Thermoelectric experiments on nanowire-based quantum dots

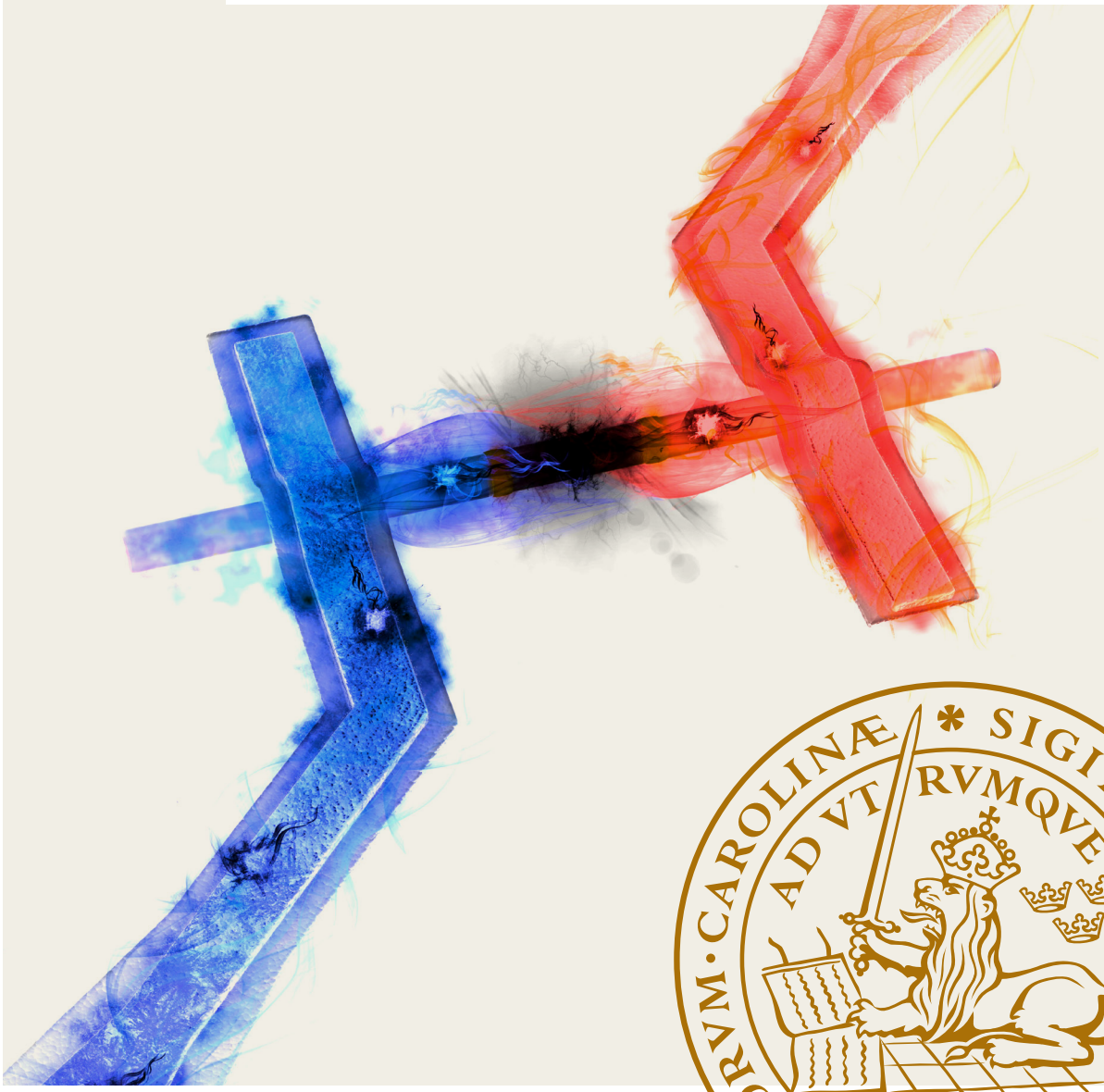
Svilans, Artis

2018

Document Version:
Publisher's PDF, also known as Version of record

Thermoelectric experiments on nanowire-based quantum dots

ARTIS SVILANS
FACULTY OF ENGINEERING | DEPARTMENT OF PHYSICS | LUND UNIVERSITY



Thermoelectric experiments on nanowire-based quantum dots

Artis Svilans

Doctoral Thesis
2018

LUND
UNIVERSITY

Thesis advisor: Prof. Heiner Linke

To be publicly defended, with the permission of the Faculty of Engineering, Lund University, on Friday, the 12th of October, 2018 at 9:15 in the Rydberg lecture hall (Rydbergsalen) at the Department of Physics, Sölvegatan 14, Lund.

Faculty opponent: Prof. Klaus Ensslin, Solid State Physics Laboratory, ETH Zurich, Switzerland

Organization	Document name
LUND UNIVERSITY Department of Physics Box 118 SE–221 00 LUND Sweden	**DOCTORAL DISSERTATION** Date of disputation 2018-10-12 Sponsoring organization
Author(s) Artis Svilans	

Title and subtitle

Thermoelectric experiments on nanowire-based quantum dots:

Abstract

This thesis experimentally investigates the possibilities of using quantum effects in semiconductor nanostructures for engineering their thermoelectric properties. More specifically, heterostructured InAs/InP nanowires are used to create short InAs quantum dots (QDs) with electronic state structure resembling that found in atoms. Recently developed top-heater architecture is used to apply a temperature differential across the QDs. The nanowire-based QD devices are used for studies of thermoelectric effects at the nanoscale and for experimental demonstration of particle-exchange heat engines.

This thesis first gives an overview of the most important physical effects governing the behavior of quantum dots (QDs). The Master equation approach to model the electronic transport in QDs is introduced in the sequential electron tunneling approximation. It is used to illustrative the transport behavior of QDs. The Landauer-Büttiker approach is also introduced as a reference and the differences with the sequential electron tunneling approximation are discussed. A summary of the most important literature on the thermoelectric properties of single QDs is given and discussed to provide the context for the experimental studies in this thesis. Finally, a description of the experimental methods used in this thesis is given.

There are three studies included in this thesis. The first investigates the nonlinear thermoelectric response of a QD with an applied thermal bias. A strongly nonlinear behavior is observed which can be fully explained by the interplay between different QD electronic states contributing to thermocurrent in opposing directions. The second study experimentally demonstrates efficient *particle-exchange heat engines* based on QDs for the first time. The analysis of the heat engines' power and efficiency indicate heat-to-electric conversion efficiencies up to 70% of Carnot efficiency. The third study investigates the thermoelectric response of QDs in the presence of Kondo correlations. It verifies a previous theoretical prediction that the sign of the thermoelectric signature in QDs inverts due to the Kondo correlations.

The experiments presented in this thesis have been successful in filling a gap between theory and experiments on several fronts. Future experiments could, for example, study Kondo-correlated QDs in the nonlinear thermoelectric response regime in the presence of magnetic field, where theory predictions are harder to obtain, or could employ thermoelectric characterization techniques to study entropy of various different QD states.

Key words

quantum dots, nanowire, thermoelectric, maximum power, efficiency, Carnot, Kondo effect

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language English
ISSN and key title	ISBN 978-91-7753-832-5 (print) 978-91-7753-833-2 (pdf)
Recipient's notes	Number of pages 154
	Security classification

Thermoelectric experiments on nanowire-based quantum dots

Artis Svilans

Doctoral Thesis
2018

LUND
UNIVERSITY

Division of Solid State Physics
Department of Physics
Lund University
Sweden

Front cover: Stylistic illustration of a quantum dot within a nanowire that is contacted between a hot (red) and a cold (blue) electrode. The quantum dots works as a heat engine and produces electric power. Image: Philip Krantz, Krantz NanoArt

Faculty of Engineering
Department of Physics
Lund University

ISBN: 978-91-7753-832-5 (print)
ISBN: 978-91-7753-833-2 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2018

Dedicated to my grandfather Laimonis

Contents

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Nonlinear thermoelectric response due to energy-dependent transport properties of a quantum dot**

A. Svilans, A. M. Burke, S. Fahlvik Svensson, M. Leijnse, H. Linke

Physica E **82** (2015) 34–38

I fabricated experimental devices, carried out electrical measurements, analyzed the data, did theory calculations, made the figures and contributed to the writing of all sections of the paper.

II **A quantum-dot heat engine operated close to thermodynamic efficiency limits**

M. Josefsson*, **A. Svilans***, A. M. Burke, E. A. Hoffmann, S. Fahlvik, C. Thelander, M. Leijnse, H. Linke

Published online in *Nature Nanotechnology* 16th of July 2018

I fabricated experimental devices, carried out electrical measurements, did SEM imaging, partly analyzed the data, partly made the figures and contributed to the writing of all sections of the paper.

* authors contributed equally to this work.

III **Thermoelectric characterization of the Kondo resonance in nanowire quantum dots**

A. Svilans, M. Josefsson, A. M. Burke, S. Fahlvik, C. Thelander, H. Linke, M. Leijnse

Submitted for journal publication

I initiated the study, fabricated experimental devices, carried out electrical measurements, analyzed the data, made all figure and led the writing of the paper with contributions from all authors.

The following papers are relevant, but are not included in this thesis

Experiments on the thermoelectric properties of quantum dots

A. Svilans, M. Leijnse, H. Linke

C. R. Physique **17** (2016) 1096–1108

I carried out the literature study, made all figures and led the writing of the paper with contributions from all authors.

Thermoelectric Power Factor Limit of a 1D Nanowire

I-Ju Chen, A. M. Burke, **A. Svilans**, H. Linke, C. Thelander

Phys. Rev. Lett. **120** (2018) 177703

I contributed to the choice of the measurement methodology and contributed to writing of the paper.

Optimum power and efficiency of quantum-dot heat engines

M. Josefsson, **A. Svilans**, H. Linke, M. Leijnse

Manuscript in preparation

I fabricated experimental devices, carried out electrical measurements and contributed to the writing of the paper.

Popular science summary

Much of the technological progress over the past century has been a result of the continuous advances in materials research. Much of the focus has been on synthesizing materials with new or improved physical properties. This engineering of the material's properties very often requires an understanding of its atomic structure and composition. It has allowed researchers to engineer the properties of semiconductor structures in great detail, which is why they play a critical role in almost every piece of technology that surrounds us in our daily lives.

Most of us are now used to hearing about the miniaturization of computer chips down to unbelievably small scales. The trend of miniaturization essentially offers reducing the cost of the technology, however, it is becoming increasingly challenging. The semiconductor technologies have reached a state in which single electrical circuit elements in computer chips can now be a few tens of nanometers in size (less then 1/1000 of a thickness of a human hair). This is only about 50 times the distance between atoms in solids. Engineering materials at such a small length scale is conceptually different because then the electrons that carry current often can not be considered as point like particles anymore.

The ability to manipulate material properties at a nanometer scale also opens up for new engineering opportunities. Therefore, one of the focuses of nanoscience is to take advantage of the effects related to the small size of nanostructures. Researchers at NanoLund have expertise in growing very thin wires (nanowires) with diameters often below 1/10000 of a millimeter with high quality. Nanowires often have improved material properties and also allow combining materials that are not possible to combine in bulk without degradation of their properties. On the one hand this platform is being used to develop nanowire based high performance transistors, light emitting diodes, photovoltaics and light detectors. On the other hand nanowires enable the creation of engineered nanostructures with unique properties for targeted studies of physical phenomena, not possible to do otherwise.

This thesis uses specific kinds of nanostructures, called *quantum dots*, created into the nanowires in order to study heat conversion into electricity at a very fundamental level. One of the highlights of this thesis is an experimental realization of a particular type of heat engine based on a quantum dot called the *particle-exchange heat engine*. Such an engine was predicted to enable approaching the ideal thermoelectric conversion efficiency, the so-called Carnot efficiency limit. The experimental results presented as a part of this thesis demonstrate heat-to-electric conversion efficiency up to 70 % of the Carnot efficiency, which is comparable to the efficiency of the best heat engines. Another highlight of this thesis is the realization of thermoelectric experiments probing quantum dots in an exotic state (Kondo regime). This regime can be reached at extremely low temperatures (below -270°C) and it enables the formation of such states of matter that can electrically

short the quantum dot as if it was no longer there. The experimental results in this regime are a direct verification of the somewhat counter-intuitive theory predictions and provide a basis for future studies.

Over the course of this thesis the nanowire-based quantum dots have enabled successful experimental implementation of studies that so far had not been realized due to experimental difficulties. The approaches used in this thesis are expected to be used in future experiments further exploring the thermoelectric effects at the nanoscale.

Abstract

This thesis experimentally investigates the possibilities of using quantum effects in semiconductor nanostructures for engineering their thermoelectric properties. More specifically, heterostructured InAs/InP nanowires are used to create short InAs quantum dots (QDs) with electronic state structure resembling that found in atoms. Recently developed top-heater architecture is used to apply a temperature differential across the QDs. The nanowire-based QD devices are used for studies of thermoelectric effects at the nanoscale and for experimental demonstration of particle-exchange heat engines.

This thesis first gives an overview of the most important physical effects governing the behavior of quantum dots (QDs). The Master equation approach to model the electronic transport in QDs is introduced in the sequential electron tunneling approximation. It is used to illustrative the transport behavior of QDs. The Landauer-Büttiker approach is also introduced as a reference and the differences with the sequential electron tunneling approximation are discussed. A summary of the most important literature on the thermoelectric properties of single QDs is given and discussed to provide the context for the experimental studies in this thesis. Finally, a description of the experimental methods used in this thesis is given.

There are three studies included in this thesis. The first investigates the nonlinear thermoelectric response of a QD with an applied thermal bias. A strongly nonlinear behavior is observed which can be fully explained by the interplay between different QD electronic states contributing to thermocurrent in opposing directions. The second study experimentally demonstrates efficient *particle-exchange heat engines* based on QDs for the first time. The analysis of the heat engines' power and efficiency indicate heat-to-electric conversion efficiencies up to 70% of Carnot efficiency. The third study investigates the thermoelectric response of QDs in the presence of Kondo correlations. It verifies a previous theoretical prediction that the sign of the thermoelectric signature in QDs inverts due to the Kondo correlations.

The experiments presented in this thesis have been successful in filling a gap between theory and experiments on several fronts. Future experiments could, for example, study Kondo-correlated QDs in the nonlinear thermoelectric response regime in the presence of magnetic field, where theory predictions are harder to obtain, or could employ thermoelectric characterization techniques to study entropy of various different QD states.

1 Introduction

Trends in semiconductor research

Material research and development is in the foundation of technological advancement. A significant part of the technological progress during the last century has been allowed by continuously increasing the understanding of electronic properties of semiconductors [1, Chap. 1-4]. The possibility to engineer semiconducting materials with desired characteristics has opened vast amounts of possibilities for applications. The most prominent example of this, of course, is the development of semiconductor transistors and integrated circuits [1, Chap. 3-4] resulting in multi-decade-long miniaturization of computer chips with exponentially increasing computation power [2]. However, the applications of semiconductor advances are far from being limited to computation. Other applications include: power and high speed electronics, photovoltaics and opto-electronics. These technologies have allowed for the modern power grids, radar and wireless communication technologies, solar cells, high speed optical communication, low power lighting, flat screen displays, digitization of photography and video, thermal imaging - the list could go on [1, Chap. 5-10]. Overall, one has to conclude that the impact of fundamental understanding of semiconductors on society has been enormous.

It is important to point out that the continuation of this trend has always relied on the introduction of new solutions for overcoming existing limitations. Therefore, material research still plays a key role in advancing semiconductor technologies. The miniaturization of modern transistors [3] is still continuing, however, it is projected that this trend will come to an end within the coming decade [4]. A paradigm shift is expected in the form of novel applications for semiconductor devices in the context of the so-called *Internet of Things* (IoT) [5]. For this class of applications with many interconnected small scale sensors, transducers and information processing units increased functionality and lower power consumption are more desirable than the metrics, like information processing speed or the absolute size.

Here, opportunities provided by bottom-up growth of nanostructures [6] can add on to the existing processing practices, provide new functionality or, perhaps, even form a basis for entirely new technologies. A relatively developed platform in this respect has been the selective growth of semiconductor nanowires [7]. It demonstrates the ability to combine high material quality with high control over composition, both in radial and axial dimensions [8]. The small size of nanowires also offers solutions for combining highly lattice mismatched materials without introducing crystal defects [9], which is not possible in bulk. Both of these characteristics in combination allow for great flexibility in the design of nanoscale devices, which has been exercised in many nanowire-based device concepts. They include opto-electronic devices like light emitting diodes (LEDs) [10], lasers [11] and infrared detectors [12], but also photovoltaic cells [13] and high performance transistors [14].

Such experimental devices have therefore already demonstrated that they can ultimately be scaled down to a size comparable to the size of modern transistors without losing functionality. The possibility of fabricating so many different devices at the nanoscale suggests a great potential for further integration of various technologies and perhaps could even lead to entirely new applications.

Engineering thermoelectric effects at the nanoscale

Among various research directions investigating potential nanowire applications for future solid state devices is engineering nanowire thermoelectronic properties. This thesis makes direct use of the small size of nanowires and the heterostructuring capabilities available in them to engineer electron transport characteristics. It allows fine-tuning of the nanowire properties in ways that benefit thermoelectric performance figures, like power or efficiency, and are not applicable to bulk thermoelectric materials. From a practical standpoint, engineering of nanomaterials' thermoelectric properties is of interest for on-chip cooling or power generation [15]. From a scientific standpoint, investigation of thermoelectric effects at a nanoscale attempts to clarify which physical mechanisms are responsible for the fundamental limits, as well as for the optimal operating conditions, of thermoelectric energy conversion devices.

Experimentally, the epitaxial growth of nanowires allows integration of different compound semiconductors with atomically sharp interfaces while ensuring minimal amount of crystal defects [8]. The ability to heterostructure nanowires in this manner is a powerful engineering tool for electron transport. This thesis uses nanowires in which composition varies along the nanowire length for obtaining desired electronic transport properties. The specific structure consists of a pair of thin indium phosphide (InP) segments integrated into an indium arsenide (InAs) nanowire (see Fig. 1a). The InP segments are relatively poorly conductive and are used to enclose a short segment of InAs separating a countable number of conduction electrons from the rest of the nanowire [16, 17]. As a consequence of their small size, these islands of conduction electrons, also called *quantum dots* (QDs), possess a set of energetically very distinct single particle orbital states, similar to those in atoms. It is precisely this quasi-discrete energy state structure of QDs that makes them good model systems for studying thermoelectric effects at a nanoscale. In this thesis it has been possible to make considerable experimental progress in the understanding of thermoelectric properties of QDs, largely owing to the high quality of InAs/InP nanowires.

The research in this thesis has been directed mainly by two ideas. The first idea driving the thermoelectric studies of QDs is based on using thermoelectric quantities as a complementary characterization tool to learn more about QDs themselves. The thermoelectric measurements can provide complementary information about physical mechanisms in the QD [18, 19] because the nature of thermal bias is different form the nature of electrical

bias. Prior experimental work on thermoelectric properties of QDs has mostly been aimed at explaining the qualitative behavior of thermoelectric signatures based on conductance behavior of the corresponding QDs. However, quantitative agreement has so far been lacking [20]. The existing lack of agreement between theory and experiments can be explained as a result of a combination of experimental limitations in creating suited experimental QD systems and the consequential lack of experience in modeling of the thermoelectric properties of realistic QDs. Work in this thesis attempts to fill in this gap between theory and experiments by making improvements on both sides. Experimental groundwork for this purpose is largely tested in Paper I and is further used all throughout the studies in Papers II and III.

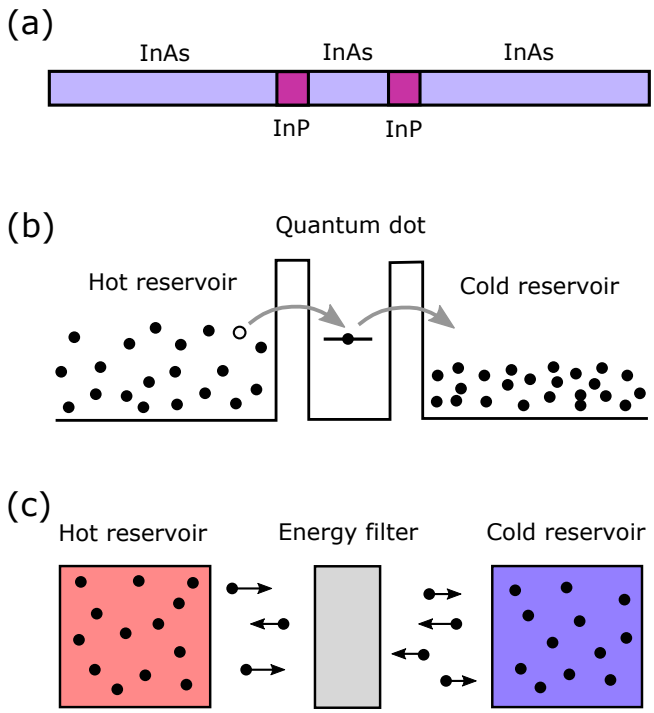


Figure 1: (a) Illustration of a heterostructured InAs/InP nanowire. The short segments (in purple) represent InP segments within InAs nanowire. (b) Schematic view of a quantum dot (QD) created between the InP segments. Electrons on the left side of the QD are at a higher temperature than electrons on the right. On the hot side the electrons are thermally excited to higher energies than those on the cold side. An electronic orbital state in the QD is able to conduct electrons at a selected energy from the hot to the cold side, thus driving a current using the temperature differential. The QD can serve as an electron energy filter in a solid state particle-exchange heat engine illustrated in (c). (c) Schematic view of a particle-exchange (PE) heat engine. Two particle reservoirs at different temperatures (red and blue) are exchanging particles with each other through a particle energy filter (gray).

The second idea is to test a two decade old theoretical prediction of QDs being ideal for efficient steady state thermal-to-electric energy conversion [21, 22]. The critical property for efficient thermoelectric conversion possessed by QDs is their ability to transport electrons at very well defined energies (illustrated in Fig. 1b), therefore behaving as energy filters for

electrons. This characteristic of QDs has been predicted to enable the implementation of a novel type of heat engine, the *particle-exchange heat engine* [23] (Fig. 1c), which would operate in a steady state by exchanging electrons at a set energy between two fermionic reservoirs of different temperatures. Following this layout, the temperature difference could be used to drive an electric current through the QD against an electrical potential difference. Mostly because of the lack of agreement between theory and experiments, no definite conclusions have been made regarding these predictions. Paper II explores this idea experimentally and demonstrates extensive agreement between theory and experiment. Throughout this study the nature of thermoelectric power production in QDs is characterized.

The approaches developed throughout Papers I and II could be applied to studies of other transport regimes where properties of QDs are less understood and more challenging to model theoretically. One of such regimes is the so-called *Kondo regime* [24, 25] where the interaction between the QD electrons and electrons in the reservoirs gives rise to quantum many-body phenomena that result in drastic modification of QDs' transport properties. Existing theoretical predictions by Costi and Zlatić [19] on the behavior of thermoelectric signals in this regime so far have not been fully mapped experimentally, therefore the theory does not stand on solid ground. Studies in Paper III are addressing this fact by conducting thermoelectric experiments on QDs in the Kondo regime. The results demonstrate good qualitative and qualitative agreement with theoretical predictions and thus constitute an early validation of the theory in Ref. [19]. The study in Paper III also opens up the field for further experimental studies in this regime as it demonstrates that the nanowire-based QDs seem to be a very good experimental system.

The thesis is constructed in the following way. Section 2 provides a review of basic physical phenomena that are important for the general theoretical understanding of QDs. These concepts are used in Section 3 where a simplified theoretical model is introduced to qualitatively explain the behavior of QDs in response to electrical and thermal biases. The applicability limits of simplifications in the model are discussed. Further, Section 4 provides a short review of the prior work on thermoelectric properties of QDs, focusing mostly on the experimental achievements. Section 5 reviews the experimental methods and specifics of devices that have been used for the experimental studies in this thesis. Section 6 gives a detailed commentary of the main experimental results of Papers I, II and III. Section 7 concludes the main body of this thesis by a short summary and outlook.

2 Background

Quantum dots (QDs) are sometimes called artificial atoms because they possess electronic state structure similar to those in atoms. Essentially, single particle kinetic (orbital) energy states in QDs are spread enough in energy not to be seen as a continuum of states at the given temperature. The energy offset between the orbital states is caused by quantum *confinement effect* (explained in Section 2.3). The aim of this section is to introduce this effect as well as effects, like charging (in Section 2.1), tunneling (in Section 2.4) and others, that play an important role in transport properties of single QDs. The principles explained in this section are further used in Section 3 to motivate the typical assumptions made in the theoretical description of QDs.

2.1 Electrostatics of small conductors

Electrostatics plays a significant role in determining the electronic properties of small conductors. The fundamental reason for this is that smaller conductors cannot screen the electric field of their charge as well as bigger conductors can, resulting in higher electric field energies. This is just another way of saying that smaller conductors have smaller capacitances. A good example for this is provided by a basic result in electrostatics - a solid metallic sphere in vacuum placed far away from other conductors. Assuming the sphere has a radius r and charge Q, its electric potential is $\phi = Q/(4\pi\varepsilon_0 r) = Q/C$, where ε_0 is the vacuum permittivity and C is the capacitance of the solid sphere to a surrounding, far away equipotential surface. This result implies that the field energy of the charged sphere is $W = \phi Q/2 = Q^2/(2C)$, and therefore proportional to $\sim 1/r$. Based on this, one can conclude that charging of small conductors is energetically more costly than of the bigger ones.

The picture becomes slightly more complicated in case of several conductors, capacitively coupled to one another. It is commonly accepted that linear response properties of such a system can be described by a capacitance matrix $\mathbf{C}$, the elements of which are defined as $C_{ij} = dQ_i/d\phi_j$. Here, Q_i is the charge of the conductor i, where as ϕ_j is the electric potential of the conductor j. The charge of each of the conductors is related to the electrical potentials via the following expression

$$Q_i = \sum_j C_{ij}\phi_j \tag{1}$$

There are two important properties of the $\mathbf{C}$ matrix that are of great importance here. First, it is a symmetric matrix ($C_{ij} = C_{ji}$), and second, that the sum of all elements in each row and column is zero $\sum_i C_{ij} = \sum_j C_{ij} = 0$. The latter implies that the total charge of the system equals zero, which is a justifiable assumption, and that only the relative potential differences

between conductors matter, which assumes the existence of a far away equipotential surface at zero potential [26].

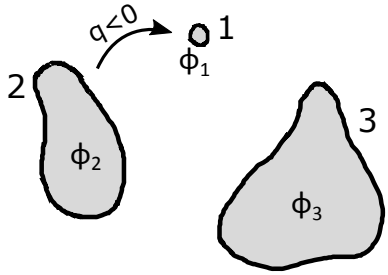


Figure 2: A schematic view of three uncharged conductors of different sizes and shapes, labeled 1, 2 and 3. Charge $q < 0$ is moved from the conductor 2 to 1. As a result of the new charge rearrangement the conductor 2 becomes positively charged whereas the conductor 1 becomes negatively charged. Resulting electrical potentials of the conductors are ϕ_1, ϕ_2 and ϕ_3 as indicated in the figure.

Let us now consider a practical example - three uncharged conductors nearby that are of various sizes and shapes (see Figure 2). Suppose that we take a charge q and move from one conductor to another. By solving Eq. 1 for a system where $(Q_1, Q_2, Q_3) = (-q, +q, 0)$, and by applying the properties of $\mathbf{C}$, it is possible to obtain the potentials ϕ_1, ϕ_2 and ϕ_3 that are fully determined by three capacitances C_{12}, C_{13} and C_{23}.

$$\begin{cases} \phi_1 = qC_{23}/(C_{13}C_{23} + C_{12}(C_{13} + C_{23})) \\ \phi_2 = -qC_{13}/(C_{13}C_{23} + C_{12}(C_{13} + C_{23})) \\ \phi_3 = 0 \end{cases} \tag{2}$$

QDs are typically the smallest conductors in the circuit, which allows to make considerable simplifications (this is illustrated in Fig. 2, where the conductor 1 is much smaller than the other conductors). The argument goes as follows. If the simple conclusion from the solid sphere case applies to all conductors, the conductor 1 capacitances to the other conductors, C_{12} and C_{13}, should be much smaller than the capacitance between the two bigger conductors C_{23}. The electric potentials can then be approximated by

$$\begin{cases} \phi_1 \approx q/(C_{12} + C_{13}) \\ \phi_2 \approx -q(C_{13}/C_{23})/(C_{12} + C_{13}) \\ \phi_3 = 0 \end{cases} \tag{3}$$

The Eq. 3 represents an important result as the assumption $C_{13}/C_{23} \ll 1$ implies that $\mid \phi_1 \mid$ must be much bigger than $\mid \phi_2 \mid$. The importance of it is in the fact that the total field energy W is then dominated by ϕ_1 and therefore the charge $-q$ on the conductor 1

$$W = \frac{1}{2}\sum_i Q_i\phi_i \approx -\frac{1}{2}q\phi_1 \tag{4}$$

In our example, considering conductors 1 and 2 as two capacitor plates (charged by $\pm q$) and using the corresponding potential difference between the conductors $\phi_2 - \phi_1$, one can also calculate the corresponding two-terminal capacitance of the effective capacitor between the two conductors.

$$\frac{q}{\phi_2 - \phi_1} = -\frac{C_{12} + C_{13}}{1 - C_{13}/C_{23}} \approx -C_{12} - C_{13} = C_\Sigma \tag{5}$$

The Eq. 5 shows that in the limit of $C_{13}/C_{23} \ll 1$ the two-terminal capacitance can be approximated by the sum of the small conductor capacitances to the other conductors (note that in Eq. 1 the non-diagonal capacitance coefficients are defined as negative numbers). What it tells us is that, first, in certain limits the electrostatics can be dominated by the properties of the smallest conductor, and second, that the corresponding capacitance of charging the small conductor is affected by the presence of uncharged objects nearby. Using C_Σ to describe the electrostatic properties of QD-like devices is a commonly used approximation [27]. The overall result is that the total electrostatic energy U for charging a small uncharged conductor by Q becomes

$$\boxed{U = \frac{Q^2}{2C_\Sigma}} \tag{6}$$

The existence of U has implications on current through small conductors like QDs. In macroscopic conductors applying an arbitrary small electric potential difference (electric bias) results in a current proportional to its conductance according to Ohm's law. However, if every electron entering the small conductor requires a significant potential energy U, the current will simply be blocked from flowing through it unless a sufficient energy is supplied to electrons from the applied electrical potential difference. This effect is widely known in mesoscopic physics as Coulomb blockade [28] and is the basis for single electron transistors [29]. The magnitude of U for QDs studied in this thesis is a few meV, therefore U plays an important role in their conductance and thermoelectric properties at low temperature.

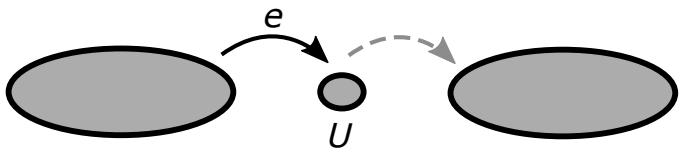


Figure 3: An illustration of an electron being moved from the left conductor to the right conductor (constituting a current) via a small conductor in between. The electron requires potential energy U to charge up the small conductor during the intermediate state.

2.2 Electrons in solids

Single electron in a solid

The development of quantum theory [30] at the first half of the 20$^{\text{th}}$ century revolutionized physics and allowed for deep insights in the wave nature of matter. The wave properties of non-relativistic particles are described by the Schrödinger equation [31]. It is a wave equation with solutions that are generally complex wave functions of coordinate and time. Each of the wave functions correspond to a state that a particle can occupy. The solid state theory applies this wave description to electrons in solids [32]. The difficulty arises because it is ultimately the large number of atomic nuclei and electrons interacting with each other that determine the electronic properties of a solid. Solving such a problem exactly for a macroscopically sized object is currently impossible, therefore, some approximations have to be made. The purpose of this section is to discuss approximations that motivate the use of single particle picture in solids, which is used throughout this thesis.

The first step towards understanding electronic structure of solids uses symmetry arguments. It is based on the fact that atoms in crystalline materials are bound at periodically located lattice sites which makes the crystal translationally symmetric. It forms basis for an assumption that an electron in a translationally symmetric electrical potential landscape is a good first approximation for a solid. In such a case, the solutions to the Schrödinger equation for spinless electrons are particular wave functions called Bloch waves [33]. These wave functions can be written as $e^{-i\mathbf{kr}}u_{\mathbf{k}}(\mathbf{r})$ where $\mathbf{k}$ is the so-called wave vector specifying the momentum of an electron and $\mathbf{r}$ is the real space position vector indicating its coordinate. The exponential part $e^{-i\mathbf{kr}}$ represents a plane wave in three dimensions whereas $u_{\mathbf{k}}(\mathbf{r})$ is a function that generally varies periodically with the periodicity of the crystal lattice.

Solving the Schrödinger equation for a periodic electrical potential gives a few insights. First, the electron energies corresponding to Bloch states fall into energy zones called *bands* [33]. So not all electron energies are possible, as for an electron in free space. Second, the dispersion relation of an electron $E(k)$ along either of the axes is modified such that $\hbar^2(\partial^2 E/\partial k^2)^{-1} = m^* \neq m$, where m is the free electron mass. This means that an electron in a periodic potential behaves as if it had a mass m^*, different from the free electron mass

m. This introduces the concept of an *effective mass*.

The emergence of the energy bands and the effective mass solely from the symmetry arguments is a very nice result in solid state theory. In fact, despite an electron in a periodic potential being a simple approximation, its conclusions can be conveniently used to explain differences between conductors, semiconductors and insulators. However, the initial approximation of a single electron in a periodic potential is not fully justified since electrons along with nuclei in a solid form an interacting many-body system. Therefore, there has to be an underlying argument of why the single particle picture is somewhat valid to start with in order to accept the explanation as a valid one.

It turns out that the effects of interactions on the behavior of a single particle can, up to some degree, be accounted for while still remaining in the single particle picture. The formal procedure for doing it is given by the many-body Green's function description of solid state theory [34]. It is done via the introduction of a concept of self-energy that essentially integrates the effects of interactions as a correction into a single particle propagator [35]. In a simplistic picture the self-energy accounts for the back-action of the other electrons and nuclei on the electron itself. In essence, the procedure approximates a single interacting particle by a single non-interacting quasi-particle with different properties, among which also a different mass (the effective mass). In this way the Green's function theory justifies the single particle picture which is why the concept is so widely used.

The single particle picture is also extensively used in this thesis in the description of QDs. Although it is not explicitly reminded throughout the text, it is important to acknowledge that the electrons discussed in this thesis are quasi-particles and therefore they have different properties in different materials. This is also used to the advantage of this work. For example, InAs (of which the QDs studied here are made) has a relatively low effective mass, which is useful for enhancing the quantum confinement effect in the QDs (discussed more in Section 2.3). The electrons in InAs also react relatively strongly to the application of magnetic field, which has made magnetic field dependence studies more convenient. The theory description of the QDs (given in Section 3) does not explicitly concern the origin of the quasi-particle properties of electrons, but rather uses them as input parameters in a simplistic theoretical description of QDs.

Electrons in heterostructures

An interface between two different crystalline materials is called a hetero-junction. In this context, a single solid made of different crystalline materials interfacing each other is called a heterostructure. Such materials do not possess a global translational symmetry and therefore cannot be described the same way as spatially uniform crystals. Even in the single particle picture no general form of solutions to the Schrödinger equation exist.

The so-called envelope function approximation (EFA) [32] is used to counter this issue. It formally assumes that a heterostructure can be modeled by a perturbative potential energy $e\phi = V_\phi(\mathbf{r})$ that is added on top of the translationally symmetric potential energy V_0 considered in the Bloch's theorem. Further, it conventionally assumes 1) that $V_\phi(\mathbf{r})$ varies slowly with respect to the periodicity of the crystal lattice a, 2) that the wave vector components are small ($k \ll \pi/a$), and 3) that the magnitude of the perturbative potential is small with respect to separation of energy bands. Given these conditions and the effective mass approximation (discussed in this context in Ref. [36]), one can write down an equation that is formally valid for a slowly varying (envelope) function $F(\mathbf{r})$ [37]

$$-\frac{\hbar^2}{2}\nabla\left[\frac{1}{m^*(\mathbf{r})}\nabla F(\mathbf{r})\right] = \left[E - E_c(\mathbf{r})\right]F(\mathbf{r}) \tag{7}$$

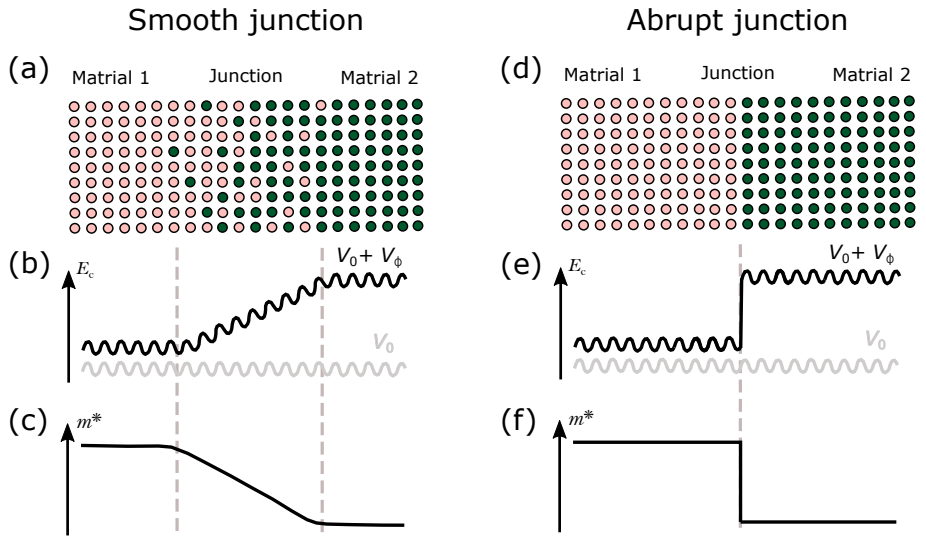


Figure 4: (a-c) Smooth heterojunction. (d-f) Atomically abrupt heterojunction. (a), (d) Illustrations of a crystal lattice made up of two types of unit cells (pink and dark green) corresponding to two materials (1 and 2) forming the hetero-junctions. In (a) the far out regions of the lattice have homogeneous composition whereas the interface (junction) region has a mixed composition. In (d) there is an atomically sharp change of the composition between the materials. (b), (e) illustrate model electric potentials for an electron in the lattice across the corresponding hetero-junctions in (a) and (d). The translationally symmetric potential V_0 is complemented by a potential V_ϕ that is varying slowly in (b) and abruptly in (e). (c), (f) illustrate an example how the effective mass m^* could vary across the corresponding hetero-junctions in (a) and (d). In (c) the change in m^* is gradual whereas in (f) the change of m^* is abrupt. Note that generally the relative effective mass change does not relate to the sign of V_ϕ change.

where $\hbar$ is the reduced Plank's constant, ∇ is the operator nabla, E is the total energy and $E_c(\mathbf{r}) = V_0 + V_\phi(\mathbf{r})$ is the potential energy of the electronic states. $E_c(\mathbf{r})$ models the band edge energy as it is the lowest energy an electron can have in a band. Note that here the effective mass $m^*(\mathbf{r})$ in the Eq. 7 is also a function of coordinate. Under the assumptions described above, the function $F(\mathbf{r})$ is an approximate substitute for the plane wave part in the Bloch wave function. The full wave function is thus approximated by $F(\mathbf{r})u_\mathbf{k}(\mathbf{r})$.

What is convenient about the form of Eq. 7 is that it reduces the many-body problem in a heterostructure down to a single electron picture in an effective potential that is set by the varying band edge energy. Such a convenience has made the EFA a widely used trick. It was even used as an approximation for atomically abrupt hetero-junctions in quantum well structures and superlattices [38], where the effective mass has an abrupt change, even though doing so was beyond the known applicability limits of EFA at the time. Further development of a more exact formulation of EFA [36, 39] has allowed to extend the applicability of the method to abrupt interfaces and to evaluate the magnitude of corrections needed. It has been argued [37] that Eq. 7 under similar criteria is valid for finding envelope functions also for atomically abrupt hetero-junctions (see section 2.1 in Ref. [37] for details). In a general case, however, (like for energies much above the band edge), the accuracy of the Eq. 7 can not be guarantied. Then, the full form of envelope function equations along with generalized boundary conditions should be used [36, 39].

The general form of EFA equations is normally discussed only in specialized literature [36, 37, 39] and is not used often. One reason for this is of course that the equations are more complicated, therefore less convenient to use. But also the general form EFA equations no longer has form of the Schrödinger equation, which makes it much less educational. Here too, mostly for the convenience, but also conforming with the literature examples, the further discussion assumes that the applicability conditions of the Eq. 7 are met. This might not be fully justified for the case discussed in Section 2.3 which considers InAs/InP heterostructure with atomically sharp interfaces (thus modeling the QDs used in the experiments). Nevertheless, since no drastic differences are expected, using Eq. 7 serves the purpose just as well, i.e. it gives a simple picture to use when thinking about electrons in heterostructures and illustrates how the behavior of electrons change when the dimensions of the heterostructure segments are changed.

2.3 Quantum confinement effect

When a solid becomes comparable in size to the electron (de Broglie) wavelength its electronic properties start to be strongly influenced by its finite size. The electrons are said to be confined by the finite size of the solid because leaving it requires additional energy. This additional energy can be classically seen as a potential barrier for electrons. The purpose of this section is to briefly look at this, so-called, *quantum confinement effect* by reviewing the results of rather simple text-book problems in quantum mechanics - particle in an infinite and a finite potential well. These problems are formulated within the single particle picture (see Section 2.2) and their correspondence to the InAs/InP nanowire-based QDs used in the experimental studies relies on the envelope wave function approximation.

Infinitely deep well

When considering an infinitely deep potential well the electrical potential energy profile V_ϕ is considered infinite outside the well and zero inside it. An additional assumption is made that the potential barriers are also infinitely sharp. Such an assumption, of course, is not physical since an infinitely sharp electrical potential barrier implies an infinity large electric field, however, these conditions can be used to conveniently introduce the physical principles and to provide order of magnitude estimates for electron energies.

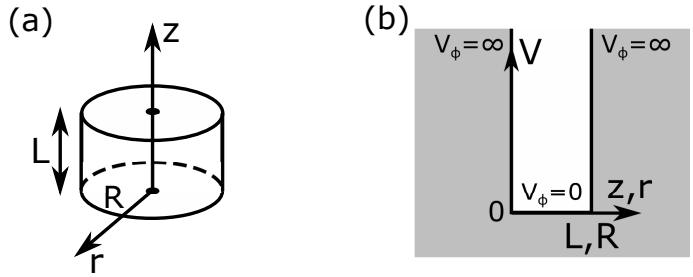


Figure 5: (a) Illustration of a cylinder with length L along the cylinder axis (z-axis) and radius R along the radial axis (r-axis). Electrical potential for an electron $V_\phi = e\phi$ is considered to be 0 inside the cylinder and infinite outside of it ($e > 0$ is the elementary charge). (b) Graphical representation of the electric potential V_ϕ as a function z and r. The plot area that is toned gray represents the area below the curves $V_\phi(z)$ and $V_\phi(r)$.

Since the QDs in the InAs/InP nanowires are cylindrical (disc like) objects, it is instructional to consider an infinite cylindrical potential well (see Fig. 5) with length L and radius R. Finding the possible electronic states in the cylinder is done by solving the envelope function equation Eq. 7 in cylindrical coordinates

$$-\frac{\hbar^2}{2m^*}\Delta F_{n,l,m} = (E_{n,l,m} - V_\phi)F_{n,l,m} \tag{8}$$

were Δ is the Laplace operator in cylindrical coordinates and $E_{n,l,m}$ is the energy of the electron states. The Eq. 8 has a simplified form because within the volume of the cylinder the effective mass m^* is assumed to be constant. The problem itself is an eigenvalue problem in three dimensions, therefore the spectrum of solutions is specified by three integer numbers n, l, m (quantum numbers). The infinite cylindrical potential well requires $F_{n,l,m}$ to be zero at the cylinder boundaries

$$F_{n,l,m}(z = 0, r) = F_{n,l,m}(z = L, r) = F_{n,l,m}(z, r = R) = 0 \tag{9}$$

The mathematical form of $F_{n,l,m}$ contains special functions called the Bessel functions of first kind [40]. However, the precise form of these functions is not of interest here. What we are interested in instead is the eigenvalues $E_{n,l,m}$, which have the following form [41]

$$E_{n,l,m} = \frac{\hbar^2}{2m^*}\left[\left(\frac{\pi n}{L}\right)^2 + \left(\frac{\alpha_{l,m}}{R}\right)^2\right] \tag{10}$$

The quantum number n is a positive integer, whereas $\alpha_{l,m}$ is the l-th root of the Bessel function $J_m(r)$ with m being a non-negative integer. The values of $\alpha_{l,m}$ are tabulated, for example in Ref. [42]. Overall, Eq. 10 defines a set of electron kinetic energies given the effective mass m^* and the dimensions of the cylinder. By picking values for parameters m^*, L and R such that they roughly correspond to the experimental conditions, we can obtain representative values for $E_{n,l,m} = E_N$ (here N is the energy level number) for various sizes of the cylinder.

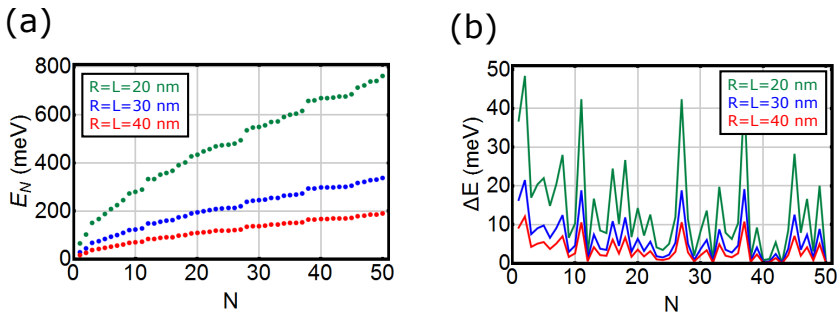


Figure 6: (a) Energy E_N as a function of the level number N, calculated using Eq. 10. (b) Energy state spacing ΔE_N as a function of the level number N corresponding to the E_N data in (a). In (a) and (b) the three different presented cases use parameters values $R = L = 20$ nm (green), $R = L = 30$ nm (blue) and $R = L = 40$ nm (red). An approximate effective mass of electrons in InAs is $m^* = 0.023m_e$, where m_e is the free electron mass.

Both the energy E_N and energy spacing $\Delta E_N = E_{N+1} - E_N$ are plotted in the Fig. 6. One can see that values of E_N scale inversely with size of the cylinder. An implied consequence of this is that the bigger is the cylinder, the more energetically dense the E_N spectrum is. Overall, ΔE_N is very sensitive to the exact dimensions of the cylinder and is irregular. ΔE_N tends to decrease gradually with N, which is consistent with the density of states in 3-dimensional materials.

An infinite cylindrical potential well is, of course, just and approximation, therefore values of ΔE_N are not expected to be in exact agreement with experimental reality. However, both the order of magnitude and the irregularity of ΔE_N are consistent with experimental observations [43]. In many ways a small piece of solid, as considered here, can be seen as an artificial atom because the conduction electrons, considered as non-interacting quasi-particles, occupy energetically very distinct single particle orbital states. This illustrates how the physics of atoms, can be captured in much bigger objects, like InAs QDs studied in this thesis.

Finite well

When considering a finite potential well instead of an infinite one, the infinite potential energy V_ϕ outside the well is replaced by a finite value. For simplicity, here the problem is reviewed in one dimension (coordinate x) which also means that the values of ΔE_N cannot be directly compared with the results in Fig. 6. This is because the states in a 1-dimensional well are numbered only by one quantum number $n = N$ and the additional states corresponding to the variations of the other quantum numbers do not show up in the analyses. Nevertheless, the conclusions obtained here are rather general and relevant for the discussion in the following Section 2.4.

$$-\frac{\hbar^2}{2}\nabla\left[\frac{1}{m^*(x)}\nabla F_N(x)\right] = \left[E_N - V_\phi(x)\right] F_N(x) \tag{11}$$

The problem of a particle in a finite potential well has analytical solutions, however, the energies E_N have to be found numerically [44]. Equation 11 has to be solved with different boundary conditions where, instead of setting $F_N = 0$ at the edge of the well, one has to ensure continuity of $F_N(x)$ and $[1/m^*(x)]\,\partial F_N(x)/\partial x$ across hetero-junctions thus ensuring the continuity of the probability current [38]. Note that this implies $\partial F_N(x)/\partial x$ not being continuous across those junctions at which $m^*(x)$ changes abruptly.

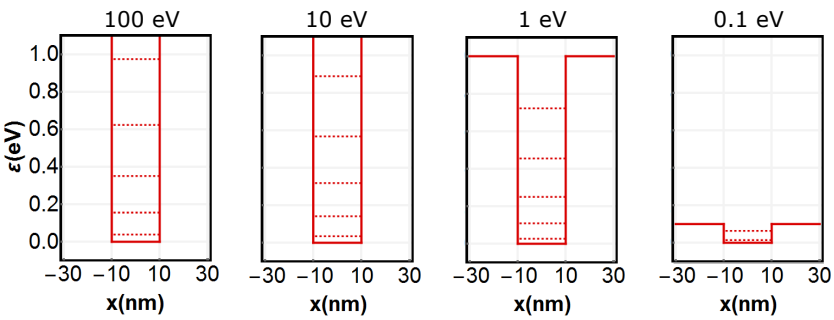


Figure 7: The lowest eigenvalues E_N in finite potential wells of different depth. The width of the wells is $L = 20$ nm, whereas the depth of the wells varies from 100 to 0.1 eV, as indicated above the plots. Continuous lines in red indicate the shape of the well. Dashed lines of red indicate the values of E_N in the corresponding wells. The calculations use effective mass $m^*_W = 0.023m_e$ inside the well and $m^*_B = 0.080m_e$ outside the well (corresponding to InAs and InP respectively).

Figure 7 considers a 20 nm wide quantum well with different depths. Energies of the lowest electronic states in the well are indicated by dashed lines crossing the well. It is apparent that the energies E_N shift slightly as the height of the barriers is changed. The number of states also shrinks with the well becoming more shallow. To explain this we examine Fig. 8 where the envelope wave functions of the two lowest states (F_1 and F_2) are sketched. It can be seen that in a deep well the functions F_N are essentially contained within the well, whereas when the well becomes more shallow they start to extend into the barrier regions.

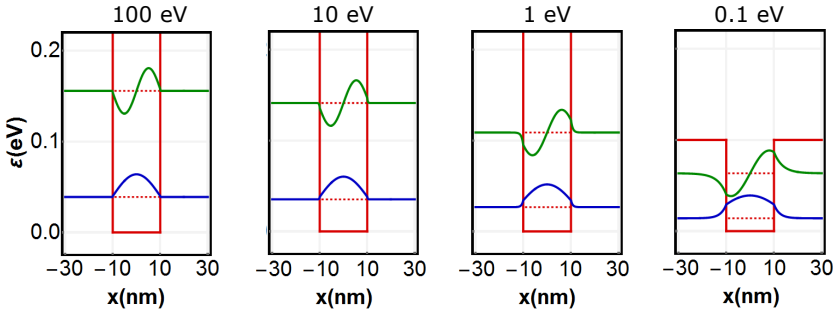


Figure 8: The two lowest eigenvalues E_N in finite potential wells of different depth along with the corresponding shape of the envelope wave functions. The width of the wells is $L = 20$ nm, whereas the depth of the wells varies from 100 to 0.1 eV, as indicated above the plots. Continuous lines in red indicate the shape of the well. Dashed lines of red indicate the values of E_N in the corresponding wells. The blue and green lines represent the envelope wave functions $F_1(x)$ and $F_2(x)$. The function values are scaled for plotting and offset by the corresponding state energies E_N. The calculations use effective mass $m^*_W = 0.023m_e$ inside the well and $m^*_B = 0.080m_e$ outside the well (corresponding to InAs and InP respectively).

What this tells us is that, by lowering the barrier height, the electrons effectively become less confined and therefore energies E_N go down. For energies above the barrier there are no bound states and therefore no localized envelope wave functions exist. The fact that functions F_N extend into the barrier regions is interpreted as a finite probability of finding the electron within the volume of space where its kinetic energy classically would not allow it to be. The implications of this fact are discussed in the following Section 2.4.

2.4 Tunneling

The Section 2.3 demonstrates that for potential wells of finite height and thickness the envelope wave functions of the electrons always somewhat penetrate the barrier. This small, but finite, probability for the electrons being into the classically forbidden barrier regions implies that there is a finite probability for them to reach the other side of finite barriers. This process of bypassing a potential barrier is called *tunneling*. The following introduces the so-called weak tunneling (weak tunnel-coupling) which assumes that the finite potential barrier thicknesses are thick enough not to affect the energies E_N too much. Also, the tunneling process is assumed to be elastic, meaning that the energy of the tunneling electron is preserved. These conditions are convenient, because they allow approaching the problem pertubatively and provide a rather straight forward way to estimate the tunnel rate Γ.

Here the problem is approached in line with Bardeen's theory for tunneling [45]. Two independent wells that are characterized by electric potentials $V_{\phi 1}(x)$ and $V_{\phi 2}(x)$ are introduced (see Figs. 9a and b). The solutions to these separate problems can be used to estimate tunnel rate Γ, similarly as done in Ref. [32]. The first well is chosen such that it hosts only one bound state with energy E_1. The tails of the corresponding envelope wave function F_1 are stretching into the barrier region, similarly as shown the previous section

(in Fig. 8). The other well is much longer and positioned at a slightly different location. It hosts many more states with energies ε_N because the electrons are not confined as much in the longer well as in the shorter one. Figure 9b shows how an envelope wave function χ_N of a state with energy $\varepsilon_N \approx E_1$ stretches into the barrier region in a similar fashion as F_1 in the shorter well.

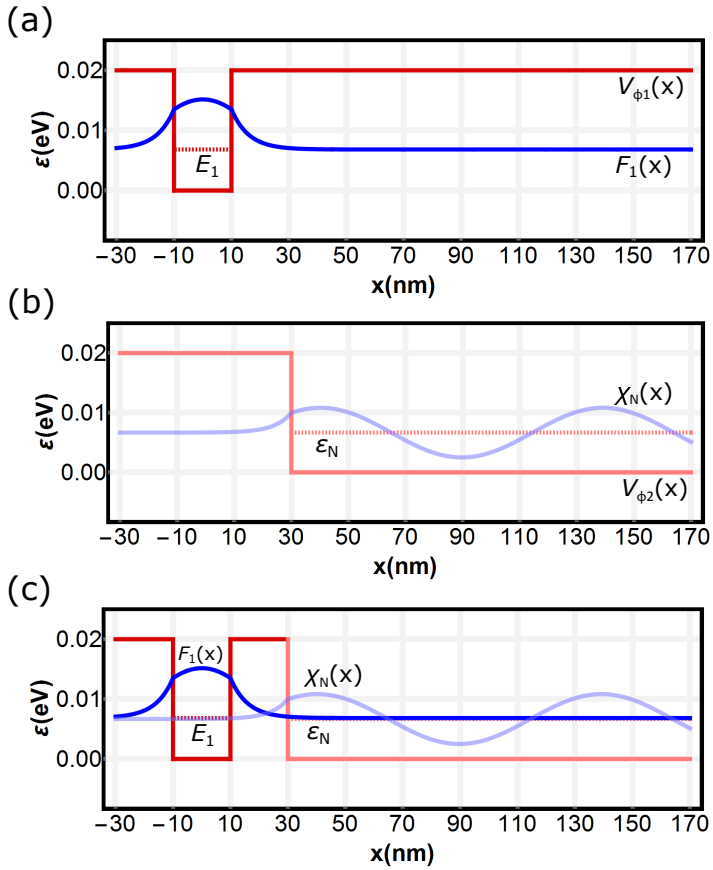


Figure 9: (a) An electronic state in a finite well, as given by the electrical potential $V_1(x)$. The length of the well is $L = 20$ nm and the depth of the well is $V_b = 20$ meV. One bound state with energy $E_1 \approx 6.8$ meV is present in the well. The state is characterized by an envelope wave function F_1 (scaled for plotting and offset from zero by E_1). (b) One of the electronic states in a finite well with a greater length $L = 2000$ nm, as given by potential $V_2(x)$. The depth of the well is $V_b = 20$ meV. A state at energy $\varepsilon_N \approx E_1$ is characterized by an envelope wave function χ_N (scaled for plotting and offset from zero by E_1). (c) A double well potential as given by $V_1 + V_2 - V_b$. The same functions F_1 and χ_N, as in (a) and (b), are plotted on top of each other. The functions F_1 and χ_N are not normalized with respect to each other. In (a), (b) and (c) the calculations use effective mass $m^*_W = 0.023m_e$ inside the well and $m^*_B = 0.080m_e$ outside the well (corresponding to InAs and InP respectively).

So both wells considered independently have certain analytical envelope wave function solutions F_1 and χ_N with numerically calculable state energies, however, these functions do not form exact basis for solutions when both wells are present at the same time (as shown in Fig. 9c). The problem is approached perturbatively by assuming that F_1 and χ_N

are approximate solutions to the double well profile when the electron is, respectively, either in the left or the right well. In this approximation, the tunnel rate Γ can be expressed [46] as

$$\Gamma = \frac{2\pi}{\hbar} \sum_N \mid M_{1N} \mid^2 \delta(E_1 - \varepsilon_N) \tag{12}$$

where δ is the Kronecker delta function (with a physical dimension of 0-dimensional density of states $[J^{-1}]$) and M_{1N} are tunneling matrix elements which measure the strength of the tunnel-coupling across the barrier. Bardeen's theory [45] in 1-dimension approximates M_{1N} by the following expression

$$M_{1N} = \frac{\hbar^2}{2m_B^*} \left(F_1 \frac{\partial \chi_N^*}{\partial x} - \chi_N^* \frac{\partial F_1}{\partial x} \right)_{x=x_0} \tag{13}$$

where $x = x_0$ is a point somewhere within the barrier region (not too close to either of the wells [47]) and the functions F_1 and χ_N are normalized such that $\int_{-\infty}^{\infty} \mid F_1 \mid^2 dx = 1$ and $\int_{-\infty}^{\infty} \mid \chi_N \mid^2 dx = 1$. The form of Eq. 13 has similarities to the probability current given in any basic quantum mechanics textbook. One can see that in order for the M_{1N} to be non-zero, the functions F_1 and χ_N have to have finite values and derivatives within the barrier region. Given the exponential decay of these functions in the barriers, the overall tunnel rate Γ is exponentially sensitive to the thickness of the barrier. It is, however, important to remember that this approximation relies on F_1 and χ_N each being close to zero at the other well, so the approximation is not expected to be valid for arbitrarily thin barriers where it would predict arbitrarily large Γ. The sum in the Eq. 12 indicates that the tunnel rate Γ increases with the number of states on the right well at energy $\varepsilon_N = E_1$ because all of them represent viable alternatives for tunneling into (or from).

In the limit where the states are packed densely enough so that they can be considered as a continuum of states, the functions χ_N are expected to be similar enough to give the same matrix element $M_{1N} = M$, so Eq. 12 can be simplified to

$$\Gamma = \frac{2\pi}{\hbar} \mid M \mid^2 \rho(E_1) \tag{14}$$

where ρ is density of states. This general form of the expressions for Γ (Eqs. 12 and 14) is also known as Fermi's golden rule and is typically applied for calculating current through tunnel-junctions in the weak coupling limit [48].

The above explanation considers one dimension, but the same general conclusions are valid for three dimensions too because the form of Eq. 12 does not change. What changes in three

dimensions is the form of the matrix element M_{1N} that is evaluated as a surface integral. Essentially the 3-dimensional envelope wave functions have to be integrated across a full cross-section of the tunnel-junction instead of calculating the value of Eq. 13 at a single point. However, it does not change the exponential sensitivity of Γ in the weak tunneling limit, as the value of the tunneling matrix element at each cross-section point is still exponentially sensitive to the barrier thickness.

In the experimental practice the tunnel rate Γ, however, is not calculated. Normally it is far easier to use theoretical models to extract the value of Γ from the experiment. The theoretical framework laid out in this section thus have a more illustrative purpose. It allows us to understand the basic principles and gives an idea of how Γ would depend on the dimensions of the tunnel-junction.

2.5 Electron spins

Even though spin is a widely known property of elementary particles, its presence in non-relativistic quantum mechanics is somewhat empirical. The spin property (possession of an intrinsic angular momentum) of elementary particles is only explained by the relativistic description of quantum mechanics [49] indicating that the concept is intrinsically related to principles of relativity [50]. The inclusion of spin effects in the non-relativistic description (used throughout this thesis) can be motivated by reviewing the non-relativistic limit of the Dirac equation [51] for a particle in electromagnetic fields and identifying the relevant terms for the Schrödinger equation.

In external magnetic field the spin property gives rise to the Zeeman effect [52] because the magnetic moment associated with the spin interacts with the field. Free electrons, similar to protons and neutrons, have a spin quantum number 1/2 and therefore have two possible spin angular momentum projections $S_z = \pm\hbar/2$ associated with the magnetic moment projections $\mu_{Sz} \approx \pm e\hbar/(2m_e) = \pm\mu_B$. This leads to a potential energy contribution $\Delta E_{Zeeman} \approx \pm\mu_B B$ called the *Zeeman energy*.

In the context of this thesis the electron spin is important because in the absence of magnetic field the spin introduces a degeneracy of electronic orbital states. In other words, each of the orbital quantum states, as described by the envelope wave functions, can be occupied by electrons of either spin projection. In combination with Coulomb blockade, the spin degeneracy plays an important role in QDs, as it increases the conductance maximum and shifts it in energy (see Section 3). The energetics of the corresponding QD state filling is explained in the following Section 2.6.

2.6 Energetics of state filling

The two contributing effects already discussed in Sections 2.1 and 2.3, namely, the charging and confinement effects, seem to produce a similar outcome in QDs - they both lead to a discretization of the electronic state energy spectrum. The net effect is that the electrochemical potential (which is the total energy needed to add an electron) of a QD increases rapidly with the number of electrons added. However, the two effects are of unrelated origin. On the one hand, adding an electron increases the electrostatic energy of the QD, which is the potential energy component of the electrochemical potential. On the other hand, the confinement effect restricts the kinetic energy of the added electron to a set of quantized values, which is the kinetic energy contribution to the electrochemical potential. The two effects also have a different dependency on the QD size. Whereas the charging effects scale with $1/C_\Sigma$ and therefore roughly inversely with the size of the QD, the confinement effects scale roughly inversely to the square of the QD size. In practice, this means that the confinement is expected to be dominant for small enough QDs and less important for bigger ones.

It seems obvious that the potential and kinetic energy contributions should add up in a *total energy* ϵ_{tot} (electrochemical potential), as is explained in Refs. [28, 53]. However, it might not be straight forward to envision how it would affect the state filling of a QD. Figure 10 attempts to clarify this graphically. Suppose we start out with a QD that has no net charge and the available kinetic energies, as set by the quantum confinement effect, are specified by a spectrum of spin-degenerate single electron orbital states energies $\{\varepsilon_i\}$ (see Fig. 10a). Adding an electron, say in the lowest orbital, requires kinetic energy ε_1. However, because adding the electron to the QD will also charge it up, there is going to be an additional potential energy cost of $U_{Q=-e} - U_{Q=0} = \Delta U_1 = e^2/(2C_\Sigma)$, according to the Eq. 6. So the total energy ϵ_{tot} is going to equal $\varepsilon_1 + e^2/(2C_\Sigma)$. Now suppose we add another electron in the same orbital state with the opposite spin. The new electron requires the same kinetic energy ε_1 as the previous one, however, because adding it increases the charge of the QD from $-e$ to $-2e$, the increase of the electrostatic energy will become $U_{Q=-2e} - U_{Q=-e} = \Delta U_2 = 3e^2/(2C_\Sigma)$. This means that adding the second electron in the same orbital requires a total energy ϵ_{tot} of $\varepsilon_1 + 3e^2/(2C_\Sigma)$. Generalizing this line of thinking one can show that the energy for adding the n-th electron in the i-th orbital requires energy ϵ_{tot} [53]

$$\epsilon_{tot} = \varepsilon_i + \frac{e^2}{C_\Sigma}\left(n - \frac{1}{2}\right) \tag{15}$$

Note that the difference in the total energies between adding the first and second electron in the same orbital state is e^2/C_Σ, which is also often referred to as the *charging energy* [28], and, according to this description, it should hold for double-filling any of the

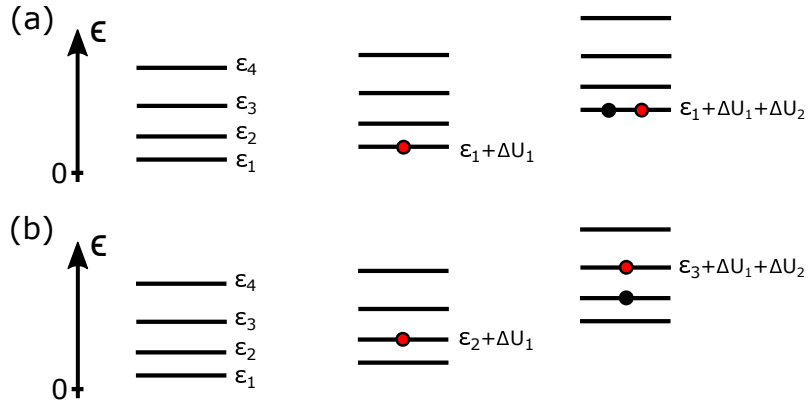


Figure 10: Electrochemical potential spectrum for zero, one and two electrons on a QD. The single particle orbitals are characterized by a spectrum for their kinetic energies $\{\varepsilon_i\}$. The newly added electrons are indicated in red. (a) considers a scenario where two electrons are sequentially added into the lowest orbital with kinetic energy ε_1. The corresponding total energies ϵ_{tot} needed to add each of those electrons are $\varepsilon_1 + \Delta U_1 = \varepsilon_1 + e^2/(2C_\Sigma)$ and $\varepsilon_1 + \Delta U_2 = \varepsilon_1 + 3e^2/(2C_\Sigma)$. (b) considers a scenario where two electrons are sequentially added into two different orbitals with kinetic energies ε_2 and ε_3. The corresponding total energies ϵ_{tot} needed to add each of those electrons are $\varepsilon_2 + \Delta U_1 = \varepsilon_2 + e^2/(2C_\Sigma)$ and $\varepsilon_3 + \Delta U_2 = \varepsilon_3 + 3e^2/(2C_\Sigma)$.

spin-degenerate orbital states (note that in this context the term *charging energy* refers to the electrostatic part of the additional energy needed rather than to the amount of energy needed to charge up a neutral object with the capacitance C_Σ). Generally, however, adding an electron may require occupying an energetically higher orbital state, in which case the additional energy cost required would also have an additional kinetic energy component corresponding to the given orbital state. The overall difference in the total energies (electrochemical potentials) between adding sequential electrons in orbitals x and y is $\varepsilon_y - \varepsilon_x + e^2/C_\Sigma$, also sometimes called the *addition energy* [53].

The principles described using the example of adding two electrons in the lowest orbital ε_1 are also valid for other situations. Suppose the first electron is added in the orbital with energy ε_2 but the second one is added in the orbital with energy ε_3 (see Fig. 10b). The total electrostatic energy increase is expected to be the same regardless of which orbital the electrons occupy, but the kinetic energy component of the electrochemical potential will change according to the orbital energies.

It is worth noting that the expression for U, as given by the Eq. 6, assumes that the capacitance C_Σ does not vary with the number of electrons added. This is a valid assumption for small metallic objects, in which the high density of conduction electrons enables effective screening of the electric field inside them. However, it is not generally true for semiconducting QDs, within which the electric field is not fully screened. As a result the total field energy of a charged semiconductor QD is generally bigger, which translates into a smaller capacitance C_Σ. Adding additional charges to a semiconducting object can effectively make it more metallic and therefore change its capacitance. In such a case the energy in-

creases due to charging cannot be calculated according to the Eq. 6 which considers only the geometric capacitance of the QD.

2.7 Electron reservoirs

Because free electrons are spin-1/2 particles, they are fermions obeying Pauli's exclusion principle. This allows a system of non-interacting electrons at thermodynamic (chemical and thermal) equilibrium to be characterized by the Fermi-Dirac distribution [54]

$$f(\epsilon) = \frac{1}{e^{(\epsilon-\mu)/(kT)} + 1}. \tag{16}$$

In terms of thermodynamic variables, electrochemical potential μ and absolute temperature T, the function $f(\epsilon)$ represents the probability of a state at energy ϵ being occupied (k is the Boltzmann constant). It predicts that the low energy states are going to be fully occupied ($f = 1$) whereas the high energy states will be unoccupied ($f = 0$). The transition between occupied and unoccupied states takes place around μ with the majority of change within a range of $\sim \mu \pm 2kT$.

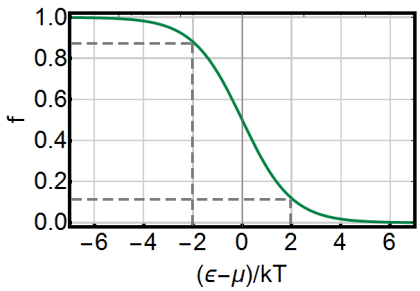


Figure 11: The value of the Fermi-Dirac distribution f as a function of energy $(\epsilon - \mu)$ relative to the typical thermal activation energy kT. The distribution has a value of $f = 0.5$ at $\epsilon = \mu$ at all temperatures T. The dashed lines indicate the values of f at $\epsilon - \mu = \pm 2kT$.

These ideas from statistical mechanics can also be directly applied to solid state theory because in solids the effective mass approximation allows us to treat electrons as non-interacting quasi-particles. All the possible ways that many electrons can occupy the available single particle quantum states in a solid can be viewed as a grand canonical ensemble with a fixed state structure and fixed μ and T. Consequently, under thermodynamic equilibrium electrons in solids can also be assumed to occupy states according to the Fermi-Dirac statistics.

Since QDs have a small number of strongly interacting conduction electrons it is unpractical to talk about thermal equilibrium. However, the contact leads, that the QD is coupled

to, are normally much bigger, possess higher densities of conduction electrons and also have bigger heat capacities. What this means is that transferring any relevant amount of energy or particles from the QD to the contact leads affects neither their μ nor T. Therefore the contact leads can be seen as relatively good fermion reservoirs close to thermal equilibrium and the Fermi-Dirac statistics can be an appropriate approximation. In fact, this a typical assumption in modeling transport through QDs and is also used in following Section 3.

3 Theoretical description of transport through quantum dots

The previous section (Section 2) reviewed some of the basics in solid state physics and electrostatics relevant to QDs. The spectrum of relevant phenomena is very rich and therefore the theoretical description is constructed based on many considerations. The aim of this section is to use the physical understanding from Section 2 as a basis for the introduction of QD models. Initially, a circuit model of a QD is introduced. Then attention is turned to the Master equation approach for modeling steady state transport through QDs in sequential electron tunneling approximation (SETA). This approach is relatively pedagogical and provides with a qualitatively accurate description of QDs. In addition, the Landauer-Büttiker approach is introduced as a reference.

The SETA was used to explain the results in Paper I. The studies in Papers II and III, however, use a more sophisticated version of the Master equation approach because the SETA fails to capture all of the relevant transport contributions at the given experimental conditions. The Landauer-Büttiker approach can model some of the features that the SETA does not capture, therefore its predictions are used as an illustrative example for developing a more complete picture of transport in QDs.

3.1 Circuit model

A single QD in a measurement circuit is normally considered to be electrostatically coupled to both contact leads and to a gate electrode. This is modeled by the absolute values of the coupling capacitances C_L, C_R and C_G (see Fig. 12) [29, 55]. Following the reasoning in Section 2.1, the overall capacitance of the QD can be expressed as a sum of all capacitances $C_L + C_R + C_G = C_\Sigma$ [29, 55]. In addition, the QD is also tunnel-coupled to the two contact leads for two-terminal current-voltage characterization. The tunnel-junctions are characterized by their tunnel rates Γ_L and Γ_R (explained in Section 2.4). The contact leads can be seen as electron reservoirs with set temperatures, T_L and T_R, and electrochemical potentials, μ_L and μ_R. When the two contact leads have different electrochemical potentials $\Delta\mu = \mu_L - \mu_R \neq 0$, the QD is said to be *electrically biased*. Similarly, if the temperatures of the two contact leads are different $\Delta T = T_L - T_R \neq 0$, the QD is said to be *thermally biased*.

In order to understand the properties of such a circuit, let us first assume that tunneling between the QD and the contact leads is not possible. In such a case the QD is electrostatically floating and its electrical potential ϕ_{QD} is set by the three potentials V_L, V_R and V_G according to the following expression.

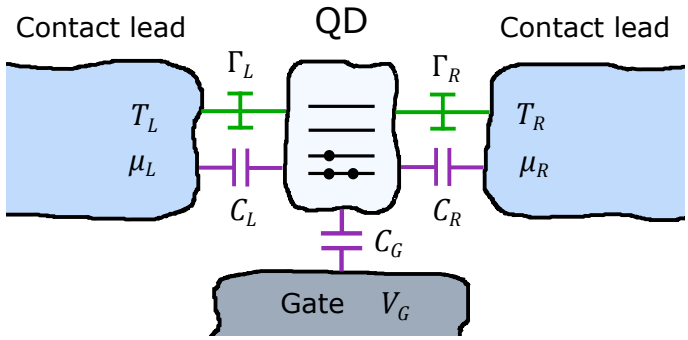


Figure 12: Graphical illustration of a QD (in white), contact leads (in blue), a gate electrode (in gray) and the effective electrical circuit elements connecting them. The QD is characterized by its electrochemical potential spectrum $\{\epsilon_{tot}\}_{i,n}$, whereas the metallic contact leads are characterized by their temperatures (T_L and T_R) and electrochemical potentials ($\mu_L = -eV_L$ and $\mu_R = -eV_R$). The gate electrode is characterized by its electrical potential V_G. The QD is tunnel-coupled to the contact leads with tunnel rates Γ_L and Γ_R (indicated by circuit elements in green) and electrostatically coupled to the contact leads with capacitances C_L and C_R (indicated by circuit elements in purple). Additionally, the QD is electrostatically coupled to the gate electrode with capacitance C_G.

$$\phi_{QD} = \frac{Q_{QD}}{C_\Sigma} = \frac{1}{C_\Sigma} \sum_{i=L,R,G} C_i V_i = \sum_{i=L,R,G} \alpha_i V_i \tag{17}$$

Here Q_{QD} is the charge induced (displaced) by the field between the all leads/electrodes and the QD (note that in Eq. 17 the net charge of the QD is assumed to be zero). The application of positive potentials V_i therefore lead to an increase of the QD potential ϕ_{QD}. The capacitance ratios $C_i/C_\Sigma = \alpha_i$ are sometimes called the lever arms because they describe the relative sensitivity of ϕ_{QD} to each V_i. The change of ϕ_{QD} in response to the lead/electrode potentials is a *gating effect*, because it allows us to modify the electron potential energies on the QD by means of an external electrical potential - exactly what the gate electrode is intended to do. The gating effect is accounted for by considering an additional term $-e\phi_{QD}$ in the expression for ϵ_{tot} so that

$$\epsilon_{tot} = \varepsilon_i + \frac{e^2}{C_\Sigma}\left(n - \frac{1}{2}\right) - e\sum_{j=L,R,G} \alpha_j V_j. \tag{18}$$

Note that although the expression for ϵ_{tot} depends on all three potentials, in practice the potentials V_L and V_R are kept at relatively low values and all gating is done by varying V_G.

When the tunneling between the gated QD and the contact leads is enabled, the QD will exchange electrons with the contact leads in a way that minimizes the total energy of the QD-lead system. The transfer of each electron will modify the electrostatic energy of the QD according to the rules described in Section 2.6. In short, the application of a positive gate potential $V_G > 0$ will lower the total energy of the states ϵ_{tot} enabling electrons to

tunnel into the QD from the contact leads. Whenever an additional electron enters the QD, it also increases the potential energy of the QD. Visa versa, whenever a negative gate potential $V_G < 0$ is applied, the total state energies ϵ_{tot} rise. This encourages electrons to tunnel out of the QD to the contact leads where they can have lower total energy, lowing the potential energy of the QD in the process.

3.2 Master equation approach

The main idea behind the Master equation approach to transport through QDs is to determine the probabilities of a QD being in various configurations (electron arrangements in orbital states). These probabilities depend on all parameters - the temperatures $T_{L,R}$, the tunnel rates $\Gamma_{L,R}$, the electrochemical potentials of the contact leads $\mu_{L,R}$ and the potential of the gate electrode V_G. The probability distribution among configurations can then be used to calculate charge and heat currents through the QD.

A convenient starting point for steady state description of QDs is the so-called sequential electron tunneling approximation (SETA). This approximation is assumed to be valid for weakly tunnel-coupled QDs that obey the condition $(\Gamma_L + \Gamma_R) \ll kT/h$, where h is the Plank's constant and T is the temperature of the contact leads [18, 28, 55]. As the name suggests, in this approximation only the sequential tunneling processes are considered and effects like co-tunneling and charge fluctuations [56, 57, 58] are neglected. This means that electrons on the QD are assumed to have set energies and the energy broadening due to the finite lifetime of electrons on the QD (according to the Heisenberg uncertainty principle [30]) is assumed to be unimportant. The success of this description is that it treats the charging and confinement effects exactly. Furthermore, the calculations of current through a QD are relatively simple because they rely on the evaluation of simple analytical algebraic expressions and their summation.

First, a relatively detailed summary of the theory for a simplified case is given. It reviews the main ideas behind the calculation of charge current (and also heat current) via electrically- and thermally-biased QDs occupied by one electron. The mathematical expressions in this case are relatively simple and therefore are easier to analyze. Further, the theoretical framework is generalized for the case of arbitrarily many single particle states on a QD. Similar expressions for the calculation of charge and heat current are given. Representative calculations for the case of two spin-degenerate orbital states are carried out in order to demonstrate the main differences from the singly occupied QD case. This is also a type of calculation that is used in Paper I to explain the thermoelectric behavior of a QD with a single spin-degenerate orbital. Finally, this section is concluded by a short discussion concerning higher order tunneling processes that are neglected in the SETA. Such processes have been considered in calculations presented in Paper II, for example, where the efficiency of QD heat engines is studied.

Single occupied spin-degenerate orbital

The simplest case that captures most of the relevant phenomena is the case of a spin-degenerate orbital state on the QD with a total energy of ϵ_1 (see Fig. 13). It is considered in one electron limit, which is equivalent to an assumption that the charging energy, for adding another electron, is much bigger than the typical thermal activation energy kT in the contact leads. In such a case there are only three different possible QD configurations - one with no electrons on the QD and two with a single electron on the QD (one for each spin state). The time-average probabilities of the QD being in these configurations are labeled P_0, P_{1u} and P_{1d} correspondingly. The sum of all probabilities has to be unity $P_0 + P_{1u} + P_{1d} = 1$ because the QD is assumed to definitely be in one of the three configurations. The probability of a single electron occupancy is given by a sum $P_1 = P_{1u} + P_{1d} = 1 - P_0$.

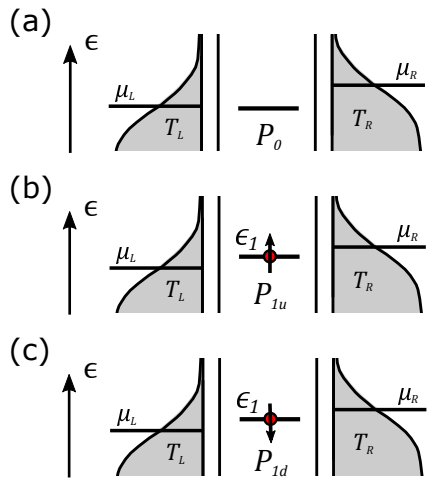


Figure 13: Schematic view of a QD with a spin-degenerate orbital in three different configurations - (a) with no electrons on the QD, (b) with a single spin up electron with energy $\epsilon_{tot} = \epsilon_1$ on the QD, and (c) with a single spin down electron with energy $\epsilon_{tot} = \epsilon_1$ on the QD. The probabilities of the QD being in the corresponding configurations are labeled by P_0, P_{1u} and P_{1d}. In (a), (b) and (c) the orbital state energy of the QD is represented by a black line in the center of the figures. The Fermi-Dirac distribution of the electronic state population in the contact leads is represented by the gray sectors on both sides of the QD. Both contact leads are characterized by their temperatures $T_{L,R}$ and electrochemical potentials $\mu_{L,R}$.

In a steady state it is assumed that the time-average probabilities of the QD being in either of the configurations are constant in time. Calculation of the probabilities requires consideration of the transition rates between the different possible QD configurations. An electron tunneling in or out of the QD changes its configuration, so the time-average probability of any given configuration being occupied should be proportional to the sum of transition rates from all other possible QD configurations to the given one. The probabilities P_0 and P_1 can thus be expressed as

$$P_0 = \frac{W_{1\to 0}}{W_{1\to 0} + W_{0\to 1}} \qquad P_1 = \frac{W_{0\to 1}}{W_{1\to 0} + W_{0\to 1}} \tag{19}$$

where $W_{1\to 0}$ is the total transition rate for an electron to tunnel out of a singly occupied QD, and $W_{0\to 1}$ is the total transition rate for an electron (of either spin) to tunnel into an empty QD. The denominator is given by the sum of both transition rates as it ensures the normalization of probabilities. The rates themselves depend not only on tunnel rates via both tunnel-junctions $\Gamma_{L,R}$ but also on the population of the electronic states in the contact leads $f_{L,R}$, as given by the values of Fermi-Dirac distributions at energy ϵ_1.

$$W_{1\to 0} = \Gamma_L(1 - f_L) + \Gamma_R(1 - f_R) \qquad W_{0\to 1} = 2(\Gamma_L f_L + \Gamma_R f_R) \tag{20}$$

Both transition rates $W_{1\to 0}$ and $W_{0\to 1}$ are sums of contributions from both contact leads, which is indicated by the indices L and R. Each term in the sums is proportional to the tunnel rate of the corresponding tunnel-junction. In addition, since tunneling out of the QD requires a free electronic state at energy ϵ_1 in a contact lead, the transition rate for electron tunneling out through either of the tunnel-junctions is also proportional to $(1-f)$ in the corresponding contact lead. Similarly, because tunneling into the QD requires an electron from a contact lead at energy ϵ_1, the transition rate for electrons tunneling in through either of the tunnel-junctions is also proportional to f in the corresponding contact lead. The factor of 2 in $W_{0\to 1}$ originates from the spin-degeneracy of the orbital state.

Charge current

The probabilities P_0 and P_1 can be used to calculate current I through the QD. Because this description already ensures current conservation, it is enough to determine the current only across one of the tunnel-junctions (in this case the left junction)

$$I_{L\to R} = e\,\Gamma_L[(1 - f_L)P_1 - 2f_L P_0] \tag{21}$$

Here, the positive I direction (opposite to the electron flow direction) is defined from left to right. The current is proportional to the value of elementary charge and the tunnel rate of the left tunnel-junction. The two contributing terms, containing $(1 - f_L)$ and f_L, represent current contributions in the opposite directions. Each of the terms can be understood as a multiplication of two probabilities; the process of an electron tunneling out requires an electron in the QD (with probability P_1) and an empty state in the left contact lead (with probability $(1 - f_L)$), whereas the opposing process requires an electron in the left contact lead (with probability f_L) and an empty state in the QD (with probability P_0). The probabilities are multiplied to find out the resulting probabilities for the tunneling in either of the directions.

In order to simplify Eq. 21 the expressions for P_0 and P_1 (Eq. 19 and 20) are plugged into Eq. 21. After the simplification one can arrive at the following expression

$$I = \frac{2e\Gamma_L\Gamma_R(f_R - f_L)}{\Gamma_R(1+f_R) + \Gamma_L(1+f_L)}. \tag{22}$$

Equation 22 clearly demonstrates that the current I is ultimately always driven by the difference in the contact lead electronic state occupancy $(f_R - f_L)$ at the energy ϵ_1 (the total energy required to occupy the spin-degenerate orbital state of the QD). In order to understand how a non-zero $(f_R - f_L)$ occurs, one has to look at how f_R and f_L depend on their parameters $T_{L,R}$ and $\mu_{L,R}$.

$$f_L(\epsilon_1) = \frac{1}{e^{(\epsilon_1-\mu_L)/(kT_L)}+1} \qquad f_R(\epsilon_1) = \frac{1}{e^{(\epsilon_1-\mu_R)/(kT_R)}+1} \tag{23}$$

So, with a fixed ϵ_1 there are essentially two ways to drive the current I: either by applying an electrical bias to the QD, so that $V = V_L - V_R = -(\mu_L - \mu_R)/e$, or, by applying a thermal bias to the QD, so that $\Delta T = T_R - T_L$. The expression for ϵ_1 is the same in both f_L and f_R and has the following form

$$\epsilon_1 = \varepsilon_1 + \epsilon_0 + \frac{e^2}{C_\Sigma}\left(n - \frac{1}{2}\right)_{n=1} - e\sum_{j=L,R,G}\alpha_j V_j \tag{24}$$

Here $n = 1$ because we consider only one electron on the QD. The parameter ϵ_0 is introduced to set a potential energy reference level. It formally allows us to make the simplifying assumption $\epsilon_0 = -\Delta U_1 = -e^2/(2C_\Sigma)$, which does nothing more than offsets the reference energy. Therefore, for the one electron case Eq. 24 can be simplified to

$$\epsilon_1 = \varepsilon_1 - e\sum_{j=L,R,G}\alpha_j V_j \tag{25}$$

Equations 22, 23, 24 provide analytic expressions for calculating I. We can use them to look at what they predict in some special cases. A good starting point is a commonly used linear response coefficient for characterizing QD - the differential conductivity $g = dI/dV$ as a function of V_G. It is possible to derive an analytic expression for g at zero electrical and thermal biases, g_0. However, because it contains trigonometric functions, there is more than one way to write it down. Here, a particular form is chosen because it illustrates the properties of g_0 in a convenient way

$$g|_{V=\Delta T=0} = g_0 = \frac{2e^2\tilde{\Gamma}}{kT\left[3+2\sqrt{2}\cosh\left(\frac{e\alpha_G V_G}{kT}+\frac{\ln 2}{2}\right)\right]} \tag{26}$$

where $\tilde{\Gamma} = \Gamma_L\Gamma_R/(\Gamma_L+\Gamma_R)$. The only V_G dependency is in the hyperbolic cosine function, which itself is an even function with its minimum at $(2e\alpha_G V_G + kT\ln 2) = 0$. Being in the denominator, the hyperbolic cosine yields a symmetric peak of g_0 around a gate potential value $V_{G0} = -kT\ln 2/(2e\alpha_G)$. This analytic result implies that peak position energy is expected to shift linearly with T, which is an interesting feature of QDs with spin-degenerate states. The peak value $g_0(V_{G0}) = g_{max}$ reaches

$$g_{max} = \frac{2(3-2\sqrt{2})e^2\tilde{\Gamma}}{kT} \tag{27}$$

where the value of the numerical constant $2(3-2\sqrt{2}) \approx 0.3431$ [28]. A few examples of $g(V_G)$ at different temperatures T are shown in Fig. 14a. The peak value g_{max} scales inversely with T, whereas the width at half maximum of the peak increases with T as $\Delta V_G = 2kT\,\mathrm{arccosh}\left(2+3\sqrt{2}/4\right)/(e\alpha_G)$. In effect, increasing temperature makes the conduction peak lower and less sharp, which is a consequence of the thermal excitation of electrons in the contact leads (thermal smearing). This is why carrying out measurements at lower T makes it easier to distinguish the conductance peaks.

What Eq. 26 also shows is that g_0 scales with $\tilde{\Gamma}$. However, because we have considered the SETA, valid for $(\Gamma_L+\Gamma_R) \ll kT/h$, as done in Ref. [28], it generally bounds the applicability of this result to low g_{max} values because the tunnel rates Γ_L and Γ_R are limited. Even in the best case, when the tunnel rates are the same ($\Gamma_L = \Gamma_R$), this condition implies that the values of g_{max} must satisfy $g_{max} \ll (3-2\sqrt{2})\,e^2/h \approx 0.17\,e^2/h$. If the tunnel rates are not the same ($\Gamma_L \neq \Gamma_R$), the applicability of Eq. 26 is bound to even lower g_{max} values, which is limiting for modeling QDs in realistic experimental conditions. Here it is important to note that sometimes a more flexible condition $(\Gamma_L+\Gamma_R) \ll kT/\hbar$ (with $\hbar$ instead of h) is also used, which yields the condition $g_{max} \ll e^2/h$ instead. This is also why $g_{max}/(e^2/h)$ is often used as a simple measure of the strength of the tunnel-coupling and a measure of whether the SETA is applicable.

Another characteristic of QDs is their unique thermoelectric response. As an example for this Fig. 14c shows the differential current dI in response to an infinitesimal thermal bias $d(\Delta T) = \delta T$, as a function of the gate potential V_G. Using the same Eqs. 22, 23, 24 it is possible to arrive at a simplified expression for $dI/\delta T$.

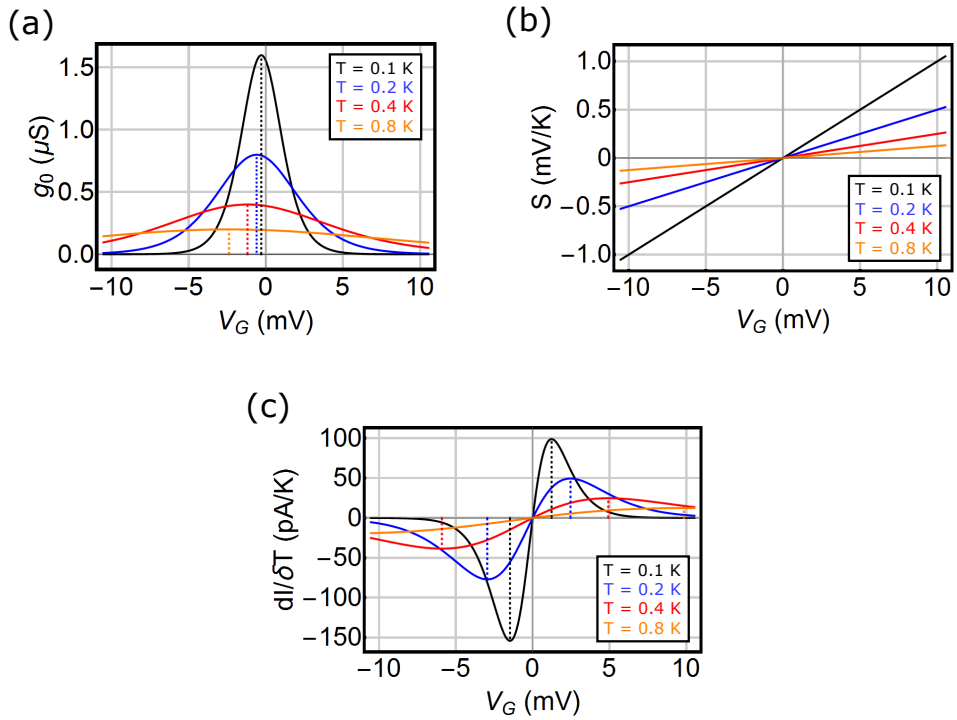


Figure 14: Calculated (a) conductivity g_0, (b) Seebeck coefficient $S = dV_{th}/\delta T$, and (c) $dI/\delta T$ of a QD with a single spin-degenerate orbital state as a function of the gate potential V_G (given by Eq. 22). In (a) the vertical dashed lines indicate the gate potential value of the peak conductivity that appears at $V_{G0} = -kT \ln 2/(2e\alpha_G)$. In (c) the vertical dashed lines indicate the gate potential value of the two peaks in $dI/\delta T$ that appear at $V_{G1,2} = B_{1,2}kT/(2e\alpha_G)$. The calculation parameter values are $\Gamma_L = \Gamma_R = 500$ MHz, $\mu_L = \mu_R = \varepsilon_1 = 0$ meV, $\Delta T = 0$, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the temperature of the contact leads $T = T_L = T_R$, which are 0.1 (black), 0.2 (blue), 0.4 (red) and 0.8 K (orange).

$$\left.\frac{dI}{\delta T}\right|_{V=\Delta T=0} = \frac{2e^2\tilde{\Gamma}\alpha_G V_G}{kT^2\left[3+2\sqrt{2}\cosh\left(\frac{e\alpha_G V_G}{kT}+\frac{\ln 2}{2}\right)\right]} = g_0 \cdot \frac{\alpha_G V_G}{T} = g_0 \cdot S \qquad (28)$$

As can be seen, the entire expression can be conveniently expressed in terms of g_0. The additional multiplier $\alpha_G V_G/T$ plays the role of the Seebeck coefficient S. The mathematical form of S is a well known result for QDs in the SETA [18]. As a linear response coefficient it measures the amount of thermovoltage dV_{th} in response to an infinitesimal thermal bias δT. It has a linear dependency on V_G and crosses zero at $V_G = 0$ where (according to assumptions made above) ϵ_1 aligns with the electrochemical potentials $\mu_L = \mu_R$ (see Fig. 14b).

One can see in Fig. 14c that the two extrema of $dI/\delta T = g_0 S$ are of different amplitude. This is a direct consequence of the shift of the g_0 peak due to the degeneracy whereas S remains

an odd function with respect to $V_G = 0$. It can also be seen from the Eq. 28 that the two extrema scale inversely with T, similar to g_{max}. The peak $dI/\delta T$ values reach $A_{1,2} \cdot e\tilde{\Gamma}/T$ at $V_{G1,2} = B_{1,2} \cdot kT/(e\alpha_G)$ where the numerical constants are given by $A_1 \approx -0.3847$, $A_2 \approx 0.2466$, $B_1 \approx -1.715$ and $B_2 \approx 1.429$.

The two linear response coefficients g_0 and S can be combined in a figure $g_0 S^2/4 = (dI/\delta T) \cdot (dV_{th}/\delta T)/4 = dP_{th}/(\delta T)^2/4$ that is called the power factor [59]. In this case it describes the QD's ability to produce thermoelectric power P_{th} per $(\Delta T)^2$ in the linear response regime where the filling factor is $1/4$. The resulting $g_0 S^2/4$ has two positive peaks of different magnitude that reach maximal values $A_{3,4} \cdot k\tilde{\Gamma}/T$ at $V_{G3,4} = B_{3,4} \cdot kT/(e\alpha_G)$ where the numerical constants are given by $A_3 \approx 0.2037$, $A_4 \approx 0.1148$, $B_3 \approx -2.533$ and $B_4 \approx 2.317$.

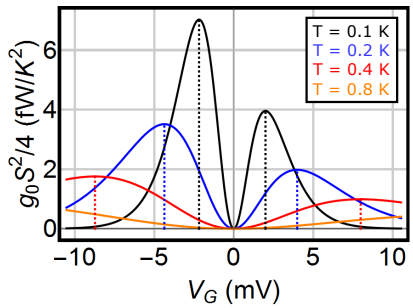


Figure 15: Calculated $g_0 S^2/4$ of a QD with a single spin-degenerate orbital state as a function of the gate potential V_G. The vertical dashed lines indicate the gate potential value of the peaks at $V_{G3,4} = B_{3,4} kT/(2e\alpha_G)$. The parameter values are the same as in Fig. 14: $\Gamma_L = \Gamma_R = 500$ MHz, $\mu_L = \mu_R = \varepsilon_1 = 0$ meV, $\Delta T = 0$, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the temperature of the contact leads $T = T_L = T_R$, which are 0.1 (black), 0.2 (blue), 0.4 (red) and 0.8 K (orange).

However, it is important to remember that both g_0 and S are linear response coefficients and therefore cannot describe a QD fully. The transport through QDs at low enough T can easily be driven into a nonlinear regime, therefore the full Eq. 22 generally has to be used. So, the quantities g_0 and S are order of magnitude predictions for I and V_{th}, given V and ΔT respectively. They, therefore, mainly serve a pedagogical purpose of illustrating the qualitative behavior of QDs under electrical and thermal biases.

Energy conversion

Within the framework of this section passing a current I through an electrically biased QD ($V \neq 0$) always relates to energy conversion from electrostatic energy to thermal energy, or visa versa. This is because an electron that has traversed a biased QD has received some thermal energy from one contact lead and has deposited a different amount of thermal energy at the other contact lead. The heat power that leaves the left contact lead through the QD is equal to the thermal energy $(\epsilon_1 - \mu_L)$ of each leaving electron multiplied by the frequency I/e they pass the QD

$$J_L = -(\epsilon_1 - \mu_L)\frac{I}{e}. \tag{29}$$

Similarly the heat power that the right contact lead receives through the QD is

$$J_R = -(\epsilon_1 - \mu_R)\frac{I}{e}. \tag{30}$$

The minus signs in the Eqs. 29 and 30 refer to the fact that the current direction is opposite to the electron flow direction. The overall converted power is the difference between J_R and J_L

$$P = J_R - J_L = (\mu_R - \mu_L)\frac{I}{e} = VI. \tag{31}$$

When $P > 0$, more heat is being deposited at the right contact lead than being subtracted from the left. The additional heat comes from the electrostatic energy being dissipated. Conversely, when $P < 0$, heat is being converted into electrostatic energy and thermoelectric power generation takes place. An example of an electrically and thermally biased QD is shown in Fig. 16. It shows that when $\Delta T = 0$, electrostatic energy is only converted into heat irrespective of the energy ϵ_1 controlled by V_G. However, when $\Delta T \neq 0$, heat-to-electrostatic energy conversion appears for a range of V_G.

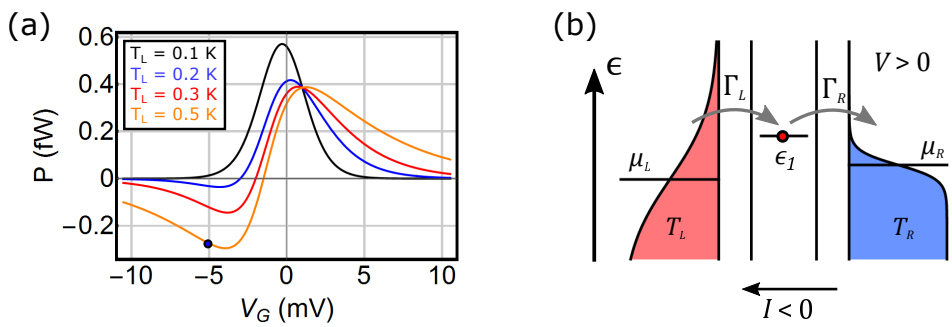


Figure 16: (a) Power conversion $P = IV$ in a single spin-degenerate orbital state of a QD as a function of the gate potential V_G. Negative conversion ($P < 0$) indicates heat-to-electrostatic energy conversion. The parameter values are $V = 20$ μV, $\Gamma_L = \Gamma_R = 500$ MHz, $\varepsilon_1 = 0$ meV, $T_R = 0.1$ K, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the temperature of the left contact lead T_L and are 0.1 (black), 0.2 (blue), 0.3 (red) and 0.5 K (orange). (b) Schematics of a QD corresponding to the conditions indicated by a dot in (a). T_L is higher than T_R, which is indicated by the different shapes for Fermi-Dirac distributions (colored in red and blue correspondingly). Electrical bias $V > 0$, however, the resulting current $I < 0$ due to the thermal bias $\Delta T \neq 0$. Note that the positive current direction is defined opposite to the electron flow direction, therefore the current, indicated in the schematics, is negative.

In the generation mode the efficiency η can be determined as a ratio between the generated power $P_{th} = -P = -IV$ and the heat power leaving the hot contact lead. For the case sketched in Fig. 16, where $T_L > T_R$, the efficiency reads

$$\eta = \frac{P_{th}}{J_L} = \frac{eV}{\epsilon_1 - \mu_L}. \tag{32}$$

The power generation persists while I runs against the electrical bias V. According to the Eq. 22 this remains the case while $f_L(\epsilon_1) > f_R(\epsilon_1)$. Using the explicit forms of f_L and f_R one can show that this condition is in fact equivalent to a condition $eV/(\epsilon_1 - \mu_L) < \Delta T/T_L = \eta_C$, where η_C is the Carnot efficiency. In other words, the fact that thermoelectric conversion efficiency is $\eta < \eta_C$ naturally comes out of the Eq. 22 and the properties of the Fermi-Dirac distributions. At the point when the given condition is violated I inverts direction and the QD enters the dissipation mode. This is precisely the limit experimentally investigated in Paper II. The results show, however, that effects beyond sequential tunneling increase the heat current and somewhat reduce the achievable η.

Multiple orbitals

A QD with N_{sps} single particle states can be in $2^{N_{sps}}$ different configurations, therefore the expressions for charge and heat currents quickly become long sums that are not convenient to write out explicitly. The probability of occupying each configuration can be represented by an element P_j in a vector $\mathbf{P}$ that is $2^{N_{sps}}$ elements in length [28, 55]. All the possible QD configurations that are different by one electron tunneling in or out are coupled. Similarly as for the singly occupied QD case, the coupling between different configurations i and f is described by transition rates $W_{i\to f} = W_{fi}$. In this case the transition rates make up a matrix $\mathbf{W}$. The elements of the matrix $\mathbf{W}$ that correspond to the transitions between the uncoupled configurations are zero. In steady state the probabilities are found by solving a homogeneous linear equation system [55]

$$\mathbf{W}\cdot\mathbf{P} = 0 \tag{33}$$

along with the normalization condition $\sum_m P_m = 1$. In a general form the transition rate matrix elements W_{fi} can be written as either

$$W_{fi}^{in} = \Gamma_{fi}^{Lin} f_L(\xi_{fi}) + \Gamma_{fi}^{Rin} f_R(\xi_{fi}) \tag{34}$$

or

$$W_{fi}^{out} = \Gamma_{fi}^{Lout}[1 - f_L(\xi_{fi})] + \Gamma_{fi}^{Rout}[1 - f_R(\xi_{fi})] \tag{35}$$

depending on whether the configurations i and f are coupled by an electron tunneling *into* or *out* of the QD. Note that the matrices $\mathbf{\Gamma^{in}}$ and $\mathbf{\Gamma^{out}}$ for either of the tunnel-junctions are transposed to each other because the role of the initial and final configurations is then inverted. Tunnel rates for transitions between the coupled configurations $\Gamma_{fi}^{L,R}$ (for tunneling in and out) are chosen constants $\Gamma_{L,R}$, whereas the tunnel rates between the uncoupled configurations are manually set to 0. In fact, it holds true for all variations of i and f that no two elements W_{fi}^{in} and W_{fi}^{out} have finite values at the same time, therefore the full matrix can also be obtained by summing the two elements

$$W_{fi} = W_{fi}^{in} + W_{fi}^{out} \tag{36}$$

The energy at which the Fermi-Dirac distributions of the elements W_{fi} must be evaluated is labeled ξ_{fi} and have the following form

$$\xi_{fi} = \epsilon_0 + \varepsilon_{fi} + \frac{e^2}{C_\Sigma}\left(n_{fi} - \frac{1}{2}\right) - e\sum_{j=L,R,G}\alpha_j V_j \tag{37}$$

where ε_{fi} is the kinetic energy of the tunneling electron (in or out) and n_{fi} is the number of electrons on the QD, inducing the electron that tunnels (in or out). The constant ϵ_0 (just as done for the single spin-degenerate orbital case) sets the potential energy reference level and is chosen as $\epsilon_0 = -\Delta U_1 = -e^2/(2C_\Sigma)$ to simplify Eq. 37 to

$$\xi_{fi} = \varepsilon_{fi} + \frac{e^2}{C_\Sigma}n_{fi} - e\sum_{j=L,R,G}\alpha_j V_j \tag{38}$$

Also similarly as it was introduced for a single spin-degenerate orbital in Eq. 21, the current can be found by summing up all relevant current contributions in both directions

$$I_{L\to R} = e\sum_{i,f} P_f\{\Gamma_{fi}^{Lout}[1 - f_L(\xi_{fi})] - \Gamma_{fi}^{Lin} f_L(\xi_{fi})\} \tag{39}$$

At the same time, the heat power leaving the left contact lead through the QD is

$$J_L = -\sum_{i,f}(\xi_{fi} - \mu_L)P_f\{\Gamma_{fi}^{Lout}[1 - f_L(\xi_{fi})] - \Gamma_{fi}^{Lin} f_L(\xi_{fi})\} \tag{40}$$

and the heat power that the right contact lead is receiving through the QD is

$$J_R = -\sum_{i,f}(\xi_{fi} - \mu_R)P_f\{\Gamma_{fi}^{Lout}[1 - f_L(\xi_{fi})] - \Gamma_{fi}^{Lin}f_L(\xi_{fi})\}. \quad (41)$$

Overall, there are a number of additional effects that come into play when one considers a larger spectrum of possible QD states. In order to illustrate this, the following considers two spin-degenerate orbital states (total number states $N_{sps} = 4$) which allow the QD to be in 16 different configurations. Figure 17a demonstrates current I through the QD as a function of both, the gate potential V_G and the electrical bias V. The most important feature of this graph is the emergence of rhombus-shaped regions in the (V_G, V) plane around $V = 0$ where current is blocked ($I = 0$). Plots of this type are sometimes referred to as charge stability diagrams. The regions with blocked current are known as *Coulomb diamonds* [60] which emerge a result of finite addition energies for adding electrons on the QD, as explained in Section 2.6.

By increasing V beyond the diamond edges, I increases in a step-wise manner. This can be best seen in Fig. 17b, which plots the differential conductance $g = dI/dV$ instead. The steps in I show up as lines of non-zero differential conductance (conduction lines). The various lines can be interpreted using a set of illustrations given in Fig. 17c. Illustration A corresponds to an unbiased condition where $\mu_L = \mu_R$ ($V = 0$). The lowest orbital state is occupied by a single electron, but no conduction is possible because all states at the same energy in the contact leads are occupied. Increasing V (symmetrically) eventually leads to a situation shown by illustration B, where μ_L approaches the energy of the occupied orbital state. This is the condition at which the conduction starts, as empty states in the left contact lead become available for tunneling out of the QD. The subsequent empty state left in the QD can then be filled by another electron tunneling in from the right contact lead, and so forth. The scenario described in illustration B continues until increasing V lets μ_R approach another, energetically higher, orbital state of the QD, corresponding to illustration C. At that point, in addition to electron transport through the energetically lower orbital state, there is now another alternative - when the lowest state is emptied, another electron from the right contact lead has an additional option of entering the higher (excited) orbital state instead. This additional transport channel increases the overall current through the QD and thus leads to another step in I. Continuing to increase V would eventually provide the right contact lead electrons with enough energy to load a second electron on the QD, which would further increase the number of QD configurations participating in the transport. Overall, the rule for conduction at low enough T is that V has to be big enough such that there is at least one available orbital state within an energy window between electrochemical potentials μ_L and μ_R.

The corresponding linear response coefficients g_0, $dI/\delta T$ and S as functions of V_G are shown in Fig. 18. First, Fig. 18a shows a set of conduction peaks in $g_0(V_G)$. At low T they resemble four independent peaks, each looking like the peak in Fig. 14a. The difference is that se-

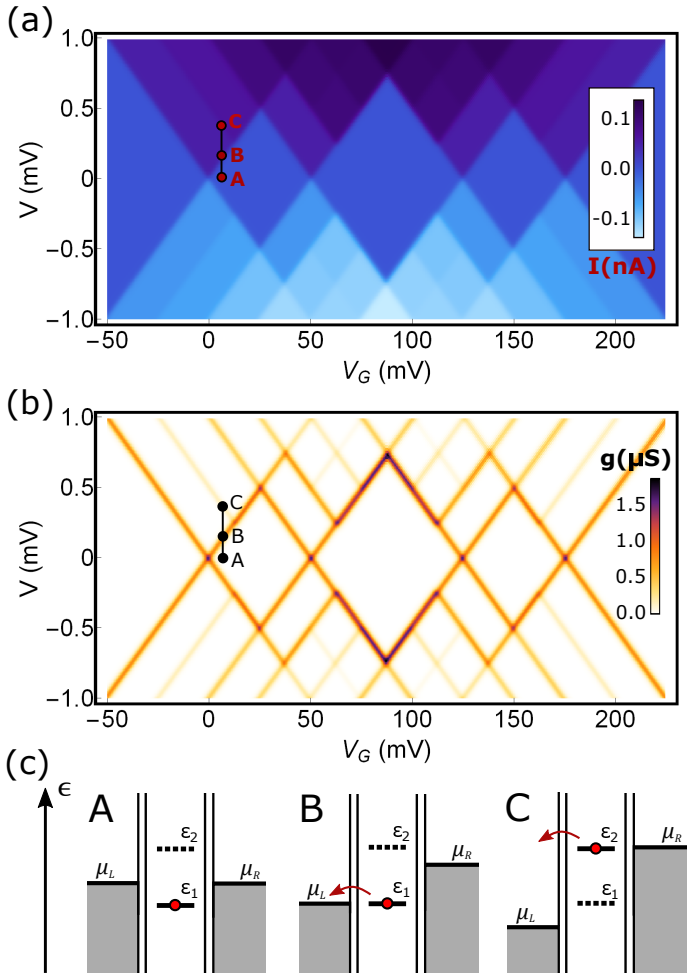


Figure 17: Calculated current I in (a) and differential conductance $g = dI/dV$ in (b) through a QD with two spin-degenerate orbital states as a function of the gate potential V_G and the electrical bias V. Calculation parameter values are $e^2/C_\Sigma = 1$ meV, $\varepsilon_1 = 0$ meV, $\varepsilon_2 = 0.5$ meV, $\Gamma_L = \Gamma_R = 500$ MHz, $T = T_L = T_R = 0.1$ K, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. (c) Three illustrations A, B and C demonstrating three different electrical bias conditions of a single occupied QD. Temperature T of the contact leads are assumed to be low enough such that electrons can be assumed to fill up the states in the contact leads up to the electrochemical potentials μ_L and μ_R. The illustration A corresponds to an unbiased QD ($\mu_L = \mu_R$) with one electron occupying ε_1. No current can flow because electronic states in the contact leads at energy ε_1 are occupied. The illustration B corresponds to a biased QD ($V = -(\mu_L - \mu_R)/e > 0$) where μ_L approaches ε_1 and the electron can be evacuated from the QD to the left contact lead and the current flow is enabled. The illustration C corresponds to a biased QD where μ_R approaches ε_2 which enables the QD to pass a current using also the higher orbital. Note that the energy needed for the second electron to enter the QD, $(\varepsilon_1 + e^2/C_\Sigma$, is not within the sketched energy range.

quential peaks shift in opposite directions. This altering behavior results from degeneracy and is a manifestation of the electron-hole symmetry for filling up the spin-degenerate orbital states. At higher T, however, the behavior of g_0 peaks is modified due to the fact that the thermally excited electrons in the contact leads have enough energy to enable transport via several QD configurations, not just one. It manifests in the fact that g_0 remains non-

negligible also in between the conduction peaks. The peak values and the magnitude of the peak shifts are also affected in a non-trivial way depending on the addition energies.

Figure 18b shows the corresponding results for $dI/\delta T$. Similarly as for g_0, the low T behavior of $dI/\delta T$ demonstrates four independent wiggles (like in Fig. 14c) with negligible magnitude in between them. The asymmetric height of the positive and negative extrema exchange roles from one wiggle to the next, which is consistent with the altering directions of the g_0 peak shifts. At higher T the magnitude of $dI/\delta T$ never settles at zero in between the wiggles and the magnitude of the wiggle extrema is also affected in a nontrivial way depending on the addition energies.

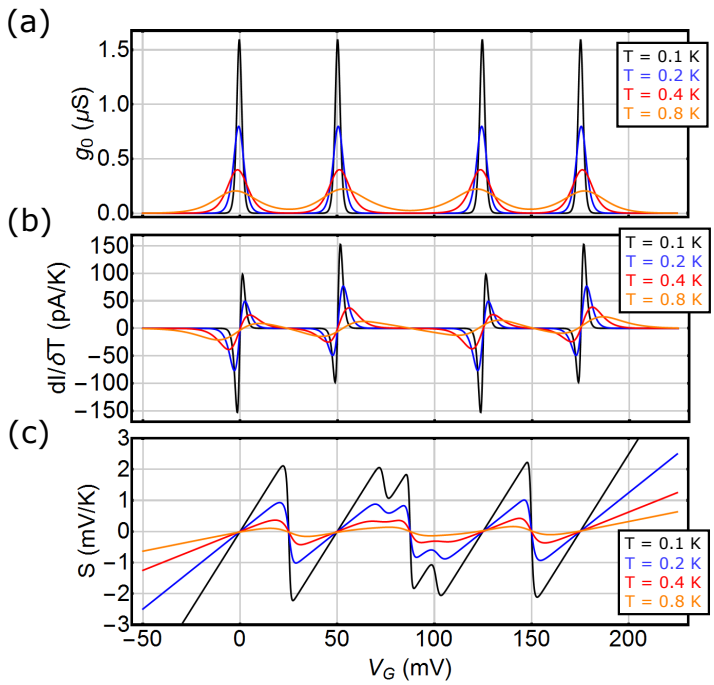


Figure 18: Calculated g_0 in (a), $dI/\delta T$ in (b) and S in (c) of a QD with two spin-degenerate orbital states as a function of the gate potential V_G. Calculation parameter values are $\epsilon^2/C_\Sigma = 1$ meV, $\varepsilon_1 = 0$ meV, $\varepsilon_2 = 0.5$ meV, $\Gamma_L = \Gamma_R = 500$ MHz, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the temperature of the contact leads $T = T_L = T_R$, which are 0.1 (black), 0.2 (blue), 0.4 (red) and 0.8 K (orange).

In a way, the most interesting behavior can be seen in the V_G dependence of S, which is shown in Fig. 18c. Again, at low T the behavior of $S(V_G)$ that shown in Fig. 14b; i.e. S scales linearly with V_G and crosses zero when the electron energy aligns with electrochemical potentials μ_L and μ_R. However, unlike for the case of g_0 and $dI/\delta T$, the low T behavior of S also shows something that is not captured when considering a singly occupied QD, namely, there are abrupt jumps in S in between the linear parts. The abruptness of these jumps are shown to converge to a saw-tooth behavior when approaching the limit of $T = 0$ [18]. Also the opposite is true, an increasing T smoothens out the jumps. In the two spin-

degenerate orbital model, that we consider as an example here, the jumps that invert the sign of S relate to transitions between different QD configurations with different numbers of electrons. The jumps that do not invert sign involve the same number of electrons on the QD and relate to transitions to excited state configurations. It is also important to note that the linear increase of S stretching away from the considered V_G region predicts an infinitely increasing S. This is a result of the finite size of the QD configuration space as there are no other orbital states above and below the energies of the two orbitals that are considered. However, detecting this increasing S fully in practice would be impossible, as it would require a rapidly increasing impedance of a voltmeter, due to the impedance of the QD itself rising rapidly away from the orbital state energies.

Further increase of T leads to more and more QD configurations contributing to transport at the same time. For comparison, the equivalent plots for linear response coefficients g_0, $dI/\delta T$ and S as functions of V_G at higher T are shown in Fig. 19. In case of g_0 the peaks start to merge together, whereas $dI/\delta T$ and S gradually lose their characteristic oscillatory behavior with the corresponding zero crossings. This also illustrates that low enough temperature $T \ll e^2/(kC_\Sigma)$ is important for QDs in order to resolve transport contributions of individual QD configurations.

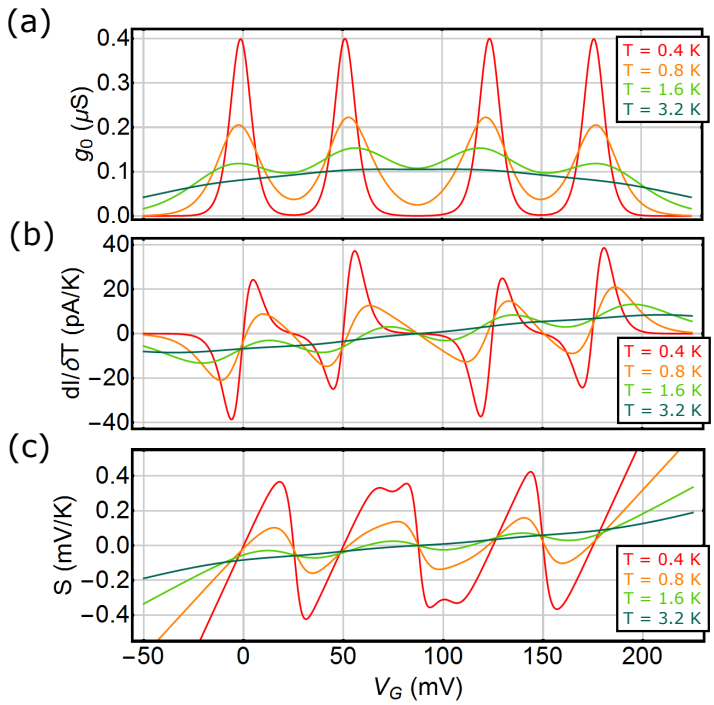


Figure 19: Calculated g_0 in (a), $dI/\delta T$ in (b) and S in (c) of a QD with two spin-degenerate orbital states as a function of the gate potential V_G. Calculation parameter values are $e^2/C_\Sigma = 1$ meV, $\varepsilon_1 = 0$ meV, $\varepsilon_2 = 0.5$ meV, $\Gamma_L = \Gamma_R = 500$ MHz, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the temperature of the contact leads $T = T_L = T_R$, which are 0.4 (red), 0.8 (orange), 1.6 (light green) and 3.2 K (dark green).

Comment on second order tunneling effects

Hopefully the discussion above has demonstrated that the Master equation approach to transport through QDs in the SETA provides a straight forward way to account for charging effects even in the presence of multiple orbital states. The spectrum of effects that can be explained by this approach is large, however, it neglects all effects related to coherent many-electron tunneling processes of which the most known are co-tunneling and charge fluctuations [56, 57, 58]. They allow for violation of energy conservation at short time scales in line with the Heisenberg uncertainty principle [30]. Such effects are responsible, for example, for an effective energy broadening of the orbital state energy and non-zero conductance within the Coulomb blockade regime.

Reference [57] demonstrates that including the second order tunneling processes in the Master equation approach yields non-trivial effects in charge and heat transport. They include conductance within the Coulomb diamonds, that agrees with experimental observations [56, 61], and negative differential heat conductance, that so far has not been observed experimentally. Unfortunately, the additional terms needed to account for the second order tunneling processes make the calculations significantly more complex - calculation time increases by roughly two orders of magnitude. Therefore giving exact calculation examples is beyond the scope of the current section.

Such calculations considering second order tunneling effects are applied in Paper II where the power and efficiency of QD heat engines is studied. The understanding developed in this study allows us to argue that the second order tunneling effects have a relatively small effect on the maximal power that a QD heat engine can produce. However, the effect on the efficiency is significantly more apparent. The effective energy broadening of the QD state essentially allows heat flow through the QD even without the presence of charge current. This means that the proportionality relation between the heat and charge flow, given for example by Eqs. 29 and 30, no longer holds. This has the consequence of reducing the maximum achievable thermal-to-electric conversion efficiency, because the second order tunneling processes increase the heat flow in the denominator of Eq. 32.

3.3 Landauer-Büttiker approach

The Landauer-Büttiker (LB) approach is a widely used technique for simulating transport in junctions that can be considered as non-interacting (at least effectively) [62]. It assumes that all electron scattering is elastic and that the electron transport is phase coherent. When applied to QDs, it considers a QD as a coherent and non-interacting scatterer in one dimension that can either transmit or reflect incident electrons. The biggest drawback of this method is that it cannot fully model the Coulomb blockade in QDs, as the Master equation approach can, because it considers interactions only within a mean field approxima-

tion. Consequently, its application to QDs, that undergo strong Coulomb blockade, has limitations. However, it can provide valuable insights that complement the SETA picture as it easily allows us to model transport through QDs in cases where the tunnel-coupling is not weak and the transport through the QD is no longer limited to very narrow energy windows.

The conventional expression for calculating current at finite T and V is given by the Landauer formula [62]

$$I = \frac{e}{h} \int_{-\infty}^{\infty} \tau(\epsilon)[f_R(\epsilon) - f_L(\epsilon)]d\epsilon \tag{42}$$

where $\tau(\epsilon)$ is an energy dependent transmission function that is assumed to fully characterize the scatterer. In the low T limit Eq. 42 reproduces the current through a perfect 1-dimensional conduction channel $I = (e^2/h)V$ if a constant transmission function $\tau = 1$ is assumed (note that the absence of a factor of 2 in Eq. 42 is intentional to consider only one spin species). More generally Green's function theory can be used to calculate the form of $\tau(\epsilon)$ in interacting systems like QDs [63]. However, a conventional approximation for $\tau(\epsilon)$ in Coulomb-blockaded QDs is a Lorentz function [64, 65]

$$\tau(\epsilon) = h\tilde{\Gamma} \cdot \frac{1}{\pi} \cdot \frac{\gamma}{\gamma^2 + (\epsilon - \xi)^2} \tag{43}$$

where $\gamma = \hbar(\Gamma_L + \Gamma_R)/2$ and ξ is the energy of the orbital state

$$\xi = \varepsilon_1 - e \sum_{j=L,R,G} \alpha_j V_j \tag{44}$$

The function in Eq. 43 defines a Lorentzian with a full width at half maximum (FWHM) equal to $\hbar(\Gamma_L + \Gamma_R)$. When integrated over all energies ϵ, it returns a pre-factor $h\tilde{\Gamma} = h\Gamma_L\Gamma_R/(\Gamma_L + \Gamma_R)$. In case of $\Gamma_L = \Gamma_R$, the value of $\tau(\epsilon = \xi)$ peaks at 1. Using such Lorentz transmission probability in Eq. 42 yields results consistent with the Master equation approach for a single non-degenerate orbital. For example, in the high bias limit $\hbar(\Gamma_L + \Gamma_R), kT \ll eV$ it predicts the same maximal current $I_{max} = e\tilde{\Gamma}$.

One of the conveniences of using the LB approach is that it provides a relatively straight forward way to calculate the linear response parameters g_0 and S. They are given by the following expressions [66, 67]

$$g_0 = \frac{e^2}{h} \int_{-\infty}^{\infty} \tau(\epsilon) \left[-\frac{\partial f}{\partial \epsilon} \right] d\epsilon \tag{45}$$

and

$$S = -\frac{k}{e}\frac{\int_{-\infty}^{\infty}\tau(\epsilon)\left[-\frac{\partial f}{\partial \epsilon}\right]\frac{\epsilon-\mu}{kT}d\epsilon}{\int_{-\infty}^{\infty}\tau(\epsilon)\left[-\frac{\partial f}{\partial \epsilon}\right]d\epsilon} \tag{46}$$

where f is the Fermi-Dirac distribution at the contact leads with $T = T_L = T_R$ and $\mu = \mu_L = \mu_R$. The differential current response dI to an infinitesimal thermal bias $d(\Delta T) = \delta T$ can be obtained by multiplying the two linear response coefficients

$$\frac{dI}{\delta T} = g_0 \cdot S \tag{47}$$

Figure 20 demonstrates calculation results for g_0, S and $dI/\delta T$ using same parameters as in Fig. 14 in the previous Section 3.2. Overall, results might seem similar, however, there are a few details that distinguish these calculation results from the ones obtained in the SETA. First, when looking at g_0 in Fig. 20a, the peak values g_{max} are lower due to the fact that the spin-degeneracy of the orbital state at ε_1 is not considered properly. The g_{max} is not predicted to scale precisely with $1/T$ and no peak shift with increasing T can be seen.

Even bigger differences can be seen in thermoelectric signatures, particularly in S. Figure 20b shows that the LB approach does not predict S to scale linearly with the gate potential V_G, as the SETA in the Master equation approach. For the lower T cases within the plotted V_G range one instead can observe saturation and eventual decay of S, similarly as seen in experiments [68, 69]. Furthermore, as a combined consequence of the behavior of g_0 and S, the response of $dI/\delta T$ is also modified accordingly - the two maxima of $dI/\delta T$ are lower, have the same amplitude and decay with T in a slightly different manner (see Fig. 20c).

The reason for the differences in results between the Master equation approach in SETA and the LB approach is partly, of course, a consequence of the failure of the LB approach to accurately account for charging effects and the degeneracy, however there is more to it. The additional differences also originate from the fact that the transport through the QD in the LB approach is not limited to a very well defined energy, but to a range of energies as specified by τ instead. In fact, as demonstrated with Fig. 21, by decreasing the width of τ (i.e. decreasing $\hbar(\Gamma_L + \Gamma_R)$), the range in which S scales linearly is extended and the agreement with the SETA is improved. However, the full linear behavior of S can only be reproduced when $\hbar(\Gamma_L + \Gamma_R)$ becomes many orders of magnitude smaller than kT (rather than just much smaller than kT, as conventionally assumed). This points to additional physics introduced by the broadening of the orbital state energy, which ends up strongly modifying the behavior of S in QDs.

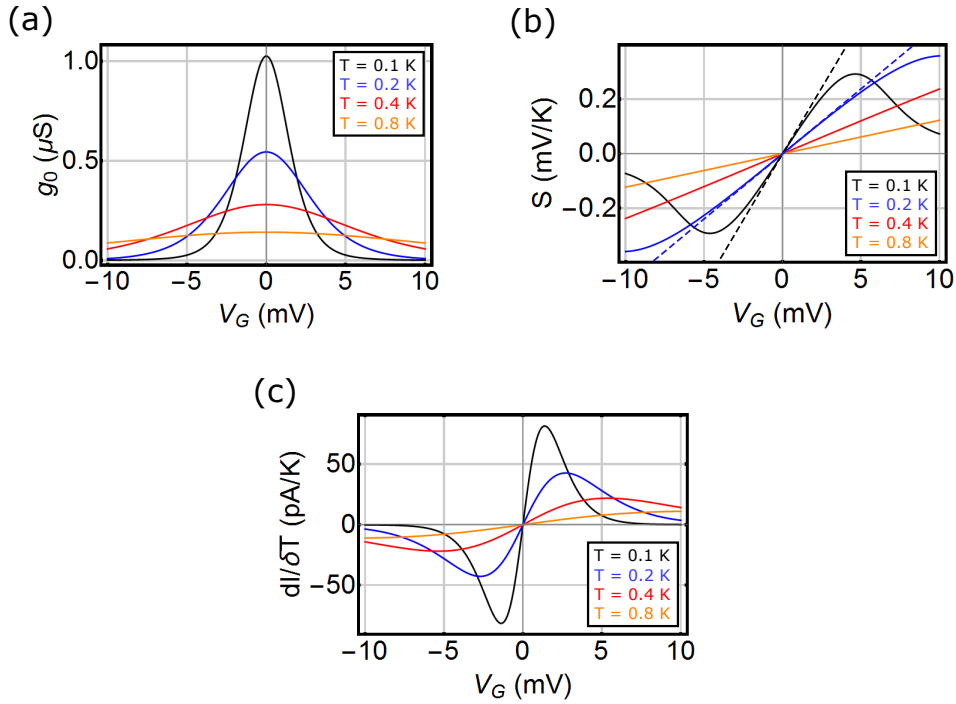


Figure 20: Calculated g_0 in (a), S in (b) and $dI/\delta T$ in (c) as a function of the gate potential V_G using the Landauer-Büttiker approach. The dashed lines in (b) represent the corresponding S obtained using the SETA in Fig. 14. The calculation parameter values are also the same as in Fig. 14: $\varepsilon = 0$ meV, $\Gamma_L = \Gamma_R = 500$ MHz, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the temperature of the contact leads $T = T_L = T_R$, which are 0.1 (black), 0.2 (blue), 0.4 (red) and 0.8 K (orange).

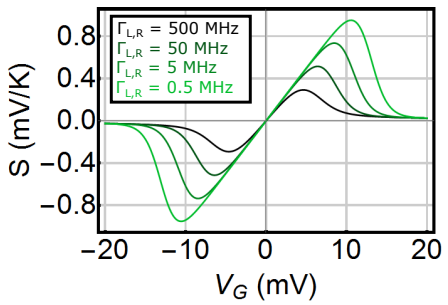


Figure 21: Calculated S as a function of the gate potential V_G using the Landauer-Büttiker approach. Calculation parameter values: $\varepsilon = 0$ meV, $T = T_L = T_R = 0.1K$, $\alpha_L = \alpha_R = 0.495$, $\alpha_G = 0.01$. Four curves differ by the $\Gamma_L = \Gamma_R = 500, 50, 5$ and 0.5 MHz. The corresponding values $kT/[\hbar(\Gamma_L + \Gamma_R)] \approx 13, 130, 1300$ and 13000.

The fact that experimental observations [68, 69] are in agreement with these predictions indicates that the SETA in the Master equation approach neglects important contributions to transport close to the open circuit conditions which are at least partly captured by the LB approach.

3.4 Summary

This section has introduced a circuit model of single QDs and two approaches for modeling their conductance and thermoelectric properties. The Master equation approach in the single electron tunneling approximation (SETA) is able to accurately account for charging and degeneracy effects and therefore can be used to qualitatively model QDs with many relevant orbital states. However, because it considers that only electrons with an exact energy of the orbital state can occupy the QD, it, neglects the effective broadening of orbital state energies due to electron finite lifetime on the QD. Further, the SETA can not account for the existence of a finite conductance within the Coulomb blockaded regions due coherent two electron scattering processes. This is where the Landauer-Büttiker (LB) approach serves a pedagogical purpose. It treats a QD as a coherent and an effectively non-interacting scatterer and can model the effects of broadening of the orbital state energy with simple calculations. It illustrates an important result - the Seebeck coefficient S of a QD continues to be sensitive to the broadening of the QD's transmission function even well within the traditionally assumed applicability limits of the SETA. Since such a behavior is consistent with experimental observations, it suggests that the SETA might not be optimal for modeling thermoelectric behavior of QDs, at least not when they are operated close to the open circuit configuration. However, because the LB approach does not accurately account for charging effects, it has difficulties modeling orbital state degeneracy and excited states.

A more accurate way to model QDs is to include second order tunneling processes into the Master equation calculations [57, 58]. This approach is used in Paper II where accurate evaluation of charge and heat currents through QDs close to open circuit conditions were needed. It is also applied in Paper III to estimate temperatures of the QD contact leads, as the tunnel-coupling was too strong to use the SETA. Unlike when only considering sequential tunneling, this method is experimentally found to model the thermoelectric behavior of QDs in the regime $(\Gamma_L + \Gamma_R) \ll kT/h$ rather well at all circuit configurations. Unfortunately, the calculation difficulty is increased and the calculation time is roughly a hundred times bigger than when using the SETA. This is why going beyond the SETA is sometimes unnecessary, as, for example, in Paper I where conditions closer to the short circuit conditions are used and the essential physics can be captured without the second order tunneling effects.

4 Prior work

The research on thermoelectric properties of single QDs has mostly been pursuing one of two directions. One of them has been using thermoelectric properties in the characterization of QDs complementing the conventional conductance characterization. This direction has dominated the experimental part of the field. The other focus of studies has been using QDs as efficient heat engines. This direction has so far been dominated by theoretical studies whereas experimental verifications have been lacking. This section aims at summarizing the key theoretical and experimental advances over the past three decades distinguishing between the two research focuses. It is partly based on a recent review [20] by the author and coworkers, which focuses more on experimental behavior of thermovoltage (Seebeck coefficient) and its comparison to theory. The current section expands the review in the direction of QD heat engines.

4.1 Development of quantum dot thermoelectric devices

Coulomb blockade phenomena have been known to increase the low temperature resistance of conductors of limited cross sectional area since the early 1950s [70]. By the end of 1980s the nanofabrication techniques had developed far enough to enable systematic studies of single electron devices [71, 72]. Famous examples are metallic single-electron transistors [29], in which charging effects are used to study regimes in which electrons transit devices one at a time. However, because the Fermi wavelength of electrons in metals is at a scale of interatomic distances, confinement effects did not play an important role in metallic devices.

In contrast to metals, the Fermi wavelength in semiconductors can be at a scale of several tens of nanometers which means that for small enough devices quantum confinement effects can no longer be neglected. The implementation of modulation-doped semiconductor heterostructures [73] enabled controlled fabrication of high mobility 2-dimensional electron gases (2DEGs), the development of which progressed rapidly throughout the 1980s. The 2DEGs could be created at the interfaces of two semiconductors, typically GaAs and AlGaAs [74], and could be shaped electrostatically by using top gates [75]. The 2DEGs allowed for convenient fabrication of resonant tunneling structures [76, 77] at the scale comparable to the Fermi wavelength. This gave access to the quantum confinement effect needed to realize QD thermoelectric devices [78, 79].

Even up till now, the majority of thermoelectric studies done on single QDs have been using 2DEG devices. However, a growing number of studies investigate thermoelectric effects in synthesized quasi 1-dimensional structures [69, 80]. Examples include carbon nanotubes [81] and semiconductor nanowires [7], and even in single molecules [82]. The

advantage of using such structures for fundamental studies is that the tunnel-junctions and the electronic state structure of small physical objects can often be defined more accurately than it is allowed by the electrostatic gates in 2DEG devices.

4.2 Thermal biasing and thermometry

In addition to the requirements of the conventional electrical characterization, mesoscopic thermoelectric experiments also require means of applying and characterizing temperature differences ΔT across sub-micrometer-sized objects. This means that both the heating and thermometry have to be local at these length scales [20]. This section aims at giving a summary of the most commonly used techniques.

Thermal biasing

A thermal bias of single QDs is achieved when the two QD contact leads are at different electronic temperatures. This is conventionally realized by dissipation of Joule heat that couples asymmetrically to the two contact leads. The Joule heat itself originates from passing an electric heating current I_H through a resistive element. Conceptually, there have been three different approaches that have so far been used in thermoelectric experiments on QDs. Figure 22 schematically illustrates the differences between the different approaches.

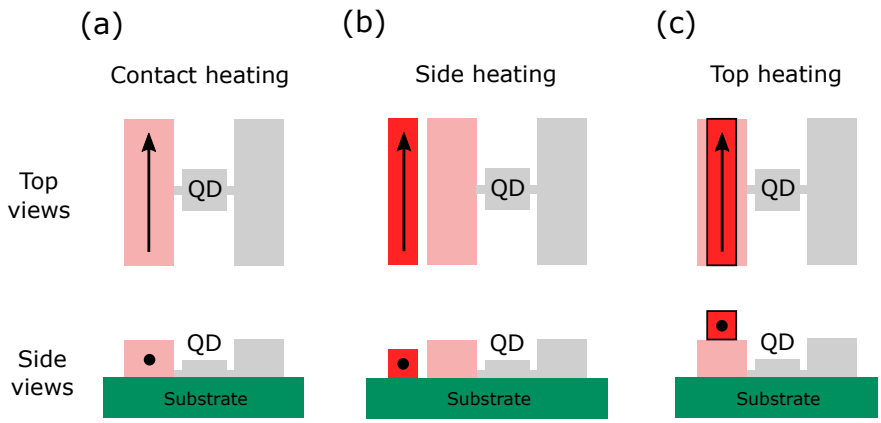


Figure 22: Schematic illustrations of different QD thermal biasing approaches. The gray rectangle in the center represents a QD between two contact leads. The contact lead that is indicated in pink is warmer than the other. The arrows represent heating currents I_H that dissipate Joule heat. (a) Contact heating: the Joule heat is dissipated by passing I_H directly through the contact lead. (b) Side heating: the Joule heat is dissipated in a dedicated heater lead that is located close to the contact lead, but does not make an electrical connection with it. The heat is thermally conducted through a substrate (in green). (c) Top heating: the heater lead is placed on top of the contact lead while being insulated from it by a thin layer of insulator. The heat is conducted to the contact lead via the thermal conduction through the insulator.

The most popular and historically the first thermal biasing approach that was implemented

in the experiments [78] heats one of the QD contact leads by passing a heating current directly through it (contact heating, see Fig. 22a). The advantages of this method are that it is local and convenient from a fabrication point of view. This approach is of particular importance at temperatures significantly lower than 100 mK or so, where electron-phonon coupling is weak. It allows realizing experiments in which heat transport is purely dominated by electrons [83]. The greatest disadvantage of this method is that harder to integrate it with measurements of thermocurrent I_{th} or thermovoltage V_{th} [69, 84]. Essentially, the heater current I_H introduces electrical potential variations along the length of the contact lead. These potential variations, if not carefully taken care of, can compromise the thermoelectric measurements by introducing unwanted electrical biases across the QD. Such an effect is very inconvenient from the point of view of conducting an experiment, because compensating for this takes a significant part of the experimental effort and introduces additional experimental uncertainties.

An alternative approach is to introduce a dedicated (side) heater lead that is placed in the proximity of the contact lead instead (see Fig. 22b) [80, 85]. Such an arrangement relies on thermal conduction through the substrate for the heat delivery to the contact lead. Because the heater lead is electrically disconnected from the electrical biasing circuit of the QD, no electrical potential balancing across the QD is needed and ΔT can therefore be freely tuned without much additional experimental effort. The disadvantage of such a method is that only a part of the heat is reaching the contact lead electrons and most of it only heats up the setup. As a result, more heating is needed to ensure a given ΔT and the overall temperature rise makes it difficult to achieve high $\Delta T/T$.

An approach that attempts to combine the advantages of the two previously described approaches while minimizing the disadvantages is the so-called top-heating approach [86] (see Fig. 22c). The heater lead is still a part of an electrical circuit separate from the electrical biasing circuit, but is placed as close as possible to the contact lead - on top of it. The contact and heater leads are insulated by a thin layer of insulating material. Such an approach still allows easy tuning of ΔT while minimizing the overall heating of the substrate. The biggest disadvantage of this approach is that it uses a two step (rather than one step) lithographic process which effectively doubles the fabrication time and reduces the device fabrication yield. However, it is because of its advantages that this approach is used for fabrication of QD thermoelectric devices in this thesis.

Thermometry

Thermometry in QD thermoelectric experiments is particularly challenging. On the one hand, the measurement of the QD contact lead temperatures has to be sufficiently local in order to represent the actual temperatures that a QD experiences on both sides rather than a convolution of temperatures at the device area. On the other hand, the accuracy in

temperature has to be significantly better than ΔT itself or otherwise it cannot be determined. Experimentally it is difficult to perform a local temperature estimation of the contact leads at the tunnel-junctions. However, the fact that the temperature of the contact leads is relatively even allows using their properties to determine ΔT.

Some experiments with 2DEGs thermoelectric devices have used quantum point contacts (QPCs) [87] for thermometry purposes [88, 89]. While using this method, a QPC is placed between a heated reservoir of a QD and another reservoir at a known reference temperature. The resulting thermovoltage that builds up due to the temperature difference across the QPC is then used to calibrate the temperature of the heated reservoir. Other experiments on 2DEGs thermoelectric devices have used Shubnikov–de Haas (SdH) resistance oscillations in order to estimate contact lead temperatures [90, 91].

For devices that are not 2DEG-based, a convenient choice has been estimation of reservoir temperatures based on the resistance of the contact lead segments. Such an approach has been used for thermoelectric studies of a QD implemented in a carbon nanotube [80, 85] and is widely used in thermoelectric studies of nanowires [59, 92]. However, the accuracy of this method becomes reduced at temperatures below 10 or 20 K where the resistance of metallic contact lead segments generally lose their temperature dependence due to impurity dominated scattering [93].

The method applied in this thesis uses the QD thermoelectric response itself to determine temperatures. The advantage of this method is that it probes the very same temperatures that the QD experiences. It has previously been applied for estimation of ΔT in nanowire QDs in Refs. [68, 69, 84, 94]. One has to point out though that the accuracy of such a method is directly linked to the applicability of the underlying theory to each particular QD regime. For this purpose, Papers II and III use a Master equation approach that accounts for the first and second order tunneling processes [57, 58]. The disadvantage here can be that accurate calculations can require relatively much computational resources and that only a subset of experimentally fabricated QDs can be modeled well by simple models.

4.3 Lineshape of the thermovoltage

State of the field

A part of this thesis work has been to identify and summarize the opened problems in the field of QD thermoelectrics. One of the early conclusions was that the lineshape of $S(V_G)$ in single QDs has so far not been confirmed to comply with the early theoretical prediction by Beenakker and Staring [18]. Instead, the Landauer-Büttiker approach, despite its disadvantages as summarized in Section 3.4, has been more successful in explaining the experimental results [20]. This is likely due to the failure of the SETA to accurately

describe thermally biased QDs close to open circuit conditions because transport contributions beyond sequential tunneling turn out to be important then. These conclusions served as a motivation for using a Master equation approach that considers transport mechanisms beyond sequential tunneling for the analyses in Papers II and III. The following provides a review of the most important work in the direction of understanding the lineshape of thermovoltage.

Review of studies

Pioneering theoretical work on quantum confinement effects in Coulomb-blockaded devices was done by Beenakker at the beginning of the 1990s [18, 28]. These works discussed the periodicity and amplitude of g_0 and S for QDs in the SETA and are the basis for the theoretical description of QDs, laid out in Section 3.2. Some of the main conclusions of this work therefore have already been illustrated in Figs. 18 and 19. The Seebeck coefficient S, as a function of gate potential V_G was predicted to invert sign two times within a period of a single Coulomb oscillation in g_0, one close to the g_0 peak and one in between the peaks. The overall *lineshape* of $S(V_G)$ was predicted to resemble a saw-tooth shape with a certain amplitude. QD excited states were predicted to show up in S as additional wiggles on top of the saw-tooth shape. This is in contrast to the linear conductance g_0 in which no signatures of the excited states were expected. The qualitative difference between g_0 and S illustrated the fact that S, as a sensing tool, provides complementary information to g_0 about the energetics of the QD state spectrum without entering the nonlinear response regime.

Soon thereafter two experiments were published investigating the lineshape of $S(V_G)$ in QDs [78, 97]. The work by Staring et al. [78] measured thermovoltage V_{th} as a function of V_G and observed the predicted periodicity, however they were unable to accurately determine the applied thermal bias ΔT, so the exact magnitude of S was not determined. This work also pioneered V_{th} measurements in the nonlinear $\Delta T \gg T$ regime, however, it is not clear how to interpret the data given that the measurements where carried out using lock-in techniques that rely on linear response assumptions. The work by Dzurak et al. [97] also reported results on thermovoltage V_{th} as a function of V_G that showed the predicted periodicity corresponding to the conduction peaks. However, since the QD was tunnel-coupled strongly, the theory developed for QDs in the SETA was not applied. The results were shown to be in agreement with the Mott relation for the Seebeck coefficient [98] instead. This left the magnitude of S in the SETA being unverified.

A few years later Dzurak et al. [79] published another, more extensive study of S on a QD, covering a wider V_G range and including more than 10 consecutive S oscillations. The lineshape of $V_{th}(V_G) = S\Delta T$ was clearly demonstrated to oscillate in agreement with the period of the conduction peaks. It also demonstrated for the first time that the effects of

QD excited states can indeed be visible in the lineshape of S. However, another important conclusion was that the observed amplitude S was roughly two orders of magnitude smaller than predicted by the SETA. The deviation was attributed to the none negligible effects of higher order tunneling effects. The experimental $V_{th}(V_G)$ data was found to be in a much better agreement with the Landauer-Büttiker approach for calculating S and used a transmission function that was fitted to conductance data. Similar findings were published by Small et al. [80, 85] who systematically investigated the amplitude of $S(V_G)$ oscillations as a function of T in a carbon nanotube QD. While the observed oscillations were irregular in shape and therefore no exact lineshape comparison could be made, their root-mean-square (RMS) amplitude of S was found to be systematically much lower than predicted by the SETA, nevertheless still scaled proportionally to $1/T$ up to $T = 30$ K.

The question of how higher order tunneling processes affect S of QD-like devices from a theoretical perspective were first investigated by Turek and Matveev [99], published in 2002. They considered a single electron transistor at low temperatures where co-tunneling could not be neglected. Their findings showed that co-tunneling indeed modifies the lineshape of S and significantly reduces its oscillation amplitude. Similar findings where published in Ref. [100] a couple of years later. Experimentally, the effect of co-tunneling on the lineshape of S was studied by Scheibner et al. [68] who observed the predicted lineshape change of S as a function of T. However, despite the fact that co-tunneling (as given by Ref. [99]) could explain most of the deviations between the experimentally observed amplitude of S and the amplitude predicted by the SETA [18], the quantitative agreement was lacking.

One of the best qualitative agreements with theoretical predictions for the lineshape of S was demonstrated in semiconductor nanowire QDs by Svensson et al. [69] in 2012. A similar lineshape evolution with T as in Ref. [68] was observed. The experimental data was compared to calculations of S using the Landauer-Büttiker approach and, after accounting for the impedances in the measurement circuit, the results were shown to be in good qualitative agreement and showed only about two times smaller S amplitude than predicted by the Landauer-Büttiker approach.

A review on QD experiments would not be complete without mentioning the strong coupling regime ($(\Gamma_L + \Gamma_R) > kT/h$). This regime in single QDs has mostly been investigated with interest in the Kondo effect [25]. Under strong coupling conditions the second order tunneling effects can become dominant [101] and QDs can undergo drastic changes in the behavior of g_0 [102]. In case of the Kondo effect the configurations of the QD that possess a finite spin value strongly interact with the electrons in the contact leads forming a many-body state. In simple terms this means that the single electron orbital picture of QDs introduced in Section 3.2 is not valid anymore. The Kondo effect has been widely studied in single QDs in several experimental systems [103, 104, 105, 106]. However, very few thermoelectric experiments have been done in this regime [89, 107, 108]. The existing experimental results neither fully confirm, nor disprove the theoretical predictions by

Costi and Zlatić [19] who predict rather sophisticated behavior of thermoelectric effect in Kondo correlated QDs. The most interesting pending prediction is a polarity inversion of the thermoelectric response when entering the Kondo regime. It has been suggested that the Kondo effect could also lead to strongly nonlinear effects in the thermoelectric response of QDs [109]. This has been mentioned as one of the possible explanations for the otherwise fully unexplained nonlinear behavior of thermovoltage and thermocurrent in some experiments [84].

4.4 Studies of quantum dots as heat engines

In 1996 Mahan and Sofo [21] predicted that in the linear response the most appropriate density of states for efficient thermoelectric conversion is a Dirac delta function. In such a hypothetical conductor electron transport can only occur at a certain energy while transport at all other energies is blocked. In the following decade the idea of such idealistic electron energy filtration was further investigated by Humphrey et al. [22, 23, 110]. First, they pointed out that such an ideal electron energy filtration implemented between two fermionic baths at different temperatures in general should allow approaching the reversible operation limit of thermoelectric conversion [22]. Second, it was suggested that this idea could be realized in practice by using QDs that possess very well defined electronic state structure [22]. And third, it was argued that engines of this type are in fact a different class of heat engines which do not operate by following a cyclical path in the state space of thermodynamic variables [23]. Instead they operate in a steady state and state variables cannot be defined in the same way as conventionally done in the context of heat engines. Therefore a concept of *particle-exchange heat engines* was introduced [23].

More recent theoretical work has also investigated the thermoelectric conversion efficiency of QDs at their maximum power production capacity. Both the Master equation approach [111] and the Landauer-Büttiker approach [65] have been used. In both cases it was found that the efficiency at maximum power in QDs should be very close to the Curzon-Ahlborn efficiency [112, 113], which is more widely known as an approximate efficiency limit of heat engines of various scales operated at maximum power [114]. Experimentally, however, it has been challenging to examine these theory predictions. The biggest problem has been the ability to reliably evaluate the heat flows through single QD heat engines. Earlier experimental work in this direction was done by E. Hoffmann who investigated ways of using QDs themselves as thermometers [94] and performed early experiments on thermoelectrically produced power [115]. However, because of the insufficient agreement between theory and experiments, detailed analyses of power and efficiency was not possible. This thesis makes additional steps to improve experiments and theoretical models to carry out the full analyses, which is presented in Paper II and discussed in Section 6.2.

5 Experimental methods

5.1 Nanowire quantum dots

This thesis has made use of heterostructured bottom-up grown InAs/InP nanowires. The possibility of incorporating the intermediate bandgap material InP ($E_g \approx 1.35$ eV) within nanowires of the low bandgap material InAs ($E_g \approx 0.35$ eV) with nearly atomically sharp interfaces yield excellent experimental systemd for QDs. The heterostructured InAs/InP nanowires used in this thesis are grown using a growth technique called chemical beam epitaxy (CBE) [116]. CBE combines the sources from the more commonly used metal-organic vapor phase epitaxy (MOVPE) with the molecular beam epitaxy (MBE) technique (injecting metal-organic molecules as beams onto the growth substrate).

The growth of semiconductor nanowires is most often catalyzed (or seeded) by gold nanoparticles [117]. At elevated temperatures the seed nanoparticles melt and the growth species from the gas phase preferably condense on it forming an alloy. The growth occurs when the alloy becomes supersaturated with growth species and the crystal material self-organizes into a solid (precipitates) at the edge of the seed particle.

The growth process of InAs/InP nanowires is briefly illustrated in Fig. 23 and a more detailed description of nanowire growth with CBE can be found in Ref. [118]. In short, the growth requires three sources providing In, As and P. Trimethyl indium (TMIn) is used as a source of In atoms. When injected onto the growth substrate TMIn decomposes releasing In atoms that can diffuse on the surface and form alloys with the gold particles. At the same time, Tertiarybutyl arsine (TBAs) is used as a source of As during growth of InAs whereas tertiarybutyl phosphine (TBP) is used as a source of P during growth of InP. However, unlike for In atoms, due to the very limited surface diffusion length of As and P atoms [119] only incident TBAs and TBP molecules onto the seed particle contribute to the growth of the nanowire. InAs/InP heterostructures are created by switching between TBAs and TBP during the growth process. More details on similar growths of InAs/InP wires on the same growth machine are given in Ref. [120].

The studies carried out in this thesis used nanowires from a specific growth (see Appendix A for the technical details of the growth). Figure 24 shows a scanning electron microscope (SEM) image of the growth substrate after the growth and a high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) image of a sample wire from the same growth. The as-grown nanowires, mostly made of InAs, contain two thin and closely spaced InP segments forming tunnel-junctions to the InAs QD. As a consequence of the growth routine, another, longer InP segment is incorporated within the InAs, which has not been used in the experiments. The purpose of this longer InP segment prior the growth of the QD is to prepare the seed particle for the growth of the thin InP segments.

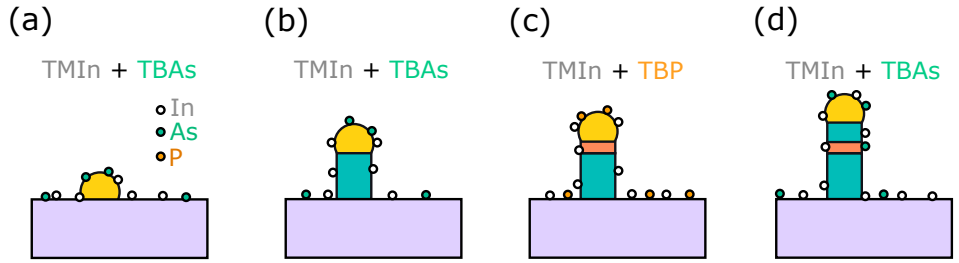


Figure 23: Illustration of InAs/InP nanowire growth process. (a) Nucleation: a gold seed particle is deposited on the growth substrate. TMIn and TBAs are supplied in the growth chamber providing In and As atoms that alloy with the gold seed particle. The growth of the InAs nanowire starts when the gold particle is supersaturated with In and As atoms. Then the InAs crystal precipitates at the interface between the seed particle and the growth substrate. (b) InAs growth: Continued supply of TMIn and TBAs in the growth chamber provides In and As atoms for continued growth of InAs. In atoms can defuse on the surface of the substrate and the nanowire side facets reaching the gold particles whereas only the As atoms released from the TBAs molecules incident to the gold particle contribute to the growth. (c) Switching to InP growth: The TBAs supply is cut and TBP is supplied in the chamber instead. In atoms can diffuse towards the gold particle whereas only the P atoms released from the TBP molecules incident to the gold particle contribute to the growth. (d) Switching back to InAs growth: The TBP supply is cut and TBAs is supplied in the chamber again. The InAs growth follows the description in (a) and (b).

The resulting heterostructure dimensions were analyzed using HAADF-STEM images.

Because of the high bandgap offset between InAs and InP, the conduction band offset is of the order of 0.5 eV. This is a significant feature of InAs/InP nanowire QDs with respect to other QD systems because it results in tunable tunnel-rates across InP tunnel-junctions and enables experiments in a wide energy range. In addition, because the tunnel-junctions are defined structurally, not by electrostatic gates, they are more electrostatically stable.

5.2 Devices

The nanowire devices are fabricated onto n-doped Si substrates covered by thermally grown 100 nm thick SiO_2. The oxide isolates the substrate from the fabricated structures on top allowing it to be used as a global back gate. Figure 25a shows 10 device fields of 12 gold contact pads that are pre-pattered onto the substrate oxide. The function of the contact pads is to enable easier wire bonding from an external point. Figure 25b shows that contact pads in each field have extensions leading towards the write field at the center where devices are fabricated. Figure 25c shows closer example of a write field. It features a coordinate grid that is used to map nanowires' positions once they are deposited onto the substrate. The grid is made of small dots separated 2.5 µm from each other and covers 100 by 100 µm area that partly overlaps with the tips of the contact pad extensions. The coordinate grid also contains a set of significant markers unique to particular coordinate points that allow easier mapping of the nanowire positions.

SEM images of deposited nanowires are used to map the nanowire positions within the write field using the coordinate grid (see Fig. 26a). A software is used to custom design

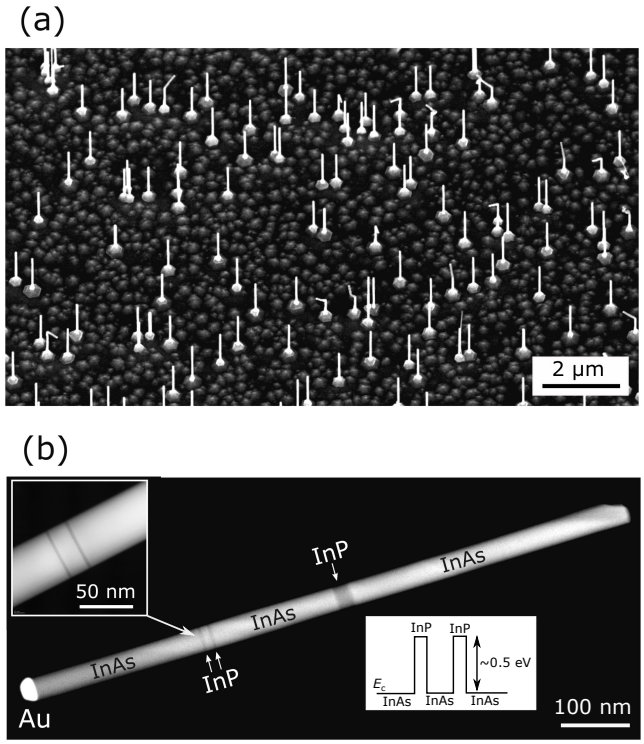


Figure 24: (a) Scanning electron microscope (SEM) image of the growth substrate after the growth. The nanowires have grown from the aerosol deposited gold seed particles. (b) high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) image of a nanowire from the growth in (a). The gold seed particle is visible (bright) on the left of the nanowire. The nanowire is continued by an InAs segment leading to a QD which is defined by the two short InP segments. The QD between the thin InP segments is made of InAs. On the other side of the QD there is another InAs segment, then a longer InP plug and the remaining InAs segment. From analyses of 11 nanowires the typical wire diameter was determined to be around 60 nm, the two thin InP segments were 2 - 8 nm and the QDs where 16 - 19 nm. The length of InAs segments leading to the QD were 290 - 440 nm of the seed particle side and 190 - 320 nm on the other. Left inset: a close up HAADF-STEM image of the QD section of another nanowire from the same growth. Right inset: an illustration of the conduction band alignment across the QD segment length. The band offset is of the order of 0.5 eV. The growth of the nanowires and SEM imaging was done by Sofia Fahlvik (prev. Svensson) and the STEM analyses was done by Sebastian Lehmann.

the layout of the device contact leads (see Fig. 26b). The contact leads are designed to make contact to the contact pad extensions therefore allowing devices to be contacted via the bigger contact pads. Description of fabrication steps can found in Appendix B. An example of a finished device within the write field can be seen in Fig. 27.

In order to interface the devices with measurement setups the processed substrates are glued onto ceramic carriers with a conductive paste (see Fig. 28). The design of the carriers allows us to make bonds between the contact pads on the substrate and the side pads on the carrier that are connected to the carrier pins. In this way the bonded device has direct electrical connections to the pins of the carrier allowing it to be placed in an electrical measurement setup for electrical characterization.

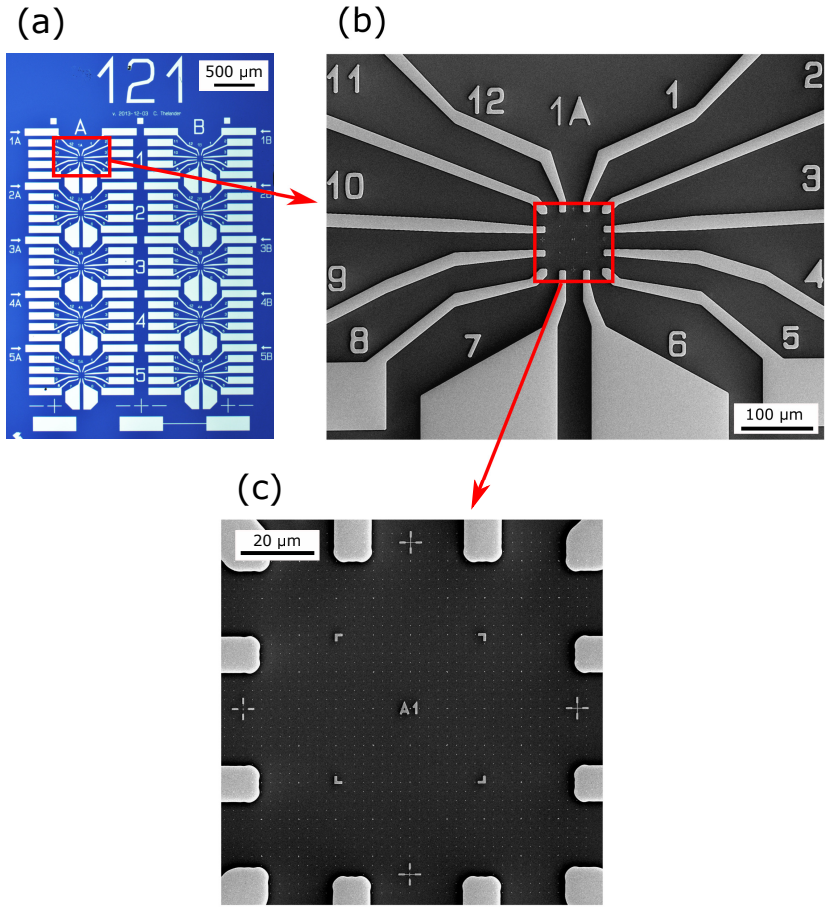


Figure 25: (a) Optical microscope image of a blank chip with 10 device fields arranged in a 2 by 5 array. The contact pad layout features 12 contact pads for each of the fields (design by Claes Thelander). (b) Scanning electron microscope (SEM) image of the contact pad extensions towards a write field. (c) SEM image of a write field featuring a coordinate grid.

5.3 Measurement setups

Resolving the physics of QDs in electrical measurements requires sufficiently low temperatures such that the typical thermal excitation energy kT is smaller than the charging energy e^2/C_Σ. In order to resolve transport through a single QD orbital state, kT also has to be below the typical orbital state energy offset resulting from the quantum confinement effect. The typical energy scale for these quantities in the nanowire QDs is of the order of 1 meV, which means that the required temperature range falls below 10 K ($kT \approx 0.862$ meV).

A dilution refrigerator (Triton 200) with electron temperature below 100 mK was used for cooling devices in all studies (Papers I, II and III). During thermoelectric measurements

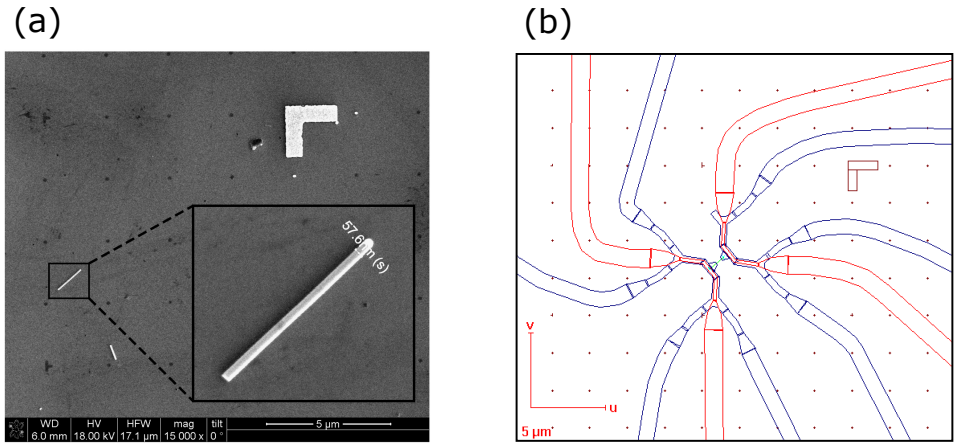


Figure 26: Designing of the contact leads to the nanowires. (a) scanning electron microscope (SEM) image of a nanowire deposited on a substrate within the write field. A significant corner mark, is used to determine the exact position of the nanowire within the write field. Inset: a close-up SEM image of the same nanowire. The InP plug is visible as a slightly darker section of the nanowire. (b) Screenshot of a finished two layer design in the Raith150 software. The layer in blue is a nanowire contact lead design and the layer in red is the heater lead design. The additional array of dots along with the significant corner mark is a guide to the eye.

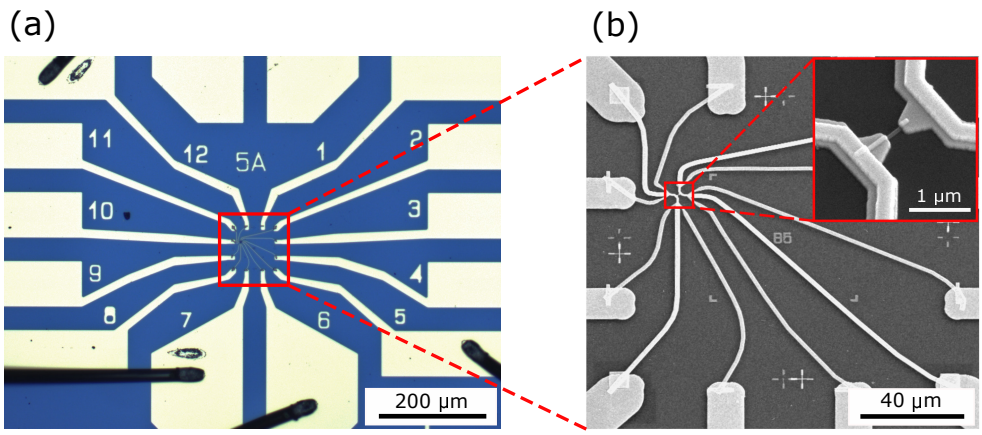


Figure 27: Sample images of a finished device. (a) Optical microscope image of the contact pads and their extensions towards a device area with a finished device. Black wires coming in from the sides of the image are wire bonds to the contact pads. (b) Scanning electron microscope (SEM) image of the device area of a finished device. The device leads stretch out to the contact pad extensions and make contact to them. Inset: close-up SEM image of the contacted nanowire. The heater leads overlay the contact leads in a top-heater architecture.

the actual device temperatures were higher due to the parasitic heating effects caused by the heater current. The typical temperature range was from a few 100 mK to several K. The dilution refrigerator was also equipped with a vector magnet able to apply magnetic field up to 1 T in all directions and up to 9 T along a specific axis of the system. This was used in the study of Paper III where magnetic field dependence of thermoelectric signatures was investigated. For most of the measurements the electrical lines in the refrigeration system

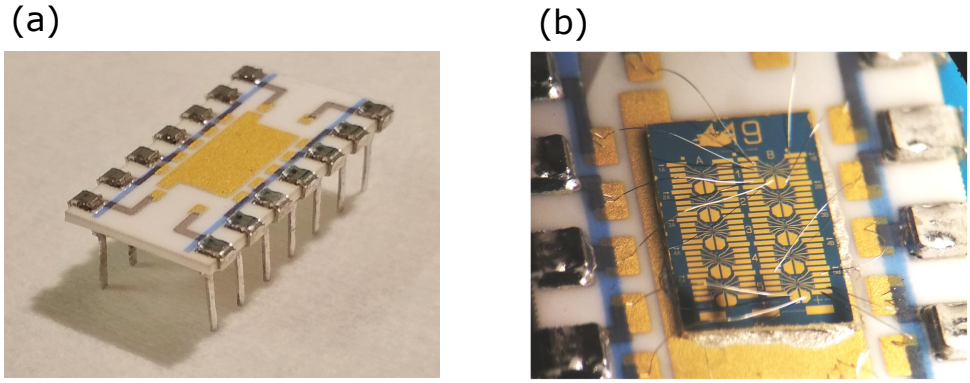


Figure 28: (a) Image of a blank 14-pin ceramic chip carrier. It features a gold coated pad in the center for hosting the substrate and 14 side pads that are electrically linked to one of the carrier pins each. (b) Optical microscope image of a wire bonded substrate onto the chip carrier. Fine wires link the contact pads on the substrate to the side pads on the carrier. In this example two devices in total are bonded for electrical characterization.

leading down to the cooled device were filtered at the 3.5 K cooling stage by RC and 7 stage pi filters with a combined cut-off frequency of 0.3 kHz. The study in Paper III also used an additional filtering stage at the sample holder with a higher cut-off frequency.

Most relevant electrical measurements were done in DC mode. Yokogawa 7651 DC voltage sources with RC low-pass output filters ($\approx$ 100 Hz cutoff) were used for gating and electrical biasing of the QD as well as for applying the heating current. The current through QDs was detected using either of the low noise current pre-amplifiers SR570 or Femto DLPCA-200 which were read out by HP 34401a multimeters. The electrical measurements were controlled and programmed via an in-house developed Labview interface.

6 Summary of research results

6.1 Paper I: Nonlinear thermoelectric transport through quantum dots

Background

Already the very first experimental study on the thermoelectric behavior of QDs (by Staring et al. [78]) probed the nonlinear ΔT regime. The experiments used a method for detection of thermoelectric response that has since been commonly applied to thermoelectric studies of QDs [68, 69, 84, 89, 121]. The method is based on applying AC heating current I_H for thermal biasing. In that case the heating power scales with I_H^2 and the corresponding variation of ΔT takes place at double the frequency of the I_H. Knowing this frequency therefore allows one to detect the thermoelectric response using a lock-in amplifier.

In this way Staring et al. [78] measured the peak-to-peak amplitude of $V_{th}(V_G)$ oscillations as the amplitude of I_H was increased. They found that V_{th} saturated, decayed and eventually inverted the sign. No explicit explanation for this behavior was provided. Similar results for V_{th} and I_{th} as a function of I_H were found in a more recent study by Svensson et al. [84] where strong nonlinearities along with sign reversals were also detected. No conventional QD model could explain these results without introducing additional fitting parameters, like temperature dependent transmission function of the QD. It was hypothesized that the origin of such a behavior could be related to the Kondo effect, similarly as predicted by Ref. [19]. However, no conventional signatures of the Kondo effect, like zero-bias anomaly [122], were shown in conductance data.

Current work

The purpose of the study in Paper I was to further investigate the nonlinear thermoelectric response in QDs. The experimental devices featured a number of improvements with respect to the previous experiments [84]. First, the heterostructured InAs/InP nanowire QDs provided strong enough charging and quantum confinement effects such that $e^2/C_\Sigma \approx \Delta\varepsilon \approx 4$ meV (or ≈ 46 K in kT). This meant that the transport contributions of QD excited states could be neglected up to relatively high temperatures. This enabled to study the behavior of a single spin-degenerate orbital in a relatively wide ΔT range. Second, the experimental device featured top-heater geometry [86] (see Fig.29). This allowed applying relatively big ΔT without affecting the electrical bias across the QD. Judging from the response of the device, the achieved ΔT was of the order of 10 K across the 20 nm long QD. However, this came at the cost of also increasing the cold contact lead temperature.

Prior experiments on nonlinear thermoelectric behavior of QDs [78, 84] have used the AC

heating method to improve the signal-to-noise ratio of the small thermoelectric signals. However, this is not optimal for studies of nonlinear effects because the signal at two times the frequency of the I_H then only contains a part of the thermoelectric response. This makes quantitative analyses very difficult because signals from all relevant harmonics should be accounted for to deconstruct the true nonlinear V_{th} or I_{th} response as a function of ΔT. An alternative is to measure the thermoelectric response to DC heating current that requires significantly better signal-to-noise ratios. Partly because of the optimized application of ΔT, the study in Paper I managed to obtain a strong enough thermoelectric response to reliably use DC detection I_{th}. In such a manner the nonlinear I_{th} dependence on ΔT could be studied without any additional data processing. The detection of I_{th} instead of V_{th} was chosen because it is affected considerably less by small voltage drifts and possible leakage currents.

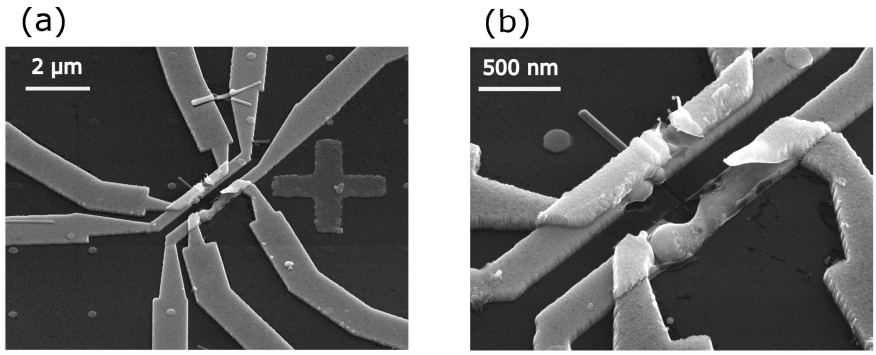


Figure 29: Scanning electron microscope (SEM) images of the device investigate in Paper I. Two metallic leads make contact to the nanowire on both sides of the QD. Additional heater leads run on top of the contact leads for thermal biasing. The device got damaged during the unloading procedure from the low-temperature measurement setup. (a) and (b) are different by the SEM magnification.

The experimental results were compared to a Master equation based model in the SETA. It is worth mentioning that the model also accounted for a resistive load in series with the QD which modeled the 1 MΩ input impedance of the current preamplifier. In practice this meant that the potential difference across the QD has to be found self-consistently, because any finite thermocurrent I_{th}, being generated solely from ΔT, then also induces a reverse voltage bias across the QD due to the resistance in series. Even though considering the load was not critical to the explanation of the nonlinear effects, the practice of including it in calculations turned out to be very important for studies in Paper II.

Results

The measurements showed a very strongly nonlinear I_{th} dependence on I_H. Typically, when beginning to increase I_H, a super-linear increase of I_{th} was seen (see Fig. 30a). Continued increase of I_H, however, yielded a saturation and a decay of I_{th}. This nonlinear behavior was

shown to vary somewhat depending on the gate potential V_G. The physical explanation for the nonlinear behavior could be found by modeling the transport through a single spin-degenerate orbital using the SETA. The experimental results were qualitatively explained by an interplay between different QD configurations (with zero, one or two electrons) contributing to I_{th} differently at different temperatures. This is illustrated in Fig. 30b which shows that at lower ΔT only one electron configurations mediate the charge transport, contributing to I_{th} in one direction. However, when continuing to increase ΔT (and also T) the two electron configuration enables a counter flow of charge that reduces I_{th} (similarly as discussed in Ref. [123]). The modeling could also explain the behavior of I_{th} noise amplitude that also showed dependency on I_H and V_G. Although it was not included in Paper I, the behavior of noise amplitude was consistent with how sensitive I_{th} was to variations in V_G at different temperature and gating conditions.

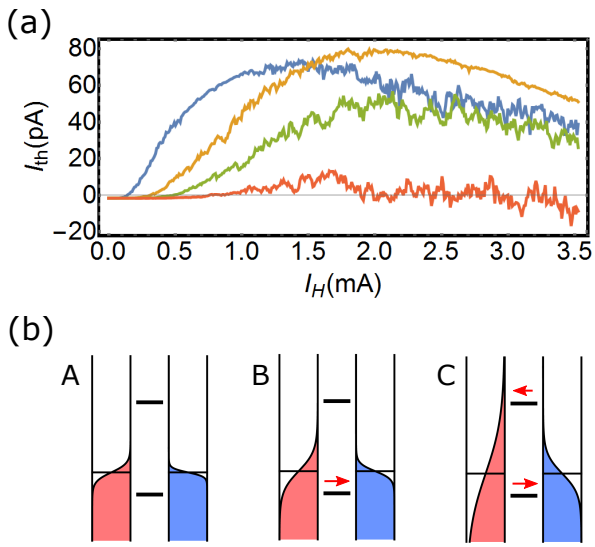


Figure 30: (a) Measured thermocurrent I_{th} as a function of heating current I_H for several gate voltages V_G. (b) Illustrations of a QD at different thermal biases ΔT. A shows a thermal bias that is insufficient to yield I_{th}. B shows a thermal bias that enables I_{th} via a single electron configurations of the QD. C shows a thermal bias that is big enough to allow single and double electron configurations, which for the sketched case yields a reduction of I_{th} due to counted flow of electrons. Note that the red arrows indicate I_{th} direction, opposite to the electron flow direction.

Realization of such an experiment showed for the first time that a nonlinear thermoelectric response of QD can in certain cases be fully explained. It also established the top-heater geometry as a useful tool for small scale thermoelectric experiments. When compared to theory, the QD's thermoelectric response to I_H suggested that ΔT significantly exceeding 10 K could be realized across the roughly 20 nm long QD.

6.2 Paper II: Quantum dots as efficient heat engines

Background

The idea of using the sharply energy dependent density of states of low-dimensional conductors for enhancing thermoelectric conversion efficiency has been around for more than two decades [124, 125]. In line with this thinking Mahan and Sofo [21] identified delta-function as the ideal density of states for efficient thermal-to-electric energy conversion. This idea was expanded upon by Humphrey et al. [22] who suggested that QDs could be used for implementation of solid state heat engines that in principle could approach the reversible operation (the Carnot efficiency limit). A point was consequently risen that such QD heat engines would be principally different from the conventional cyclical heat engines because they would operate in a steady state, therefore a concept of *particle-exchange heat engine* was introduced [23].

Despite the experimental progress on QD thermoelectric devices (discussed in Section 4), in particular experimental efforts by Hoffmann et al. [94, 115], such a QD heat engine had not been successfully realized. Among experimental challenges were difficulties to reliably estimate heat flows through a single QD as well as difficulties applying and characterizing thermal bias. An approach that was initiated by Hoffmann used the QD thermoelectric response for characterizing ΔT and evaluating the electronic heat flow based on theoretical modeling. However, given that the quantitative agreement between theory and experiments had so far been lacking [20], no definite conclusions were reached on whether this approach is sufficiently accurate and which theory is most appropriate to use.

Current work

The study in Paper II investigated the performance of thermally biased QDs as heat engines by making use of the methodology and device design used in Paper I. Again, the quality of the InAs/InP nanowire QDs played a critical. They provided with electrostatically stable QDs that were tunnel-coupled to contact leads weakly enough ($\hbar\Gamma$ of a few tens of μeV) while at the same time possessing strong enough charging and confinement energies (of a few meV). Such energy scales enable operation of a single-orbital QD in the quantum regime ($\hbar\Gamma \ll kT \ll \Delta\varepsilon, e^2/C_\Sigma$) at temperatures below 5 K - a requirement for efficient operation of QD heat engines.

Also similarly as in Paper I, the experimental devices used the top-heating architecture [86] (Fig. 31) that allowed studies of QDs outside of linear thermal bias regimes, while providing easy means of tuning the thermal bias. The device architecture yielded thermal biases up to such $\Delta T/T$, where the value of the Carnot efficiency parameter was relatively sizable $\eta_C \approx 0.4$. In this regime the thermoelectric performance is not necessarily predictable

by the linear response parameters, S and g_0. Therefore an important part of the studies in Paper II was to directly measure the thermoelectrically produced power P_{th} of the QDs heat engines. The power was determined solely from measurements of I_{th} using a variable load R in series with the QD that simulated a consumer of the thermoelectrically produced power $P_{th} = I_{th}^2 R$.

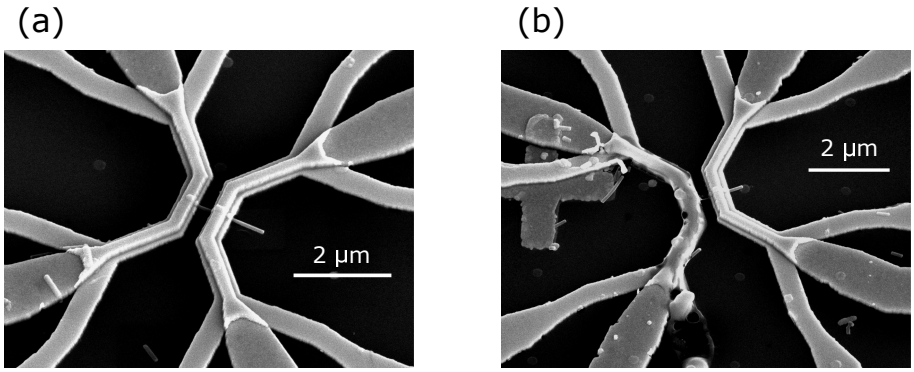


Figure 31: (a) and (b) are scanning electron microscope (SEM) images of QD devices used for studies of QD heat engines. Devices feature top-heater architecture where heater leads run on top of the leads contacting the nanowires. In each device only one of four contact leads was used on each side of the QD. The device in (b) got damaged during the unloading procedure from the measurement setup.

Another important advancement during these studies was the application of a more accurate theory approach that until then had not been used for modeling thermoelectric experiments on QDs. This work was done in collaboration with in-house theoreticians M. Josefsson and M. Leijnse who ran calculations. The theoretical description used to model the experimental results was based on a Master equation approach that goes beyond the SETA [57, 58]. It considers all terms in the expansion of the tunnel-Hamiltonian up to quadratic dependency on tunnel-rates. This approach fully accounted for nonlinear effects and electron-electron interactions on the QD as well as modeled co-tunneling and charge fluctuations. Consideration of these effects were most important when approaching open circuit conditions (high R conditions) because this is where the non-sequential tunneling effects become increasingly more important.

An important purpose of the theory was to perform thermometry on QDs. The theory was fit to the measured $I_{th}(V_G)$ allowing us to extract temperatures of the contact leads as the only two fit parameters. The fits showed consistency over a wide range of parameters V_G, R and ΔT, therefore indicating that the theory captured all of the important physics of the QD heat engines. Obtaining the temperature data allowed for a calculation of the associated electronic heat flows through the QDs, corresponding to the measured I_{th} at the given load conditions. The information about the electronic heat flows was then used to estimate the thermal-to-electric conversion efficiency.

Results

Experimental results in Paper II demonstrate that the QDs' ability to produce thermoelectric power P_{th} depends strongly on the magnitude of the load resistance R. When plotted as a function of R, the power P_{th} shows a distinct peak, which suggests that load matching is important also in QD heat engines (see Fig. 32a). The precise values of the power optimizing loads are hard to determine as they depend on tunnel rates and temperatures of the reservoirs in a complex way. However, what can be seen by comparing different devices at various temperature is that the power optimizing load seem to increase with temperatures and decrease with tunnel rates.

Access to the QD reservoir temperatures in this study allowed estimation of the electronic heat flow leaving the hot reservoir J_Q through the QD at various operation conditions. J_Q can be interpreted as the amount of heat power consumed during the operation of the heat engine, therefore the thermal-to-electric conversion efficiency can be calculated as $\eta = P_{th}/J_Q$. An interesting self consistent conclusion of the efficiency analysis was that the efficiency close to the maximum obtainable P_{th} was close to the so-called Curzon-Ahlborn efficiency $\eta_{CA} = 1 - \sqrt{T_C/T_H}$ [126], which is a widely known efficiency figure for heat engine operation at maximal power [127].

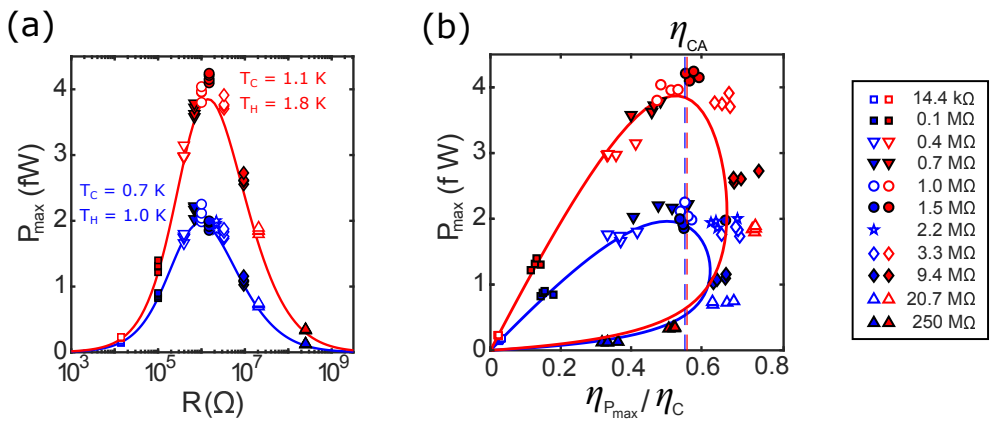


Figure 32: Results on one of the PE heat engines studied in Paper II. (a) Data points are the measured power P_{th} plotted as a function of load resistance R for two different thermal bias conditions, which are color coded in the figure. The solid lines are theory predictions for the estimated parameters: tunnel rates $\Gamma_L = \Gamma_R = \Gamma = 8.9$ GHz and temperatures of the hot and cold reservoirs T_H and T_C. (b) Data points represent the measured power P_{th} and the estimated efficiency $\eta = P_{th}/J_Q$ plotted parametrically as a function of load resistance R for two different thermal bias conditions (color coded the same way as in (a)). The solid lines are theory predictions for the estimated parameters, same as in (a).

An interesting behavior of maximum efficiency in the presence of higher order tunneling effects was observed (see Fig. 32b). Because these processes allow for heat transport through QDs also in absence of current (when $P_{th} = 0$), they reduce the efficiency at the vanishing power production, where the SETA predicts the efficiency to approach η_C. It results in the maximal efficiency η_{max} being reached at a finite P_{th} instead. In the main experimental

device efficiency $\eta_{max} \approx 0.7\eta_C$ was achieved at roughly one half of the maximal P_{th}.

The experiments in Paper II thus demonstrate a full power and efficiency analysis of QD heat engines for the first time. Although the electronic heat flow was only estimated by using theory, instead of being measured, the bare fact that the theoretical framework describes the experimental results so well can be taken as a strong indication that second order tunneling effects are very important in the thermoelectric transport through QDs, even in the quantum regime ($\hbar\Gamma \ll kT \ll \Delta\varepsilon, e^2/C_\Sigma$). The results of Paper II suggest that cotunneling and fluctuations are playing a limiting role in the thermal-to-electrical efficiency in QDs at high efficiency modes of operation, preventing the heat engines from achieving η_C. This is an intriguing observation given that these transport processes are only allowed due to Heisenberg uncertainty principle (due to wave characteristics of electrons) which is a fundamental principle of quantum mechanics. It makes one wonder up to which extent the Heisenberg uncertainty principle fundamentally limits the ability to reach η_C in particle-exchange heat engines.

6.3 Paper III: Thermoelectric transport through Kondo-correlated quantum dots

Background

The spin-1/2 Kondo effect [25] in QDs can be observed at low temperatures where $kT \ll \hbar\Gamma < \Delta\varepsilon, e^2/C_\Sigma$ [103, 122, 128, 129]. These requirements make it challenging to reach the Kondo regime experimentally for a number of reasons. First, increasing the parameters $\Delta\varepsilon$ and e^2/C_Σ require reducing the size of a QD, therefore these parameters normally have finite experimentally achievable values. Second, the finite values of $\Delta\varepsilon$ and e^2/C_Σ put rather strict conditions on T, that ideally has to be as low as possible, so that the much larger $\hbar\Gamma$ can still be smaller than $\Delta\varepsilon$ and e^2/C_Σ.

Additional difficulties arise while running thermoelectric measurements in the Kondo regime. First, during such experiments T is often risen many times above the base temperature of the cryostat because of thermal biasing. In practice this means that the Kondo effect has to be strong enough to be seen at temperature significantly above the typical cryostat base temperatures. Second, unlike for thermoelectric experiments on QDs in the weak coupling regime (presented by Papers I and II), those done in the Kondo regime cannot make use of the same theory based thermometry approach because the Master equation approach is not valid in the strong coupling regime. Given these considerations, it is not surprising that there have been only a very limited number of experimental studies on Kondo-correlated QDs [89, 107, 108].

On the theory side there have been a number of works [19, 109, 130, 131, 132, 133, 134, 135, 136] that investigate the Kondo regime, none of which so far have had a solid experimental verification. One of the most noteworthy studies by Costi & Zlatić [19] investigates the linear response behavior of Kondo correlated QDs as a function of T. According to their calculations the onset of Kondo correlations at low T should be accompanied by a sign inversion of V_{th} (and therefore also of I_{th}). Some signatures of this behavior have been previously seen by Thierschmann et al. [108] in 2DEG systems, however, the results have been inconclusive.

Current work

The experimental work in Paper III aims to fill this gap between theory and experiments. InAs/InP nanowire QDs were used to implement $\Delta\varepsilon$ and e^2/C_Σ of the order of a few meV, while ensuring a strong tunnel coupling ($\hbar\Gamma \sim 1$ meV). Such metrics allowed the spin-1/2 Kondo correlations to exist up to 1 K in temperature. This high characteristic Kondo temperatures, T_K, were important for the experiments as they provided enough flexibility in T range to apply a sufficiently big ΔT while remaining in the Kondo regime. This in turn increased I_{th} and enabled its reliable detection using DC measurements methods.

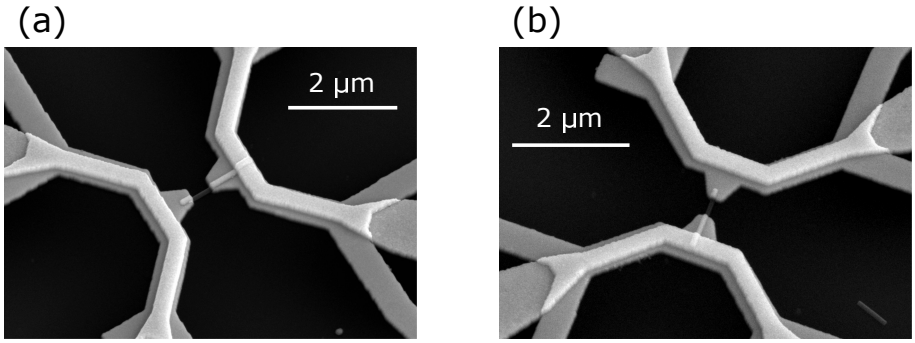


Figure 33: Scanning electron microscope (SEM) images of two devices used to study Kondo-correlated QDs. Devices feature the top-heater architecture where heater leads run on top of the leads contacting the nanowires. In each device only one of four contact leads was used on each side of the QD.

Thanks to the architecture of the top-heaters [86] the temperature control of devices during thermoelectric measurements could be realized using the heaters alone. This is because application of heating current I_H through a heater lead not only elevates the temperature of the contact lead beneath, but also elevates it to a lesser degree in the other contact lead. In the studied devices the application of the ΔT in this manner ensured elevation of T with an estimated ratio $\Delta T/T \approx 1/3$. Controlling the device temperature in this way enabled thermoelectric experiments at device temperatures up to several K while minimally affecting the cryostat temperature as measured by the built in temperature sensor.

The study in Paper III also used theory fits to experimental data to estimate the temperatures of the contact leads, in the same way as done in Paper II. However, because the corresponding theory is not applicable in the Kondo regime, a QD had to be tuned to a regime where the tunnel-coupling is weaker. This was achieved by applying a more negative gating potential V_G and depleting the QD. Unfortunately, only one out of three QDs showed transport behavior that was suited for thermometry (weak enough tunnel coupling and large enough quantum confinement effect). However, since all devices were identical by design, the thermometry results were assumed to provide good temperature estimates for all devices.

The work in Paper III also investigated the behavior of the thermoelectric signals in presence of a magnetic field. The relatively big g-factors of InAs nanowire QDs allowed to realize large enough Zeeman splitting to destroy Kondo correlations. Thanks to this, the transition from a Kondo correlated state to a normal state could be also studied using the thermoelectric signals.

Results

The results in Paper III show that the thermoelectric signature of QDs is indeed altered in the presence of the Kondo correlations as predicted by Costi & Zlatić in Ref. [19]. The measured thermocurrent I_{th} changes direction within the Kondo region (indicated in the Fig. 34a) as the temperature T is lowered. As a result of the I_{th} inversion within the Kondo region, the curves lose two zero-crossings. The systematic observation of this behavior constitutes a verification of the Kondo-related sign reversal of the thermoelectric signal predicted in Ref. [19]. Furthermore, the temperature values at which the sign reversals were detected in all three devices were also found to be consistent with the calculation results in Ref. [19].

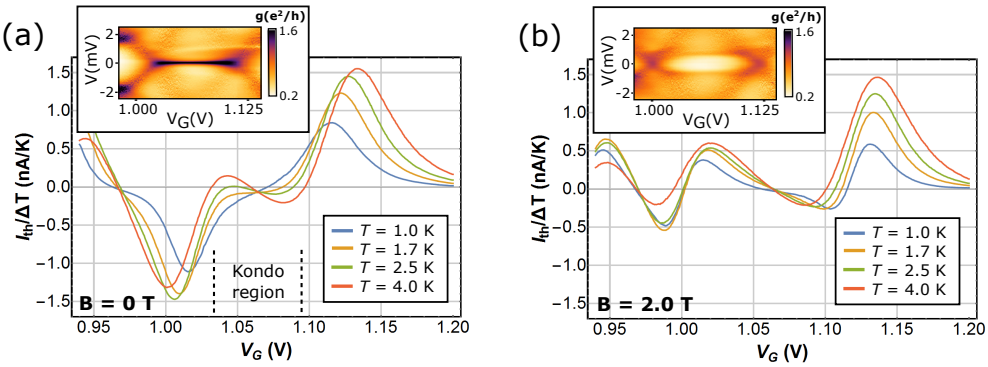


Figure 34: Thermocurrent I_{th} as a function of gate voltage V_G at different temperatures T, measured on the device shown in Fig. 33. The I_{th} is normalized by the thermal bias ΔT. (a) and (b) are different by the value of the magnetic field B, which is indicated in the figures. The color plot insets show the charge stability diagram at the corresponding magnetic field values.

In addition to the I_{th} measurements at zero magnetic field, identical measurements were repeated in the presence of a finite field B. As can be seen in the stability diagrams in Fig. 34, the magnetic field is found to Zeeman split, and eventually destroy, the Kondo resonance. A strong magnetic field effectively turns Kondo-correlated QDs into QDs with non-degenerate spin states. As a result, the corresponding I_{th} signal shows the conventional number of zero-crossings and no qualitative changes in $I_{th}(V_G)$ with temperature T can be observed. Unexpectedly, the measurements indicate that Zeeman energies ΔE_{Zeeman} above kT_K are needed to destroy the Kondo related I_{th} inversion.

The experimental work in Paper III shows that InAs/InP nanowire QDs can be used to carry out state of the art thermoelectric experiments on Kondo correlated QDs. The nanowire QDs used in the thermoelectric devices with the top-heater geometry enable convenient studies of thermoelectric effects in Kondo correlated QDs, also in the presence of magnetic field. This means that more studies of thermal bias and magnetic field dependencies on Kondo-correlated QDs are likely to be realized in similar experimental systems. Such studies could be used to further test the applicability limits of theoretical models and their

approximations in the Kondo regime.

7 Summary and outlook

This thesis presents experimental work on thermoelectric properties of QDs. The heterostructured InAs/InP nanowire-based QDs enable thermoelectric studies in both the weak and the strong coupling regimes thanks to the high band-gap offset between InAs and InP (about 0.5 eV). Aside from the material system, a very important role in this thesis is played by the top-heater architecture, that enables the application of significant ΔT across the short QD segments, therefore increasing the thermoelectric signal levels.

The thermocurrent measurements suggest that ΔT in a range of several tens of K is possible to apply across the roughly 20 nm long QDs at cryogenic temperatures. Unfortunately, in the current design devices typically achieve only $\Delta T/T \approx 1/3$, meaning that there is significant heating of the cold QD reservoir. It would be of interest to minimize this parasitic heating effect thus increasing the maximal explorable range of $\Delta T/T$. This is likely to require more detailed studies of the nature of heat flow from the device heater into the substrate in order to understand how the temperature gradients are shaped. Enabling higher $\Delta T/T$ would allow going deeper into the nonlinear ΔT response regime. This would be of particular importance for regimes in which theory predictions are generally harder to obtain, such as for the Kondo regime.

The detailed studies of QD heat engines presented in this thesis have shown that they indeed can be operated with heat-to-electric conversion efficiencies comparable to the best cyclic heat engines. The magnitude of the electronic heat flow through the QDs in these studies is evaluated using Master equation calculations that account for all first and second order tunneling processes [57] because none of the known methods for measuring so small heat flows are applicable to this QD system. The interest of realizing analogous heat flow measurements experimentally would be to verify the predictions of Ref. [57]. The calculations show that certain phenomena observable in heat transport should be absent in charge transport. More specifically, they predict the existence of negative differential heat conductivity at the edges and within the Coulomb blockaded regions.

The possibility of realizing thermoelectric experiments in the strong coupling regime, where Kondo correlations can be observed, opens up a range of possible studies with minor experimental modifications. This could include thermoelectric measurements on two level QDs [134] or more detailed studies in the presence of magnetic fields. Perhaps the most intriguing experimental conditions would be possible if $\Delta T/T$ could be further increased. High ΔT would allow one to conduct transport measurements with one of the QD reservoirs at a temperature within the Kondo regime and the other reservoir at a much higher temperature (similarly as discussed in Ref. [109]).

More generally, thermoelectric measurements have recently been demonstrated to enable the determination of entropy of electronic states in QDs with a simple orbital state struc-

ture [137]. It is foreseeable that similar measurements to those presented in this thesis could be used to learn more about QDs with more complex electronic state structures or about QDs in the Kondo regime.

Appendices

A Growth details of nanowires 0251

All experimental studies presented in Papers I, II and III use nanowires from the same growth (grown by Sofia Fahlvik [prev. Svensson]). Aerosol deposited gold seed particles of 50 nm nominal size where used. The growth substrate was baked at 320°C for 25 min for decomposition of contaminants on the substrate surface. The growth of InAs was carried out at approximately 390°C (measued by an infra-red sensor in the growth system) providing precursor pressures 0.15 mbar and 1.5 mbar for TMIn and TBAs, respectively. For the growth of InP segments 1.0 mbar pressure of TBP was provided instead of TBAs.

B Device fabrication details

Prior to device fabrication the substrates are cleaned in acetone in an ultrasonic bath for a few minutes. Acetone helps to dissolve organic contaminants on the substrate whereas the ultrasound helps to shake off any left over dust from scribing of the substrate. The fabrication starts with a deposition of the nanowires within the write field area. Once the wires are deposited, their positions are mapped and the devices are designed, the fabrication process is followed by a standard electron beam lithography (EBL) process for fabrication of contact leads to InAs nanowires.

A resist PMMA 950 A5 is spin-coated onto the substrate at 5000 rpm for 60 s and then baked for 5 min at 180°C. Typically used parameters for the EBL exposure are 20 keV acceleration of electrons in the electron beam with a typical resist exposure dose of 260 $\mu C/cm^2$. The exposed pattern is developed in a developer for PMMA (MIBK:IPA 1:3) for 45 s. In order to to get rid of any residual resist within the developed pattern, the development is complemented by an oxygen plasma ash at 5 mbar oxygen pressure for 30 s. Prior to the evaporation of contact lead metals, the nanowire segments, that are opened for contacting, are passivated in a bath of dilute ammonium polysulfide $(NH_4)_2S_x$ for removal of native oxide. The ammonium polysulfide is dilute in water approximately 1:130 and the passivation takes place for 2 min at 40°C. Immediately after the passivation, the sample is placed in a vacuum chamber of the evaporator. Thermal evaporation of metals is used for deposition of contact leads. Typically, the composition of contact leads is 25 nm of Ni (for making a better contact to the InAs nanowire) with 75 nm of Au on top. After the evaporation, a lift-off of the resist is carried out in acetone. The remaining resist residues are ashed away using a complementary oxygen plasma ash for 30 s at 5 mbar pressure.

After fabrication of the contact leads, the next step is the fabrication of the top-heaters. The nanowires along with the contact leads are covered in approximately 10 nm of hafnium

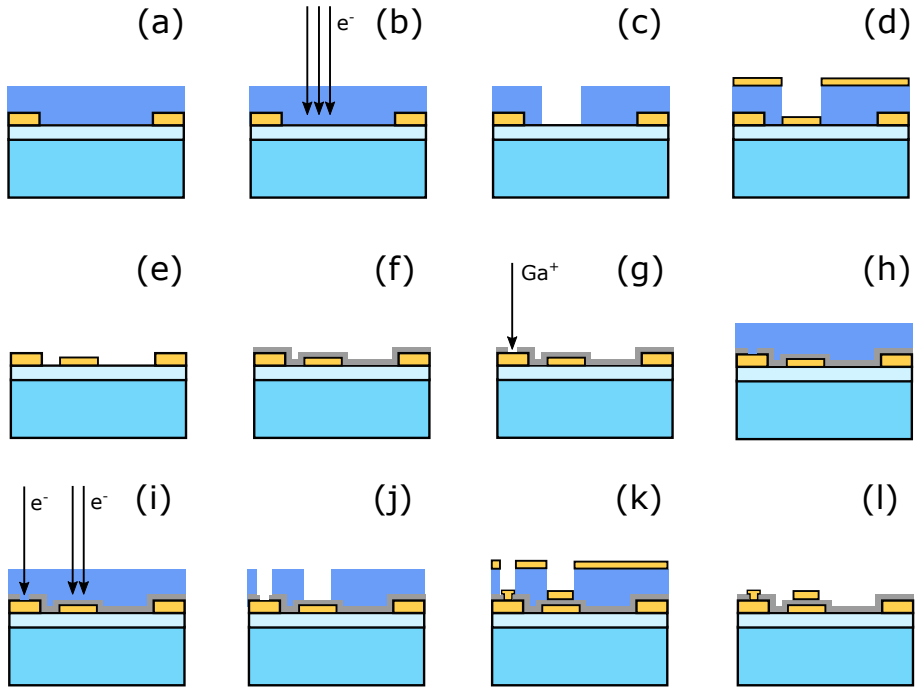


Figure 35: Illustration of a fabrication process. The fabrication stats with a Si/SiO_2 substrate (light blue) with gold contact pads (yellow) on top. (a) The resist (dark blue) is spin-coated on the substrate. (b) A pattern in the resist is exposed to an electron beam. (c) The exposed pattern in the resist is developed. (d) The developed pattern is metallized. (e) The resist is lifted off leaving the metallized pattern on the substrate. (f) An oxide is deposited covering the pattern and the contact pads. (g) Openings in the oxide are opened by milling with a focused Ga ion beam. (h) The resist is spin-coated on the substrate again (i) A new pattern in the resist is exposed to an electron beam. (j) The new pattern in the resist is developed. (k) The new pattern is metallized. (l) The resist is lifted of leaving a second layer of pattern on the substrate that insulated form the first layer and making a contact to certain contact pads.

oxide (HfO_2) using atomic layer deposition (ALD). This is done to electrically insulate the contact leads from the heater leads. In order to open a hole in HfO_2 for contacting the bonding pad extensions, the oxide is milled by a focused Ga ion beam. The following steps repeat the EBL process from the contact lead fabrication, just applied to the fabrication of the heater leads. The only differences being that no passivation is done and the thickness of the evaporated Au leads is 100 nm.

References

[1] J. W. Orton, *Semiconductors and the information revolution: magic crystals that made it happen*. Elsevier, 2009.

[2] M. Golio, "Fifty years of moore's law," *Proceedings of the IEEE*, vol. 103, pp. 1932–1937, oct 2015.

[3] D. L. Pulfrey, *Understanding Modern Transistors and Diodes*. Cambridge University Press, 2009.

[4] S. Greengard, "The future of semiconductors," *Communications of the ACM*, vol. 60, pp. 18–20, feb 2017.

[5] I. C. Ng and S. Y. Wakenshaw, "The internet-of-things: Review and research directions," *International Journal of Research in Marketing*, vol. 34, pp. 3–21, mar 2017.

[6] X. Li, C. Wang, and G. Yang, "Thermodynamic theory of growth of nanostructures," *Progress in Materials Science*, vol. 64, pp. 121–199, jul 2014.

[7] J. Johansson and K. A. Dick, "Recent advances in semiconductor nanowire heterostructures," *CrystEngComm*, vol. 13, no. 24, p. 7175, 2011.

[8] J. K. Hyun, S. Zhang, and L. J. Lauhon, "Nanowire heterostructures," *Annual Review of Materials Research*, vol. 43, pp. 451–479, jul 2013.

[9] P. D. Kanungo, H. Schmid, M. T. Björk, L. M. Gignac, C. Breslin, J. Bruley, C. D. Bessire, and H. Riel, "Selective area growth of III–v nanowires and their heterostructures on silicon in a nanotube template: towards monolithic integration of nanodevices," *Nanotechnology*, vol. 24, p. 225304, may 2013.

[10] C. Zhao, T. K. Ng, R. T. ElAfandy, A. Prabaswara, G. B. Consiglio, I. A. Ajia, I. S. Roqan, B. Janjua, C. Shen, J. Eid, A. Y. Alyamani, M. M. El-Desouki, and B. S. Ooi, "Droop-free, reliable, and high-power InGaN/GaN nanowire light-emitting diodes for monolithic metal-optoelectronics," *Nano Letters*, vol. 16, pp. 4616–4623, jun 2016.

[11] X. Duan, Y. Huang, R. Agarwal, and C. M. Lieber, "Single-nanowire electrically driven lasers," *Nature*, vol. 421, pp. 241–245, jan 2003.

[12] M. Karimi, M. Heurlin, S. Limpert, V. Jain, X. Zeng, I. Geijselaers, A. Nowzari, Y. Fu, L. Samuelson, H. Linke, M. T. Borgström, and H. Pettersson, "Intersubband quantum disc-in-nanowire photodetectors with normal-incidence response in the long-wavelength infrared," *Nano Letters*, vol. 18, pp. 365–372, dec 2017.

[13] G. Otnes and M. T. Borgström, "Towards high efficiency nanowire solar cells," *Nano Today*, vol. 12, pp. 31–45, feb 2017.

[14] C. Zota, F. Lindelöw, L.-E. Wernersson, and E. Lind, "High-frequency InGaAs tri-gate MOSFETs with f max of 400 GHz," *Electronics Letters*, vol. 52, pp. 1869–1871, oct 2016.

[15] A. Shakouri, "Nanoscale thermal transport and microrefrigerators on a chip," *Proceedings of the IEEE*, vol. 94, pp. 1613–1638, aug 2006.

[16] M. T. Björk, B. J. Ohlsson, T. Sass, A. I. Persson, C. Thelander, M. H. Magnusson, K. Deppert, L. R. Wallenberg, and L. Samuelson, "One-dimensional heterostructures in semiconductor nanowhiskers," *Applied Physics Letters*, vol. 80, pp. 1058–1060, feb 2002.

[17] M. T. Björk, C. Thelander, A. E. Hansen, L. E. Jensen, M. W. Larsson, L. R. Wallenberg, and L. Samuelson, "Few-electron quantum dots in nanowires," *Nano Letters*, vol. 4, pp. 1621–1625, sep 2004.

[18] C. W. J. Beenakker and A. A. M. Staring, "Theory of the thermopower of a quantum dot," *Physical Review B*, vol. 46, pp. 9667–9676, oct 1992.

[19] T. A. Costi and V. Zlatić, "Thermoelectric transport through strongly correlated quantum dots," *Physical Review B*, vol. 81, p. 235127, 2010.

[20] A. Svilans, M. Leijnse, and H. Linke, "Experiments on the thermoelectric properties of quantum dots," *Comptes Rendus Physique*, vol. 17, no. 10, pp. 1096 – 1108, 2016.

[21] G. D. Mahan and J. O. Sofo, "The best thermoelectric.," *Proceedings of the National Academy of Sciences*, vol. 93, pp. 7436–7439, jul 1996.

[22] T. E. Humphrey, R. Newbury, R. P. Taylor, and H. Linke, "Reversible quantum brownian heat engines for electrons," *Physical Review Letters*, vol. 89, p. 116801, aug 2002.

[23] T. Humphrey and H. Linke, "Quantum, cyclic, and particle-exchange heat engines," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 29, pp. 390–398, oct 2005.

[24] J. Kondo, "Resistance minimum in dilute magnetic alloys," *Progress of Theoretical Physics*, vol. 32, pp. 37–49, jul 1964.

[25] A. Hewson, *The Kondo Problem to Heavy Fermions*. New York: Cambridge University Press, 1993.

[26] R. A. Diaz and W. J. Herrera, "The positivity and other properties of the matrix of capacitance: Physical and mathematical implications," *Journal of Electrostatics*, vol. 69, pp. 587–595, dec 2011.

[27] K. Uchida, K. Matsuzawa, J. Koga, R. Ohba, S. ichi Takagi, and A. Toriumi, "Analytical single-electron transistor (SET) model for design and analysis of realistic SET circuits," *Japanese Journal of Applied Physics*, vol. 39, pp. 2321–2324, apr 2000.

[28] C. W. J. Beenakker, "Theory of coulomb-blockade oscillations in the conductance of a quantum dot," *Physical Review B*, vol. 44, p. 1646, 1991.

[29] K. Likharev, "Single-electron devices and their applications," *Proceedings of the IEEE*, vol. 87, pp. 606–632, apr 1999.

[30] W. Heisenberg, *The Physical Principles of the Quantum Theory*. Dover Publications, 1949.

[31] F. Berezin and M. Shubin, *The Schrödinger Equation (Mathematics and its Applications)*. Springer, 2012.

[32] T. Ihn, *Semiconductor Nanostructures: Quantum states and electronic transport*. Oxford University Press, 2010.

[33] D. K. Ferry, *Semiconductors: Bonds and bands (IOP Expanding Physics)*. Institute of Physics Publishing, 2013.

[34] J. C. Inkson, *Many-Body Theory of Solids: An Introduction*. Springer, 2012.

[35] R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem: Second Edition (Dover Books on Physics)*. Dover Publications, 1992.

[36] M. G. Burt, "The justification for applying the effective-mass approximation to microstructures," *Journal of Physics: Condensed Matter*, vol. 4, no. 32, p. 6651, 1992.

[37] B. K. Ridley, *Electrons and Phonons in Semiconductor Multilayers*. Cambridge University Press, 1996.

[38] G. Bastard, "Superlattice band structure in the envelope-function approximation," *Physical Review B*, vol. 24, pp. 5693–5697, nov 1981.

[39] M. G. Burt, "An exact formulation of the envelope function method for the determination of electronic states in semiconductor microstructures," *Semiconductor Science and Technology*, vol. 3, no. 8, p. 739, 1988.

[40] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, vol. 55. National Bureau of Standards, 1964.

[41] A. S. Baltenkov and A. Z. Msezane, "Electronic quantum confinement in cylindrical potential well," *The European Physical Journal D*, vol. 70, p. 81, apr 2016.

[42] C. L. Beattie, "Table of first 700 zeros of bessel functions - jl(x) and j'l(x)," *Bell System Technical Journal*, vol. 37, pp. 689–697, may 1958.

[43] M. T. Björk, C. Thelander, A. E. Hansen, L. E. Jensen, M. W. Larsson, L. R. Wallenberg, and L. Samuelson, "Few-electron quantum dots in nanowires," *Nano Letters*, vol. 4, pp. 1621–1625, sep 2004.

[44] B. C. Hall, *Quantum Theory for Mathematicians (Graduate Texts in Mathematics)*. Springer, 2013.

[45] J. Bardeen, "Tunnelling from a many-particle point of view," *Physical Review Letters*, vol. 6, pp. 57–59, jan 1961.

[46] M. C. Payne, "Transfer hamiltonian description of resonant tunnelling," *Journal of Physics C: Solid State Physics*, vol. 19, no. 8, p. 1145, 1986.

[47] C. Noguera, "Validity of the transfer hamiltonian approach : application to the STM spectroscopic mode," *Journal de Physique*, vol. 50, no. 18, pp. 2587–2599, 1989.

[48] T. Heinzel, ed., *Mesoscopic Electronics in Solid State Nanostructures*. Wiley-VCH Verlag GmbH, nov 2006.

[49] W. Greiner, *Relativistic Quantum Mechanics. Wave Equations*. Springer, 2000.

[50] R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. John Wiley & Sons, 1985.

[51] B. Thaller, *The Dirac Equation (Theoretical and Mathematical Physics)*. Springer, 1993.

[52] J. B. Green and R. A. Loring, "Zeeman effect and structure in the spark spectra of tin," *Proceedings of the National Academy of Sciences*, vol. 13, pp. 492–495, jul 1927.

[53] L. P. Kouwenhoven, D. G. Austing, and S. Tarucha, "Few-electron quantum dots," *Reports on Progress in Physics*, vol. 64, p. 701, 2001.

[54] L. D. Landau and E. Lifshitz, *Statistical Physics, Third Edition, Part 1: Volume 5 (Course of Theoretical Physics, Volume 5)*. Butterworth-Heinemann, 1980.

[55] E. Bonet, M. M. Deshmukh, and D. C. Ralph, "Solving rate equations for electron tunneling via discrete quantum states," *Physical Review B*, vol. 65, p. 045317, 2002.

[56] S. D. Franceschi, S. Sasaki, J. M. Elzerman, W. G. van der Wiel, S. Tarucha, and L. P. Kouwenhoven, "Electron cotunneling in a semiconductor quantum dot," *Physical Review Letters*, vol. 86, pp. 878–881, jan 2001.

[57] N. M. Gergs, C. B. M. Hörig, M. R. Wegewijs, and D. Schuricht, "Charge fluctuations in nonlinear heat transport," *Physical Review B*, vol. 91, p. 201107(R), may 2015.

[58] N. M. Gergs, S. A. Bender, R. A. Duine, and D. Schuricht, "Spin Switching via Quantum Dot Spin Valves," *Physical Review Letters*, vol. 120, p. 017701, 2018.

[59] I.-J. Chen, A. Burke, A. Svilans, H. Linke, and C. Thelander, "Thermoelectric power factor limit of a 1d nanowire," *Physical Review Letters*, vol. 120, p. 177703, apr 2018.

[60] H. Grabert and M. Devoret, *Single Charge Tunneling: Coulomb Blockade Phenomena In Nanostructures*. Nato Science Series B:, Springer, 2013.

[61] D. M. Zumbühl, C. M. Marcus, M. P. Hanson, and A. C. Gossard, "Cotunneling spectroscopy in few-electron quantum dots," *Physical Review Letters*, vol. 93, p. 256801, dec 2004.

[62] D. Ryndyk, *Theory of Quantum Transport at Nanoscale: An Introduction (Springer Series in Solid-State Sciences)*. Springer, 2015.

[63] Y. Meir and N. S. Wingreen, "Landauer formula for the current through an interacting electron region," *Physical Review Letters*, vol. 68, p. 2512, 1992.

[64] J. H. Davies, *The Physics of Low-Dimensional Semiconductors*. Cambridge University Press, 1997.

[65] N. Nakpathomkun, H. Q. Xu, and H. Linke, "Thermoelectric efficiency at maximum power in low-dimensional systems," *Physical Review B*, vol. 82, p. 235428, dec 2010.

[66] C. R. Proetto, "Thermopower oscillations of a quantum-point contact," *Physical Review B*, vol. 44, pp. 9096–9099, oct 1991.

[67] B. Wang, J. Zhou, R. Yang, and B. Li, "Ballistic thermoelectric transport in structured nanowires," *New Journal of Physics*, vol. 16, p. 065018, jun 2014.

[68] R. Scheibner, E. G. Novik, T. Borzenko, M. König, D. Reuter, A. D. Wieck, H. Buhmann, and L. W. Molenkamp, "Sequential and cotunneling behavior in the temperature-dependent thermopower of few-electron quantum dots," *Physical Review B*, vol. 75, p. 041301(R), jan 2007.

[69] S. F. Svensson, A. I. Persson, E. A. Hoffmann, N. Nakpathomkun, H. A. Nilsson, H. Q. Xu, L. Samuelson, and H. Linke, "Lineshape of the thermopower of quantum dots," *New Journal of Physics*, vol. 14, p. 033041, mar 2012.

[70] C. Gorter, "A possible explanation of the increase of the electrical resistance of thin metal films at low temperatures and small field strengths," *Physica*, vol. 17, pp. 777–780, aug 1951.

[71] K. K. Likharev, "Correlated discrete transfer of single electrons in ultrasmall tunnel junctions," *IBM Journal of Research and Development*, vol. 32, pp. 144–158, jan 1988.

[72] D. V. Averin and K. K. Likharev, "Single electronics: A correlated transfer of single electrons and cooper pairs in systems of small tunnel junctions," in *Mesoscopic Phenomena in Solids*, pp. 173–271, 1991.

[73] R. Dingle, H. L. Störmer, A. C. Gossard, and W. Wiegmann, "Electron mobilities in modulation-doped semiconductor heterojunction superlattices," *Applied Physics Letters*, vol. 33, pp. 665–667, oct 1978.

[74] L. Pfeiffer, K. W. West, H. L. Stormer, and K. W. Baldwin, "Electron mobilities exceeding 107cm2/v s in modulation-doped GaAs," *Applied Physics Letters*, vol. 55, pp. 1888–1890, oct 1989.

[75] T. J. Thornton, M. Pepper, H. Ahmed, D. Andrews, and G. J. Davies, "One-dimensional conduction in the 2d electron gas of a GaAs-AlGaAs heterojunction," *Physical Review Letters*, vol. 56, pp. 1198–1201, mar 1986.

[76] C. G. Smith, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, "The transition from one- to zero-dimensional ballistic transport," *Journal of Physics C: Solid State Physics*, vol. 21, no. 24, p. L893, 1988.

[77] B. J. van Wees, L. P. Kouwenhoven, C. J. P. M. Harmans, J. G. Williamson, C. E. Timmering, M. E. I. Broekaart, C. T. Foxon, and J. J. Harris, "Observation of zero-dimensional states in a one-dimensional electron interferometer," *Physical Review Letters*, vol. 62, pp. 2523–2526, may 1989.

[78] A. A. M. Staring, L. W. Molenkamp, B. W. Alphenaar, H. van Houten, O. J. A. Buyk, M. A. A. Mabesoone, C. W. J. Beenakker, and C. T. Foxon, "Coulomb-blockade oscillations in the thermopower of a quantum dot," *Europhysics Letters*, vol. 22, p. 57, 1993.

[79] A. S. Dzurak, C. G. Smith, C. H. W. Barnes, M. Pepper, L. Martín-Moreno, C. T. Liang, D. A. Ritchie, and G. A. C. Jones, "Thermoelectric signature of the excitation spectrum of a quantum dot," *Physical Review B*, vol. 55, pp. R10197–R10200, apr 1997.

[80] J. P. Small, K. M. Perez, and P. Kim, "Modulation of thermoelectric power of individual carbon nanotubes," *Physical Review Letters*, vol. 91, p. 256801, dec 2003.

[81] G. D. Nessim, "Properties, synthesis, and growth mechanisms of carbon nanotubes with special focus on thermal chemical vapor deposition," *Nanoscale*, vol. 2, no. 8, p. 1306, 2010.

[82] P. Reddy, S.-Y. Jang, R. A. Segalman, and A. Majumdar, "Thermoelectricity in molecular junctions," *Science*, vol. 315, pp. 1568–1571, mar 2007.

[83] E. Sivre, A. Anthore, F. D. Parmentier, A. Cavanna, U. Gennser, A. Ouerghi, Y. Jin, and F. Pierre, "Heat coulomb blockade of one ballistic channel," *Nature Physics*, vol. 14, pp. 145–148, oct 2017.

[84] S. F. Svensson, E. A. Hoffmann, N. Nakpathomkun, P. M. Wu, H. Q. Xu, H. A. Nilsson, D. Sánchez, V. Kashcheyevs, and H. Linke, "Nonlinear thermovoltage and thermocurrent in quantum dots," *New Journal of Physics*, vol. 15, p. 105011, oct 2013.

[85] J. Small, "Thermopower measurement of individual single walled carbon nanotubes," *Microscale Thermophysical Engineering*, vol. 8, pp. 1–5, jan 2004.

[86] J. G. Gluschke, S. F. Svensson, C. Thelander, and H. Linke, "Fully tunable, non-invasive thermal biasing of gated nanostructures suitable for low-temperature studies," *Nanotechnology*, vol. 25, no. 38, p. 385704, 2014.

[87] H. van Houten, L. W. Molenkamp, C. W. J. Beenakker, and C. T. Foxon, "Thermoelectric properties of quantum point contacts," *Semiconductor Science and Technology*, vol. 7, no. 3B, p. B215, 1992.

[88] A. Dzurak, C. Smith, M. Pepper, D. Ritchie, J. Frost, G. Jones, and D. Hasko, "Observation of coulomb blockade oscillations in the thermopower of a quantum dot," *Solid State Communications*, vol. 87, pp. 1145–1149, sep 1993.

[89] R. Scheibner, H. Buhmann, D. Reuter, M. N. Kiselev, and L. W. Molenkamp, "Thermopower of a kondo spin-correlated quantum dot," *Physical Review Letters*, vol. 95, p. 176602, oct 2005.

[90] T. Ando, "Theory of quantum transport in a two-dimensional electron system under magnetic fields II. single-site approximation under strong fields," *Journal of the Physical Society of Japan*, vol. 36, pp. 1521–1529, jun 1974.

[91] R. Fletcher, J. J. Harris, C. T. Foxon, and R. Stoner, "Hot-electron temperatures of two-dimensional electron gases using both de haas–shubnikov oscillations and the electron-electron interaction effect," *Physical Review B*, vol. 45, pp. 6659–6669, mar 1992.

[92] T.-Y. Ko, M. Shellaiah, and K. W. Sun, "Thermal and thermoelectric transport in highly resistive single sb2se3 nanowires and nanowire bundles," *Scientific Reports*, vol. 6, oct 2016.

[93] M. E. Brett and J. E. Black, "Two OPW calculations of electronic transport properties of copper below 20 k," *Canadian Journal of Physics*, vol. 55, pp. 521–527, mar 1977.

[94] E. A. Hoffmann, N. Nakpathomkun, A. I. Persson, H. Linke, H. A. Nilsson, and L. Samuelson, "Quantum-dot thermometry," *Applied Physics Letters*, vol. 91, p. 252114, dec 2007.

[95] B. L. Gallagher, T. Galloway, P. Beton, J. P. Oxley, S. P. Beaumont, S. Thoms, and C. D. W. Wilkinson, "Observation of universal thermopower fluctuations," *Physical Review Letters*, vol. 64, pp. 2058–2061, apr 1990.

[96] L. W. Molenkamp, H. van Houten, C. W. J. Beenakker, R. Eppenga, and C. T. Foxon, "Quantum oscillations in the transverse voltage of a channel in the nonlinear transport regime," *Physical Review Letters*, vol. 65, pp. 1052–1055, aug 1990.

[97] A. Dzurak, C. Smith, M. Pepper, D. Ritchie, J. Frost, G. Jones, and D. Hasko, "Observation of coulomb blockade oscillations in the thermopower of a quantum dot," *Solid State Communications*, vol. 87, pp. 1145–1149, sep 1993.

[98] N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys*. Dover Publications, 1958.

[99] M. Turek and K. A. Matveev, "Cotunneling thermopower of single electron transistors," *Physical Review B*, vol. 65, p. 115332, mar 2002.

[100] J. Koch, F. von Oppen, Y. Oreg, and E. Sela, "Thermopower of single-molecule devices," *Physical Review B*, vol. 70, p. 195107, nov 2004.

[101] J. R. Schrieffer and P. A. Wolff, "Relation between the anderson and kondo hamiltonians," *Physical Review*, vol. 149, p. 491, 1966.

[102] C. H. L. Quay, J. Cumings, S. J. Gamble, A. Yazdani, H. Kataura, and D. Goldhaber-Gordon, "Transport properties of carbon nanotube c60 peapods," *Physical Review B*, vol. 76, p. 073404, 2007.

[103] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, "Kondo effect in a single-electron transistor," *Nature*, vol. 391, p. 156, 1998.

[104] J. Nygård, D. H. Cobden, and P. E. Lindelof, "Kondo physics in carbon nanotubes," *Nature*, vol. 408, p. 342, 2000.

[105] G. D. Scott, Z. K. Keane, J. W. Ciszek, J. M. Tour, and D. Natelson, "Universal scaling of nonequilibrium transport in the kondo regime of single molecule devices," *Physical Review B*, vol. 79, p. 165413, Apr 2009.

[106] A. V. Kretinin, R. Popovitz-Biro, D. Mahalu, and H. Shtrikman, "Multimode fabry-perot conductance oscillations in suspended stacking-faults-free InAs nanowires," *Nano Letters*, vol. 10, pp. 3439–3445, sep 2010.

[107] R. Scheibner, *Thermoelectric properties of few-electron quantum dots*. PhD thesis, University of Würzburg, 2007.

[108] H. Thierschmann, *Heat conversion in quantum dot systems*. PhD thesis, University of Würzburg, 2014.

[109] M. A. Sierra, R. López, and D. Sánchez, "Fate of the spin-12 kondo effect in the presence of temperature gradients," *Physical Review B*, vol. 96, p. 085416, aug 2017.

[110] T. E. Humphrey and H. Linke, "Reversible thermoelectric nanomaterials," *Physical Review Letters*, vol. 94, p. 096601, mar 2005.

[111] M. Esposito, K. Lindenberg, and C. V. den Broeck, "Universality of efficiency at maximum power," *Physical Review Letters*, vol. 102, p. 130602, apr 2009.

[112] F. L. Curzon and B. Ahlborn, "Efficiency of a Carnot engine at maximum power output," *American Journal of Physics*, vol. 43, pp. 22–24, Jan. 1975.

[113] C. Van den Broeck, "Thermodynamic efficiency at maximum power," *Physical Review Letters*, vol. 95, p. 190602, nov 2005.

[114] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, "Efficiency at maximum power of low-dissipation carnot engines," *Physical Review Letters*, vol. 105, p. 150603, Oct 2010.

[115] E. A. Hoffmann, *The thermoelectric efficiency of quantum dots in indium arsenide/indium phosphide nanowires*. PhD thesis, University of Oregon, 2009.

[116] W. T. Tsang, "Chemical beam epitaxy of InP and GaAs," *Applied Physics Letters*, vol. 45, pp. 1234–1236, dec 1984.

[117] M. E. Messing, K. Hillerich, J. Bolinsson, K. Storm, J. Johansson, K. A. Dick, and K. Deppert, "A comparative study of the effect of gold seed particle preparation method on nanowire growth," *Nano Research*, vol. 3, pp. 506–519, jun 2010.

[118] L. Fröberg, *Growth, physics, and device applications of InAs-based nanowires*. PhD thesis, Lund University, 2008.

[119] C. Foxon and B. Joyce, "Interaction kinetics of as4 and ga on {100} GaAs surfaces using a modulated molecular beam technique," *Surface Science*, vol. 50, pp. 434–450, jun 1975.

[120] S. F. Svensson, S. Jeppesen, C. Thelander, L. Samuelson, H. Linke, and K. A. Dick, "Control and understanding of kink formation in InAs–InP heterostructure nanowires," *Nanotechnology*, vol. 24, p. 345601, jul 2013.

[121] A. S. Dzurak, C. G. Smith, C. H. W. Barnes, M. Pepper, L. Martín-Moreno, C. T. Liang, D. A. Ritchie, and G. A. C. Jones, "Thermoelectric signature of the excitation spectrum of a quantum dot," *Physical Review B*, vol. 55, pp. R10197–R10200, Apr 1997.

[122] S. M. Cronenwett, T. H. Oosterkamp, and L. P. Kouwenhoven, "A tunable kondo effect in quantum dots," *Science*, vol. 281, p. 540, 1998.

[123] M. A. Sierra and D. Sánchez, "Strongly nonlinear thermovoltage and heat dissipation in interacting quantum dots," *Physical Review B*, vol. 90, p. 115313, sep 2014.

[124] L. D. Hicks and M. S. Dresselhaus, "Effect of quantum-well structures on the thermoelectric figure of merit," *Physical Review B*, vol. 47, pp. 12727–12731, may 1993.

[125] L. D. Hicks and M. S. Dresselhaus, "Thermoelectric figure of merit of a one-dimensional conductor," *Physical Review B*, vol. 47, pp. 16631–16634, jun 1993.

[126] F. L. Curzon and B. Ahlborn, "Efficiency of a carnot engine at maximum power output," *American Journal of Physics*, vol. 43, pp. 22–24, jan 1975.

[127] C. V. den Broeck, "Thermodynamic efficiency at maximum power," *Physical Review Letters*, vol. 95, p. 190602, nov 2005.

[128] J. Schmid, J. Weis, K. Eberl, and K. von Klitzing, "A quantum dot in the limit of strong coupling to reservoirs," *Physica B*, vol. 256, p. 182, 1998.

[129] F. Simmel, R. H. Blick, J. P. Kotthaus, W. Wegscheider, and M. Bichler, "Anomalous kondo effect in a quantum dot at nonzero bias," *Physical Review Letters*, vol. 83, p. 804, 1999.

[130] D. Boese and R. Fazio, "Thermoelectric effects in kondo-correlated quantum dots," *Europhysics Letters*, vol. 56, p. 576, 2001.

[131] P. Roura-Bas, L. Tosi, A. A. Aligia, and P. S. Cornaglia, "Thermopower of an SU(4) kondo resonance under an SU(2) symmetry-breaking field," *Physical Review B*, vol. 86, p. 165106, oct 2012.

[132] J. Azema, A.-M. Daré, S. Schäfer, and P. Lombardo, "Kondo physics and orbital degeneracy interact to boost thermoelectrics on the nanoscale," *Physical Review B*, vol. 86, p. 075303, aug 2012.

[133] I. Weymann and J. Barnaś, "Spin thermoelectric effects in kondo quantum dots coupled to ferromagnetic leads," *Physical Review B*, vol. 88, p. 085313, aug 2013.

[134] L. Ye, D. Hou, R. Wang, D. Cao, X. Zheng, and Y. Yan, "Thermopower of few-electron quantum dots with kondo correlations," *Physical Review B*, vol. 90, p. 165116, oct 2014.

[135] A. Dorda, M. Ganahl, S. Andergassen, W. von der Linden, and E. Arrigoni, "Thermoelectric response of a correlated impurity in the nonequilibrium kondo regime," *Physical Review B*, vol. 94, p. 245125, dec 2016.

[136] D. B. Karki and M. N. Kiselev, "Thermoelectric transport through a SU(n) kondo impurity," *Physical Review B*, vol. 96, p. 121403, sep 2017.

[137] N. Hartman, C. Olsen, S. Lüscher, M. Samani, S. Fallahi, G. C. Gardner, M. Manfra, and J. Folk, "Direct entropy measurement in a mesoscopic quantum system," *Nature Physics*, 2018.

Acknowledgements

First and foremost, I would like to thank my supervisor Heiner Linke for everything he has done to make my work here successful. Already starting from the interview I was impressed by how well you manage the dual role of being a scientist and a director. It requires an enormous amount of passion and dedication, which we all can look after. In many ways, the project you took me on board, has been your "baby" for well over a decade. I think that it matched my skillset and interests much better than I would have hoped, which is partly why the sailing has been relatively smooth. I am grateful for being able to contribute to this project. I have had a lot of exciting moments while working on it, which I will surely remember with smile and appreciation for the rest of my life. I have learned so much from you as a scientist and a leader, but also from you as a personality.

Second, I would like to thank to my co-supervisor Adam M. Burke. You have been of such importance for my work here. Starting from the very beginning you took up a leading role in setting me on track and taught me almost everything I lacked to handle the practical matters on my own. You have been accessible and helpful, often after the working hours, all throughout my time here in Lund, despite the responsibilities of becoming a parent (twice). I dare to say that our relationship extends well beyond the formal title of a "co-supervisor" and I hope to remain in close contact with you in the future.

Many thanks also goes to my other supervisor Claes Thelander for the enormous help with the electrical measurement control. Your Labview interface with all the options has saved so much time and energy for everyone using it. I would even say that some of the most critical measurements in certain studies would not have happened without the user friendly automatization simply because of the time constraints. Your advice and help in device fabrication has also been of importance for running things smooth.

I would also like to thank Martin Josefsson and Martin Leijnse who have been critical collaborators during my work and studies. Martin L., in many ways it has felt like you have been my informal co-supervisor. Thanks to you, over the past years I have learned a lot about the theory behind the experiments I did, which has significantly contributed to the end results. Martin J., the work we have done together measures in many months worth of working hours, perhaps even over a year. I enjoy working with you and am glad that all our collaborations have been successful. I think it is fair to say that I have never fully understood the "magic" behind all your calculations, but you are definitely very good at it.

Thanks also go to Sofia Fahlvik (prev. Svensson) for an extremely important contribution to almost all my work here. You managed to grow such a great set of quantum dots (the growth 0251) that even up until this day my first choice would be stretching some of the remaining nanowires off that growth substrate rather than trying to use other growths.

Unfortunately, our time here overlapped very little and we never got to know each other better, but I really hope you are doing well.

Along the same lines, I also thank Sebastian Lehmann. It is my understanding that you assisted Sofia with imaging of her growth attempts before I even came to Sweden. Therefore, the success of the "legendary" growth 0251 also somewhat rests on your shoulders. Also thanks for the nanowire TEM images, they have made very good subfigures in publications!

There are many people behind administrative, technical and organizational matters who make sure things run smoothly. It would be hard to imagine all this work being done here without their assistance. Therefore, I would like to send my thanks to Sara Ataran, Jason Beech, Louise Baldetorp, Peter Blomqvist, David Fitzgerald, Mariusz Graczyk, Anders Gustavsson, Abdul-Rehman Hakim, Mia Hedin, Dan Hessman, Maria Huffman, Håkan Ivansson, Sören Jeppesen, Andreas Johansson, Jonas Johansson, Anders Kvennefors, Håkan Lapovski, Line Lundfald, Anneli Löfgren, Leif Magnusson, Ivan Maximov, Sarah McKibbin, Maria Messing, Bengt Meuller, Johanna Mosgeller, Janne Mårtensson, Gerda Rentschler, George Rydnemalm, Charlotte Solberg, Martin Stankovski, Johan Stjernholm, Dmitry Suyatin and others that have directly or indirectly contributed to ensure the great state of the research environment at Solid State Physics and NanoLund.

In the everyday rush it is sometimes hard to appreciate what has been done before one's time. While being here, I have had the privilege to meet gentlemen like Hermann Grimmeiss, Sören Jeppesen, Bo Monemar, Knut Deppert and Lars Samuelson who have played important roles in the development of what is now known to the world as NanoLund. I would like to emphasize how much I respect what you have done and continue doing. I think it is important for any institution to preserve and build upon its core identity. Therefore, I would encourage to maintain the tradition of reminding the new generations about the history of the Solid State Physics by sharing the interesting stories and showing more historic pictures. I always enjoyed it!

I would like to thank Andreas Fuhrer who hosted my secondment at IBM Research Zürich in Switzerland. Despite the fact that we did not get far in two months, mostly for reasons beyond anyone's control, I really enjoyed my time at Rüschlikon. It was very valuable for me to experience the environment at IBM and I learned a fair amount about the topic. The way you approach challenging tasks with enthusiasm and immediate, pragmatic action is something that most, including me, can learn from.

I also want to thank I-Ju, Pierre-Adrien and Steven for all the interesting scientific and nonscientific discussions we have had fun over these years. It has been a pleasure to see your great ideas and personal lives develop. Above all, I think we have become good friends which is what I appreciate the most. Also you, Mats-Erik, have often been in the mix. It has always made it more fun!

I thank my other office mates Vilgaile, Xulu, Marcus and Trung. We have gone through a lot while sitting side by side. The working atmosphere has been great all along. Sometimes I wish we had more time to chat about what is going on, but then again, the work is not going to do itself, right? It has been a pleasure guarding the doorstep for you!

My thanks also go to Stephanie Reimann, Peter Samuelson, Andreas Wacker from Mathematical Physics. I found the opened seminars and joint meetings very useful for both learning more about what you do and for obtaining your feedback on what I was doing. This interaction between divisions is something that I can strongly encourage to expand in the future.

I would like to thank to my closest fellows at the division Kushagr, Damiano, Steve, Adam, Laura, I-Ju, Pyry, Regina, Neimantas, Malin, Vishal, Irene, Oskar, Reza, Frida, Luna, Jonatan, Sven, Lukas and others who all together have been a huge support during the challenging times of being a doctoral student. Special thanks go to Kushagr and Damiano who have been very supportive and reliable friends. Without you this time would have looked very different for me.

The warmest thanks go all my Latvian friends and family members. Despite you not being here, talking to you on regular basis has been important to me. The few times per year I got to see some of you have motivated me to keep going. You are so important to me and I am truly grateful for having you around!

Finally, I particularly would like to thank to my close friends that have been of the most importance over this period - Oskars, Marco, Sendijs, Janis and Roberts. There is no substitute for who you are to me and I deeply thank you for your friendship and support! I also thank you, Alina. You were a big part of my life here and supported me the most. I am very thankful to you for it and wish you the best in all that you do!

Paper I

Physica E 82 (2016) 34–38

Contents lists available at ScienceDirect

Physica E

journal homepage: www.elsevier.com/locate/physe

Nonlinear thermoelectric response due to energy-dependent transport properties of a quantum dot

Artis Svilans *, Adam M. Burke, Sofia Fahlvik Svensson, Martin Leijnse, Heiner Linke

NanoLund and Solid State Physics, Lund University, Box 118, 221 00 Lund, Sweden

HIGHLIGHTS

- Nonlinear thermocurrent in a semiconductor quantum dot.
- Excellent agreement with a Master equation based theoretical model.
- Measurements enabled by novel heater architecture for efficient thermal biasing.

ARTICLE INFO

Article history:
Received 30 June 2015
Received in revised form
30 September 2015
Accepted 5 October 2015
Available online 20 October 2015

Keywords:
Nanowire
Thermoelectrics
Nonlinear
Coulomb blockade
Thermocurrent
Top-heating

ABSTRACT

Quantum dots are useful model systems for studying quantum thermoelectric behavior because of their highly energy-dependent electron transport properties, which are tunable by electrostatic gating. As a result of this strong energy dependence, the thermoelectric response of quantum dots is expected to be nonlinear with respect to an applied thermal bias. However, until now this effect has been challenging to observe because, first, it is experimentally difficult to apply a sufficiently large thermal bias at the nanoscale and, second, it is difficult to distinguish thermal bias effects from purely temperature-dependent effects due to overall heating of a device. Here we take advantage of a novel thermal biasing technique and demonstrate a nonlinear thermoelectric response in a quantum dot which is defined in a heterostructured semiconductor nanowire. We also show that a theoretical model based on the Master equations fully explains the observed nonlinear thermoelectric response given the energy-dependent transport properties of the quantum dot.

1. Introduction

Quantum dots (QDs) are known for their tunable and strongly energy-dependent electron transport properties, which result in a nonlinear response to an applied electrical bias V_{SD}. Nonlinear conductance due to the Coulomb blockade [1] is perhaps the most well known example of such nonlinear behavior. It is also well established that the energy-dependent electron transport properties of QDs strongly influence their thermoelectric behavior [2,3], which has made them attractive model systems for fundamental studies of quantum thermoelectric effects [4–10]. Nonlinear response to an applied thermal bias ΔT, in particular, has been theoretically investigated in various mesoscopic systems, including resonant tunneling structures [11,12], multi-terminal quantum conductors [12–14] and Kondo-correlated devices [15,16]. For QDs, one can expect that the quasi-discrete resonance energy spectrum of a QD alone should lead to nonlinear thermoelectric response [17,18]. This behavior was explored in detail by Sierra and Sanchez who predicted a strongly nonlinear regime behavior in QDs when ΔT is about an order of magnitude larger than the background temperature T_0 [19].

In experiments, a nonlinear thermovoltage as a function of thermal bias ΔT has been observed in semiconductor QDs [3,20,21] and in molecular junctions [22]. Most recent studies using a tunable thermal bias have shown a strongly nonlinear thermovoltage and thermocurrent in semiconductor nanowire QDs that could not be fully explained by the energy-dependence of the QD resonance energy spectrum alone, and was attributed to a renormalization of resonance energies as a function of heating [18].

The key experimental challenge in the observation of nonlinear thermoelectric behavior in QDs is the ability to apply a tunable and large enough thermal bias ΔT across a nanoscale object without significant overall heating of the device. The latter can prevent the ability to perform low-temperature experiments, and makes it difficult to distinguish temperature-dependent transport effects

* Corresponding author.
E-mail addresses: artis.svilans@ftf.lth.se (A. Svilans), heiner.linke@ftf.lth.se (H. Linke).

http://dx.doi.org/10.1016/j.physe.2015.10.007
1386-9477/

from the true nonlinear response to the thermal bias ΔT. Here, we report measurements of a strongly nonlinear thermocurrent as a function of ΔT across a QD that is defined by two InP segments within an InAs nanowire. To a large extent the measurements presented here were enabled by a recently developed heater architecture that allows local and electrically non-invasive thermal biasing of a nanowire [23]. This architecture enables tuning of ΔT over a wide range by applying a relatively small heating power, thus minimizing the parasitic heating effects. We also use theoretical calculations based on Master equations to demonstrate that the experimentally measured thermocurrent can be fully understood from the QD resonance energy spectrum, and is consistent with the previously presented theory in Ref. [19].

2. Experiment

2.1. Device fabrication

The device consists of a heterostructured InAs/InP nanowire with a 60 nm diameter (see Fig. 1a) that was grown by chemical beam epitaxy seeded by a gold particle [24,25]. Based on transmission electron microscopy (TEM) analyses of 11 nanowires from the same growth, the InAs/InP nanowire (starting from the seed particle) consists of a 350 ± 70 nm InAs segment, followed by a 17 ± 1.5 nm long InAs QD defined by two, 4 ± 3 nm thick, InP segments, and a second InAs segment of 265 ± 60 nm in length. The remaining nanowire, which is not used in the device, consists of a 25 nm InP plug incorporated for growth reasons and another InAs segment.

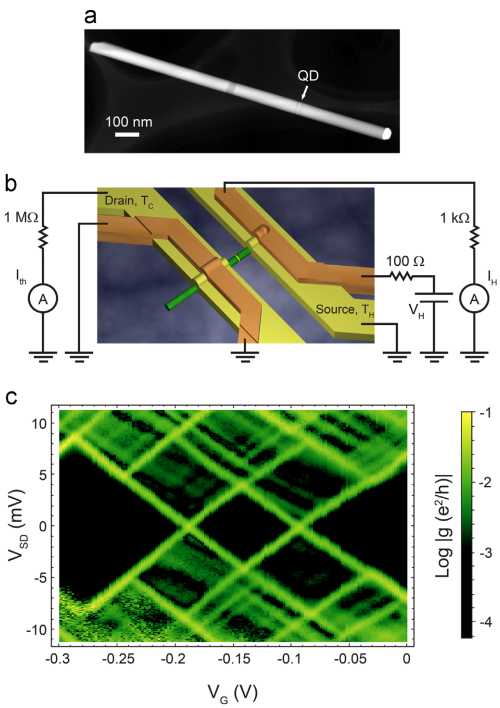


Fig. 1. (a) Transmission electron microscope image of a nanowire nominally identical to the one used in our thermoelectric device. (b) Device schematic with circuitry diagram for the I_{th} measurement setup. The source and drain contacts in yellow, top-heaters in orange, InAs/InP nanowire in green, quantum dot in light green. The heater over the drain lead is unused. (c) Stability diagram of the InAs quantum dot. Magnitude of differential conductivity, $dI/dV_{SD} = g$, in $\log_{10}$-scale as a function of back-gate bias, V_G, and source–drain bias, V_{SD}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The nanowire is contacted to metallic source and drain contacts, as illustrated in Fig. 1b. Electrically isolated metallic top-heaters pass over the source and drain contacts enabling local dissipation of Joule heat directly on top of the contacts; ensuring heat transfer to the nanowire. Only the heater on top of the source contact was used in the experiments presented here. The device fabrication followed the process developed by Gluschke et al. [23]. In brief, electron-beam lithography (EBL) was used to define a pair of source and drain contacts centered around the QD and separated by 300 nm. A dilute sulfur passivation is performed before source and drain contacts are deposited on the nanowire [26]. A 10 nm thick layer of HfO_2 was deposited via atomic layer deposition to insulate the metallic contacts from the overlying heaters, which were aligned and exposed in a second EBL step. Both the contacts and the heaters were deposited thermally with a metal stack of 25 nm Ni and 75 nm Au for the contacts and 25 nm Ni and 125 nm Au for the heaters. The heater layer was thicker to ensure continuity as the heater steps onto the contact region. The entire device rests on 100 nm of thermally grown SiO_2, allowing the underlying doped Si substrate to be used as a global back gate.

2.2. Electrical characterization

Measurements were conducted in a cryostat in which the estimated electron temperature in the device, T_0, was below 1 K without heating. Bias spectroscopy of the device was carried out using a Stanford Research SRS-830 lock-in amplifier. The voltage from the oscillation output was reduced using a 1:20,000 voltage divider circuit to provide a stable AC source–drain bias amplitude $dV_{SD} = 25\ \mu V \ll k_B T_0/e$ (k_B – Boltzmann constant, e – elementary charge). To measure the differential conductance $g = dI/dV_{SD}$ as a function of a DC source–drain bias V_{SD}, the differential current amplitude, dI, was measured in response to dV_{SD}, while adding the AC and DC source–drain bias components in a summing box.

To measure Coulomb oscillations (Fig. 2a), a source–drain current, I_{SD}, was measured in DC mode using Yokogawa 7651 voltage source to bias the source lead at 100 μV and a SR570 current preamplifier with 1 MΩ input impedance.

The set-up used for thermoelectric characterization of the QD nanowire device is shown in Fig. 1b. A thermal bias, ΔT, was applied by running a current I_H through the heater on top of the source contact using a Yokogawa 7651 DC voltage source. The dissipated Joule heat mostly heats the underlying source contact, but is expected to also create a fractional temperature rise in the drain contact [23]. The resulting thermocurrent through the QD nanowire device, I_{th}, was amplified via the SR570 current preamplifier.

2.3. Experimental results and discussion

The QD's stability diagram, measured as a function of the source–drain voltage, V_{SD}, and a back-gate voltage, V_G, is shown in Fig. 1c. The dark diamond-like regions represent bias conditions at which the conductivity is suppressed due to Coulomb blockade. From the bias spectroscopy data we estimate a charging energy E_C of 4.0 ± 0.2 meV, which is a measure of electron–electron interaction strength in the QD. We also determine the value of the coupling constant $\alpha_G = 0.042 \pm 0.04$, which characterizes the capacitive coupling strength between the QD and the back-gate electrode.

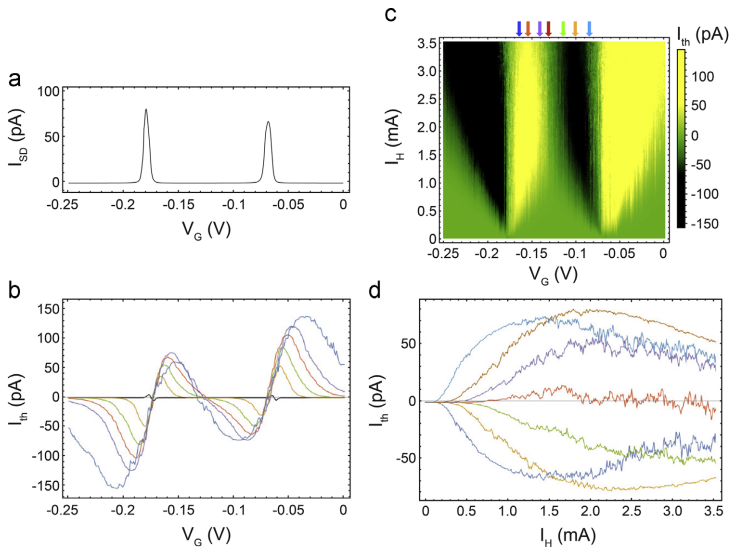


Fig. 2. (a) Coulomb oscillations in source–drain current I_{SD} as a function of back-gate voltage V_G, with the source potential set to 100 μV. (b) Thermocurrent, I_{th}, as a function of back-gate bias for different heater currents I_H = (0, 0.35, 0.70, 1.06, 1.41, 3.17 mA). (c) Thermocurrent (color) as a function of back-gate voltage, V_G, and heating current I_H. Arrows along the top correspond to V_G values for traces in (d) as indicated by their color. (d) Thermocurrent as a function of heating current I_H for different V_G values (− 0.165, − 0.154, − 0.141, − 0.131, − 0.115, − 0.101, − 0.085 V) taken from data in (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 2b shows I_{th} as a function of V_G. The data confirms that our device's thermoelectric response is typical for QDs [2,18,27], where I_{th} goes to zero and changes direction at those V_G values where the Coulomb peaks in Fig. 2a are centered. The locations of these thermocurrent zeros do not depend on the heating current, as can be seen in Fig. 2c, which shows I_{th} as a function of V_G and I_H. This independence of the I_{th} zeros from I_H is in contrast to previous studies [18], where the nonlinear behavior of I_{th} was strongly influenced by a heating dependent renormalization (shift) of the resonance energies of the QD. The stability of the resonances in the present study is attributed to the benefits of the top-heater architecture where a higher ΔT can be applied with much less overall background heating of the device [23].

The core observation of our experiments is the strongly nonlinear behavior of the thermocurrent as a function of ΔT. This nonlinearity is clearly apparent in Fig. 2d where several back-gate voltage traces, taken from the data in the Fig. 2c, are plotted as a function of I_H.

Several key features can be identified in the observed nonlinear behavior of I_{th}, all of which can be understood in terms of the QD's resonance energy spectrum at different thermal biases. In the following we base our discussion on Ref. [19] and use phenomenological sketches of a QD resonance spectrum and Fermi-Dirac distributions in the leads to illustrate how the increase in ΔT can lead to nonlinear effects (Fig. 3). The currents $I_{\varepsilon 1}$ and $I_{\varepsilon 2}$ in Fig. 3b combine to give the overall thermocurrent I_{th} through the QD.

First, we observe that the I_H at which Ith starts to rapidly increase depends on V_G (Fig. 2d). As shown in sketch A in Fig. 3a, this behavior can be understood based on the energy of the QD resonances, ε_1 and ε_2. Until the temperature on the hot side reaches a certain value, there is no net current because the electronic states at energies ε_1 and ε_2 in both leads are equally occupied-

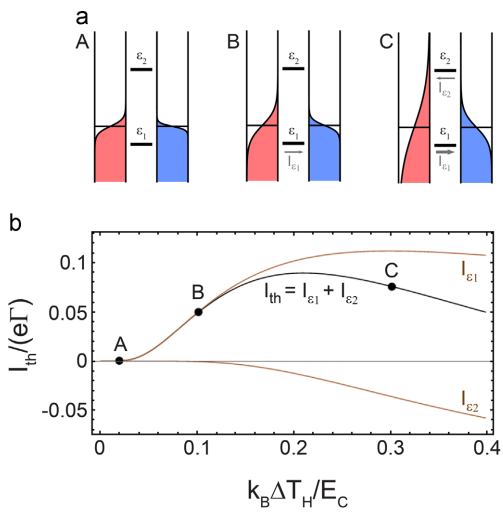


Fig. 3. (a) Schematic representation of electron distribution in source (red) and drain (blue) leads when the thermal bias is (A) $k_B\Delta T_H/E_C = 0.02$, (B) $k_B\Delta T_H/E_C = 0.1$ and (C) $k_B\Delta T_H/E_C = 0.3$. Current direction through resonances of a quantum dot is indicated with arrows. Electron energy increases up the vertical axis. (b) Simulated thermocurrent as a function of thermal bias for the back-gate voltage $e\alpha_G V_G/E_C = 0.24$ (black). Brown curves are thermocurrent contributions through each resonance of the quantum dot. See Fig. 4 for simulation parameters and Section 3.2 for a detailed description. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

either completely full or completely empty. This is reflected in point A in Fig. 3b.

The second interesting experimental feature in Fig. 2d is the nonlinear increase of I_{th}, as a function of thermal bias. Sketch B in Fig. 3a illustrates how increased heating on the source side leads to a misbalance of the electronic state occupancy in the leads at ε_1. This misbalance leads to a net current as indicated by an arrow in the sketch and by point B in Fig. 3b. Thus, the origin of the nonlinear increase in I_{th} is the nonlinear change of the electronic state occupancy in the leads due to heating.

Finally, I_{th} tends to decrease at higher I_H. Ref. [19] predicts such behavior due to an increasing backflow of electrons at large thermal bias values ($\Delta T/T_0 \geq 10$). We believe that the same is true for I_{th} in our experiment, except we expect that we also parasitically heat the drain lead when aiming for high ΔT. Sketch C in Fig. 3a illustrates that the major current contribution, $I_{\varepsilon 1}$, is still provided by the electron transport through ε_1, however, the thermally excited electrons on the source side also leak back through ε_2, thus contributing to the decrease in I_{th}. We note that any decrease of the current through ε_1 in the sketch is, in fact, caused by the overall increase in temperature; e.g. slight heating of the drain. However, the backflow of electrons through ε_2 is caused purely by the thermal bias.

3. Theory

3.1. Model description

We model electron transport through the InAs/InP nanowire by considering a QD which is tunnel-coupled to two electron reservoirs (source and drain leads). Following the experimental setup showed in Fig. 1b the QD is considered in series with a resistive load R to model the input impedance of the current preamplifier. The source and drain leads are characterized by their electrochemical potentials, $u_S = E_F - eV_S$ and $u_D = E_F - eV_D$, where E_F is Fermi energy, and their temperatures, T_S and T_D. Electrons in the leads are assumed to occupy states according to the Fermi-Dirac distribution $f_r(E) = \{1 + \exp[(E - u_r)/(k_B T_r)]\}^{-1}$ and the density of states in the leads is assumed to be a constant. The QD is capacitively coupled to the leads with capacitances C_S and C_D, and to the global back-gate with a capacitance C_G, giving rise to a charging energy $E_C = e^2/(C_S + C_D + C_G)$. In order to model resonance energies we consider a QD in which adding the N_{th} electron changes its state from i to f and that has an electrochemical potential of the form

$$\mu_{fi} = \epsilon_{fi} + (N-1)E_C - e\sum_{r=G,S,D} \alpha_r V_r .$$

Here ϵ_{fi} is energy of the single-electron orbital in which the electron is added and $\alpha_r = C_r/(C_S + C_D + C_G)$ are dimensionless coupling constants. We label the probability of the f^{th} state to be occupied p_f. Steady-state probabilities for each state occupancy can be represented by a vector $\mathbf{P}$ and are found using the Master equation for a stationary case

$$\mathbf{WP} = 0.$$

Here $\mathbf{W}$ is a matrix with elements W_{fi} given by

$$W_{fi} = \begin{cases} \Gamma_{fi}^{S,in} f_S(\mu_{fi}) + \Gamma_{fi}^{D,in} f_D(\mu_{fi}) + \\ \qquad + \Gamma_{fi}^{S,out}[1 - f_S(\mu_{fi})] + \Gamma_{fi}^{D,out}[1 - f_D(\mu_{fi})], & \text{if } i \neq f \\ -\sum_m W_{mf}, & \text{if } i = f \end{cases}$$

where $\Gamma^{S,in}$, $\Gamma^{D,in}$, $\Gamma^{S,out}$ and $\Gamma^{D,out}$ are matrices containing tunnel rates for single electron tunneling *in* or *out* of the QD, involving *source* or *drain* leads. Here non-diagonal matrix elements W_{fi} express physical rates at which the QD changes its state from i to f. Probability normalization requires that the sum of all occupancy probabilities p_f must be 1.

The current I_{SD} through the QD is then found by adding up current contributions from all possible QD states given the calculated steady state occupancies p_f

$$I_{SD} = -e\sum_{i,f} p_f\{\Gamma_{fi}^{S,in} f_S(\mu_{fi}) - \Gamma_{fi}^{S,out}[1 - f_S(\mu_{fi})]\}.$$

In order to calculate the current I_{SD} through the circuit with the QD and the load R in series, a bias value on the drain side V_D is calculated self-consistently using the Ohms law $V_D = I_{SD}R$.

For the purpose of comparing with our experimental results it is sufficient to consider a QD with only one single electron orbital, in which N can take values 0, 1 or 2. Including electron spin this gives four possible QD states i, f={ 0, ↑, ↓, ↑↓ }. In this case, the phenomenological resonance energies ε_1 and ε_2 discussed in the experimental section (Fig. 3) thus correspond to the electrochemical potentials $\mu_{\sigma 0} = \varepsilon_1$ and $\mu_{\uparrow\downarrow\sigma} = \varepsilon_2$, with σ=↑, ↓. For qualitative comparison with experiment we consider the tunnel-barriers to be identical and characterized by a constant tunnel rate Γ.

3.2. Simulation results

We now calculate the thermocurrent as a function of temperature in source and drain leads. Since in our experiment the source lead is heated, we label the source temperature $T_S = T_H = T_0 + \Delta T_H$ and the drain temperature $T_D = T_C = T_0 + \Delta T_C$. In simulations the base temperature T_0 is chosen such that $k_B T_0/E_C = 0.01$, which is close to the experimental value. Because in the experiments the drain lead is also expected to be somewhat heated we assume $\Delta T_C = \Delta T_H/3$. The ratio between ΔT_H and ΔT_C is chosen to obtain a qualitative agreement with the experimental data, but the precise value is not important for the discussed physics.

In Fig. 4 we sum up our thermocurrent simulation results. Thermocurrent as a function of the back-gate voltage for different thermal bias values is shown in Fig. 4a (compare with the corresponding experimental data in Fig. 2b). Similarly, we plot the simulated thermocurrent as a function of the thermal bias for different back-gate voltage values in Fig. 4b. The dimensionless range of thermal bias shown is chosen based on the similarity to Fig. 2d. Finally, the color plot in Fig. 4c is produced using the ranges of the electrochemical potential and the thermal bias used in Fig. 4a and b, and closely matches the experimental result shown in Fig. 2c.

According to our simulations, the source–drain bias V_{SD} that develops across the QD due to the series load at peak thermocurrents is estimated to be below $\pm 0.04\ E_C/e$ and therefore does not significantly influence the behavior of the thermocurrent. Note that it is very challenging to measure the temperature in the leads leading up to the QD directly and this was not attempted in the experiment. However, given the qualitative agreement between the experimental thermocurrent data in Fig. 2 and the simulated thermocurrent in Fig. 4, one can conclude that the relation between I_H and ΔT must be close to linear. Moreover, the agreement also suggests that 1 mA of I_H gives rise to a thermal bias ΔT of several Kelvin between the source and drain leads.

4. Conclusions

In summary, we have reported measurements of a strongly nonlinear thermocurrent in a QD. By comparing our measurements to simulation results, we show that the nonlinear behavior can be fully explained in terms of the QD's energy-dependent transport properties [19]. This is in contrast to earlier experiments [18] where this behavior was masked by effects that can also be explained by

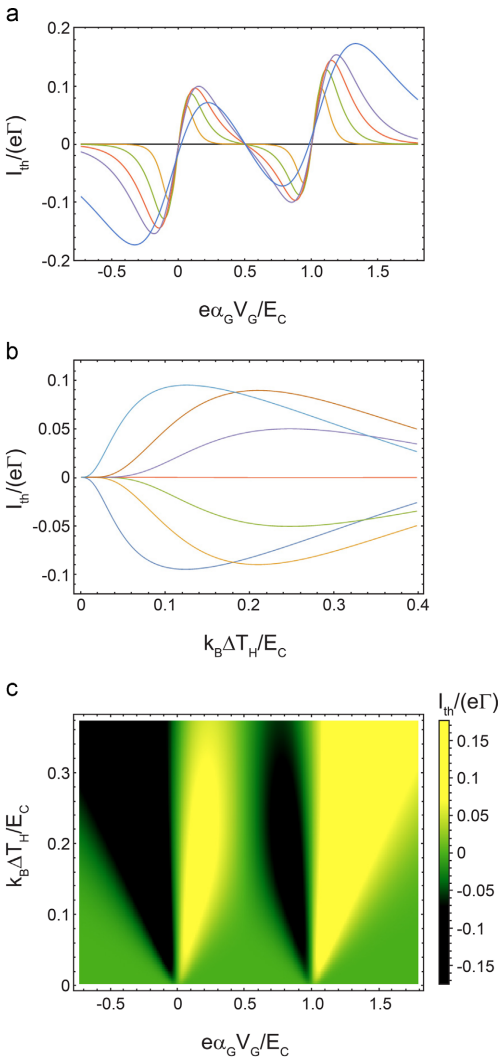


Fig. 4. (a) Simulated thermocurrent as a function of back-gate voltage for different thermal biases $k_B\Delta T_H/E_C = (0,\ 0.04,\ 0.08,\ 0.12,\ 0.16,\ 0.32)$. (b) Simulated thermocurrent as a function of the thermal bias for several back-gate voltage values $e\alpha_C V_C/E_C = (0.11,\ 0.24,\ 0.37,\ 0.50,\ 0.63,\ 0.76,\ 0.89)$. (c) Simulated thermocurrent (color) as a function of both, back-gate voltage and thermal bias. Other parameters: $\Gamma = 5\,\mathrm{GHz}$, $R = 1\,\mathrm{M\Omega}$, $T_0 = 0.01E_C$.

the overall heating of the device. Our results were enabled by use of a novel heating technique [23] that allows the application of very large ΔT across a nanoscale device with minimal overall heating of the sample space, even at low temperatures. The ability demonstrated here opens a wide range of quantum thermoelectric experiments in mesoscopic systems.

Acknowledgments

We thank Sebastian Lehmann for the TEM image in Fig. 1a. This work was supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7-People-2013-ITN) under REA Grant agreement no. 608153, by the Swedish Energy Agency (Project P38331-1), by the Swedish Research Council (Project 621-2012-5122) and by NanoLund.

References

[1] H. van Houten, C.W.J. Beenakker, A.A.M. Staring, Single Charge Tunneling, in: H. Grabert, M.H. Devoret (Eds.), Plenum, New York, 1992.
[2] C.W.J. Beenakker, A.A.M. Staring, Phys. Rev. B 46 (15) (1992) 9667–9676.
[3] A.A.M. Staring, L.W. Molenkamp, B.W. Alphenaar, H. van Houten, O.J.A. Buyk, M.A. A. Mabesoone, C.W.J. Beenakker, C.T. Foxon, EPL (Europhys. Lett.) 22 (1) (1993) 57.
[4] T.E. Humphrey, R. Newbury, R.P. Taylor, H. Linke, Phys. Rev. Lett. 89 (11) (2002) 116801.
[5] H.L. Edwards, Q. Niu, A.L. de Lozanne, Appl. Phys. Lett. 63 (13) (1993) 1815–1817.
[6] M.F. O'Dwyer, T.E. Humphrey, H. Linke, Nanotechnology 17 (11) (2006) S338–S343.
[7] M. Esposito, K. Lindenberg, C. Van den Broeck, Europhys. Lett. 85 (6) (2009) 60010.
[8] N. Nakpathomkun, H.Q. Xu, H. Linke, Phys. Rev. B 82 (23) (2010) 235428.
[9] A.N. Jordan, B. Sothmann, R. Sanchez, M. Büttiker, Phys. Rev. B 87 (7) (2013) 075312.
[10] X. Zianni, J. Electron. Mater. 39 (9) (2010) 1996–2001.
[11] J. Wang, L. Wan, Y. Wei, Y. Xing, J. Wang, Mod. Phys. Lett. B 20 (05) (2006) 215–223.
[12] D. Sánchez, R. López, Phys. Rev. Lett. 110 (2) (2013) 026804.
[13] J. Meair, P. Jacquod, J. Phys.: Condens. Matter 25 (8) (2013) 082201.
[14] R.S. Whitney, Phys. Rev. B 87 (11) (2013) 115404.
[15] D. Boese, R. Fazio, EPL (Europhys. Lett.) 56 (4) (2001) 576.
[16] J. Azema, A.M. Daré, S. Schäfer, P. Lombardo, Phys. Rev. B 86 (7) (2012) 075303.
[17] N. Nakpathomkun, Thermoelectric properties of quantum dots and other low-dimensional systems (Ph.D. thesis), University of, Oregon, United States, 2010.
[18] S.F. Svensson, E.A. Hoffmann, N. Nakpathomkun, P.M. Wu, H.Q. Xu, H.A. Nilsson, D. Sánchez, V. Kashcheyevs, H. Linke, New J. Phys. 15 (10) (2013) 105011.
[19] M.A. Sierra, D. Sánchez, Phys. Rev. B 90 (11) (2014) 115313.
[20] A.G. Pogosov, M.V. Budantsev, R.A. Lavrov, A.E. Plotnikov, A.K. Bakarov, A. I. Toropov, J.C. Portal, JETP Lett. 83 (3) (2006) 122–126.
[21] E.A. Hoffmann, The thermoelectric efficiency of quantum dots in InAs/InP nanowires (Ph.D. thesis), University of, Oregon, United States, 2009.
[22] P. Reddy, S.-Y. Jang, R.A. Segalman, A. Majumdar, Science 315 (5818) (2007) 1568–1571.
[23] J.G. Gluschke, S.F. Svensson, C. Thelander, H. Linke, Nanotechnology 25 (38) (2014) 385704.
[24] L.E. Fröberg, B.A. Wacaser, J.B. Wagner, S. Jeppesen, B.J. Ohlsson, K. Deppert, L. Samuelson, Nano Lett. 8 (11) (2008) 3815–3818.
[25] A.I. Persson, L.E. Fröberg, S. Jeppesen, M.T. Björk, L. Samuelson, J. Appl. Phys. 101 (3) (2007) 034313.
[26] D.B. Suyatin, C. Thelander, M.T. Björk, I. Maximov, L. Samuelson, Nano-technology 18 (10) (2007) 105307.
[27] S.F. Svensson, A.I. Persson, E.A. Hoffmann, N. Nakpathomkun, H.A. Nilsson, H. Q. Xu, L. Samuelson, H. Linke, New J. Phys. 14 (3) (2012) 033041.

Paper II

nature nanotechnology

LETTERS

https://doi.org/10.1038/s41565-018-0200-5

A quantum-dot heat engine operating close to the thermodynamic efficiency limits

Martin Josefsson[1,2], Artis Svilans[1,2], Adam M. Burke[1], Eric A. Hoffmann[1], Sofia Fahlvik[1], Claes Thelander[1], Martin Leijnse[1] and Heiner Linke[1]*

Cyclical heat engines are a paradigm of classical thermodynamics, but are impractical for miniaturization because they rely on moving parts. A more recent concept is particle-exchange (PE) heat engines, which uses energy filtering to control a thermally driven particle flow between two heat reservoirs[1,2]. As they do not require moving parts and can be realized in solid-state materials, they are suitable for low-power applications and miniaturization. It was predicted that PE engines could reach the same thermodynamically ideal efficiency limits as those accessible to cyclical engines[3–6], but this prediction has not been verified experimentally. Here, we demonstrate a PE heat engine based on a quantum dot (QD) embedded into a semiconductor nanowire. We 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. We find that at the maximum power conditions, η is in agreement with the Curzon–Ahlborn efficiency[6–9] and that the overall maximum η is in excess of 70% of the Carnot efficiency while maintaining a finite power output. Our results demonstrate that thermoelectric power conversion can, in principle, be achieved close to the thermodynamic limits, with direct relevance for future hot-carrier photovoltaics[10], on-chip coolers or energy harvesters for quantum technologies.

Traditional closed-cycle heat engines are based on an enclosed working medium that exchanges heat, but not particles, with hot and cold thermal reservoirs at temperatures T_H and T_C, respectively. The engines' thermal efficiency is bounded by the fundamental Carnot limit $\eta_C = (T_H - T_C)/T_H$ (ref. [11]). In practice, however, the goal is usually to optimize the efficiency at a maximum power: that is, to operate near the lower Curzon–Ahlborn efficiency $\eta_{CA} = 1 - \sqrt{T_C/T_H}$ (ref. [7]). The best Stirling engines, for example, reach thermal efficiencies slightly above $0.5\eta_C$ (ref. [12]), comparable to the η_{CA} (Supplementary Section A).

A drawback of cyclical engines is that they require moving parts, which severely limits low-power applications, for example, in sensors or wearables. By contrast, PE heat engines[2] require no moving elements as they operate by exchanging particles (for example, photons[1] or electrons[4]) between two heat reservoirs. Theory predicts that PE heat engines can be operated near η_C provided that, first, the energy at which particles are exchanged between reservoirs is limited to an energy band much narrower than kT_H (refs [3,5,13]) and, second, that said energy is adjusted such that the particle transfer produces no entropy[1–4,8]. These conditions describe an ideal solid-state thermoelectric system[3,5].

One way to achieve the required energy filtering in the solid state is to use a QD that is tunnel-coupled to two electron reservoirs[5,14]. Single-electron orbital states that act as energetically sharp transmission channels (resonances) for electrons at energy ε_0 can be used as energy filters. According to theory, by adjusting ε_0 one can operate the system either near η_C (refs [4,14]) or near η_{CA} (ref. [15]). Experimentally, there has been significant progress in the study of the thermoelectric properties of QDs[16,17] and QD-based solid-state cooling devices[18]. However, predictions about the achievable efficiencies in PE heat engines have not been experimentally investigated before as it is a challenge to fulfil all the requirements for quantitative tests simultaneously: an accurate reservoir thermometry, tunable and electrically non-invasive reservoir heating and a QD that approximates an ideal energy filter.

In this work we explore whether it is possible to reach η_C and η_{CA} in PE heat engines based on QDs formed by thin InP segments embedded into InAs nanowires (Fig. 1a), as proposed previously[14]. This system offers small and electrostatically stable QDs defined with atomic precision[19] and makes use of well-established device-fabrication techniques. The energy width of the resonance is determined by the tunnel rates Γ across the InP segments, and an electrostatic gate can be used to control the resonance energy.

The operation principle of our PE heat engine is illustrated in Fig. 1b. The resonance energy ε_0 is positioned relative to the chemical potentials of the electron reservoirs, μ_C and μ_H, such that electronic state occupancy at ε_0 is higher in the hot (red) reservoir than in the cold (blue) reservoir. In this configuration, the temperature difference $\Delta T = (T_H - T_C)$ can drive an electric current I against an electrical potential difference $V = (\mu_C - \mu_H)/e$ (ref. [4]). In the limit $\hbar\Gamma \to 0$, electrons are transferred only at ε_0. Each transferred electron then produces electric work (eV) at the cost of removing heat $Q_H = \varepsilon_0 - \mu_H$ from the hot reservoir and depositing $Q_C = \varepsilon_0 - \mu_C$ in the cold reservoir. For finite Γ, this coupling between charge and heat is no longer exact because of a finite resonance width and effects such as co-tunnelling[20], thus resulting in an increased heat flow. Also, the existence of a well-defined resonance energy ε_0 requires the single-particle level spacing to be much larger than kT_H to avoid transport through the excited states of the QD.

In the presence of a load R in series with the QD (Fig. 1c) and zero external bias V_{ext}, the circuit self-consistently satisfies the relation $V = -I_{th}R$, where $I_{th} = I(V_{ext} = 0)$ is the thermocurrent. The thermoelectrically produced power in the steady state is $P_{th} = -I_{th}V = I_{th}^2R$ and the electronic efficiency is $\eta = P_{th}/J_Q$, where J_Q is the electronic heat flow that leaves the hot reservoir through the QD. We emphasize that P_{th} and η depend on R, which can thus be used to optimize either P_{th} or η.

[1]NanoLund and Solid State Physics, Lund University, Lund, Sweden. [2]These authors contributed equally: Martin Josefsson, Artis Svilans.
*e-mail: heiner.linke@ftf.lth.se

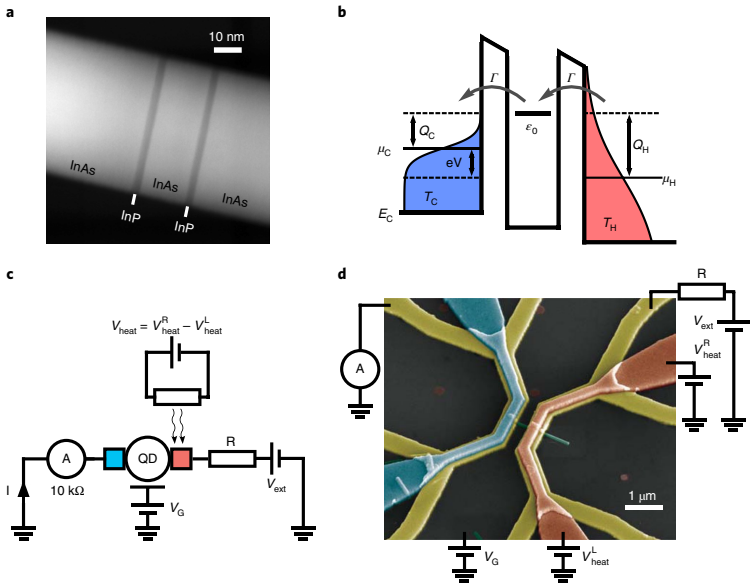


Fig. 1 | Experimental device and its operational principle. a, Scanning transmission electron microscopy with a high-angle annular dark field (STEM-HAADF) image of a representative InAs/InP/InAs/InP/InAs heterostructured nanowire from the same growth as the nanowires in the devices. Methods gives details on the nanowire dimensions. **b**, Illustration of a QD-based PE heat engine with resonance energy ε_0. The QD is tunnel-coupled (rate Γ) to hot and cold electron reservoirs with Fermi distributions characterized by T_H, μ_H and T_C, μ_C, respectively. An electron that traverses the QD at energy ε_0 removes heat Q_H from the hot reservoir, converts part of it into useful work, eV, and deposits the remaining part as heat, Q_C, in the cold reservoir. **c**, The circuit used for thermoelectric characterization features a tunable resistor R (this also includes a 10 kΩ input impedance of the current preamplifier and a 4.5 kΩ resistance of the RC filters (not shown)), a current preamplifier and a voltage source V_{ext}. A separate voltage source, $V_{heat} = V^L_{heat} - V^R_{heat}$ is applied in a push–pull configuration for running a current through a heater that is electrically decoupled from the hot electron reservoir. **d**, False-coloured SEM image of a nominally identical device to the one used in the experiment. Metallic leads (yellow) make contact to the nanowire (green). Heaters (blue and red) 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. The resulting $\Delta T = T_H - T_C$ is set by the temperature profile of the phonon bath.

Our experimental device consisted of an InAs/InP nanowire QD in contact with metallic leads, as shown in Fig. 1d. We used top heaters[21] for the effective thermal biasing of the QD. The differential conductance $dI/dV_{ext} = G$ of the QD as a function of V_G and V_{ext} shows that the QD had a well-defined resonance located at $V_G \approx 0.13$ V, as indicated by the intersecting G lines at $V_{ext} = 0$ V (Fig. 2a). This resonance was separated from others by the QD's charging energy of 4.9 meV, which is much larger than $kT_H = 0.17$ meV at the highest electronic temperature used in the experiment, $T = 2$ K. No transport via excited states is evident (Fig. 2a). All the results discussed in the following were obtained using only this resonance as the energy filter. Data from additional devices are presented in Supplementary Section B.

To estimate the engine's efficiency, we calculated the heat flow J_Q based on experimentally determined parameters. This task required a theoretical description that includes full non-linear effects, large electron–electron interactions (Coulomb blockade) and goes beyond the sequential-tunnelling approximation generally used for small Γ. Our theoretical approach used the real-time diagrammatic (RTD) technique to expand the Liouville equation for the density matrix in Γ and solved the generalized master equations[22–24] that resulted for a single spin-degenerate energy level. We included all contributions to the current up to order Γ^2, which includes co-tunnelling, level broadening and energy renormalization processes. We accounted for R by solving the self-consistent equation for V across the QD (Supplementary Section C).

Our analysis of the PE heat-engine performance was based primarily on current measurements that allowed for the accurate determination of Γ, T_H and T_C (Fig. 2b–d). We determined Γ by fitting the RTD theory to the measured peak in $G(V_{ext} = 0)$ as a function of V_G with $\Delta T = 0$ (Fig. 2b). In total, four independent measurements were performed at elevated temperatures of around 0.5 and 1.0 K to ensure that $kT \gg \hbar\Gamma$, a required condition for the validity of our theory. We assumed equal tunnelling rates across both InP segments, which yielded Γ values in the range 8.3–9.3 GHz (Supplementary Sections D and E give details). For further analysis, we used the average value $\Gamma = 8.9$ GHz ($\hbar\Gamma = 5.9$ μeV).

We determined T_H and T_C by measuring $I_{th}(V_G)$ as a response to an applied heater bias V_{heat} (Fig. 2c). The amplitude of I_{th} is sensitive to T_H and T_C, and characteristically reverses direction at the resonance ($V_G \approx 0.13$ V)[17]. Using T_H and T_C as free parameters, we found excellent fits of the RTD theory (black lines in Fig. 2c) to the experimental data points, and observed an approximately linear increase of T_H and T_C with V_{heat} (Fig. 2d) (Supplementary Section D). Note that the positive and negative I_{th} peak amplitudes are not the same, which is correctly reproduced by our theory. This asymmetry is due to electron–electron interactions in the spin-degenerate QD orbital at ε_0.

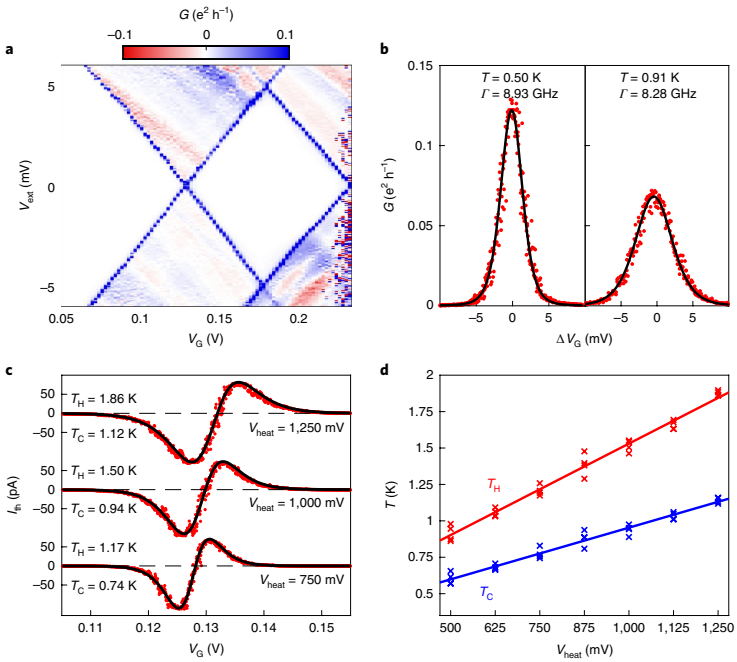


Fig. 2 | Electric and thermoelectric characterization of the device. a, Colour plot of differential conductance G of the QD measured at $T < 200$ mK and $V_{heat} = 0$ as a function of V_{ext} and V_G. **b**, $G(V_{ext} = 0)$ as a function of ΔV_G (defined relative to the resonance at $V_G \approx 0.13$ V) at two elevated temperatures close to 0.5 and 1.0 K. Solid lines are fits of the RTD theory with free parameters $T = T_C = T_H$ and Γ, which yields the fit values shown. **c**, Measured I_{th} as a function of V_G for different values of V_{heat}. Solid curves are fits of the RTD theory with temperatures T_C and T_H as the fitting parameters and $\Gamma = 8.9$ GHz. The shift in the I_{th} reversal point is caused by a parasitic gating effect (the first part of Supplementary Section D gives more information). **d**, Results from all the temperature fits. $I_{th}(V_G)$ is measured four times for each V_{heat}, and T_H and T_C are fitted for each individual measurement. All the data shown in this figure are taken with $R = 14.5$ kΩ (accounted for in our theoretical approach).

Unlike the measured P_{th}, the calculated heat flow J_Q does not go to zero at the I_{th} reversal point (Fig. 3a,b). This is because the tunnelling effects of second order in Γ effectively decouple the charge current from the heat current, which also reduces the maximum achievable efficiency for the PE heat engine. Such effects could be pictured as contrapropagating charges at slightly different energies that results in a J_Q with no net I_{th}.

By varying ε_0, we were able to optimize either P_{th} or η at each given load R (Fig. 3c). The maximum η was achieved between the peak and the reversal point of I_{th} (black symbols in Fig. 3). However, fluctuations in I_{th} led to a significant scatter of the data points in this range and we therefore focused on the maximum P_{th} at each R, denoted P_{max}, for which the signal-to-noise ratio was better. We found that P_{max} peaked in the R range between 0.7 and 1.5 MΩ, depending on the V_{heat} used, but this value will, in general, also depend on Γ (Fig. 4a). We note that no simple analytic expression exists for the optimal R for the maximum power production[15].

We denote the estimated η at P_{max} as $\eta_{P_{max}}$. We found $\eta_{P_{max}} \approx \eta_{CA}$ for the R that produces the overall maximum P_{max} (Fig. 4b), which confirms theoretical predictions[9,13,15]. Optimizing R for maximal $\eta_{P_{max}}$ yielded efficiencies in excess of $0.7\eta_C$ but still maintained a finite steady-state power output, roughly equal to one-half of the overall maximum power for the same V_{heat} (Fig. 4b). Deviations between the data points and the RTD theory curves for $\eta_{P_{max}}$ in Fig. 4b originate from the measured P_{max} being slightly higher than that predicted in theory. Note that the sign and magnitude of the underlying deviations in the I_{th} peak values are consistent with a small thermoelectric effect in the contact leads, which our model does not account for.

Our results demonstrate an electronic efficiency at a finite power output in excess of 70% of the Carnot limit, comparable to traditional cyclical heat engines (Supplementary Section A), and they confirm that QDs can be operated close to the Curzon–Ahlborn efficiency at maximum power. We achieved this by combining the use of high-quality, epitaxially defined QDs with a novel technique for non-invasive thermal biasing in immediate proximity to the QD[21], and by directly measuring the power produced by the QD as a function of an external load R. To determine the electronic heat flow, and thus the efficiency, we used a theory that includes electron–electron interactions (Coulomb blockade), full non-linear effects and higher-order tunnelling. The reliability of the model is validated by its agreement with the experimental data (Supplementary Section D.III). In future work it would be desirable to measure the heat flow directly—a difficult task in a QD heat engine under operating conditions that is subject to both electrical and thermal biases at the same time.

Our experiment approximates 'the best thermoelectric'[3] by realizing a system in which particle exchange between heat baths takes place only within a very narrow energy window. Our analysis is limited to electronic thermal reservoirs and does not consider phonon-mediated heat flow, which is a parasitic effect that reduces

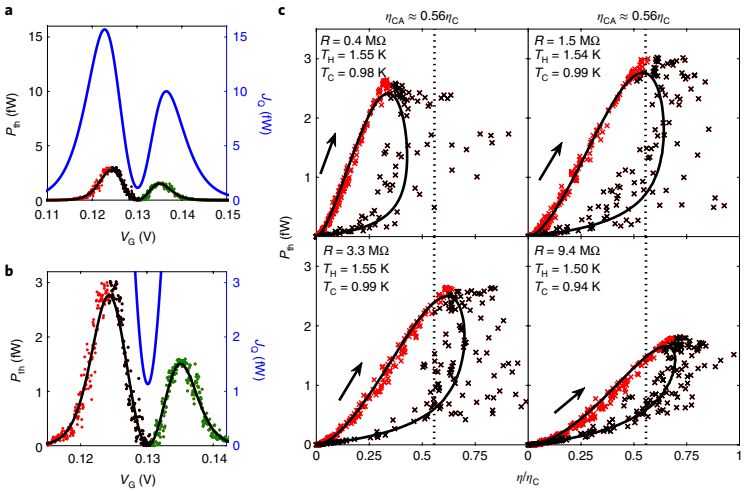


Fig. 3 | Thermoelectric performance of the PE engine. a,**b**, Data points are the measured P_{th} at V_{heat} = 1,000 mV and R = 1.5 MΩ. The solid lines are the RTD calculations of P_{th} (black) and J_Q (blue) using T_C = 0.99 K, T_H = 1.54 K and Γ = 8.9 GHz, as obtained from the measurements of I_{th} and $G(V_{ext} = 0\,V)$ (Supplementary Section D gives more details). Data points are coloured to identify the different V_G ranges. **c**, Parametric plot of P_{th} and $\eta = P_{th}/J_Q$ of the data in **a** and **b** with R = 1.5 MΩ as well as for three other loads, R = 0.4, 3.3 and 9.4 MΩ, as indicated. The same colour code as in **a** and **b** is used (the data from the smaller of the two peaks in P_{th}, marked with green dots in **a** and **b**, are not plotted). Red and black markers identify data points from the corresponding V_G ranges indicated in **a** and **b**. Data points are based on the measured values of P_{th} and the calculated J_Q using the experimentally determined parameters. The solid lines are based purely on RTD calculations using the same parameters. The arrow indicates the direction for increasing V_G. The dashed line is the theoretical prediction for $\eta_{CA} = 0.56\eta_C$, which aligns well with the maximum power P_{th} observed at R = 1.5 MΩ. The spread in the data points marked in black is due to fluctuations in the measured current when approaching the I_{th} reversal point in V_G. Note that η_{CA} is expected to be an approximate upper limit for η at the overall maximal power of the engine, which for the V_{heat} = 1,000 mV used in this figure is obtained for $R \approx 1.5$ MΩ (top-right panel). For other R values for which lower powers are achieved, the maximum η can be smaller (top-left panel) or larger (bottom panels) than η_{CA}.

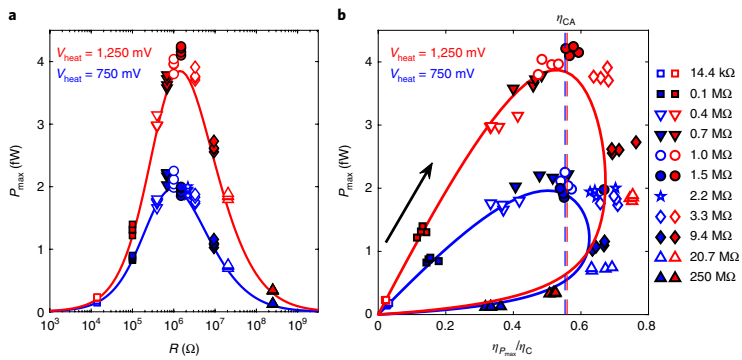


Fig. 4 | Device operation at maximum power. a, Power P_{max} (symbols), maximized with respect to V_G, compared with the theoretical predictions (solid lines) as a function of the R for two values of V_{heat} as indicated in the figure. The theoretical parameters are Γ = 8.9 GHz and T_C = 0.69 K, T_H = 1.02 K for V_{heat} = 750 mV and T_C = 1.13 K, T_H = 1.83 K for V_{heat} = 1,250 mV as obtained from the fits of the RTD theory to I_{th} and $G(V_{ext} = 0\,V)$. **b**, The same data points for the measured power, P_{max}, as in **a**, plotted parametrically for the same range of R values, as a function of the associated $\eta_{P_{max}}$. The solid lines are the theoretical model's predictions. The arrow indicates the direction of increasing R and the vertical dashed lines indicate the corresponding η_{CA}, both at $\eta_{CA} = 0.56\eta_C$.

the efficiency[25]. Nevertheless, our results are directly applicable to emerging non-equilibrium devices, such as hot-carrier solar cells, that seek to harvest thermal energy stored by photogenerated carriers that are out of thermal equilibrium with phonons[10,26].

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41565-018-0200-5.

Received: 2 October 2017; Accepted: 13 June 2018;
Published online: 16 July 2018

References

1. Scovil, H. & Schulz-DuBois, E. Three-level masers as heat engines. *Phys. Rev. Lett.* **2**, 262–263 (1959).
2. Humphrey, T. E. & Linke, H. Quantum, cyclic, and particle-exchange heat engines. *Physica E* **29**, 390–398 (2005).
3. Mahan, G. & Sofo, J. The best thermoelectric. *Proceed. Natl Acad. Sci. USA* **93**, 7436–7439 (1996).
4. Humphrey, T. E., Newbury, R., Taylor, R. P. & Linke, H. Reversible quantum Brownian heat engines for electrons. *Phys. Rev. Lett.* **89**, 116801 (2002).
5. Humphrey, T. E. & Linke, H. Reversible thermoelectric nanomaterials. *Phys. Rev. Lett.* **94**, 096601 (2005).
6. Van den Broeck, C. Thermodynamic efficiency at maximum power. *Phys. Rev. Lett.* **95**, 190602 (2005).
7. Curzon, F. & Ahlborn, B. Efficiency of a Carnot engine at maximum power output. *Am. J. Phys.* **43**, 22–24 (1975).
8. Esposito, M., Lindenberg, K. & Van den Broeck, C. Universality of efficiency at maximum power. *Phys. Rev. Lett.* **102**, 130602 (2009).
9. Esposito, M., Kawai, R., Lindenberg, K. & Van den Broeck, C. Efficiency at maximum power of low-dissipation Carnot engines. *Phys. Rev. Lett.* **105**, 150603 (2010).
10. Limpert, S., Bremner, S. & Linke, H. Reversible electron–hole separation in a hot carrier solar cell. *New J. Phys.* **17**, 095004 (2015).
11. Callen, H. B. *Thermodynamics and an Introduction to Thermostatistics* 2nd edn, Ch. 4 (John Wiley & Sons, New York, 1985).
12. Wong, W. A., Wilson, S., Collins, J. & Wilson, K. *Advanced Stirling Converter (ASC) Technology Maturation* Report NASA/TM-2016-218908 (National Aeronautics and Space Agency, 2016).
13. Nakpathomkun, N., Xu, H. Q. & Linke, H. Thermoelectric efficiency at maximum power in low-dimensional systems. *Phys. Rev. B* **82**, 235428 (2010).
14. O'Dwyer, M. F., Humphrey, T. E. & Linke, H. Concept study for a high-efficiency nanowire based thermoelectric. *Nanotechnology* **17**, S338–S343 (2006).
15. Esposito, M., Lindenberg, K. & Van den Broeck, C. Thermoelectric efficiency at maximum power in a quantum dot. *Europhys. Lett.* **85**, 60010 (2009).
16. Staring, A. A. M. et al. Coulomb-blockade oscillations in the thermopower of a quantum dot. *Europhys. Lett.* **22**, 57 (1993).
17. Svilans, A., Leijnse, M. & Linke, H. Experiments on the thermoelectric properties of quantum dots. *C. R. Phys.* **17**, 1096–1108 (2016).
18. Prance, J. R. et al. Electronic refrigeration of a two-dimensional electron gas. *Phys. Rev. Lett.* **102**, 146602 (2009).
19. Björk, M. T. et al. Few-electron quantum dots in nanowires. *Nano Lett.* **4**, 1621 (2004).
20. Turek, M. & Matveev, K. Cotunneling thermopower of single electron transistors. *Phys. Rev. B* **65**, 115332 (2002).
21. Gluschke, J. G., Svensson, S. F., Thelander, C. & Linke, H. Fully tunable, non-invasive thermal biasing of gated nanostructures suitable for low-temperature studies. *Nanotechnology* **25**, 385704 (2014).
22. König, J., Schoeller, H. & Schön, G. Cotunneling at resonance for the single-electron transistor. *Phys. Rev. Lett.* **78**, 4482 (1997).
23. Leijnse, M. & Wegewijs, M. Kinetic equations for transport through single-molecule transistors. *Phys. Rev. B* **78**, 235424 (2008).
24. Gergs, N. M., Hörig, C. B., Wegewijs, M. R. & Schuricht, D. Charge fluctuations in nonlinear heat transport. *Phys. Rev. B* **91**, 201107 (2015).
25. Dresselhaus, M. S. et al. New directions for low-dimensional thermoelectric materials. *Adv. Mat.* **19**, 1043–1053 (2007).
26. Conibeer, G., Jiang, C. W., Green, M., Harder, N. & Straub, A. Selective energy contacts for potential application to hot carrier PV cells. *Proc. 3rd World Conf. Photovolt. En. Conv.* **3**, 2730–2733 (2003).

Acknowledgements

We thank S. Lehmann for the structural imaging of the nanowires used in this study. We acknowledge financial support by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7-People-2013-ITN) under REA grant agreement no. 608153 (PhD4Energy), by the Swedish Energy Agency (project P38331-1), by the Swedish Research Council (projects 621-2012-5122, 2014-5490, 2015-00619 and 2016-03824), by the Knut and Alice Wallenberg Foundation (project 2016.0089), Marie Sklodowska Curie Actions, Cofund, Project INCA 600398 and by NanoLund. The computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at LUNARC.

Author contributions

H.L. and M.L. designed and guided the study. E.A.H. and S.F. performed preliminary experiments. S.F. grew the nanowires. A.S., A.M.B. and C.T. designed and fabricated the devices and carried out the experiments. M.J. and M.L. performed the theoretical calculations. M.J. and A.S. analysed the data. All the authors contributed to writing and editing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41565-018-0200-5.

Correspondence and requests for materials should be addressed to H.L.

Methods

Device specifications. The QDs in this study are defined in axially heterostructured InAs/InP nanowires[19] (Fig. 1b) grown by chemical beam epitaxy using gold nanoparticles as catalysts[27,28]. From the analyses of the transmission electron microscopy images of 11 nanowires from the same growth we found that, on average, the diameter of the nanowires was 60 nm, the thickness of the thin InP segments that defined the tunnel junctions was 4 ± 3 nm and the length of the InAs QD segment was 17 ± 1.5 nm. The outer InAs segments usable for contacting the nanowire were 350 ± 70 nm and 265 ± 60 nm long. The device fabrication procedure was identical to that reported previously[21], with the only difference being that Ti (instead of Ni) was used for the electrode adhesion layers. A key element in this architecture is the so-called top heater[21], that is, heaters fabricated directly on top of the contact leads rather than next to them or using the contact leads themselves as heaters[17]. This geometry allowed us to combine two important features of nanoscale heaters—a higher thermal bias ΔT with little overall heating of the device and the cryostat, and an electrically non-invasive thermal biasing that allows for the easy tuning of ΔT. As I_{th} roughly scales with ΔT, a large ΔT improves the signal-to-noise ratio of I_{th}.

Measurements. The experiment was carried out in a Triton 200 dilution refrigerator with resistor–capacitor (RC) cold-filtering at 3.5 K and a cut-off frequency of 300 Hz. All the measurements were performed with d.c. using Yokogawa 7651 voltage sources for the electrical biasing of the QD contacts, the top heater and the global back gate. A Femto DLPCA-200 low-noise current preamplifier with an input impedance of 10 kΩ (at the gain mode of 1 nA V^{-1}) was used to measure the current through the QD. Due to changes in the virtual ground potential of the DLPCA-200 (of the order of 10 μV), whenever the measurement circuit was reconfigured (for example, when changing R and/or grounding and ungrounding the sample) a more accurate value of the V_{ext} zero bias point was determined using $I(V_G)$ of the QD's conductance peak as a probe. We used a Femto DLPVA voltage preamplifier to record voltage V (voltage across the QD with the RC filter in series) simultaneously with all the I_{th} measurements. This allowed us to characterize drifts in the applied V_{ext} that were smaller than ± 1 μV over a period of a single I_{th} measurement trace. Before and after the I_{th} measurements at each V_{heat} (at every R), four additional measurements of $V(V_G)$ and $I(V_G)$ with $V_{heat} = 0$ V were performed. This allowed us to ensure that the overall V_{ext} drift was less than ± 2 μV within the measurement sessions at each R. The $V(V_G)$ data (recorded simultaneously with all the $I_{th}(V_G)$ measurements) also gave us a consistent set of redundant data with higher noise levels. Due to the filter resistance (4.5 kΩ) of the refrigerator measurement lines, electrical biasing of the top heater was done by setting the two potentials V^{L}_{heat} and V^{R}_{heat} (see Fig. 1d) on both ends such that the potential of the heater at the device level remained closer to the ground potential to avoid unwanted gating of the QD (the first part of Supplementary Section D gives more details). The heater current is approximately $(V^{L}_{heat} - V^{R}_{heat})/4.5\,\text{k}\Omega$.

Data availability. The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

References

27. Persson, A. I., Fröberg, L. E., Jeppesen, S., Björk, M. T. & Samuelson, L. Surface diffusion effects on growth of nanowires by chemical beam epitaxy. *J. Appl. Phys.* **101**, 034313 (2007).
28. Fröberg, L. E. et al. Transients in the formation of nanowire heterostructures. *Nano Lett.* **8**, 3815–3818 (2008).

nature nanotechnology

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41565-018-0200-5

In the format provided by the authors and unedited.

A quantum-dot heat engine operating close to the thermodynamic efficiency limits

Martin Josefsson[1,2], Artis Svilans[1,2], Adam M. Burke[1], Eric A. Hoffmann[1], Sofia Fahlvik[1], Claes Thelander[1], Martin Leijnse[1] and Heiner Linke[1]*

[1]NanoLund and Solid State Physics, Lund University, Lund, Sweden. [2]These authors contributed equally: Martin Josefsson, Artis Svilans.
*e-mail: heiner.linke@ftf.lth.se

A quantum-dot heat engine operating close to the thermodynamic efficiency limits

Supplementary information

Martin Josefsson*, Artis Svilans*, Adam M. Burke, Eric A. Hoffmann, Sofia Fahlvik, Claes Thelander, Martin Leijnse and Heiner Linke+

A. Heat engines at various scales

The key results from our work are that (i) we achieve a thermal-to-electric energy conversion efficiency (for electrons) in excess of 70% of the Carnot efficiency, and (ii) an efficiency at maximum power of about the Curzon-Ahlborn efficiency. These are the first tests of the performance limit of particle-exchange heat engines. To place our results into context, in this section we offer an overview of the efficiencies achieved in other systems.

Power plants. As conveniently summarized in a table by M. Esposito *et al.* [1] efficiencies for large scale power plants built already decades ago achieved overall efficiencies in the 40% range and correspond to up to 70% of Carnot efficiency. Currently the most efficient heat engine technology is realized in so called combined cycle power plants where the waste heat, produced as a by-product of the gas turbine burning gas at high temperature, is captured and used for boiling water and running another steam turbine. According to a recent review [2] such plants can reach up to 61% overall efficiency and with a simple estimated gas entry temperature (also from Ref. 2) of 1600 °C and ambient temperature of 20 °C, this gives a Carnot efficiency of 84.4 % and the plant efficiency at 72.3% of the Carnot efficiency.

Stirling engines. An example of recent developments in Stirling engine technology is NASA's Advanced Stirling Radioisotope Generator program where Advanced Stirling Convertors are developed for integration with General Purpose Heat Sources for generation of electric power [3]. The report (Ref. 3) refers to earlier achievements reported in 2007 when 88 W of AC electric power was achieved with a total efficiency of 38% using 850 °C heater and 90 °C cooler temperatures, which puts it at 56.1% of the Carnot efficiency.

Thermoelectric materials. The highest figure of merit *ZT* to date reported for bulk thermoelectric materials (SnSe crystals at 650 °C) is 2.6 [4]. Assuming *ZT* remains at 2.6 when operating between baths at temperatures 650 °C and 23 °C would theoretically allow achieving an overall efficiency of 27.3%, that is 40.4% of Carnot efficiency. Whereas this is significantly lower than the efficiency of existing combined cycle power plants [2] and Stirling engines [3], thermoelectric materials are more suitable for generating small amounts of power [5,6], for example for wearables [7] and self-powered sensors in general [8].

Quantum dot based cooling devices. The possibility to use energy-selective electron transport through QDs for cryogenic cooling was first proposed in Ref. 9, recognizing the potential for high efficiency. Experimentally, cryogenic cooling of an electron gas has been realized using the peaked superconducting density of states in superconducting-insulator-normal metal tunnel-junctions [10], using charging effects [11], using QDs [12], and using a single electron transistor (similar to a QD) [13]. However, the electronic efficiency of such devices has to date not been tested.

B. Reproducibility – additional devices

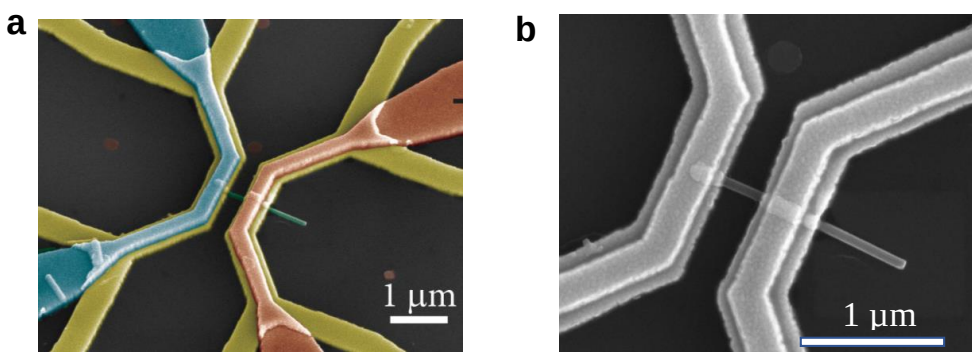
The data presented in the main text are all based on a single QD device, using a single resonance. This device was chosen because it had very good quality, a value of Γ that allowed reaching high efficiency and because we managed to collect the most comprehensive data, including a large range of load resistances that included loads optimized for maximum power production and for maximum efficiency.

The experiments have however been performed on several devices, often using more than one resonance per device. Data on other devices include QDs with stronger as well as weaker tunnel couplings compared with the one in the main text. Fits of the RTD theory to the measured data in order to obtain Γ and $T_{C,H}$ were performed for each one of them. The data also includes examples of two energetically close resonances characterized together. Below follows a summary of the other devices used in similar experiments. For every device, the same assumptions regarding Γ were made as described in the main text (see Supplemental section D). All nanowires used in fabrication of devices described in this material come from the same growth.

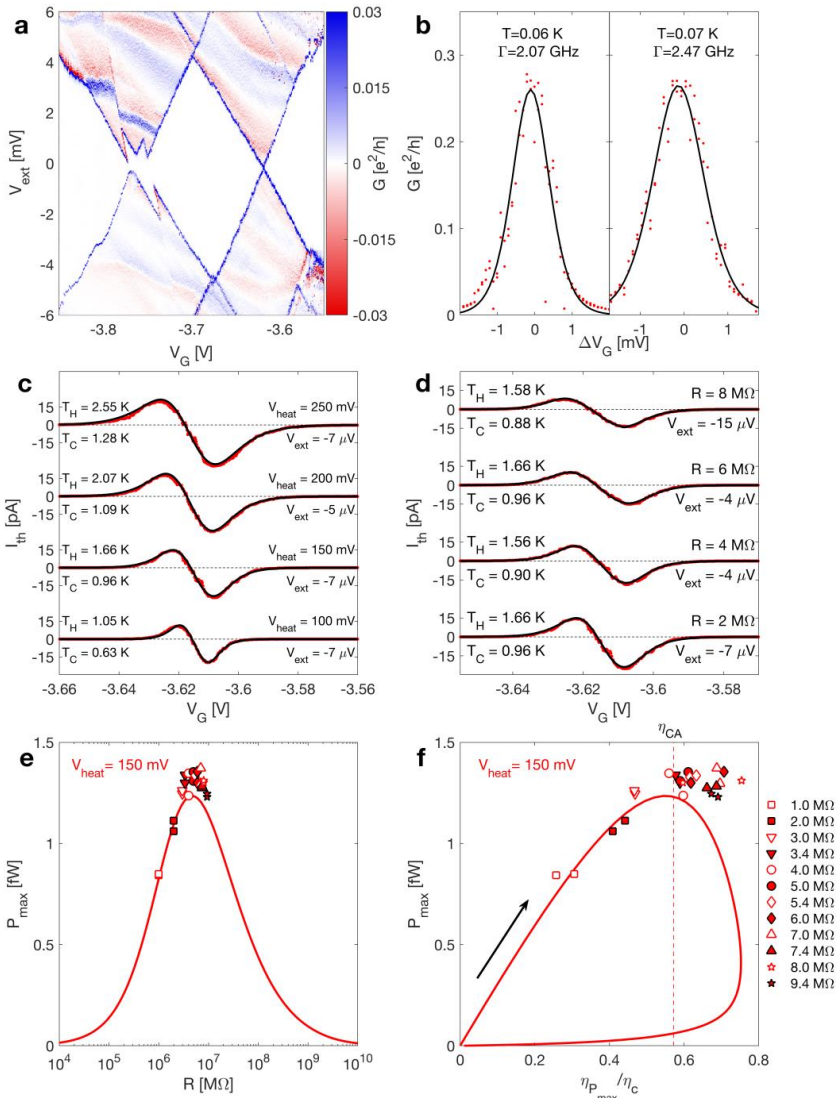
The results from all devices are consistent with the findings presented in the main text. In particular, we show parametric loop graphs for two resonances from device I, which show results consistent with those presented in Fig. 4 of the main text, albeit not for the full range of *R*.

I. Device I

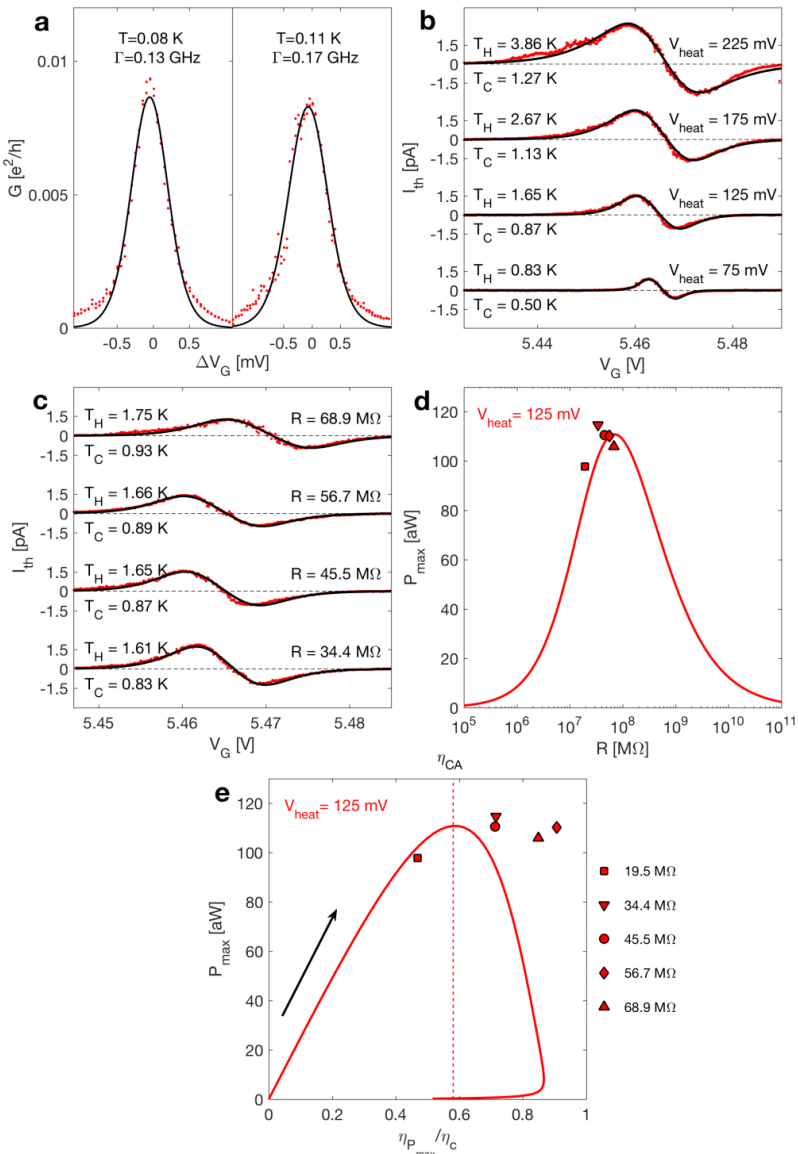
Two separate resonances were characterized for Device I. Measurements of I_{th} on Resonance 1 were done using loads values R = 1, 2, 3, 3.4, 4, 5, 5.4, 6, 7, 7.4, 8 and 9.4 MΩ. As the Γ value for Resonance 2 was found to be much smaller than for any other resonance we measured, the measurements of I_{th} on the Resonance 2 were done using higher loads R = 19.5, 34.4, 44.5, 56.7 and 68.9 MΩ. G peaks for both resonances were characterized before and after I_{th} measurements. The total heater circuit resistance was ≈ 0.6 kΩ such that a relatively high V_{heat} was needed. A typical problem when measuring this device, particularly for resonance 2, was a drifting V_{ext} which made it harder to ensure zero electrical bias of the device. Therefore, measurements of $I(V_G)$ with V_{heat} = 0 mV were taken between each V_{heat} setting. The nonzero V_{ext} showed up in the $I(V_G)$ traces as small G peaks, the magnitude of which was used to estimate a more exact V_{ext} value by fitting it to the RTD theory. The existence of an external bias, V_{ext}, also means that the dissipated power in the load is not equal to the produced power of the QD heat engine. Thus, we choose to investigate the engine's effective power, $-IV_{QD} = I(RI + V_{ext})$, and maximum effective power, $I_{max}(RI_{max} + V_{ext})$, instead.



Supplementary Fig. 1: SEM images of Device I after the measurements (also used in the main text). **a** Tilted view image with false colours, **b** top view. **Both:** The two parallel metallic (gold) strips, deposited on the substrate oxide, were used as source and drain contact leads. Heater leads are running on top of the source and drain contact leads. Only one of the heaters (coloured in red) was functional at the time of the measurements and was used for thermal biasing, while the other (blue) was not contacted.



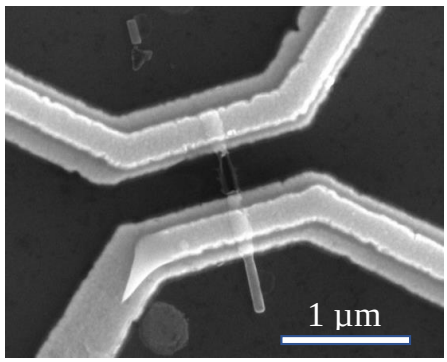
Supplementary Fig. 2: Characterization of resonance 1. a Differential conductance $G = dI/dV_{ext}$ as a function of V_G and V_{ext} of Device I, resonance 1. G data is used to determine E_C = 4.4 meV and α_G = 0.032. **b** The measured G as a function of ΔV_G (red dots) for the G peak (at V_G = -3.617 V), characterized before and after the I_{th} measurements. Fit to the RTD theory of the G peak (black lines) yielding Γ values as indicated in the figure. **c** Measured I_{th} as a function of V_G (red dots) for four V_{heat} settings using transport through the same resonance as characterized in b. All four measurements use R = 2 MΩ. The corresponding fits of the RTD theory (black curves) use Γ = 2.27 GHz, α_G = 0.032 and small offset V_{ext}, as obtained from G peaks at V_{heat} = 0 mV (values indicated in the figure). The values for T_H and T_C obtained from the fits are indicated in the figure together with the corresponding V_{heat} settings. **d** Measured I_{th} (red dots) for four different R values (as indicated in the figure) using the same device and V_{heat} = 150 mV. The small offset V_{ext}, was obtained from G peaks at V_{heat} = 0 mV (values indicated in the figure). **e** Measured (markers) and calculated (solid line) maximum effective power, $-V_{QD}I_{max} = I_{max}(RI_{max}+V_{ext})$, plotted against the external load. V_{ext} is obtained from G peaks at V_{heat} = 0 mV. **f** Parametric plot of maximum effective power, $-V_{QD}I_{max} = I_{max}(RI_{max}+V_{ext})$, and efficiency at maximum effective power when varying the external load. The arrow indicates the direction of increasing load and the dashed line shows the Curzon-Ahlborn efficiency, $\eta_{CA} \approx 0.57\eta_C$.



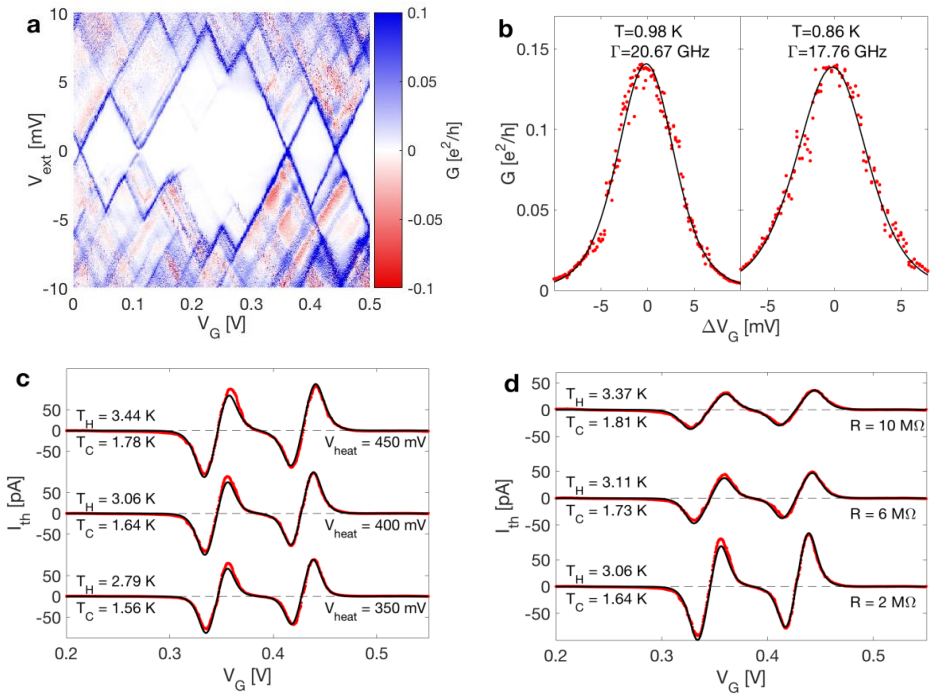
Supplementary Fig. 3: Characterization of resonance 2. a The measured G as a function of ΔV_G (red dots) for the G peak (at V_G = -5.462 V) of Device I, resonance 2. This characterization was done after the I_{th} measurements. α_G = 0.05 is obtained from characterization of Coulomb blockade peaks at finite V_{ext}. Fit of the RTD theory to the measured G peak (black lines) yielding Γ values 165 and 170 MHz (before the I_{th} measurements), 132 and 148 MHz (after the I_{th} measurements) giving an average of 154 MHz. **b** Measured I_{th} as a function of V_G (red dots) for four V_{heat} settings using transport through the same resonance as characterized in **a**. All four measurements use R = 45.5 MΩ. The corresponding fits of the RTD theory (black curves) use Γ = 154 MHz, α_G = 0.05. The values for T_H and T_C obtained from the fits are indicated in the figure together with the corresponding V_{heat} settings. **c** Measured I_{th} (red dots) for four different R values (as indicated in the figure) using the same device and V_{heat} = 125 mV. **d** Measured (markers) and calculated (solid line) maximum power plotted against external load with V_{heat} as indicated in the figure. **e** Parametric plot showing efficiency at maximum power and maximum power when the external load is varied. The solid line represents the theoretical predictions and markers the experimental values. The arrow indicates direction of increasing load and the dashed line shows the Curzon-Ahlborn efficiency, $\eta_{CA} \approx 0.58\eta_C$.

II. Device II

A set of two resonances was characterized during the measurements on Device II. Several load values were used (R = 0.01, 0.5, 1, 2, 4, 6, 8, 10 MΩ). It can be seen from Fig. S5b that both resonances were likely featuring slightly different tunnelling rates, which lowers the quality of the RTD theory fits to I_{th}, Fig. S5c-d, where the average rate of Γ = 19.22 GHz was used. To achieve a comparable thermal bias, a relatively large heater bias V_{heat} was needed because of a relatively high heater circuit total resistance ≈ 1.1 kΩ.



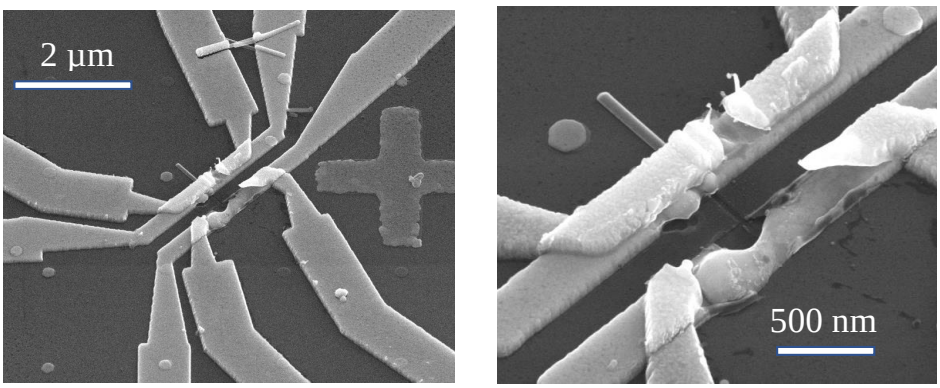
Supplementary Fig. 4: SEM image of Device II. The device was damaged in the process of preparing for imaging. The two metallic (gold) strips, deposited on the substrate oxide, were source and drain contact leads. Heater leads are running on top of the source and drain contact leads. Only one of the heaters was functional at the time of the measurements and was used for thermal biasing, the other featured a leak to the back-gate and was kept ungrounded. A shadow-like area was left between the two contact leads where the nanowire containing the QD used to be located.



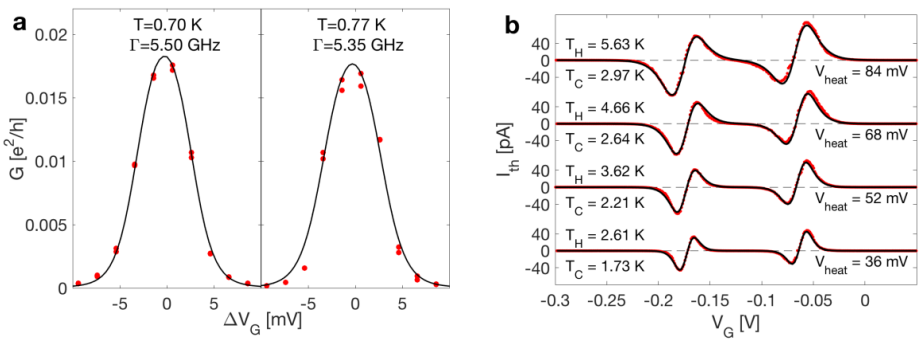
Supplementary Fig. 5: Characterization of device II. a Differential conductance $G = dI/dV_{ext}$ as a function of V_G and V_{ext} of Device II. G data is used to determine E_C = 3.9 meV and α_G = 0.048. **b** The measured G as a function of ΔV_G (red dots) for two neighbouring G peaks (at V_G = 0.430 V and V_G = 0.347 V), characterized before the I_{th} measurements. Fits of the RTD theory to the peaks (black lines) yielding Γ values, as indicated in the figure. **c** Measured I_{th} as a function of V_G (red dots) for three V_{heat} settings using transport through the same resonances as characterized in **b**. All three measurements use R = 2 MΩ. The corresponding fits of the RTD theory (black curves) use Γ = 19.22 GHz and α_G = 0.048. The values for T_H and T_C obtained from the fits are indicated in the figure together with the corresponding V_{heat} settings. **d** Measured I_{th} (red dots) for three different R values (as indicated in the figure) using the same device and V_{heat} 400 mV.

III. Device III

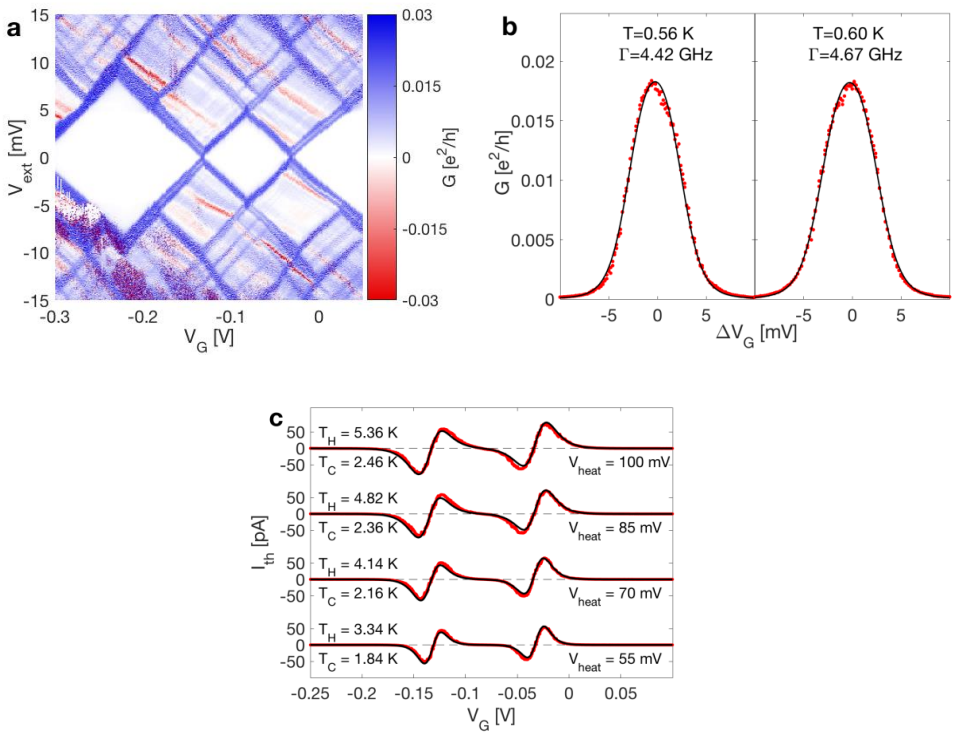
All measurements on this device were done with R = 1 MΩ which was the input impedance of the current preamplifier SR570. A set of two resonances is characterized twice. Slight charge rearrangements in the device lead to the resonances being shifted in V_G as well as to slight changes in tunnelling rates Γ. Data from this device is published in Ref.14 without the quantitative analyses of the temperatures and tunnelling rates. The total resistance of the heater circuit was ≈ 115 Ω.



Supplementary Fig. 6: SEM images of Device III. Tilted-view SEM images of Device III after the measurements. The device was damaged in the process of preparing for imaging. The two parallel metallic (gold) strips, positioned on the substrate oxide perpendicular to the nanowire, were source train contact leads. Running on top of the source and drain contact leads are the heater leads. Both heaters were functioning at the time of the measurements. Visible in the right-hand picture is a shadow-like area that was left between the two contact leads where the QD-containing nanowire used to be located before it was unintentionally blown away after the experiments, likely by some electrostatic shock.



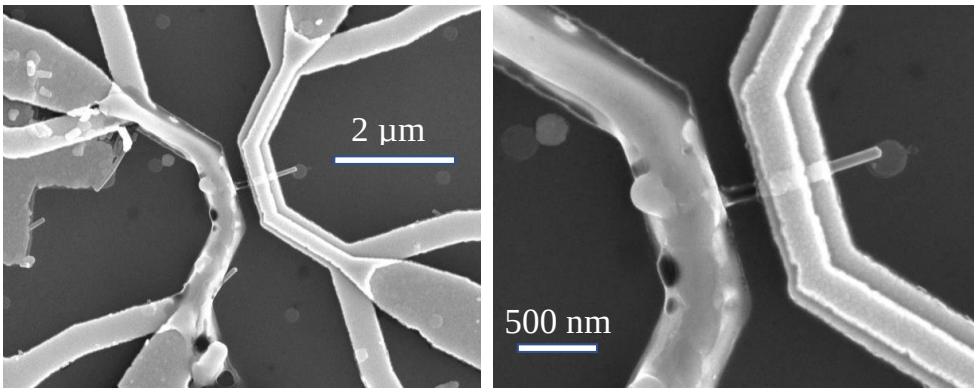
Supplementary Fig. 7: Characterization 1 of device III. **a** The measured G as a function of ΔV_G (red dots) for two neighbouring G peaks (at V_G = -0.174 V and V_G = -0.070V) of Device III. Fits of the RTD theory to the measured peaks (black lines) yielding similar Γ values, as indicated in the figure. α_G = 0.050 is obtained from characterization of Coulomb blockade at finite V_{ext}. **b** Measured I_{th} as a function of V_G (red dots) for four V_{heat} settings using transport through the same resonances as characterized in **a**. The corresponding fits of the RTD theory (black curves) use Γ = 5.43 GHz and α_G in a range between 0.050 and 0.052 (set using the I_{th} reversal points in V_G). The values for T_H and T_C obtained from the fits are indicated in the figure together with the corresponding V_{heat} settings.



Supplementary Fig. 8: Characterization 2 of device III. **A** Differential conductance $G = dI/dV_{ext}$ as a function of V_G and V_{ext} of Device III. G data is used to determine E_C = 4.7 meV and α_G = 0.0472 **b** The measured G as a function of ΔV_G (red dots) for two neighbouring G peaks at V_G = -0.132 V and V_G = -0.03V, characterized before the I_{th} measurements. Fits of the RTD theory to the peaks (black lines) yielding similar Γ values, as indicated in the figure. G characterization before and after I_{th} measurements provided four values Γ = 4.67, 4.42, 4.33, 3.40 GHz, yielding an average of 4.21 GHz. **C** Measured I_{th} as a function of V_G (red dots) for four V_{heat} settings using transport through the same resonances as characterized in **b**. The corresponding fits of the RTD theory (black curves) use Γ = 4.21 GHz and α_G in a range between 0.047 and 0.048 (set using the I_{th} reversal points in V_G that changed slightly with temperature). The values for T_H and T_C obtained from the fits are indicated in the figure together with the corresponding V_{heat} setting.

IV. Device IV

Device IV was the device used for the experiments in the main paper. In Fig. S9 we show SEM images of the device taken after the measurements.



Supplementary Fig. 9: SEM images of Device IV. Top view SEM images of Device IV after the measurements. The device was damaged in the process of preparing for imaging. The two parallel metallic (gold) strips, deposited on the substrate oxide perpendicular to the nanowire, were used as source and drain contact leads. The leads running on top of the source and drain contact leads are the heater leads. Only one of the heaters was functional at the time of the measurements and was used for thermal biasing, the other featured a leak to the back-gate and was kept ungrounded. A shadow-like area was left between the two contact leads where the nanowire containing the QD used to be located.

C. Theory

Below we describe the theoretical method used in the main paper. In short, our theory is based on the real-time diagrammatic techniques originally developed in Refs. [15,16]. To model the experiment we need to consider a finite voltage bias as well as a temperature difference, and we calculate both the charge and heat current, for which most aspects of the underlying theory have been developed in Refs. 17,18,19.

I. Hamiltonian

We are interested in open quantum systems where a central quantum dot (QD) is coupled to macroscopic reservoirs via hybridization terms in the Hamiltonian. The Fock-space of the central QD is typically small. The reservoirs are assumed to be described by the grand canonical ensemble.

The total system is modelled as the Hamiltonian

$$H = H_D + \sum_r H_r + \sum_r H_{T,r} = H_D + H_R + H_T. \tag{1}$$

Here, H_D is the Hamiltonian describing the QD, which is assumed to be diagonal in the many body states of the QD

$$H_D = \sum_\sigma \varepsilon_\sigma n_\sigma + E_C n_\uparrow n_\downarrow, \tag{2}$$

with single particle energy ε_σ for an electron with spin $\sigma = \uparrow, \downarrow$, and energy cost E_C associated with double occupation (because of electron-electron interactions). $n_\sigma = d_\sigma^\dagger d_\sigma$ is the number operator made up from the fermion creation ($d_\sigma^\dagger$) and annihilation (d_σ) operators. In the following, field operators acting on the QD subspace are denoted by the letter d and those acting on the reservoir subspace by the letter c. A reservoir r is described by

$$H_r = \sum_{k,\sigma} \omega_{k,\sigma,r}\, n_{k,\sigma,r}, \quad n_{k,\sigma,r} = c_{k,\sigma,r}^\dagger c_{k,\sigma,r} \tag{3}$$

where $\omega_{k,\sigma,r}$ is the eigenenergy for a state in reservoir r with quantum numbers k, σ. Finally, the QD and reservoir r are coupled by the tunnelling Hamiltonian

$$H_{T,r} = \sum_{k,\sigma} t_{k,\sigma,r}\, d_\sigma^\dagger c_{k,\sigma,r} + h.c.. \tag{4}$$

The amplitude for electron tunnelling is given by $|t_{k,\sigma,r}|^2$, which allows us to define a tunneling rate

$$\Gamma_r = \frac{2\pi \nu_r |t_{k,\sigma,r}|^2}{\hbar}. \tag{5}$$

Here ν_r is the density of states of the reservoir and $\hbar$ is Planck's reduced constant. In the sections to come we utilize the wide band approximation assuming that ν_r is constant over an energy D much larger than any other involved energy scale, and for convenience we set $\hbar = e = k_B = 1$.

II. Liouville-von Neumann equation

Our aim is to calculate the nonequilibrium state of the QD when the reservoirs are kept at different temperatures and chemical potentials, and then to find the charge and energy currents flowing as a result of this nonequilibrium condition. The starting point is the Liouville-von Neumann equation for the time evolution of the full system density matrix (ρ)

$$\frac{d}{dt}\rho = -i[H,\rho]_- = -iL\rho, \tag{6}$$

where $L\bullet \equiv [H,\bullet]_-$ is the Liouville super-operator acting on an arbitrary operator $\bullet$ (we similarly define $L_T\bullet \equiv [H_T,\bullet]_-$ and $L_D\bullet \equiv [H_D,\bullet]_-$). Here the term super-operator is used for mathematical objects operating on normal operators. The matrix elements of such a super-operator are evaluated as

$$A_{cd}^{ab} = \langle c|(A|a\rangle\langle b|)|d\rangle. \tag{7}$$

There are many ways to find an approximate solution to eq. 6, we here use the real time diagrammatic technique [15] formulated in Liouville space [20]. The first step is performing a Laplace transformation and expanding the resulting expression in L_T. This expansion can then be re-summed by identifying it as a geometrical series to obtain

$$\rho_D(z) = \frac{i}{z - L_D - W(z)}\rho_D(0), \tag{8}$$

where $\rho_D = \mathrm{Tr}_R\{\rho\}$ is the QD (reduced) density matrix and the kernel $W(z)$ is given by

$$W(z) = \sum_{k=1}^{\infty} \mathrm{Tr}_R\left(L_T \frac{1}{z - L_D - L_R}\right)^k L_T\rho_R|_{irred.}. \tag{9}$$

irred here refers to the irreducible diagrams in the perturbative series [15,20,21].

Here, we are only interested in the stationary state (long time limit and time-independent Hamiltonians). Re-arranging eq. 8 and taking the stationary (zero frequency) limit yields

$$0 = (-iL_D + W)\rho_D. \tag{10}$$

The generalized master equation in eq. 10 is valid for the full QD density matrix, but for our model the QD density matrix will be diagonal in spin and charge because these quantities are conserved by the total Hamiltonian.

Next we will introduce a convenient super-operator notation in Liouville space which will make the evaluation of W fairly straight forward.

III. Kernel evaluation

In order to evaluate the series in eq. 9 it is convenient to find super-operators in Liouville space whose properties resemble those of fermion field operators in Fock space.

We define the notation for our fermionic creation and annihilation operators acting on the QD subspace as

$$d_{\eta\sigma} = \begin{cases} d_\sigma & \text{if } \eta = - \\ d_\sigma^\dagger & \text{if } \eta = + \end{cases}. \tag{11}$$

We combine all indices to a single index represented by a positive integer, $1 = \eta_1\sigma_1$, and denote by a bar the sign change of appropriate numbers i.e. $\overline{1} = \overline{\eta}_1\sigma_1 = (-\eta_1)\sigma_1$ and $\overline{q}_1 = -q_1$. Furthermore, a sum over such a multi-index (Σ_1) is implicitly understood as sums over all involved indices if nothing else is stated. The field operators acting on the reservoir subspace are defined in a similar manner where the indices k and r are also included in the multi-index.

Using the notation in eq. 11 we define Liouville field super-operators as

$$\mathcal{G}_1^{q_1} \bullet = \frac{1}{\sqrt{2}}(d_1 \bullet + q_1(-1)^N \bullet (-1)^N d_1), \quad q_1 \in \{-,+\} \tag{12}$$

with $N = \sum_\sigma n_\sigma$. These operators have properties similar to those of fermionic operators in Hilbert space

$$(\mathcal{G}_1^{q_1})^\dagger = \mathcal{G}_{\overline{1}}^{\overline{q}_1}, \tag{13}$$

$$[\mathcal{G}_2^{q_2}, \mathcal{G}_1^{q_1}]_+ = \delta_{q_2\overline{q}_1}\delta_{2\overline{1}}\mathcal{I}. \tag{14}$$

Analogously for the field operators acting on the reservoir subspace we use the notation $\mathcal{J}_1^{q_1}$.

Using these operator definitions the kernel in eq. 9 can be evaluated. However, this expression contains an infinite sum of terms and needs to be truncated. Only terms of even order in k gives a non-vanishing contribution and we include terms up to order $k = 4$, which account for processes $\propto \Gamma$ and Γ^2. The remaining terms are evaluated by collecting all reservoir super-operators and calculating their expectation value using Wick's theorem. The leading order terms of the kernel now reads

$$W^{(2)} = \sum_{q_{11}} \frac{\Gamma_1}{2\pi} \mathcal{G}_{\overline{1}}^{+} \frac{q_1 \gamma_1^{q_1}}{i0 - L_D + \eta_1\omega_1} \mathcal{G}_1^{\overline{q}_1}, \tag{15}$$

and the next to leading order [17,18,19].

$$\begin{aligned} W^{(4)} = \sum_{12q_1q_2} \frac{\Gamma_1\Gamma_2}{(2\pi)^2} & \left(\mathcal{G}_{\overline{1}}^{+} \frac{1}{\eta_1\omega_1 + i0 - L_D} \mathcal{G}_{\overline{2}}^{+} - \mathcal{G}_{\overline{2}}^{+} \frac{1}{\eta_2\omega_2 + i0 - L_D} \mathcal{G}_{\overline{1}}^{+}\right) \\ & \times \frac{\overline{q}_2 \gamma_2^{\overline{q}_2}}{\eta_1\omega_1 + \eta_2\omega_2 + i0 - L_D} \mathcal{G}_2^{q_2} \frac{\overline{q}_1 \gamma_1^{\overline{q}_1}}{\eta_1\omega_1 + i0 - L_D} \mathcal{G}_1^{q_1}, \end{aligned} \tag{16}$$

Where

$$\gamma^{q_1} = \delta_{q_1+} + \delta_{q_1-}\tanh\left(\frac{\eta_1(\omega_1 - \mu_1)}{2T_1}\right). \tag{17}$$

The resulting integrals in equation 15 and 16 are solved in the Supplementary Information of Ref.17.

IV. Observables

Any observable can now be calculated using the density matrix

$$\langle A \rangle = \mathrm{Tr} A\rho. \tag{18}$$

When evaluating this expectation value it is however possible to make several simplifications by writing

$$\langle A \rangle = \frac{1}{2}\mathrm{Tr}\, L_A^+ \rho, \quad L_A^+ \bullet = [A, \bullet]_+. \tag{19}$$

IV.a Charge current

The charge current is given by the time derivative of the particle number in one of the reservoirs

$$I_r = -\frac{d}{dt}N_r = -i[H, N_r], \quad N_r = \sum_{k,\sigma} n_{k,\sigma,r}. \tag{20}$$

The calculation of the charge current can be simplified by using the fact that the charge is conserved in all tunnelling processes [22]. Using this conservation law the current leaving reservoir r is given by

$$I_r = -\frac{1}{2} i\, \mathrm{Tr} L_N^+ W_r \tag{21}$$

where $N = \sum_\sigma n_\sigma$ and W_r is similar to W in eq. 9 with the only difference that the left-most L_T is replaced by $L_{T,r}$.

IV.b Energy and heat current

A similar treatment as the charge current is possible for the energy current. It is however non-trivial because the tunnelling Hamiltonian introduces additional dynamics which cannot be ignored [19, 23]. The energy current thus consists of two parts

$$J_{E_r} = i\underset{D}{\mathrm{Tr}} L_{H_D}^+ W_r \rho_D - i\underset{D}{\mathrm{Tr}} W_{\Gamma,r}\rho_D \tag{22}$$

where the first term on the right-hand side is evaluated in the same manner as eq. 21. In order Γ only the first term in 22 contributes and in the next to-leading order the second term is given by [19]

$$-i\underset{D}{\mathrm{Tr}} W_{\Gamma,r}\rho_D = \frac{\Gamma_r \Gamma_{r'}}{2\pi} \sum_{\sigma,p=\pm} [\, f_{r\sigma p} + p f_{r\sigma p} \underset{D}{\mathrm{Tr}}(-\mathcal{I})^{n_{\sigma'}} \rho_D], \tag{23}$$

where $r' \neq r$, $\mathcal{I}$ the identity matrix and

$$f_{r\sigma p} = \mathrm{Re}\left[\Psi\left(\frac{1}{2} - i\frac{\frac{E_C}{2(p-1)} - \epsilon + \mu_r}{2\pi T_r}\right) - \Psi\left(\frac{1}{2} - i\frac{\frac{E_C}{2(p-1)} - \epsilon + \mu_{r'}}{2\pi T_{r'}}\right)\right], \tag{24}$$

with Ψ denoting the digamma function.

The heat flow leaving the hot reservoir is related to the above quantities through the first law of thermodynamics

$$J_{Q_H} = J_{E_H} - \frac{\mu_H}{e} I_H. \tag{25}$$

V. Self-consistent solution

The introduction of a serial load R introduces additional computation steps because the electrical bias V_{QD} across the QD is set by the current flowing through the dot, $V_{QD} = I_{QD}R$, rather than by an external bias applied to the reservoirs (see circuit diagram in Fig. 1c in the main paper). V_{QD} thus needs to be obtained by solving the self-consistent equation

$$I_{QD}(V_{QD}) - \frac{V_{QD}}{R} = 0, \tag{26}$$

which is done numerically in an iterative manner. $I_{QD}(V_{QD})$ is obtained from eq. 21 with $eV_{QD} = \mu_H - \mu_C$. This means that for each iteration in the root finding algorithm (eq. 26) H_R is changed and thus a new density matrix and new currents need to be calculated by going through the steps described above.

D. Details of the measurement and fitting process

The parameters Γ, T_C and T_H were obtained from fitting the RTD theory to the measured conductance and current using least square fits. Data in Fig. 2a in the main text allowed determination of the gate coupling $\alpha_G = 0.049$ which was further used as an input parameter. Motivated by the straight Coulomb diamonds (see Fig. 2a in the main paper) we let the potential eV drop symmetrically over both tunnel barriers. Furthermore, the value of ε_0 was adjusted to compensate for electrostatic potential variations, not considered by the model. On a measurement-to-measurement basis, one source of such variations in general is charge rearrangements in the proximity of the QD. In our experimental setup there was an additional factor, namely finite electrical potential changes at the top-heater as a function of V_{heat}. The origin of this effect can be understood in the following way. Whenever the heater is biased by V_{heat} in order to run a heating current through it, the filter resistances in series with the heater on both sides develop voltages across them. In case this effect is not counterbalanced by applying $V_{heat} = V^{R}_{heat} - V^{L}_{heat}$ in a push-pull manner (by setting positive bias V^{R}_{heat} on one side of the heater and negative bias V^{L}_{heat} on the other, or visa versa) it results in an overall potential change at a heater electrode that can slightly gate the QD. A characteristic of this effect is that this parasitic "gating" scales linearly with V_{heat}. Unfortunately, due to different resistances of the voltage sources' output filters the balancing of the potentials was not done perfectly and a slight shift in the I_{th} reversal point with increasing ΔT remained and was compensated for when fitting.

I. Tunnel rate characterization

The tunnelling rate Γ can be related to the $G(V_{ext} = 0)$ of the Coulomb peaks (Fig 2b in the main text). The V_G dependence of the G peaks is sensitive to Γ and temperature T. Experimentally we determined $G(V_G)$ by measuring I-V_{ext} curves and finding linear fits at each V_G. The measurements were done in DC and used the range of V_{ext} in which the I-V curves are linear, typically $V_{ext} \leq \pm$ 50 μV. We believe that most of the scatter of the data points in G originates from the variations in the electrostatic potential landscape due to charge rearrangements in the proximity of the QD, having an effect similar to V_G noise.

The explained G peak characterization was done at the beginning as well as at the end of the thermoelectric characterization of the resonance. For better consistency $G(V_G)$ was measured at the cryostat base temperature (estimated electron temperature slightly below 0.2 K) as well as at two elevated temperatures obtained by rising the cryostat mixing chamber temperature to 0.5 and 1.0 K. From the resulting fits of the RTD theory to the measured data we could conclude that at the cryostat base temperature $\hbar\Gamma/kT \approx 0.34$ was too large for the theory to quantitatively reproduce the measurements with high accuracy because we were not deep enough in the perturbative regime ($\hbar\Gamma/kT << 1$). Thus only G measurements at elevated temperatures were used for determining Γ. The four remaining conductance measurements yielded Γ = 8.93 GHz (and T = 503 mK), Γ = 9.25 GHz (and T = 480 mK), Γ = 8.28 GHz (and T = 906 mK) and Γ = 9.06 GHz (and T = 953 mK). This gives an average Γ = 8.88 ± 0.42 GHz which was used for all calculations in the main text. One standard deviation is used as the error interval.

II. Temperature characterization

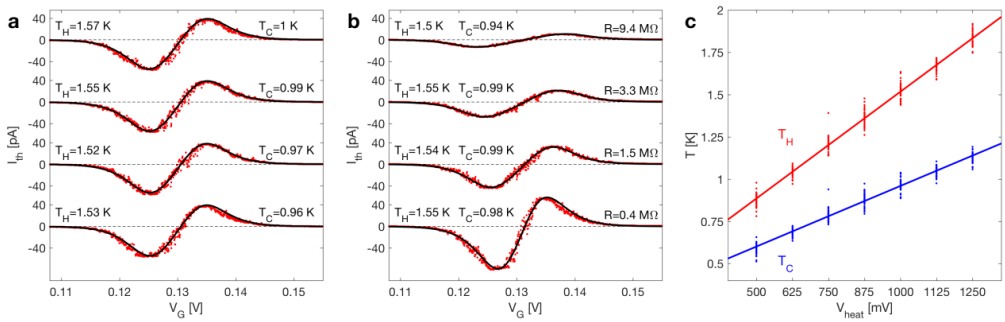
In our experimental device the electron reservoirs were, effectively, the short (~ 100 nm) InAs nanowire segments, and our analysis requires determination of their electronic temperatures T_H and T_C. Fitting the RTD theory to the characteristic behaviour of $I_{th}(V_G)$ is a non-invasive method with good sensitivity given that the thermal bias $\Delta T = T_H - T_C$ is big enough to produce a sufficiently large I_{th} for accurate measurements. We preferred using $I_{th}(V_G)$ over the more conventionally used $I(V_{ext})$ [24], because of the energy dependency of the tunnelling rates at higher V_{ext} conditions (manifesting in a V_{ext} dependent saturation current Fig. S13). This is likely due to the small size of the nanowire lead

segments that affect the density of states in the reservoirs. Using the $I_{th}(V_G)$ ensures that the temperatures are characterized at the conditions in which the engine itself is characterized.

T_H and T_C were determined from least-square fits of the RTD theory to $I_{th}(V_G)$ using the temperatures as the only two free parameters. We used Γ = 8.88 GHz and α_G = 0.049 as determined from the G measurements.

Small measurement-to-measurement variations in the V_G-location of the resonant energy ε_0 were compensated for by setting the theoretical I_{th} reversal point equal to the measured one before fitting. We also accounted for a small offset in the current (ca -0.3 pA), determined from the current reading in the Coulomb blockade regime without any biases, V or ΔT, applied. The external load was included in the model by self-consistently solving for V across the QD.

$I_{th}(V_G)$ was measured four times for each V_{heat} configuration (two times sweeping V_G in one direction, and two in the other). The two-parameter RTD theory fit for T_H and T_C was done for each $I_{th}(V_G)$ trace independently. An example of four such fits is shown in Fig. S10a. This procedure was repeated for all 11 R values used in the experiment of the presented device and similar temperatures for all R values were found. For an example, the Fig. S10b shows the result for four I_{th} using V_{heat} = 1000 mV, but for different R. In Fig. S10c we summarize thermometry results performed at 7 different thermal bias conditions 4 times for each of the 11 loads. Overall both T_H and T_C are found to increase approximately linearly with V_{heat}.



Supplementary Fig. 10: Temperature characterization. The fits of the RTD theory to I_{th} using the two temperatures are free parameters yield similar results for all measurements using a given V_{heat}. **a** Results from the RTD theory fits to $I_{th}(V_G)$ for all measurements with V_{heat} = 1000 mV and R = 1.5 MΩ. **b** Resulting RTD theory fits to $I_{th}(V_G)$ using V_{heat} = 1000 mV and different R as indicated. **c** The temperatures from fits to all measurements including V_{heat} in a range between 500 and 1250 mV, and R in a range between 14 kΩ and 250 MΩ.

III. Validity of the model and heat flow estimate

In order to estimate the heat-to-electric energy conversion efficiency $\eta = P_{th}/J_Q$, our analysis required a quantitative estimate of the heat flow J_Q. To achieve this, we used microscopic modeling based on the RTD theory, described in Section C. This theoretical approach includes full nonlinear effects, strong electron-electron interactions and properly takes into account all second order tunneling processes. Furthermore, our QD model has only a few experimentally determined parameters (Γ and gate coupling are extracted from G, and the $T_{C,H}$ at each V_{heat} from I_{th}).

By using the theory described in Section C we found an excellent agreement between the calculated charge current, eq. 21, and the measured I_{th} over wide value ranges of V_G, V_{heat} (see fig. 2.c in the main paper) and R (see fig. S10.b). This quantitative and qualitative agreement provides a strong indication

that our model fully captures the physics of the experiment. Even though we used modelling – rather than a direct measurement – to estimate the electronic heat flow J_Q using eq. 25 we are confident about the quantitative accuracy of our calculations. An alternative, model-independent way of determining the electronic heat flow, used in Refs. 25 and 26 does not allow for independent control of the thermal and electric biases and is therefore incompatible with our performance analysis at constant ΔT.

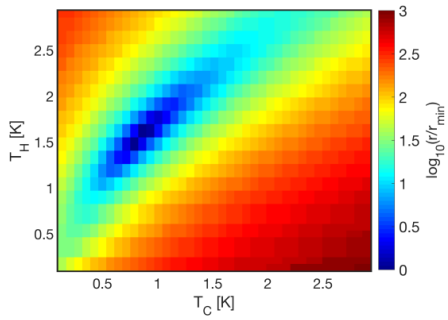
IV. Finding the best fit

When fitting the measured current and conductance to the RTD theory in order to obtain the tunnel rate and the temperatures, the best fit is found by minimizing the sum of squares of the residues

$$r = \sum_i \left[I_{theory}(V_{G,i}) - I_{experiment}(V_{G,i}) \right]^2. \quad (27)$$

An example of such a residue is shown in Fig. S11 where the current is measured for V_H = 1000 mV and R=14.4 kΩ. The termination condition used for the temperature fits is that the minimum is found with a resolution of 0.003 K and when fitting the tunnel rate the resolution is 0.01 GHz.

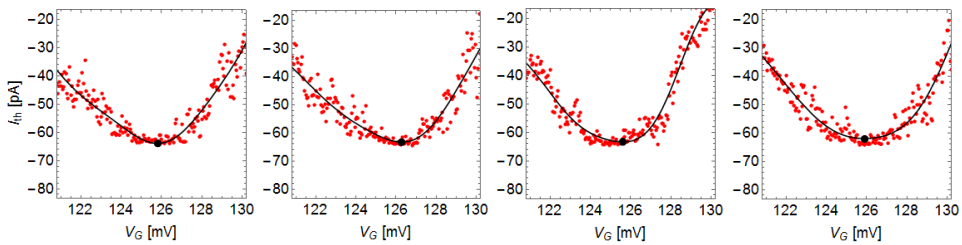
Since the residue in Fig. S11 shows one clear, global minimum the fit results in two unique temperatures and the fitting procedure is a valid thermometry tool.



Supplementary Fig. 11: Residues from temperature fit. Logarithm of the sum of residues (see eq. 27) normalized by its minimum value when fitting the RTD theory to the measured thermocurrent using V_{heat} = 1000 mV and R = 14.4 kΩ. The best fit is given by T_C = 0.89 K and T_H = 1.46 K. A global minimum means that the fitting technique produces two unique temperatures.

V. Obtaining the maximal power value from the data

In Fig. 3a,b of the main text P_{th} was calculated using $P_{th} = I_{th}^2 R$ on a point-to-point basis. A polynomial fit was used to obtain the maximal power value from the each I_{th} trace. The gate position of the I_{th} extremum was then used to calculate the electronic heat flow J_Q and thus also the efficiency at maximum power, η_{Pmax}.



Supplementary Fig. 12: Obtaining the I_{th} maximal amplitude for the maximal $P_{th} = I_{th}^2 R$. All four traces show I_{th} data around the maximal power peak for R = 1 MΩ and V_{heat} = 1250 mV. In total 400 data points (red dots) around the I_{th} peak are fitted using a polynomial (black line). The peak value and position in V_G is determined by the extremum of the fit (black dot).

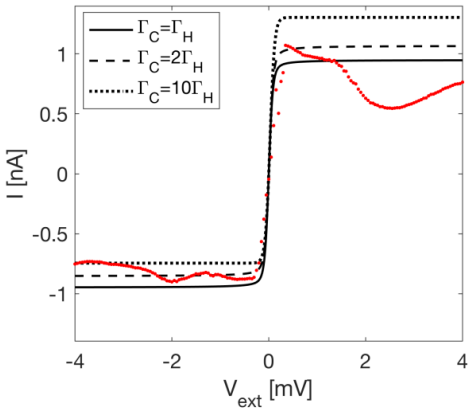
E. Symmetric tunnel rates

The procedure by which we determined Γ from conductance data is described in Section D.I. During the fitting process we assumed that the QD is coupled equally strong to both reservoirs, i.e. $\Gamma = \Gamma_C = \Gamma_H$. Here we motivate why we are safe to assume symmetric and energy independent tunnelling rates.

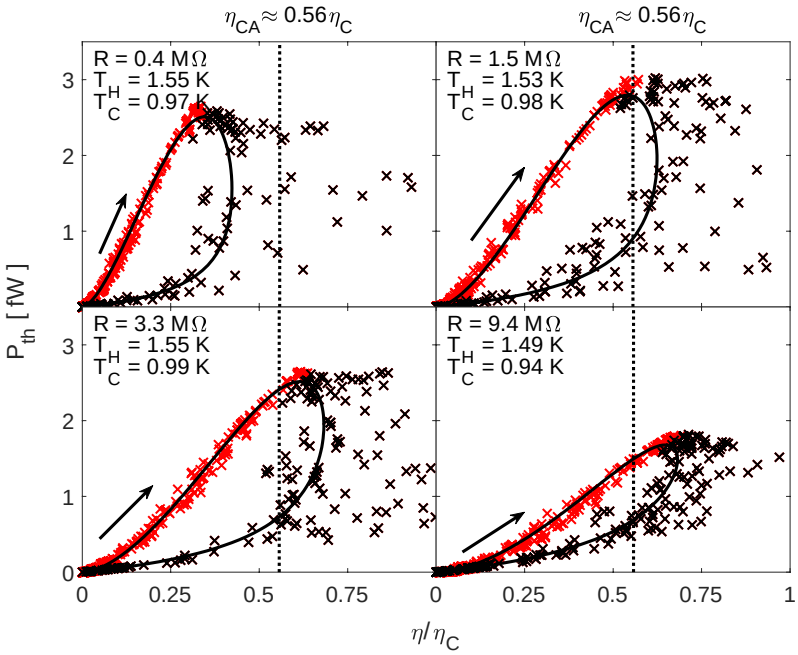
Performing the RTD theory fits to the conductance $G(V_{ext} = 0)$ alone does not provide any information about any asymmetry in tunnelling rates, because the conductance scales with the parameter $\Gamma_C\Gamma_H/(\Gamma_C+\Gamma_H)$ and does not directly depend on Γ_C and Γ_H. Information about asymmetry is however found in the saturation current, i.e. the maximum current the corresponding resonance can carry in either direction. This is easiest seen when gating the device to the resonant condition ($V_G \sim 0.13$ V in Fig 2a of the main text) and applying a large enough bias V_{ext} of both polarities. For example, for a two-fold spin degenerate QD energy level, one finds that a large asymmetry, $\Gamma_{C/H} >> \Gamma_{H/C}$, results in a factor two for the difference in the saturation currents in opposite directions [27].

In Fig. S13 we compare the measured saturation currents $I(V_{ext})$ in our device to the theoretical predictions using $\Gamma_C = \Gamma_H$, $\Gamma_C = 2\Gamma_H$ and $\Gamma_C = 10\Gamma_H$ while keeping $\Gamma_C\Gamma_H/(\Gamma_C+\Gamma_H)$ fixed. The measured plateau current is not constant, which we attribute to energy dependent density of states (DOS) of the short nanowire leads segments. However, we note that the scale within which these variations manifest themselves is on the order of several mV while the voltages that were developed across the QD during the thermoelectric experiments were limited to much smaller values, less than 0.25 meV, within which it is safe to assume an energy-independent DOS.

The saturation current levels in Fig. S12 suggest that there might be a small Γ asymmetry in our device corresponding to Γ_C/Γ_H between 1 and 2, but it is not possible to obtain a good quantitative estimate of its magnitude. Such an asymmetry will slightly change the thermally induced current and thus also the temperatures extracted form fitting the RTD theory to the current, and more importantly increase the heat current since $\Gamma_H + \Gamma_C$ is increased. In Fig. S14 we show the same type of analyses as in Fig. 3c in the main paper with the exception that the analysis is performed using an asymmetry factor of 2. The difference between η at maximum P obtained when using $\Gamma_C/\Gamma_H = 2$, instead of $\Gamma_C/\Gamma_H = 1$, is at most $0.02\,\eta_C$, showing that the asymmetry has little effect on the analysis of our dot and it is safe to assume symmetric couplings.



Supplementary Fig. 13: Saturation current. Measured (red dots) and calculated (lines) current as a function of V_{ext} at the resonance (V_G = 0.126 V). The RTD calculations use different asymmetries as indicated in the figure. The magnitude of the saturation current is used to probe tunnel coupling asymmetry.



Supplementary Fig. 14: Thermoelectric performance of the QD including tunnel rate asymmetry. Same as Fig 3 c in the main paper but using $\Gamma_C = 2\Gamma_H$, $\Gamma_C\Gamma_H/(\Gamma_C+\Gamma_H) = 4.44$ GHz. Parametric plot of P_{th} and $\eta = P_{th}/J_Q$ as V_G is varied at V_{heat}= 1000 mV for four different R = 0.4, 1.5, 3.3 and 9.4 MΩ. Red and black markers identify data points from the corresponding V_G ranges as indicated in FIG 3 a,b in the main paper**.** Data points are based on the measured values of P_{th} and the calculated J_Q using the experimentally determined parameters. The solid lines are based purely on RTD calculations using the same parameters. The arrow indicates the direction for increasing V_G. The plots show close resemblance to those in Fig. 3c. The difference between η at maximum P obtained when using $\Gamma_C/\Gamma_H = 2$, instead of $\Gamma_C/\Gamma_H = 1$, is at most $0.02\eta_C$, which is within the measurement uncertainties.

Any additional resistance in series with the QD would modify the Coulomb peak, which would in turn lead to an incorrect Γ value obtained from the G measurements. However, since the magnitude of

saturation currents is supposed to be independent of the series load it can be used to check whether any additional resistances, such as contact resistances between the nanowire and the metallic leads or nanowire lead resistances, are negligible. This is done by comparing theory predictions for saturation currents given the obtained Γ value. Indeed, from Fig. S12 we see that $\Gamma = 8.9$ GHz, as determined from $G(V_{ext} = 0)$, predicts the saturation current magnitude quite well, and we thus conclude that there are no significant contributions from parasitic resistances.

Further evidence for the negligibility of the contact resistances is provided by a comparison of the measured and calculated thermoelectric power produced by the dot. Our measurement of $P_{th} = I_{th}^2 R$ is determined using the known load R. Any additional power dissipated, say in the nanowire leads (a few kΩ at most), is not accounted for and would lead to an error in P_{th} proportional to the relative contribution of the parasitic resistance to the overall resistance. If this error was significant it would manifest itself as a significant deviation from the predicted power at smaller loads (Fig. 4a). However, this is not the case, further confirming the negligibility of parasitic resistances.

References

[1] Esposito, M. Kawai, R. Lindenberg, K. & Van den Broeck, C. Efficiency at Maximum Power of Low-Dissipation Carnot Engines. *PRL* **105**, 150603 (2010).

[2] Ol'khovskii, G. G. Combined cycle plants: yesterday, today, and tomorrow. *Thermal Engineering* **63**, 488–494 (2016).

[3] Wong, W. A. Wilson, S. Collins, J. & Wilson, K. Advanced Stirling convertor (ASC) technology maturation. *NASA/TM-2016-218908* (2016).

[4] Zhao, L. D. *et al.* Ultralow thermal conductivity and high thermoelectric figure of merit in SnSe crystals. *Nature* **508,** 373 (2014).

[5] Vining, C. B. An inconvenient truth about thermoelectrics. *Nature Materials* **8,** 83-85 (2009).

[6] Snyder, G. J. Small thermoelectric generators. *Electrochemical Society Interface* **17,** 54-56 (2008).

[7] Kim, M. K. Kim, M. S. Lee, S. Kim, C. & Kim, Y.J. Wearable thermoelectric generator for harvesting human body heat energy. *Smart Mater. Struct.* **23,** 105002 (2014).

[8] Kuchle, J. J. & Love, N. D. Self-powered wireless thermoelectric sensors. *Measurement* **47,** 26–32 (2014).

[9] Edwards, H. L. Niu, Q. & de Lozanne, A. L. A quantum-dot refrigerator. *Appl. Phys. Lett.* **63,** 1815 (1993).

[10] Nahum, M. Eiles, T. M. & Martinis, J. M. Electronic microrefrigerator based on a normal-insulator-superconductor tunnel junction. *Appl. Phys. Lett.* **65,** 3123 (1994).

[11] Saira, O. P. *et al.* Heat transistor: demonstration of gate-controlled electronic refrigeration. *Phys. Rev. Lett.* **99,** 027203 (2007).

[12] Prance, J. R. *et al.* Electronic refrigeration of a two-dimensional electron gas. *Phys. Rev. Lett.* **102,** 146602 (2009).

[13] Feshchenko, A. V. Koski, J. V. & Pekola, J. P. Experimental realization of a Coulomb blockade refrigerator. *Phys. Rev. B* **90,** 201407 (2014).

[14] Svilans, A. Burke, A. M. Fahlvik Svensson, S. Leijnse, M. & Linke, H. Nonlinear thermoelectric response due to energy-dependent transport properties of a quantum dot. *Physica E* **82,** 34-38 (2015).

[15] König, J. Schoeller H. & Schön, G. Resonant tunneling and Coulomb oscillations. *Eur. Phys. Lett.* **31**, 31 (1995)

[16] Schoeller, H. & Schön, G. Mesoscopic quantum transport: Resonant tunneling in the presence of a strong Coulomb interaction. *Phys. Rev. B* **50,** 18436 (1994)

[17] Gergs, N.M, Bender, SA, Duine, & and Schuricht, D. Spin switching via quantum dot spin valves. *Phys. Rev. Lett.* **120,** 017701 (2018).

[18] Gergs, N.M., Hörig, C.B.M, Wegewijs, M. R. & Schuricht, D. Charge fluctuations in nonlinear heat transport. *Phys. Rev. B* **91,** 201107 (2015).

[19] Gergs, N.M. Transport and topological states in strongly correlated nanostructures. Ph.D. thesis, Utrecht University (2017).

[20] Schoeller, H. A perturbative nonequilibrium renormalization group method for dissipative quantum mechanics. *Eur. Phys. J. Spec. Top*. **168**, 179-266 (2009).

[21] Leijnse, M. & Wegewijs, M.R. Kinetic equations for transport through single-molecule transistors. *Phys. Rev. B* **78,** 235424 (2008).

[22] Saptsov, R.B. & Wegewijs, M.R. Fermionic superoperators for zero-temperature nonlinear transport: Real-time perturbation theory and renormalization group for Anderson quantum dots. *Phys. Rev. B* **86,** 235432 (2012).

[23] Gergs, N. M., Saptsov, R., Wegewijs, M. R. & Schuricht, D. Unpublished.

[24] Giazotto F. Heikkilä, T. T. Luukanen, A. Savin, A. M. & Pekola, J. P. Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications. *Rev. Mod. Phys.* **78,** 217 (2009)

[25] Jezouin, S. *et al.* Quantum limit of heat flow across a single electronic channel. *Science* **342**, 601 (2013).

[26] Banerjee, M. *et al.* Observed quantization of anyonic heat flow. *Nature* 545, 75 (2017)

[27] Bonet, E. Deshmukh, M. M. & Ralph, D.C. Solving rate equations for electron tunneling via discrete quantum states. *Phys. Rev. B* **65,** 045317 (2002).

Paper III

Thermoelectric characterization of the Kondo resonance in nanowire quantum dots

Artis Svilans, Martin Josefsson, Adam M. Burke, Sofia Fahlvik, Claes Thelander, Heiner Linke, and Martin Leijnse
Division of Solid State Physics and NanoLund, Lund University, Box 118,S-221 00 Lund, Sweden

We experimentally verify hitherto untested theoretical predictions about the thermoelectric properties of Kondo correlated quantum dots (QDs). The specific conditions required for this study are obtained by using QDs epitaxially grown in nanowires, combined with a recently developed method for controlling and measuring temperature differences at the nanoscale. This makes it possible to obtain data of very high quality both below and above the Kondo temperature, and allows a quantitative comparison with theoretical predictions. Specifically, we verify that Kondo correlations can induce a polarity change of the thermoelectric current, which can be reversed either by increasing the temperature or by applying a magnetic field.

PACS numbers: 72.10.Fk, 72.20.Pa, 73.63.Kv, 81.07.Gf

Measurements of electric and thermoelectric transport properties can be used to reveal and characterize novel strongly correlated phases, which often appear in meso- and nano-scale systems. The Kondo effect [1] is a prominent example where interactions between conduction electrons and magnetic impurities result in a many-body singlet state involving the impurity spin and a large number of conduction electrons. In metals it leads to increased resistivity at low temperatures where the magnetic impurity scattering dominates. More recently, quantum dots (QDs) tunnel-coupled to two leads have provided a platform for more detailed experimental studies of the Kondo effect [2–4]. In QDs, the Kondo scattering lifts the Coulomb blockade [5–7] and gives rise to a peak in the differential conductance $g = dI/dV$ around $V = 0$ (I is the current and V is the bias voltage).

Several theoretical works (see, e.g., Refs [8–16]) have proposed that additional insights into Kondo physics can be gained from thermoelectric measurements. Here, a temperature difference $\Delta T = T_c - T_h$ is applied between a hot (h) and a cold (c) lead and one measures either the resulting thermocurrent I_{th} (measured under closed-circuit conditions), or the thermovoltage V_{th} (measured under open-circuit conditions). In QDs without Kondo correlations, I_{th} and V_{th} have characteristic shapes as functions of the gate voltage, V_G, exhibiting a sign reversal (zero crossing) at each charge degeneracy point as well as in the center of each Coulomb valley [17–19]. It has been theoretically predicted [9] that Kondo correlations would significantly change this behavior by removing some zero crossings and consequently reversing the polarity of I_{th} and V_{th} over a finite V_G range. Whether a QD shows the typical Kondo or non-Kondo behavior depends sensitively on several system parameters. Therefore, by observing the qualitative change in thermoelectric response as Kondo correlations are suppressed, e.g., by increased average temperature $T = (T_h + T_c)/2$ or magnetic field B, one can not only gain insights into Kondo physics, but also probe the internal QD energy scales.

Despite such clear theoretical predictions, experimental studies of the thermoelectric properties of Kondo correlated QDs remain rather limited [20, 21]. Experimentally uncontrolled internal QD degrees of freedom often complicate even a qualitative comparison with theory. Therefore, the predicted reversal of I_{th} and V_{th} has been difficult to observe (although some unpublished data exist [22]) and the response to a B field has, to the best of our knowledge, not been investigated.

In this Letter, we take important steps towards filling this gap between experiments and theory by presenting thermoelectric measurements on several Kondo correlated QDs. We measure I_{th} and $g(V \approx 0) = g_0$ over consecutive Kondo and non-Kondo Coulomb valleys. We observe the sign reversal of I_{th} theoretically predicted for the Kondo regime, and measure the transition between Kondo and non-Kondo behavior as T is increased, finding quantitative agreement with theory [9]. Furthermore, we apply an external B field which also destroys Kondo correlations and find that a surprisingly large B is needed to recover the typical non-Kondo behavior.

Our observations necessitate overcoming significant experimental challenges to access the parameter regimes which most clearly reveal the Kondo correlations and allow detailed comparison with theory predictions. The requirements include: *(i)* strong quantum confinement such that transport is dominated by a single orbital ($\delta\epsilon \gg k_B T, \Gamma$, where $\delta\epsilon$ is the orbital spacing and Γ is the tunnel coupling); *(ii)* large charging energy U and tunnel coupling, and low temperature, such that the Kondo regime is reached ($U \gg \Gamma \gg k_B T$ and $T < T_K$, where T_K is the Kondo temperature); *(iii)* application and characterization of ΔT across a very small QD; *(iv)* the ability to controllably tune the system in and out of the Kondo regime, e.g., by gating or by varying T or B.

To achieve the above requirements we use QDs epitaxially defined in axially heterostructured InAs/InP nanowires grown by chemical beam epitaxy [24] [see inset in Fig. 1(a)]. Each InAs nanowire from the same growth is about 60 nm in diameter and contains two thin InP segments that confine an approximately 20 nm long InAs QD, similar to those used in our previous studies [25, 26]. The small QD size and the small effective mass of InAs

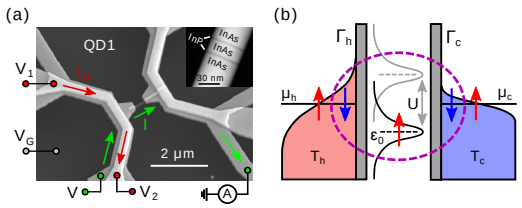


FIG. 1: (a) Scanning electron microscope (SEM) image of device QD3. An InAs/InP nanowire containing a QD is contacted to metallic leads for electrical biasing with voltage V (see SM [23] for details on circuitry). Additional heater leads (lighter gray) enable application of a thermal bias ΔT to the QD by running a current I_H resulting from a heater bias $V_H = V_1 - V_2$. Only one heater is used in the experiment. Inset: Close-up scanning transmission electron microscope with high angle annular dark field (STEM-HAADF) image of an InAs/InP nanowire from the same growth. (b) Sketch of an unbiased spin-1/2 QD tunnel-coupled to two leads (h and c).

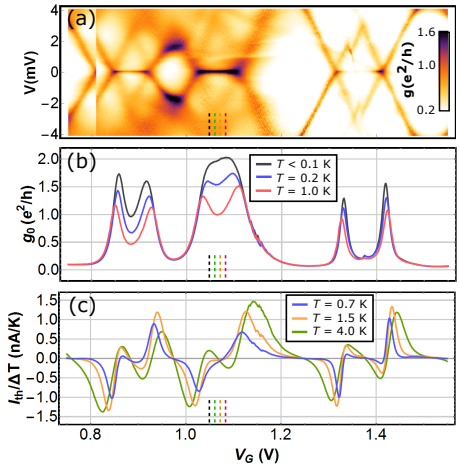


FIG. 2: (a) g as a function of V and V_G, measured at $T = T_0 < 100$ mK. (b) The corresponding g_0 as a function of V_G measured at three different T. (c) I_{th} normalized by ΔT as a function of V_G measured at three different T. The horizontal V_G axis is the same in (a)–(c). Vertical dashed lines refer to V_G values in Fig. 3.

give sufficiently large U and $\delta\epsilon$, and the large g-factor allows tuning the Zeeman energy over a wide range. We use the fabrication process developed in Ref. [27] to fabricate thermoelectric devices. Figure 1(a) shows a scanning electron microscope (SEM) image of the device QD1. In short, the devices are fabricated on an n-doped Si wafer coated with SiO_2. Two Ni/Au leads are used to contact the outer InAs segments on each side of the QD. The nanowires along with the contacting leads are coated with HfO_2 in order to electrically isolate the heater leads from the electrical biasing circuit [25, 27]. A back contact to the Si wafer is at a voltage V_G and allows for electrostatic gating of the epitaxially defined QDs. We let T_h and T_c denote the temperatures of the nanowire leads contacting the QD, which might differ from those in the metallic leads further away. Application of a heating current I_H increases both T_h and (to a lesser degree) T_c, which in our devices gives control over ΔT and T while maintaining a roughly constant $\Delta T/T \approx 0.30 - 0.35$. T_h and T_c are estimated based on QD thermometry (see Supplemental Material (SM) [23] and Ref. [26] for details). During this study we characterized three QDs (QD1, QD2 and QD3) showing similar behavior. Only the data from QD1 is presented with figures in the main text. Results on all devices are summarized in Table I (see SM [23] for the corresponding data on other devices). The characterization was done in a dilution refrigerator with electron base temperature $T_0 < 100$ mK.

Figure 1(b) shows a sketch of a single-level QD with orbital energy ε_0 and onsite Coulomb repulsion U, coupled to leads by tunnel couplings Γ_h and Γ_c (Anderson model). The Kondo effect occurs when the level is occupied by a single electron. It originates from anti-ferromagnetic exchange interaction due to virtual exchange of electrons between the leads and the QD. Kondo correlations give rise to the formation of a singlet-like state (with binding energy $\sim k_B T_K$), involving the QD spin and a large number of electron spins in the leads. Below this energy the system behaves as a Fermi liquid and Coulomb blockade is lifted.

We use Fig. 2 to identify the effects of Kondo correlations in the experimental data. The measured charge stability diagram at $T_0 < 100$ mK in Fig. 2(a) shows an increased g_0 inside Coulomb diamonds corresponding to odd electron numbers on the QD. In the absence of Kondo correlations one expects $g_0 < e^2/h$, but Fig. 2(b) shows that (at $T = T_0$) g_0 approaches the limit $2e^2/h$, as expected in the Kondo regime. Increasing T reduces g_0 in the odd occupancy Coulomb valleys but has little effect on valleys with even occupancy.

Figure 2(c) shows $I_{th}/\Delta T$ measured over the same gate range. We note that around $V_G = 1.06$ V, where the strongest Kondo correlations are seen in (a) and (b), there is a qualitative change in $I_{th}(V_G)$ with increasing T with two sign reversals being absent at low T. Therefore, we focus our analysis on this particular V_G range and come back to a detailed discussion of the thermoelectric behavior later.

Figure 3 presents the analysis for determining T_K and Γ. We use the T dependence of g_0 at the chosen V_G range to determine T_K using the phenomenological ex-

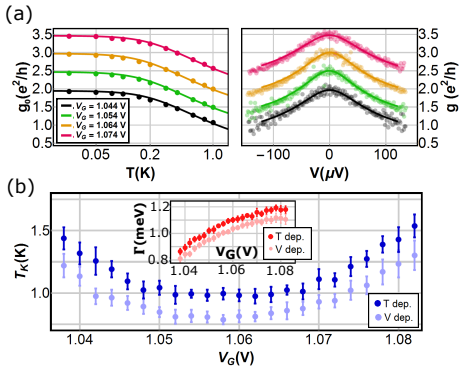


FIG. 3: (a) Dots are the measured values of $g_0(T)$ in the left panel and of $g(V)$ at T_0 in the right panel, both for four different values of V_G, also indicated by vertical dashed lines in Fig. 2 with the corresponding colors. The solid lines for $g_0(T)$ are fits to Eq. (1) whereas the solid lines for $g(V)$ are fits to Eq. (2) in the SM [23]. The curves and data points at different V_G are offset by 0.5 e^2/h in g from each other. (b) T_K as a function of V_G determined from fitting the measured $g_0(T)$ to Eq. (1) (dark blue points) and the measured $g(V)$ to Eq. (2) in the SM [23] (light blue points). The error bars represents a 95 % confidence interval for T_K as a fit parameter. Inset to (b): the corresponding estimates of Γ as a function of V_G using Eq. (2).

pression [2, 28]

$$g_0(T) = g_0(T=0)\left[1+\left(2^{1/s}-1\right)\left(\frac{T}{T_K}\right)^2\right]^{-s}, \quad (1)$$

where $s = 0.22$ for a spin-1/2 QD and $g_0(T = 0)$ and T_K are used as free fit parameters. Examples of the fits can be seen in the left panel of Fig. 3(a) while the corresponding T_K fit values are plotted in Fig. 3(b). We also cross check the T_K values by fitting the V dependence of g instead [23, 29]. Examples of those fits are shown in the right panel of Fig. 3(a) while the corresponding fit values are plotted in Fig. 3(b). Although the V dependence of g yields somewhat lower T_K values, there is a qualitative agreement between the two methods. In all further analysis we use the values extracted using the more well-established method based on the T-dependency of g_0.

For a single-orbital model, T_K is given by [7]

$$k_B T_K = \frac{1}{2}\sqrt{\Gamma U}\exp\left(\frac{\pi\varepsilon_0(\varepsilon_0+U)}{\Gamma U}\right), \quad (2)$$

where ε_0 is the energy of the QD orbital relative to the Fermi level of the leads and varies from 0 to $-U$ across the Coulomb valley. Equation (2) is strictly valid only in the Kondo regime where $-U + \Gamma/2 < \varepsilon_0 < -\Gamma/2$ [28, 30], i.e., far enough from the charge degeneracy points into the Coulomb valley. We estimate $U \approx 3.5$ meV which is used to calculate Γ from the estimated T_K values using Eq. (2) [see inset of Fig. 3(c)]. We find that Γ has a slight V_G dependence, which is commonly observed in nanowire QDs because of the quasi one-dimensional density of states in the leads.

We now turn our attention to the thermoelectric properties of QDs in the Kondo regime. The $T = 4$ K trace in Fig. 2(c) illustrates the expected behavior of $I_{th}(V_G)/\Delta T$ in the absence of Kondo correlations, where it undergoes twice as many sign reversals as there are charge degeneracy points – one when passing through zero at each of the charge degeneracy points and one in the middle of every Coulomb valley [17–19]. It was theoretically predicted in Ref. [9] that this behavior is qualitatively different in the presence of Kondo correlations, which cause the zero crossings at the degeneracy points to disappear. Our experimental data verifies this prediction as the two sign reversals in the gate range between $V_G = 1.02$ and 1.10 V disappear at low T. This behavior is in contradiction with what the Mott relation [31] would suggest for the measured $g_0(V_G)$ as the center of the Coulomb valley in Fig. 2(b) develops already at $T = 0.2$ K (see SM [23] for details on Mott relation). Two additional Kondo resonances are also seen in Fig. 2 (close to $V_G = 0.9$ and 1.4 V) but because T_K is lower in those cases the sign reversal is not observed in Fig. 2(c).

Figure 4 shows the sign reversal of I_{th} more closely and focuses on the effects of increasing T and B. Both are known to destroy Kondo correlations and it is therefore intuitive that also the sign reversal should be affected. We aim to quantify the values of B and T below which the Kondo-induced sign reversal takes place. Figures 4(a)–(d) show data at different B values, each for several different T. The corresponding charge stability diagrams for the same values of B are displayed in Fig. 4(e). The sign reversal of I_{th} as a function of T is best seen in Fig. 4(a) where $B = 0$ T. The trace at $T = 1$ K $\approx T_K$ shows a single zero-crossing within the Coulomb valley marked as the "Kondo region". By raising the temperature $T > T_K$ the two additional zero-crossings are recovered, indicating a reversal of the direction of I_{th} within the Kondo region. This observation is a clear verification of theoretical predictions in Ref. [9].

Based on the splitting of the Kondo peak, observable in Fig. 4(e), we estimate the electron g-factor $|g_Z| \approx 9$ [23]. Thus, $T_K \approx 1$ K corresponds to $B \approx 0.17$ T. Interestingly, however, the behavior of I_{th} at $B = 0.5$ T, as shown in Fig. 4(b), remains qualitatively and quantitatively similar to the zero field case. Only when increasing B to 1.0 T and 2.0 T the two additional zero-crossings are recovered at all accessible T [see Figs. 4(c) and (d)].

Closer examination of the sign reversal requires analysis of the small I_{th} within the Coulomb valley which is sensitive to the experimental uncertainties in the applied electrical bias ($\delta V \approx \pm 1$ μV). We therefore fit the

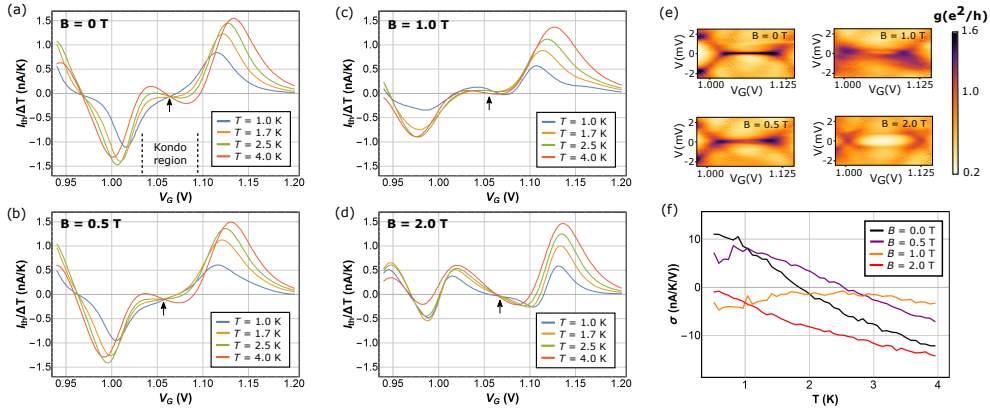


FIG. 4: (a)–(d) Measured I_{th} normalized by ΔT for different T as indicated in the figure. The black arrows indicate V_G positions around which the values of σ in (f) are determined. The magnetic field is increased from $B = 0$ to $B = 2$ T from (a)–(d) as indicated in the figures. The dashed lines in (a) indicate the V_G range that corresponds to the Kondo regime. (e) Charge stability diagrams showing g as a function of V and V_G, measured at the base temperature $T_0 < 100$ mK for values of B corresponding to those used in (a)–(d). (f) Thermocurrent slope, $\sigma = (dI_{th}/dV_G)/\Delta T$, as a function of T at B field values corresponding to those used in (a)–(d).

gate-slope of the thermocurrent, $(dI_{th}/dV_G)/\Delta T = \sigma$, at the center of the Coulomb valley [marked by arrows in Figs. 4(a)–(d)] and use its sign as an alternative indicator for the sign reversal of I_{th}. Figure 4(f) shows $\sigma(T)$ measured at different B values. We let T_1 denote the temperature at which $\sigma(T)$ changes from positive to negative. Interestingly, we find that T_1 is larger at $B = 0.5$ T ($T_1 \approx 2.5$ K) than at $B = 0$ T ($T_1 \approx 1.8$ K), however this result does not seem to be reproduced in other devices and we do not have an explanation for it. In contrast, at field values $B = 1.0$ T and $B = 2.0$ T σ no longer reverses sign as a function of T. Therefore, we conclude that the crossover happens between $B = 0.5$ T and $B = 1.0$ T. These field values correspond to $|g_Z|\mu_B B/k_B \approx 3T_K - 6T_K$, which is consistent with measurements under B field on other devices [23].

Table I summarizes our results from all devices, see SM [23] for the corresponding data and analysis. We estimate the relative errors for U and T_K to be in the range $\pm 10\%$, which translates into a similar error for Γ. The accuracy of T_1 depends mostly on the accuracy of the thermometry, which we have not been able to quantify. However, we do not expect it to be a source of significant error. For all resonances we find $T_1/T_K \approx 1.2-1.8$. This is in quantitative agreement with theory predictions in Ref. [9] where $T_1/T_K \approx 1.6$ for $U/\Gamma = 3$ at $B = 0$ T.

In conclusion, we have presented a detailed experimental study of the thermoelectric properties of Kondo correlated QDs. Our measurements confirm the theoretical prediction [9] that sufficiently strong Kondo correlations can reverse the direction of I_{th} over a finite range in V_G. We find quantitative agreement with theoretical predictions for the temperature T_1 at which sign the reversal takes place. We have also investigated the magnetic field dependence of I_{th} and conclude that, unlike other transport quantities which change behavior at $|g_Z|\mu_B B/(k_B T_K) \approx 1$ [32], the sign reversal of I_{th} remains until this ratio is significantly larger than 1. This raises new questions and opens up for further theoretical and experimental studies. More generally, our work demonstrates that the use of thermoelectric measurements can be a sensitive probe of Kondo physics and other strong correlation effects. An interesting direction for future works is to investigate more complex QDs with additional symmetries [16] or the nonlinear, large ΔT, regime [14, 15], where theoretical predictions are much more challenging.

TABLE I: Summary of data from several devices. QD1a (in bold) represents results obtained on the device QD1 for which the data is shown in this Letter. QD1b represents results obtained on the same device QD1 but in a different V_G range. QD2 and QD3 represents results obtained on devices QD2 and QD3.

	U(meV)	T_K(K)	Γ(meV)	T_1(K)
QD1a	**3.5**	**1.0**	**1.1**	**1.8**
QD1b	2.2	0.6	0.7	0.7
QD2	2.6	0.6	0.8	0.8
QD3	3.0	0.8	1.0	1.2

Acknowledgements – We gratefully acknowledge funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7-People-2013-ITN) under REA grant agreement no. 608153 (PhD4Energy), from the Swedish Research Council (projects 2012-5122 and 2016-03824), from the Knut and Alice Wallenberg Foundation (project 2016.0089), from the Swedish Energy Agency project (project 38331-1), from NanoLund, and computational resources from the Swedish National Infrastructure for Computing (SNIC) at LUNARC (projects SNIC 2017/4-10 and SNIC 2018/6-3).

[1] J. Kondo, Progr. Theor. Phys. **32**, 37 (1964).
[2] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, Nature **391**, 156 (1998).
[3] S. M. Cronenwett, T. H. Oosterkamp, and L. P. Kouwenhoven, Science **281**, 540 (1998).
[4] J. Nygård, D. H. Cobden, and P. E. Lindelof, Nature **408**, 342 (2000).
[5] T. K. Ng and P. A. Lee, Phys. Rev. Lett. **61**, 1768 (1988).
[6] L. I. Glazman and M. E. Raikh, JETP Lett. **47**, 452 (1988).
[7] A. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, New York, 1993).
[8] D. Boese and R. Fazio, Eur. Phys. Lett. **56**, 576 (2001).
[9] T. A. Costi and V. Zlatic, Phys. Rev. B **81**, 235127 (2010).
[10] P. Roura-Bas, L. Tosi, A. A. Aligia, and P. S. Cornaglia, Phys. Rev. B **86**, 165106 (2012).
[11] J. Azema, A.-M. Daré, S. Schäfer, and P. Lombardo, Phys. Rev. B **86**, 075303 (2012).
[12] I. Weymann and J. Barnaś, Phys. Rev. B **88**, 085313 (2013).
[13] L. Z. Ye, D. Hou, R. Wang, D. Cao, X. Zheng, and Y. J. Yan, Phys. Rev. B **90**, 165116 (2014).
[14] A. Dorda, M. Ganahl, S. Andergassen, W. von der Linden, and E. Arrigoni, Phys. Rev. B **94**, 245125 (2016).
[15] M. A. Sierra, R. López, and D. Sánchez, Phys. Rev. B **96**, 085416 (2017).
[16] D. B. Karki and M. N. Kiselev, Phys. Rev. B **96**, 121403 (2017).
[17] C. W. J. Beenakker and A. A. M. Staring, Phys. Rev. B **46**, 9667 (1992).
[18] A. A. M. Staring, L. W. Molenkamp, B. W. Alphenaar, H. van Houten, O. J. A. Buyk, M. A. A. Mabesoone, C. W. J. Beenakker, and C. T. Foxon, Eur. Phys. Lett. **22**, 57 (1993).
[19] A. Dzurak, C. Smith, M. Pepper, D. Ritchie, J. Frost, G. Jones, and D. Hasko, Solid State Commun. **87**, 1145 (1993).
[20] R. Scheibner, H. Buhmann, D. Reuter, M. N. Kiselev, and L. W. Molenkamp, Phys. Rev. Lett. **95**, 176602 (2005).
[21] S. F. Svensson, E. A. Hoffmann, N. Nakpathomkun, P. Wu, H. Q. Xu, H. A. Nilsson, D. Sánchez, V. Kashcheyevs, and H. Linke, New J. Phys. **15**, 105011 (2013).
[22] H. Thierschmann, Ph.D. thesis, University of Würzburg (2014).
[23] See Supplemental Material at [URL will be inserted by publisher] for data from additional devices and description of temperature control and measurements.
[24] M. T. Björk, C. Thelander, A. E. Hansen, L. E. Jensen, M. W. Larsson, L. R. Wallenberg, and L. Samuelson, Nano Lett. **4**, 1621 (2004).
[25] A. Svilans, M. Leijnse, and H. Linke, C. R. Physique **17**, 1096 (2016).
[26] M. Josefsson, A. Svilans, A. M. Burke, E. A. Hoffmann, S. Fahlvik, C. Thelander, M. Leijnse, and H. Linke, arXiv:1710.00742 (to appear in Nature Nanotechnology) (2017).
[27] J. G. Gluschke, S. F. Svensson, C. Thelander, and H. Linke, Nanotechnology **25**, 385704 (2014).
[28] A. V. Kretinin, H. Shtrikman, D. Goldhaber-Gordon, M. Hanl, A. Weichselbaum, J. von Delft, T. Costi, and D. Mahalu, Phys. Rev. B **84**, 245316 (2011).
[29] M. Pletyukhov and H. Schoeller, Phys. Rev. Lett. **108**, 260601 (2012).
[30] F. D. M. Haldane, Phys. Rev. Lett. **40**, 416 (1978).
[31] N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Dover Publications, 1958), ISBN 048660456X.
[32] M. Filippone, C. P. Moca, A. Weichselbaum, J. von Delft, and C. Mora (2018), arXiv:1609.06165.

Supplementary Material

Thermoelectric characterization of the Kondo resonance in nanowire quantum dots

Artis Svilans, Martin Josefsson, Adam M. Burke, Sofia Fahlvik, Claes Thelander, Heiner Linke, and Martin Leijnse
Division of Solid State Physics and NanoLund, Lund University, Box 118,S-221 00 Lund, Sweden

PACS numbers: 72.10.Fk, 72.20.Pa, 73.63.Kv, 81.07.Gf

DATA ON OTHER DEVICES

The main paper presents data and analysis on a certain resonance of device quantum dot 1 (labeled QD1a). In this section we provide additional data and the corresponding analyses on devices QD2 and QD3, as well as additional data on quantum dot 1 (labeled QD1b) measured over a different range of gate voltages. Figure S1 shows SEM images of the devices (QD1, QD2 and QD3). They all show clear signs of Kondo correlations in the charge stability diagrams (Fig. S2) when characterized at the base temperature of the cryostat $T_0 < 100$ mK. They also show similar behavior of the thermocurrent I_{th} as a function of average temperature $T = (T_h + T_c)/2$ (Fig. S4), in particular a Kondo-related inversion of the direction of I_{th} as a function of T. Complementary data on the behavior of I_{th} in the presence of finite magnetic field B is also given for QD1b and QD3. Unfortunately no data in the presence of magnetic field is available on QD2. All three devices are identical by design and are fabricated in the same process on the same sample chip. Figure S1 shows SEM images of the corresponding devices.

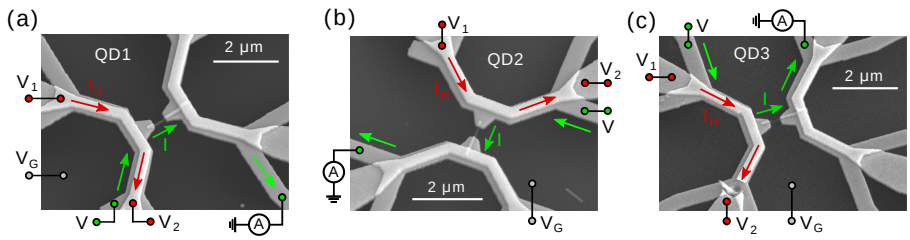


Fig. S1: Scanning electron micrograph (SEM) of devices QD1, QD2 and QD3. InAs/InP nanowires containing QDs are contacted by metallic leads on both sides of the QD. A bias voltage V can be applied to the QDs via the leads. The path of the current I, resulting from $V > 0$, is indicated by green arrows. The leads along with the nanowire are covered in HfO_2 to electrically insulate the overlaying thermal biasing circuit from the electrical circuit. A heater bias $V_H = V_2 - V_1$ is applied to one of the heater leads of every device resulting in a heating current I_H that dissipates heat locally. The path and direction of I_H resulting from $V_H > 0$ is indicated by the red arrows. The damage to the heater visible on Device QD3 occurred during the process of unloading from the cryostat after the experiment was finalized.

Accounting for series resistances in the measured differential conductance

The current I though the QD is measured by a current preamplifier with input impedance $R_I = 1$ kΩ. Both the contact and the heater leads are connected to the external electrical setup via DC lines with RC filters (total of $R_{RC} = 3.26$ kΩ per line) designed to cut off frequencies above 300 Hz. In order to account for the serial resistances originating from the DC line filters and the input impedance of the current preamplifier when determining g, a distinction is made between V, which is the voltage across the QD, and V_{ext}, which is the voltage applied to the entire circuit (including the filter resistances and the input impedance of the preamplifier). Therefore, g is calculated as $g = dI/(dV_{ext} - RdI) = dI/dV$, where $R = 2R_{RC} + R_I = 7.52$ kΩ.

When taking the series resistance into account the resulting array of g data for plotting charge stability diagrams is no longer rectangular in the coordinates of V and V_G, therefore, polynomials are used to interpolate $g(V)$ in between

the data points [concerns Fig. 3(c) in the main paper and Figs. S2 and S5 here]. Figure S2 plots the charge stability diagrams measured on QD1b, QD2 and QD3 that show clear signs of Kondo correlations in the form of zero-bias conductance peaks (similar to resonance QD1a discussed in the main paper).

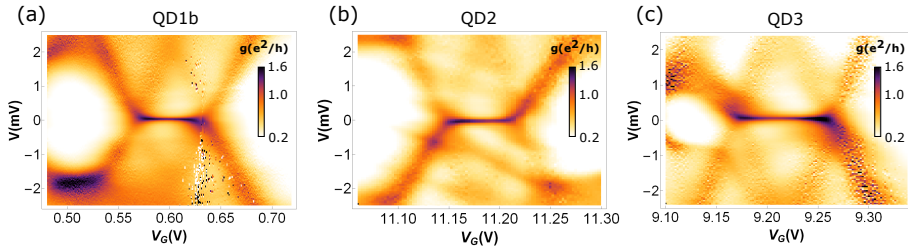


Fig. S2: g as a function of V_G and V for Kondo resonances QD1b (a), QD2 (b) and QD3 (c).

Characterization of the Kondo temperature T_K and the strength of the tunnel-coupling Γ.

A characterization of the Kondo effect, equivalent to that presented in the main paper for QD1a, is also carried out for QD1b, QD2 and QD3. The characterization presented in Figs. 3 and S3 is based on analyzing the behavior of g as a function of T and V. For the T dependence we use the notation $g(V \approx 0) = g_0$ to distinguish it from $g(V)$. The quantity g_0 in Fig. 3 and Fig. S3 is determined using linear fits to $I(V)$ within a range of $V \approx \pm 10$ μV, whereas $g(V)$ values are calculated as the differential increase in current ΔI in response to the voltage change ΔV between two sequential steps in V. We note that in the Fig. 2 the quantity g_0 is approximated by I/V_{ext} where $V_{ext} = 25$ μV

The first row in Fig. S3 shows how $g_0(V_G)$ evolves with increasing T. Within the Coulomb valleys g_0 demonstrates a gradual decrease with T indicating the melting of the Kondo correlations. Examples of fits to $g_0(T)$ and $g(V)$ are shown in the second row of Fig. S3. The two dependencies are used to determine the value of the Kondo temperature T_K by fitting it to empirical analytic expressions. For analyzing g_0, we use a standard expression

$$g_0(T) = g_0(T=0)\left[1 + \left(2^{1/s} - 1\right)\left(\frac{T}{T_K}\right)^2\right]^{-s}, \tag{1}$$

where $s = 0.22$ for a spin-1/2 QD and g_0 and T_K are used as free fit parameters [1, 2]. For analyzing the V dependence of g we use the recently proposed expression [2]

$$g(V) = g_0(T=0)\left[1 + \left(\frac{2^{1/s} - 1}{1 - b + b\left(\frac{e(V-V_0)}{kT_K\prime}\right)^{s'}}\right)\left(\frac{e(V-V_0)}{kT_K\prime}\right)^2\right]^{-s}, \tag{2}$$

where $s = 0.32$, $b = 0.05$, $s' = 1.26$ and the parameters $g_0(T=0)$, V_0 and T_K are used as free fit parameters. Note that the qualitative shapes of $g_0(T)$ and $g(V)$, as defined by in Eqs. (1) and (2), have a slightly different qualitative behavior. Also note that the two temperature constants, T_K and $T_K\prime$, do not have the same meaning ($T_K \neq T_K\prime$). However, as pointed out in Ref. [2] using $T_K\prime \approx 1.8T_K$ yields a definition of T_K in Eq. (2) that is consistent with the definition of T_K in Eq. (1), therefore we plot the results from $g(V)$ as $T_K \approx T_K\prime/1.8$.

The third row in Fig. S3 shows the results for T_K as a function of V_G. As expected, T_K has a minimum in the center of the Coulomb valley. For a single-orbital model, T_K is given by [3] [Eq. (2) in the main paper]

$$k_B T_K = \frac{1}{2}\sqrt{\Gamma U}\exp\left(\frac{\pi\varepsilon_0(\varepsilon_0 + U)}{\Gamma U}\right), \tag{3}$$

which can be used to extract the corresponding tunnel-coupling strength Γ. The results of this calculation are shown as insets in the third row of the Fig. S3. The V_G dependence of Γ might be explained by an energy dependent density of states in the leads, which are effectively quasi one-dimensional pieces of nanowires.

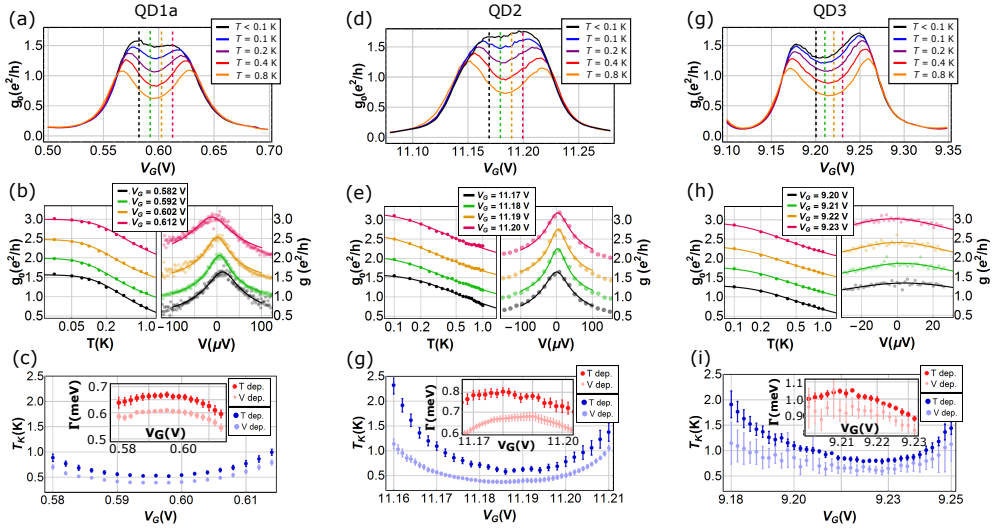


Fig. S3: Characterization of T_K and Γ. (a)–(c) present data for QD1b, (d)–(f) present data for QD2, and (g)–(i) present data for QD3. (a), (d) and (g) show g as a function of V_G at different T. (b), (e) and (h) show g as a function of T in left panels and as a function of V in right panels, both for several different V_G settings which are indicated in (a), (d) and (g) as dashed lines with the same color. The data points and curves are offset by 0.5 e^2/h. (c), (f) and (i) show T_K results using the data from the T-dependence (dark blue points) and using the data from the V-dependence (light blue points). Error bars represent a 95% confidence interval for the T_K value. Insets: results of numerical calculation of Γ using the results for T_K and Eq. (3) where $\varepsilon_0 = -e\alpha_G(V_G - V_{G0})$ with α_G being the gate lever arm. When calculating Γ, we used the following parameter values extracted from the charge stability diagrams: **QD1b**: $U = 2.2$ meV, $\alpha_G = 0.0349$, $V_{G0} = 0.564$ V; **QD2**: $U = 2.6$ meV, $\alpha_G = 0.0382$, $V_{G0} = 11.152$ V; **QD3**: $U = 3.0$ meV, $\alpha_G = 0.0326$, $V_{G0} = 9.169$ V.

Thermoelectric characterization

Thermoelectric characterization is realized by applying ΔT and measuring the thermocurrent I_{th}. In the linear response regime, and with no parasitic resistances in series with the QD, it is related to the thermovoltage V_{th} via the relation $I_{th} = g_0 V_{th}$, therefore the detection of the sign inversion in I_{th} also implies a sign inversion in V_{th}. Applying ΔT in our experimental devices also heats the colder reservoir. We use this effect to our advantage in order to study the $T = (T_h + T_c)/2$ dependence of I_{th} as a function of V_G. All $I_{th}(V_G)$ traces (at each T) have been measured by sweeping V_G in both directions a number of times (4 or 8) in order to average out current fluctuations related to the instability of the electrical bias V across the QD with a typical magnitude of ± 1 μV. Plots and the data analyses use the mean values of I_{th} at every V_G.

Figure S4 shows results of $I_{th}/\Delta T$ for QD1b, QD2 and QD3. In the absence of magnetic field ($B = 0$) all three devices show a similar behavior to the one presented in the main paper, i.e., $I_{th}/\Delta T$ shows a sign inversion as a function of T within the Coulomb valley where Kondo correlations can be observed at low T. We let T_1 denote the temperature at which the sign inversion takes place. Similarly as done in the main paper, we examine $\sigma = (dI_{th}/dV_G)/\Delta T$ as a function of T to study the sign inversion. To complement the data in the main paper, we also present data on QD1b and QD3 in presence of magnetic field B. For QD1b we present data at $B = 0.2$ T and $B = 1.5$ T [$\sim 1.3 k_B T_K/(|g_Z|\mu_B)$ and $\sim 10 k_B T_K/(|g_Z|\mu_B)$ respectively]. For QD3 we present data at $B = 0.25$ T ($\sim 2.3 k_B T_K/(|g_Z|\mu_B)$). One can observe that for both (QD1b and QD3) the presence of a weak field [B between 1 and 3 $k_B T_K/(|g_Z|\mu_B)$] does not reverse back the direction of I_{th} which is consistent with observations in QD1a. However, the presence of the weak field seem to reduce T_1 instead of increasing it, as seen in presence of $B \approx 3 k_B T_K/(|g_Z|\mu_B)$ for QD1a. In addition we show that under application of a strong field ($B = 1.5$ T) to QD1b, the sign inversion is no longer observable and a

behavior of $I_{th}/\Delta T$ characteristic to QDs without Kondo correlations is recovered, similarly as in the case of a strong field for QD1a [$B = 2.0$ T, corresponding to $\sim 12k_BT_K/(|g_Z|\mu_B)$] presented in the main paper.

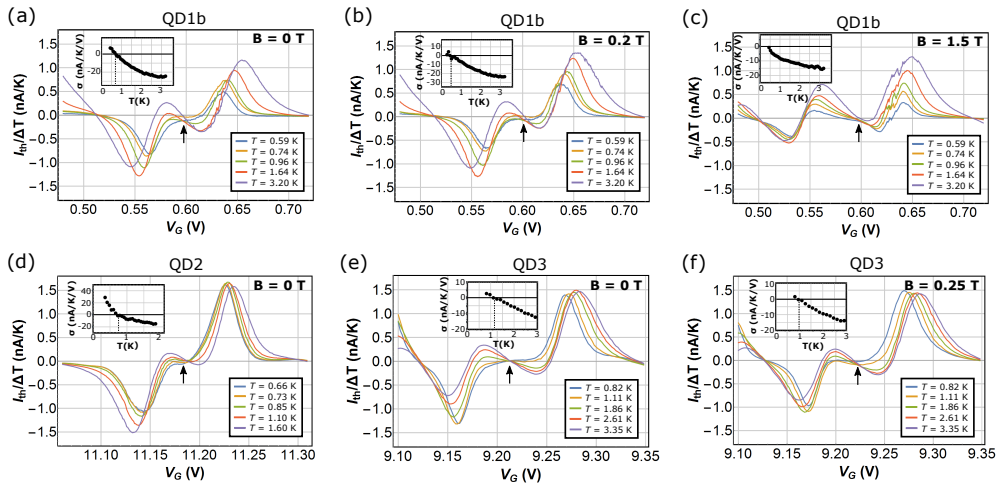


Fig. S4: Results of thermoelectric characterization. Results from QD1b are presented in (a) – (c), results on QD2 are shown in (d), and results from QD3 are shown in (e) and (f). All plots present results on the measured I_{th} normalized by $\Delta T = T_h - T_c$ for various $T = (T_h + T_c)/2$, as indicated in the figures. B values are indicated in the right top corners of all subfigures (a)–(f). The insets present data for the corresponding thermocurrent slopes $\sigma = (dI_{th}/dV_G)/\Delta T$ that are determined at the V_G values indicated by the black arrows. The temperature T_1 at which σ crosses 0 is estimated from the graph and indicated by the vertical dashed lines.

Measurements of g-factors

Figure S5 shows the splitting of the Kondo peaks in magnetic field, based on which we estimate the g-factors. In all cases, we find values slightly below the bulk value $g_Z^{\text{bulk}} \approx -14.9$ for InAs [4]. This is consistent with other studies showing a confinement-induced reduction and variation of g-factors [5, 6].

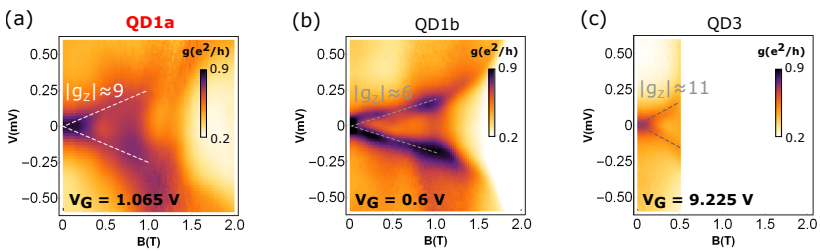


Fig. S5: Splitting of the Kondo peaks in magnetic field. All plots show differential conductance g as a function of V and B (at $T_0 < 100$ mK). (a) Data on QD1a, corresponding to the data the main paper. (b) Data on QD1b. (c) Data in QD3. The corresponding gate voltage settings for are indicated in figures. The dashed lines are rough estimates of the Zeeman splitting corresponding to the g-factor values, $|g_Z|$, that are also indicated in the figures.

COMPARISON OF THERMOELECTRIC DATA ON QD1A WITH THE MOTT RELATION

The Mott relation is an approximate relation between the conductivity G as a function of carrier energy E and the Seebeck coefficient S_{Mott} [7]

$$S_{Mott} = \frac{\pi^2 k^2 T}{3e} \frac{1}{G(E)} \frac{\partial G(E)}{\partial E}. \tag{4}$$

Equation (4) is a low-temperature approximation (based on a leading order Sommerfeld expansion), valid only when $G(E)$ does not vary significantly over an energy range $k_B T$ [8]. In nanoscale systems, $G(E)$ can vary rapidly because of confinement effects, and strong electron-electron interactions further complicate the picture. Therefore, the Mott relation might not even hold at extremely low temperatures. A further, commonly used, approximation is to assume that there is a linear dependence between the energy dependence and the gate-voltage dependence of $G(E)$, such that $E \propto -e\alpha_G V_G$, where α_G is the gate coupling.

In our experimental setup we measure the thermocurrent $I_{th}/\Delta T$ that in the linear-response regime would correspond to the Seebeck coeficient $S = -g_0 V_{th}/\Delta T$, where we replaced $G(E)$ with g_0 because of linear response. For the purpose of comparing the predictions of the Mott relation with our experimental results we therefore plot $-S_{Mott} g_0/T$ which is given by

$$-\frac{S_{Mott} g_0}{T} = \frac{1}{3\alpha_G} \left(\frac{\pi k}{e}\right)^2 \frac{\partial g_0}{\partial V_G}. \tag{5}$$

We use g_0 data from Fig. 2 of the main paper as well as corresponding data at additional temperatures to evaluate the quantity $-S_{Mott} g_0/T$. The results are shown in Fig. S6(a).

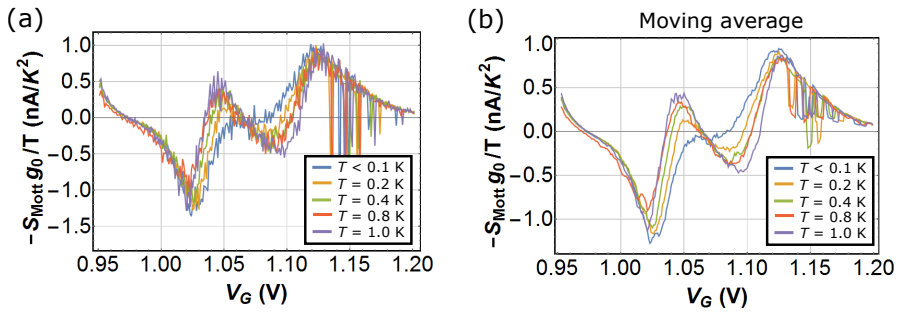


Fig. S6: Prediction for $-S_{Mott} g_0/T$ based on the Mott relation using conductance data from Fig. 2(b) (as well as additional analogous g_0 traces at $T = 0.4$ and 0.8 K). In (a) $\partial g_0/\partial V_G$ is calculated using the sequentially measured g_0 data points in V_G, (b) shows the moving average (over 5 points) of the curves plotted in (a). Here $\alpha_G = 0.039$.

Figure S6 demonstrates that, given the measured g_0, the Mott relation predicts the sign reversal of the thermoelectric signature within the Coulomb valley to occur below $T = 0.2$ K. This is in contradiction with the experimental observations in which the reversal of $I_{th}/\Delta T$ is detected at $T = T_1 \approx 1.8$ K, or at almost an order of magnitude higher temperature than the Mott relation predicts. This, in fact, can be concluded already from Fig. 2(b) in the main paper, as well as from the g_0 plots in Fig. S3. The sign inversion of S_{Mott} originates from the sign inversion of $\partial g_0/\partial V_G$ within the Coulomb valley region, and measurements on all devices show that there is a valley in g_0 at temperatures way below the measured T_1, meaning that the disagreement with the Mott relation is systematic. This demonstrates that measuring I_{th} (or V_{th}) indeed provides additional information which cannot be extracted from g_0.

THERMOMETRY OF QD3 IN THE WEAK COUPLING REGIME

To estimate the elevated electronic temperatures induced by the current I_H though the heater lead we use the thermocurrent I_{th} as a function of V_G in the weak coupling regime ($\Gamma \ll k_B T$) where Kondo correlations are absent.

The measured $I_{th}(V_G)$ is sensitive to both T_c and T_h and therefore can be compared to theoretical calculations to estimate the temperatures, in the same way as done in Ref. [9]. Such measurements were performed on all QDs, however, only QD3 displayed a resonance which exhibited the type of clear single-level behavior and weak enough coupling needed for the thermometry approach to be valid. Hence, this data is used to estimate T_h and T_c as a function of the heater bias V_H. We find an approximately linear dependence of temperatures on V_H which is consistent with observations in previous studies [9] (see Fig. S7). Unfortunately, the thermometry data on QD3 does not cover the full range of V_H used in measurements on other devices, and we therefore base our temperature estimates for large V_H on a linear extrapolation of T_h and T_c. Because all devices are identical by design we apply the thermometry result of QD3 also to devices QD1 and QD2. The estimated ΔT yields consistent results for $I_{th}/\Delta T$ on all devices which increases our confidence that this is a good approximation.

Our theoretical approach [10, 11] for calculating the current through the weakly coupled QD is based on the real time diagrammatic (RTD) approach [12], in which one expands the Liouville-von Neumann equation in Γ in order to calculate the stationary state reduced density matrix of the QD, as well as the stationary charge current. We keep all terms up to order Γ^2 in the expansion which, in addition to sequential tunneling, also accounts for co-tunneling, fluctuations and energy renormalization processes. The resistive load in series with the QD is included in the modeling by solving the self-consistent equation $I_{th}(V_G, V) = -V/R$ (see Supplementary Information of Ref. [9]).

We model the QD as a single-orbital Anderson model. This Hamiltonian can be broken into the respective Hamiltonians for the electronic reservoirs (H_R), the QD (H_D) and the tunnel couplings (H_T)

$$H = H_R + H_D + H_T. \tag{6}$$

Here the reservoirs are assumed to be non-interacting

$$H_R = \sum_{r=h,c} \epsilon_{r,\sigma,k} c^\dagger_{r,\sigma,k} c_{r,\sigma,k}, \tag{7}$$

and they are also assumed to be in local equilibrium at all times such that they can be characterized by the Fermi-Dirac distribution $f_r(\epsilon) = [e^{(\epsilon-\mu_r)/kT_r} + 1]^{-1}$. $c^\dagger_{r,\sigma,k}$ ($c_{r,\sigma,k}$) is the creation (annihilation) operator for an electron in reservoir r with wave vector k and spin σ, and $\epsilon_{r,\sigma,k}$ is the eigenenergy for the same electron. The QD is modeled as a single spin-degenerate energy level with on-site electron-electron interactions

$$H_D = \sum_\sigma \epsilon_\sigma n_\sigma + U n_\uparrow n_\downarrow, \quad n_\sigma = d^\dagger_\sigma d_\sigma, \tag{8}$$

with single particle energy ϵ_σ and interaction strength U. The field operators acting on the QD subspace are denoted by the letter d. Finally the tunneling Hamiltonian is given by

$$H_T = \sum_{r=h,c} t_{r,\sigma,k} c^\dagger_{r,\sigma,k} d_\sigma + \ h.c. \ . \tag{9}$$

The amplitude for an electron making a tunneling transition $t_{r,\sigma,k}$ is related to the tunneling rate Γ as

$$\Gamma_{r,\sigma} = 2\pi\nu_r |t_{r,\sigma,k}|^2, \tag{10}$$

where ν_r is the density of states of reservoir r.

The analysis required to obtain the temperatures is a three-step process. First, basic QD parameters are extracted from current measurements without a thermal bias. From the stability diagram in Fig. S7(a) we find $U = 4.2$ meV and $\alpha_G = 0.061$. Since the Coulomb diamonds in Fig. S7(a) are almost perfectly straight we consider symmetric voltage drop across both reservoirs, $\mu_{h,c} = \pm V/2$. Next we determine Γ_h and Γ_c by fitting the RTD theory to the measured current as a function of applied voltage as the QD is gated to the resonant condition $V_G = 2.385$ V, see Fig. S7(b). Determination of Γs is possible since the saturated current values for large V_{ext}, and thus also large V, are independent of other parameters, like temperature or series resistance [13]. In the final step we fit the RTD theory to the measured $I_{th}(V_G)$ using the previously extracted values for U, α_G, Γ_h and Γ_c, see Fig. S7(c). This is done for all traces of $I_{th}(V_G)$, two for each V_H setting.

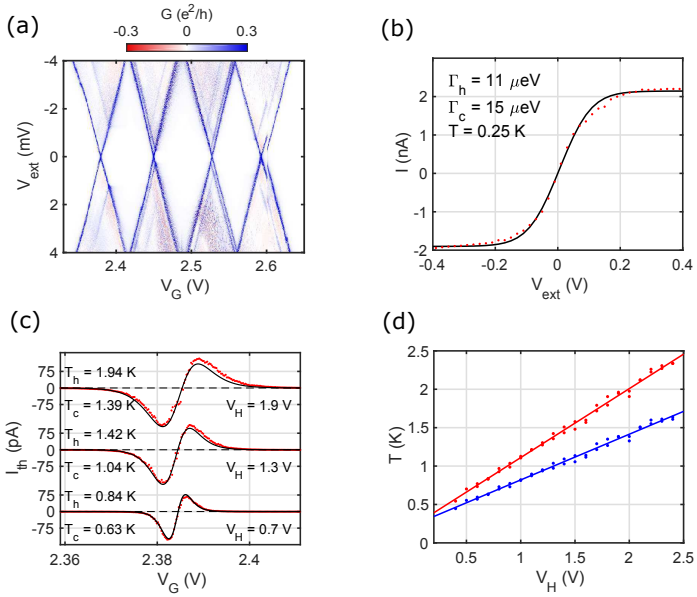


Fig. S7: Temperature estimates using a weakly coupled resonance. (a) $G = dI/dV_{ext}$ as a function of V_G and V_{ext} show four possible resonances for temperature characterization. Only the resonance at $V_G = 2.385$ V was well reproduced by a single-level QD model with weak tunnel couplings. $U = 4.2$ meV and $\alpha_G = 0.061$ are determined from (a). (b) The tunnel rates are determined by fitting the RTD theory to the measured $I(V_{ext})$ at the resonance without thermal bias using Γ_h and Γ_c as free parameters. (c) Using the Γs from (b) the temperatures of the electronic reservoirs are estimated by fitting the RTD theory to the measured $I_{th}(V_G)$ with T_h (plotted in red) and T_c (plotted in blue) as free parameters. The resulting temperatures and the applied V_H are indicated in the figure. (d) The estimated temperatures for all $I_{th}(V_G)$ at this resonance show a linear increase as a function of V_H. Linear regression for the two temperatures yield $T_h = 0.90V_H + 0.21$ and $T_c = 0.60V_H + 0.22$. All measurements and calculations in the figure include a series resistance $R = 7.52$ kΩ.

[1] A. V. Kretinin, H. Shtrikman, D. Goldhaber-Gordon, M. Hanl, A. Weichselbaum, J. von Delft, T. Costi, and D. Mahalu, Phys. Rev. B **84**, 245316 (2011).
[2] M. Pletyukhov and H. Schoeller, Phys. Rev. Lett. **108** (2012).
[3] A. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, New York, 1993).
[4] C. R. Pidgeon, D. L. Mitchell, and R. N. Brown, Physical Review **154**, 737 (1967).
[5] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer, Heidelberg, 2003).
[6] H. A. Nilsson, P. Caroff, C. Thelander, M. Larsson, J. B. Wagner, L.-E. Wernersson, L. Samuelson, and H. Q. Xu, Nano Lett. **9**, 3151 (2009).
[7] N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Dover Publications, 1958), ISBN 048660456X.
[8] A. M. Lunde and K. Flensberg, J. Phys.: Condens. Matter **17**, 3879 (2005).
[9] M. Josefsson, A. Svilans, A. M. Burke, E. A. Hoffmann, S. Fahlvik, C. Thelander, M. Leijnse, and H. Linke, arXiv:1710.00742 (2017).
[10] M. Leijnse and M. R. Wegewijs, Phys. Rev. B **78**, 235424 (2008).
[11] N. M. Gergs, S. A. Bender, R. A. Duine, and D. Schuricht, Physical Review Letters **120** (2018).
[12] J. Konig, H. Schoeller, and G. Schön, Phys. Rev. Lett. **78**, 4482 (1997).
[13] E. Bonet, M. M. Deshmukh, and D. C. Ralph, Phys. Rev. B **65**, 045317 (2002).

Printed by Media-Tryck, Lund 2018 NORDIC SWAN ECOLABEL 3041 0903

Lund University
Faculty of Engineering
Department of Physics
ISBN 978-91-7753-832-5