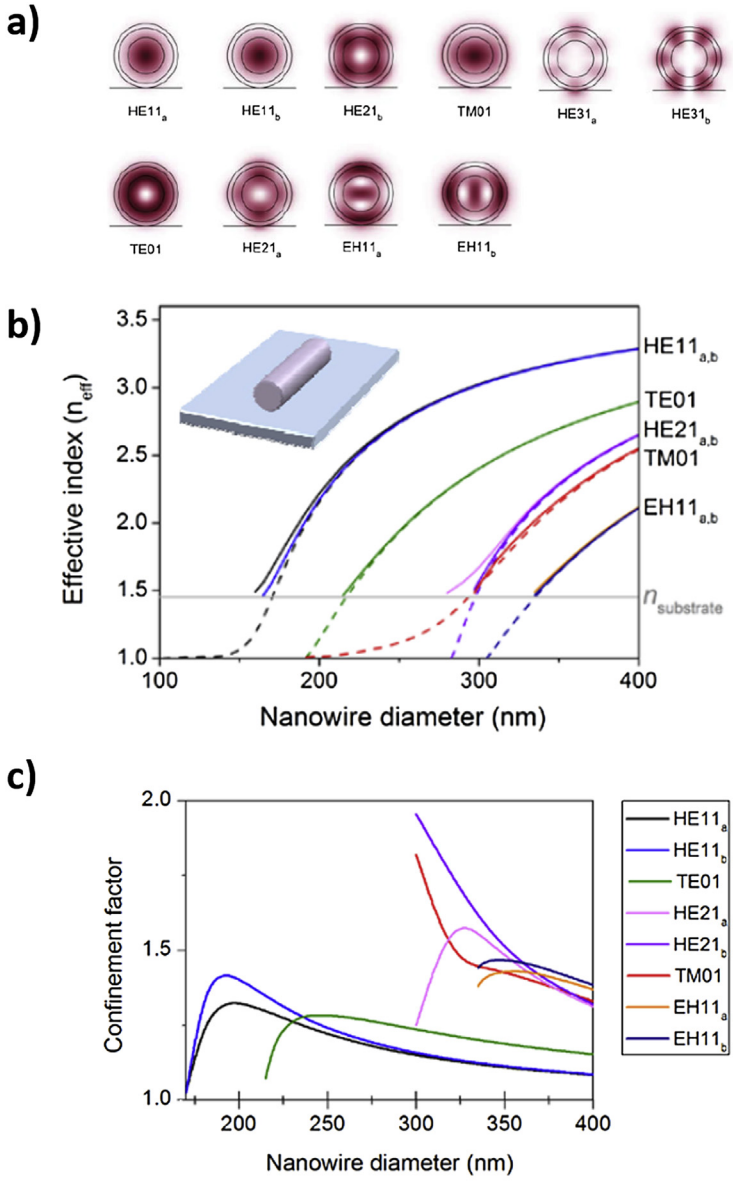


Figure 12.2 Calculations for the impact of transverse mode on reflectivity and confinement factor in a gallium arsenide nanowire. (a) Electric field intensities for a 300 nm diameter nanowire. (b) The effective refractive index (n_{eff}) of each mode as a function of nanowire diameter. (c) Dielectric mode confinement factor for eight optical modes as a function of diameter. *Adapted from Saxena et al. Nature Photonics 7, 963 (2013) Copyright © Springer Nature 2013 [64].*

transverse mode determines the confinement factor Γ which in turn can have a significant effect on the modal gain of the structure [66]. A full discussion of this factor is given in Ref. [67], but for a nanowire laser this can be given by the following: [68]

$$\Gamma = \frac{v_{ph}}{V_{E,z}} \frac{\int_{\text{active}} U_e dA}{\int_{\text{waveguide}} (U_e + U_m)/2dA} = \frac{v_{ph}}{V_{E,z}} \Gamma_0 \tag{12.7}$$

where v_{ph} is the phase velocity in the active medium, $V_{E,z}$ is the energy velocity in the waveguide, $U_{e,m}$ are the energy stored in the electric and magnetic fields, respectively, and Γ_0 is the conventional definition of the confinement factor (which is smaller than 1). The integrals are calculated over the transverse area (dA) of the active and waveguide regions. It is noted that for a nanowire, however, Γ can be greater than 1 if the energy velocity is lower (or group index is higher) than the phase velocity (or index) [68]. This provides a further opportunity to decrease the laser threshold in nanowire lasers when compared with their bulk counterparts, and typical confinement factors are shown in Fig. 12.2c.

Whilst the opportunities presented by nanoscale cavities are of interest, it is noted that specific design decisions in the fabrication of nanowire laser cavities must incorporate the realities of nanowire synthesis. These include imperfect control of nanowire length, diameter, end-facet quality and tapering as well as specific challenges arising from fabricating devices using nanowires.

2.1.3 Gain material

Gain in a semiconductor material requires the stimulated emission rate to exceed the absorption and spontaneous emission rates. The strength and hence rate of an interaction of light and matter in the weak perturbation limit can be approximated at the quantum mechanical level by Fermi's Golden Rule

$$k = \frac{2\pi}{\hbar} \left| \langle \Psi_f | H' | \Psi_i \rangle \right|^2 g(E) \tag{12.8}$$

where k is the transition rate, $\Psi_{i,f}$ are the initial and final states of the system (respectively), H' is the interaction Hamiltonian and $g(E)$ is the density of final states at energy E. For efficient light–matter interaction in materials with free charge carriers, we require a large matrix element $M = \left| \langle \Psi_f | H' | \Psi_i \rangle \right|$ and a large density of states at the transition energy; these guiding principles lead to the requirement for spatial overlap of electrons and holes for recombination, and quantum confinement of carriers as a mechanism for increasing $g(E)$ and providing efficient

band-edge recombination. The macroscopic rates of light–matter processes for a transition were described by Einstein, and for band-like systems, the rates between a conduction band state at E_c and a valence band state at E_v can be written as follows for spontaneous (k_{spon}), stimulated (k_{stim}) and absorption (k_{abs}) rates: [69]

$$k_{spon} = A_{cv}F_c(E_c)[1 - F_v(E_v)]\,g(E_c)g(E_v) \tag{12.9}$$

$$k_{stim} = B_{cv}F_c(E_c)[1 - F_v(E_v)]\,g(E_c)g(E_v)\rho(\Delta E) \tag{12.10}$$

$$k_{abs} = B_{vc}F_v(E_v)[1 - F_c(E_c)]\,g(E_c)g(E_v)\rho(\Delta E) \tag{12.11}$$

Here, $F_{c,v}$ are the Fermi distributions in the conduction band or valence band, respectively, and ρ is the photon density at an energy $\Delta E = (E_c - E_v)$. The pre-factors A and B are constants; in the simplified Bernard–Duraffourg analysis, these can be linked such that $A_{cv} \propto B_{cv}$ and $B_{vc} = B_{cv}$ [70]. In other words, under certain limited conditions, the absorption rate is equal to the stimulated rate, and the spontaneous rate is proportional to the stimulated rate, unless we have the freedom to set $F_c \neq F_v$, to have a different Fermi distribution describing the carrier densities in the conduction and valence bands. Specifically, population inversion for lasing can be achieved when the Fermi distributions for electron and holes have separate Fermi energies (described as quasi-Fermi energies) which lie within their respective bands leading to degenerate carrier densities, such that

$$E_g < E_\gamma < \left(\Delta E_f = E_{fc} - E_{fv}\right) \tag{12.12}$$

Here, E_g is the material bandgap, E_γ is the photon energy and $E_{fc,fv}$ are the quasi-Fermi energies for electrons in the conduction band and holes in the valence band, respectively. We may achieve this quasi-Fermi energy separation in nonequilibrium conditions, which can be achieved through injection pumping or optical excitation.

2.1.4 Density of states engineering

Achieving population inversion for lasing in bulk-like homojunction materials – those with a three-dimensional density of states – requires high excitation or injection conditions because more states are available. However, even when population inversion is achieved, the spontaneous emission rate can be low, as the density of states tends to zero at energies just above the bandgap. For a parabolic dispersion relationship, where the carrier energy can be described by the nonrelativistic equation $E \propto k^2$, the density of states as a function of dimensionality n is given by

$$g_n(E) \propto E^{(n-2)/2} \tag{12.13}$$

as shown schematically in Fig. 12.3. It can be seen that two-dimensional (quantum well), one-dimensional (quantum wire) and zero-dimensional (0D) (quantum dot) systems have a relatively large density of states close to the band edge when compared with bulk (three-dimensional) systems; such systems provide opportunities for reduced lasing threshold and improved temperature stability, once growth challenges can be overcome [9].

Three common approaches to achieve lasing from systems with reduced dimensionality are given here, using excitons, quantum wells and quantum dots.

Excitonic gain. Materials which have stable room-temperature excitons provide a simpler route to high recombination rates [71]. Excitons, the quasi-particles formed by the Coulomb interaction of an electron and a hole, act as 0D or *atomic-like* systems. Furthermore, excitonic lasers behave as quasi-four-level laser systems, with inter-band excitation, carrier relaxation and capture to form an exciton, and efficient monomolecular stimulated emission. The binding energy of an unconfined Wannier–Mott type exciton is given by:

$$E(n) = -\frac{\mu R_\infty}{m_0 \varepsilon_r^2 n^2} \tag{12.14}$$

where $\mu = (1/m_e + 1/m_h)^{-1}$ is the reduced mass of the exciton, m_0 is the bare mass of the electron, R_∞ is the Rydberg constant, ε_r the relative dielectric permittivity of the material and n the state of the exciton. For materials with high effective mass and low permittivity, it can be seen that

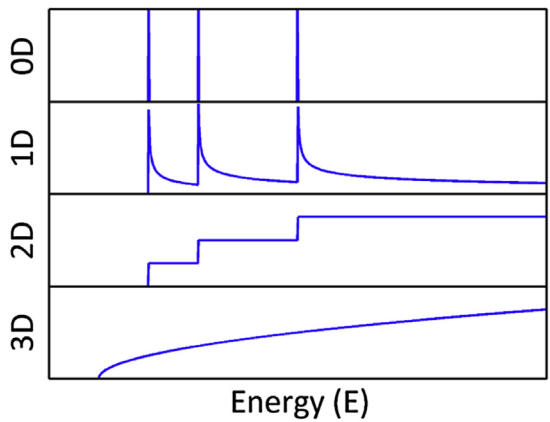


Figure 12.3 A schematic of the density of states as a function of energy for three-, two-, one- and zero-dimensional freedom, neglecting degeneracy.

the binding energy can be large; for ZnO, the binding energy of 60 meV is significantly larger than thermal energy ($kT \approx 25$ meV) at room temperature, leading to stable exciton formation and the potential for lasing [72]. Early studies of nanowire lasers made use of excitonic recombination as the gain mechanism in ZnO [73] and chalcogenides [74].

Quantum wells. The development of engineered quantum wells [75] enabled the first solid-state laser revolution [8,76]. By quantum confining carriers, the recombination energy can be tuned independently from the intrinsic bandgap of the material used. Quantum-well lasers are more efficient as they have a larger density of states close to the band edge (see Fig. 12.3), and are more thermally stable. Finally, excitons can be stabilised in quantum wells potentially providing the advantages described above (Section 2.1.4). In the infinite barrier approximations, the nth confined state energy for a carrier in a quantum well with width L is given by

$$E_n = \frac{n^2 h^2}{8 m_{e,h} L^2} \tag{12.15}$$

where $m_{e,h}$ is the carrier effective mass, leading to a lowest confined electron to hole transition (the $n_e = 1$ to $n_h = 1$ transition) energy of

$$\Delta E = E_g + \frac{h^2}{8L^2}\left(\frac{1}{m_h} + \frac{1}{m_e}\right) \tag{12.16}$$

Nanowire lasers with a quantum-well gain region have been fabricated in III-nitrides [77] and III-V materials [78] to achieve tunable lasing wavelength with low threshold, which are discussed in Section 3.1.

Quantum dots. By confining carriers to a region with dimensions smaller than the de Broglie wavelength of the carriers, 0D structures can be formed. These artificial atoms provide a sharp and narrow density of states for high efficiency and low thermal drift lasing [9]. In planar optoelectronics, the formation of quantum dots often relies on self-assembled processes such as Stransky–Krastanov growth [79] to produce these confined structures. In semiconductor nanowires, a new opportunity is provided; if axial quantum wells are produced in nanowires with a sufficiently small radius, arrays of quantum dots can be formed with deterministic control [34,53] or through thermodynamically driven processes [80].

2.2 Nanoscale laser cavity design

To form an effective Fabry–Pérot cavity, we must consider the reflectivity of the end-facets. The naïve Fresnel equation for the power reflected

perpendicular to a planar mirror with reflectivity R in a wave traveling from medium 1 with refractive index n_1 to medium 2 with n_2

$$R = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|^2 \tag{12.17}$$

does not hold where the cavity size approaches that of the wavelength of light. A fuller analysis [24,81] shows that the reflectivity depends strongly on the transverse mode type, wavelength and radius of the cavity. Early growth approaches had little control over nanowire diameter; in their first report, Huang and colleagues grew ZnO nanowires with diameters ranging from 20 to 150 nm [2], and reports with zinc sulfide nanoribbons fared little better [82] with thicknesses of 40–200 nm.

The adoption of seeded growth developed for silicon [83,84] and III-V materials [85] provides a route to diameter selective synthesis of optoelectronic nanowires. By appropriate tuning of the colloidal preparation and deposition, narrow diameter distributions (with a dispersion of less than 5% of the wire width) are achievable [86]. Well-controlled nanowire diameters provide one of the key requirements for controllable reflectivity; in addition, the diameter of the nanowire impacts on the supported transverse modes of the laser. Whilst nanowires with deep sub-wavelength dimensions can support the HE_{11} mode, those with larger dimensions can lead to the emergence of higher order modes (i.e., TM_{01} or TE_{01}) which change the far-field emission [87–89], the reflectivity [24] and the electric field distribution within the wire.

Control over the electric field distribution within a nanowire provides an opportunity to create efficient lasers based on single [90] or multiple radial quantum wells [78,91]. In these structures, the gain lies in the region hosting the quantum wells, and higher-order donut-type modes such as the TE_{01} mode can provide better spatial overlap between the gain and cavity mode [78]. These higher order modes can exhibit less leakage of the field into the nanowire core or the surrounding media, although it is noted that the field in the core remains nonzero and re-absorption can occur [92]. For high-efficiency lasing, this confinement is beneficial; however, the ease of access to evanescent waves at the surface of nanowire lasers also provides an opportunity to modify the emission by coupling to plasmons [93–95], for optical pumping [96], for coupling of light into waveguides [26,30] or for more complex coupling schemes [97,98].

Most nanowire laser systems use the reflectivity from the end-facets of the nanowire as Fabry–Pérot mirrors to provide feedback. Whilst the true

reflectivity in these systems can be high – for instance, reflectivity of over 70%, around double the naïve Fresnel value has been measured [99] – integration of external cavity elements has been studied by use of distributed Bragg reflectors. Here, the nanowire lasers can be grown directly onto a high reflector with over 90% power reflectivity [29], or post-growth fabrication steps can be used to impose a periodic structure onto a wire using focussed ion-beam milling [100] for reflectivities over 90%.

2.2.1 Array lasers

Whilst many applications call for single nanowire lasers, arrays of nanowire lasers provide an opportunity to create feedback through photonic coupling of the cavities without exploiting Fabry–Pérot structures. By spreading the photonic field over many wires, increased gain can be produced at the cost of a larger device volume. Two approaches have been used: random lasing based on localisation of light in a scattering medium [101,102] and designed arrays of nanowires for integration [48,103]. In the latter, a one-dimensional photonic crystal is typically used to direct the light energy into an adjacent waveguide for on-chip integration.

2.3 Laser threshold

The threshold of a laser is the point at which stimulated emission becomes significant in output. Whilst parameters such as material gain and the threshold for transparency are well defined (see Section 2.1.3), the laser threshold relies on measurements of light power emitted by a device as a function of power provided either via optical pumping or electrical injection. Conventionally, three regions are considered for a cavity in the weak-coupling regime:

Spontaneous emission. When spontaneous emission dominates, the laser acts as simple gain material with emission that couples weakly to cavity modes (Fig. 12.4). Emission is uncorrelated, and the spectrum is primarily determined by that of the material rather than the cavity.

Amplified spontaneous emission. In the intermediate pumping regime, stimulated amplification of spontaneous emission occurs, but below the level required to compensate for round-trip losses. Emission becomes supercorrelated, or bunched, with an emission spectrum which shows evidence for low-Q-factor modes without coherent emission.

Lasing. Above threshold, stimulated emission dominates and lasing occurs, with round-trip losses compensated by the gain of the material.

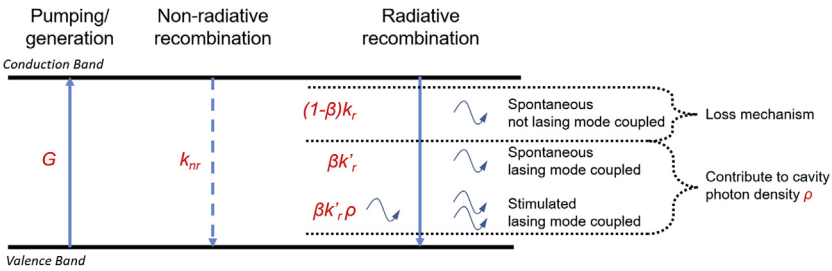


Figure 12.4 Schematic description of the rate equation model for determining laser threshold in a simple two-band system. Transitions are shown in *blue* (*grey* in print version), with radiative and nonradiative as solid and dashed lines respectively. Transition rates are shown in *red* (light *grey* in print version).

The emission is uncorrelated, with an emission spectrum showing one or more longitudinal or transverse modes which are coherent.

The transition between these regimes can be experimentally challenging to determine, particularly in pulsed systems where emission from all regimes can occur within one excitation cycle. A simplified model shown in Fig. 12.5 depicts the transition from spontaneous emission dominated to lasing dominated behaviour.

Eq. (12.10) showed that laser threshold is intrinsically linked to the density of excited states n and the intra-cavity photon density ρ. A simplified

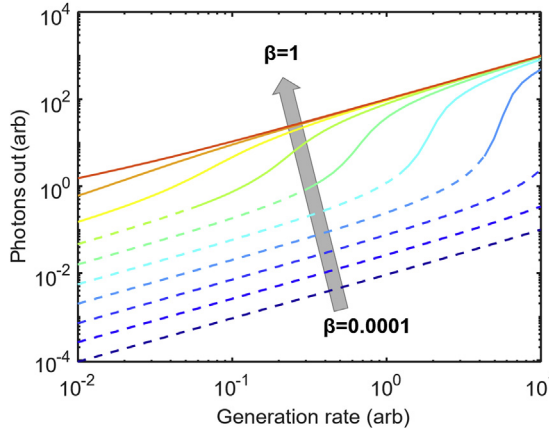


Figure 12.5 Schematic of photons out as a function of generation rate (or equivalently pumping) for a laser for different β factors. Dashed lines represent regions where spontaneous emission dominates, whereas solid lines are regions where stimulated emission dominates.

coupled rate equation analysis is often used to model the light-out versus light-in data [104,105] as depicted schematically in Fig. 12.4. The carrier density is determined by a generation and three recombination terms – radiative nonlaser mode coupled, radiative laser mode coupled and nonradiative. The photon density comprises a generation rate from radiative laser mode coupled emission and a loss mechanism:

$$\frac{dn}{dt} = G - (1-\beta)k_r n - \beta k_r'(1+\rho)n - k_{nr}n \tag{12.18}$$

$$\frac{d\rho}{dt} = \beta k_r'(1+\rho)n - k_\rho \rho \tag{12.19}$$

Here, G is the generation rate, k_r is the intrinsic spontaneous emission rate, β is the coupling efficiency of spontaneous emission to the cavity mode, k_{nr} is the nonradiative carrier depopulation rate and k_ρ is the damping rate for photons (including photon outcoupling). Solutions to these coupled equations for a range of β factors are shown in Fig. 12.5. The factor k_r' is the cavity-enhanced spontaneous emission rate, which is related to the free space rate by $k_r' = Fk_r$; here, F is the enhancement factor, arising from Purcell-type enhancement in high quality (strongly coupled) cavities. For low Q cavities, including most single nanowire lasers, $F \approx 1$. The factor β is of particular importance for nanoscale lasers. Wavelength-scale laser cavities have longitudinal and transverse modes which are widely spaced in energy and have strong optical confinement; as such, photons emitted are more likely to couple to the lasing mode, providing a route to near-unity β such that the second term in Eq. (12.18) tends to zero. In this regime, stimulated emission will dominate over nonradiative processes and Eq. (12.18) can be simplified to contain only terms in n_ρ, producing a so-called *thresholdless laser* [106]. Thresholdless lasers can be extremely efficient and have high switching speed; engineering high- β nanolasers requires high control over the geometry of the cavity, and remains an active area of study for nanowire lasers [107].

2.3.1 Figures of merit

The metrics or figures of merit used to define a "good" nanowire laser are dependent upon the application. However, a number of important metrics do exist to compare nanowire lasers. These are as follows:

- **Threshold pump density:** Either the current density (Am^{-2}) for electrical injection, optical pump density (kWm^{-2}) for continuously

operating optically pumped devices or pulse fluence (Jm^{-2}) for pulsed optically pumped devices required to produce lasing emission

- **Laser centre wavelength:** The first or strongest laser mode supported by the nanowire laser
- **Laser linewidth:** The width of the primary laser emission, which is related to the Q-factor of the cavity
- **β factor:** The strength of coupling of spontaneous emission to the lasing mode
- **Single/multimode operation:** The number of laser modes present above threshold. Often single mode operation is preferred
- **Maximum operating temperature:** Room temperature lasing is essential for many applications. However, elevated temperature operation is sought for real-world use [108].
- **Operational lifetime:** Continuous or pulsed operation can lead to degradation of the nanowire cavity or gain material, particularly for air-sensitive materials such as hybrid perovskites [12].
- **Wire-to-wire disorder:** For integration of multiple nanowire lasers on a chip, tight control over performance homogeneity is required [109].
- **Device volume:** For the ultimate in scaling or for near-field applications, extremely small device volumes are sought.

Within each nanowire laser architecture and for each application, different metrics are often optimised. For instance, for telecoms band operation, the centre wavelength is critical [53]; for perovskite nanolasers, stability is sought [12] and for on-chip applications, wavelength uniformity might be demanded [40,92]. However, room-temperature operation at low threshold remains the key metrics for nanowire laser development [64,110].

Some of the key developments in the history of nanowire laser development are listed in Table 12.1. Methods to optimise nanowire lasers based on design and growth are described in Section 3.

3. Nanowires as laser elements

The design of nanowire lasers builds on the two key features of a laser system: the choice of gain material, and the design of the optical cavity. The advantages of the nanowire platform – where the gain material forms the cavity – lead to a challenge in fabrication and control. In this section, a brief background to semiconductor nanowire growth is provided, and the specific design approaches developed for nanowire lasing are described.

Table 12.1 A table of key reports in the development of pulsed nanowire laser systems. Only, reports where laser thresholds could be converted to μJcm^{-2} are included.

Year	Material	λ (nm)	P_{th} (μ J cm^{-2})	Temp. (K)	Note	Ref.
10/2001	ZnO	380	43	RT	First single-wire lasers	[1]
5/2002	ZnO	675	25,641	RT	Solution-phase nanowire lasers	[111]
9/2002	GaN	375	8	RT	First nitride nanowire laser	[112]
3/2005	CdS	500	4	250	Excitonic lasing	[74]
4/2006	GaSb	1555	50	30	Near-infrared lasing	[46]
2/2007	Organic	448	2829	RT	First polymeric nanowire lasers	[113]
9/2008	GaN	380–480	7074	RT	First multiple-quantum-well nanolaser	[77]
4/2008	Organic	375	16,977	RT	Ultraviolet organic laser	[114]
10/2008	GaAs/GaAsP	815	270	75	Core-shell nanowire laser	[115]
2/2009	Organic	585	60	RT	Electrospun lasers	[116]
8/2009	CdS	489	4	10	Plasmonic laser	[117]
2/2011	CdSe	750	40	RT	Loop-mirror demonstration	[118]
8/2012	CdS	515	85	RT	Quasi-continuous lasing	[119]
1/2013	GaN	372	400	RT	Cleave-coupled lasers	[98]
10/2013	CdSe	725	60	RT	Room temperature plasmonic laser	[11]
10/2013	CdS	490	8	RT	Surface-plasmon polariton	[120]
11/2013	GaAs	840	90	RT	Room-temperature lasing	[64]
12/2013	GaAs	911	146	RT	Room-temperature lasing	[110]
8/2014	InP	900	120	RT	Wurtzite InP laser	[121]

(*Continued*)

Table 12.1 A table of key reports in the development of pulsed nanowire laser systems. Only, reports where laser thresholds could be converted to μJcm^{-2} are included.—cont'd

Year	Material	λ (nm)	P_{th} (μ J cm^{-2})	Temp. (K)	Note	Ref.
9/2014	GaN	370	35000	RT	Plasmonic laser	[122]
11/2014	ZnO	378	200	RT	Ultrafast plasmonic lasing	[123]
6/2015	GaAs	875	180	RT	First quantum dot nanowire laser	[124]
1/2016	GaAs	792	46	19	Multiple quantum-well arsenide laser	[91]
1/2016	Perovskite	560–820	6	RT	Formamidinium perovskite laser	[125]
2/2016	Perovskite	530	5	RT	Robust CsPbBr perovskite laser	[12]
6/2016	GaAs	880	165	RT	Doping-enhanced nanowire laser	[126]
7/2016	GaAs	800	100	RT	Multiple quantum-well arsenide lasers	[78]
1/2017	InGaAs	860	286	7	Nanowire integrated with a DBR	[29]
1/2017	GaAs	875	145	RT	Wire in plasmonic cavity	[127]
1/2017	InAsP/InP	1259	63	4	Photonic crystal	[50]
2/2017	GaAs	870	15	5	Dilute nitride laser	[128]
3/2017	InP	1210	450	4	Site-controlled growth	[31]
4/2017	Organic	650	10	RT	Phosphorescent laser	[129]
4/2017	CdS	510	34	RT	High temperature operation	[108]
5/2017	InGaAs	1120	100	RT	Telecoms wavelength	[103]

7/2017	InGaAs	1100–1500	100	RT	Telecoms wavelength	[47]
7/2017	GaAs	750	43	RT	Yield analysis	[92]
10/2017	Organic	720	1	RT	Near-infrared lasing	[130]
4/2018	GaAsSb	950	75	RT	Axial superlattices	[131]
5/2018	GaAs	870	1400	10	Wires on flexible membranes	[34]
5/2018	Perovskite	521	16	RT	CsPbX - dual colour lasing	[132]
9/2018	GaAs/InGaAs	905	3600	RT	Tunable lasing	[42]
10/2018	Perovskite	530	30	RT	Plasmonic laser	[133]
11/2018	InGaAs	1290	155	RT	Array lasers	[48]
12/2018	GaAs	855	10	RT	Optimised doping	[134]
1/2019	GaAs/GaNAs	1000	20	5	Multiple quantum-well	[52]
2/2019	InP/InAs	1570	2000	RT	Telecoms band	[23]
5/2019	GaAsP	745	20	6	Strain engineering	[23]
10/2019	InAs	2500	90	4	Mid-infrared lasing	[135]
3/2020	GaAsP	775	6	RT	Q-factor measurement	[99]
3/2020	ZnO	380	1	RT	Photonic crystal	[136]

3.1 Nanowire growth

The growth of semiconductor nanowires using a bottom-up or *self-directed* mechanism has been explored for over 50 years since the development of vapour–liquid–solid growth (VLS) in the 1960s [83]. Bottom-up growth is achieved by the use of either a seed material or a template, with the goal of directing the growth of a semiconductor crystal. In either approach, precursor materials are introduced to a chamber and growth conditions are tuned such that crystal growth preferentially occurs at defined sites and along preferred orientations. Nanowires have been grown using metal-organic vapour phase epitaxy (MOVPE, also known as MOCVD) [137], chemical or molecular beam epitaxy (CBE/MBE) [138], solution phase growth [139], laser ablation [84], physical evaporation [140] and other emerging methods [25]. Modern growth for scaled applications has predominantly coalesced around two approaches – MOVPE and MBE – with growth directed by either VLS [141,142] or selective area growth (SAG) [143–146]. For a full description of nanowire growth, the interested reader is directed to review articles for silicon [147] and III-V nanowires [148], as well as general history [149].

A simplified schematic of the two primary growth modes for nanowires is shown in Fig. 12.6. Both methods start with a prepared epitaxial substrate – this can be the same material, as the nanowires for homoepitaxial growth, or a different material (for instance, III-V nanowire growth on silicon [150]). The small diameter of nanowires allows for efficient strain relaxation, enabling the defect-free growth of materials with significant lattice mismatch [151,152] allowing for a wide range of unconventional heteroepitaxial systems to be prepared. Following this, the mechanisms diverge.

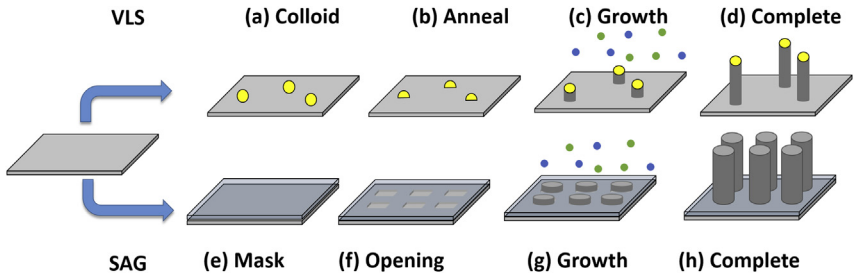


Figure 12.6 A schematic of two growth modes for epitaxial nanowire synthesis. In VLS growth (a–d), seed nanoparticles are deposited from solution or via evaporation, which are annealed, before precursor gases are introduced and growth is initiated. In SAG (e–h), a hard mask is deposited before holes are lithographically defined and opened, which act as a template for nanowire growth.

3.1.1 Vapour-liquid-solid growth

A catalyst seed is used in VLS growth. This can be a specific metal – often gold, as this is available as highly monodisperse colloids – or can be one of the precursors for self-catalysed growth [153,154], which can avoid foreign metal contamination. The metallic nanoparticle is heated under a precursor flux, saturating the metal and forming a eutactic. By optimising the growth conditions, crystal formation can be made favourable at the interface between the nanoparticle and the substrate leading to selective growth under the catalyst. The diameter of the resultant nanowire is controlled by the diameter of the growth catalyst and the thermodynamic balance between axial and radial growth (known as tapering). Precursors are fed into the growth vessel until an appropriate length is achieved. By varying the composition of the precursors, axial heterostructures can be produced, such as quantum discs or dots [124]. Alternatively, by varying the growth conditions, surface growth (radial growth) can be performed to produce radial heterostructures such as surface passivation [155] or radial quantum wells [77]. Control of growth conditions is essential to achieve defect-free growth [156] with low surface recombination [157]. Whilst VLS growth using colloids typically leads to randomly positioned nanowires, nanoimprint lithography has been explored for site position of growth [158].

3.1.2 Selective area growth

In SAG, a hard template is employed to limit the crystal surface available for epitaxy. Whilst different mask materials have been studied [159], a 20–50 nm thick SiO_2 or Si_3N_4 mask is typically used. Holes are defined on the surface using electron beam lithography and opened using a buffered etch. Material re-growth is performed using MOVPE or MBE; GaAs [143] and nitrides [160] were first studied, which has since expanded to include telecom wavelength materials [47]. Similar to VLS growth, by changing the growth conditions, both radial [115] and axial [124] structures can be produced. Whilst SAG is more complex and has a lower throughput than VLS growth, the results can be more uniform and allow for alignment with pre-existing substrate features.

More recently, a substrate-free method of continuous growth of nanowires has been reported using *aerotaxy* [25]. In this approach, the growth is performed using catalysts falling freely through precursor gases, which can produce both binary and doped [161] nanowires.

3.2 Material systems for nanowire lasers

A wide range of materials systems have been explored for the fabrication of semiconductor nanowire lasers. The choice of materials is determined by a number of factors: the ease of fabrication of bottom-up nanowires, the bandgap of the material, integration with existing electronic or photonic platforms and availability of approaches for heterostructuring or electrical injection. Early studies focussed on materials for which nanowires were readily available (such as ZnO and gallium nitride), whilst subsequent development moved to materials with high gain and ease of processibility (chalcogenides and perovskites) or with a large technology ecosystem (III–V materials). Key features of these materials systems are described here.

3.2.1 Zinc oxide

Many of the earliest studies on nanowire lasers were performed using ZnO nanowires. This was driven both by a desire to produce short-wavelength ultraviolet lasing – ZnO has a bandgap of 3.4 eV, corresponding to a wavelength of 365 nm – and by the existence of stable excitons at room temperature – with an exciton binding energy of around 60 meV – leading to efficient stimulated emission [162]. In 2001, the Yang group at Berkeley reported the first room-temperature ultraviolet lasing in self-organised ZnO nanowire arrays [73] and single nanowires [1]. It was noted that the laser threshold, at around 120 kW cm^{-2}, was three times higher for single nanowires than for arrays: this indication that evanescent fields play an important role in nanowire lasing led to later developments in coupled structures [163] and optical-pumping techniques [96] (see Section 4.2).

The production of ZnO nanowires via solution-phase or *hydrothermal* growth [164] provided opportunities for low-cost and wafer-scale fabrication of nanowires [139], which in turn have shown room-temperature lasing [111]. However, conventional high-quality growth approaches such as pulsed laser deposition continued to dominate for high efficiency nanolasers, with internal and external quantum efficiencies of 85% and 60% being demonstrated respectively, both significantly above the spontaneous emission efficiencies of around 10% [165].

ZnO has proven to be a reliable material for studying new effects, particularly for plasmonic nanowire lasers (see Section 4.3). Plasmonic lasers, which use optical confinement originating from surface plasmon polaritons to reduce the spatial scale of the electronic interaction, have been promoted as a route to ultra-high speed laser dynamics. With their high gain, ZnO

plasmonic lasers were first demonstrated in 2014 [123], and plasmonic operation at elevated temperatures was shown by Chou in 2016 [166].

Despite its strengths as a gain material, the lack of facile p-type doping [167] and the challenges in heterostructure engineering with crude hydrothermal growth have led to the sidelining of ZnO as a practical lasing material for device applications.

3.2.2 Nitrides

The development of III-nitride materials – (B,In,Ga,Al)N – has revolutionised semiconductor applications throughout science and technology, most notably leading to the 2014 Nobel Prize in Physics for Akasaki, Amano, and Nakamura "*for the invention of efficient blue light-emitting diodes*" [168]. The enabling technology behind this development was the use of high-quality growth techniques such as MOVPE to produce heterostructured materials which allowed for strain relaxed growth onto sapphire [169] and ultimately high-brightness LEDs [170] by the mid-1990s.

Like ZnO, GaN is a wide bandgap semiconductor with a band-to-band energy gap of 3.45 eV (360 nm) at room temperature, and a large exciton binding energy of 20 meV (close to kT at room temperature) [171]. Unlike ZnO, heterostructure growth and doping are relatively straightforward to achieve using conventional growth approaches, and ternary alloys (including AlGaN and InGaN) are readily available for bandgaps spanning the UV to visible spectrum [172,173]. As such, GaN-based nanowire lasers were a key target for early development, with room-temperature lasing observed in 2002 [112]. Whilst the growth of planar nitride layers often requires a buffer to achieve lattice matching to common substrates (such as sapphire), the growth of nanowires allows for easy strain relief [151] leading to facile growth along semi-polar directions, which in turn provide high-quality cavities for low threshold lasing [174].

The exciton binding energy in GaN is below kT at room temperature, and as such recombination occurs through electron-hole processes under lasing conditions. A low surface recombination velocity in GaN nanowires [175] means that surface passivation is generally not required. However, by exploiting an understanding of heterostructure growth from planar structures, quantum wells provide an opportunity both to tune the emission wavelength and provide confinement to increase the exciton binding energy and produce more efficient nanowire lasers. The first quantum-well nanowire lasers were produced in InGaN/GaN systems [77], providing a breakthrough for the use of radial heterostructures for lasing which

continues to this day [99]. By exploiting methods to grow along nonpolar directions, reduced thresholds have been demonstrated [176]. Conversely, the nanowire platform provides an opportunity to grow challenging bulk-like materials – such as AlN – opening the possibility for deep ultraviolet emission with potential for array-type lasing [173].

As a relatively well-understood material with low lasing thresholds and an etch-resistant structure, nitride-based nanowires have been used to explore novel fabrication approaches, including cleave coupling [98] and integrated Bragg reflectors [100] (see Section 4). Notably, novel fabrication approaches to make large area arrays have been proposed, including interferometric lithography [177], which take advantage of the understanding developed in the planar nitride materials community.

3.2.3 Chalcogenides

Chalcogenide containing compounds such as cadmium sulfide (CdS) or cadmium selenide (CdSe) have long been studied as optoelectronic materials, due to their high optical gain arising from stable excitons at room temperature [178]. It is therefore unsurprising that structures based on CdS were used for the first demonstration of electrically injected nanowire lasing [35] and for continuous wave lasing [179]. After photo-excitation, excitons are formed [178] which are able to rapidly cool through exciton–exciton [180] and exciton–phonon scattering, leading to a sharp and narrow gain spectrum with emission in the blue-green region [74].

Ternary alloys of CdS_xSe_{1-x} have been explored due to the ability to tune emission from 1.8 to 2.5 eV (490–690 nm) by controlling alloy fraction (x); this ease of tuning has led to the demonstration of lasing across the visible spectrum [181]. The ability to manipulate the local bandgap of the nanowire lasers has led to the design of axial heterostructures of CdS and CdSe; in nanolasers, symmetric graded bandgap structures can provide a passive higher bandgap cavity [182], whilst abrupt axial heterostructures have been demonstrated as dual-wavelength lasers [183].

Semiconductor nanowires are known to be mechanically flexible; silicon nanowires have an increasing strain to failure with decreasing diameter [184], whilst for GaAs nanowires, a failure strain up to 10% has been measured [185,186]. In particular, chalcogenide nanowires are exceptionally flexible and can support strain gradients of 1% per micron length [187]. This provides an opportunity to create novel cavity architectures through shape engineering, including single-mode lasing by replacing one cavity end-mirror with a loop mirror [118]. This flexibility also makes

chalcogenide nanowires attractive for incorporation into waveguide circuitry [43]; however, it has been noted that highly strained CdSe nanowires have an increased threshold and decreased efficiency when compared with unstrained wires [188].

3.2.4 III-V

Since the earliest development of optoelectronics, semiconductors based on compound group III-V materials have dominated the field [4,5,7]. Light–matter interaction is well understood in these materials, and the availability of binary substrates and advanced deposition techniques for ternary and quaternary materials with near monolayer control provides an ecosystem of techniques and knowledge. The first demonstration of near-infrared nanowire lasing, based on gallium antimonide (GaSb) wires at low temperatures and under pulsed excitation, was reported in 2006 [46], with lasing occurring at the telecommunications band ($\sim$1550 nm).

Whilst telecommunications band lasing is of great interest, the GaSb material system is relatively uncommon; lasing based on the well-understood (In,Ga)–(As,P) nanowires was sought. However, GaAs has a large density of surface defects, leading to a very short minority carrier lifetime in the surface recombination-dominated nanowire geometry [189,190]. The development of effective surface passivation using type-I radial heterostructure [155,157,191] for GaAs nanowires enabled the demonstration of near-infrared lasing, initially under pulsed excitation at low temperatures [115] and later at room temperature [64,110] and continuously at low temperature [107]. Nanowires based on indium phosphide (InP) are known to have lower surface recombination [190], but suffer from greater challenges in controlling the crystal phase during growth (as well as opportunities for crystal phase engineering [192]). Control of the growth mechanism was demonstrated by using SAG (templated), which enabled room-temperature lasing [121].

Lasing based on recombination in bulk materials is less efficient than recombination in quantum confined structures (see Section 2.1.4). Heterostructure growth in the III-V system is well developed; as such, nanowire lasers based on axial quantum dots [29,53,124] and radial quantum wells [23,31,78,90,91] have been shown to have tunable lasing emission and low lasing thresholds; however, room temperature continuously operating nanowire lasers have still not been demonstrated. Alternative approaches for optimising and tuning nanowire emission include doping – either electronic [126,134] or with dilute nitrides [52,193] – or exploiting strain [194,195].

A key strength of the III-V growth approach is the relative ease of incorporation onto silicon [158] for applications in on-chip or chip-to-chip communication (see Section 1) [17,31,47,51,103]. Much contemporary research is in the integration of nanowire lasers with on-chip waveguides [48,196] or in plasmonic systems [127,197]; whilst nanowire lasers based on chalcogenides and perovskites (Sections 2.3.2.3 and 3.2.5) have shown lower lasing thresholds; III-V nanowire lasers (and technology based on nanowire growth [198]) remain industrially relevant.

3.2.5 Perovskites

Since their demonstration as a sensitising layer for dye-sensitised solar cells [199,200], research on hybrid perovskites materials for optoelectronics has grown exponentially due to their tunability, high internal quantum efficiency and ease of fabrication [201]. Perovskite materials span a wide range of properties due to the tunability conferred by the palette of cation and anion materials that can be chosen. The growth of nanowires using the prototypical material, methylammonium lead iodide ($CH_3NH_3Pbl_3$ or MAPI), was demonstrated using solution-mediated crystallisation in 2014 [202]; in 2015, single nanowire lasing was demonstrated in MAPI [203], followed by other halide and mixed halide materials [204].

As had previously occurred in the photovoltaic field, these results represented a significant advancement in nanolaser research. The first demonstrated nanowire lasers exhibited thresholds as low as 200 $nJcm^{-2}$ and lasing quantum yields approaching 100% [203], spurring widespread study. Whilst MAPI perovskites are known to suffer from environmental-induced degradation under high photoexcitation densities, replacing methylammonium with formamidinium ($CH(NH_2)_2^+$) provided a tenfold improvement in stability [125]. All-inorganic $CsPbBr_3$ nanowire lasers remained stable for at least 10^9 pulses [12] as well as providing for tunability through choice of halide ion [205].

Whilst the choice of anion is conventionally made at growth time, the large surface area to volume ratio of nanowires provides an opportunity to change the anion (and hence emission wavelength) of nanowire lasers at the fabrication stage through gas-phase anion exchange [206]. The control of positioning and growth direction for perovskite nanowires for lasing applications has parallels with more traditional materials; template-guided growth [207] can provide lasing arrays, whilst orientation with crystallographic facets can provide large-area arrays [208] of low-threshold nanolasers.

The photo-physical origin of lasing action in perovskite materials has been intensively studied. Whilst excitons are stable at low excitation densities, when the exciton density passes the Mott density, screening leads to emergence of an electron-hole plasma [209]. Development of $CsPbBr_3$ perovskite nanowires initially attributed the gain to polariton modes [210] with continuously operating lasers demonstrated with this material [211]. However, time-resolved measurements on the same nanowire systems indicated that the stimulated emission is provided by recombination in an electron-hole plasma present at early times after excitation [212].

Nano- and micro-wire laser research based on perovskites continues. The development of new materials – such as graded nanowires with tunable emission [132] – and applications – such as 100 GHz switching speed via mode switching [213] – is likely to be an emerging area for future nanoscale coherent light emission [19].

3.2.6 Organic materials

The definition of a nanowire laser – one where the gain material and the cavity is provided by a single solid state structure – allows for a wide variety of material systems. Whilst the primary materials in contemporary research are described in Sections 2.3.2.1–3.2.5, lasing has been observed in a much wider range of materials. Based on developments in organic laser systems [214,215], organic nanowire lasers have been investigated since the first reports of optical pumped lasing in 2007 [113]. Organic solid state lasers make use of a quasi-four-level lasing scheme based on excitonic levels $S_{n,\nu}$ where n and ν enumerate the electronic and vibration states, respectively. In a typical scheme, absorption from a ground electronic and vibrational state ($S_{0,0}$) occurs to an excited electronic and vibrational state ($S_{1,\geq 0}$), followed by rapid relaxation to the ground vibrational state of the excited manifold ($S_{1,0}$), and the lasing transition to a higher lying vibrational state of the ground electronic state ($S_{0,\geq 0}$). This process provides for a low lasing threshold, but organic materials suffer from intersystem crossing from singlet (emissive) excitons to triplet excitons at high exciton density.

In the first study, Fabry–Pérot type cavities were formed using a porous alumina template producing wires with a range of diameters; unlike crystalline materials described above, nanowires of organics made in this way have geometries governed by their growth template, though with a polymer chain preferentially oriented along the long dimension. A very low lasing threshold of around 100 $nJcm^{-2}$ was found for these structures, with lasing at around 450 nm. Nanowires produced using small molecules

of 2,4,5-triphenylimidazole in adsorbent-assisted physical vapour deposition were observed to support lasing towards the ultraviolet, at 375 nm [114]. Whilst these structures had even lower threshold, the challenges of fabrication and poor environmental stability of organic materials have limited their application.

Despite these challenges, improvements in production reproducibility through improved fabrication such as electrospun fibres [116], in designer small molecules for near-infrared lasing [130], and the development of phosphorescence based lasing [129] have followed. This latter study used a novel compound with efficient intersystem crossing and a high phosphorescence quantum efficiency to lase on triplet transitions.

4. Contemporary topics in nanowire laser technology

In the development of single or array nanowire laser technology, a number of key design techniques have led to significant advances. Amongst these, quantum confinement within nanowires – such as radial quantum wells or axial quantum dots – has led to low threshold and tunable lasing, waveguide coupling has advanced prospects for on-chip integration and exploitation of plasmonic effects have enabled extremely small and fast coherent light sources. These contemporary areas of research are discussed in this section.

4.1 Quantum confinement

Fabry–Pérot type semiconductor nanowires for lasing applications have diameters on the scale of 100–1000 nm, designed to support one (or a few) transverse mode at the lasing wavelength. Whilst small, this is significantly larger than the de Broglie wavelength of carriers at room temperature in common laser materials. As such, early nanowire lasers tended to make use of excitonic transitions for gain [73] or required high excitation fluences [64,110]. It is well known than population inversion and hence gain is easier to achieve in quantum-confined systems (see Section 2.1.4). To achieve quantum confinement of carriers in the nanowire, heterostructuring is required. Fig. 12.7 shows two typical heterostructure approaches used for nanowire lasers.

4.1.1 Radial quantum-well nanowire lasers

Radial quantum-well nanowire laser structures were first demonstrated by Qian in 2008 [77]. To achieve sufficient gain, multiple quantum wells are

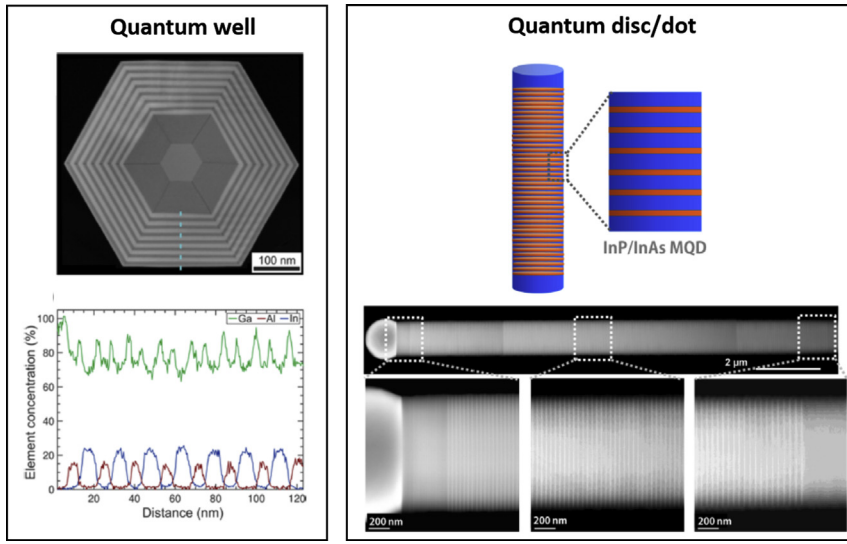


Figure 12.7 The characterisation of (left) a multiple radial quantum-well structure, showing electron microscopy and elemental mapping for an (In,Al)GaAs structure. (right) An axial multiple quantum-dot structural design with TEM imaging. *Left: Adapted with permission from Nano Lett. 18, 6292. (Copyright 2018 American Chemical Society) [42]; Right: adapted under CC-BY license from Science Advances 5, eaat8896 (2019). Copyright ©2019 The Authors [53].*

typically required. Not only do quantum wells provide an opportunity to reduce threshold, they also provide for independent tunability through well composition or thickness. At low temperature, quantum wells stabilise excitons (see Section 2.1.4) providing spectrally narrow and strong gain [91]. As radial quantum wells are offset from the nanowire core, the best matched transverse mode is typically a higher order TE01 or "doughnut" mode [78]. In the well-explored GaAs-AlGaAs material system, this provides an opportunity to confine the laser mode away from the core, to avoid re-absorption losses [92]. More recent approaches have sought to replace the potentially re-absorbing core with a higher bandgap material such as GaAsP [99,195], or to move to lower bandgap ternary compounds for the quantum well such as GaAsSb [216] or InGaAs [42]. The latter approach is particularly important for tuning laser emission towards the telecommunications band; however, this remains an outstanding challenge in quantum-well systems.

The growth of high-quality radial quantum wells relies on the abrupt switching of composition during growth. Heterostructure growth onto nanowires is more challenging than conventional planar growth as

nanowires have multiple facets with potentially different growth kinetics [217], or due to inter-diffusion during growth [157]. Nevertheless, recent growth approaches have achieved near constant alloy composition in the wells, and with highly uniform well thicknesses [42].

4.1.2 Axial quantum dot nanowire lasers

By altering the composition of precursors during nanowire growth, axially structured nanowires can be produced [218,219]. Axial heterostructure with quantum confinement is challenging to fabricate, as very thin low-bandgap inclusions must be produced. In the VLS method (Section 3.1.1), growth occurs in the presence of a metallic nanoparticle which can act as a reservoir for precursors, inhibiting the formation of abrupt interfaces. As such, early demonstrations of axial quantum confinement for lasing were produced through SAG (catalyst-free) [124]. More recently, VLS-grown nanowire lasers incorporating quantum dots or discs have been demonstrated, where the reservoir effect has been managed [53].

It is notable that axial nanowire quantum dots have applications beyond lasing, and early work explored these as single photon sources [220]. In contrast, for high efficiency nanowire lasers, multiple quantum dots must be produced. It is important that these remain uncoupled, and that they are as uniform composition and thickness as possible. This is best achieved by use of binary wells to avoid issues with ternary composition; for example, an InP nanowire laser containing InAs quantum dots has been demonstrated operating at room temperature, albeit with a high operating threshold (>2 mJcm^{-2}) [53].

4.2 Optical coupling

In Fabry–Pérot-type nanowire lasers, the electric field is confined to the nanowire cavity. However, the nanoscale confinement leads to evanescent fields penetrating the surrounding medium, providing a route to manipulate laser operation or provide new functionality. Coupling of two nanowire lasers through evanescent fields was theoretically explored for parallel nanowires [163]; very high efficiencies of over 95% were predicted for optimised coupling in silicon. However, this geometry is challenging to fabricate as it requires sub-wavelength positioning.

To experimentally study coupling effects, Gao and colleagues [98] used ion-beam milling to cleave long nanowire lasers leaving a small gap between segments. This approach was low loss for gaps of less than 50 nm, allowing for lasing across the two coupled nanowire lasers for true single mode operation. More recently, advanced pick-and-place methods have been proposed to

physically place nanowire lasers [33]. With sub-100 nm precision, it is possible to couple wires with antennas [221] to manipulate far-field behaviour. However, the removal of nanowire lasers from their host substrate and their transfer to a receiver substrate can lead to increased disorder in end-facet reflectivity [32] or other mechanical changes which impact the laser threshold or centre wavelength [109].

More generally, the monolithic integration of nanowires with waveguides is required for scalable on-chip photonic applications. Using SAG, nanowire lasers have been directly coupled onto silicon waveguides [30] and insulating waveguides [47,103]. Theoretical work based on finite-difference time-domain modelling of nanowire lasers suggests that waveguide coupling with 50% efficiency might be achieved through optimisation of the nanowire–waveguide interface and dielectric properties [196]; this holds great promise for on-chip optical interconnects.

Nanowire lasers can be used to couple light into on-chip photonic crystal cavities. Sergent and colleagues have investigated the theoretical [222] and experimental approaches to fabricating nanowire lasers coupled to photonic-crystal nanocavities [223], work that exploits the efficiency and tunability of these sources. AFM can be used to both image the structures and to manipulate nanowires into the cavities, and these structures have been seen to host quasi-continuous ultraviolet lasing at room temperature [136]. Notably, the wires used may have a diameter below that required for Fabry–Pérot lasing, as the mode is supported by the photonic crystal rather than the nanowire cavity.

4.3 Plasmonics

A plasmon refers to the quantum of a coherent oscillation of charge; when charge is coupled to an optical field, a surface-plasmon polariton (SPP) can be formed. SPPs have properties of both photons and plasmons, and allow the focussing of electromagnetic radiation in deep sub-wavelength regions. Coupling to SPPs is in general challenging, as it requires both energy and momentum matching. In 2003, Bergman and Stockman [93] introduced the concept of a surface-plasmon laser, a device that can generate coherent SPPs [224]. Nanowires lasers with proximal metallic structures were recognised as a building block for plasmonics [10], which was realised in the CdS system [117] for a plasmonic mode area of $1/400\ \lambda^2$ (where $\lambda = 489$ nm).

Since the first development, plasmonic-based lasers have been demonstrated in CdS [120], ZnO [123], GaAs [94], GaN [122,225] and perovskite materials [133]. In general, nanowires are transferred to planar dielectric-

coated metal (or a structured plasmonic waveguide [127]). Plasmonic lasers demonstrate deep sub-wavelength confinement [117], lasing on leaky modes [95], ultrafast switching behaviour [123] and for waveguiding [127]. Importantly, the field is confined to a small volume providing an opportunity for so-called "thresholdless" lasers, where the coupling of light to the lasing mode approaches 100%. The development of plasmonic applications remains an active area of research [226].

5. State of the art and outlook

After 20 years of research on nanowire laser physics, materials and application, many of the original proposals have been achieved in a range of material systems. Single nanowire continuous lasing [211], telecommunications wavelength lasers [53], silicon-integrated [31,41], room-temperature [64], laser arrays [177] and electrically pumped lasers [35] have been independently demonstrated. Along the way, novel physical approaches (such as SPP lasers [117]), materials (such as perovskites [19]) and applications (such as photonic integrated circuitry [13] and as intracellular probes [60]) have been explored.

However, some of the original goals for this architecture are not yet met. In particular, single nanowire, addressable, silicon-integrated and electrically pumped nanowire lasers have not been demonstrated [17,23]. Whilst not discussed in this chapter, key issues remain the production of high quality p-i-n type junctions for electrical injection, and issues of heat and subsequent damage arising from any inefficiencies at the nanoscale. A focus on industrial yield is also needed for future applications [92,134,227]; here, templated nanowire growth provides an opportunity area for developing high-quality site-controlled materials [228] with research on doping [229] and nanowire-seeded growth continuing at pace [230].

The promises of nanowires as tunable sub-micron sources of coherent light have been demonstrated. Integration of nanowire lasers in scientific and commodity applications for robust and reliable performance is the remaining frontier for development, as research on nanowire lasers enters its third decade.

Acknowledgements

We thank Prof. Phillip Dawson and Dr Juan Arturo Alanis for useful discussions. PP acknowledges support from the UKRI under the Future Leaders Fellowship scheme (MR/T021519/1).

References

[1] Johnson JC, Yan H, Schaller RD, Haber LH, Saykally RJ, Yang P. Single nanowire lasers. J Phys Chem B 2001;105:11387–90. https://doi.org/10.1021/jp012304t. https://pubs.acs.org/doi/10.1021/jp012304t.

[2] Huang MH, Mao S, Feick H, Yan H, Wu Y, Kind H, Weber E, Russo R, Yang P. Room-temperature ultraviolet nanowire nanolasers. Science 2001;292:1897–9. https://doi.org/10.1126/science.1060367. http://www.ncbi.nlm.nih.gov/pubmed/11397941. http://www.sciencemag.org/cgi/doi/10.1126/science.1060367.

[3] Javan A, Bennett WR, Herriott DR. Population inversion and continuous optical maser oscillation in a gas discharge containing a He-Ne mixture. Phys Rev Lett 1961;6:106–10. https://doi.org/10.1103/PhysRevLett.6.106. https://link.aps.org/doi/10.1103/PhysRevLett.6.106.

[4] Hall RN, Fenner GE, Kingsley JD, Soltys TJ, Carlson RO. Coherent light emission from GaAs junctions. Phys Rev Lett 1962;9:366–8. https://doi.org/10.1103/PhysRevLett.9.366. https://link.aps.org/doi/10.1103/PhysRevLett.9.366.

[5] Holonyak N, Bevacqua SF. Coherent (visible) light emission from Ga(As1-xPx) junctions. Appl Phys Lett 1962;1:82–3. https://doi.org/10.1063/1.1753706. http://aip.scitation.org/doi/10.1063/1.1753706.

[6] Kroemer H. A proposed class of hetero-junction injection lasers. Proc IEEE 1963;51:1782–3. https://doi.org/10.1109/PROC.1963.2706. http://ieeexplore.ieee.org/document/1444636/.

[7] Hayashi I, Panish MB, Foy PW, Sumski S. Junction lasers which operate continuously at room temperature. Appl Phys Lett 1970;17:109–11. https://doi.org/10.1063/1.1653326. http://aip.scitation.org/doi/10.1063/1.1653326.

[8] Holonyak N, Kolbas RM, Dupuis RD, Dapkus PD. Quantum-well heterostructure lasers. IEEE J Quant Electron 1980;16:170–86. https://doi.org/10.1109/JQE.1980.1070447. http://ieeexplore.ieee.org/document/1070447/.

[9] Ledentsov NN, Grundmann M, Heinrichsdorff F, Bimberg D, Ustinov VM, Zhukov AE, Maximov MV, Alferov ZI, Lott JA. Quantum-dot heterostructure lasers. IEEE J Sel Top Quant Electron 2000;6:439–51. https://doi.org/10.1109/2944.865099. http://ieeexplore.ieee.org/document/865099/.

[10] Maslov AV, Ning CZ. Size reduction of a semiconductor nanowire laser by using metal coating. In: Osinski M, Henneberger F, Arakawa Y, editors. Physics and simulation of optoelectronic devices XV, vol. 6468. International Society for Optics and Photonics; 2007. 64680I. https://doi.org/10.1117/12.723786. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.723786.

[11] Wu X, Xiao Y, Meng C, Zhang X, Yu S, Wang Y, Yang C, Guo X, Ning CZ, Tong L. Hybrid photon-plasmon nanowire lasers. Nano Lett 2013;13:5654–9. https://doi.org/10.1021/nl403325j. https://pubs.acs.org/doi/10.1021/nl403325j.

[12] Eaton SW, Lai M, Gibson NA, Wong AB, Dou L, Ma J, Wang L-W, Leone SR, Yang P. Lasing in robust cesium lead halide perovskite nanowires. Proc Natl Acad Sci USA 2016;113:1993–8. https://doi.org/10.1073/pnas.1600789113. http://www.pnas.org/lookup/doi/10.1073/pnas.1600789113.

[13] Shen Y, Harris NC, Skirlo S, Prabhu M, Baehr-Jones T, Hochberg M, Sun X, Zhao S, Larochelle H, Englund D, Soljačić M. Deep learning with coherent nanophotonic circuits. Nat Photonics 2017;11:441–6. https://doi.org/10.1038/nphoton.2017.93. http://arxiv.org/abs/1610.02365. http://www.nature.com/doifinder/10.1038/nphoton.2017.93. http://www.nature.com/articles/nphoton.2017.93.

[14] Wang ZL. ZnO nanowire and nanobelt platform for nanotechnology. Mater Sci Eng R Rep 2009;64(3–4):33–71. https://doi.org/10.1016/j.mser.2009.02.001. https://www.sciencedirect.com/science/article/pii/S0927796X09000229.

[15] Vanmaekelbergh D, Van Vugt LK. ZnO nanowire lasers. Nanoscale 2011;3:2783–800. https://doi.org/10.1039/c1nr00013f. http://xlink.rsc.org/?DOI=c1nr00013f.

[16] Zimmler MA, Capasso F, Müller S, Ronning C. Optically pumped nanowire lasers: invited review. Semicond Sci Technol 2010;25:024001. https://doi.org/10.1088/0268–1242/25/2/024001. http://stacks.iop.org/0268–1242/25/i=2/a=024001?key=crossref.19fed2250a6e91e6a00155a2b2.
[17] Koblmiiller G, Mayer B, Stettner T, Abstreiter G, Finley JJ. GaAs-AlGaAs core-shell nanowire lasers on silicon: invited review. Semicond Sci Technol 2017;32:053001. https://doi.org/10.1088/1361–6641/aa5e45. http://stacks.iop.org/0268–1242/32/i=5/a=053001?key=crossref.ba4e12afb515deb7c083b6a69c.
[18] Zhang J, Yang X, Deng H, Qiao K, Farooq U, Ishaq M, Yi F, Liu H, Tang J, Song H. Low-dimensional halide perovskites and their advanced optoelectronic applications. Nanomicro Lett 2017;9(3):36. https://doi.org/10.1007/s40820–017–0137–5. http://link.springer.com/10.1007/s40820–017–0137-5.
[19] Dong H, Zhang C, Liu X, Yao J, Zhao YS. Materials chemistry and engineering in metal halide perovskite lasers. Chem Soc Rev 2020;49:951–82. https://doi.org/10.1039/c9cs00598f. http://xlink.rsc.org/?DOI=C9CS00598F.
[20] Arafin S, Liu X, Mi Z. Review of recent progress of III-nitride nanowire lasers. J Nanophotonics 2013;7:074599. https://doi.org/10.1117/1.jnp.7.074599. http://nanophotonics.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JNP.7.074599.
[21] Huang Y, Lieber CM. Integrated nanoscale electronics and optoelectronics: exploring nanoscale science and technology through semiconductor nanowires. Pure Appl Chem 2004;76:2051–68. https://doi.org/10.1351/pac200476122051. http://www.degruyter.com/view/j/pac.2004.76.issue-12/pac200476122051/pac200476122051.xn.
[22] Zapf M, Sidiropoulos T, Röder R. Tailoring spectral and temporal properties of semiconductor nanowire lasers. Adv Opt Mater 2019;7(17):1900504. https://doi.org/10.1002/adom.201900504. https://onlinelibrary.wiley.com/doi/abs/10.1002/adom.201900504.
[23] Zhang Y, Saxena D, Aagesen M, Liu H. Toward electrically driven semiconductor nanowire lasers. Nanotechnology 2019;30:192002. https://doi.org/10.1088/1361-6528/ab000d. http://stacks.iop.org/0957-4484/30/i=19/a=192002?key=crossref.7251158d04d4bddab0ec7bed.
[24] Maslov AV, Ning CZ. Reflection of guided modes in a semiconductor nanowire laser. Appl Phys Lett 2003;83:1237–9. https://doi.org/10.1063/1.1599037. http://aip.scitation.org/doi/10.1063/1.1599037.
[25] Heurlin M, Lindgren D, Deppert K, Samuelson L, Magnusson MH, Ek ML, Wallenberg R. Continuous gas-phase synthesis of nanowires with tunable properties. Nature 2012;492(7427):90–4. https://doi.org/10.1038/nature11652. http://www.nature.com/articles/nature11652.
[26] Jevtics D, Hurtado A, Guilhabert B, McPhillimy J, Cantarella G, Gao Q, Tan HH, Jagadish C, Strain MJ, Dawson MD. Integration of semiconductor nanowire lasers with polymeric waveguide devices on a mechanically flexible substrate. Nano Lett 2017;17:5990–4. https://doi.org/10.1021/acs.nanolett.7b02178. https://pubs.acs.org/doi/10.1021/acs.nanolett.7b02178.
[27] Yang P, Yan R, Fardy M. Semiconductor nanowire: what's next? Nano Lett 2010;10:1529–36. https://doi.org/10.1021/nl100665r. http://pubs.acs.org/doi/abs/10.1021/nl100665r.
[28] Zhang Y, Aagesen M, Holm JV, Jørgensen HI, Wu J, Liu H. Self-catalyzed GaAsP nanowires grown on silicon substrates by solid-source molecular beam epitaxy. Nano Lett 2013;13:3897–902. https://doi.org/10.1021/nl401981u. http://pubs.acs.org/doi/abs/10.1021/nl401981u.
[29] Tatebayashi J, Kako S, Ho J, Ota Y, Iwamoto S, Arakawa Y. Growth of InGaAs/GaAs nanowire-quantum dots on Al- GaAs/GaAs distributed Bragg reflectors for

laser applications. J Cryst Growth 2017;468:144–8. https://doi.org/10.1016/j.jcrysgro.2016.12.022. https://linkinghub.elsevier.com/retrieve/pii/S0022024816308818.

[30] Stettner T, Kostenbader T, Ruhstorfer D, Bissinger J, Riedl H, Kaniber M, Koblmüller G, Finley JJ. Direct coupling of coherent emission from site-selectively grown III-V nanowire lasers into proximal silicon waveguides. ACS Photonics 2017;4:2537–43. https://doi.org/10.1021/acsphotonics.7b00805. https://pubs.acs.org/doi/10.1021/acsphotonics.7b00805.

[31] Schuster F, Kapraun J, Malheiros-Silveira GN, Deshpande S, Chang-Hasnain CJ. Site-controlled growth of monolithic InGaAs/InP quantum well nanopillar lasers on silicon. Nano Lett 2017;17:2697–702. https://doi.org/10.1021/acs.nanolett.7b00607. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.7b00607.

[32] Alanis JA, Chen Q, Lysevych M, Burgess T, Li L, Liu Z, Tan HH, Jagadish C, Parkinson P. Threshold reduction and yield improvement of semiconductor nanowire lasers: via processing-related end-facet optimization. Nanoscale Adv 2019;1:4393–7. https://doi.org/10.1039/c9na00479c. http://xlink.rsc.org/?DQI=C9NA00479C.

[33] Guilhabert B, Hurtado A, Jevtics D, Gao Q, Tan HH, Jagadish C, Dawson MD. Transfer printing of semiconductor nanowires with lasing emission for controllable nanophotonic device fabrication. ACS Nano 2016;10:3951–8. https://doi.org/10.1021/acsnano.5b07752. https://pubs.acs.org/doi/10.1021/acsnano.5b07752.

[34] Tatebayashi J, Ota Y, Ishida S, Nishioka M, Iwamoto S, Arakawa Y. Nanowire-quantum-dot lasers on flexible membranes. APEX 2018;11:065002. https://doi.org/10.7567/APEX.11.065002. http://stacks.iop.org/1882-0786/11/i=6/a=065002?key=crossref.0610217fca1b234429c84cb0f.

[35] Duan X, Huang Y, Agarwal R, Lieber CMC, Fast CG. Single-nanowire electrically driven lasers. Nature 2003;421:241–5. https://doi.org/10.1038/nature01353. http://www.nature.com/articles/nature01353.

[36] Chu S, Wang G, Zhou W, Lin Y, Chernyak L, Zhao J, Kong J, Li L, Ren J, Liu J. Electrically pumped waveguide lasing from ZnO nanowires. Nat Nanotechnol 2011;6:506–10. https://doi.org/10.1038/nnano.2011.97. http://www.nature.com/doifinder/10.1038/nnano.2011.97.

[37] Zhao S, Liu X, Wu Y, Mi Z. An electrically pumped 239 nm AlGaN nanowire laser operating at room temperature. Appl Phys Lett 2016;109:191106. https://doi.org/10.1063/1.4967180. http://aip.scitation.org/doi/10.1063/1.4967180.

[38] Young IA, Mohammed E, Liao JT, Kern AM, Palermo S, Block BA, Reshotko MR, Chang PL. Optical I/O technology for tera-scale computing. IEEE J Solid State Circ 2010;45:235–48. https://doi.org/10.1109/JSSC.2009.2034444. http://ieeexplore.ieee.org/document/5357567/.

[39] Tanaka S, Jeong S-H, Sekiguchi S, Kurahashi T, Tanaka Y, Morito K. High-output-power, single-wavelength silicon hybrid laser using precise flip-chip bonding technology. Opt Express 2012;20:28057. https://doi.org/10.1364/oe.20.028057. https://www.osapublishing.org/oe/abstract.cfm?uri=oe-20-27-28057.

[40] Sun C, Wade MT, Lee Y, Orcutt JS, Alloatti L, Georgas MS, Waterman AS, Shainline JM, Avizienis RR, Lin S, Moss BR, Kumar R, Pavanello F, Atabaki AH, Cook HM, Ou AJ, Leu JC, Chen Y-H, Asanović K, Ram RJ, Popović MA, Stojanović VM. Single-chip microprocessor that communicates directly using light. Nature 2015;528:534–8. https://doi.org/10.1038/nature16454. http://www.nature.com/doifinder/10.1038/nature16454.

[41] Mayer B, Janker L, Loitsch B, Treu J, Kostenbader T, Lichtmannecker S, Reichert T, Morkötter S, Kaniber M, Abstreiter G, Gies C, Koblmüller G, Finley JJ. Monolithically integrated high- β nanowire lasers on silicon. Nano Lett 2016;16:152–6. https://

doi.org/10.1021/acs.nanolett.5b03404. https://pubs.acs.org/doi/10.1021/acs.nanolett.5b03404. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.5b03404.

[42] Stettner T, Thurn A, Döblinger M, Hill MO, Bissinger J, Schmiedeke P, Matich S, Kostenbader T, Ruhstorfer D, Riedl H, Kaniber M, Lauhon LJ, Finley JJ, Koblmüller G. Tuning lasing emission toward long wavelengths in GaAs-(in,Al)GaAs core-multishell nanowires. Nano Lett 2018;18:6292–300. https://doi.org/10.1021/acs.nanolett.8b02503. http://pubs.acs.org/doi/10.1021/acs.nanolett.8b02503. https://pubs.acs.org/doi/10.1021/acs.nanolett.8b02503.

[43] Bao Q, Li W, Xu P, Zhang M, Dai D, Wang P, Guo X, Tong L. On-chip single-mode CdS nanowire laser. Light Sci Appl 2020;9(42). https://doi.org/10.1038/s41377-020-0277-0. http://www.nature.com/articles/s41377-020-0277-0.

[44] Chen B, Wu H, Xin C, Dai D, Tong L. Flexible integration of free-standing nanowires into silicon photonics. Nat Commun 2017;8:20. https://doi.org/10.1038/s41467-017-00038-0. http://www.nature.com/articles/s41467-017-00038-0.

[45] ITU-T. ITU-T G. 652: characteristics of a single-mode optical fibre and cable. Technical Report. Geneva, Switzerland: ITU; 2016. https://doi.org/10.1002/1000/13076. https://www.itu.int/itu-t/recommendations/rec.aspx?rec=13076.

[46] Chin AH, Vaddiraju S, Maslov AV, Ning CZ, Sunkara MK, Meyyappan M. Near-infrared semiconductor subwavelength-wire lasers. Appl Phys Lett 2006;88:163115. https://doi.org/10.1063/1.2198017. http://aip.scitation.org/doi/10.1063/1.2198017.

[47] Kim H, Lee WJ, Farrell AC, Balgarkashi A, Huffaker DL. Telecom-wavelength bottom-up nanobeam lasers on silicon-on-insulator. Nano Lett 2017;17:5244–50. https://doi.org/10.1021/acs.nanolett.7b01360. https://pubs.acs.org/doi/10.1021/acs.nanolett.7b01360.

[48] Kim H, Lee W-J, Chang T-Y, Huffaker DL. Room-temperature InGaAs nanowire array band-edge lasers on patterned silicon-on-insulator platforms. Phys Status Solidi (RRL) 2019;13:1800489. https://doi.org/10.1002/pssr.201800489. http://doi.wiley.com/10.1002/pssr.201800489.

[49] Kim H, Chang TY, Lee WJ, Huffaker DL. III-V nanowire array telecom lasers on (001) silicon-on-insulator photonic platforms. Appl Phys Lett 2019;115:213101. https://doi.org/10.1063/1.5126721. https://aip.scitation.org/doi/10.1063/1.5126721.

[50] Yokoo A, Takiguchi M, Birowosuto MD, Tateno K, Zhang G, Kuramochi E, Shinya A, Taniyama H, Notomi M. Subwave-length nanowire lasers on a silicon photonic crystal operating at telecom wavelengths. ACS Photonics 2017;4:355–62. https://doi.org/10.1021/acsphotonics.6b00830. https://pubs.acs.org/doi/10.1021/acsphotonics.6b00830.

[51] Deshpande S, Bhattacharya I, Malheiros-Silveira G, Ng KW, Schuster F, Mantei W, Cook K, Chang-Hasnain C. Ultracompact position-controlled InP nanopillar LEDs on silicon with bright electroluminescence at telecommunication wavelengths. ACS Photonics 2017;4:695–702. https://doi.org/10.1021/acsphotonics.7b00065. https://pubs.acs.org/doi/10.1021/acsphotonics.7b00065.

[52] Chen S, Yukimune M, Fujiwara R, Ishikawa F, Chen WM, Buyanova IA. Near-infrared lasing at 1 μm from a dilute-nitride-based multishell nanowire. Nano Lett 2019;19:885–90. https://doi.org/10.1021/acs.nanolett.8b04103. http://pubs.acs.org/doi/10.1021/acs.nanolett.8b04103. https://pubs.acs.org/doi/10.1021/acs.nanolett.8b04103.

[53] Zhang G, Takiguchi M, Tateno K, Tawara T, Notomi M, Gotoh H. Telecom-band lasing in single InP/InAs heterostructure nanowires at room temperature. Sci Adv 2019;5:eaat8896. https://doi.org/10.1126/sciadv.aat8896. https://advances.sciencemag.org/lookup/doi/10.1126/sciadv.aat8896. http://advances.sciencemag.org/content/5/2/eaat8896?rss=1.

[54] Patolsky F, Zheng G, Lieber CM. Nanowire sensors for medicine and the life sciences. Nanomedicine 2006;1:51–65. https://doi.org/10.2217/17435889.1.1.51. http://www.futuremedicine.com/doi/10.2217/17435889.1–1.51.
[55] He B, Morrow TJ, Keating CD. Nanowire sensors for multiplexed detection of biomolecules. Curr Opin Chem Biol 2008;12:522–8. https://doi.org/10.1016/j.cbpa.2008.08.027. https://linkinghub.elsevier.com/retrieve/pii/S1367593108001245.
[56] Gautam V, Naureen S, Shahid N, Gao Q, Wang Y, Nisbet D, Jagadish C, Daria VR. Engineering highly interconnected neuronal networks on nanowire scaffolds. Nano Lett 2017;17:3369–75. https://doi.org/10.1021/acs.nanolett.6b05288. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.6b05288. https://pubs.acs.org/doi/10.1021/acs.nanolett.6b05288.
[57] Liu R, Chen R, Elthakeb AT, Lee SH, Hinckley S, Khraiche ML, Scott J, Pre D, Hwang Y, Tanaka A, Ro YG, Matsushita AK, Dai X, Soci C, Biesmans S, James A, Nogan J, Jungjohann KL, Pete DV, Webb DB, Zou Y, Bang AG, Dayeh SA. High density individually addressable nanowire arrays record intracellular activity from primary rodent and human stem cell derived neurons. Nano Lett 2017;17:2757–64. https://doi.org/10.1021/acs.nanolett.6b04752. https://pubs.acs.org/doi/10.1021/acs.nanolett.6b04752.
[58] Fikouras AH, Schubert M, Karl M, Kumar JD, Powis SJ, Di Falco A, Gather MC. Non-obstructive intracellular nanolasers. Nat Commun 2018;9:4817. https://doi.org/10.1038/s41467-018-07248-0. http://www.nature.com/articles/s41467-018-07248-0.
[59] Yan R, Park J-H, Choi Y, Heo C-J, Yang S-M, Lee LP, Yang P. Nanowire-based single-cell endoscopy. Nat Nanotechnol 2012;7:191–6. https://doi.org/10.1038/nnano.2011.226. http://www.nature.com/articles/nnano.2011.226.
[60] Wu X, Chen Q, Xu P, Chen YC, Wu B, Coleman RM, Tong L, Fan X. Nanowire lasers as intracellular probes. Nanoscale 2018;10:9729–35. https://doi.org/10.1039/c8nr00515j. http://xlink.rsc.org/?DOI=C8NR00515J.
[61] Gordon JP, Zeiger HJ, Townes CH. The maser-new type of microwave amplifier, frequency standard, and spectrometer. Phys Rev 1955;99:1264–74. https://doi.org/10.1103/PhysRev.99.1264. https://link.aps.org/doi/10.1103/PhysRev.99.1264.
[62] Siegman AE. Lasers. 1st ed. Sausalito: University Science Books; 1986.
[63] Ning CZ. Semiconductor nanolasers. Phys. Status Solidi B 2010;247:774–88. https://doi.org/10.1002/pssb.200945436. http://doi.wiley.com/10.1002/pssb.200945436.
[64] Saxena D, Mokkapati S, Parkinson P, Jiang N, Gao Q, Tan HH, Jagadish C. Optically pumped room-temperature GaAs nanowire lasers. Nat Photonics 2013;7:963–8. https://doi.org/10.1038/nphoton.2013.303. http://www.nature.com/doifinder/10.1038/nphoton.2013.303. http://www.nature.com/articles/nphoton.2013.303.
[65] Kennedy SA, Szabo MJ, Teslow H, Porterfield JZ, Abraham ER. Creation of Laguerre-Gaussian laser modes using diffractive optics. Phys Rev, A Atom Mol Opt Phy 2002;66:438011–5. https://doi.org/10.1103/PhysRevA.66.043801.
[66] Maslov AV, Ning CZ. Modal gain in a semiconductor nanowire laser with anisotropic bandstructure. IEEE J Quant Electron 2004;40:1389–97. https://doi.org/10.1109/JQE.2004.834767. http://ieeexplore.ieee.org/document/1337019/.
[67] Visser TD, Blok H, Demeulenaere B, Lenstra D. Confinement factors and gain in optical amplifiers. IEEE J Quant Electron 1997;33:1763–6. https://doi.org/10.1109/3.631280. http://ieeexplore.ieee.org/document/631280/.
[68] Ning C-Z. Semiconductor nanolasers and the size-energy-efficiency challenge: a review. Adv Photonics 2019;1:1. https://doi.org/10.1117/1.AP.1.1.014002. https://www.spiedigitallibrary.org/journals/advanced-photonics/volume-1/issue-01/014002.

[69] Hilborn RC. Einstein coefficients, cross sections, f values, dipole moments, and all that. Am J Phys 1982;50:982–6. https://doi.org/10.1119/1.12937. http://aapt.scitation.org/doi/10.1119/1.12937.
[70] Bernard MG, Duraffourg G. Laser conditions in semiconductors. Phys Status Solidi (B) 1961;1:699–703. https://doi.org/10.1002/pssb.19610010703. http://doi.wiley.com/10.1002/pssb.19610010703.
[71] Wegscheider W, Pfeiffer LN, Dignam MM, Pinczuk A, West KW, McCall SL, Hull R. Lasing from excitons in quantum wires. Phys Rev Lett 1993;71:4071–4. https://doi.org/10.1103/PhysRevLett.71.4071. https://link.aps.org/doi/10.1103/PhysRevLett.71.4071.
[72] Kawasaki M, Ohtomo A, Ohkubo I, Koinuma H, Tang ZK, Yu P, Wong GK, Zhang BP, Segawa Y. Excitonic ultraviolet laser emission at room temperature from naturally made cavity in ZnO nanocrytal thin films. Mater Sci Eng B 1998; 56:239–45. https://doi.org/10.1016/S0921-5107(98)00248-7.
[73] Huang Y, Duan X, Cui Y, Lauhon LJ, Kim K-H, Lieber CM. Logic gates and computation from assembled nanowire building blocks. Science 2001;294:1313–7. https://doi.org/10.1126/science.1066192. http://www.sciencemag.org/cgi/doi/10.1126/science.1066192.
[74] Agarwal R, Barrelet CJ, Lieber CM. Lasing in single cadmium sulfide nanowire optical cavities. Nano Lett 2005;5:917–20. https://doi.org/10.1021/nl050440u. https://pubs.acs.org/doi/10.1021/nl050440u.
[75] Dingle R, Wiegmann W, Henry CH. Quantum states of confined carriers in very thin AlxGa1-xAs-GaAs-AlxGa1-xAs heterostructures. Phys Rev Lett 1974;33: 827–30. https://doi.org/10.1103/PhysRevLett.33.827. https://link.aps.org/doi/10.1103/PhysRevLett.33.827.
[76] Arakawa Y, Sakaki H. Multidimensional quantum well laser and temperature dependence of its threshold current. Appl Phys Lett 1982;40:939–41. https://doi.org/10.1063/1.92959. http://aip.scitation.org/doi/10.1063/1.92959.
[77] Qian F, Li Y, Gradečak S, Park H-GG, Dong Y, Ding Y, Wang ZL, Lieber CM. Multi-quantum-well nanowire heterostructures for wavelength-controlled lasers. Nat Mater 2008;7:701–6. https://doi.org/10.1038/nmat2253. http://www.nature.com/articles/nmat2253. http://www.nature.com/doifinder/10.1038/nmat2253.
[78] Saxena D, Jiang N, Yuan X, Mokkapati S, Guo Y, Tan HH, Jagadish C. Design and room-temperature operation of GaAs/AlGaAs multiple quantum well nanowire lasers. Nano Lett 2016;16:5080–6. https://doi.org/10.1021/acs.nanolett.6b01973. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.6b01973. https://pubs.acs.org/doi/10.1021/acs.nanolett.6b01973.
[79] Leonard D, Pond K, Petroff PM. Critical layer thickness for self-assembled InAs islands on GaAs. Phys Rev B 1994;50:11687–92. https://doi.org/10.1103/PhysRevB.50.11687. https://link.aps.org/doi/10.1103/PhysRevB.50.11687.
[80] Heiss M, Fontana Y, Gustafsson A, Wüst G, Magen C, O'Regan DD, Luo JW, Ketterer B, Conesa-Boj S, Kuhlmann AV, Houel J, Russo-Averchi E, Morante JR, Cantoni M, Marzari N, Arbiol J, Zunger A, Warburton RJ, Fontcuberta I Morral A. Self-assembled quantum dots in a nanowire system for quantum photonics. Nat Mater 2013;12:439–44. https://doi.org/10.1038/nmat3557. http://www.nature.com/articles/nmat3557.
[81] Wang S, Hu Z, Yu H, Fang W, Qiu M, Tong L. Endface reflectivities of optical nanowires. Opt Express 2009;17:10881. https://doi.org/10.1364/0E.17.010881. http://www.ncbi.nlm.nih.gov/pubmed/19550488https://www.osapublishing.org/oe/abstract.ci.
[82] Zapien JA, Jiang Y, Meng XM, Chen W, Au FC, Lifshitz Y, Lee ST. Room-temperature single nanoribbon lasers. Appl Phys Lett 2004;84:1189–91. https://doi.org/10.1063/1.1647270. http://aip.scitation.org/doi/10.1063/1.1647270.

[83] Wagner RS, Ellis WC. Vapor-liquid-solid mechanism of single crystal growth. Appl Phys Lett 1964;4:89–90. https://doi.org/10.1063/1.1753975. http://aip.scitation.org/doi/10.1063/1.1753975.
[84] Morales AM, Lieber CM. A laser ablation method for the synthesis of crystalline semiconductor nanowires. Science 1998;279:208–11. https://doi.org/10.1126/science.279.5348.208. http://www.ncbi.nlm.nih.gov/pubmed/9422689.
[85] Trentler TJ, Hickman KM, Goel SC, Viano AM, Gibbons PC, Buhro WE. Solution-liquid-solid growth of crystalline III-V semiconductors: an analogy to vapor- liquid-solid growth. Science 1995;270:1791–4. https://doi.org/10.1126/science.270.5243.1791. https://www.sciencemag.org/lookup/doi/10.1126/science.270.5243.1791.
[86] Gudiksen MS, Lieber CM. Diameter-selective synthesis of semiconductor nanowires. J Am Chem Soc 2000;122:8801–2. https://doi.org/10.1021/ja002008e. https://pubs.acs.org/doi/10.1021/ja002008e.
[87] Maslov AV, Ning CZ. Far-field emission of a semiconductor nanowire laser. Opt Lett 2004;29:572–4. https://doi.org/10.1364/OL.29.000572. https://www.osapublishing.org/abstract.cfm?URI=ol-29-6-572. http://www.ncbi.nlm.nih.gov/pubmed/15035474.
[88] Sun L, Ren ML, Liu W, Agarwal R. Resolving parity and order of fabry-pérot modes in semiconductor nanostructure waveguides and lasers: Youngs interference experiment revisited. Nano Lett 2014;14:6564–71. https://doi.org/10.1021/nl503176w. https://pubs.acs.org/doi/10.1021/nl503176w.
[89] Saxena D, Wang F, Gao Q, Mokkapati S, Tan HH, Jagadish C. Mode profiling of semiconductor nanowire lasers. Nano Lett 2015;15:5342–8. https://doi.org/10.1021/acs.nanolett.5b01713. https://pubs.acs.org/doi/10.1021/acs.nanolett.5b01713.
[90] Yan X, Wei W, Tang F, Wang X, Li L, Zhang X, Ren X. Low-threshold room-temperature AlGaAs/GaAs nanowire/single-quantum-well heterostructure laser. Appl Phys Lett 2017;110:061104. https://doi.org/10.1063/1.4975780. http://aip.scitation.org/doi/10.1063/1.4975780.
[91] Stettner T, Zimmermann P, Loitsch B, Döblinger M, Regler A, Mayer B, Winnerl J, Matich S, Riedl H, Kaniber M, Abstreiter G, Koblmuller G, Finley JJ. Coaxial GaAs-AlGaAs core-multishell nanowire lasers with epitaxial gain control. Appl Phys Lett 2016;108:011108. https://doi.org/10.1063/1.4939549. http://aip.scitation.org/doi/10.1063/1.4939549.
[92] Alanis JA, Saxena D, Mokkapati S, Jiang N, Peng K, Tang X, Fu L, Tan HH, Jagadish C, Parkinson P. Large-scale statistics for threshold optimization of optically pumped nanowire lasers. Nano Lett 2017;17:4860–5. https://doi.org/10.1021/acs.nanolett.7b01725. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.7b01725. http://pubs.acs.org/doi/10.1021/acs.nanolett.7b01725.
[93] Bergman DJ, Stockman MI. Surface plasmon amplification by stimulated emission of radiation: quantum generation of coherent surface plasmons in nanosystems. Phys Rev Lett 2003;90:4. https://doi.org/10.1103/PhysRevLett.90.027402. https://link.aps.org/doi/10.1103/PhysRevLett.90.027402.
[94] Ho J, Tatebayashi J, Sergent S, Fong C, Iwamoto S, Arakawa Y. Demonstration of GaAs based nanowire plasmonic laser. In: Extended abstracts of the 2014 international conference on solid state devices and materials. Ibaraki: The Japan Society of Applied Physics; 2014. https://doi.org/10.7567/SSDM.2014.B-3-1. 3–1, https://confit.atlas.jp/guide/organizer/ssdm/ssdm2014/subject/B-3-1/detail.
[95] Wuestner S, Hamm JM, Pusch A, Hess O. Plasmonic leaky-mode lasing in active semiconductor nanowires. Laser Photon Rev 2015;9:256–62. https://doi.org/10.1002/lpor.201400231. http://doi.wiley.com/10.1002/lpor.201400231.
[96] Wei W, Liu Y, Zhang X, Wang Z, Ren X. Evanescent-wave pumped room-temperature single-mode GaAs/AlGaAs core-shell nanowire lasers. Appl Phys Lett

2014;104:223103. https://doi.org/10.1063/1.4881266. http://aip.scitation.org/doi/10.1063/1.4881266.
[97] Buschlinger R, Lorke M, Peschel U. Coupled-mode theory for semiconductor nanowires. Phys Rev Appl 2017;7:034028. https://doi.org/10.1103/PhysRevApplied.7.034028. http://link.aps.org/doi/10.1103/PhysRevApplied.7.034028. http://arxiv.org/abs/1611.09672.
[98] Gao H, Fu A, Andrews SC, Yang P. Cleaved-coupled nanowire lasers. Proc Nat Acad Sci USA 2013;110(3):865–9. https://doi.org/10.1073/pnas.1217335110. http://www.ncbi.nlm.nih.gov/pubmed/23284173. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3549097.
[99] Skalsky S, Zhang Y, Alanis JA, Fonseka HA, Sanchez AM, Liu H, Parkinson P. Heterostructure and Q-factor engineering for low-threshold and persistent nanowire lasing. Light Sci Appl 2020;9:43. https://doi.org/10.1038/s41377- 020-0279-y. http://www.nature.com/articles/s41377-020-0279-y.
[100] Fu A, Gao H, Petrov P, Yang P. Widely tunable distributed Bragg reflectors integrated into nanowire waveguides. Nano Lett 2015;15:6909–13. https://doi.org/10.1021/acs.nanolett.5b02839. https://pubs.acs.org/doi/10.1021/acs.nanolett.5b02839.
[101] Sakai M, Inose Y, Ema K, Ohtsuki T, Sekiguchi H, Kikuchi A, Kishino K. Random laser action in GaN nanocolumns. Appl Phys Lett 2010;97:151109. https://doi.org/10.1063/1.3495993. http://aip.scitation.org/doi/10.1063/1.3495993.
[102] Li KH, Liu X, Wang Q, Zhao S, Mi Z. Ultralow-threshold electrically injected AlGaN nanowire ultraviolet lasers on Si operating at low temperature. Nat Nanotechnol 2015;10:140–4. https://doi.org/10.1038/nnano.2014.308. http://www.nature.com/articles/nnano.2014.308.
[103] Kim H, Lee W-JJ, Farrell AC, Morales JSD, Senanayake P, Prikhodko SV, Ochalski TJ, Huffaker DL. Monolithic InGaAs nanowire array lasers on silicon-on-insulator operating at room temperature. Nano Lett 2017;17:3465–70. https://doi.org/10.1021/acs.nanolett.7b00384. https://pubs.acs.org/doi/10.1021/acs.nanolett.7b00384. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.7b00384.
[104] Yokoyama H, Brorson SD. Rate equation analysis of microcavity lasers. J Appl Phys 1989;66:4801–5. https://doi.org/10.1063/1.343793. http://aip.scitation.org/doi/10.1063/1.343793.
[105] Björk G, Yamamoto Y. Analysis of semiconductor microcavity lasers using rate equations. IEEE J Quant Electron 1991;27:2386–96. https://doi.org/10.1109/3.100877.
[106] Khajavikhan M, Simic A, Katz M, Lee JH, Slutsky B, Mizrahi A, Lomakin V, Fainman Y. Thresholdless nanoscale coaxial lasers. Nature 2012;482:204–7. https://doi.org/10.1038/nature10840.
[107] Mayer B, Janker L, Rudolph D, Loitsch B, Kostenbader T, Abstreiter G, Koblmüller G, Finley JJ. Continuous wave lasing from individual GaAs-AlGaAs core-shell nanowires. Appl Phys Lett 2016;108:071107. https://doi.org/10.1063/1.4942506. http://scitation.aip.org/content/aip/journal/apl/108/7/10.1063/1.4942506.
[108] Zapf M, Ronning C, Röder R. High temperature limit of semiconductor nanowire lasers. Appl Phys Lett 2017;110:173103. https://doi.org/10.1063/1.4982629. http://aip.scitation.org/doi/10.1063/1.4982629.
[109] Jevtics D, McPhillimy J, Guilhabert B, Alanis JA, Tan HH, Jagadish C, Dawson MD, Hurtado A, Parkinson P, Strain MJ. Characterization, selection, and microassembly of nanowire laser systems. Nano Lett 2020;20:1862–8. https://doi.org/10.1021/acs.nanolett.9b05078. https://pubs.acs.org/doi/10.1021/acs.nanolett.9b05078.
[110] Mayer B, Rudolph D, Schnell J, Morkötter S, Winnerl J, Treu J, Müller K, Bracher G, Abstreiter G, Koblmüller G, Finley JJ. Lasing from individual

GaAs-AlGaAs core-shell nanowires up to room temperature. Nat Commun 2013; 4:2931. https://doi.org/10.1038/ncomms3931. http://www.nature.com/doifinder/10.1038/ncomms3931. http://www.nature.com/ncomms/2013/131205/ncomms3931/full/ncomms3931.html. http://www.nature.com/articles/ncomms3931.

[111] Govender K, Boyle D, O'Brien P, Binks D, West D, Coleman D. Room-temperature lasing observed from ZnO nanocolumns grown by aqueous solution deposition. Adv Mater 2002;14:1221–4. https://doi.org/10.1002/1521-4095(20020903)14:171221::AID- ADMA12213.0.CO;2-1.

[112] Johnson JC, Choi H-J, Knutsen KP, Schaller RD, Yang P, Saykally RJ. Single gallium nitride nanowire lasers. Nat Mater 2002;1:106–10. https://doi.org/10.1038/nmat728. http://apps.webofknowledge.com/full. http://www.nature.com/articles/nmat728.

[113] O'Carroll D, Lieberwirth I, Redmond G. Microcavity effects and optically pumped lasing in single conjugated polymer nanowires. Nat Nanotechnol 2007;2:180–4. https://doi.org/10.1038/nnano.2007.35. http://www.nature.com/articles/nnano.2007.35.

[114] Zhao YS, Peng A, Fu H, Ma Y, Yao J. Nanowire waveguides and ultraviolet lasers based on small organic molecules. Adv Mater 2008;20:1661–5. https://doi.org/10.1002/adma.200800123. http://doi.wiley.com/10.1002/adma.200800123.

[115] Hua B, Motohisa J, Kobayashi Y, Hara S, Fukui T. Single GaAs/GaAsP coaxial core-shell nanowire lasers. Nano Lett 2009;9:112–6. https://doi.org/10.1021/nl802636b. https://pubs.acs.org/doi/10.1021/nl802636b.

[116] Camposeo A, Benedetto FD, Stabile R, Neves AA, Cingolani R, Pisignano D. Laser emission from electrospun polymer nanofibers. Small 2009;5:562–6. https://doi.org/10.1002/smll.200801165. http://doi.wiley.com/10.1002/smll.200801165.

[117] Oulton RF, Sorger VJ, Zentgraf T, Ma RM, Gladden C, Dai L, Bartal G, Zhang X. Plasmon lasers at deep subwavelength scale. Nature 2009;461:629–32. https://doi.org/10.1038/nature08364.

[118] Xiao Y, Meng C, Wang P, Ye Y, Yu H, Wang S, Gu F, Dai L, Tong L. Single-nanowire single-mode laser. Nano Lett 2011;11:1122–6. https://doi.org/10.1021/nl1040308. https://pubs.acs.org/doi/10.1021/nl1040308.

[119] Geburt S, Thielmann A, Röder R, Borschel C, McDonnell A, Kozlik M, Kühnel J, Sunter KA, Capasso F, Ronning C. Low threshold room-temperature lasing of CdS nanowires. Nanotechnology 2012;23:365204. https://doi.org/10.1088/0957-4484/23/36/365204. https://iopscience.iop.org/article/10.1088/0957-4484/23/36/365204. http://stacks.iop.org/0957-4484/23/i=36/a=365204?key=crossref.3b26e193e669410451fcb1ac6.

[120] Liu X, Zhang Q, Yip JN, Xiong Q, Sum TC. Wavelength tunable single nanowire lasers based on surface plasmon polariton enhanced Burstein-Moss effect. Nano Lett 2013;13:5336–43. https://doi.org/10.1021/nl402836x. https://pubs.acs.org/doi/10.1021/nl402836x.

[121] Gao Q, Saxena D, Wang F, Fu L, Mokkapati S, Guo Y, Li L, Wong-Leung J, Caroff P, Tan HH, Jagadish C. Selective-area epitaxy of pure wurtzite InP nanowires: high quantum efficiency and room-temperature lasing. Nano Lett 2014;14:5206–11. https://doi.org/10.1021/nl5021409. https://pubs.acs.org/doi/10.1021/nl5021409.

[122] Zhang Q, Li G, Liu X, Qian F, Li Y, Sum TC, Lieber CM, Xiong Q. A room temperature low-threshold ultraviolet plasmonic nanolaser. Nat Commun 2014;5:4953. https://doi.org/10.1038/ncomms5953. http://www.nature.com/articles/ncomms5953.

[123] Sidiropoulos TPH, Röder R, Geburt S, Hess O, Maier SA, Ronning C, Oulton RF. Ultrafast plasmonic nanowire lasers near the surface plasmon frequency. Nat Phys

2014;10:870–6. https://doi.org/10.1038/nphys3103. http://www.nature.com/articles/nphys3103.

[124] Tatebayashi J, Kako S, Ho J, Ota Y, Iwamoto S, Arakawa Y. Room-temperature lasing in a single nanowire with quantum dots. Nat Photonics 2015;9:501–5. https://doi.org/10.1038/nphoton.2015.111. http://www.nature.com/doifinder/10.1038/nphoton.2015.111. http://www.nature.com/articles/nphoton.2015.111.

[125] Fu Y, Zhu H, Schrader AW, Liang D, Ding Q, Joshi P, Hwang L, Zhu XY, Jin S. Nanowire lasers of formamidinium lead halide perovskites and their stabilized alloys with improved stability. Nano Lett 2016;16:1000–8. https://doi.org/10.1021/acs.nanolett.5b04053. https://pubs.acs.org/doi/10.1021/acs.nanolett.5b04053.

[126] Burgess T, Saxena D, Mokkapati S, Li Z, Hall CR, Davis JA, Wang Y, Smith LM, Fu L, Caroff P, Tan HH, Jagadish C. Doping-enhanced radiative efficiency enables lasing in unpassivated GaAs nanowires. Nat Commun 2016;7:11927. https://doi.org/10.1038/ncomms11927. http://www.nature.com/doifinder/10.1038/ncomms11927.

[127] Bermúdez-Ureña E, Tutuncuoglu G, Cuerda J, Smith CLC, Bravo-Abad J, Bozhevolnyi SI, Fontcuberta I Morral A, García-Vidal FJ, Quidant R, FontcubertaMorral A, García-Vidal FJ, Quidant R. Plasmonic waveguide- integrated nanowire laser. Nano Lett 2017;17:747–54. https://doi.org/10.1021/acs.nanolett.6b03879. https://pubs.acs.org/doi/10.1021/acs.nanolett.6b03879.

[128] Chen HL, Himwas C, Scaccabarozzi A, Rale P, Oehler F, Lemaître A, Lombez L, Guillemoles JF, Tchernycheva M, Harmand JC, Cattoni A, Collin S. Determination of n-type doping level in single GaAs nanowires by cathodoluminescence. Nano Lett 2017;17:6667–75. https://doi.org/10.1021/acs.nanolett.7b02620.

[129] Yu Z, Wu Y, Xiao L, Chen J, Liao Q, Yao J, Fu H. Organic phosphorescence nanowire lasers. J Am Chem Soc 2017;139:6376–81. https://doi.org/10.1021/jacs.7b01574. https://pubs.acs.org/doi/10.1021/jacs.7b01574.

[130] Wang X, Li ZZ, Zhuo MP, Wu Y, Chen S, Yao J, Fu H. Tunable near-infrared organic nanowire nanolasers. Adv Funct Mater 2017;27:1703470. https://doi.org/10.1002/adfm.201703470. http://doi.wiley.com/10.1002/adfm.201703470.

[131] Ren D, Scofield AC, Farrell AC, Rong Z, Haddad MA, Laghumavarapu RB, Liang B, Huffaker DL. Exploring time-resolved photoluminescence for nanowires using a three- dimensional computational transient model. Nanoscale 2018;10:7792–802. https://doi.org/10.1039/C8NR01908H. http://xlink.rsc.org/?DOI=C8NR01908H.

[132] Huang L, Gao Q, Sun LD, Dong H, Shi S, Cai T, Liao Q, Yan CH. Composition-graded cesium lead halide perovskite nanowires with tunable dual-color lasing performance. Adv Mater 2018;30:1800596. https://doi.org/10.1002/adma.201800596. http://doi.wiley.com/10.1002/adma.201800596.

[133] Wu Z, Chen J, Mi Y, Sui X, Zhang S, Du W, Wang R, Shi J, Wu X, Qiu X, Qin Z, Zhang Q, Liu X. All-inorganic CsPbBr 3 nanowire based plasmonic lasers. Adv Opt Mater 2018;6:1800674. https://doi.org/10.1002/adom.201800674. http://doi.wiley.com/10.1002/adom.201800674.

[134] Alanis JA, Lysevych M, Burgess T, Saxena D, Mokkapati S, Skalsky S, Tang X, Mitchell P, Walton AS, Tan HH, Jagadish C, Parkinson P. Optical study of p-doping in GaAs nanowires for low-threshold and high-yield lasing. Nano Lett 2019;19:362–8. https://doi.org/10.1021/acs.nanolett.8b04048. http://pubs.acs.org/doi/10.1021/acs.nanolett.8b04048.

[135] Sumikura H, Zhang G, Takiguchi M, Takemura N, Shinya A, Gotoh H, Notomi M. Mid-infrared lasing of single wurtzite InAs nanowire. Nano Lett 2019;19:8059–65. https://doi.org/10.1021/acs.nanolett.9b03249. https://pubs.acs.org/doi/10.1021/acs.nanolett.9b03249.

[136] Sergent S, Takiguchi M, Tsuchizawa T, Taniyama H, Notomi M. Low-threshold lasing up to 360 K in all-dielectric subwavelength-nanowire nanocavities. ACS Photonics 2020:0c00166. https://doi.org/10.1021/acsphotonics.0c00166. acsphotonics, https://pubs.acs.org/doi/10.1021/acsphotonics.0c00166.

[137] Seifert W, Borgström M, Deppert K, Dick KA, Johansson J, Larsson MW, Mårtensson T, Sköld N, Patrik C, Svensson T, Wacaser BA, Reine Wallenberg L, Samuelson L. Growth of one-dimensional nanostructures in MOVPE. J Cryst Growth 2004;272:211–20. https://doi.org/10.1016/j.jcrysgro.2004.09.023. https://linkinghub.elsevier.com/retrieve/pii/S0022024804011169.

[138] Persson AI, Larsson MW, Stenström S, Ohlsson BJ, Samuelson L, Wallenberg LR. Solid-phase diffusion mechanism for GaAs nanowire growth. Nat Mater 2004;3:677–81. https://doi.org/10.1038/nmat1220. http://www.nature.com/articles/nmat1220.

[139] Greene LE, Law M, Goldberger J, Kim F, Johnson JC, Zhang Y, Saykally RJ, Yang P. Low-temperature wafer-scale production of ZnO nanowire arrays. Angew Chem Int Ed 2003;42:3031–4. https://doi.org/10.1002/anie.200351461. http://doi.wiley.com/10.1002/anie.200351461.

[140] Wang Y, Meng G, Zhang L, Liang C, Zhang J. Catalytic growth of large-scale single-crystal CdS nanowires by physical evaporation and their photoluminescence. Chem Mater 2002;14:1773–7. https://doi.org/10.1021/cm0115564. https://pubs.acs.org/doi/abs/10.1021/cm0115564.

[141] Givargizov EI. Fundamental aspects of VLS growth. J Cryst Growth 1975;31:20–30. https://doi.org/10.1016/0022-0248(75)90105-0. https://www.sciencedirect.com/science/article/pii/0022024875901050.

[142] Harmand JC, Patriarche G, Péré-Laperne N, Mérat-Combes M-N, Travers L, Glas F. Analysis of vapor-liquid-solid mechanism in Au-assisted GaAs nanowire growth. Appl Phys Lett 2005;87:203101. https://doi.org/10.1063/1.2128487. http://aip.scitation.org/doi/10.1063/1.2128487.

[143] Noborisaka J, Motohisa J, Fukui T. Catalyst-free growth of GaAs nanowires by selective-area metalorganic vapor-phase epitaxy. Appl Phys Lett 2005;86:1–3. https://doi.org/10.1063/1.1935038. http://aip.scitation.org/doi/10.1063/1.1935038.

[144] Kanungo PD, Schmid H, Björk MT, Gignac LM, Breslin C, Bruley J, Bessire CD, Riel H. Selective area growth of III-V nanowires and their heterostructures on silicon in a nanotube template: towards monolithic integration of nano-devices. Nanotechnology 2013;24:225304. https://doi.org/10.1088/0957-4484/24/22/225304. https://iopscience.iop.org/article/10.1088/0957-4484/24/22/225304.

[145] Tomioka K, Kobayashi Y, Motohisa J, Hara S, Fukui T. Selective-area growth of vertically aligned GaAs and GaAs/AlGaAs core–shell nanowires on Si(111) substrate. Nanotechnology 2009;20:145302. https://doi.org/10.1088/0957-4484/20/14/145302. https://iopscience.iop.org/article/10.1088/0957-4484/20/14/145302. http://stacks.iop.org/0957-4484/20/i=14/a=145302?key=crossref.84f864f89e24fa54c1915b7d6.

[146] Ikejiri K, Sato T, Yoshida H, Hiruma K, Motohisa J, Hara S, Fukui T. Growth characteristics of GaAs nanowires obtained by selective area metal-organic vapour-phase epitaxy. Nanotechnology 2008;19:265604. https://doi.org/10.1088/0957-4484/19/26/265604. https://iopscience.iop.org/article/10.1088/0957-4484/19/26/265604.

[147] Schmidt V, Wittemann JV, Senz S, Gósele U. Silicon nanowires: a review on aspects of their growth and their electrical properties. Adv Mater 2009;21(25–26):2681–702. https://doi.org/10.1002/adma.200803754. http://doi.wiley.com/10.1002/adma.200803754.

[148] Dick KA. A review of nanowire growth promoted by alloys and non-alloying elements with emphasis on Au-assisted III-V nanowires. Prog Cryst Growth Charact

Mater 2008;54(3–4):138–73. https://doi.org/10.1016/j.pcrysgrow.2008.09.001. https://www.sciencedirect.com/science/article/pii/S0960897408000181.
[149] Dasgupta NP, Sun J, Liu C, Brittman S, Andrews SC, Lim J, Gao H, Yan R, Yang P. 25th anniversary article: semiconductor nanowires - synthesis, characterization, and applications. Adv Mater 2014;26:2137–83. https://doi.org/10.1002/adma.201305929. http://www.ncbi.nlm.nih.gov/pubmed/24604701. http://doi.wiley.com/10.1002/adma.201305929.
[150] Mårtensson T, Svensson CPT, Wacaser BA, Larsson MW, Seifert W, Deppert K, Gustafsson A, Wallenberg LR, Samuelson L. Epitaxial III-V nanowires on silicon. Nano Lett 2004;4:1987–90. https://doi.org/10.1021/nl0487267. https://pubs.acs.org/doi/10.1021/nl0487267.
[151] Knelangen M, Consonni V, Trampert A, Riechert H. Insitu analysis of strain relaxation during catalyst-free nucleation and growth of GaN nanowires. Nanotechnology 2010;21:245705. https://doi.org/10.1088/0957-4484/21/24/245705. http://stacks.iop.org/0957-4484/21/i=24/a=245705?key=crossref.37b5a3e971f0c4ec90bab835c.
[152] De La Mata M, Magén C, Caroff P, Arbiol J. Atomic scale strain relaxation in axial semiconductor III-V nanowire heterostructures. Nano Lett 2014;14:6614–20. https://doi.org/10.1021/nl503273j. https://pubs.acs.org/doi/10.1021/nl503273j.
[153] Wei M, Zhi D, MacManus-Driscoll JL. Self-catalysed growth of zinc oxide nanowires. Nanotechnology 2005;16:1364–8. https://doi.org/10.1088/0957-4484/16/8/064. https://iopscience.iop.org/article/10.1088/0957-4484/16/8Z064.
[154] Colombo C, Spirkoska D, Frimmer M, Abstreiter G, Fontcuberta I Morral A. Ga-assisted catalyst-free growth mechanism of GaAs nanowires by molecular beam epitaxy. Phys Rev B Condens Matter 2008;77:155326. https://doi.org/10.1103/PhysRevB.77.155326. https://link.aps.org/doi/10.1103/PhysRevB.77.155326.
[155] Parkinson P, Joyce HJ, Gao Q, Tan HH, Zhang X, Zou J, Jagadish C, Herz LM, Johnston MB. Carrier lifetime and mobility enhancement in nearly defect-free core-shell nanowires measured using time-resolved terahertz spectroscopy. Nano Lett 2009;9:3349–53. https://doi.org/10.1021/nl9016336. http://pubs.acs.org/doi/abs/10.1021/nl9016336.
[156] Joyce HJ, Gao Q, Tan HH, Jagadish C, Kim Y, Zhang X, Guo Y, Zou J. Twin-free uniform epitaxial GaAs nanowires grown by a two-temperature process. Nano Lett 2007;7:921–6. https://doi.org/10.1021/nl062755v.
[157] Jiang N, Gao Q, Parkinson P, Wong-Leung J, Mokkapati S, Breuer S, Tan HH, Zheng CL, Etheridge J, Jagadish C. Enhanced minority carrier lifetimes in GaAs/AlGaAs core-shell nanowires through shell growth optimization. Nano Lett 2013;13:5135–40. https://doi.org/10.1021/nl4023385.
[158] Mårtensson T, Carlberg P, Borgström M, Montelius L, Seifert W, Samuelson L. Nanowire arrays defined by nanoimprint lithography. Nano Lett 2004;4:699–702. https://doi.org/10.1021/nl035100s. https://pubs.acs.org/doi/abs/10.1021/nl035100s.
[159] Yamaguchi KI, Okamoto K, Imai T. Selective epitaxial growth of gaas by metalorganic chemical vapor deposition. Jpn J Appl Phys 1985;24:1666–71. https://doi.org/10.1143/JJAP.24.1666. https://iopscience.iop.org/article/10.1143/JJAP.24.1666.
[160] Hersee SD, Sun X, Wang X. The controlled growth of GaN nanowires. Nano Lett 2006;6:1808–11. https://doi.org/10.1021/nl060553t. https://pubs.acs.org/doi/10.1021/nl060553t.
[161] Metaferia W, Sivakumar S, Persson AR, Geijselaers I, Wallenberg LR, Deppert K, Samuelson L, Magnusson MH. n-type doping and morphology of GaAs nanowires in Aerotaxy. Nanotechnology 2018;29:285601. https://doi.org/10.1088/1361-6528/aabec0. https://iopscience.iop.org/article/10.1088/1361-6528/aabec0.

[162] Chen Y, Bagnall D, Yao T. ZnO as a novel photonic material for the UV region. Mater Sci Eng, B 2000;75:190–8. https://doi.org/10.1016/S0921-5107(00)00372-X. https://www.sciencedirect.com/science/article/pii/S092151070000372X.
[163] Huang K, Yang S, Tong L. Modeling of evanescent coupling between two parallel optical nanowires. Appl Opt 2007;46:1429–34. https://doi.org/10.1364/AO.46.001429. https://www.osapublishing.org/abstract.cfm?URI=ao-46-9-1429.
[164] Vayssieres L, Keis K, Hagfeldt A, Lindquist SE. Three-dimensional array of highly oriented crystalline ZnO microtubes. Chem Mater 2001;13:4395–8. https://doi.org/10.1021/cm011160s. https://pubs.acs.org/doi/10.1021/cm011160s.
[165] Zhang Y, Russo RE, Mao SS. Quantum efficiency of ZnO nanowire nanolasers. Appl Phys Lett 2005;87:043106. https://doi.org/10.1063/1.2001754. http://aip.scitation.org/doi/10.1063/1-2001754.
[166] Chou YH, Wu YM, Hong KB, Chou BT, Shih JH, Chung YC, Chen PY, Lin TR, Lin CC, Lin SD, Lu TC. High-operation-temperature plasmonic nanolasers on single-crystalline aluminum. Nano Lett 2016;16:3179–86. https://doi.org/10.1021/acs.nanolett.6b00537. https://pubs.acs.org/doi/10.1021/acs.nanolett.6b00537.
[167] Look DC, Claflin B. P-type doping and devices based on ZnO. Phys Status Solidi (B) 2004;241:624–30. https://doi.org/10.1002/pssb.200304271. http://doi.wiley.com/10.1002/pssb.200304271.
[168] Nobel Media AB. The Nobel prize in physics 2014. 2020. https://www.nobelprize.org/prizes/physics/2014/summary/.
[169] Amano H, Sawaki N, Akasaki I, Toyoda Y. Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer. Appl Phys Lett 1986;48:353–5. https://doi.org/10.1063/1.96549. http://aip.scitation.org/doi/10.1063/1.96549.
[170] Nakamura S, Mukai T, Senoh M. Candela-class high-brightness InGaN/AlGaN double-heterostructure blue-light-emitting diodes. Appl Phys Lett 1994;64:1687–9. https://doi.org/10.1063/1.111832. http://aip.scitation.org/doi/10.1063/1.111832.
[171] Muth JF, Lee JH, Shmagin IK, Kolbas RM, Casey HC, Keller BP, Mishra UK, DenBaars SP. Absorption coefficient, energy gap, exciton binding energy, and recombination lifetime of GaN obtained from transmission measurements. Appl Phys Lett 1997;71:2572–4. https://doi.org/10.1063/1.120191. http://aip.scitation.org/doi/10.1063/1.120191.
[172] Singh R, Doppalapudi D, Moustakas TD, Romano LT. Phase separation in InGaN thick films and formation of InGaN/GaN double heterostructures in the entire alloy composition. Appl Phys Lett 1997;70:1089–91. https://doi.org/10.1063/1.118493. http://aip.scitation.org/doi/10.1063/1.118493.
[173] Mi Z, Zhao S, Woo SY, Bugnet M, Djavid M, Liu X, Kang J, Kong X, Ji W, Guo H, Liu Z, Botton GA. Molecular beam epitaxial growth and characterization of Al(Ga)N nanowire deep ultraviolet light emitting diodes and lasers. J Phys Appl Phys 2016;49:364006. https://doi.org/10.1088/0022-3727/49/36/364006. http://stacks.iop.org/0022-3727/49/i=36/a=364006?key=crossref.c766ac74e3d7790cc603c7749.
[174] Gradečak S, Qian F, Li Y, Park H-G, Lieber CM. GaN nanowire lasers with low lasing thresholds. Appl Phys Lett 2005;87:173111. https://doi.org/10.1063/1.2115087. http://aip.scitation.org/doi/10.1063/1.2115087.
[175] Parkinson P, Dodson C, Joyce HJ, Bertness KA, Sanford NA, Herz LM, Johnston MB. Noncontact measurement of charge carrier lifetime and mobility in GaN nanowires. Nano Lett 2012;12:4600–4. https://doi.org/10.1021/nl301898m.
[176] Li C, Wright JB, Liu S, Lu P, Figiel JJ, Leung B, Chow WW, Brener I, Koleske DD, Luk TS, Feezell DF, Brueck SR, Wang GT. Nonpolar InGaN/GaN core-shell single

nanowire lasers. Nano Lett 2017;17:1049–55. https://doi.org/10.1021/acs.nanolett.6b04483. https://pubs.acs.org/doi/10.1021/acs.nanolett.6b04483. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.6b04483.

[177] Behzadirad M, Nami M, Wostbrock N, Zamani Kouhpanji MR, Feezell DF, Brueck SR, Busani T. Scalable top-down approach tailored by interferometric lithography to achieve large-area single-mode GaN nanowire laser arrays on sapphire substrate. ACS Nano 2018;12:2373–80. https://doi.org/10.1021/acsnano.7b07653. https://pubs.acs.org/doi/10.1021/acsnano.7b07653.

[178] Voigt J, Spiegelberg F, Senoner M. Band parameters of CdS and CdSe single crystals determined from optical exciton spectra. Phys Status Solidi (B) 1979;91:189–99. https://doi.org/10.1002/pssb.2220910120. http://doi.wiley.com/10.1002/pssb.2220910120.

[179] Röder R, Wille M, Geburt S, Rensberg J, Zhang M, Lu JG, Capasso F, Buschlinger R, Peschel U, Ronning C. Continuous wave nanowire lasing. Nano Lett 2013;13:3602–6. https://doi.org/10.1021/nl401355b. https://pubs.acs.org/doi/10.1021/nl401355b.

[180] Fischer T, Bille J. Recombination processes in highly excited CdS. J Appl Phys 1974;45:3937–42. https://doi.org/10.1063/1.1663891. http://aip.scitation.org/doi/10.1063/1.1663891.

[181] Pan A, Zhou W, Leong ESP, Liu R, Chin AH, Zou B, Ning CZ. Continuous alloy-composition spatial grading and superbroad wavelength-tunable nanowire lasers on a single chip. Nano Lett 2009;9:784–8. https://doi.org/10.1021/n1803456k. https://pubs.acs.org/doi/10.1021/nl803456k.

[182] Guo YN, Xu HY, Auchterlonie GJ, Burgess T, Joyce HJ, Gao Q, Tan HH, Jagadish C, Shu HB, Chen XS, Lu W, Kim Y, Zou J. Phase separation induced by Au catalysts in ternary InGaAs nanowires. Nano Lett 2013;13:643–50. https://doi.org/10.1021/nl304237b. http://pubs.acs.org/doi/10.1021/nl304237b.

[183] Zhang Q, Liu H, Guo P, Li D, Fan P, Zheng W, Zhu X, Jiang Y, Zhou H, Hu W, Zhuang X, Liu H, Duan X, Pan A. Vapor growth and interfacial carrier dynamics of high-quality CdS-CdSSe-CdS axial nanowire heterostructures. Nano Energy 2017;32:28–35. https://doi.org/10.1016/j.nanoen.2016.12.014. https://www.sciencedirect.com/science/article/pii/S221128551630578X.

[184] Steighner MS, Snedeker LP, Boyce BL, Gall K, Miller DC, Muhlstein CL. Dependence on diameter and growth direction of apparent strain to failure of Si nanowires. J Appl Phys 2011;109:033503. https://doi.org/10.1063/1.3537658. http://aip.scitation.org/doi/10.1063/1.3537658.

[185] Wang YB, Wang LF, Joyce HJ, Gao Q, Liao XZ, Mai YW, Tan HH, Zou J, Ringer SP, Gao HJ, Jagadish C. Super deformability and young's modulus of gaAs nanowires. Adv Mater 2011;23:1356–60. https://doi.org/10.1002/adma.201004122. http://doi.wiley.com/10.1002/adma.201004122.

[186] Wang Y, Joyce HJ, Gao Q, Liao X, Tan HH, Zou J, Ringer SP, Shan Z, Jagadish C. Self-healing of fractured GaAs nanowires. Nano Lett 2011;11:1546–9. https://doi.org/10.1021/nl104330h. https://pubs.acs.org/doi/10.1021/nl104330h.

[187] Fu Q, Zhang ZY, Kou L, Wu P, Han X, Zhu X, Gao J, Xu J, Zhao Q, Guo W, Yu D. Linear strain-gradient effect on the energy bandgap in bent CdS nanowires. Nano Res 2011;4:308–14. https://doi.org/10.1007/s12274-010-0085-6. http://link.springer.com/10.1007/s12274-010-0085-6.

[188] Yang W, Ma Y, Wang Y, Meng C, Wu X, Ye Y, Dai L, Tong L, Liu X, Yang Q. Bending effects on lasing action of semiconductor nanowires. Opt Express 2013;21:2024. https://doi.org/10.1364/oe.21.002024. https://www.osapublishing.org/oe/abstract.cfm?uri=oe-21-2-2024.

[189] Parkinson P, Lloyd-Hughes J, Gao Q, Tan HH, Jagadish C, Johnston MB, Herz LM. Transient terahertz conductivity of GaAs nanowires. Nano Lett 2007;7:2162–5. https://doi.org/10.1021/nl071162x. https://pubs.acs.org/doi/10.1021/nl071162x.

[190] Joyce HJ, Docherty CJ, Gao Q, Tan HH, Jagadish C, Lloyd-Hughes J, Herz LM, Johnston MB. Electronic properties of GaAs, InAs and InP nanowires studied by terahertz spectroscopy. Nanotechnology 2013;24:214006. https://doi.org/10.1088/0957-4484/24/21/214006. http://stacks.iop.org/0957-4484/24/i=21/a=214006?key=crossref.3ddd730e84f137f4d4de09739.

[191] Jiang N, Parkinson P, Gao Q, Breuer S, Tan HH, Wong-Leung J, Jagadish C. Long minority carrier lifetime in Au-catalyzed GaAs/Al_xGa_1-x as core-shell nanowires. Appl Phys Lett 2012;101:023111. https://doi.org/10.1063/1.4735002. http://aip.scitation.org/doi/10.1063/1.4735002.

[192] Yuan X, Liu K, Skalsky S, Parkinson P, Fang L, He J, Tan HH, Jagadish C, Tan HH, Jagadish C. Carrier dynamics and recombination mechanisms in InP twinning superlattice nanowires. Opt Express 2020;28:16795. https://doi.org/10.1364/OE.388518. https://www.osapublishing.org/abstract.cfm?URI=oe-28-11-16795.

[193] Chen S, Jansson M, Stehr JE, Huang Y, Ishikawa F, Chen WM, Buyanova IA. Dilute nitride nanowire lasers based on a GaAs/GaNAs core/shell structure. Nano Lett 2017;17:1775–81. https://doi.org/10.1021/acs.nanolett.6b05097. http://pubs.acs.org/doi/abs/10.1021/acs.nanolett.6b05097.

[194] Balaghi L, Bussone G, Grifone R, Hübner R, Grenzer J, Ghorbani-Asl M, Krasheninnikov AV, Schneider H, Helm M, Dimakis E. Widely tunable GaAs bandgap via strain engineering in core/shell nanowires with large lattice mismatch. Nat Commun 2019;10. https://doi.org/10.1038/s41467-019-10654-7.

[195] Zhang Y, Davis G, Fonseka HA, Velichko A, Gustafsson A, Godde T, Saxena D, Aagesen M, Parkinson PW, Gott JA, Huo S, Sanchez AM, Mowbray DJ, Liu H. Highly strained III-V-V coaxial nanowire quantum wells with strong carrier confinement. ACS Nano 2019;13:5931–8. https://doi.org/10.1021/acsnano.9b01775. http://pubs.acs.org/doi/10.1021/acsnano.9b01775.

[196] Bissinger J, Ruhstorfer D, Stettner T, Koblmüller G, Finley JJ. Optimized waveguide coupling of an integrated III-V nanowire laser on silicon. J Appl Phys 2019;125:243102. https://doi.org/10.1063/1.5097405. http://aip.scitation.org/doi/10.1063/1.5097405.

[197] Ho J, Tatebayashi J, Sergent S, Fong CF, Iwamoto S, Arakawa Y. Low-threshold near-infrared GaAs-AlGaAs core-shell nanowire plasmon laser. ACS Photonics 2015;2:165–71. https://doi.org/10.1021/ph5003945. http://pubs.acs.org/doi/abs/10.1021/ph5003945.

[198] Mauthe S, Schmid H, Moselund KE, Trivino NV, Sousa M, Staudinger P, Baumgartner Y, Tiwari P, Stoferle T, Caimi D, Scherrer M. Monolithic integration of III-V microdisk lasers on silicon. In: International conference on optical MEMS and Nanophotonics, volume 2019-July. IEEE; 2019. p. 32–3. https://doi.org/10.1109/OMN.2019.8925128. https://ieeexplore.ieee.org/document/8925128/.

[199] Kojima A, Teshima K, Shirai Y, Miyasaka T. Organometal halide perovskites as visible-light sensitizers for photovoltaic cells. J Am Chem Soc 2009;131:6050–1. https://doi.org/10.1021/ja809598r. https://pubs.acs.org/doi/10.1021/ja809598r.

[200] Lee MM, Teuscher J, Miyasaka T, Murakami TN, Snaith HJ. Efficient hybrid solar cells based on meso-superstructured organometal halide perovskites. Science 2012;338:643–7. https://doi.org/10.1126/science.1228604. https://www.sciencemag.org/lookup/doi/10.1126/science.1228604.

[201] Zhao Y, Zhu K. Organic-inorganic hybrid lead halide perovskites for optoelectronic and electronic applications. Chem Soc Rev 2016;45:655–89. https://doi.org/10.1039/c4cs00458b. http://xlink.rsc.org/?D0I=C4CS00458B.

[202] Horváth E, Spina M, Szekrényes Z, Kamarás K, Gaal R, Gachet D, Forró L. Nanowires of Methylammonium Lead Iodide ($CH_3NH_3PbI_3$) prepared by low temperature solution- mediated crystallization. Nano Lett 2014;14:6761–6. https://doi.org/10.1021/nl5020684. https://pubs.acs.org/doi/10.1021/nl5020684.

[203] Zhu H, Fu Y, Meng F, Wu X, Gong Z, Ding Q, Gustafsson MV, Trinh MT, Jin S, Zhu X-YY. Lead halide perovskite nanowire lasers with low lasing thresholds and high quality factors. Nat Mater 2015;14:636–42. https://doi.org/10.1038/nmat4271. http://www.nature.com/doifinder/10.1038/nmat4271. http://www.nature.com/articles/nmat4271.

[204] Xing J, Liu XF, Zhang Q, Ha ST, Yuan YW, Shen C, Sum TC, Xiong Q. Vapor phase synthesis of organometal halide perovskite nanowires for tunable room- temperature nanolasers. Nano Lett 2015;15:4571–7. https://doi.org/10.1021/acs.nanolett.5b01166. https://pubs.acs.org/doi/10.1021/acs.nanolett.5b01166.

[205] Zhang D, Yang Y, Bekenstein Y, Yu Y, Gibson NA, Wong AB, Eaton SW, Kornienko N, Kong Q, Lai M, Alivisatos AP, Leone SR, Yang P. Synthesis of composition tunable and highly luminescent cesium lead halide nanowires through anion-exchange reactions. J Am Chem Soc 2016;138:7236–9. https://doi.org/10.1021/jacs.6b03134. https://pubs.acs.org/doi/10.1021/jacs.6b03134.

[206] He X, Liu P, Wu S, Liao Q, Yao J, Fu H. Multi-color perovskite nanowire lasers through kinetically controlled solution growth followed by gas-phase halide exchange. J Mater Chem C 2017;5:12707–13. https://doi.org/10.1039/c7tc03939e. http://xlink.rsc.org/?DOI=C7TC03939E.

[207] Liu P, He X, Ren J, Liao Q, Yao J, Fu H. Organic-inorganic hybrid perovskite nanowire laser arrays. ACS Nano 2017;11:5766–73. https://doi.org/10.1021/acsnano.7b01351. https://pubs.acs.org/doi/10.1021/acsnano.7b01351.

[208] Wang X, Shoaib M, Wang X, Zhang X, He M, Luo Z, Zheng W, Li H, Yang T, Zhu X, Ma L, Pan A. High-quality in-plane aligned CsPbX3 perovskite nanowire lasers with composition-dependent strong exciton-photon coupling. ACS Nano 2018;12:6170–8. https://doi.org/10.1021/acsnano.8b02793. https://pubs.acs.org/doi/10.1021/acsnano.8b02793.

[209] Saba M, Cadelano M, Marongiu D, Chen F, Sarritzu V, Sestu N, Figus C, Aresti M, Piras R, Geddo Lehmann A, Cannas C, Musinu A, Quochi F, Mura A, Bongiovanni G. Correlated electron-hole plasma in organometal perovskites. Nat Commun 2014;5:5049. https://doi.org/10.1038/ncomms6049. http://www.nature.com/articles/ncomms6049.

[210] Du W, Zhang S, Shi J, Chen J, Wu Z, Mi Y, Liu Z, Li Y, Sui X, Wang R, Qiu X, Wu T, Xiao Y, Zhang Q, Liu X. Strong exciton-photon coupling and lasing behavior in all-inorganic $CsPbBr_3$ micro/nanowire fabry-pérot cavity. ACS Photonics 2018;5:2051–9. https://doi.org/10.1021/acsphotonics.7b01593. https://pubs.acs.org/doi/10.1021/acsphotonics.7b01593.

[211] Evans TJ, Schlaus A, Fu Y, Zhong X, Atallah TL, Spencer MS, Brus LE, Jin S, Zhu XY. Continuous-wave lasing in cesium lead bromide perovskite nanowires. Adv Opt Mater 2018;6:1700982. https://doi.org/10.1002/adom.201700982. http://doi.wiley.com/10.1002/adom.201700982.

[212] Schlaus AP, Spencer MS, Miyata K, Liu F, Wang X, Datta I, Lipson M, Pan A, Zhu XY. How lasing happens in $CsPbBr_3$ perovskite nanowires. Nat Commun 2019;10:265. https://doi.org/10.1038/s41467-018-07972-7. http://www.nature.com/articles/s41467-018-07972-7.

[213] Zhang N, Fan Y, Wang K, Gu Z, Wang Y, Ge L, Xiao S, Song Q. All-optical control of lead halide perovskite microlasers. Nat Commun 2019;10:1770. https://doi.org/10.1038/s41467-019-09876-6. http://www.nature.com/articles/s41467-019-09876-6.

[214] Chénais S, Forget S. Recent advances in solid-state organic lasers. Polymer International 2012;61(3):390–406. https://doi.org/10.1002/pi.3173. http://doi.wiley.com/10.1002/pi.3173.
[215] Baldo MA, Holmes RJ, Forrest SR. Prospects for electrically pumped organic lasers. Phys Rev B Condens Matter 2002;66:353211–3532116. https://doi.org/10.1103/PhysRevB.66.035321. https://link.aps.org/doi/10.1103/PhysRevB.66.035321.
[216] Yuan X, Saxena D, Caroff P, Wang F, Lockrey M, Mokkapati S, Tan HH, Jagadish C. Strong amplified spontaneous emission from high quality $GaAs_{1-x}Sb_x$ single quantum well nanowires. J Phys Chem C 2017;121:8636–44. https://doi.org/10.1021/acs.jpcc.7b00744. http://pubs.acs.org/doi/abs/10.1021/acs.jpcc.7b00744. https://pubs.acs.org/doi/10.1021/acs.jpcc.7b00744.
[217] Kempa TJ, Kim SK, Day RW, Park HG, Nocera DG, Lieber CM. Facet-selective growth on nanowires yields multi-component nanostructures and photonic devices. J Am Chem Soc 2013;135:18354–7. https://doi.org/10.1021/ja411050r. https://pubs.acs.org/doi/10.1021/ja411050r.
[218] Wen CY, Reuter MC, Bruley J, Tersoff J, Kodambaka S, Stach EA, Ross FM. Formation of compositionally abrupt axial heterojunctions in silicon-germanium nanowires. Science 2009;326:1247–50. https://doi.org/10.1126/science.1178606. http://www.ncbi.nlm.nih.gov/pubmed/19965471.
[219] Dick KA, Bolinsson J, Borg BM, Johansson J. Controlling the abruptness of axial heterojunctions in III-V nanowires: beyond the reservoir effect. Nano Lett 2012;12:3200–6. https://doi.org/10.1021/nl301185x. https://pubs.acs.org/doi/10.1021/nl301185x.
[220] Claudon J, Bleuse J, Malik NS, Bazin M, Jaffrennou P, Gregersen N, Sauvan C, Lalanne P, Gérard J-M. A highly efficient single-photon source based on a quantum dot in a photonic nanowire. Nat Photonics 2010;4:174–7. https://doi.org/10.1038/nphoton.2009.287x. http://www.nature.com/articles/nphoton.2009.287x.
[221] Xu W-Z, Ren F-F, Jevtics D, Hurtado A, Li L, Gao Q, Ye J, Wang F, Guilhabert B, Fu L, Lu H, Zhang R, Tan HH, Dawson MD, Jagadish C. Vertically emitting indium phosphide nanowire lasers. Nano Lett 2018;18:3414–20. https://doi.org/10.1021/acs.nanolett.8b00334. http://pubs.acs.org/doi/10.1021/acs.nanolett.8b00334.
[222] Sergent S, Takiguchi M, Taniyama H, Shinya A, Kuramochi E, Notomi M. Design of nanowire-induced nanocavities in grooved 1D and 2D SiN photonic crystals for the ultraviolet and visible ranges. Opt Express 2016;24:26792. https://doi.org/10.1364/oe.24.026792. https://www.osapublishing.org/abstract.cfm?URI=oe-24-23-26792.
[223] Sergent S, Takiguchi M, Tsuchizawa T, Taniyama H, Kuramochi E, Notomi M. Nanomanipulating and tuning ultraviolet ZnO- nanowire-induced photonic crystal nanocavities. ACS Photonics 2017;4:1040–7. https://doi.org/10.1021/acsphotonics.7b00116.
[224] Zheludev NI, Prosvirnin SL, Papasimakis N, Fedotov VA. Lasing spaser. Nat Photonics 2008;2:351–4. https://doi.org/10.1038/nphoton.2008.82. http://www.nature.com/articles/nphoton.2008.82.
[225] Lu Y-J, Kim J, Chen H-Y, Wu C, Dabidian N, Sanders CE, Wang C-Y, Lu M-Y, Li B-H, Qiu X, Chang W-H, Chen L-J, Shvets G, Shih C-K, Gwo S. Plasmonic nanolaser using epitaxially grown silver film. Science 2012;337:450–3. https://doi.org/10.1126/science.1223504. http://www.sciencemag.org/cgi/doi/10.1126/science.1223504.
[226] Zhu T, Zhou Y, Lou Y, Ye H, Qiu M, Ruan Z, Fan S. Plasmonic computing of spatial differentiation. Nat Commun 2017;8:15391. https://doi.org/10.1038/ncomms15391. http://www.nature.com/articles/ncomms15391.
[227] Parkinson P, Alanis JA, Skalsky S, Zhang Y, Liu H, Lysevych M, Tan HH, Jagadish C. A needle in a needlestack: exploiting functional inhomogeneity for optimized nanowire

lasing. In: Proceedings SPIE 11291, quantum dots, nanostructures, and quantum materials: growth, characterization, and modeling XVII; 2020. 112910K. https://doi.org/10.1117/12.2558405. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11291/2558405/A-need.

[228] Schmid H, Borg M, Moselund K, Gignac L, Breslin CM, Bruley J, Cutaia D, Riel H. Template-assisted selective epitaxy of III-V nanoscale devices for co-planar heterogeneous integration with Si. Appl Phys Lett 2015;106:233101. https://doi.org/10.1063/1.4921962. http://aip.scitation.org/doi/10.1063/1.4921962.

[229] Bologna N, Wirths S, Francaviglia L, Campanini M, Schmid H, Theofylaktopoulos V, Moselund KE, Fontcuberta I Morral A, Erni R, Riel H, Rossell MD. Dopant-induced modifications of $Ga_xIn_{(1-x)}P$ nanowire-based p-n junctions monolithically integrated on Si(111). ACS Appl Mater Interfaces 2018; 10:32588–96. https://doi.org/10.1021/acsami.8b10770.

[230] Wirths S, Mayer BF, Schmid H, Sousa M, Gooth J, Riel H, Moselund KE. Room-temperature lasing from monolithically integrated GaAs microdisks on silicon. ACS Nano 2018;12:2169–75. https://doi.org/10.1021/acsnano.7b07911. https://pubs.acs.org/doi/10.1021/acsnano.7b07911.

CHAPTER THIRTEEN

Nitride single photon sources

Mark J. Holmes[1] **and Rachel A. Oliver**[2,*]
[1]Institute for Nano Quantum Information Electronics, The University of Tokyo, Meguro-ku, Tokyo, Japan
[2]Department of Materials Science and Metallurgy, The University of Cambridge, Cambridge, United Kingdom
*Corresponding author: rao28@cam.ac.uk

1. Introduction

1.1 The concept of a single photon source

An ideal single photon source is a device which generates exactly one photon per excitation trigger pulse – i.e., a device which can provide single photons reliably on demand. Such a device differs fundamentally from more typical nitride light emitters, such as light emitting diodes (LEDs) or laser diodes (LDs). It is not simply a dim and inefficient version of these less esoteric devices. For an LED or LD operating in pulsed mode, one could attenuate the emission such that, for any given excitation pulse, the most likely number of photons to reach a detection system would be zero. However, whilst under these circumstances there would be a finite probability of just one single photon arriving at the detection system following an excitation, there would also be a finite probability of more than one photon reaching that detector. True single photon emission on the other hand may be obtained by exploiting the emission of a single quantum system. Such a system may be excited from its ground state to a higher excited state, and relax, emitting a photon. After relaxation, the system is no longer in an excited state and cannot emit another photon until a further excitation trigger is provided [1]. In an ideal technological implementation of such a scheme, a single photon would reach the detection system for every trigger pulse, and the occurrence rate of multiple photon events would be zero. The probabilities for the emission of different numbers of photons by different light sources are illustrated in Fig. 13.1. Relevant single quantum objects include atoms or ions in the gas phase, certain fluorescent organic molecules, atom-like defects (most famously nitrogen-vacancy (NV) centres in diamond) and semiconductor quantum dots (which may either be formed in solution by chemical synthesis or by epitaxial growth from the gas phase).

Semiconductor Nanodevices
ISBN: 978-0-12-822083-2
https://doi.org/10.1016/B978-0-12-822083-2.00011-3
© 2021 Elsevier Ltd.
All rights reserved.

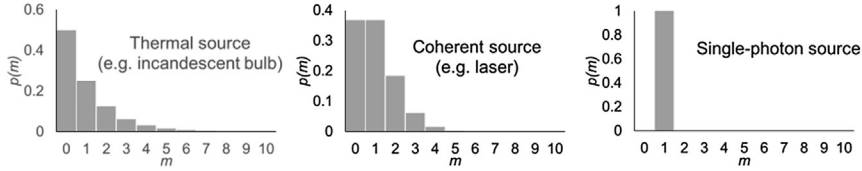


Figure 13.1 Probability, $p(m)$, of the number, m, of photons emitted simultaneously for three different light sources each sources each emitting on average a single photon (Left) Thermal source. (Centre) Coherent source. (Right) Single photon source.

Such a light source has no role to play in general illumination. However, the potential applications of single photon sources are many and varied and include the measurement of weak absorptions (where the use of classical light sources results in measurement sensitivities limited by photon noise, i.e., fluctuations of the number of photons) and the generation of random numbers for conventional computing and cryptography. Probably, the most exciting family of applications, however, are in quantum communications and information processing. Using single photons to carry information in an optical communication system provides a means to reliably test for eavesdropping on the data channel. Hence, by using single photon sources, one could assure secure transfer of digital cryptographic keys. Such quantum key distribution (QKD) schemes have been demonstrated using weak laser pulses, but the use of such classical light sources compromises the security of the key. Efficient single photon sources could also enable linear optical quantum computation, a potentially scalable quantum computing paradigm which uses an all-optical architecture, in which the qubits (quantum bits) are represented by photons and manipulated by mirrors and beam splitters.

1.2 Key measurements on single photon sources

Single photon sources, being optical devices, are characterised using optical measurement techniques. That is, the single photon emission must be collected (usually using the objective lens of a microscope or related system), filtered from spectral contamination at other wavelengths, and then directed to some detector. In this section, we will discuss the measurements and typical setup that are used for characterising a single photon emitter. The single photon emitter can be excited to its emission ready state either optically (by way of pulsed or continuous wave [CW] lasers) or electronically (via the direct injection of electrons, which can also be performed continuously, or pulsed). Here, we will not concern ourselves with the exact nature of the excitation, but will discuss the general differences between continuous and pulsed excitation.

As a given objective lens will cover a certain numerical aperture (typically of order ~ 0.5), we will not be able to collect all of the photons emitted from the device. However, the emission that is collected is then typically directed to a spectrometer where it is dispersed using a diffracting element and focussed onto a detector (typically a charge coupled device). At this stage, inspection of the emission spectrum can be performed. For solid state single emitters, it is experimentally challenging to obtain a purely spatially and spectrally isolated emitter. Even when collecting the emission from a small area of the sample using a microscope objective, there are still likely to be various emissions at different energies (corresponding to different emission centres, and different band gaps in the material system). However, the emission from a single photon emitter will typically be a sharp peak in the spectrum corresponding to the energy of the two-level transition that gives rise to the single photon emission. Spectral filtering can be employed to cut away the light at energies not corresponding to the emitter. This filtering can be performed with combinations of short- and long-pass filters, or by using the spectrometer as a monochromator and directing the light though a narrow exit slit.

After filtering has been performed and the emitter has been both spectrally and spatially isolated to a sufficient degree, the nature of the emission can finally be determined. Due to poor photon number-resolving properties of typical light detection systems, single detectors are not usually employed to determine if the measured emission consists of a stream of individually emitted photons or not. With current technology, characterising single photon emission is not as simple as measuring the emission intensity on a detector. In order to characterise the photon statistics of a given emitter, the second-order correlation function (intensity autocorrelation)

$$g^{(2)}(\tau) = \frac{\langle I(t+\tau) \cdot I(t) \rangle}{\langle I(t) \rangle^2}$$

must be measured. Here, τ is the time delay, and I is the emission intensity. A measurement of $g^{(2)}(\tau)$ is performed using *two* detectors (typically, photomultiplier tubes, avalanche photodiodes or recently superconducting nanowire detectors) and a 50/50 beamsplitter (half-silvered mirror) in a setup first designed by Hanbury Brown and Twiss (and known as an HBT setup). The electrical signals from the detectors are connected to the start and stop inputs of a Time-to-Amplitude converter, which measures the duration of time between detection events at the two detectors. See Fig. 13.2 for a

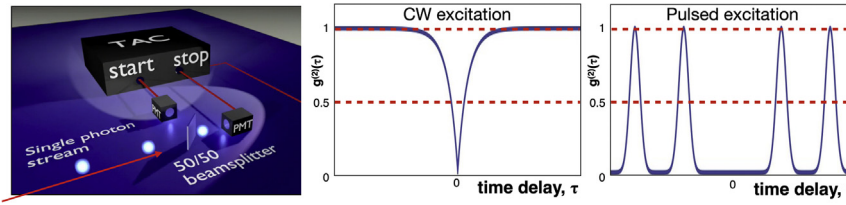


Figure 13.2 (Left) Schematic of a typical single photon detection HBT setup. (Centre/Right) idealised autocorrelation data from a single photon source showing anti-bunching in the photon statistics (a supressed signal at time delay 0) for both continuous wave and pulsed excitation conditions.

schematic of such a setup, and an example of typical measurement data. For large values of τ, the correlation is normalised to one, and the important point for characterising the emission is at time delay zero: $g^{(2)}(0)$. For the case of measurement of a pure single photon emitter without background contamination, and using a setup with very high temporal resolution detectors, the number of instantaneous counts on the two detectors will be zero: $g^{(2)}(0) = 0$ (as a single photon cannot be detected at two detectors at the same time). This is in contrast to coherent light (such as laser emission), for which $g^{(2)}(0) = 1$, and thermal light (such as sunlight), for which $g^{(2)}(0) > 1$. Experiments to measure $g^{(2)}(0)$ can be performed under either CW excitation or pulsed excitation. A CW measurement of $g^{(2)}(0)$ from a single photon emitter typically produces data with a 'dip' at $\tau = 0$, whereas a pulsed excitation measurement results in a series of peaks separated by the time between excitation pulses with no peak at $\tau = 0$. See for Fig. 13.2 for some examples of the data in both cases.

As mentioned above, the pure extraction of single photons without any contaminating emission is technically challenging. Accordingly, in reality, most experimental verifications of single photon emission do not measure $g^{(2)}(0) = 0$. Indeed, photon *pairs* impinging on the HBT setup will yield a measurement of $g^{(2)}(0) = 0.5$ (it can be shown that a $g^{(2)}(\tau)$ measurement of a Fock number state of n photons will result in $g^{(2)}(0) = 1 - 1/n$, see Ref. [2]). Therefore, any measurement of $g^{(2)}(0) < 0.5$ is actually enough to claim the presence of a single photon emitter (with some level of contaminating background that degrades the purity of the extracted single photon emission). For practical application of single photon emitters, it will be required that $g^{(2)}(0)$ be very close to zero, and indeed recent experiments

[3] using InAs/InP quantum dots have yielded $g^{(2)}(0)$ values as low as 4×10^{-4}. Further properties of an *ideal* single photon source are outlined in the following section.

1.3 Basic properties of the 'ideal' single photon source

By their very nature, single photon sources are not bright sources of light. It is therefore of paramount importance to realise efficient emitters: every photon counts! Although the development of single photon emitters is still at an early stage (particularly with III-nitride based devices), the important properties are clear. Ideally, one would need a site-controlled structure, such that device arrays and coupling to down-stream photonic components can be facilitated. The emitter itself should be have an internal quantum efficiency of unity, emit single photons with high purity (characterised by $g^{(2)}(0) = 0$) and emit into a narrow solid angle such that the emitted photons can be collected efficiently and coupled to some additional optics/photonic circuit. If we imagine some environments in which such emitters may be used in future (CPUs, data centers, inside biological systems, satellites), then the operational temperature also becomes an important concern. Ideally, one would like these emitters to operate at room temperature or even hotter environments, and also be robust to thermal fluctuations in those environment conditions. Finally, depending on the intended application of the emitter, the degree of coherence of the emission is also an important concern. For some applications, indistinguishability between each photon is required, meaning that each photon has exactly the same energy, phase and polarisation. Under these conditions, it is possible to perform two-photon interference, and realise more advanced applications for quantum information processing.

2. Basic principles of quantum dot fabrication

Typical nitride LEDs are based on quantum wells: a layer of a narrow bandgap material sandwiched between two layers of wider bandgap material, with the layer thickness being only a few nanometres, confining carriers to within the layer, but allowing them to move within the plane of the well. Whilst in a quantum well, only one dimension is on the few nanometre scale, in a quantum dot it is desirable that all three dimensions are of the order of the size of the exciton Bohr radius, leading to quantum confinement in all three dimensions. This implies that the spatial extent of the

narrow bandgap material should be only a few nanometres in all three dimensions. In practice, quantum dots may be up to a few tens of nanometres in one or more of their spatial dimensions and nevertheless exhibit quantum properties, but the fact remains that making quantum dots is intrinsically a nanofabrication problem. Whilst modern techniques for semiconductor processing permit the formation of structures of these very small sizes, the use of lithography to define quantum dots often leads to damage to the material which degrades quantum dot properties. Hence, rather than using the 'top-down' lithographic approach, quantum dots are often formed using self-assembly processes, an approach referred to as 'bottom up'. A simple analogy for the difference between a top-down and a bottom-up process is the growth of a bonsai tree. If, in the top-down paradigm, it was required to make a very small tree, one might start with the trunk of a large tree and some woodworking tools. One could saw and chisel, carve and sand to produce a much smaller chunk of wood in the shape of a very small tree. How much more elegant is it, however, to take a seed, and engineer the conditions in which the tree grows (for example, confining the trees roots in a small pot) in order to allow a fully functioning miniature tree to self-assemble naturally? Nonetheless, there are disadvantages to the self-assembly approach to quantum dot growth, as we shall discuss, and opportunities are presented by combining the top-down and bottom-up paradigms.

2.1 Self-assembly on planar surfaces

Let us consider the available growth modes in the epitaxial growth of a thin film on an unlike substrate, where the thin film material has a lower bandgap than the substrate material, as is required for the formation of a quantum well or a quantum dot. Typically, the substrate surface will be terraced, and atoms impinging on the surface may be able to diffuse across the surface and become attached to the surface at step edges. Step edges represent a low energy site for the diffusing atom because a greater number of bonds may be formed by an atom at a step edge than an atom in the middle of a terrace. When growth proceeds via the attachment of atoms at step edges, this is termed 'step flow growth' (Fig. 13.3). However, if the step edges are too widely spaced, or the temperature is too low, limiting the diffusion distance on the surface, then an impinging atom will not typically be able to reach a step edge. In this case, growth may proceed by the formation of monolayer islands on the step edges, which grow and merge to reform a smooth

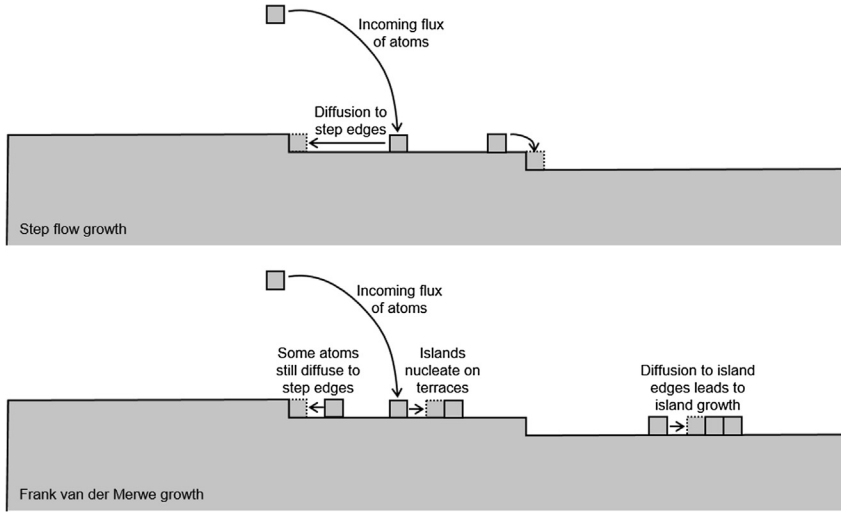


Figure 13.3 Schematic illustrations of two-dimensional growth modes: (Top) Step flow growth. (Bottom) FvdM growth.

terraced surface, prior to the formation of further islands, which once again grow and merge (Fig. 13.3). This growth mode is termed 'island nucleation and growth' and is also referred to as 'Frank-van der Merwe' (FvdM) growth [4]. A consequence of FvdM growth is that isolated 2D islands of the narrow bandgap material can be formed either on the surface of the wide bandgap substrate or on the surface of the growing narrow bandgap film. Following capping with a further layer of wide bandgap material, either the isolated monolayer islands or the monolayer thickness fluctuations of a quantum well grown in the FvdM mode can act as quantum dots. Indeed, we will discuss in Section 13.4 the characteristics of quantum dots formed due to monolayer thickness fluctuations in GaN/AlN quantum wells (These structures are sometimes referred to as 'interface fluctuation quantum dots').

It should be noted, however, that the formation of such quantum dots can be difficult to monitor and control, particularly in MOVPE. In MBE, the use of reflection high-energy electron diffraction (RHEED) provides the means to monitor the island nucleation, growth and coalescence process: the electron beam incident upon the surface may be scattered away from the expected diffraction spots by surface steps. Hence, for a surface free from monolayer islands, where the step density is lowest, the intensity of the RHEED streaks will be highest. However, once islands nucleate and grow, the step density increases causing scattering and a decrease in the

intensity of the RHEED streak. As the islands start to coalesce, the number of scattering sites decreases, and the RHEED streak becomes brighter again. Once the 2D islands have fully coalesced to form a smooth surface, the cycle begins again. Hence, the intensity of the RHEED spots oscillates with time, as the cycle of island nucleation, growth and coalescence repeats [5].

Our consideration of growth modes up to this point has ignored the influence of strain on the growth of the film. Whilst matching of the in-plane lattice constants can be achieved between a substrate and a heteroepitaxial layer for certain material compositions, in the general case, the film and substrate have different lattice constants (most commonly the wider bandgap material has a larger in plane lattice constant than the narrow bandgap material). For the growth of a coherent thin film (i.e., one in which all the lattice planes are continuous across the film–substrate interface), the strain energy of the film will increase as the film thickness increases, providing a driving force to activate strain relief mechanisms. Strain can be relieved by plastic relaxation – i.e., by the introduction of misfit dislocations at the film substrate interface – but may also be relieved elastically. The formation of three-dimensional islands represents an elastic strain relief mechanism since, unlike in a continuous two-dimensional layer where any given region of the material is constrained to conform to the lattice parameter of the underlying substrate by the presence of adjacent material, the material in a three-dimensional island has space around it into which it can relax elastically, relieving some of the strain in the film [6], as illustrated schematically in Fig. 13.4.

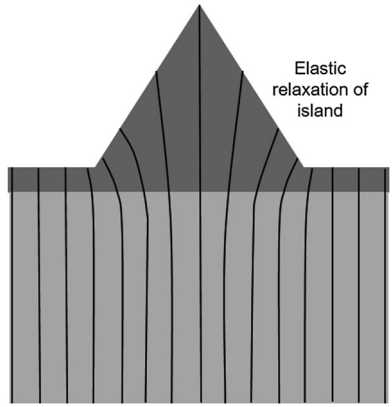


Figure 13.4 Schematic illustration of the elastic relaxation of three dimensional island during the growth of an epitaxial layer (dark grey) on a mismatched substrate (lighter grey) with a smaller in plane lattice parameter. Black lines symbolise lattice planes.

Let us briefly consider the energetics of thin film growth. As illustrated in Fig. 13.5, if we consider the growth of a lattice matched 2D island on a substrate, we must consider the balance of the energy of the surface of the substrate, γ_s, the energy associated with the surface of the thin film, γ_f, and the energy of the film–substrate interface, γ_{fs}. The growth of the island removes an area of clean substrate surface, replacing it with an equal area of film–substrate interface and an equal area of film surface. If $\gamma_f + \gamma_{fs} < \gamma_s$, then the film will tend to wet the surface, covering it uniformly. If $\gamma_f + \gamma_{fs} > \gamma_s$, then there will be a tendency towards dewetting and the formation of three dimensional islands. Even in the absence of strain, this can lead to a growth mode where no flat two dimensional film is formed, and the material grows as facetted three-dimensional islands. This is termed Volmer–Weber (VW) growth. In the case where $\gamma_f + \gamma_{fs} < \gamma_s$ but where the film is also strained, the film will initially grow as a two-dimensional layer. However, as the film thickness increases and the strain energy builds up, there will be a driving force for the formation of three-dimensional islands. Once this driving force is sufficient to overcome the energetic cost of forming an island, then such islands will form on top of the two-dimensional layer. The two-dimensional layer is referred to as the 'wetting layer' and the overall growth mode in which growth initially proceeds by FvdM growth prior to the formation of coherently strained islands is known as Stranski–Krastanov (SK) growth. The VW and SK growth modes are illustrated schematically in Fig. 13.6.

The net energetic cost of forming a new island is determined by the reduction in strain energy upon formation of the island, the increase in the total energy due to the presence of the island facets (which will have a larger area and a different energy per unit area) and the additional energetic cost of the presence of the island edges, where the facets meet. As a first approximation, the change in strain energy will scale as the volume of the island (roughly proportional to the cube of its height, h^3), the energy related

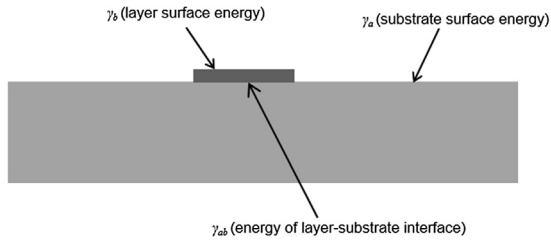


Figure 13.5 Schematic illustration of the surfaces and interfaces contributing towards the energy of formation of a two dimensional epitaxial islands on an unlike substrate.

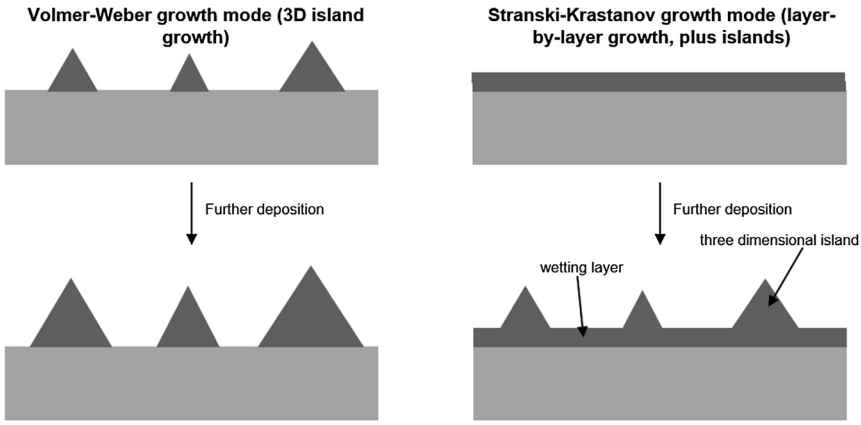


Figure 13.6 Schematic illustration of the main three-dimensional island growth modes in heteroepitaxy, and their evolution with deposition time: (left) VW growth and (right) SK growth.

to the facets will scale roughly as h^2, and the energy of the edges will scale very roughly as h. Hence, the overall change in energy upon formation of an individual island, E_{island}, can be expressed as follows:

$$E_{island} = ah^3 + bh^2 + ch + d$$

where a, b, c and d are constants [7]. There will be a minimum energy for this function at some specific island height, h, suggesting that an array of islands self-assembling in the SK growth mode will have some preferred size. (The relevant value(s) of h are only positive for certain sets of values of the constants $a - d$). For growth of an island array at finite temperature, there will inevitably be a distribution of sizes, which can be expressed by a function $f(h)$:

$$f(h) \propto e^{-(E_{island}(h))/k_B T}$$

Whilst this *very* simple model for the energetics of island growth is incomplete, and even other, more sophisticated versions [8] have been disputed [9], it is a common experimental observation that self-assembled island arrays can display remarkably good size uniformity, making SK growth a favoured method for the growth of quantum dots, since good control over the island size can typically be achieved through control of the layer thickness and growth temperature. A much more complete exposition of the energetics of island growth and the theories exploring the expected equilibrium size distribution can be found in Ref. [8], whilst the specific application of this theory to the nitrides has been discussed by Ref. [7].

More subtle points such as the impact of island–island interactions on the energetics are addressed in these papers and elsewhere [10].

Whilst this simplified theoretical picture provides a useful framework within which SK growth may be understood, the relevant parameters relating the change in E_{island} to the dot height are extremely difficult to predict accurately from first principles theory. They are likely to be dependent on a range of growth conditions, particularly the temperature and stoichiometry of the growth environment. Furthermore, the model assumes that during growth the system will reach an equilibrium state, and ignores kinetic constraints. The equilibrium state that is described here is not, in general, the true equilibrium state of the system, since if the film material were soluble in the substrate material, the true equilibrium between the two would often see the atoms of the film uniformly distributed in solution throughout the substrate. Whilst this is unlikely to happen in practice, it is common to see some inter-diffusion between film and substrate and the extent of this will vary with temperature. All these factors mean that achieving SK growth of quantum dots in practice is usually achieved by trial and error: exploring a range of growth conditions and looking for signatures of three-dimensional nucleation, which can be monitored in situ in MBE using RHEED (When three dimensional islands nucleate on the surface, the RHEED beam is transmitted through them, resulting in a transition from a streaky to a spotty RHEED pattern). Alternatively *ex situ* techniques such as atomic force microscopy can be used to monitor the growth mode.

In the nitrides specifically, SK growth has been achieved in MBE for both InGaN/GaN and GaN/AlN [11], and there is some evidence that the equilibrium model presented here is relevant to the trends in quantum dot size distribution seen when the thickness or temperature is changed [7]. Such quantum dots grown by MBE often achieve high areal densities ($>10^{11}$ cm^{-2}), making it difficult to isolate a single structure for optical characterisation. This has limited the number of studies of single photon emission from these systems, although [12] have successfully developed InGaN/GaN single photon sources emitting in the red using MBE-grown dots, as we shall discuss later. In MOVPE, there is a lack of consensus on whether an appropriate set of growth conditions to induce SK growth can be found for growth of InGaN on the (0001) plane of GaN. In the earliest growth of InGaN quantum dots by MOVPE, it was suggested that the growth proceeded somewhat similarly to the SK mode [13], but some unusual features were noted, and later authors have provided evidence that SK growth does not occur [14]. This issue has not been fully explored, but it is possible that

the presence of hydrogen during MOVPE growth lowers the energy of the (0001) InGaN surface, making three-dimensional growth less favourable. For growth on the nonpolar a-plane, recent studies have suggested that the SK mode may be achievable in MOVPE, but again insufficient detailed analysis of the growth has been performed to fully support this suggestion [15]. For GaN/AlN growth in MOVPE, the SK growth mode has been reliably achieved, however [16]. Even where SK growth is successfully achieved, the completion of the quantum dot structure requires the overgrowth of a wide bandgap cap, and care must be taken to ensure that any change in growth conditions needed to complete the cap growth does not lead to damage to the quantum dot layer. Ref. [17], for example, showed that in MBE, capping can result in complete destruction of the three-dimensional dot structures if inappropriate growth conditions are selected.

SK growth is the most popular method for quantum dot growth in the well-established arsenide materials system in which the greatest successes for quantum dot single photon sources have been achieved. The forgoing discussion of this growth mode also provides a useful example of the thermodynamic factors which influence nanostructure formation across the board (and whilst other growth methods will now be discussed in somewhat less detail, the reader should be aware than many of the same considerations play a role). However, even in the arsenide system, SK growth has its limitations; for example, it can only be used when there is sufficient lattice mismatch between the quantum dot and the substrate. In nitride MOVPE, the difficulty of reliably achieving SK growth has led to a plethora of alternative methods being considered. The most relevant of these to single photon sources is droplet epitaxy, a modified version of which has been extensively explored for InGaN/GaN single photon sources [18].

In droplet epitaxy, nanoscale metallic droplets (e.g., indium or gallium) are first formed on a substrate surface, and then reacted with a nonmetal (e.g., arsenic) or a precursor for a nonmetal (e.g., ammonia as a nitrogen precursor) to form nanocrystals of a compound semiconductor. There are thus two main aspects to development of a growth process: control of the formation of the nanoscale metallic droplets and control of their crystallisation. In deposition of a metallic layer on a III-V semiconductor surface, growth usually proceeds in the VW mode. Pragmatically, the density and size of the droplets can both be controlled (although not independently) by varying the growth temperature, with increasing temperature tending to encourage the formation of larger nanostructures with a lower density. More details

about control of nanodroplet formation for droplet epitaxy can be found in a fairly recent review [19]. Following metallic droplet formation, the supply of a flux of a nonmetallic element leads to the formation of semiconductor nanostructures which can take various forms including quantum dots similar to those formed in SK growth, nanorings and double nanoring structures. In GaAs, droplets can also be used to etch away the underlying material via dissolution of the GaAs into the Ga droplet, creating nanoholes, which can be used to template quantum dot growth at specific locations.

Both GaN [20] and InN [21] nanostructures have been grown using droplet epitaxy approaches similar to those which are popular in arsenide research but such structures have not in general demonstrated clear QD optical signatures, and little progress has been made in applying this approach for single photon source fabrication. However, in the InGaN/GaN system, a slightly different form of droplet epitaxy (sometimes referred to as modified droplet epitaxy or MDE) has been extensively explored, in which (instead of forming metallic droplets using a metallic flux) an InGaN layer is deposited by MOVPE and then annealed in an atmosphere free from active nitrogen, resulting in partial decomposition of the InGaN layer and formation of indium-rich metallic droplets [22]. The InGaN layer is not completely destroyed but rather remains in a somewhat damaged form with large pits appearing in the layer which penetrate through its entire thickness. The metallic droplets may rest either on the remains of the InGaN film, or in the pits, directly on the GaN surface. Rather than then supplying an active nitrogen flux in order to form nitride QDs, the droplets react with ammonia during the overgrowth of the InGaN layer by GaN. The resulting structure is somewhat disordered, with some of the QDs resting on an InGaN quantum well, in a structure akin to the QD atop a wetting layer which forms in SK growth, whilst other QDs, arising from droplets formed in pits in the InGaN layer, are entirely surrounded by GaN and have are not in contact with a quantum well structure. Nonetheless, there are similarities with the standard droplet epitaxy approach described for GaAs, including the formation of nanoring structures under some circumstances [23].

A number of other self-assembly methods for nitride QDs have also been attempted in the InGaN/GaN system, possibly due to the difficulty of achieving reliable SK growth in this system by MOVPE, which is the most popular epitaxy technique. These include various approaches to pretreatment [24] or passivation [25] of the substrate surface and approaches which attempt to exploit phase separation phenomena in InGaN [26]. These approaches are less widely used, however, than MDE or SK growth in the context of single photon sources.

2.2 Self-assembly of nanorods

Another widely used approach to the self-assembly of QDs in the arsenides and many other semiconductors is based on the vapour–liquid–solid (VLS) mechanism for the formation of nanorods. The VLS mechanism was originally developed as a method for the fabrication of micron-sized single-crystalline silicon whiskers in the 1960s [27]. However, it took almost 30 years for researchers to begin to recognise its potential for the growth of very fine nanowires [28]. The VLS mechanism is most straightforwardly explained in the context of its original incarnation: the growth of Si structures using a gold catalyst. If we consider the Au–Si phase diagram (see Fig. 13.7), then we see that when an Au nanoparticle is heated to above the eutectic temperature (363°C) in a Si-containing atmosphere it is possible to form a liquid

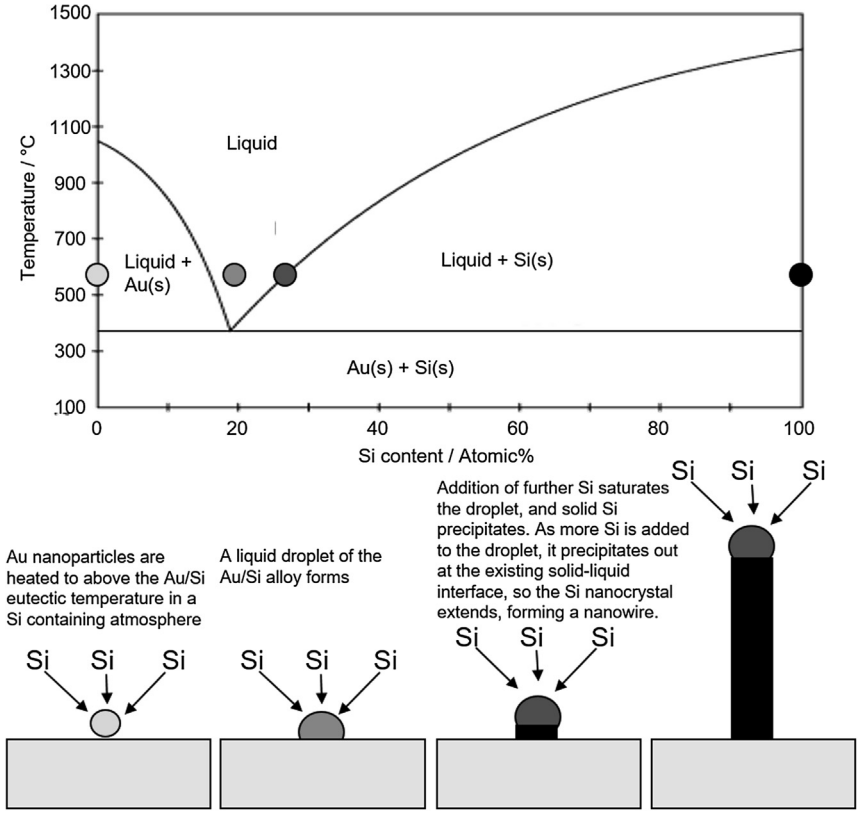


Figure 13.7 VLS growth of Si nanowires using an Au catalyst particle. (Upper) Au–Si phase diagram (after [78]). Various compositions are marked with grey circles, the shade of which matches material of that composition illustrated in the schematic diagram (lower).

Ai–Si alloy. If the amount of Si in the droplet increases, the solution of Si in Au can become supersaturated, and Si will precipitate out, usually at the droplet–substrate interface, since this minimises the increase in surface/interface energy. As more Si is added to the droplet, it then precipitates out at the existing solid–liquid interface, so the Si nanocrystal extends, forming a nanowire. This mechanism results in a nanowire diameter that is controlled by the size of the original Au catalyst particle, so that smaller catalyst particles result in narrower wires.

Nanoparticles of a range of metals including Co, Ni and Fe can be used to catalyse GaN nanowire growth [29]. However, the possibility of a self-catalytic mechanism employing liquid gallium droplets has also been suggested [30] (Other III-V semiconductor nanowires can also be grown by self-catalytic methods). This is an example of an extrinsic particle free method, and has the advantage of avoiding the risk of contamination of the sample or the growth system by the foreign catalyst material. In the nitrides, such methods need not in fact involve a metallic particle at all. For example, in MBE, it has been suggested that GaN nanowires can form due to thermodynamically driven variations in surface sticking coefficients on different crystallographic planes [31].

Once the growth of a nanowire structure has been established, a QD may be formed by briefly switching the precursors used from those required for the nanowire stalk to those required for a related narrower bandgap material, for example, switching from an Al precursor flux to a Ga precursor flux for a short period during the growth of an AlN nanowire to form a GaN region sandwiched between two AlN regions (see Fig. 13.8). This 'axial heterostructure' acts as a quantum dot [32].

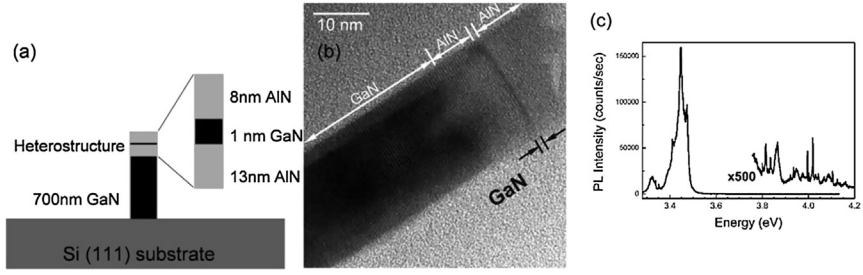


Figure 13.8 Axial heterostructure in an MBE-grown nitride nanowire: (a) Schematic illustration of structure, (b) High-resolution TEM image of a single GaN inclusion which acts as a quantum dot, as illustrated by the micro-photoluminescence spectrum in (c) which illustrates the emission from an ensemble of such structures at 5 K, with sharp peaks originating from individual quantum dots being observed at higher energies than the emission from the GaN nanowire stem at ca 3.45 eV due to quantum confinement. *Reprinted (adapted) with permission from [32]. Copyright 2008 American Chemical Society.*

2.3 Lithographic approaches to quantum dot formation

An alternative approach to quantum dot fabrication which is closely related to the exploitation of self-assembled nanowires defined above is the use of selective area epitaxy (SAE) techniques to facilitate the growth of nanowires or other relevant micro- or nanostructures. In SAE of nitrides, a dielectric layer (typically SiO_2 or Si_3N_4) is deposited either onto a pre-grown GaN template or directly onto a foreign substrate such as sapphire. Lithographic techniques are used to create a pattern of holes in this layer, which then acts as a mask for subsequent nanowire growth. Initial growth conditions must be carefully selected, to avoid parasitic growth on the masking material, as well as in the openings. Thereafter, growth conditions which encourage vertical rather than lateral growth are required for the formation of nanorods [33]. In MOVPE, the growth conditions must be particularly carefully selected if nanorods are required, since within much of the growth window, hexagonal pyramid structures tend to form rather than extended rods.

For the purposes of quantum dot formation, however, both hexagonal pyramids and nanorods with a pencil-like structure (where the rod narrows to a point at the top often in a hexagonal pyramidal morphology) have proved useful. A core consisting of either a nitride nanorod or a pyramid may be completely overgrown with a shell of a narrower bandgap nitride which coats all exposed facets. Over this, an outer shell of the wide bandgap material may be grown. In this core–shell morphology, the thickness and (where an alloy is deposited) composition of the narrow bandgap layer will depend on the orientations of the different facets of the core structure. Additionally, edges and apices exhibit different growth rates and local compositions to the planar regions. One upshot of this is that a core–shell nanorod may have quantum wells emitting at different wavelengths on differently oriented facets [34]. Furthermore, when (as is often the case) the edges or apices result in a locally altered growth rate or composition, quantum wires and quantum dots may form. Fig. 13.9 illustrates one configuration in the formation of InGaN/GaN core–shell structures in which the apex region at the top of a pyramid or pencil-like nanorod is enriched in indium and, as long as that apex forms a sharp point, this results in the formation of a quantum dot [35]. Similar quantum dot structures can also be formed at the apex of core–shell heterostructures incorporating AlGaN and GaN [36].

Such quantum dots formed at the apex of pyramid or nanorod structures are increasingly widely studied. The approach has a number of advantages: the quantum dots are formed at well-defined locations, usually with a fairly

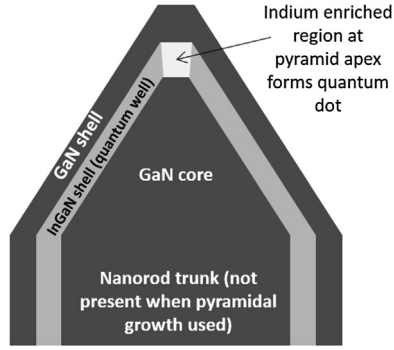


Figure 13.9 Schematic illustration of quantum dot formation at the apex of a core–shell structure based on a nanorod or pyramid grown by SAE.

low density and are thus easily located. Pyramids with sharp apices can be fairly straightforwardly formed on (0001) oriented GaN by using growth conditions which favour semi-polar facet formation. Since the quantum dot is formed at the apex, it is not necessary to use nanoscale lithography. Micro-scale pyramids with sharp points can successfully template quantum dot formation and the relevant patterning of the mask is reasonably straightforward.

Despite the disadvantages mentioned above, the direct lithographic definition of nitride quantum dots has also been attempted by, for example, defining structures of few-nm lateral dimension on the surface of an InGaN/GaN quantum well sample using electron beam lithography and then etching to form nanopillars containing isolated InGaN disks. Although there are a few reports [37] of successful single photon source fabrication using this method, damage to the rod surface during etching remains a concern, although the low surface recombination velocities of the nitrides render the impact of the resulting surface states less severe than in other materials such as the III-arsenides.

3. Basic properties of quantum dots for single photon emission

3.1 The physics of 3-dimensional confinement

A quantum dot is, in essence, a localised region of space of low electrical potential energy, $V(r)$, in which electronic charges can be confined [38]. The confinement is three dimensional in nature, and the physical size of the quantum dot is on the order of a few nanometres such that the system is described by quantum mechanics. Accordingly, the energies, E_i,

and wavefunctions, $\psi_i(r)$, of the electrons and holes trapped in the conduction band (CB) and valence band (VB) of the quantum dot are in principle defined by the Schrödinger equation:

$$\frac{-\hbar^2}{2m}\nabla^2\psi_i(r) + V(r)\psi_i(r) = E_i\psi_i(r)$$

The three-dimensional nature of the quantum confinement leads to complete quantisation of energy such that the density of states corresponds to a series of delta functions at the energies E_i, rather like that of an atom. For this reason, quantum dots are also known as 'artificial atoms', and the emission spectra of quantum dots contain series of sharp lines corresponding to transitions between these energy levels (from electrons in the CB to holes in the VB). Each of the levels, as defined above, is doubly degenerate in spin. However, in most real quantum dots the degeneracy of the spin states is usually lifted by some anisotropy – leading to a fine structure splitting (FSS) in the emission spectrum.

A transition between two given states $i \rightarrow j$ (otherwise known as recombination between an electron occupying state i in the CB and a hole occupying state j in the VB) can occur if the optical transition dipole moment is finite, i.e., the transition obeys selection rules and is not forbidden. When the dot is occupied with a single electron/hole pair, the transition proceeds via the emission of a single photon at an energy $E_\gamma = E_i - E_j$, less any other perturbations such as Coulomb interactions (the exciton binding energy) and spin-related interactions, which are much smaller than E_γ. The energy states of the quantum dot must then be repopulated with another electron/hole pair before a second photon can be emitted. This 'dead time' in the emission during the re-population process allows for the emission, one-by-one, of individual photons of light.

Due to fast energy relaxation of charges (through the emission of phonons), any electron (hole) that becomes trapped in a quantum dot will soon relax to the lowest (highest) energy level in the conduction (valence) band. Therefore, the single photon-emitting transition process usually occurs between the lowest confined level in the CB and the highest energy level in the VB. This process of (nonresonant) excitation, relaxation and recombination is shown schematically in Fig. 13.10.

It is important to consider the situation in which more than one electron/hole pair occupies the dot. Indeed, under conditions of stronger excitation, such that the exciton population of a quantum dot becomes two, a bi-exciton state (two electrons and two holes in the same dot) can be formed. For conventional bi-exciton states, both charges of a given type

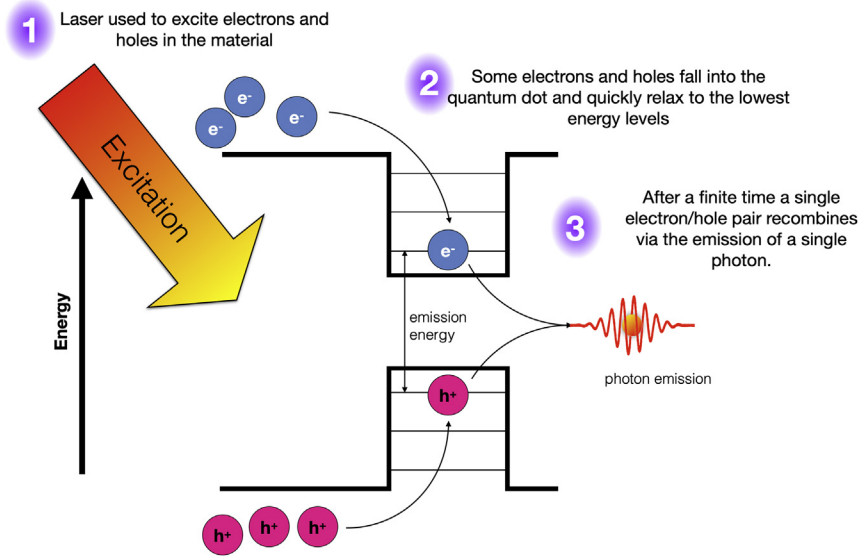


Figure 13.10 Schematic showing the main processes in single photon emission.

occupy the same energy level (with opposite spin), but the mutual Coulomb and exchange interactions between the four carriers act to perturb the total energy of the system, such that the total energy of the bi-exciton state is usually not equal to two times the average exciton energy. The energy difference, ΔE_{XX}, is known as the bi-exciton binding energy. The energy levels of the whole system (including the FSS in the exciton state) are shown in Fig. 13.11. Although there are four possible ways to select one electron and one hole from two pairs, only two of the combinations have a finite transition matrix element (due to conservation of angular momentum), and therefore only two exciton levels (the so-called 'bright excitons') are shown in the figure. In the figure, it can be seen that when one of the constituent electron/hole pairs of the bi-exciton recombines (also a single photon emission process in itself), the energy of the emitted photon is different to the aforementioned excitonic emission energy. After the first recombination process (the bi-exciton to exciton state transition), the quantum dot is occupied with a single exciton, which is then free to recombine to emit another single photon. Such a bi-exciton–exciton cascade process results in the emission of two individual photons with an energy separation equal to ΔE_{XX}. Due to the quantum confinement effect in the quantum dot, ΔE_{XX} can be either positive or negative depending on the strengths of

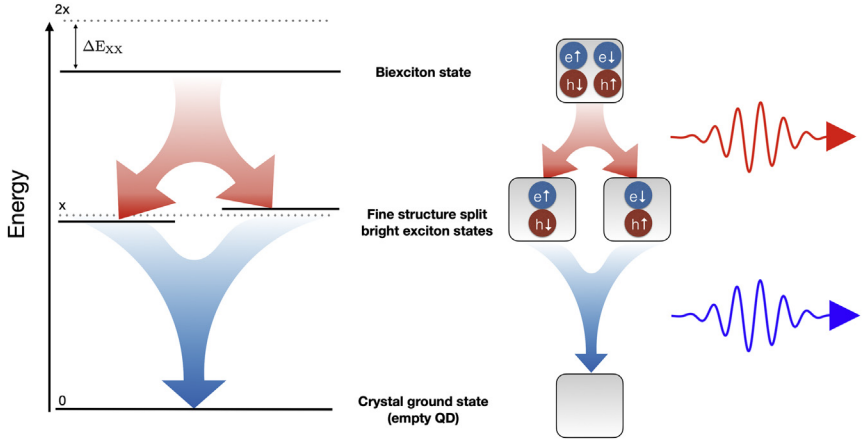


Figure 13.11 Energy diagram showing the bi-exciton state, fine-structure split bright exciton states, and the crystal ground state of a quantum dot. The cascade from the bi-exciton state to the crystal ground state can occur via either of the two exciton states.

various interactions between the constituent charges. The presence of an additional peak in the emission spectrum of the dot can lead to problems when attempting to isolate a single transition for single photon emission purposes. To that end, it is desirable to have a large bi-exciton binding energy, such that the emission peaks are well separated in energy and can be filtered. However, it should be noted that in some cases, a bi-exciton binding energy of zero can be used for the generation of entangled photon pairs [39].

3.2 Specific properties of III-nitride quantum dots and considerations for single photon emission

Whilst, at the most fundamental level, quantum dots of any material are described by the same physics, III-nitride QDs have a few special properties that make them stand out. In general, one of the most important features of the III-nitride material system is the large range of bandgaps that are available. These in principle allow the realisation of emitters that can operate over a wide range of emission wavelengths (from the ultraviolet to the infra-red [IR]), and also the fabrication of confinement potentials that are so high as to allow for the confinement of charge at even room temperature. These effects will be discussed further in the next section.

III-nitride semiconductors most commonly form with the Wurtzite crystal structure (with space group $P6_3mc$), which, due to a lack of inversion

symmetry, induces a piezoelectric nature in the material. It is therefore an inherent property of III-nitride heterostructures to exhibit an internal electric field (strain exists in the system due to the coherent growth of materials with different lattice constants, and polarisation charges thus build up on the heterostructure interfaces). The internal electric field in III-nitride nanostructures (which in essence consists of both pyroelectric and piezoelectric components) is of order $\sim \mathrm{MeVcm}^{-1}$, and lies along the crystal *c*-axis (the [0001] direction). This electric field leads to *three* related effects that are pertinent to the optical properties of III-nitride quantum dots:

(1) A strong band bending, which shifts the emission energies dramatically depending on the quantum dot size (and/or orientation with respect to the field).

(2) A significant reduction of the electron-hole wavefunction overlap due to the fact the electrons and holes are forced in opposite directions by the field (leading to suppressed recombination and corresponding long emission lifetimes).

(3) The existence of a permanent dipole moment of the electron-hole pair in the quantum dot.

The first two effects are linked and commonly known as the quantum confined Stark effect. The resulting emission energy and emission lifetime are both strongly dependent on the actual dot size. Indeed, the emission energy of GaN quantum dots can vary over the range of 3.5–4.5 eV, depending on the QD size, with carrier lifetimes spanning four orders of magnitude from a few microseconds to a few hundreds of picoseconds [40].

To illustrate the band bending and wavefunction separation, a calculation of the ground state electron and hole density functions $\left|\psi_{e,h}(r)\right|^2$, calculated using 8 band k.p theory [41,42], of a GaN/AlN QD is shown in Fig. 13.12. For simplicity, the dot geometry was chosen to be a hexagonal disk for this calculation. The electron and hole are spatially separated, forced to opposite sides of the quantum dot by the electric field.

The resulting permanent dipole moment $\overrightarrow{\mu}$: (defined as $\overrightarrow{\mu} = q\langle\psi_h|r|\psi_h\rangle - q\langle\psi_e|r|\psi_e\rangle$) interacts strongly with itinerant and trapped charges in the material surrounding the dot, and the energy of the confined electron-hole pair fluctuates in time. This random fluctuation of the emission energy is known as spectral diffusion, and is a major issue affecting the stability of emission properties of III-nitride quantum dots (see Section 4 for further discussion on this topic).

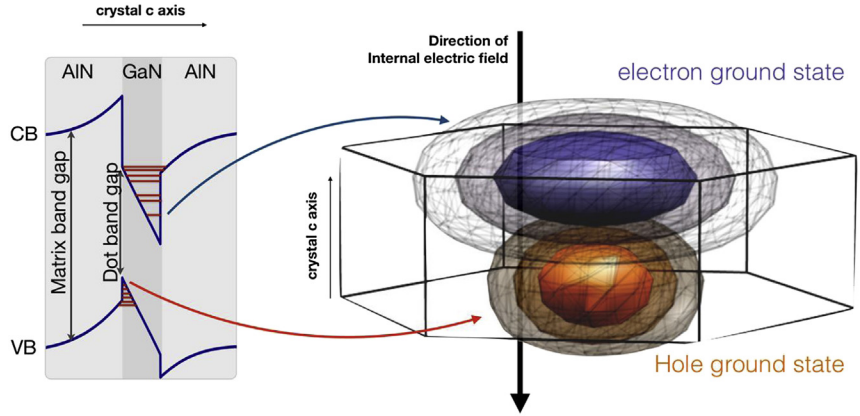


Figure 13.12 (Left) Bandstructure of a III-nitride QD, showing the band bending due to the electric field. (Right) calculated ground state electron (blue [dark grey in print version]) and hole (red [grey in print version]) density functions, $|\psi_{e,h}(r)|^2$, of such a dot.

4. Strengths and weaknesses of nitride quantum dot single photon sources

As discussed above, The III-nitride material system exhibits several properties that make it an interesting system for the generation of single photons. In this section, we provide a brief history of III-nitride QD-based single photon research, and give a guide to the current state of the art, highlighting both the strengths and weaknesses of the material system using examples from the literature.

In recent years, III-nitride quantum dots have been shown to operate as single photon emitters that can work at comparatively high temperatures, but early research into the optical properties of III-nitride QDs was largely performed at cryogenic temperatures. Indeed, the first example of single photon emission from III-nitride quantum dots (at temperatures of ~4K) was presented in 2005 [43], whereby both pulsed and CW excitation were employed to excite GaN QDs that had been grown by MOCVD. Due to the use of GaN as the QD material, the measured emission was in the UV, at wavelengths in the region of 355 nm. This work was swiftly followed with further research into extending the operation wavelength into the blue (435 nm) by using InGaN QDs [44] grown using the MDE approach in MOVPE. These structures incorporated a distributed Bragg Reflector to enhance the extraction efficiency. After these initial studies,

work branched into three main (though somewhat intertwining) avenues towards trying to realise the potential of III-nitride QDs:

(1) Exploiting the range of material bandgaps to realise single photon emission over a wide range of wavelengths.

(2) Exploiting the strong exciton confinement to realise single photon emission at temperatures where cryogenic cooling is not required.

(3) Attempts to realise faster-emitting and more efficient devices by controlling the effect of the internal electric field.

The main way to control the emission wavelength is to use materials with different bandgaps. Indeed, this technique was used for the aforementioned extension of the operation wavelength range from 350 to 435 nm via changing from GaN/AlGaN to InGaN/GaN QDs. Further extension into emission wavelengths of ~600 nm has also been possible using InGaN/GaN QDs with a larger indium content [12]. It should also be noted that a fair degree of control can also be obtained by utilising quantum dots of different size (whereby the change in curvature of the charge carrier wavefunctions will result in different energies: smaller dots emit at higher energies, larger dots emit at lower energies, all other things being equal). For example, by using small GaN/AlGaN QDs (of order 1 nm in height), emission in the ultraviolet at wavelengths of around 280 nm has been realised [45].

In addition to exploring this impressive range of available emission wavelengths, several groups have recently succeeded in generating single photons at ambient conditions using III-nitride QDs. Recent research has shown that InGaN QDs emitting in the red (~620 nm) can be used for single photon emission at temperatures up to 280 K, and even under the direct injection of current via use of an LED structure [12]. Recent work on blue-emitting InGaN QDs (~485 nm) has shown single photon emission at 220 K, which is a temperature that can be reached with thermo-electric cooling. GaN-based QDs emitting in the UV at ~350 nm have also been used to show high temperature single photon emission, at 200 K in 2006 [46], again within the operational range of thermoelectric cooling. Moreover, single photon emission has been realised at 300 K in 2014 [47], and 350 K in 2016 [36] by using small, low-density site-controlled, nanowire quantum dots. The use of such structures enabled the isolation and measurement of a quantum level transition at high temperatures even when the emission was thermally broadened by interactions with phonons. These recent high temperature results at various wavelengths are showcased in Fig. 13.13.

In addition to the strong confinement potentials of III-nitride QDs, it is also important to note that for some dot geometries, it is possible to observe very large bi-exciton binding energies. Indeed, values of ΔE_{XX} ranging from

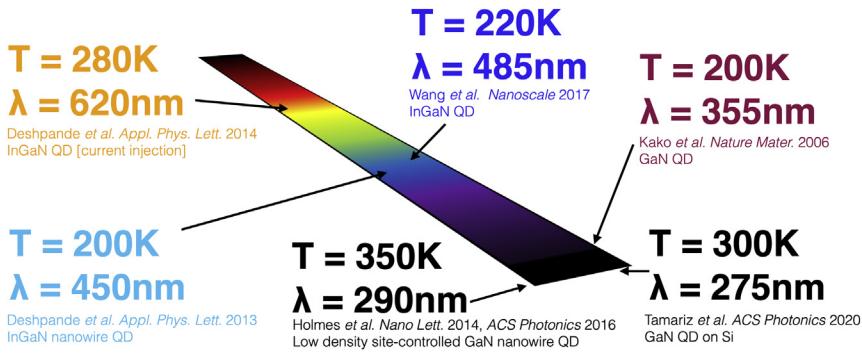


Figure 13.13 Examples of high temperature single photon emission from III-nitride quantum dots in the literature.

approximately −50 meV to +50 meV have been reported in the literature. The exact theory explaining the size and variation of ΔE_{XX} in nitride QDs is beyond the scope of this chapter, and the reader is referred to the work of Hönig et al. for further information [52]. The large bi-exciton binding energies possible with III-nitride QDs allow for spectral isolation of the emission from a single transition, even in the presence of phonon-assisted linewidth broadening.

Considerable effort has been made to suppress the effects of the internal electric field in III-nitride QDs by exploring the fabrication of dots using nonpolar [53,79] and semi-polar [54] material growth (Nonpolar growth implies the growth axis is perpendicular to the polar c-axis, whilst semi-polar growth implies the growth axis is at an acute angle to the semi-polar c-axis). Such quantum dots were indeed shown to have short exciton lifetimes, indicating a suppression of the wavefunction separation. Other researchers have experimented with the growth of nitride quantum dots with the less-common zinc blende crystal structure [55,56]. The zinc blende structure is centro-symmetric and hence nonpolar, and along certain crystal orientations, piezoelectric fields can also be avoided. Zinc blende nitride quantum dots have also been shown to exhibit single photon emission [57] with emission lifetimes about one order of magnitude faster than typical polar wurtzite quantum dots emitting at the same energy [58]. Another method to successfully suppress the effect of the electric field is to use quantum dots with a small extension along the polar crystal axis, forcing an increased electron/hole wavefunction overlap. Such quantum dots grown

in nanowires have been shown to exhibit single photon emission with lifetimes on the order a few hundred picoseconds, with the caveat that the emission is in the ultraviolet range of the spectrum at wavelengths below 300 nm [67]. Recent studies have also shown that it is possible to suppress the internal electric field in III-nitride nanostructures by using the interfaces formed in additional growth layers to provide a counteracting field [59].

Another strength of the III-nitride quantum dots for the emission of single photons lies in the polarisation of the emission. Due to effects such as QD shape anisotropy [35], strong VB mixing effects [60] and suppressed population of the high energy states of emission doublets [52], the emission from nitride QDs can be shown to exhibit a strong degree of linear polarisation. Linearly polarised single photons are required for the realisation of QKD using the BB84 protocol [61]. However, correct realisation of the protocol would require the generation of single photons in orthogonal states of two bases. Therefore, linear polarisation by itself is insufficient: control over the axis of the polarisation is also required. To that end, several researchers have been working on controlling the emission polarisation by several means. These have included positioning QDs within structures that exhibit overall controlled anisotropy such as micro-pyramids [62], and nanowires [50,51], and also directly controlling the shape of the dot via etching techniques to generate four individual emitting elements with different polarisation at the same time [63]. Recently, it has also been shown that the use of InGaN QDs formed in a nonpolar crystal plane also leads to polarised single photon emission [48,49] along the crystal m-axis. Such structures, in combination with a rotating polarisation element, would also be sufficient for implementation of BB84 QKD.

As the III-nitride material system is becoming an increasingly important technological material system, growth techniques are constantly being improved. These have led to the aforementioned development of single photon LEDs, site-controlled nanostructures and control over the polarisation of the single photon emission. However, there are still a few drawbacks in the use of III-nitride QDs for single photon emission that we discuss below. One particular drawback is that, whilst emission has been achieved in the UV and visible region of the spectrum, IR single photon emission has yet to be realised. In particular, single photon emission in the telecoms band would be required for the development of QKD systems using optical

fibers. Moreover, due to high defect densities in the material system, and typical strong exciton permanent dipole moments in III-nitride-based QDs, the emission from these structures exhibits a large degree of spectral diffusion. This diffusion can act on long time scales [64,65], whereby the emission line is observed to move over time, and also on fast timescales (faster than is measureable with typical spectrometers), such that the measured emission linewidth is broadened [66,67]. The timescales of these fast fluctuations have recently been measured to be ~10ns in GaN QDs, and around 200ns in InGaN QDs [68,69]. In the case of GaN QDs, the linewidth broadening can vary from a few 100s of μeV to a few meV depending on the emission energy [66,70,71]. In either case, the spectral instability of the emission energy will have detrimental effects on the possibility of generating indistinguishable photons (as each emitted photon has a different frequency). Recently, this has been somewhat overcome via the development of GaN interface fluctuation quantum dots that have been reported to suppress the spectral diffusion effect due to a cleaner environment of the QD. Indeed, using such QDs, linewidths lower than 100 μeV have been observed [72]. These structures also exhibit relatively clean emission spectra, leading to single photon emission with comparatively high degrees of purity: $g^{(2)}(0) \sim 0.02$.

5. Single photon sources based on defects in nitrides

Recently, a new alternative to the use of QDs for single photon emitters has emerged: the exploitation of emission from defects. Point defects in wide bandgap semiconductors, for example, NV centres in diamond, have long been known to act as two-level systems and hence as single photon emitters [73]. However, until recently, such a phenomenon had not been observed in GaN, a material in which point defects are very poorly understood. Berhane et al. [74] however, succeeded in identifying single photon emitters in a range of GaN samples, grown on both sapphire and SiC. These initial studies demonstrated single photon emission at wavelengths from 640 to 740 nm. Later studies [75] identified emitters with wavelengths into the IR, between 1085 and 1340 nm, accessing one of the key windows for communication using optical fibre.

These new defect-based emitters have a range of impressive properties [76]. They exhibit emission at room temperature, with raw $g^{(2)}(0)$ values

for the IR emitters as low as 0.05. They are also notably bright, again at room temperature, achieving count rates in excess of 700,000 s^{-1}. Such high count rates offer the potential of high data transmission rates in QKD, for example. However, it should be noted that count rate is not an absolute measure of the efficiency or emission rate of the emitter, being dependent on the excitation power density, and on the lifetime of the emission process. Comparing lifetimes extracted from time-resolved photoluminescence measurements on the defect emitters with roughly comparable data from QDs suggests that the defects have lifetimes fairly similar to many c-plane QD samples (around 1 to a few ns), and rather longer than nonpolar QDs (both wurtzite and zinc blende structures) or QDs with small extent in the *c*-direction at nanowire apices. Like many nitride QDs, the defect emitters exhibit a high degree of linear polarisation, in excess of 90%. Whilst the number of emitters studied to date is much lower than in the QD case, the available data suggest some tendency for preferential polarisation along the [1–100] direction. One less favourable property which the defect emitters share with QDs is the tendency to spectral diffusion, particularly at fast timescales, resulting in linewidths of a few meV.

Understanding the physical processes underlying spectral diffusion from the defect emitters requires an understanding of the actual structural nature of these defects. This understanding will also be vital if such emitters are to be engineered deterministically and incorporated into cavity structures or other devices. Ref. [74] suggested that the emitters consisted of a point defect enclosed in a cubic inclusion, and peak shifts in the near band edge cathodoluminescence peak towards the lower zinc blende bandgap were suggested to support this contention. Such peak shifts might arise for other reasons, however, and it has not yet been possible to find definitive evidence for the cubic inclusions such as TEM images of such defects anywhere in the relevant samples, or X-ray diffraction data illustrating the presence of cubic material. Indeed, stacking faults and cubic inclusions are rarely observed in high-quality hexagonal GaN, and since the emitters appear to occur rather commonly in a range of GaN samples, this casts some doubt upon this structural model. Attempts to correlate these emitters with more common defects, such as dislocations, have yet to allow any further understanding of their origin [77]. Such an understanding is a key next step in this rather new sub-field, and once it is achieved may allow the engineering of GaN defect-based devices which compete not only with GaN QD single photon emitters but with devices engineered from the more conventional red and IR-emitting III-Vs.

6. Outlook

Nitride-based single photon sources have a number of advantages, in particular the wide range of emission wavelengths available, the relatively high temperatures at which they operate and the ability to achieve polarised emission straightforwardly from self-assembled structures. However, much remains to be done if they are going to achieve widespread use in QKD or other applications. Although some recent QD-based devices have started to overcome these limitations, looking across the accessible spectral range, the majority of devices exhibit poor degrees of purity of the single photon emission and often also significant spectral diffusion. For a single photon emitting structure to find application outside the laboratory context, it will not only have to overcome these limitations in an experimental photoluminescence setup, but in an electrically pumped device with excellent extraction efficiency, requiring development of not only the straightforward single photon LEDs achieved to date, but high-quality cavity structures which also allow electrical injection.

Significant progress is being made towards such devices in the more developed arsenide system, but in the nitrides, limited options for etching the material, the low conductivity of p-type GaN and a generally less mature set of processing protocols make such achievements very challenging. Nonetheless, the burgeoning international interest in quantum technologies provides strong motivation for continued studies of these fascinating devices. With the recent emergence of single photon emitting defects in the nitrides, providing an alternative to the self-assembled QD systems, the broad nitride materials systems can now provide single photon emission across a very wide range of emission wavelengths, from the IR to the UV. The exciting properties of such sources coupled with further progress in materials growth and processing may inspire future technologies as the development of quantum systems continues to accelerate.

References

[1] Lounis B, Orrit M. Single-photon sources. Rep Prog Phys 2005;68(5):1129. https://doi.org/10.1088/0034-4885/68/5/R04.

[2] Buckley S, Rivoire K, Vuckovic J. Engineered quantum dot single-photon sources. Rep Prog Phys 2012;75(12):126503. https://doi.org/10.1088/0034-4885/75/12/126503.

[3] Miyazawa T, Takemoto K, Nambu Y, Miki S, Yamashita T, Terai H, et al. Single-photon emission at 1.5 μm from an InAs/InP quantum dot with highly suppressed multi-photon emission probabilities. Appl Phys Lett 2016;109(13):132106-1–132106-4. https://doi.org/10.1063/1.4961888.

[4] Bauer E. Phänomenologische theorie der Kristallabscheidung an Oberflächen. Z Kristallogr 1958;110:372.
[5] Woodruff D, Delchar T. Modern techniques of surface science. 2nd ed. Cambridge, GBR: Cambridge University Press; 1994.
[6] Shchukin V, Ledentsov N, Bimberg D. Epitaxy of nanostructures. Berlin: Springer; 2011.
[7] Adelmann C, Daudin B, Oliver R, Briggs G, Rudd R. Nucleation and growth of GaN/AlN quantum dots. Phys Rev B 2004;70(12):125427. https://doi.org/10.1103/PhysRevB.70.125427.
[8] Williams R, Medeiros-Ribeiro G, Kamins T, Ohlberg D. Thermodynamics of the size and shape of nanocrystals: epitaxial Ge on Si(001). Annu Rev Phys Chem 2000;51(1):527–51. https://doi.org/10.1021/ar970236g.
[9] Ross F, Tersoff J, Tromp R. Coarsening of self-assembled Ge quantum dots on Si(001). Phys Rev Lett 1998;80(5):984–7. https://doi.org/10.1103/PhysRevLett.80.984.
[10] Rudd R, Briggs G, Sutton A, Medeiros-Ribeiro G, Williams R. Equilibrium model of bimodal distributions of epitaxial island growth. Phys Rev Lett 2003;90(14):146101. https://doi.org/10.1103/PhysRevLett.90.146101.
[11] Daudin B, Widmann F, Feuillet G, Samson Y, Arlery M, Rouvière J. Stranski-Krastanov growth mode during the molecular beam epitaxy of highly strained GaN. Phys Rev B 1997;56(12):R7069–72. https://doi.org/10.1103/PhysRevB.56.R7069.
[12] Deshpande S, Frost T, Hazari A, Bhattacharya P. Electrically pumped single-photon emission at room temperature from a single InGaN/GaN quantum dot. Appl Phys Lett 2014;105(14):141109. https://doi.org/10.1063/1.4897640.
[13] Tachibana K, Someya T, Arakawa Y. Nanometer-scale InGaN self-assembled quantum dots grown by metalorganic chemical vapor deposition. Appl Phys Lett 1999;74(3):383–5. https://doi.org/10.1063/1.123078.
[14] van der Laak NK, Oliver RA, Barnard JS, Cherns PD, Kappers MJ, Humphreys CJ. Towards a better understanding of nano-islands formed during atmospheric pressure MOVPE. Phys Status Solidi C 2006;3(6):1544–7. https://doi.org/10.1002/pssc.200565164.
[15] Griffiths J, Zhu T, Oehler F, Emery R, Fu W, Reid B, Taylor R, Kappers M, Humphreys C, Oliver R. Growth of non-polar (11-20) InGaN quantum dots by metal organic vapour phase epitaxy using a two temperature method. APL Mater 2014;2(12):126101. https://doi.org/10.1063/1.4904068.
[16] Simeonov D, Feltin E, Demangeot F, Pinquier C, Carlin J, Butté R, Frandon J, Grandjean N. Strain relaxation of AlN epilayers for Stranski–Krastanov GaN/AlN quantum dots grown by metal organic vapor phase epitaxy. J Cryst Growth 2007;299(2):254–8. https://doi.org/10.1016/j.jcrysgro.2006.12.005.
[17] Pretorius A, Yamaguchi T, Kübel C, Kröger R, Hommel D, Rosenauer A. TEM analyses of wurtzite InGaN islands grown by MOVPE and MBE. Phys Status Solidi C 2006;3(6):1679–82. https://doi.org/10.1002/pssc.200565333.
[18] Zhu T, Oliver R. Nitride quantum light sources. EPL (Europhys Lett) 2016;113(3):38001. https://doi.org/10.1209/0295-5075/113/38001.
[19] Wu J, Wang Z. Droplet epitaxy for advanced optoelectronic materials and devices. J Phys Appl Phys 2014;47(17):173001. https://doi.org/10.1088/0022-3727/47/17/173001.
[20] Kawasaki K, Yamazaki D, Kinoshita A, Hirayama H, Tsutsui K, Aoyagi Y. GaN quantum-dot formation by self-assembling droplet epitaxy and application to single-electron transistors. Appl Phys Lett 2001;79(14):2243–5. https://doi.org/10.1063/1.1405422.

[21] Kumar M, Roul B, Bhat TN, Rajpalke MK, Sinha N, Kalghatgi AT, Krupanidhi SB. Droplet epitaxy of InN quantum dots on Si(111) by RF plasma-assisted molecular beam epitaxy. Adv Sci Lett 2010;3(4):379–84. https://doi.org/10.1166/asl.2010.1163.
[22] Oliver RA, Briggs GAD, Kappers MJ, Humphreys CJ, Yasin S, Rice JH, Smith JD, Taylor RA. InGaN quantum dots grown by metalorganic vapor phase epitaxy employing a post-growth nitrogen anneal. Appl Phys Lett 2003;83(4):755. https://doi.org/10.1063/1.1595716.
[23] Springbett H, Griffiths J, Ren C, O'Hanlon T, Barnard J, Sahonta S, Zhu T, Oliver R. Structure and composition of non-polar (11-20) InGaN nanorings grown by modified droplet epitaxy (Phys. Status Solidi B 5/2016). Phys Status Solidi B 2016;253(5):793. https://doi.org/10.1002/pssb.201670530.
[24] Tanaka S, Suemune I, Ramvall P, Aoyagi Y. GaN quantum structures with fractional dimension — from quantum well to quantum dot. Phys Status Solidi B 1999;216(1):431–4. https://doi.org/10.1002/(SICI)1521-3951(199911)216:1<431::AID-PSSB431>3.0.CO;2-3.
[25] Chen Z, Lu D, Yuan H, Han P, Liu X, Li Y, Wang X, Lu Y, Wang Z. A new method to fabricate InGaN quantum dots by metalorganic chemical vapor deposition. J Cryst Growth 2002;235(1–4):188–94. https://doi.org/10.1016/S0022-0248(01)02091-7.
[26] Tessarek C, Figge S, Aschenbrenner T, Bley S, Rosenauer A, Seyfried M, Kalden J, Sebald K, Gutowski J, Hommel D. Strong phase separation of strained InxGa1−xN layers due to spinodal and binodal decomposition: formation of stable quantum dots. Phys Rev B 2011;83(11). https://doi.org/10.1103/PhysRevB.83.115316.
[27] Wagner R, Ellis W. Vapor-liquid-solid mechanism of single crystal growth. Appl Phys Lett 1964;4(5):89–90. https://doi.org/10.1063/1.1753975.
[28] Yazawa M, Koguchi M, Hiruma K. Heteroepitaxial ultrafine wire-like growth of InAs on GaAs substrates. Appl Phys Lett 1991;58(10):1080–2. https://doi.org/10.1063/1.104377.
[29] Chen C, Yeh C, Chen C, Yu M, Liu H, Wu J, Chen K, Chen L, Peng J, Chen Y. Catalytic growth and characterization of gallium nitride nanowires. J Am Chem Soc 2001;123(12):2791–8. https://doi.org/10.1021/ja0040518.
[30] Stach E, Pauzauskie P, Kuykendall T, Goldberger J, He R, Yang P. Watching GaN nanowires grow. Nano Lett 2003;3(6):867–9. http://dor.org/10.1021/nl034222h.
[31] Bertness K, Roshko A, Mansfield L, Harvey T, Sanford N. Mechanism for spontaneous growth of GaN nanowires with molecular beam epitaxy. J Cryst Growth 2008;310(13):3154–8. https://doi.org/10.1016/j.jcrysgro.2008.03.033.
[32] Renard J, Songmuang R, Bougerol C, Daudin B, Gayral B. Exciton and biexciton luminescence from single GaN/AlN quantum dots in nanowires. Nano Lett 2008;8(7):2092–6. http://pubs.acs.org/doi/abs/10.1021/nl0800873.
[33] Li S, Waag A. GaN based nanorods for solid state lighting. J Appl Phys 2012;111(7):071101. https://doi.org/10.1063/1.3694674.
[34] Le Boulbar E, Edwards P, Vajargah S, Griffiths I, Gîrgel I, Coulon P, Cherns D, Martin R, Humphreys C, Bowen C, Allsopp D, Shields P. Structural and optical emission uniformity of m-plane InGaN single quantum wells in core–shell nanorods. Cryst Growth Des 2016;16(4):1907–16. https://doi.org/10.1021/acs.cgd.5b01438.
[35] Hsu C-W, Lundskog A, Karlsson KF, Forsberg U, Janzén E, Holtz PO. Single excitons in InGaN quantum dots on GaN pyramid arrays. Nano Lett 2011;11(6):2415–8. https://doi.org/10.1021/nl200810v.
[36] Holmes M, Kako S, Choi K, Arita M, Arakawa Y. Single photons from a hot solid-state emitter at 350 K. ACS Photonics 2016;3:543. https://doi.org/10.1021/acsphotonics.6b00112.

[37] Zhang L, Teng C, Hill T, Lee L, Ku P, Deng H. Single photon emission from site-controlled InGaN/GaN quantum dots. Appl Phys Lett 2013;103(19):192114. https://doi.org/10.1063/1.4830000.
[38] Arakawa Y, Sakaki H. Multidimensional quantum well laser and temperature dependence of its threshold current. Appl Phys Lett 1982;40(11):939. https://doi.org/10.1063/1.92959.
[39] Tomić S, Pal J, Migliorato MA, Young RJ, Vukmirović N. Visible spectrum quantum light sources based on $In_xGa_{1-x}N$/GaN quantum dots. ACS Photonics 2015;2(7):958–63. https://doi.org/10.1021/acsphotonics.5b00159.
[40] Kako S, Miyamura M, Tachibana K, Hoshino K, Arakawa Y. Size-dependent radiative decay time of excitons in GaN/AlN self-assembled quantum dots. Appl Phys Lett 2003;83(5):984. https://doi.org/10.1063/1.1596382.
[41] Chuang SL, Chang CS. k–p method for strained wurtzite semiconductors. Phys Rev B 1996;4:2491.
[42] Andreev AD, O'Reilly EP. Theory of the electronic structure of GaN/AlN hexagonal quantum dots. Phys Rev B 2000;62(23):15851.
[43] Santori C, Götzinger S, Yamamoto Y, Kako S, Hoshino K, Arakawa Y. Photon correlation studies of single GaN quantum dots. Appl Phys Lett 2005;87(5):051916.
[44] Jarjour AF, Taylor RA, Oliver RA, Kappers MJ, Humphreys CJ, Tahraoui A. Cavity-enhanced blue single-photon emission from a single InGaN/GaN quantum dot. Appl Phys Lett 2007;91(5):052101.
[45] Choi K, Kako S, Holmes M, Arita M, Arakawa Y. Strong exciton confinement in site-controlled GaN quantum dots embedded in nanowires. Appl Phys Lett 2013;103(17):171907. https://doi.org/10.1063/1.4826931.
[46] Kako S, Santori C, Hoshino K, Götzinger S, Yamamoto Y, Arakawa Y. A gallium nitride single-photon source operating at 200 K. Nat Mater 2006;5(11):887–92.
[47] Holmes M, Choi K, Kako S, Arita M, Arakawa Y. Room-temperature triggered single photon emission from a III-nitride site-controlled nanowire quantum dot. Nano Lett 2014;14(2):982–6.
[48] Wang T, Puchtler TJ, Patra SK, Zhu T, Jarman JC, Oliver RA, et al. Deterministic optical polarisation in nitride quantum dots at thermoelectrically cooled temperatures. Sci Rep 2017:1–9. https://doi.org/10.1038/s41598-017-12233-6.
[49] Wang T, Puchtler TJ, Zhu T, Jarman JC, Nuttall LP, Oliver RA, Taylor RA. Polarisation-controlled single photon emission at high temperatures from InGaN quantum dots. Nanoscale 2017;10:631–7.
[50] Deshpande S, Das A, Bhattacharya P. Blue single photon emission up to 200 K from an InGaN quantum dot in AlGaN nanowire. Appl Phys Lett 2013;102(16):161114. https://doi.org/10.1063/1.4803441.
[51] Deshpande S, Heo J, Das A, Bhattacharya P. Electrically driven polarized single-photon emission from an InGaN quantum dot in a GaN nanowire. Nat Commun 2013;4:1675–8. https://doi.org/10.1038/ncomms2691.
[52] Hönig G, Callsen G, Schliwa A, Kalinowski S, Kindel C, Kako S, et al. Manifestation of unconventional biexciton states in quantum dots. Nat Commun 2014;5:5721. https://doi.org/10.1038/ncomms6721.
[53] Zhu T, Oehler F, Reid BPL, Emery RM, Taylor RA, Kappers MJ, Oliver RA. Nonpolar (11-20) InGaN quantum dots with short exciton lifetimes grown by metal-organic vapor phase epitaxy. Appl Phys Lett 2013;102(25):251905. https://doi.org/10.1063/1.4812345.
[54] Gačević Ž, Holmes M, Chernysheva E, Müller M, Torres-Pardo A, Veit P, et al. Emission of linearly polarized single photons from quantum dots contained in nonpolar, semipolar, and polar sections of pencil-like InGaN/GaN nanowires. ACS Photonics 2017;4(3):657–64. https://doi.org/10.1021/acsphotonics.6b01030.

[55] Daudin B. Polar and nonpolar GaN quantum dots. J Phys Condens Matter 2008; 20(47):473201. https://doi.org/10.1088/0953-8984/20/47/473201.
[56] Martinez-Guerrero E, Adelmann C, Chabuel F, Simon J, Pelekanos NT, Mula G, et al. Self-assembled zinc blende GaN quantum dots grown by molecular-beam epitaxy. Appl Phys Lett 2000;77(6):809. https://doi.org/10.1063/1.1306633.
[57] Kako S, Holmes M, Sergent S, Bürger M, As DJ, Arakawa Y. Single-photon emission from cubic GaN quantum dots. Appl Phys Lett 2014;104(1):011101. https://doi.org/10.1063/1.4858966.
[58] Sergent S, Kako S, Bürger M, As DJ, Arakawa Y. Narrow spectral linewidth of single zinc-blende GaN/AlN self-assembled quantum dots. Appl Phys Lett 2013;103(15): 151109. https://doi.org/10.1063/1.4824650.
[59] Hönig GMO, Westerkamp S, Hoffmann A, Callsen G. Shielding electrostatic fields in polar semiconductor nanostructures. Phys Rev Appl 2017;7(2):024004–12. https://doi.org/10.1103/PhysRevApplied.7.024004.
[60] Bardoux R, Guillet T, Gil B, Lefebvre P, Bretagnon T, Taliercio T, et al. Polarized emission from GaN/AlN quantum dots: single-dot spectroscopy and symmetry-based theory. Phys Rev B 2008;77(23):235315. https://doi.org/10.1103/PhysRevB.77.235315.
[61] Bennett CH, Brassard G. Quantum cryptography: public key distribution and coin tossing. In: Presented at the proc. of IEEE int. conf. on computers, systems and signal processing; 1984. p. 175–9.
[62] Lundskog A, Hsu C-W, Fredrik Karlsson K, Amloy S, Nilsson D, Forsberg U, et al. Direct generation of linearly polarized photon emission with designated orientations from site-controlled InGaN quantum dots. Light Sci Appl 2014;3(1):e139. https://doi.org/10.1038/lsa.2014.20.
[63] Teng C-H, Zhang L, Hill TA, Demory B, Deng H, Ku P-C. Elliptical quantum dots as on-demand single photons sources with deterministic polarization states. Appl Phys Lett 2015;107(19):191105. https://doi.org/10.1063/1.4935463.
[64] Rice JH, Robinson JW, Jarjour A, Taylor RA, Oliver RA, Briggs GAD, et al. Temporal variation in photoluminescence from single InGaN quantum dots. Appl Phys Lett 2004;84(20):4110. https://doi.org/10.1063/1.1753653.
[65] Bardoux R, Guillet T, Lefebvre P, Taliercio T, Bretagnon T, Rousset S, et al. Photoluminescence of single GaN/AlN hexagonal quantum dots on Si(111): spectral diffusion effects. Phys Rev B 2006;74(19):195319. https://doi.org/10.1103/PhysRevB.74.195319.
[66] Kindel C, Callsen G, Kako S, Kawano T, Oishi H, Honig G, et al. Spectral diffusion in nitride quantum dots: emission energy dependent linewidths broadening via giant built-in dipole moments. Phys Status Solidi Rapid Res Lett 2014;8(5):408–13. https://doi.org/10.1002/pssr.201409096.
[67] Holmes M, Kako S, Choi K, Arita M, Arakawa Y. Spectral diffusion and its influence on the emission linewidths of site-controlled GaN nanowire quantum dots. Phys Rev B 2015;92(11):115447. https://doi.org/10.1103/PhysRevB.92.115447.
[68] Gao K, Solovev I, Holmes M, Arita M, Arakawa Y. Nanosecond-scale spectral diffusion in the single photon emission of a GaN quantum dot. AIP Adv 2017;7(12): 125216. https://doi.org/10.1063/1.4997117.
[69] Gao K, Holmes M, Arita M, Arakawa Y. Measurement of the emission lifetime of a GaN interface fluctuation quantum dot by power dependent single photon dynamics. Phys Status Solidi A 2018;215(9):1700630. https://doi.org/10.1002/pssa.201700630.
[70] Schmidt G, Berger C, Veit P, Metzner S, Bertram F, Bläsing J, et al. Direct evidence of single quantum dot emission from GaN islands formed at threading dislocations using nanoscale cathodoluminescence: a source of single photons in the ultraviolet. Appl Phys Lett 2015;106(25):252101. https://doi.org/10.1063/1.4922919.

[71] Demangeot F, Simeonov D, Dussaigne A, Butté R, Grandjean N. Homogeneous and inhomogeneous linewidth broadening of single polar GaN/AlN quantum dots. Phys Status Solidi 2009;6(S2):S598–601. https://doi.org/10.1002/pssc.200880971.
[72] Arita M, Le Roux F, Holmes M, Kako S, Arakawa Y. Ultraclean single photon emission from a GaN quantum dot. Nano Lett 2017;17(5):2902–7. https://doi.org/10.1021/acs.nanolett.7b00109.
[73] Kurtsiefer C, Mayer S, Zarda P, Weinfurter H. Stable solid-state source of single photons. Phys Rev Lett 2000;85(2):290–3. https://doi.org/10.1103/PhysRevLett.85.290.
[74] Berhane A, Jeong K, Bodrog Z, Fiedler S, Schröder T, Triviño N, Palacios T, Gali A, Toth M, Englund D, Aharonovich I. Bright room-temperature single-photon emission from defects in gallium nitride. Adv Mater 2017;29(12):1605092. https://doi.org/10.1002/adma.201605092.
[75] Zhou Y, Wang Z, Rasmita A, Kim S, Berhane A, Bodrog Z, Adamo G, Gali A, Aharonovich I, Gao W. Room temperature solid-state quantum emitters in the telecom range. Sci Adv 2018;4(3):eaar3580. https://doi.org/10.1126/sciadv.aar3580.
[76] Berhane A, Jeong K, Bradac C, Walsh M, Englund D, Toth M, Aharonovich I. Photophysics of GaN single-photon emitters in the visible spectral range. Phys Rev B 2018; 97(16). https://doi.org/10.1103/PhysRevB.97.165202.
[77] Nguyen M, Zhu T, Kianini M, Massabuau F, Aharonovich I, Toth M, Oliver R, Bradac C. Effects of microstructure and growth conditions on quantum emitters in gallium nitride. APL Mater 2019;7:081106. https://doi.org/10.1063/1.5098794.
[78] Okamoto H, Massalski TB. The Au−Si (Gold-Silicon) system. Bull Alloy Phase Diagr 1983;4(2):190–8.
[79] Griffiths JT, Zhu T, Oehler F, Emery RM, Fu WY, Reid BPL, Taylor RA, Kappers MJ, Humphreys CJ, Oliver RA. Growth of non-polar (11-20) InGaN quantum dots by metal organic vapour phase epitaxy using a two temperature method. APL Materials 2014;2:126101. https://doi.org/10.1063/1.4904068.

Index

'*Note:* Page numbers followed by "f" indicate figures and "t" indicate tables.'

A

Absorption noise, 162–163
Aharonov-Bohm (AB) effects, 247, 258–260, 259f
Amplified spontaneous emission, 401
Andreev bound states, 254–256
Angle resolved photoemission spectroscopy (ARPES), 243
Anti-phase boundaries (APBs), 354
Array lasers, 401
Asymmetric quantum interference device (SQUID), 256
Atomic point contacts (APCs), 177
Axial quantum dot nanowire lasers, 418

B

Backward scattering, 34–35
Ballistic transport, 252–253, 252f
Bloch's theory of sound waves, 32
Bosonisation method, 36

C

Capacitor charging experiment, 123–125
Carrier density variation, 21–22, 22f
Carr–Purcell–Meiboom–Gill sequence, 278
C-band QD lasers, 368–369
Chalcogenide, 412–413
Charge neutrality point (CNP), 244–245
Charge sensors, 283
Chip-to-chip communication, 390
Chiral Luttinger liquid theory, 200–201
Cleaved-edge overgrowth (CEO), 47
Cold atoms, 63–64
Complementary metal – oxide – semiconductor (CMOS) circuits, 265
Complementary metal oxide semiconductor (CMOS), 353
Coulomb blockade thermometer, 93
Coulomb drag, 62–63
Coulomb interaction, 8, 14
Coulomb oscillations, 73–74
Cryogenic capacitor, 124–125
Cryogenic microwave amplifiers, 171
Current heating technique, 75–76

D

1D drift-diffusion method, 322
1D interaction effects
  cold atoms, 63–64
  Coulomb drag, 62–63
  helical current, 63
  linear Tomonaga–Luttinger liquid (TLL) model. *See* Tomonaga–Luttinger liquid (TLL)
  Nonlinear Tomonaga–Luttinger liquid (TLL) approximation. *See* Nonlinear Tomonaga–Luttinger liquid (TLL) approximation
DC shot noise
  annihilation operator, 157
  conductance measurements, 161
  current fluctuations, 159
  current operator, 158–159
  electron kinetic energy, 156
  fermion operator, 158
  finite temperature low frequency shot noise, 160, 161f
  general conductor, 156f
  partition noise, 160
  scattering matrix, 157
  scattering state, 157
  scattering states, 158
  thermal noise, 160
Dense wavelength division multiplexing (DWDM) systems, 371
Device volume, lasers, 404
Dorokhov distribution, 153–154
Dynamic structure factor (DSF), 62

E

Electric dipole spin resonance (EDSR), 277–278
Electric field control, quantum dots
  charged transitions
    energy splitting, 305–306
    Voigt geometry g-factors, 306–307, 306f
  heterostructures, 295
    arsenides, 296
    device designs, 298f
    electron (hole) charging device, 299
    emission spectrum, 299
    layer sequences and band structures, 297
    photoluminescence *vs.* electric field plots, 298f
    photon collection efficiency, 300–302
    quadratic energy level dependence, 300
    Quantum Confined Stark Effect, 299–300
    Schottky contacts, 299
    self-assembled InGaAs quantum dots, 295–296
    tunnelling rates, electrons and holes, 296f
    two diode heterostructures, 296–297
  hybrid approaches
    coherent optical control, 309–310, 310f
    strain and electric field tuneable quantum dots, 311, 311f
    tuneable electroluminescent quantum dot light source, 308, 309f
  neutral transitions
    entangled photon emission, 305f
    excitonic eigenstates, 304
    fine structure splitting (FSS), 302–303
    s-shell emission, 302
  planar-micro-cavity LEDs, 301
Electron beam lithography, 3
Electron hole emission, 117, 118f
Electron quantum optics
  de-coherence and relaxation
    capacitive interactions, 138
    charge fractionalisation, 138, 139f
    electron transport, 139
    energy spectroscopy, 140f
    hot-electron wave packets, 140f
    Mach–Zehnder interferometer, 139–140, 141f
  experimental setups, 129
  indistinguishability test, 131–134, 133f
  partition noise measurements
    cross-correlation, 132f
    current splitting, 130
    experimental configuration, 131f
    Hanbury Brown and Twiss effect, 130–131
  quantum tomography
    experimental setup, 136f
    fermionic statistics, 135–137
    leviton excitation, 137
    off-diagonal components, 137
    partially transmitting gates, 134
    projected distributions, 135
    quadrature amplitude, 135
    Wigner function, 135, 138f
  single-electron emission, 129
Electron relaxation models, 328–329
Electron transport, 4
Emission noise, 162
Entangled photon emission, electric field tuneable device, 304–305, 305f
Exchange-driven spin repulsion, 12
Excitonic gain, 398–399
External quantum efficiency (EQE) spectra, 335–336, 336f

F

Fabry–Pérot edge-emitting lasers
  C-band QD lasers, 368–369
  high-performance QD lasers, 364–365
  O-band QD lasers, 365–368
Fano factor, 152–153, 155
Fermi liquid theory failure, 31–32
Fermi–Dirac distribution, 69
Field-damping layer (FDL), QDSCs
  band structure engineering
    calculated energy band structures, 335f
    electrical characteristics, 333–335
    equilibrium energy band structures, 334f

external quantum efficiency (EQE) spectra, 335–336
field-damping layer (FDL), 331–332, 332f
light current–voltage curves, 334f
photocurrent production, 336–338, 337f
wide-gap material capping
AlAsSb fence structure, 340–341
InGaAs QDs, 340f
light current-voltage curve, 338–339, 338f, 341f
wide-gap fence structure, 339–340, 339f
Fine structure splitting (FSS), 302–303, 303f, 456
Finite frequency shot noise
absorption noise, 162–163
emission noise, 162
energy independent transmission, 161–162
high-frequency regime, 161
photon Bose–Einstein distribution, 162
Forward scattering, 34–35
Fourier transform infrared (FTIR) photocurrent spectroscopy, 321
benefits, 342–343
two-step photo-absorption (TSPA) spectroscopy
IR absorption edge fitting analysis, 344, 344f
IR light power intensity spectrum, 342–343
low temperature IR photocurrent spectra, 343–344
spectroscopy setup, 342–343, 343f
temperature dependence, 343f, 344, 345f
universal linear relationship, 345–346, 345f
Fractional quantum Hall effect (FQHE), 2, 7–8, 193–195
edge channels
backscattering, 198f
chiral Luttinger liquid theory, 200–201
cyclotron orbit, 196
finite energy, 196–198
fractional charges, 199–200
Landau level, 196–198
strong backscattering regime (SB), 198
uncertainty relation, 196
low frequency DC shot noise measurements, 201–206
auto-correlation noise measurements, 201
conductance *versus* bias voltage, 204–205
differential conductance, 202
empirical approach, 204
filling factor, 205–206
Luttinger approach, 202
quasi-particle charge, 202
Nyquist bandwidth, 195
photo-assisted shot noise and high frequency shot noise, 207–210
variance, 195
Full Counting Statistics (FCS), 151–152

G

Gate-based quantum computation, 267
Green's function, 39

H

Hahn-echo sequence, 278
Half-integer QH effect, 245–247, 245f
Hall conductivity, 245–247
Hanbury Brown and Twiss (HBT) experiment, 212f
electron quantum optics, 218–223
two-electron collision experiment, 211–213
two-particle coincidence probability, 213–215
two-particle interference, 213, 214f
Heavy impurity model, 53–54
Heisenberg interaction, 278–279
Helical current, 63
High electron mobility transistor (HEMT), 168–169
Hikami–Larkin–Nagaoka formula, 244–245
Hong Ou Mandel experiments, 223–227
Hong–Ou–Mandel effect, 131–134, 133f
Hot-electron wave packets, 140f

I

III–V lasers employing quantum dots (QDs), 354–355
Incipient Wigner crystal, 14–16
In-plane magnetic field, 18–19, 22–23
Inverse photoemission spectroscopy, 42
I-V characteristics
  three-junction pump, 106f
  turnstile quantum dot single-electron transfer, 103–104, 104f

J

Johnson Nyquist noise, 147–148, 169f
Josephson effect, 122

K

Kondo effect, 7–8

L

Landauer-Büttiker formalism
  linear response, 67
  Lorenz ratio, 68
  Onsager relation, 67
  Peltier coefficient, 68
  thermal conductance, 68
  thermopower, 68
Laser centre wavelength, 404
Lasers. *See also* Semiconductor nanowire lasers
  array lasers, 401
  cavity and waveguides
    cavity length, 393
    confinement factor, 394–396
    refractive index, 393–394
    single homogeneous gain material, 394–396
    transverse modes, 394, 395f
  density of states engineering
    excitonic gain, 398–399
    quantum dots, 399
    quantum wells, 399
    three-dimensional, 397–398
  gain material, 396–397
  nanoscale laser cavity design, 399–401
  propagation constant, 392
  threshold
    coupled rate equation analysis, 402–403
    figures of merit, 403–404
    intrinsic spontaneous emission rate, 403
    lasing, 401–402
    photon density, 402–403
    photons, 402f
    rate equation model, 402f
    spontaneous emission, 401
  threshold gain, 393
Leviton excitation
  electron and hole excitations, 119–120, 120f
  Fermi energy, 119–120
  Lorentzian-shaped pulses, 117–119
  voltage pulse, 117–119
Leviton Single Electron Source (L-SES), 216–217, 218f
Lithographic approaches, 454–455
Longitudinal resistance, 245f
Long-period oscillations, 74
Lorenz ratio, 68, 81
Low frequency shot noise
  diffusive conductors
    coherent regime, 180–182
    conductivity, 183–184
    electron–electron interaction, 182–183
    evanescent mode transmissions, 184
    Fano factor, 182–185, 183f, 185f
  and full counting statistics, 187–189
  large quantum dots, 185–187
Lyapunov exponent, 186

M

Mach–Zehnder interferometer, 139–140, 141f
Magnetoconductance, 249–251, 250f
Many-body interactions, 31
Maximum operating temperature, laser, 404
Mesoscopic capacitor, 117
  electron hole emission, 117, 118f
  time-domain measurements, 117
Mesoscopic capacitor single electron source (MC-SES), 216–217, 218f–219f

Mesoscopic solid-state physics, 90
Metal split-gate technique, 3
Metrological triangle, 122–123, 123f
Micro-cavity lasers, 373–376, 378
Micro-disk (MD) lasers, 374–375
Micro-ring lasers, 375–376
Misfit dislocations (MDs), 359
Mobile impurity model
  1D spectral function, 53–54
  heavy impurity model, 53–54
  momentum-dependent power law, 54
  perturbation analysis, 52
  plasmon's lifetime, 54
  spinful fermions, 53–54
  X-ray scattering, 52
Mode-locked lasers (MLLs), 372–373
Molecular beam epitaxy (MBE) system, 2, 362–363
Momentum-dependent power law, 58–60, 59f
Monolithic III-V quantum dot lasers, silicon
  Fabry–Pérot edge-emitting lasers
    C-band QD lasers, 368–369
    high-performance QD lasers, 364–365
    O-band QD lasers, 365–368
  heteroepitaxial growth
    APBs formation, 358–359
    misfit dislocations (MDs), 359
  high-quality III-V/Si epitaxy
    APB-free III–V growth, 359–363
    buffer layer, MOCVD/MOVPE, 360–362
    molecular beam epitaxy (MBE) system, 362–363
    patterned Si substrates, 360
    thermal cracks, 364
    threading dislocation density reduction, 363–364
  micro-cavity lasers, 373–376
    micro-disk (MD) lasers, 374–375
    micro-ring lasers, 375–376
  photonic crystal lasers, 376–377
  QD mode-locked lasers, 372–373
  single mode QD edge-emitting lasers
    QD distributed-feedback lasers, 369–371
    QD wavelength tunable lasers, 371–372
Moore's law, 265
Mott relation, 75–76
Multi-junction single-electron pump, 104–105, 105f

N

Nanodevices, 1
Nanowire transistor device, 275f
National Institute of Standards and Technology (NIST)
  seven-junction pump, 124f
Neutral transitions, electric field control
  entangled photon emission, 305f
  excitonic eigenstates, 304
  fine structure splitting (FSS), 302–303
  s-shell emission, 302
Nitride single photon sources
  confinement potentials, 461–462
  defect-based emitters, 464–465
  emission wavelength control, 461
  III-nitride semiconductors
    band bending and wavefunction separation, 459
    optical properties, 458–459
    permanent dipole moment, 459
    Wurtzite crystal structure, 458–459
  internal electric field, 462–463
  polarisation, 463
  quantum dot fabrication
    droplet epitaxy, 450–451
    elastic relaxation, 446f
    energetics of island growth, 448–449
    first principles theory, 449
    GaN/InGaN layer, 451
    growth modes, 446
    MOVPE, 449–450
    quantum wells, 443–455
    reflection high-energy electron diffraction (RHEED), 445–446
    self-assembly, planar surfaces, 444–451
    SK growth, 450
    thin film growth, 447
    three-dimensional island growth modes, 447–448, 448f

Nitride single photon sources (*Continued*)
top-down and a bottom-up process, 443–444
two-dimensional growth modes, 444–445, 445f
quantum dot formation, 454–455, 455f
self-assembly
nanorods, 452–453, 452f
planar surfaces, 444–451
strengths and weaknesses, 460–464
technological material system, 463–464
Noise measurement techniques. *See also* Shot noise measurement
high frequency shot noise measurement
cryogenic microwave amplifiers, 171
emission noise, 171–172
high frequency regime, 170–171
impedance transformation, 172–173
Johnson–Nyquist noise calibration, 169f
Kubo formula, 171
noise power, 170–171
noise temperature, 171
low frequency shot noise measurement techniques
amplifier noise temperature, 166
auto-correlation, 168
coherent regime, 165
commercial low noise amplifiers, 166–167, 166t
cryogenic amplification, 168–169
current noise spectral density, 165
Poissonian noise, 165
resistor, 165
total voltage noise, 165–166
Non-Fermi liquids, 31–32
Nonlinear carbon nanotubes (CNTs), 60–62, 61f
Nonlinear Luttinger plasmons, 61f
Nonlinear Tomonaga–Luttinger liquid (TLL) approximation
carbon nanotubes (CNTs), 60–62, 61f
1D 'replica' modes
higher-order mode, 58
length dependence, 58
mode hierarchy model, 56–57
tunnelling conductance, 57f–58f
mobile impurity model
1D spectral function, 53–54
heavy impurity model, 53–54
momentum-dependent power law, 54
perturbation analysis, 52
plasmon's lifetime, 54
spinful fermions, 53–54
X-ray scattering, 52
mode hierarchy
arbitrary nonlinear excitations, 54
hydrodynamic modes, 56
spectral function, 55
spectral threshold, 56
spinless fermions, 54–55, 55f
momentum-dependent power law, 58–60
properties, 52
spinon and holons lifetime, 60
Non-magnetic fractional states
asymmetry effect, 20–21, 20f
quantum Hall state, 19–20
ring exchange, 21
tight-binding model, 19–20
zigzag phase, 20–21
Normal Fermi liquid, 31

O

O-band QD lasers, 365–368
Odd-peak splitting, 11–12
On-chip photonics, 390
One-dimensional electron system, 32, 32f
Onsager relation, 67
Operational lifetime, laser, 404
Optical coupling, nanowire laser, 418–419
Optoelectronic nanodevices, 4

P

Partition function, 70–71
Pauli exclusion principle, 46
Peltier coefficient, 68, 85–86
Perovskites, 414–415
Perturbation theory, 32
Photo-assisted shot noise (PASN), 155
diffusive metallic wires, 192
low frequency shot noise, 191–192
microwave radiation, 193
nonequilibrium situation, 190–191
origin, 163

periodic phase, 164
physics, 193–195
QPC, 192f
THz radiation detection, 193, 194f
zero frequency, 164
Photonic crystal lasers, 376–377
Photonic integrated circuits (PICs), 353–354
Plasmonic-based lasers, 419–420
Poissonian noise, 165, 173
Power factor, 79f
Power-law behaviour, 39–40
Pulsed nanowire laser systems, 405t–407t

Q

QD distributed-feedback lasers, 369–371
Quantum billiard, 186, 187f
Quantum coherent conductors, 147
Quantum Confined Stark Effect, 299–300
Quantum confinement, 74
Quantum dot
3-dimensional confinement
bi-exciton states, 456–458, 458f
physical size, 455–456
single photon emission, 457f
three-dimensional nature, 456
transition and dead time, 456
formation, 455f
Quantum dot lasers on silicon
vs. quantum wells
defect-tolerable properties, 357
Gaussian energy distribution, 356
linewidth enhancement factor, 356
relative intensity noise (RIN), 356
semiconductor quantum dots, 355–356
Quantum dot, thermoelectric properties
Coulomb oscillations and thermopower
conductance, 72–73, 73f
current heating technique, 75–76
fine structure, 74, 75f
lattice temperature dependence, 76
Mott relation, 75–76
oscillations, 74
quantum confinement, 74
double barriers, 93
energy filters, energy harvesting, 92–93
figure-of-merit, 78–79
dimensionless, 81
electric power, 78–79
quantum limit, 78–80
sawtooth oscillations, 80
heat engine experiment, 90–92, 91f–92f
linear response, 85
low temperature thermometers, 93
model
electron thermal conductance, 69
equilibrium probabilities, 70–71
equivalent circuit, 69, 69f
Fermi–Dirac distribution, 69
heat flux, 70–71
partition function, 70–71
stationary current, 70–71
transport coefficients, 71
multi-terminal heat engines, 94–95
nanoscale conductors, 90
nonlinear regime
average reservoir temperature, 83–84
heat-to-work conversion, 83
Peltier coefficient, 85–86
thermopower, 84–85
two-terminal QD system, 83–84
output power
and efficiency, 87, 89
maximum efficiency, 87–89, 88f
maximum output power, 88f, 89
steady-state conditions, 87
two-terminal system, 87
power factor, 79f
quantum limit
discreteness, 71–72
equidistant energy-level spectrum, 72
nondegenerate energy levels, 72
spin degeneracy, 77–78, 77f, 93–94
Wiedemann–Franz law, 81–83
Quantum dot-based based intermediate band solar cells (QD-IBSCs)
band structure engineering
calculated energy band structures, 335f
electrical characteristics, 333–335
equilibrium energy band structures, 334f
external quantum efficiency (EQE) spectra, 335–336

Quantum dot-based based intermediate band solar cells (QD-IBSCs) (*Continued*)
field-damping layer (FDL), 331–332, 332f
light current–voltage curves, 334f
photocurrent production, 336–338, 337f
drift-diffusion analysis
calculated band diagram, 326f
calculated quantum efficiencies, 327, 328f–330f
current-voltage (J-V) curves, 322, 327f
1D drift-diffusion method, 322
electron relaxation models, 328–329
electron transition rate, 323–324
electrostatic potential, 323
employed device structure, 325–326, 325f
incident photon excitation, 328
optical carrier generation rates, 324
optical generation rates, 324
photon flux density, 324–325
simulation method, 322–326
sun illumination, 330–331
TSPA process, 322
fabrication, 320–321
Fourier transform infrared (FTIR) photocurrent spectroscopy, 321
implementation, 320
schematic structure, 332–333
strain-balanced growth technique, 320–321
structural configuration, 320
sub-bandgap IR photons, 321
sun condition, 321
two-step photo-absorption (TSPA), 320
two-step photo-absorption (TSPA) spectroscopy. *See* Two-step photo-absorption (TSPA) spectroscopy
Quantum Hall effect (QHE), 7–8, 241–242
Quantum key distribution (QKD) schemes, 440
Quantum point contacts (QPCs), 147, 173–175, 175f–176f, 178f
Quantum shot noise, 147–148
binomial statistics
binomial probability distribution, 151–152
conductance quantisation and noiseless current, 150
current, 149–151
diffusive system, 155
Dorokhov distribution, 153–154
Fano factor, 155
Fermi statistics, 149–150
Landauer formula, 151
one-dimensional conductor, 149–150
Pauli principle, 149–150
Schottky noise, 152
shot noise, 154f, 153
single wall carbon nanotube (SWCN), 149
spin degeneracy, 152–153
zero frequency current fluctuations, 152
quantum scattering approach
DC shot noise, 155–161
finite frequency shot noise, 161–163
photon-assisted shot noise, 163–165
Quantum state tomography (QST), 227–230
Quantum wells, 399
Quasi one-dimensional electron gas
carrier density, 8
1D device cooldown
carrier density variation, 21–22, 22f
conductance characteristics, 9–10, 10f
incipient Wigner crystal, 14–16
in-plane magnetic field, 18–19, 18f, 22–23
non-linear transport, fractional states, 23–25
non-magnetic fractional states, 19–21
source drain bias, 16–17
split gate voltage, 9–10
transverse electron focussing, 11–14
GaAs/AlGaAs heterostructures, 9
weakly confined 1D quantum wire, 8

R

Radial quantum-well nanowire lasers, 416–418
Reflection high-energy electron diffraction (RHEED), 445–446

S

Schottky contacts, 299
Schottky noise, 152
Schrieffer–Wolff transformation, 278–279
Selective area epitaxy (SAE) techniques, 454
Semiconductor nanowire lasers
  advantages, 404
  biological study, 390
  chalcogenide, 412–413
  chip-to-chip communication, 390
  Fabry–Pérot laser cavity, 390f
  III-V growth approach, 413–414
  nanowire growth
    bottom-up growth, 408
    growth modes, 408, 408f
    selective area growth, 409
    vapour-liquid-solid growth, 409
  nanowire lasers approach, 389
  nitrides, 411–412
  on-chip photonics, 390
  optical coupling, 418–419
  optical pumping, 390
  organic materials, 415–416
  perovskites, 414–415
  plasmonics, 419–420
  pulsed nanowire laser systems, 405t–407t
  quantum confinement
    axial quantum dot nanowire lasers, 418
    radial quantum-well nanowire lasers, 416–418
  zinc oxide, 410–411
Semiconductor quantum dot solar cells
  nonradiative recombination, 319
  quantum dot-based based intermediate band solar cells (QD-IBSCs). *See* Quantum dot-based based intermediate band solar cells (QD-IBSCs)
  Shockley–Queisser limit, 319–320
Short-period oscillations, 74
Shot noise measurement
  0.7 QPC anomaly, 178–180
    conductance and shot noise, 180
    Fermi statistics, 178–180
    Landauer-Buttiker picture, 178–180
    spin degeneracy, 180
  electron quantum optics
    electronic Hong Ou Mandel experiments, 223–227
    Hanbury Brown and Twiss experiments, 218–223
    on-demand coherent single electron sources, 216–217
    quantum state tomography, 227–230
    SAW-based electron pump, 216f
    single electron pump and sources, 215–218
  fractional quantum hall effect
    edge channels, 195–201
    Nyquist bandwidth, 195
    variance, 195
  high frequency effect
    high frequency quantum suppression, 189–190
    photo-assisted quantum shot noise, 190–195
    quantum detection, 191f
  low frequency shot noise. *See* Low frequency shot noise
  quantum suppression
    atomic point contacts (APCs), 177
    Fano factor, 176–177
    low frequency shot noise, 173–177
    noise temperature, 175–176
    Poissonian noise, 173
    SWNTs, 177
  two-particle correlation and interference
    four contact phase coherent conductor, 211f
    Hanbury Brown and Twiss (HBT) experiment, 211–215, 212f
    quantum interference, 212f
Shot noise thermometer, 93
Si photonics, 353–354

Silicon qubit devices
  fabrication
    silicon host material, 269–270
    silicon-on-insulator (SOI) nanowire, 272–275
    Si-MOS devices, 270–272
    Si/SiGe heterostructure, 272
  Moore's law, 265
  quantum computation
    gate-based quantum computation, 267
    implementations, 268
    parallelism, 267
    platforms, 267–268
    quantum mechanical two-level system, 266–267
    quantum state coherence and gate fidelity, 269–270
    qubit state, 266, 266f
    superposition and entanglement, 267
  single-spin qubit
    control and operation, 276–278, 277f
    effective Hamiltonian, 276–277
    electric dipole spin resonance (EDSR), 277–278
    electron spin resonance (ESR), 276
    pulsed experiments, 278
    quantum information processing, 275–276
    spin manipulation, 276–277
    time-dependent Hamiltonian, 276
    two-qubit gate, 278–281
    Zeeman energy split, 275–276
  singlet-triplet qubit, 281–282
  spin readout
    charge sensors, 283
    gate-based readout, 285–286
    magnetic moment, 282–283
    spin-to-charge conversion, 283–285, 284f
Silicon-on-insulator (SOI)
  fabrication, 273–274, 274f
  transmission electron microscope images, 275
Silicon-on-insulator (SOI) nanowire, 272–275, 273f
single electron transport (SET), 122
Single photon sources
  basic properties, 443
  diode designs, 295–302
  host semiconductor platform, 295
  light emitting diodes (LEDs), 439
  measurements
    intensity autocorrelation, 441–442
    objective lens, 441
    single photon detection HBT setup, 442f
    single photon emitter, 440
    spectral filtering, 441
  quantum key distribution (QKD) schemes, 440
  trapped carriers, 295
  true single photon emission, 439
Single wall carbon nanotube (SWCN), 149
Single wall nanotube shot noise, 179f
Single-electron source
  capacitor charging experiment, 123–125
  electron quantum optics. *See* Electron quantum optics
  Leviton, 117–120
    electron and hole excitations, 119–120, 120f
    Fermi energy, 119–120
    Lorentzian-shaped pulses, 117–119
    voltage pulse, 117–119
  mesoscopic capacitor, 117
  multi-junction single-electron pump, 104–105
  operating mode, 102
  quantum metrological triangle, 122–123, 123f
  SI ampere realisation, 121–122
  superconductor-normal metal hybrid turnstile, 106–108
  surface acoustic wave single-electron transfer, 108–113
  tunable barrier quantum dot pump, 113–116
  turnstile quantum dot single-electron transfer
    electrochemical potential, 102–103
    I-V characteristics, 103–104, 104f
    tunnel barriers, 102–103
    turnstile operation, 102, 103f

Singlet-triplet qubit, 281–282
Si-on-insulator (SOI), 353
Source drain bias, 16–17, 17f, 23–25, 24f
Spectral density, current fluctuation, 159
Spectral filtering, 441
Spin degeneracy, 11–12, 77–78, 77f
Spinon and holons lifetime, 60
Spin-orbit coupling (SOC), 242
Spin-orbit interaction (SOI), 11–12
Spin-to-charge conversion, 284f
Stark effect, 300
Strain and electric field tuneable quantum dots, 311, 311f
Strain-balanced growth technique, 320–321
Superconductor-normal metal hybrid turnstile
  charge stability diagram, 106–108, 107f
  current plateaux, 108f
  pump errors, 106–108
  SiNiS turnstile device, 106, 107f
Surface acoustic wave single-electron transfer
  acoustoelectric current quantisation, 109–111, 111f
  electrostatic confinement, 108–109
  electrostatic potential, 110f
  GaAs/AlGaAs heterostructure, 109–111, 109f
  quantum point contact charge detector conductance, 111–113, 113f
  scanning electron micrograph, 111–113, 112f

T

Thermal conductance, 68
Thermal cracks, 364
thermal noise, 147–148
Thermoelectricity
  Landauer-Büttiker formalism, 67–68
  quantum dot. *See* Quantum dot, thermoelectric properties
Thermopower, 84–85, 84f
  conductance, 72–73, 73f
  current heating technique, 75–76
  fine structure, 74, 75f
  lattice temperature dependence, 76
  Mott relation, 75–76
  oscillations, 74
Threading dislocation density reduction, 363–364
Threading dislocations, 354
Threshold gain, 393
Threshold pump density, 403–404
Tight-binding model, 19–20
Tomonaga–Luttinger liquid (TLL)
  Bloch's theory of sound waves, 32
  bosonisation method, 36
  1D sub-bands, 39
  dispersion, 32–33
  electron–hole pair spectrum, 35f
  energy spectra, 38–39
  interaction energy, 34–35
  kinetic energy, 34
  Lorentzian function, 38–39
  low-energy excitations, 33–34
  magnetotunnelling spectroscopy
    cleaved-edge overgrowth (CEO), 47
    2D electron system (2DES), 45–46
    device configuration, 45–46, 45f
    interaction parameter, 51–52
    Lorentz force, 46
    Pauli exclusion principle, 46
    tunnelling probability, 46
    zero-bias anomaly (ZBA), 46–47
  momentum distribution function, 39
  perturbation theory, 32
  photoemission spectroscopy
    direct, 40–41, 41f
    inverse photoemission spectroscopy, 42
    photoemission matrix element, 42
    quasi-1D metal $Li_{0.9}Mo_6O_{17}$, 42
    $SrCuO_2$ spectrum, 42
  power-law behaviour, 39–40
  single-particle excitations, 35f
  spectral functions, 38–39, 38f
  spinful
    bosonisation method, 37
    Hamiltonian, 36–37
    spin-charge separation, 37
  spinless scenario, 37–38
  total TLL Hamiltonian, 35–36
  transport measurements

Tomonaga–Luttinger liquid (TLL) (*Continued*)
conductance measurements, 44
contact geometries, 44–45
individual nanotubes, 43
Luttinger liquid parameter, 42–43
source-drain bias, 42–43
Topological insulator nanoribbons (TINRs)
cross-section, 248f
3D topological insulators, 242
high magnetic field range, 245–247
low magnetic field range, 244–245
perpendicular magnetic field, 244–247
weak anti-localisation (WAL), 244–245
electrical transport
ballistic nature, 251
ballistic transport, 252–253, 252f
inter-reference patterns, 247
linear Dirac dispersion, 247–249
magnetoconductance, 249–251, 250f
sheet carrier density, 249–251
sheet resistance and 3D resistivity, 243f
temperature dependence, 244
transversal modes, 247–249
transverse modes, 247–249
spin-helical surface energy dispersion, 242f
strong spin-orbit coupling (SOC), 242
superconducting transport
Aharonov–Bohm effect, 258–260, 259f
Andreev bound states, 254–256, 255f
asymmetric quantum interference device (SQUID), 256
energy dispersion, 253
energy-phase relation, 254
temperature dependence, 256–258
Topological transition, 241–242
Transverse electron focussing (TEF)
dynamic spin polarisation, 11–12
experimental setup, 11–12
spin degeneracy, 11–12
spin repulsion
exchange repulsive interaction, 13–14
exchange-driven, 12
injector conductance, 13–14, 13f
spin-split peaks, 13–14
Zeeman effect, 12
spin-orbit interaction (SOI), 11–12
Tunable barrier quantum dot pump
circular cutout, 114–115
entrance barrier, 114–115
gigahertz operation, 113
operation mechanism, 114–115, 114f
perpendicular magnetic field, 115f
quantisation accuracy, 115–116
frequencies, 125–126
high-precision measurements, 126, 127f
multi-junction pumps, 128
pump errors detection, 128–129
traceable current measurements, 126–128, 126f
scanning electron micrograph, 114f
silicon, 116f
Tuneable electroluminescent quantum dot light source, 308, 309f
Turnstile quantum dot single-electron transfer
electrochemical potential, 102–103
I-V characteristics, 103–104, 104f
tunnel barriers, 102–103
turnstile operation, 102, 103f
Two-dimensional electron gas (2DEG), 7–8
Two-particle coincidence probability, 213–215
Two-particle interference, 214f
Two-qubit gate
energy level diagram, 280f
exchange gate, 280
exchange interaction strength, 280–281
gate fidelity, 280–281
Hamiltonian, 278–279
Heisenberg interaction, 278–279
phase gate, 281
resonant gate, 281
two electron spins, 278–279
Two-step photo-absorption (TSPA) spectroscopy, 320
2D photocurrent excitation spectroscopy

atmospheric absorption, 347, 348f
2D-ΔEQE spectral maps, 346–347, 347f
IR photocurrent spectra, 347–348
photocurrent spectral maps, 346
FTIR spectroscopy, In(Ga)As QDSCs
benefits, 342–343
IR absorption edge fitting analysis, 344, 344f
IR light power intensity spectrum, 342–343
low temperature IR photocurrent spectra, 343–344
spectroscopy setup, 342–343, 343f
temperature dependence, 343f, 344, 345f
universal linear relationship, 345–346, 345f
IBSC operation principle, 341–342
photocurrent enhancement, 341–342

V

Vapour-liquid-solid (VLS) growth, 409

W

Wannier–Mott type exciton, 398–399
Weak anti-localisation (WAL), 244–245
Wiedemann–Franz law
charging energy gap, 83
Coulomb blockade oscillations, 82f
deviations, 83
Lorenz ratio, 81–83
violation, 82
Wigner function, 135
Wigner lattice, 8
Wire-to-wire disorder, 404

Z

Zeeman effect, 12
Zero-bias anomaly (ZBA), 46–47
Zinc oxide nanowires, 410–411